SLAC-R-494
CONF-9507258–
UC-414

PROCEEDINGS OF THE SUMMER INSTITUTE ON PARTICLE PHYSICS

July 10-21, 1995

The Top Quark & The Electroweak Interaction

Program Directors:
David Burke
Lance Dixon
David W.G.S. Leith

Edited by Jennifer Chan & Lilian DePorcel

Sponsored by Stanford University and the Stanford Linear Accelerator Center under contract with the U.S. Department of Energy, contract DE-AC03-76SF00515.

January 1997

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

TABLE OF CONTENTS

Part I. SUMMER SCHOOL LECTURES

J. H. Kühn
"Theory of Top Quark Production and Decay" 1

P. K. Sinervo
"Top Quark Studies at Hadron Colliders" 65

K. Fujii
"Top at Future Linear e^+e^- Colliders" 99

M. Swartz
"Review of Precision Electroweak Data" 101

C. J. S. Damerell
"Vertex Detectors: The State of the Art and Future Prospects" 103

J. L. Hewett
"The Role of Top in Heavy Flavor Physics" 187

T. Ypsilantis
"Techniques for Particle Identification" 225

L. J. Hall
"The Heavy Top Quark and Supersymmetry" 261

Part II. TOPICAL CONFERENCE

R. E. Hughes
"Top Physics at CDF" 291

N. J. Hadley
"Observation of the Top Quark with the D0 Detector" 309

M. Peskin
"Monopoles from Heaven, or S-Duality, and What It Means to You" 327

T. Schalk
"Precision Electroweak Experiments at SLD" 329

D. Strom
"LEP Precision Electroweak Measurements from the Z^0 Resonance" 331

R. Enomoto
"Recent Results from TRISTAN" 351

D. Wood
"Electroweak Results from the Tevatron" 359

R. S. Chivukula
"Dynamical Electroweak Symmetry Breaking and the Top Quark 375

U. Mallik
"Photons and Pomerons in Photoproduction at HERA" 385

K. Meier
"News from the Proton—Recent DIS Results from HERA" 403

L. M. Stuart
"Spin Structure Measurements from E143 at SLAC" 413

R. Plano
"QCD at SLD" 427

K. Lang
"Searches for Very Rare Decays of Kaons" 443

P. Kasper
"An Overview of the Fermilab Fixed Target Program" 461

E. P. Solodov
"Recent Results from the CMD-2 Detector at the VEPP-2M Collider" 463

W. G. J. Langeveld
"Search for Milli–Charged Particles at SLAC" 479

C. Bula
"Test of QED at Critical Field Strength" 495

R. Messner
"Heavy Quark Physics from SLD" 507

P. J. Dornan
"Heavy Quark Physics from LEP" 525

S. Menary
"New Results on CLEO's Heavy Quarks—Bottom and Charm" 547

P. Sphicas
"Heavy Quark Physics from the Tevatron" 567

M. Wise
"Recent Advances in Heavy Quark Theory" 569

APPENDICES

List of Participants 579

Previous SLAC Summer Institute Titles and Speakers 589

Author Index 593

PREFACE

The XXIII SLAC Summer Institute on Particle Physics addressed the physics of the recently discovered top quark, and its connection to the electroweak interaction and to physics beyond the Standard Model. The Institute attracted 227 physicists from 13 countries to SLAC, from July 10 to 21, 1995. The seven-day school portion of the Institute covered many avenues for studying the top quark, from its direct production at hadron colliders and at future electron-positron colliders, to its virtual effects in precision electroweak quantities, in heavy flavor physics, and in the renormalization of supersymmetric theories. Vertex detectors—critical for identifying the b quark decay products of the top—and Cherenkov techniques for particle identification were also reviewed. The traditional format of the school, with morning lectures followed by afternoon discussion sessions, continued to work well, and there was much lively interaction between lecturers and students.

The Institute concluded with a three-day topical conference covering recent developments in theory and experiment; this year, the highlights were the CDF and DØ top quark discovery. Also featured were updated precision electroweak measurements from SLC, LEP, and the Tevatron, heavy quark results from these facilities as well as CLEO, and new photoproduction and deep-inelastic scattering data from HERA.

We are grateful to all speakers for their efforts in preparing clear and stimulating lectures. We also thank the provocateurs for their assistance at the afternoon discussion sessions. Finally, we are indebted to Lilian DePorcel and Jennifer Chan for their hard work in putting together such a smoothly run Institute, as well as these Proceedings.

David Burke
Lance Dixon
David Leith

TTP96–18

May 1996

Theory of Top Quark Production and Decay[†]

J.H. Kühn

Institut für Theoretische Teilchenphysik, Universität Karlsruhe
D-76128 Karlsruhe, Germany
email: johann.kuehn@physik.uni-karlsruhe.de

Abstract

Direct and indirect information on the top quark mass and its decay modes is reviewed. The theory of top production in hadron- and electron-positron-colliders is presented.

[†]Supported by BMBF 057KA92P and Volkswagen-Stiftung grant I/70452.

Contents

Introduction **3**

1 The Profile of the Top Quark **5**

1.1 Indirect information 6

1.1.1 Indirect evidence for the top quark 6

1.1.2 Mass limits and indirect mass determinations 11

1.1.3 The quadratic top mass dependence of $\delta\rho$ 13

1.2 Top Decays 20

1.2.1 Qualitative aspects – Born approximation 20

1.2.2 Radiative corrections to the rate 23

1.2.3 Decay spectra and angular distributions 26

1.2.4 Non–standard top decays 27

2 Top quarks at hadron colliders **35**

2.1 Lowest order predictions and qualitative features 35

2.2 QCD and electroweak corrections 40

2.2.1 Next to leading order (NLO) corrections and resummation of large logarithms 41

2.2.2 Threshold behaviour 47

2.2.3 Electroweak corrections 50

2.2.4 Gluon radiation 50

2.3 Single top production 52

2.4 Quarkonium production 57

3 Top quarks in e^+e^- annihilation **62**

3.1 Top production above threshold 64

3.1.1 Born predictions 64
3.1.2 Radiative corrections 65
3.1.3 Top quark fragmentation 69
3.1.4 Static t parameters 72
3.1.5 Normal polarization of the top quarks 72
3.1.6 Angular correlations of $t\bar{t}$ decay products 72
3.1.7 Testing the Yukawa Coupling 73
3.2 Threshold behaviour 73
3.2.1 Introductory remarks 76
3.2.2 The QCD potential 80
3.2.3 Realistic predictions for $\sigma_{t\bar{t}}$. 84
3.2.4 Momentum distributions of top quarks 87
3.2.5 Angular distributions and polarization 96
3.2.6 Rescattering 103
3.2.7 Relativistic corrections 103

Acknowledgments **106**

Bibliography **107**

Introduction

An extensive search for top quarks has been performed at electron-positron and hadron colliders for more than a decade. First evidence for top quark production in proton-antiproton collisions has been announced by the CDF collaboration in the spring of 1994. After collecting more luminosity subsequently both the CDF and the D0 experiments presented the definite analysis [1] which demonstrated not only the existence of the anticipated quark but at the same time also provided a kinematic determination of the top quark mass around 180 GeV and a production cross section consistent with the QCD predictions. The mass value is in perfect agreement with the indirect mass determinations based on precision measurements [2-7] of the electroweak parameters in e^+e^-annihilation and in lepton-nucleon scattering. Exploiting the quadratic top mass dependence of radiative corrections an indirect mass measurement of 180 GeV with a present uncertainty of roughly 20 GeV has been achieved.

The top quark completes the fermionic spectrum of the Standard Model. All its properties are uniquely fixed after the mass has been determined. However, as a consequence of its large mass and decay rate it will behave markedly different compared to the remaining five lighter quarks.

It is not just the obvious aim for completion which raises the interest in the top quark. With its mass comparable to the scale of the electroweak symmetry breaking it is plausible that top quark studies could provide an important clue for the understanding of fermion mass generation and the pattern of Yukawa couplings. In fact, it has been suggested that a top quark condensate could even be responsible for the mechanism of spontaneous symmetry breaking [8].

These lectures will be mainly concerned with top quark phenomenology within the framework of the Standard Model (SM). The precise understanding of its production and decay constitutes the basis of any search for deviations or physics beyond the SM.

The properties of the top quark will be covered in chapter 1. Direct and indirect determinations of its decay rates, decay distributions including QCD and electroweak corrections and decay modes predicted in supersymmetric extensions will be discussed. Top quark production at hadron colliders will be the subject of chapter 2. The production cross section and momentum distributions are important ingredients in any of the present analysis. An alternative reaction, namely top quark production through W-b-fusion allows to determine the W-b-$\bar{t}$ coupling and thus indirectly the top quark decay rate.

Perspectives for top studies at a future e^+e^-collider will be presented in chapter 3. An

accurate determination of the top quark mass and its width to better than 1 GeV with a relative accuracy of about 10% seems feasible, and the electroweak couplings of the top quark could be precisely measured with the help of polarized beams. Of particular interest is the interplay between the large top quark decay rate and the binding through the QCD potential which will be also covered in chapter 3.

Chapter 1

The Profile of the Top Quark

Hadron collider experiments at the TEVATRON have firmly established the existence of the top quark and already provide a fairly accurate determination of its mass. The couplings of the top quark to the gauge bosons are uniquely fixed by the SM. Thus all its properties — its production cross section and its decay rate and distributions — can be predicted unambiguously.

The study of real top quarks at high energy colliders, in particular the observation of a peak in the invariant mass of its decay products, is certainly the most impressive proof of existence. Nevertheless, the indirect evidence for a top quark and the determination of its mass is not only of historical interest. The arguments which anticipated the existence of the top quark and its mass around 180 GeV illustrate the rigid structure of the SM, its selfconsistency and beauty. They will be presented in section 1.1.

These theoretical arguments have inspired the experimental searches. The upper limit on the top mass around 200 GeV deduced already relatively early from electroweak precision studies has provided encouragement that energies of present colliders were suited to complete this enterprise. The agreement between the most recent indirect mass determinations both through radiative corrections and through direct observation strengthens the present belief into the quantum field aspect of the theory. It furthermore justifies the corresponding line of reasoning concerning the search for the last remaining ingredient of the SM, the Higgs boson. Section 1.1 of this first chapter will, with this motivation in mind, be devoted to a discussion of the indirect information on the top quark, its existence and its mass. Top decays, including aspects of radiative corrections, polarisation effects and decays induced by physics beyond the SM will be covered in section 1.2.

1.1 Indirect information

1.1.1 Indirect evidence for the top quark

Several experimental results already prior to its discovery did provide strong evidence that the fermion spectrum of the Standard Model

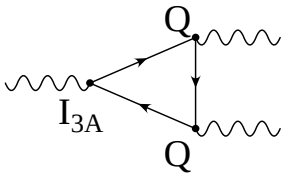
$$\begin{bmatrix}\nu_e\\ e^-\end{bmatrix}_L \quad e_R^- \qquad \begin{bmatrix}\nu_\mu\\ \mu^-\end{bmatrix}_L \quad \mu_R^- \qquad \begin{bmatrix}\nu_\tau\\ \tau^-\end{bmatrix}_L \quad \tau_R^-$$

$$\begin{bmatrix}u\\ d\end{bmatrix}_L \quad u_R \qquad \begin{bmatrix}c\\ s\end{bmatrix}_L \quad c_R \qquad \begin{bmatrix}t\\ b\end{bmatrix}_L \quad t_R$$

does include the top quark, imprinting the same multiplet structure on the third family as the first two families. The evidence is based on theoretical selfconsistency (absence of anomalies), the absence of flavour changing neutral currents (FCNC) and measurements of the weak isospin of the b quark which has been proved to be non-zero, $I_3 = -1/2$, thus demanding an $I_3 = +1/2$ partner in this isospin multiplet.

Absence of triangle anomalies

A compelling argument for the existence of top quarks follows from a theoretical consistency requirement. The renormalizability of the Standard Model demands the absence of triangle anomalies. Triangular fermion loops built-up by an axialvector charge $I_{3A} = -I_{3L}$ combined with two electric vector charges Q would spoil the renormalizability of the gauge theory. Since the anomalies do not depend on the masses of the fermions circulating in the loops, it is sufficient to demand that the sum



$$\sim \sum_L I_{3A} Q^2 = -\sum_L I_3 \left[I_3 + \frac{1}{2}Y\right]^2$$
$$\sim \sum_L Y \sim \sum_L Q$$

of all contributions be zero. Such a requirement can be translated into a condition on the electric charges of all the left-handed fermions

$$\sum_L Q = 0. \tag{1.1}$$

This condition is met in a complete standard family in which the electric charges of the lepton plus those of all color components of the up and down quarks add up to zero,

$$\sum_L Q = -1 + 3 \times \left[\left(+\frac{2}{3}\right) + \left(-\frac{1}{3}\right)\right] = 0.$$

If the top quark were absent from the third family, the condition would be violated and the Standard Model would be theoretically inconsistent.

Absence of FCNC decays

Mixing between quarks which belong to different isospin multiplets

$$\begin{bmatrix} c \\ s' \end{bmatrix}_L \quad b'_L \qquad\qquad \begin{aligned} s'_L &= s_L \cos\vartheta' + b_L \sin\vartheta' \\ b'_L &= -s_L \sin\vartheta' + b_L \cos\vartheta' \end{aligned}$$

generates non-diagonal neutral current couplings, i.e. the breaking of the GIM mechanism

$$\begin{aligned} < I_3 > &= +\frac{1}{2}(\bar{c}_L, c_L) - \frac{1}{2}\left(\bar{s}'_L, s'_L\right) \\ &= \frac{1}{2}(\bar{c}_L, c_L) - -\frac{1}{2}(\bar{s}_L, s_L)\cos^2\vartheta' - \frac{1}{2}\left(\bar{b}_L, b_L\right)\sin^2\vartheta' \\ &\quad -\frac{1}{2}\sin\vartheta'\cos\vartheta'\left((\bar{s}_L, b_L) + \left(\bar{b}_L, s_L\right)\right). \end{aligned}$$

The non-diagonal current induces flavor-changing neutral lepton pair decays $b \to s + l^+l^-$ which have been estimated to be a substantial fraction of all semileptonic B meson decays. The relative strenth of neutral versus charged current induced rate is essentially given by

$$\frac{\Gamma_{\rm NC}}{\Gamma_{\rm CC}} \sim \frac{1}{2}\left(\frac{M_W^2}{M_Z^2}\right)^2 \frac{(v_b^2 + a_b^2)(v_e^2 + a_e^2)}{(1+1)(1+1)} \sim 0.06. \tag{1.2}$$

Taking the proper momentum dependence of the matrix element and the phase space into account one finds [9]

$$\frac{\mathrm{BR}\,(B \to l^+l^- X)}{\mathrm{BR}\,(B \to l^+\nu_l X)} \geq 0.12. \tag{1.3}$$

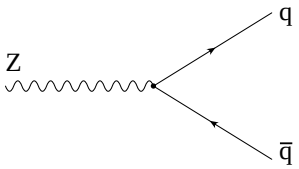
This ratio is four orders of magnitude larger than a bound set by the UA1 Collaboration [10, 11]

$$\frac{\mathrm{BR}\,(B \to \mu^+\mu^- X)}{\mathrm{BR}\,(B \to \mu\nu_\mu X)} < \frac{5.0 \times 10^{-5}}{0.103 \pm 0.005}. \tag{1.4}$$

so that the working hypothesis of an isosinglet b quark is clearly ruled out experimentally also by this method.

Partial width $\Gamma(Z \to b\bar{b})$ and forward-backward asymmetry of b quarks

The Z boson couples to quarks through vector and axial–vector charges with the well–known strength



$$= \sqrt{\frac{G_F m_Z^2}{2\sqrt{2}}}\gamma_\mu\left[v_q - a_q\gamma_5\right].$$

Depending on the isospin assignment of righthanded and lefthanded quark fields these charges are defined as

$$v_q = 2(I_q^{3L} + I_q^{3R}) - 4e_q \sin^2 \vartheta_W \tag{1.5}$$

$$a_q = 2(I_q^{3L} - I_q^{3R}) \tag{1.6}$$

For the present application the *Born approximation* in the massless limit provides an adequate representation of the partial Z decay rate

$$\Gamma_B(Z \to b\bar{b}) \approx \frac{G_F m_Z^3}{8\sqrt{2}\pi} \beta \left(v_b^2 + a_b^2\right). \tag{1.7}$$

In the Standard Model $2I_q^{3R} = 0$ and $2I_q^{3L} = \pm 1$ for up/down quarks respectively.

The ratio between the predictions in the context of a topless model and the SM amounts to

$$\frac{\Gamma^{\text{topless}}}{\Gamma^{\text{SM}}} = \frac{(4Q_b \sin^2 \theta_W)^2}{(1 + 4Q_b \sin^2 \theta_W)^2 + 1} \approx 1/13 \tag{1.8}$$

whereas theory and LEP experiments are well consistent

$$\frac{\Gamma(Z \to b\bar{b})}{\Gamma(Z \to \text{had})} = \begin{cases} 0.2155 \pm 0.0004 & \text{theory} \quad [3] \\ 0.2219 \pm 0.0017 & \text{experiment} \quad [3] \end{cases} \tag{1.9}$$

ruling out the $I = 0$ assignement for the b-quark. The forward-backward asymmetry at the Z resonance

$$A^{\text{FB}} = \frac{3}{4} A_e A_b \tag{1.10}$$

with

$$A_f \equiv \frac{2 v_f a_f}{v_f^2 + a_f^2} \tag{1.11}$$

is sensitive toward the relative size of vector and axial b quark couplings. Up to a sign, the first of these factors, $A_e \approx 0.15$, can be interpreted as the degree of longitudinal Z polarisation, $P_Z = -A_e$, which is induced by the electron coupling even for unpolarized beams. For longitudinally polarized beams it can be replaced by unity. The second factor represents essentially the analyzing power of b quarks. With a predicted value of 0.93 it is close to its maximum in the SM. In fact, this high analysing power is the reason for the large sensitivity of A^{FB} toward $\sin^2 \theta_W$ [12]. For a fictitious topless model A_b is zero. The most recent experimental results from LEP and SLC are displayed in Fig. 1.1.

A remaining sign ambiguity is finally resolved by the interference between NC and electromagnetic amplitude. It leads to a forward backward asymmetry at low energies

$$A_{\text{FB}} = -\frac{3 G_F s}{16\sqrt{2}\pi\alpha} \frac{a_e a_b}{Q_e Q_b} \tag{1.12}$$

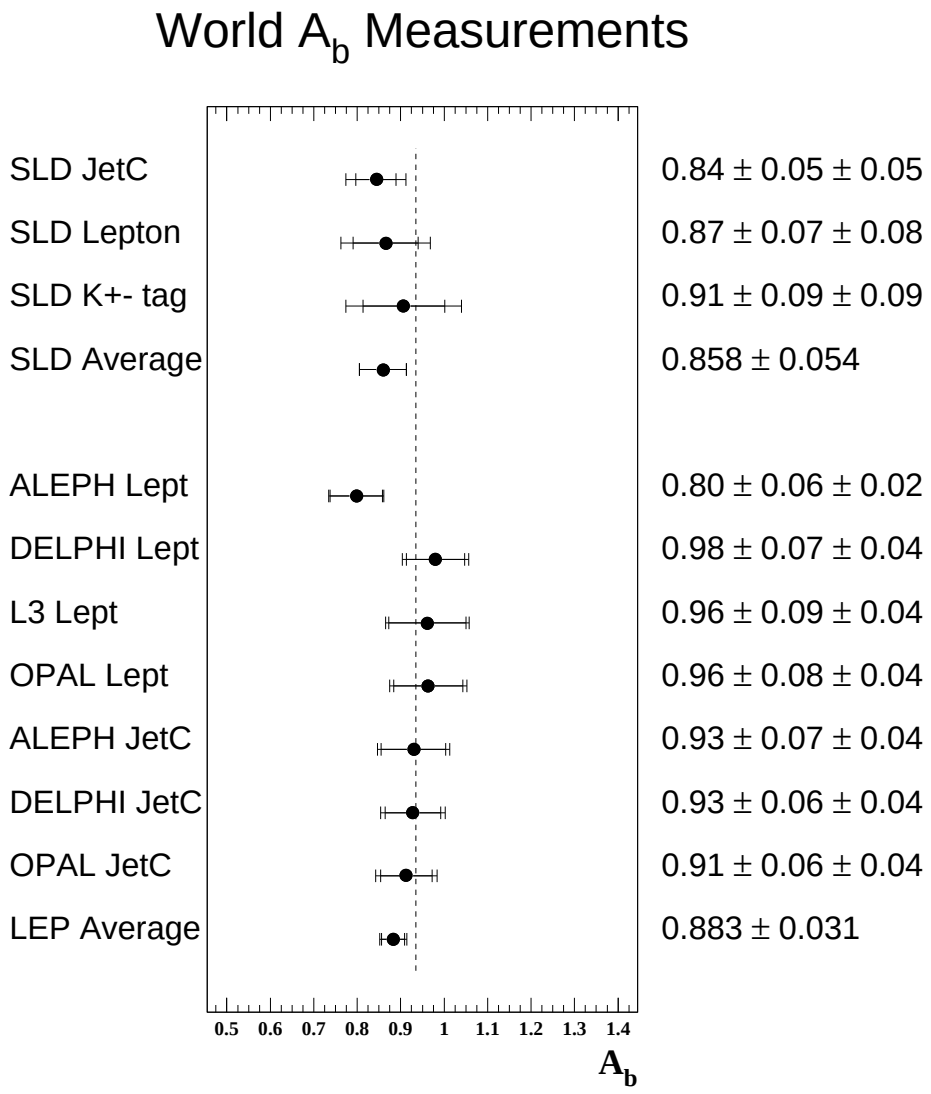


Figure 1.1: Experimental results for the asymmetry parameter A_b [13].

which has been studied in particular at PEP, PETRA, and most recently with highest precision at Tristan at a cm energy of 58 GeV [14]. Using the data available shortly after the turnon of LEP and combining $\Gamma_{b\bar{b}}$, A_b, and $A_{\rm FB}$

$$\begin{aligned} \{I_3^L(b)\}_{exp} &= -0.490^{+0.015}_{-0.012} \quad \rightarrow \quad I_3^L(b) = -1/2 \\ \{I_3^R(b)\}_{exp} &= -0.028 \pm 0.056 \quad \rightarrow \quad I_3^R(b) = 0 \end{aligned}$$

has been obtained already some time ago [15]. As shown in Fig. 1.2 all measurements are nicely consistent with the predictions of the SM[1]. The isospin assignment of the Standard Model is thus well confirmed.

Figure 1.2: The weak isospins $I_3^L(b)$ and $I_3^R(b)$ of the left– and right–handed b quark components, extracted from the data on $\Gamma(Z \rightarrow b\bar{b})$ and $A_{FB}(b)$ at LEP, and PETRA/PEP and TRISTAN, Ref.[15].

[1]For a discussion of the most recent results for R_b, however, see [3].

1.1.2 Mass limits and indirect mass determinations

Theoretical constraints

Present theoretical analyses of the Standard Model are based almost exclusively on perturbation theory. If this method is assumed to apply also to the top-quark sector, in particular when linked to the Higgs sector, the top mass must be bounded as the strength of the Yukawa-Higgs-top coupling is determined by this parameter. The following consistency conditions must be met:

Perturbative Yukawa coupling $g_Y\,(ttH)$

Defining the Yukawa coupling in the Standard Model through

$$\mathcal{L}_Y = g_Y \left(\frac{v+H}{\sqrt{2}}\right) (\bar{t}_L t_R + \text{h.c.}) \tag{1.13}$$

the coupling constant g_Y is related to the top mass by

$$g_Y\,(ttH) = m_t \sqrt{2\sqrt{2} G_F}. \tag{1.14}$$

Demanding the effective expansion parameter $g_Y^2/4\pi$ to be smaller than 1, the top mass is bounded to

$$m_t < \sqrt{\frac{\sqrt{2}\pi}{G_F}} \approx 620\text{GeV}. \tag{1.15}$$

For a top mass of 180 GeV the coupling $g_Y^2/4\pi \approx 0.085$ is comfortably small so that perturbation theory can safely be applied in this region.

Unitarity bound

At asymptotic energies the amplitude of the zeroth partial wave for elastic $t\bar{t}$ scattering in the color singlet same-helicity channel [16]

$$\begin{aligned} a_0\,(t\bar{t} \to t\bar{t}) &= -\frac{3g_Y^2}{8\pi} \\ &= -\frac{3G_F m_t^2}{4\sqrt{2}\pi} \end{aligned} \tag{1.16}$$

grows quadratically with the top mass. Unitarity however demands this real amplitude to be bounded by $|\text{Re}\, a_0| \leq 1/2$. This condition translates to

$$m_t < \sqrt{\frac{2\sqrt{2}\pi}{3G_F}} \approx 500\,\text{GeV}. \tag{1.17}$$

The bound improves by taking into account the running of the Yukawa coupling [17]. These arguments are equally applicable for any additional species of chiral fermions with mass induced via spontaneous symmetry breaking.

Stability of the Higgs system: top-Higgs bound

The quartic coupling λ in the effective Higgs potential

$$V = \mu^2 |\phi|^2 + \frac{\lambda}{2} |\phi|^4$$

depends on the scale at which the system is interacting. The running of λ is induced by higher-order loops built-up by the Higgs particles themselves, the vector bosons and the fermions in the Standard Model [17, 18]. For moderate values of the top mass, $m_t \leq 77$ GeV, these radiative corrections would have generated a lower bound on the Higgs mass of 7 GeV. With the present experimental lower limit $m_H > 65$ GeV and the top quark mass determined around 180 GeV this bound is of no practical relevance any more. At high energies the radiative corrections make λ rise up to the Landau pole at the cut-off parameter Λ beyond which the Standard Model in the present formulation cannot be continued ["triviality bound", as this bound could formally be misinterpreted as requiring the low energy coupling to vanish]. If for a fixed Higgs mass the top mass is increased, the top loop radiative corrections lead to negative values of the quartic coupling λ

$$\frac{\partial \lambda}{\partial \log s} = \frac{3}{8\pi^2} \left[\lambda^2 - 4 g_Y^4 + \text{gauge couplings} \right] \tag{1.18}$$

Since the potential is unbounded from below in this case, the Higgs system becomes instable. Thus the stability requirement defines an upper value of the top mass m_t for a given Higgs mass m_H and a cut-off scale Λ. The result of such an analysis is presented in Fig.1.3. Depending on the cut-off scale Λ where new physics may set in, the top mass is bounded to $m_t \leq 200$ GeV if Λ exceeds the Planck scale but it rises up to 400 to 500 GeV if the cut-off is reached at a level of 1 TeV and below. The estimates are similar to the unitarity analysis in the preceeding subsection. Lattice simulations of the Yukawa model have arrived at qualitatively similar results (see e.g. [19]).

These theoretical analyses have shown that for the top mass around 180 GeV the Standard Model may be valid up to a cut-off at the Planck scale. [The hierarchy problem, that is not touched in the present discussion, may enforce nevertheless new additional physical phenomena already in the TeV range.]

In the context of the SM the top Yukawa coupling is simply present as a free parameter. In the minimal supersymmetric model, however, the picture is changed completely. A large Yukawa coupling may play the role of a driving term for the spontaneous breaking of $SU(2) \times U(1)$, as discussed in [20] and in fact the large mass of the top quark had been predicted on the basis of these arguments prior to its experimental discovery.

Mass estimates from radiative corrections

First indications of a high top quark mass were derived from the rapid $(B - \overline{B})$ oscillations observed by ARGUS [21]. However, due to the uncertainties of the CKM matrix element V_{td} and of the $(b\bar{d})$ wave function, not more than qualitative conclusions can be drawn from such an analysis as the oscillation frequency $\Delta m \sim |V_{td}|^2 f_B^2 m_t^2$ depends on three [unknown] parameters.

Figure 1.3: Bounds on the Higgs and top masses following from triviality of the Higgs's quartic self-coupling and the stability of the Higgs system; from [18].

The analysis of the radiative corrections to high precision electroweak observables is much more advanced [2-7]. Since Higgs mass effects are weak as a result of the screening theorem [22], the top mass is the dominant unknown in the framework of the Standard Model. Combining the high precision measurements of the Z mass with $\sin^2\theta_W$ from the Z decay rate, from the forward-backward asymmetry and from LR polarization measurements, the top quark mass has been determined up to a residual uncertainty of less than 10 GeV plus an additional uncertainty of about 20 GeV induced through the variation of m_H

$$m_t = 178 \pm 8^{+17}_{-20}\ \text{GeV}. \tag{1.19}$$

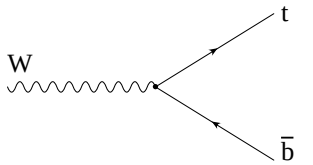
The close agreement between direct and indirect top mass determination can be considered a triumph of the Standard Model. Its predictions are not only valid in Born approximation, as expected for any effective theory, also quantum corrections play an important role, and are indirectly confirmed. Encouraged by this success and in view of the improved accuracy of theory and experiment it is conceivable that the same strategy can lead to a rough determination of the Higgs mass, or, at least, to a phenomenologically relevant upper limit.

1.1.3 The quadratic top mass dependence of $\delta\rho$

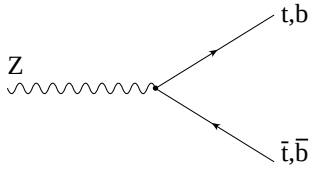
The quadratic top mass dependence of $\delta\rho$ is a cornerstone of the present precision measurements [23]. In view of its importance and the pedagogical character of these lectures

it is perhaps worthwhile to present a fairly pedestrian derivation of this result.

Let us first consider the definition of the weak mixing angle in Born approximation. It can be fixed through the relative strength of charged vs. neutral current couplings:



$$= \frac{g}{2\sqrt{2}}\gamma_\mu(1-\gamma_5) \tag{1.20}$$



$$= \frac{g}{4\cos\bar{\theta}_W}\left[\left(2I_3^f - 4Q_f \sin^2\bar{\theta}_W\right)\gamma_\mu + 2I_3^f\gamma_\mu\gamma_5\right] \tag{1.21}$$

with the SU(2) coupling related to the electromagnetic coupling through $g = e/\sin\bar{\theta}_W$. Alternatively $\sin^2\theta_W$ is defined through the mass ratio

$$M_W^2/M_Z^2 = 1 - \sin^2\theta_W \tag{1.22}$$

These two definitions coincide in Born approximation

$$\theta_W = \bar{\theta}_W\Big|_{\text{Born}} \tag{1.23}$$

However, the self energy diagrams depicted in Fig. 1.4 lead to marked differences between the two options, in particular if $m_t \gg M_{W,Z} \gg m_b$.

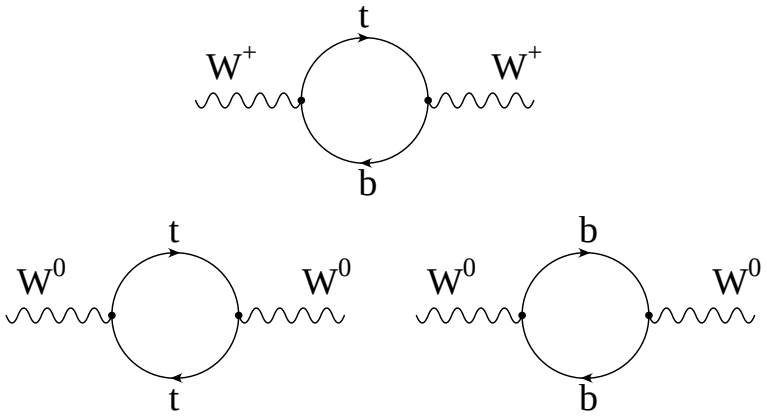
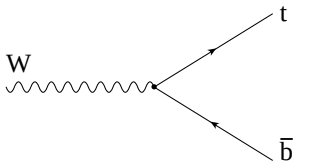


Figure 1.4: Self energy diagrams which induce the mass splitting between $W^\pm$ and W^0 and a ρ parameter different from one.

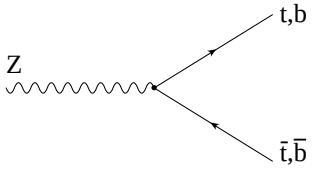
This difference can be traced to a difference in the mass shift for the W and the Z boson. For a simplified discussion consider, in a *first step*, $\sin^2\theta_W = 0$ and hence the $SU(2)$ part of the theory only. The neutral boson will be denoted by W^0. In the lowest order this implies

$$M_- = M_0 \equiv M \tag{1.24}$$

and the couplings simplify to



$$= \frac{g}{2\sqrt{2}}\gamma_\mu(1-\gamma_5) \tag{1.25}$$



$$= \frac{g}{4}2I_3^f\gamma_\mu(1-\gamma_5) \tag{1.26}$$

In order g^2 the propagators of charged and neutral bosons are modified by the self energies

$$\begin{aligned} W^+ : \quad & \frac{1}{M^2-s} \;\Rightarrow\; \frac{1}{M^2-s-\Pi_+(s)} \\ W^0 : \quad & \frac{1}{M^2-s} \;\Rightarrow\; \frac{1}{M^2-s-\Pi_0(s)} \end{aligned} \tag{1.27}$$

The mass shifts individually are given by

$$\delta M^2_{+,0} = M^2_{+,0} - M^2 = -Re\,\Pi_{+,0}(M^2). \tag{1.28}$$

They are most easily calculated through dispersion relations from their respective imaginary parts. These can be interpreted as the "decay rate" of a fictitious virtual boson of mass κ:

$$\begin{aligned} \kappa\Gamma(W^*_+ \to t+\bar{b}) &= Im\,\Pi_+(\kappa^2) \\ \delta M^2_+ &= \frac{1}{\pi}\int \frac{d\kappa^2}{\kappa^2-M^2} Im\,\Pi_+(\kappa^2). \end{aligned} \tag{1.29}$$

The decay rates of the virtual bosons are easily calculated ($m \equiv m_{top}$)

$$\kappa\Gamma(W^*_+ \to t\bar{b}) = \frac{3}{12\pi}\left(\frac{g}{2\sqrt{2}}\right)^2 \kappa^2 2\left(1-\frac{m^2}{2\kappa^2}-\frac{m^4}{2\kappa^4}\right)\left(1-\frac{m^2}{\kappa^2}\right)\Theta(\kappa^2-m^2). \tag{1.30}$$

The factor 3 originates from color, the factor 2 from the identical vector and axial contributions, the squared matrix element and the phase space are responsible for the second and third factors in brackets respectively.

Similarly one finds

$$\begin{aligned} &\kappa\Gamma(W^*_0 \to b\bar{b}) + \kappa\Gamma(W^*_0 \to t\bar{t}) = \\ &\quad \frac{3}{12\pi}\left(\frac{g}{4}\right)^2 \kappa^2 \left\{ 2\Theta(\kappa^2) + 2\left(1-\frac{m^2}{\kappa^2}\right)\left(1-\frac{4m^2}{\kappa^2}\right)^{1/2}\Theta(\kappa^2-4m^2)\right\} \\ &\delta M^2_0 = \frac{1}{\pi}\int \frac{d\kappa^2}{\kappa^2-M^2} Im\,\Pi_0(\kappa^2). \end{aligned} \tag{1.31}$$

With the large κ^2 behaviour of $\mathrm{Im}\,\Pi$ given by $\mathrm{Im}\,\Pi \sim \kappa^2$ the dispersive integral eq. (1.29) is evidently quadratically divergent. In the limit of large κ the leading ($\sim \kappa^2$) and next-to-leading ($\to$ const) terms of eqs. (1.30) and (1.31) coincide. The leading and next to leading divergences can therefore be absorbed in a $SU(2)$ invariant mass renormalization. The relative mass shift, however, the only quantity accessible to experiment, remains finite and is given by

$$\begin{aligned}
\delta M^2 &= \delta M_+^2 - \delta M_0^2 \\
&= \frac{3}{12\pi}\frac{g^2}{8}\frac{1}{\pi}\int_0^\infty \frac{d\kappa^2}{\kappa^2 - M^2}\kappa^2 \\
&\quad \times \begin{cases} -2\left(1-\frac{m^2}{\kappa^2}\right)\left(1-\frac{m^2}{2\kappa^2}-\frac{m^4}{2\kappa^4}\right) & \times\Theta(\kappa^2-m^2) \\ +1 & \times\Theta(\kappa^2) \\ \left(1-\frac{m^2}{\kappa^2}\right)\left(1-\frac{4m^2}{\kappa^2}\right)^{1/2} & \times\Theta(\kappa^2-4m^2) \end{cases}
\end{aligned} \tag{1.32}$$

We are only interested in the leading top mass dependence: $m^2 \gg M^2$. The leading term is obtained by simply setting $M^2 \to 0$ in the integrand. Introducing a cutoff $\Lambda^2 \gg m^2$ the leading contributions to the three integrals are given by

$$\begin{cases} -2\Lambda^2 + 3m^2 \ln\dfrac{\Lambda^2}{m^2} + \dfrac{3}{2}m^2 + \ldots \\ +\Lambda^2 \\ +\Lambda^2 - 3m^2 \ln\dfrac{\Lambda^2}{m^2} \end{cases} \tag{1.33}$$

and hence

$$\delta M^2 = \frac{3}{12\pi^2}\frac{g^2}{8}\frac{3}{2}m^2 \tag{1.34}$$

Up to the proportionality constant this result could have been guessed on dimensional grounds from the very beginning.

It has become customary to express the $SU(2)$ coupling in terms of G_F and the W mass

$$\frac{g^2}{8} = \frac{G_F}{\sqrt{2}}M_W^2 \tag{1.35}$$

such that

$$\frac{\delta M^2}{M^2} = \frac{3}{16\pi^2}\sqrt{2}G_F m^2. \tag{1.36}$$

The ratio of neutral versus charged current induced amplitude at small momentum transfers is thus corrected by a factor

$$\frac{M_+^2}{M_0^2} \equiv 1 + \delta\rho = \left(1 + \frac{\delta M^2}{M^2}\right) \tag{1.37}$$

with $\delta\rho$ given in eq. (1.36).

To discuss the phenomenological implications of this result it is now necessary to reintroduce the weak mixing between the neutral $SU(2)$ and $U(1)$ gauge bosons. The gauge boson masses are induced by the squared covariant derivative acting on the Higgs field

$$D_\mu \phi \to -i \left(g \vec{W}_\mu \frac{\vec{\tau}}{2} + g' B Y \right) \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{1.38}$$

giving rise to the following mass terms in the Lagrangian

$$\begin{aligned} \mathcal{L}_M &= \frac{v^2}{8}(g^2+g'^2)\left(\frac{g}{\sqrt{g^2+g'^2}}W_3 - \frac{g'}{\sqrt{g^2+g'^2}}B\right)^2 \\ &\quad + \frac{v^2}{8} g^2 (W_1^2 + W_2^2)(1+\delta\rho) \end{aligned} \tag{1.39}$$

The last term has been added to represent a contribution from a non vanishing $\delta\rho$, induced e.g. by the large top mass. The finite mass shift has been without loss of generality entirely attributed to the charged W boson.

The mass eigenstates are easily identified from eq. (1.39)

$$\begin{aligned} W^\pm &= (W_1 \mp i W_2)/\sqrt{2} \\ Z &= (\bar{c} W_3 - \bar{s} B) \\ A &= (\bar{c} B + \bar{s} W_3) \end{aligned} \tag{1.40}$$

with the weak mixing angle $\bar{\theta}_W$

$$\begin{aligned} \bar{c} \equiv \cos\bar{\theta}_W &\equiv \frac{g}{\sqrt{g^2+g'^2}} \\ \bar{s} \equiv \sin\bar{\theta}_W &\equiv \frac{g'}{\sqrt{g^2+g'^2}} \end{aligned} \tag{1.41}$$

defined through the *couplings*. This definition is, of course, very convenient for measurements at LEP, where couplings are determined most precisely. The couplings of the photon and the Z boson are thus also given in terms of $\bar{s}$. The masses are read off from eq. (1.39)

$$\begin{aligned} m_A^2 &= 0 \\ m_Z^2 &= \frac{v^2}{4}(g^2+g'^2) \\ m_W^2 &= \frac{v^2}{4} g^2 (1+\delta\rho) \end{aligned} \tag{1.42}$$

and

$$\frac{m_W^2}{m_Z^2} = \bar{c}^2 (1+\delta\rho) \tag{1.43}$$

which constitutes the standard definition of the ρ parameter. Alternatively one may define the mixing angle directly through the mass ratio

$$s^2 \equiv \sin^2\theta_W \equiv 1 - M_W^2/M_Z^2 \tag{1.44}$$

The two definitions coincide in the Born approximation; they differ, however, for $\delta\rho \neq 0$:

$$s^2 = 1 - \bar{c}^2(1+\delta\rho) = \bar{s}^2 - \bar{c}^2\delta\rho \tag{1.45}$$

It is, of course, a matter of convention and convenience, which of the two definitions (or their variants) are adopted. The choice of input parameters and observables will affect the sensitivity towards $\delta\rho$ — and hence towards m_t. The observables which are measured with the highest precision at present and in the forseable future are the fine-structure constant α, the muon life time which provides a value for G_F and the Z boson mass. To obtain the dependence of $\sin^2\theta_W$ on $\delta\rho$ we predict M_W^2 from these observables. We start from

$$\frac{g^2}{8}\frac{1}{M_W^2} = \frac{G_F}{\sqrt{2}} \tag{1.46}$$

and express g^2 through G_F and M_Z^2

$$g^2 = 4\sqrt{2}G_F M_Z^2 c^2 \tag{1.47}$$

Alternatively g can be related to the fine structure constant

$$g^2 = e^2/\bar{s}^2 \tag{1.48}$$

Note the appearance of c^2 in eq. (1.44) and of $\bar{s}^2$ in eq. (1.41).

One thus arrives at

$$e^2 = 4\sqrt{2}G_F^2 M_Z^2 c^2(s^2 + c^2\delta\rho) \tag{1.49}$$

or, equivalently, at

$$\frac{4\pi\alpha}{4\sqrt{2}G_F M_Z^2}\frac{1}{1+\cot^2\theta_W\delta\rho} = s^2c^2. \tag{1.50}$$

Solving for $\sin^2\theta_W$ (defined through the mass ratio) one obtains on one hand

$$\sin^2\theta_W \approx \frac{1}{2}\left[1 - \sqrt{1 - \frac{4\pi\alpha}{\sqrt{2}G_F M_Z^2}}\right] - \frac{c^4}{c^2-s^2}\delta\rho \tag{1.51}$$

where the Born values for c^2 and s^2 can be taken in the correction term. The definition of $\sin^2\bar{\theta}_W$ through the relative strength of $SU(2)$ and $U(1)$ couplings leads on the other hand to

$$\sin^2\bar{\theta}_W \approx \frac{1}{2}\left[1 - \sqrt{1 - \frac{4\pi\alpha}{\sqrt{2}G_F M_Z^2}}\right] - \frac{s^2c^2}{c^2-s^2}\delta\rho \tag{1.52}$$

For the actual evaluation the running coupling $\alpha(M_Z)^{-1} \approx 129$ must be employed [4]. Eq. (1.51) and eq. (1.52) exhibit rather different sensitivity towards a variation of $\delta\rho$ and

hence of m_t, with a ratio between the two coefficients of $c^2/s^2 \approx 3.3$. For a precise comparison between theory and experiment subleading one-loop corrections must be included, and subtle differences between variants of $\bar{\theta}_W$ must be taken into consideration, with $\theta_{\overline{\rm MS}}$ and $\theta_{\rm eff}^{\rm lept}$ as most frequently used options [2].

With increasing experimental accuracy numerous improvements must be and have been incorporated into the theoretical predictions.

- The full one loop corrections including all subleading terms are know since long and are certainly indispensible (see [4, 2, 24, 25] and references therein).
- Two loop purely weak corrections increase proportional $(G_F m_t^2/16\pi^2)^2$. A detailed discussion can be found in [7].
- QCD corrections are available at the two- and even three-loop level [5, 24, 25].

With α, G_F, and M_Z fixed one may determine m_t either from M_W or alternatively from $\sin^2\bar{\theta}_W$ (corresponding to a measurement of the left right asymmetry with polarised beams, the τ polarisation or forward backward asymmetries of unpolarised beams). These measurements are in beautiful agreement with the determination of m_t in production experiments at the TEVATRON (Fig. 1.5).

Figure 1.5:

1.2 Top Decays

Various aspects of top decays have been scrutinized in the literature. The large top decay rate predicted in the SM governs top quark physics. Radiative correctons from QCD and electroweak interactions have been calculated for the decay rate and for differential distributions of the decay products. Non-standard top decays are predicted in SUSY extensions of the SM, with $t \to Hb$ and $t \to \tilde{t}\tilde{\gamma}$ as most promising and characteristic signatures. Born predictions and radiative corrections (at least in part) have been worked out also for these decay modes. Beyond that a number of even more exotic decay modes, in particular FCNC decays, have been suggested.

1.2.1 Qualitative aspects – Born approximation

The decay of the top quark into $b + W$ is governed by the following amplitude

$$\mathcal{M}(t \to bW) = \frac{ig}{\sqrt{2}} \bar{b} \not{\varepsilon}^W \frac{1-\gamma_5}{2} t \tag{1.53}$$

Adopting the high energy limit ($m_t^2 > M_W^2$) for the polarisation vector ϵ_L of the longitudinal W (corresponding to helicity $h^W = 0$)

$$\varepsilon_L^W = \begin{pmatrix} p_3^W \\ 0 \\ 0 \\ p_0^W \end{pmatrix} \frac{1}{M_W} = \begin{pmatrix} p_0^W \\ 0 \\ 0 \\ p_3^W \end{pmatrix} \frac{1}{M_W} + \mathcal{O}(M_W/m_t) \tag{1.54}$$

the amplitude is dominated by contribution from longitudinal W's

$$\begin{aligned} \mathcal{M}_L &= \frac{ig}{\sqrt{2}} \bar{b} \not{\varepsilon}_L^W \frac{1-\gamma_5}{2} t \approx \frac{ig}{\sqrt{2}} \frac{m_t}{M_W} \bar{b} \frac{1+\gamma_5}{2} t \\ &= i\sqrt{2} \frac{m_t}{v} \bar{b}(1+\gamma_5) t \end{aligned} \tag{1.55}$$

This part is thus proportional to the Yukawa coupling

$$g_Y = \sqrt{2} \frac{m_t}{v} \tag{1.56}$$

with a rate growing proportional m_t^3. In contrast, the amplitude for the decay into transverse W's, is obtained with the polarisation vectors

$$\varepsilon_T^{\pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm i \\ 0 \end{pmatrix} \tag{1.57}$$

and remains constant in the high mass limit. The rate is governed by the gauge coupling g and increases only linearly with m_t. The longitudinal or transversal W is produced in

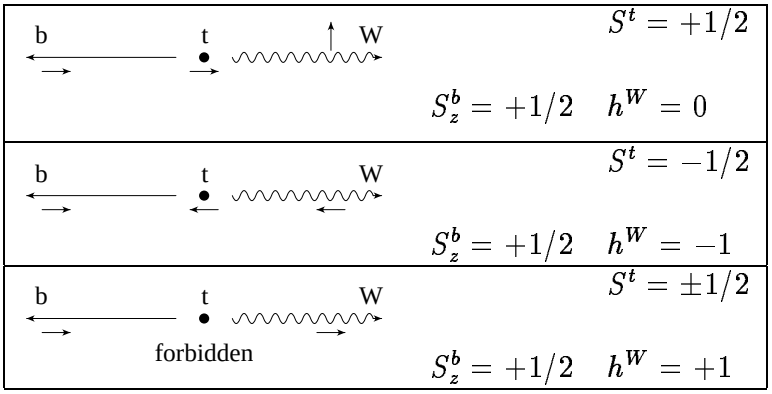


Figure 1.6: Top decays: angular momentum conservation

conjunction with a lefthanded b quark. The production of W's with helicity $h^W = +1$ is thus forbidden by angular momentum conservation (see fig. 1.6).

In total one finds

$$(h^W = -1) : (h^W = 0) : (h^W = +1) = 1 : \frac{m_t^2}{2M_W^2} : 0. \tag{1.58}$$

The implications for the angular distributions of the decay products will be discussed below. The decay rate

$$\begin{aligned} \Gamma &= \frac{G_F m_t^3}{8\sqrt{2}\pi}\left(1 - \frac{M_W^2}{m_t^2}\right)^2\left(1 + 2\frac{M_W^2}{m_t^2}\right) \\ &\approx 175\,\mathrm{MeV}\left(\frac{m_t}{M_W}\right)^3 \end{aligned} \tag{1.59}$$

increases with the third power of the quark mass and, for a *realistic* top mass around 180 GeV amounts to more than 1.5 GeV, exceeding significantly all hadronic scales. Before we discuss the implications of this fact let us briefly pursue the close similarity between the coupling of the longitudinal W to the tb system and the decay into a charged Higgs boson in a two Higgs doublet model (THDM). The decay rate is given by (see also section 1.2.4)

$$\Gamma(t \to H^+ b) = \frac{G_F m_t^3}{8\sqrt{2}\pi}\left(1 - \frac{m_H^2}{m_t^2}\right)^2\left[\left(\frac{m_b}{m_t}\right)^2 \tan^2\beta + \cot^2\beta\right] \tag{1.60}$$

The similarity between this rate and the rate for the decay into longitudinal W's is manifest from the cubic top mass dependence. The minimal value of the term in brackets is assumed for $\tan\beta = \sqrt{m_t/m_b}$. Adopting $m_b(\text{running}) \approx 3$ GeV, the minimal value of the last factor amounts to about 1/30. On the other hand, in any plausible THDM the value of $\tan\beta$ should not exceed m_t/m_b. The W decay mode will hence never be swamped by the Higgs channel. (This fact is of course also implied by the actual observation of the top quark at the TEVATRON.) Up to this point we have, tacitely, assumed the CKM

matrix element V_{tb} to be close to one. In fact, in the three generation SM one predicts (90% CL)

$$\begin{aligned} V_{tb} = 0.9990 \pm 0.0004 &\Rightarrow BR(b) \approx 1 \\ V_{ts} = 0.044 \pm 0.010 &\Rightarrow BR(s) \approx 0.2\% \\ V_{td} = 0.011 \pm 0.009 &\Rightarrow BR(d) \approx 0.01\% \end{aligned} \tag{1.61}$$

on the basis of CKM unitarity. In a four generation model, however, sizeable mixing between third and fourth generation could arise. Methods to determine the strength of V_{tb} either through single top production at a hadron collider or through a direct measurement of Γ_t in e^+e^- colliders will be discussed below in sects. 2.3 and 3.2.

The large top decay rate provides a cutoff for the long distance QCD dynamics. The implications can be summarised in the statement: "t quarks are produced and decay like free quarks" [26]. In particular the angular distributions of their decay products follow the spin 1/2 predictions. This is in marked contrast to the situation for b quarks, with B mesons decaying isotropically. The arguments for this claim are either based on a comparison of energy scales, or, alternatively, on a comparison of the relevant time scales.

Let us start with the first of these two equivalent viewpoints: The mass difference between B^{**} and B mesons amounts to 450 MeV. In the nonrelativistic quark model the B^{**} is interpreted as orbitally excited $b\bar{q}$ state. With increasing mass of the heavy quark this splitting remains approximately constant: it is essentially governed by light quark dynamics. The hyperfine splitting between B^* and B, in contrast, is proportional to the color magnetic moment and hence decreases $\sim 1/m_Q$. Given a decay rate of about 1.5 GeV it is clear that T-, T^*-, and T^{**}- mesons merge and act coherently, rendering any distinction between individual mesons meaningless. In fact even individual toponium states cease to exist. From the perturbative QCD potential an energy difference between $1S$ and $2S$ states around 1.2–1.5 GeV is predicted. This has to be contrasted with the toponium decay rate $\Gamma_{t\bar{t}} \sim 2\Gamma_t \sim 3$ GeV. All resonances merge and result in an excitation curve which will be discussed in chapter 3.

A similar line of reasoning is based on the comparison between different characteristic time scales: The formation time of a hadron from a locally produced t quark is governed by its size which is significantly larger than its lifetime

$$\tau_{\text{Formation}} \approx \text{size} \approx 1/0.5\text{GeV} \gg \tau_{\text{Decay}} \approx 1/\Gamma_t \tag{1.62}$$

Top quarks decay before they have time to communicate hadronically with light quarks and dilute their spin orientation. For sufficiently rapid top quark decay even $t\bar{t}$ bound states cease to exist. The classical time of revolution T_{rev} for a Coulombic bound state is given by $\left(\alpha_{\text{eff}} = \frac{4}{3}\alpha_s\right)$

$$T_{\text{rev}} = \pi\alpha_{\text{eff}}\sqrt{\frac{m_{\text{red}}}{2|E|^3}} \tag{1.63}$$

With $E = -\alpha_{eff}^2 m_{\text{red}}/2$ for the ground state

$$T_{\text{rev}} = \pi/|E| \gg 1/\Gamma_{t\bar{t}} \tag{1.64}$$

is obtained. The lifetime of the $t\bar{t}$ system is too small to allow for the proper definition of a bound state with sharp binding energy.

1.2.2 Radiative corrections to the rate

Perturbative corrections to the lowest-order result affect the total decay rate as well as differential distributions. Their inclusion is a necessary prerequisite for any analysis that attempts a precision analysis of top decays. Both QCD and electroweak corrections are well under control and will be discussed in the following.

QCD corrections

The correction to the *decay rate* is usually written in the form

$$\Gamma = \Gamma_{\rm Born}\,(1+\delta_{QCD}) = \Gamma_{\rm Born}\left(1-\frac{2}{3}\frac{\alpha_s}{\pi}f\right) \tag{1.65}$$

The correction function f has been calculated in [27] for nonvanishing and vanishing b mass. In the limit $m_b^2/m_t^2 \to 0$ the result simplifies considerably, but remains a valid approximation (Fig.1.7):

$$\begin{aligned} f &= \mathcal{F}_1/\mathcal{F}_0 \\ \mathcal{F}_0 &= 2(1-y)^2(1+2y) \\ \mathcal{F}_1 &= \mathcal{F}_0\left[\pi^2 + 2Li_2(y) - 2Li_2(1-y)\right] \\ &\quad +4y(1-y-2y^2)\ln y + 2(1-y)^2(5+4y)\ln(1-y) \\ &\quad -(1-y)(5+9y-6y^2), \end{aligned} \tag{1.66}$$

where $y = m_W^2/m_t^2$. In the limit $m_t^2 \gg m_W^2$ f is well approximated by $f(y) = 2/3\pi^2 - 5/2 - 3y + 9y^2/2 - 3y^2\ln y \approx 4$. For $m_t \approx 180\ GeV$ the QCD correction amounts to

$$\delta_{QCD} \approx -3.7\alpha_s/\pi \tag{1.67}$$

and lowers the decay rate by about 10%. This has a non-negligible impact on the height and width of a toponium resonance or its remnant.

The α_s^2 corrections are presently unknown, and the scale μ in $\alpha_s(\mu^2)$ is uncertain. Indications for a surprisingly large correction of order α_s^2, corresponding to a rather small scale have been obtained recently. Diagrams with light fermion insertions into the gluon propagator have been calculated numerically [28] and analytically [29] in the limit $m_t \gg M_W$

$$\begin{aligned} \Gamma &= \Gamma_{\rm Born}\left[1 - \frac{2}{3}\frac{\alpha_{\overline{\rm MS}}(m_t^2)}{\pi}\left(4\zeta_2 - \frac{5}{2}\right)\right. \\ &\quad \left. + \left(\frac{\alpha_{\overline{\rm MS}}}{\pi}\right)^2\left(-\frac{2n_f}{3}\right)\left(\frac{4}{9} - \frac{23}{18}\zeta_2 - \zeta_3\right)\right] \end{aligned} \tag{1.68}$$

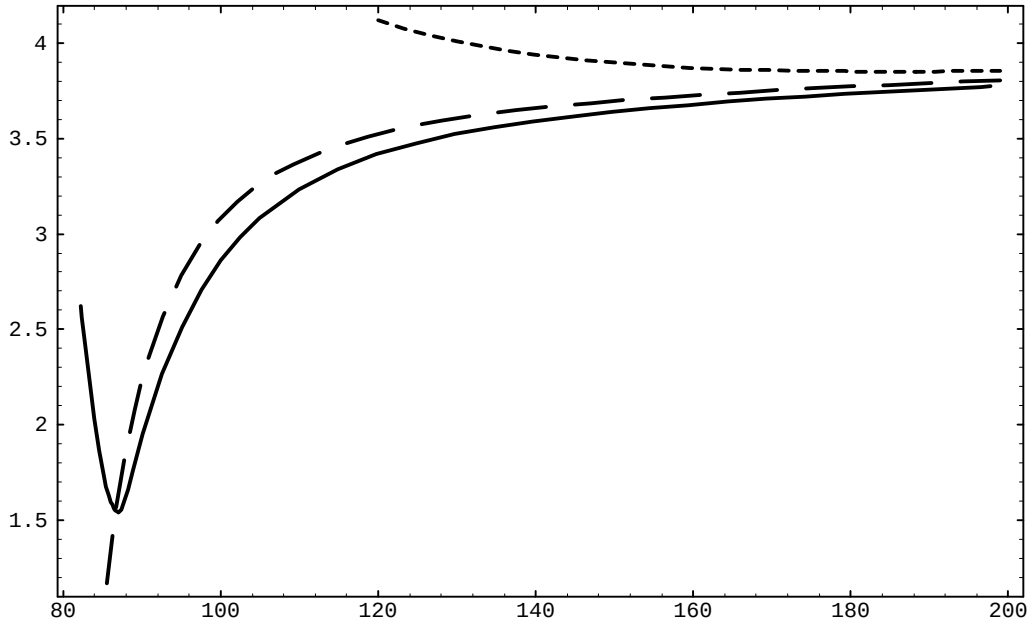


Figure 1.7: QCD correction function f for the top quark decay rate. Solid line: exact form; dashes: $m_b = 0$, dotted: approximate form for large m_t/M_W.

The BLM prescription [30] suggests that the dominant α_s^2 coefficients can be estimated through the replacement

$$-\frac{2n_f}{3} \to \left(11 - \frac{2n_f}{3}\right) \tag{1.69}$$

and absorbed through a change in the scale. For the problem at hand this corresponds to a scale $\mu = 0.122 m_t$ resulting in a fairly large effective value of α_s of 0.15 instead of 0.11 for $\mu = m_t$.

Electroweak corrections

Electroweak corrections to the top quark decay rate can be found in [31, 32]. They involve a large number of diagrams. For asymptotically large top masses the Higgs exchange diagram provides the dominant contribution. Defining the Born term by means of the Fermi coupling G_F, one derives in this limit

$$\begin{aligned} \Gamma &= \Gamma(G_F)_{\rm Born}[1 + \delta_{EW}], \\ \delta_{EW} &= \frac{G_F m_t^2}{4\sqrt{2}\pi^2}\left[\frac{17}{4} + \log\frac{m_H^2}{m_t^2}\right] + \text{subleading terms}. \end{aligned} \tag{1.70}$$

While the Higgs–top coupling is the origin of the strong quadratic dependence on the top mass, the Higgs itself is logarithmically screened in this limit. However, the detailed

analysis reveals that the subleading terms are as important as the leading terms so that finally one observes only a very weak dependence of δ_{EW} on the top and the Higgs masses, Fig.1.8. The numerical value of the corrections turns out to be small, $\delta_{EW} \approx +2\%$. Electroweak corrections in the context of the two Higgs doublet model can be found in [33] and are of comparable magnitude.

Figure 1.8: QCD and electroweak radiative corrections to the top decay width; adapted from Ref.[31].

The positive correction δ_{EW} is nearly cancelled by the negative correction δ_Γ of -1.5% from the nonvanishing finite width of the W. The complete prediction taken from [34] is displayed in table 1.1 for the choice $\alpha_s(\mu^2 = m_t^2)$. For $\mu = 0.112 m_t$ the QCD correction

m_t [GeV]	$\alpha_s(m_t)$	$\Gamma^{\rm Born}_{\rm nar.w.}$ [GeV]	$\delta_\Gamma^{(0)}$ [%]	$\delta^{\rm nw}_{\rm QCD}$ [%]	$\delta_{\rm EW}$ [%]	Γ [GeV]
170	.108	1.41	-1.52	-8.34	1.67	1.29
180	.107	1.71	-1.45	-8.35	1.70	1.57
190	.106	2.06	-1.39	-8.36	1.73	1.89

Table 1.1: Top width as a function of top mass and the comparison of the different approximations.

amounts to -11.6 % instead of -8.3%. This variation characterises the present theoretical uncertainty, which could be removed by a full α_s^2 calculation only. Additional uncertainties, e.g. from the input value of α_s ($\sim$ 1%) or from the fundamental uncertainty in the relation between the pole mass m_t and the experimentally measured excitation curve (assuming perfect data) of perhaps 0.5 GeV can be neglected in the forseable future.

Hence, it appears that the top quark width (and similarly the spectra to be discussed below) are well under theoretical control, including QCD and electroweak corrections. The remaining uncertainties are clearly smaller than the experimental error in Γ_t, which will amount to 5-10% even at a linear collider [35].

1.2.3 Decay spectra and angular distributions

Born predictions

Arising from a two body decay, the energy of the W and of the hadronic system ($\equiv b$ jet) are fixed to

$$\begin{aligned} E_W &= \frac{m_t^2 + m_W^2 - m_b^2}{2m_t} \\ E_h &= \frac{m_t^2 + m_b^2 - m_W^2}{2m_t} \end{aligned} \tag{1.71}$$

as long as gluon radiation is ignored. The smearing of this δ spike by the combined effects of perturbative QCD and from the finite width of the W will be treated below.

Top quarks will in general be polarized through their electroweak production mechanism. For unpolarized beams and close to threshold their polarization is given by the right/left asymmetry which would be measured with longitudinally polarized beams [36]:

$$P_t = A_{RL} \tag{1.72}$$

For fully longitudinally polarized electron (and unpolarized positron) beams the spin of both t and $\bar{t}$ is aligned with the spin of the e^-. Quark polarization then leads to angular distributions of the decay products which allow for various tests of the chirality of the tbW vertex.

The angular distribution of the longitudinal and transverse W's is analogous to those of ρ mesons from τ decay ($m_\tau \to m_t, m_\rho \to m_W$)

$$\frac{dN_{T/L}}{d\cos\theta} = \frac{1}{2}(1 \mp P_t \cos\theta) \quad for \quad h_W = \left\{ \begin{pmatrix} +1 \\ 0 \end{pmatrix} \right. \tag{1.73}$$

and, after summation over the W polarizations

$$\frac{dN}{d\cos\theta} = \frac{1}{2}\left(1 + \frac{m_t^2 - 2m_W^2}{m_t^2 + 2m_W^2} P_t \cos\theta\right) \tag{1.74}$$

The angle between top quark spin and direction of the W is denoted by θ. In the limit of $m_t \gg M_W$ the coefficient of the $P_t \cos\theta$ term rises to 1, for $m_t = 180$ GeV, however, it amounts to 0.43 only.

The angular distribution of leptons from the chain $t \to b + W(\to \ell^+\nu)$ will in general follow a complicated pattern with an energy dependent angular distribution

$$\frac{dN}{dx\, d\cos\theta} = f(x) + g(x) P_t \cos\theta. \tag{1.75}$$

In the SM, however, a remarkable simplification arises. Energy and angular distribution factorize [36, 37]

$$\frac{dN}{dx d\cos\theta} = f(x)(1 + P_t \cos\theta)/2. \tag{1.76}$$

This factorisation holds true for arbitrary m_t and even including the effect of the nonvanishing b-quark mass [34].

QCD corrections

The δ spike in the energy distribution of the hadrons from the decay $t \to b+W$ is smeared by quark fragmentation (not treated in this context).

Hard gluon radiation leads to a slight shift and distortion of the energy spectra with a tail extending from the lower limit given by two-body kinematics upwards to $m_t - m_W$

$$\frac{m_t^2 + m_b^2 - m_W^2}{2m_t} \leq E_{had} < m_t - m_W \tag{1.77}$$

Including finite W-width effects and $m_b \neq 0$ the differential hadron energy distribution has been calculated in [38]. The hadron energy distribution is shown in Fig.1.9 for $m_t = 180$ GeV.

The lepton spectrum (as well as the neutrino spectrum) receives its main correction close to the end points where the counting rates are fairly low.

Including QCD corrections [37, 39, 40, 41] the spectrum of both charged lepton and neutrino can be cast into the form

$$\frac{\mathrm{d}N}{\mathrm{d}x \mathrm{d}\cos\theta} = A(x) + B(x)\cos\theta \tag{1.78}$$

The shape of the charged lepton spectrum is hardly different from the lowest order result [37] with main corrections towards the end point. $B_e(x) \approx A_e(x)$ remains valid to extremely high precision [39]. The charged lepton direction is thus a perfect analyser of the top spin, even after inclusion of QCD corrections. A small admixture of $V + A$ couplings will affect spectrum and angular distributions of electron and neutrino as well. Assuming a $V + A$ admixture of relative rate $\kappa^2 = 0.1$, the functions A_e, B_e and A_ν are only marginally modified (Fig. 1.10 and 1.11). The angular dependence part of the neutrino spectrum B_ν, however, is changed significantly (Fig. 1.11). This observation could provide a useful tool in the search for new couplings.

1.2.4 Non–standard top decays

The theoretical study of non–standard top decays is motivated by the large top quark mass which could allow for exciting novel decay modes, even at the Born level. A few illustrative, but characteristic examples will be discussed in some detail in the following section.

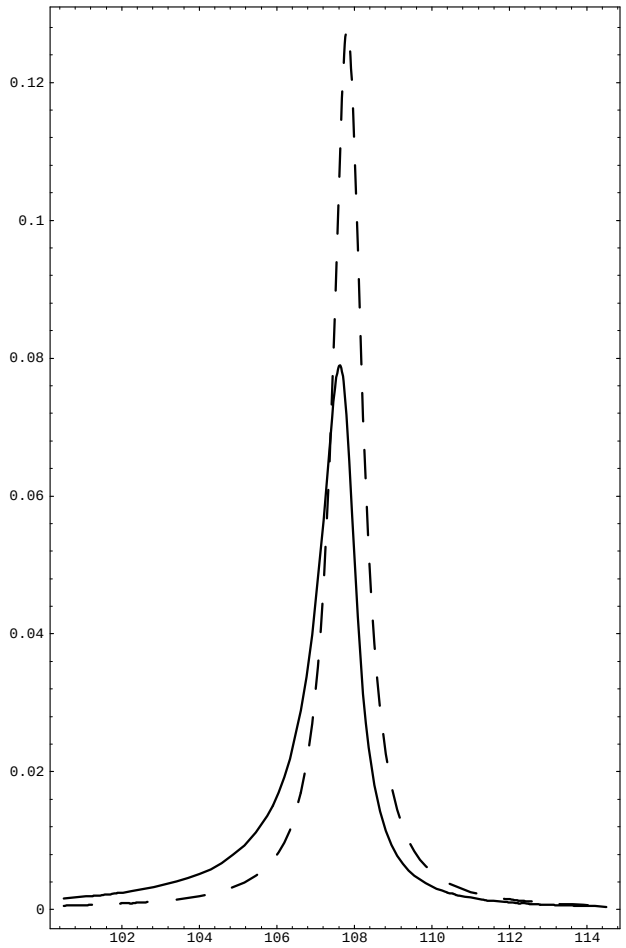


Figure 1.9: Distribution of the W energy for $m_t = 180\ GeV$ without (dashed) and with (solid curve) QCD corrections.

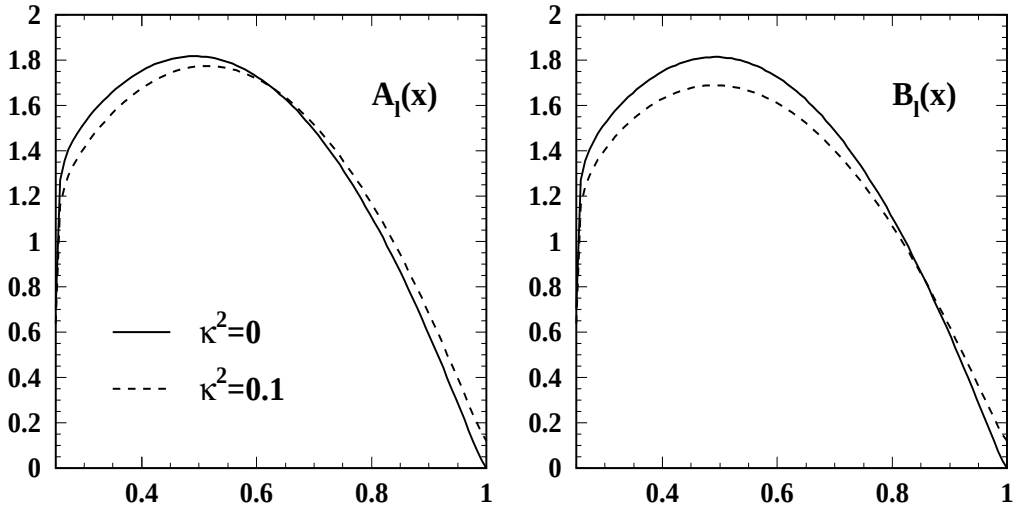


Figure 1.10: The coefficient functions a) $A_\ell(x)$ and b) $B_\ell(x)$ defining the charged lepton angular-energy distribution for $y = 0.25$ and $\alpha_s(m_t) = 0.11$: $\kappa^2 = 0$ – solid lines and $\kappa^2 = 0.1$ – dashed lines.

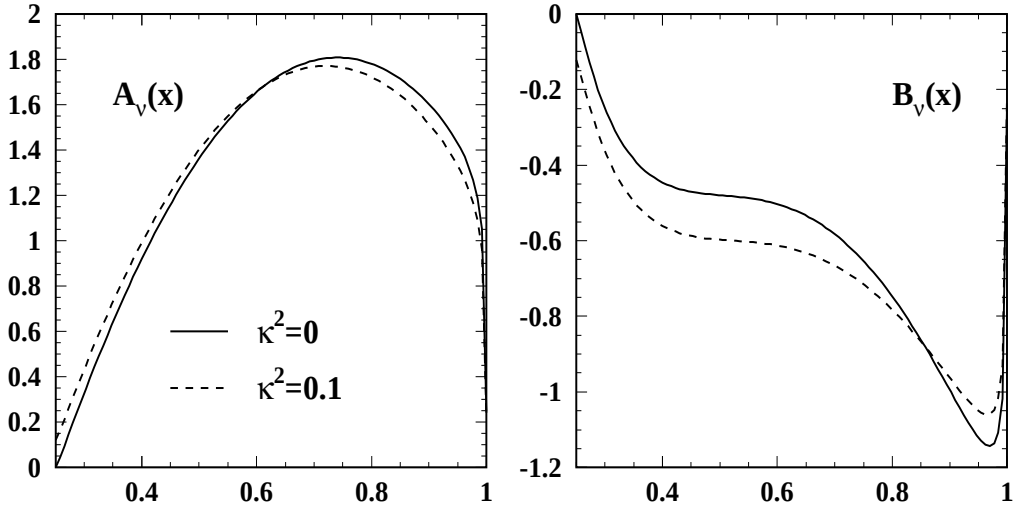


Figure 1.11: The coefficient functions a) $A_\nu(x)$ and b) $B_\nu(x)$ defining the neutrino angular-energy distribution for $y = 0.25$ and $\alpha_s(m_t) = 0.11$: $\kappa^2 = 0$ – solid lines and $\kappa^2 = 0.1$ – dashed lines.

Charged Higgs decays

Charged Higgs states $H^{\pm}$ appear in 2–doublet Higgs models in which out of the eight degrees of freedom three Goldstone bosons build up the longitudinal states of the vector bosons and three neutral and two charged states correspond to real physical particles. A strong motivation for this extended Higgs sector is provided by supersymmetry which requires the Standard Model Higgs sector to be doubled in order to generate masses for the up and down–type fermions. In the minimal version of that model the masses of the charged Higgs particles are predicted to be larger than the W mass, mod. radiative corrections,

$$m(H^{\pm}) > m(W^{+}) \text{ [mod. rad. corr.]}$$

We shall adopt this specific model for the more detailed discussion in the following paragraphs.

If the charged Higgs mass is lighter than the top mass, the top quark may decay into H^{+} plus a b quark [42],

$$t \to b + H^{+}$$

The coupling of the charged Higgs to the scalar (t, b) current is defined by the quark masses and the parameter $\tan\beta$,

$$J(b,t) = \frac{1}{\sqrt{2}v}\left[(m_b \tan\beta + m_t \cot\beta) - \gamma_5\,(m_b \tan\beta - m_t \cot\beta)\right]. \tag{1.79}$$

The parameter $\tan\beta$ is the ratio of the vacuum expectation values of the Higgs fields giving masses to up and down–type fermions, respectively. For the sake of consistency, related to grand unification, we shall assume $\tan\beta$ to be bounded by

$$1 < \tan\beta = \frac{v_2}{v_1} < \frac{m_t}{m_b} \sim 60 \tag{1.80}$$

with $v = \sqrt{v_1^2 + v_2^2} = (\sqrt{2}G_F)^{-1/2}$ corresponding to the ground state of the Standard Model Higgs field. The width following from the coupling (1.79) has a form quite similar to the Standard Model decay mode [see e.g. [43]],

$$\Gamma(t \to b + H^{+}) = \frac{G_F m_t^3}{8\sqrt{2}\pi}\left[1 - \frac{m_H^2}{m_t^2}\right]^2\left[\left(\frac{m_b}{m_t}\right)^2 \tan^2\beta + \cot^2\beta\right] \tag{1.81}$$

The branching ratio of this novel Higgs decay mode is compared with the W decay mode in Fig.1.12a (The behaviour is qualitatively similar for $m_t = 180$ GeV.) In the parameter range eq. (1.80) the W decay mode is dominant; the Higgs decay branching ratio is in general small, yet large enough to be clearly observable [45]. The Higgs branching ratio is minimal at $\tan\beta = \sqrt{m_t/m_b} \sim 6 - 8$. QCD corrections to the $t \to Hb$ mode have been calculated in [46] and electroweak corrections in [47].

The detection of this scalar decay channel is facilitated by the characteristic decay pattern of the charged Higgs bosons

$$H^{+} \to \tau^{+} + \nu_\tau \quad \text{and} \quad c + \overline{s}$$

Figure 1.12: **(a)** Branching ratios for the decays $t \to bW^+$ and $t \to bH^+$ in two–doublet Higgs models [44].

Figure 1.12: **(b)** Branching ratios for charged Higgs decays to τ leptons and quarks [44].

Since $H^\pm$ bosons couple preferentially to down–type fermions [48] for $\tan\beta > 1$,

$$\Gamma(H^+ \to \tau^+\nu_\tau) = \frac{G_F m_\tau^2}{\sqrt{2}} \frac{m_H}{4\pi} \tan^2\beta \tag{1.82}$$

$$\Gamma(H^+ \to c\bar{s}) = \frac{3G_F m_c^2}{\sqrt{2}} \frac{m_H}{4\pi} \left[\left(\frac{m_s}{m_c}\right)^2 \tan^2\beta + \cot^2\beta \right] \tag{1.83}$$

the τ decay mode wins over the quark decay mode [Fig.1.12b], thus providing a clear experimental signature. A first signal of top decays into charged Higgs particles would therefore be the breakdown of μ, e *vs.* τ universality in semileptonic top decays.

An interesting method for a determination of $\tan\beta$ is based on an analysis of the angular distribution of Higgs bosons in the decay of polarized top quarks

$$\frac{\mathrm{d}N}{\mathrm{d}\cos\theta} \sim m_t^2 \cot^2\beta(1+\cos\theta_H) + m_b^2 \tan^2\beta(1-\cos\theta_H), \tag{1.84}$$

an immediate consequence of the couplings given in (1.79).

Top decay to stop

Another exciting decay mode in supersymmetry models is the decay of the top to the SUSY scalar partner stop plus neutralinos, mixtures of neutral gauginos and higgsinos [49, 50]. This possibility is intimately related to the large top mass which leads to novel phenomena induced by the strong Yukawa interactions. These effects do not occur in light–quark systems but are special to the top.

The mass matrix of the scalar SUSY partners $(\tilde{t}_L, \tilde{t}_R)$ to the left– and right–handed top–quark components (t_L, t_R) is built–up by the following elements [51]

$$\mathcal{M}^2 = \left\| \begin{array}{cc} m_{\tilde{t}_L}^2 + m_t^2 & \delta\tilde{m}_{LR}^2 \\ \delta\tilde{m}_{LR}^2 & m_{\tilde{t}_R}^2 + m_t^2 \end{array} \right\|$$

Large Yukawa interactions lower the diagonal matrix elements $\sim -m_t^2$ with respect to the common squark mass value in supergravity models, and they mix the $\tilde{t}_L$ and $\tilde{t}_R$ states with the strength $\sim m_t$ to form the mass eigenstates $\tilde{t}_1, \tilde{t}_2$. Unlike the five light quark species, these Yukawa interactions of $\mathcal{O}(m_t)$ can be so large in the top sector that after diagonalizing the mass matrix, the smaller eigenvalue may fall below the top quark mass,

$$m_{\tilde{t}_1} < m_t \quad [\ : \text{possible}\ ].$$

The decay modes

$$t \to \tilde{t} + \text{neutralinos}$$

then compete with the ordinary W decay mode. Identifying the lightest SUSY particle with the photino $\tilde{\gamma}$ (the mass of which we neglect in this estimate) one finds

$$\frac{\Gamma(t \to \tilde{t}\tilde{\gamma})}{\Gamma(t \to bW)} \approx \frac{8\sqrt{2}\pi\alpha}{9G_F m_t^2} \frac{[1-m_{\tilde{t}}^2/m_t^2]^2}{[1-m_W^2/m_t^2][1+2m_W^2/m_t^2]} \tag{1.85}$$

This ratio is in general less than 10%. The subsequent $\tilde{t}$ decays

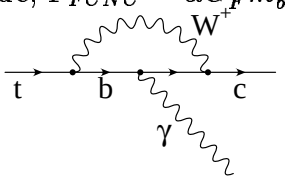
$$\begin{aligned} \tilde{t} &\rightarrow b\tilde{W}, \tilde{W} \rightarrow W\tilde{\gamma} \text{ or } l\tilde{\nu} \text{ } etc. \\ \tilde{t} &\rightarrow c\tilde{\gamma} \end{aligned}$$

lead to an overall softer charged lepton spectrum and, as a result of the escaping photinos, to an increase of the missing energy, the characteristic signature for SUSY induced phenomena.

Depending on the SUSY parameters however, stop decays could even be more enhanced if the top is heavy. Decays into strongly coupled, fairly light higgsinos could thus occur frequently.

FCNC decays

Within the Standard Model, FCNC decays like $t \rightarrow c\gamma$ are forbidden at the tree level by the GIM mechanism. However, they do occur in principle at the one–loop level, though strongly suppressed. The suppression is particularly severe for top decays since the quarks building up the loops, must be down–type quarks with m_b^2 setting the scale of the decay amplitude, $\Gamma_{FCNC} \sim \alpha G_F^2 m_b^4 m_t$. A sample of branching ratios is given below [52]:



$$\begin{aligned} BR(t \rightarrow cg) &\sim 10^{-10} \qquad BR(t \rightarrow cZ) \sim 10^{-12} \\ BR(t \rightarrow c\gamma) &\sim 10^{-12} \qquad BR(t \rightarrow cH) \sim 10^{-7} \end{aligned}$$

At this level, no Standard Model generated t decays can be observed, even given millions of top quarks in proton colliders. On the other hand, if these decay modes were detected, they would be an undisputed signal of new physics beyond the Standard Model. From such options we select one illustrative, though very speculative example for brutal GIM breaking. It is tied to the large top mass and holds out faint hopes to be observable even in low rate e^+e^- colliders.

The GIM mechanism requires all L and R quark components of the same electric charge in different families to carry identical isospin quantum numbers, respectively. This rule is broken by adding quarks in LR symmetric vector representations [53] to the "light" chiral representations or mirror quarks [54]:

$$\begin{aligned} &\underline{\text{vector quarks}}: \quad \cdots \quad \begin{bmatrix} t \\ b \end{bmatrix}_L \; t_R \qquad \begin{bmatrix} U \\ D \end{bmatrix}_L \begin{bmatrix} U \\ D \end{bmatrix}_R \\ &\underline{\text{mirror quarks}}: \quad \cdots \quad \begin{bmatrix} t \\ b \end{bmatrix}_L \; t_R \qquad U_L \; \begin{bmatrix} U \\ D \end{bmatrix}_R \\ & \qquad\qquad\qquad\qquad\qquad\quad b_R \qquad D_L \end{aligned}$$

Low energy phenomenology requires the masses M of the new U, D quarks to be larger than 300 GeV.

Depending on the specific form of the mass matrix, mixing between the normal chiral states and the new states may occur at the level $\sim \sqrt{m/M}$, so that FCNC (t, c) couplings of the order $\sim \sqrt{m_t m_c/M^2}$ can be induced. FCNC decays of top quarks, for example,

$$BR(t \to cZ) \sim \text{fraction of } \%$$

are therefore not excluded. Such branching ratios would be at the lower edge of the range accessible at e^+e^- colliders.

Chapter 2

Top quarks at hadron colliders

The search for new quarks and the exploration of their properties has been a most important task at hadron colliders in the past. The recent observation of top quarks with a mass of around 180 GeV at the TEVATRON has demonstrated again the discovery power of hadron colliders in the high energy region. Several ten's of top quarks have been observed up to now. The significant increase of luminosity toward the end of this decade will sharpen the picture. The branching ratios of the dominant decay modes will be determined and the uncertainty in the top mass reduced significantly. For a detailed study of the top quark properties the high energy collider LHC will provide the required large number of top events [order 10^7].

The main production mechanisms for top quarks in proton-antiproton collisions, Fig.2.1, are quark-antiquark fusion supplemented by a small admixture of gluon-gluon fusion [55].

$$gg \text{ and } q\overline{q} \rightarrow t\overline{t}$$

Top production at the LHC is of course dominated by the second reaction. The W-gluon fusion process [56]

$$Wg \rightarrow t\overline{b}$$

is interesting on its own. It is about a factor $0.1 - 0.2$ below the dominant reaction and thus well accessible at the high energy pp colliders — and perhaps even at the TEVATRON.

2.1 Lowest order predictions and qualitative features

The dominant Born terms for the *total top cross section* in $gg/q\overline{q} \rightarrow t\overline{t}$ fusion are well-known to be of the form [55]

$$\sigma_{gg}(\hat{s}) = \frac{4\pi\alpha_s^2}{12\hat{s}} \left[(1+\rho+\frac{\rho^2}{16}) \ln\frac{1+\beta}{1-\beta} - \beta\,(\frac{7}{4}+\frac{31}{16}\rho) \right] \qquad (1a)$$

$$\sigma_{q\bar{q}}(\hat{s}) = \frac{8\pi\alpha_s^2}{27\hat{s}}\,\beta\,[1+\frac{\rho}{2}] \qquad (1b)$$

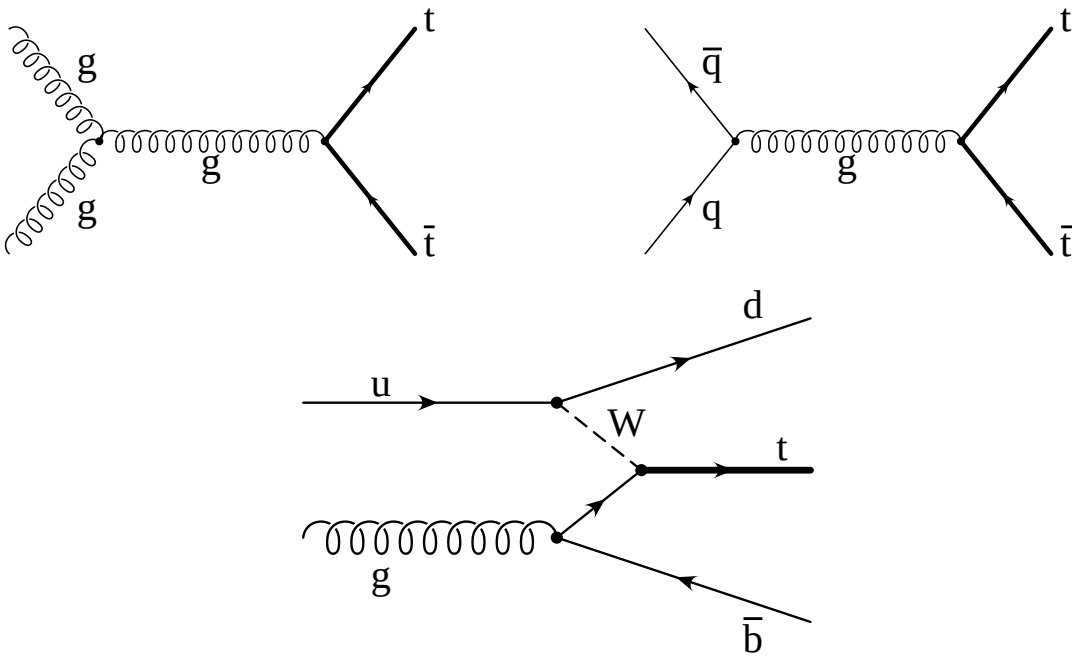


Figure 2.1: The main production mechanisms for top quarks in $p\bar{p}$ and pp colliders [generic diagrams].

with $\rho = 4m_t^2/\hat{s}$ and $\beta = \sqrt{1-\rho}$ being the velocity of the t quarks in the $t\bar{t}$ cm frame with invariant energy $\sqrt{\hat{s}}$. The total $p\bar{p}$ cross sections then follow by averaging the partonic cross sections over the $q\bar{q}$ and gg luminosities in $p\bar{p}$ (and similarly in pp) collisions,

$$\sigma\left(p\bar{p}\rightarrow t\bar{t}\right) = \int_{4m_t^2/s}^{1} d\tau \, \frac{d\mathcal{L}\left(gg\right)}{d\tau}\sigma_{gg}\left(\tau s\right) + \left[q\bar{q}\right] \tag{2.2}$$

The relative enhancement of the $q\bar{q}$ cross section by about a factor 3, as evident from the threshold behaviour

$$\begin{aligned} \sigma_{q\bar{q}} &\approx \frac{4}{9}\frac{\pi\alpha_s^2}{\hat{s}}\beta \\ \sigma_{gg} &\approx \frac{7}{48}\frac{\pi\alpha_s^2}{\hat{s}}\beta \end{aligned} \tag{2.3}$$

has to be combined with the prominent $q\bar{q}$ luminosity at the TEVATRON. As shown in Fig. 2.2

$$\mathrm{d}\mathcal{L}_{q\bar{q}} : \mathrm{d}\mathcal{L}_{gg}|_{\sqrt{\hat{s}}=2m} \approx \begin{cases} 5 & \text{TEVATRON} \\ 0.1 & \text{LHC} \end{cases} \tag{2.4}$$

which implies the dominance of $q\bar{q}$ annihilation, in contrast to the situation at the LHC, where gluon fusion is the dominant reaction.

A number of important features can be read off from this lowest order result:

- Since the parton luminosities rise steeply with decreasing τ, the production cross sections increase dramatically with the energy (Fig. 2.3)

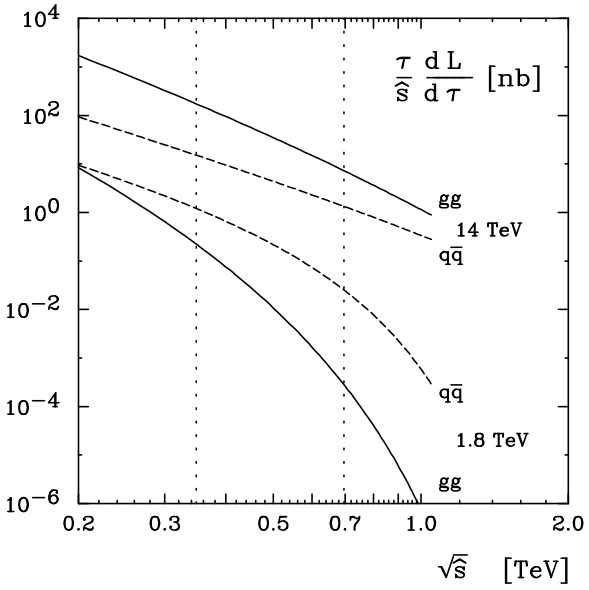


Figure 2.2: Parton luminosities for TEVATRON and LHC energies.

Figure 2.3: Production cross section for $t\bar{t}$ pairs in $p\bar{p}$ and pp colliders: Tevatron (1.8 GeV); Tevatron II (3.6 TeV); LHC (16 TeV); SSC (40 TeV). Ref.[44].

- Structure functions and quark-antiquark luminosities in the region of interest for the TEVATRON, i.e. for $\sqrt{\tau}$ between 0.2 and 0.4 are fairly well known from experimental measurements at lower energies (combined with evolution equations) and collider studies. The predictions are therefore quite stable with respect to variations between different sets of phenomenologically acceptable parton distributions. The near tenfold increase of the energy at the LHC and the corresponding decrease of x and $\sqrt{\tau}$ by nearly a factor of ten leads to the dominance of gluon-gluon fusion and results in a significantly enhanced uncertainty in the production cross section.

- With the cross section proportional to α_s^2 and uncertainties in $\alpha_s(M_Z)$ which may be stretched up to $\pm 10\%$ one might naively expect a resulting uncertainty in the predicted cross section. However, the increase in the parton cross section with increasing α_s is, to some extent, compensated by a decrease in the parton luminosity (with increasing α_s) for the kinematical region of interest at the TEVATRON. This compensation mechanism has been studied in [57] for inclusive jet production (Fig. 2.4) and applies equally well for top quark production.

- At the TEVATRON the rapidity distribution is strongly dominated by central production, $|y| \leq 1$, a consequence of the balance between the steeply falling proton and antiproton parton distributions. At the LHC, however, a rapidity plateau develops gradually and the distribution spans nearly four units in rapidity (Fig. 2.5).

- The transverse momentum distribution is relatively flat, dropping down to half its peak value at around $p_t \approx m_t/2$, again a consequence of the competition between the increase of the phase space factor $\propto \beta$ in the parton cross section and the steeply decreasing parton luminosity (Fig. 2.6). At the LHC the distribution will be even flatter and p_t values around $2m_t$ are well within reach (Fig. 2.7). This corresponds to CMS energies between 0.5 and 1 TeV in the parton subsystem and extremely large subenergies are therefore accessible. This opens the possibility to search for the radiation of W, Z, or Higgs bosons in this reaction. For high energies the suppression of the cross section through electroweak virtual corrections (cf. sect. 2.2.3) is, at least partially, compensated by the large logarithm $\ln \hat{s}/M^2_{W,Z,H}$.

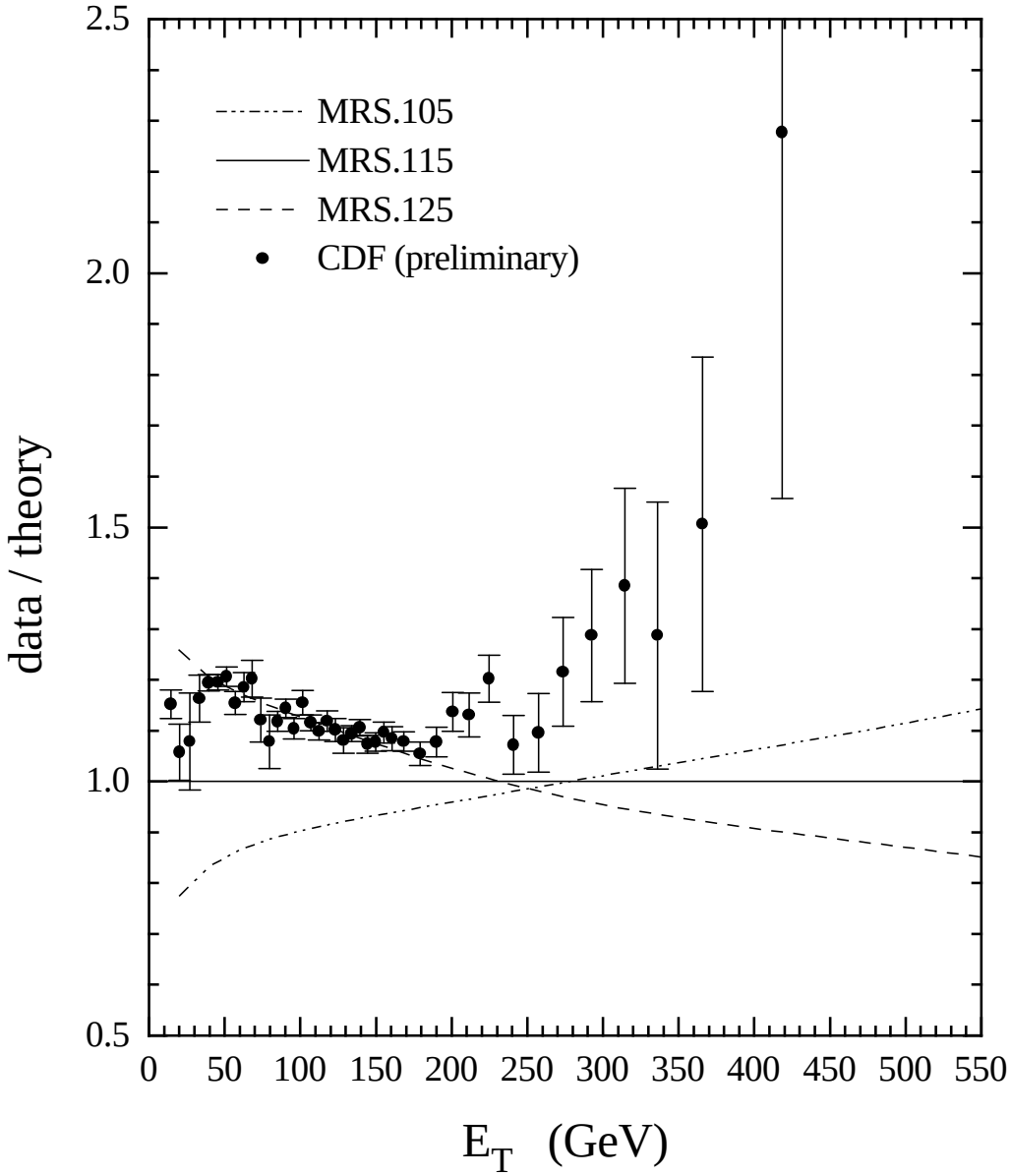


Figure 2.4: The $p\bar{p}$–initiated jet E_T distribution at $\sqrt{s} = 1.8$ TeV normalized to the prediction from partons with $\alpha_S = 0.115$ (i.e. MRS.115). The data are the CDF measurements of $d^2\sigma/dE_T d\eta$ averaged over the rapidity interval $0.1 < |\eta| < 0.7$. The curves are obtained from a leading-order calculation evaluated at $\eta = 0.4$. The data are preliminary and only the statistical errors are shown. The systematic errors are approximately 25% and are correlated between different E_T points. (From [57]).

Figure 2.5: Rapidity distribution of top quarks in $gg/q\bar{q}$ fusion at $\sqrt{s} = 16$ TeV [44].

2.2 QCD and electroweak corrections

The observation of top quarks has been well established during the last year. One of the tools to study its properties, in particular its mass and its decay modes, is a precise experimental determination of its production cross section and subsequent decay in the $t \to b + W$ channel. A large deficiency in the comparison between theory and experiment would signal the presence of new decay modes which escape the canonical experimental cuts; with $t \to bH^+$ as most prominent example. However, the early round of experiments had indicated even an excess of top events when compared to the theoretical prediction for $m_t \approx 180$ GeV. This observation was difficult to interpret and the original calculations were scrutinized again by various authors. In particular, the resummation of leading logarithms and the influence of the Coulomb threshold enhancement was investigated — in the end, however, the prediction remained fairly stable.

In these lectures we will, therefore, in a first step, present the results from a complete NLO calculation (sect. 2.2.1). This is supplemented by a qualitative discussion of the resummation of higher order leading logarithmic terms. The influence of the Coulomb enhancement is studied in section 2.2.2, electroweak corrections are presented in sect. 2.2.3. Radiation of gluons may have a sizeable effect on the aparent mass of top quarks as observed in the experiment (sect. 2.2.4), with distinct differences between initial and final state radiation.

Figure 2.6: The differential cross section for $p + \bar{p} \rightarrow t(\bar{t}) + X$ with $m_t = 120$ GeV/c^2 and $\mu_R = Q = \sqrt{m_t^2 + p_t^2}$ at $\sqrt{s} = 1.8$ TeV. The cross section is shown at different values of rapidity for (1) dashed lines: lowest order contribution scaled by an arbitrary factor (2) solid lines: full order α_s^2 calculation. (From [58].)

2.2.1 Next to leading order (NLO) corrections and resummation of large logarithms

Higher-order QCD corrections [59, 60, 58, 61] include loop corrections to the Born terms and $2 \rightarrow 3$ contributions like $gg \rightarrow t\bar{t}g$, $q\bar{q} \rightarrow t\bar{t}g$ etc. For $q\bar{q}$ annihilation a few characteristic diagrams are displayed in Fig. 2.8. Real and virtual initial state radiation (Fig. 2.8a,b) dominate, final state radiation from the slow top quarks (Fig. 2.8c) is unimportant, virtual gluon exchange at the $t\bar{t}$ vertex (Fig. 2.8d) leads to the Coulomb enhancement and will be discussed in sect. 2.2.2. The separation between $q\bar{q}$ annihilation and qg reactions (Fig. 2.9) depends on the choice of the so-called factorisation scale μ^2 which effectively enters the definition of the structure functions.

The differential as well as the total production cross section can be cast into the following form

$$\mathrm{d}\sigma = \sum_{i,j=partons} \int \mathrm{d}x_1 \mathrm{d}x_2 F_i^p(x_1, \mu^2) F_j^{\bar{p}}(x_2, \mu^2) \mathrm{d}\sigma_{ij}\left(\hat{s}, \ldots, \mu^2\right) \tag{2.5}$$

The renormalization scale μ_R and the factorisation scale μ_F are in general identified, $\mu_R = \mu_F \rightarrow \mu$, a matter of convention and convenience more than a matter of necessity.

Figure 2.7: Transverse momentum distribution of top quarks in $gg/q\bar{q}$ at $\sqrt{s} = 16$ TeV. (From [44].)

The parton distributions are extracted from structure functions as measured in deep inelastic scattering, and the analysis has to be taylored to the order of the calculation, i.e. to the NLO in the present case. The integrated expressions for the total cross sections can still be cast into a simple form

$$\sigma_{ij}(\hat{s}, m_t^2, \mu^2) = \frac{\alpha_s^2(\mu^2)}{m_t^2}\left[f_{ij}^{(0)}(\rho) + 4\pi\alpha_s(\mu^2)\left(f_{ij}^{(1)}(\rho) + \bar{f}_{ij}^{(1)}(\rho)\ln\frac{\mu^2}{m_t^2}\right)\right] \qquad (2.6)$$

where $\hat{s} = x_1 x_2 s$ and the dominant lowest-order contributions $f_{ij}^{(0)}(\rho)$ are given by the parton cross sections above; in addition $f_{gq}^{(0)} = f_{g\bar{q}}^{(0)} = 0$. The subleading higher-order expressions for $f_{ij}^{(1)}$ and $\bar{f}_{ij}^{(1)}$ are given in Refs.[59], [60]. The heavy quarks are treated within the on-shell renormalization scheme with m_t being the "physical" mass of the top quark. Outside the heavy quark sector, the $\overline{MS}$ scheme has been employed. These higher-order terms have to be used in conjunction with the running coupling $\alpha_s(\mu^2)$ and the gluon/light-quark parton densities evolved in 2-loop evolution equations. μ is the renormalization scale, identified here also with the factorization scale; typical scales that have been chosen are $\mu = m_t$ and $\sqrt{m_t^2 + p_T^2}$. More technical details are discussed in Ref.[44].

The lowest- and higher-order predictions are compared with each other in Fig.2.3. In [44]

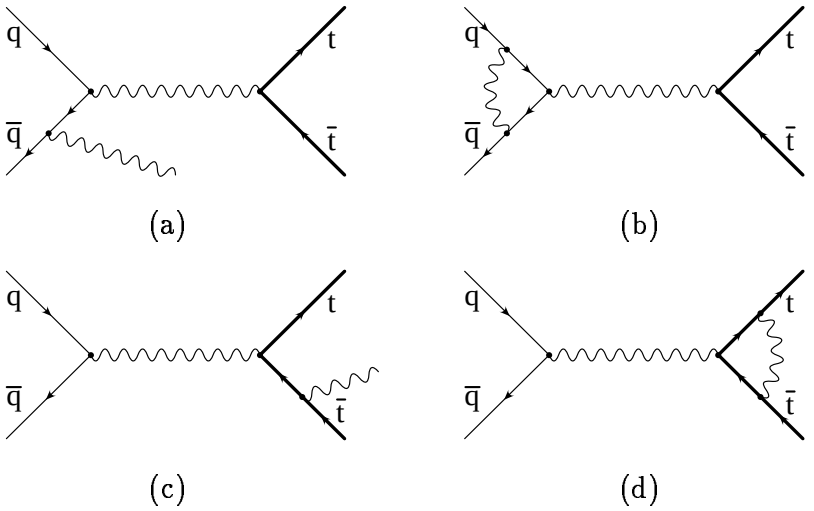


Figure 2.8: Initial and finite state radiation in the reaction $q\bar{q} \to t\bar{t}$

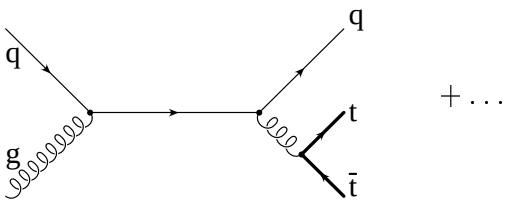


Figure 2.9: Reaction $qg \to t\bar{t}q$

it has been argued that the subdominant $2\to 3$ contributions add up to less than 10% of the dominant lowest-order results. The theoretical uncertainties of the predictions for the LHC due to different parton distributions [62] were estimated about $\pm 10\%$ plus a $\pm 10\%$ variation due to the scale ambiguity μ. The impact of the additional shift from the resummation of large logs arising in higher orders will be discussed below. Note that the "K factor", defined formally by the higher order corrections to the LO parton cross section, but the parton distributions and α_s kept fixed, amounts to an $\approx 50\%$ correction of the Born terms.

It is also instructive to study separate, physically distinct components of the α_s^3 results [61]. The initial state bremsstrahlung (ISGB) processes, illustrated for the gluon initiated reactions in Fig.2.10, dominate around threshold [$\sqrt{\hat{s}} \geq 2m_t$ or $p_t < m_t$)], the case of relevance at the TEVATRON. The gluon splitting (GS) and the flavor excitation (FE) contributions become increasingly important for $\sqrt{\hat{s}} \gg 2m_t$, the situation anticipated for the LHC.

Let us concentrate in the remainder of this section on the predictions for TEVATRON energies. Initial state radiation reduces the effective energy in the partonic subsystem, requiring larger initial parton energies to reach the threshold for top pair production. Considering the steeply falling parton distributions $F_j(x)$ one might, therefore, expect a reduction of σ through NLO contributions. However, the same effect is operative in

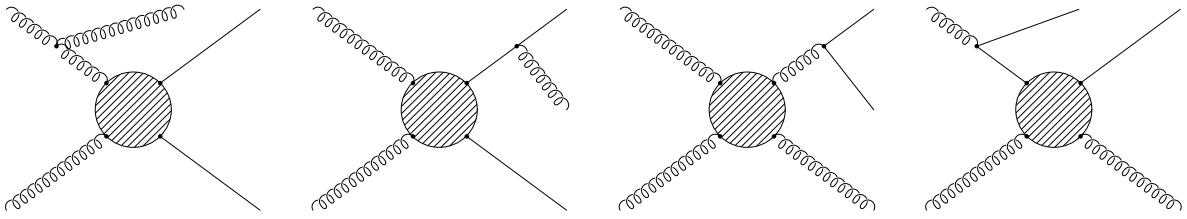


Figure 2.10: Generic QCD diagrams of the dominant higher order mechanisms.

the very definition of F_j (Fig. 2.11) through deep inelastic scattering, including NLO

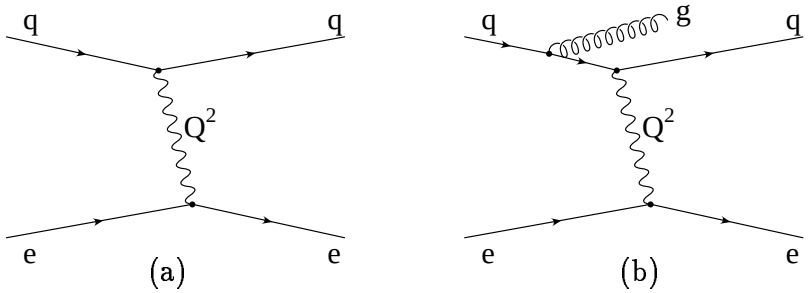


Figure 2.11: Definition of quark distributions, including NLO.

corrections. In fact, without this compensation mechanism the result would not even be finite. However, the magnitude or even the sign of the correction cannot be guessed on an intuitive basis, and, not quite unexpected, even the precise form of $f^{(1)}$ and $\bar{f}^{(1)}$ depends on the definition of the structure functions. The most prominent examples are the $\overline{MS}$ scheme where $1/\epsilon$ poles (plus $\ln(4\pi) - \gamma$) from collinear singularities are simply dropped [more precisely, they are combined with the corresponding singular terms which arise in the NLO definition of the structure function] and finite corrections have to be applied when comparing to deep inelastic scattering, and the DIS scheme, where F_j are defined through deep inelastic scattering to all orders.

Let us illustrate the qualitative aspects in the simpler example of NLO contributions to the Drell Yan process. The dominance of initial state radiation in the corrections to $t\bar{t}$ production will allow to apply the same reasoning to the case of interest in these lectures. Including NLO corrections one obtains

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2} &= \sigma_0 \sum_{i,j} \int_0^1 \frac{\mathrm{d}x_i}{x_i} \frac{\mathrm{d}x_j}{x_j} F_i^p(x_i, Q^2) F_j^{\bar{p}}(x_j, Q^2) \cdot \omega_{ij}\left(\frac{\tau}{x_i x_j}, \alpha_s\right) \\
&= \sigma_0 \sum_{i,j} \int_\tau^1 \frac{\mathrm{d}z}{z} \mathcal{F}_{ij}(\tau/z) \cdot \omega_{ij}(z, \alpha_s) \qquad (2.7)
\end{aligned}$$

with

$$\begin{aligned}
\sigma_0 &\equiv \frac{4\pi\alpha^2}{9Q^2 s} \qquad \tau \equiv \frac{Q^2}{s} \\
\omega_{q\bar{q}}^{(0)}(z, \alpha_s) &= e_q^2 \delta(1-z)
\end{aligned}$$

$$\omega^{(1)}_{q\bar{q}}(z,\alpha_s) = e_q^2\frac{\alpha_s}{2\pi}C_F\left[2(1+z^2)\left(\frac{\ln(1-z)}{1-z}\right)_+ + 3\left(\frac{1}{1-z}\right)_+ + \left(1+\frac{4\pi^2}{3}\right)\delta(1-z) + \text{regular functions}\right] \quad (2.8)$$

(The quark-gluon induced reactions will not be discussed in this connection.) The plus prescription which regulates the singularity of the distributions at $z = 1$ arises from the subtraction of collinear singularities. It can be understood by considering the limit

$$\left(\frac{\ln^n(1-z)}{1-z}\right)_+ \equiv \lim_{\epsilon\to 0}\left(\frac{\ln^n(1-z)}{1-z}\Theta(1-\epsilon-z) + \delta(1-z)\frac{\ln^{n+1}\epsilon}{n+1}\right) \quad (2.9)$$

with the coefficient of the δ function adjusted such that the integral from zero to one vanishes.

Equivalently the plus-distribution can be defined through an integral with test functions $f(z)$. If $f(z)$ vanishes outside the interval $\langle\tau, 1\rangle$ a convenient formula which will be of use below reads as follows

$$\int_\tau^1 \mathrm{d}z f(z)\left(\frac{\ln^n(1-z)}{1-z}\right)_+ = \int_\tau^1 \mathrm{d}z\frac{f(z)-f(1)}{1-z}\ln^n(1-z) - f(1)\int_0^\tau dz\,\frac{\ln^n(1-z)}{1-z} \quad (2.10)$$

The Born term $\omega^{(0)}$ is simply given by a δ function peak at $z = 1$, corresponding to the requirement that the squared energy of the partonic system $\tau/z \times s$ and the squared mass of the muon pair $\tau \cdot s = Q^2$ be equal. $\mathcal{O}(\alpha_s)$ corrections contribute to the δ function through vertex corrections and a continuous part from initial state radiation extending through the range

$$\tau \le z \le 1 \quad (2.11)$$

The upper limit $z = 1$ corresponds to the kinematic endpoint without radiation, the requirement $\tau \le z$ originates from the fact that the parton luminoisities $\mathcal{F}_{ij}(\tau/z)$ vanish for $\tau/z > 1$. The regular and the subleading pieces of ω are process dependent, the leading singularity $\sim (1+z^2)\left(\frac{\ln(1-z)}{1-z}\right)_+$ is universal (and closely related to the $q \to qg$ splitting function) and equally present in $t\bar{t}$ production.

The suppression of final state radiation in top pair production allows to extend the analogy to the Drell Yan process and to employ resummation techniques that were successfully developed and applied for muon pair production [63]. A complete treatment of this resummation is outside the scope of these lectures. Nevertheless we shall try to present at least qualitative arguments which allow to understand the origin of these large logarithms. (For a similar line of argument see [63].) With the energies of the partonic reaction $\sqrt{\hat{s}} = \mathcal{O}(2-4m_t) \approx 350-700$ GeV and the CMS energy $\sqrt{s} = 1800$ GeV of comparable magnitude it is clear that the ratio $\hat{s}/s$ will not give rise to large logs. However, large logarithms can be traced to the interplay between the collinear singularity in the subprocess and the rapidly falling parton luminosity $\mathcal{F}_{ij}$ (cf. eq. (2.7)). This rapid

decrease leads to a reduction in the effective range of integration. Let us, for the sake of argument assume a range reduced from

$$\tau \leq z \leq 1 \tag{2.12}$$

to

$$1 - \delta \leq z \leq 1 \tag{2.13}$$

and evaluate the leading term. For a constant luminosity $\mathcal{F}(\tau/z)/z = \mathcal{F}$ one would obtain

$$\int_\tau^1 \mathrm{d}z \; 2 \left(\frac{\ln(1-z)}{1-z} \right)_+ \mathcal{F} = \ln^2(1-\tau)\mathcal{F} \tag{2.14}$$

If the region of integration extended through the full kinematic range and with $\tau = 4m_t^2/s = 0.04$ there would be no large log. For the restricted range of integration, however, one finds

$$\int_{1-\delta}^1 \mathrm{d}z \; 2 \left(\frac{\ln(1-z)}{1-z} \right)_+ \mathcal{F} = \ln^2(\delta)\mathcal{F} \tag{2.15}$$

For small δ, corresponding in practice to steeply falling luminosities one thus obtains large, positive (!) corrections from the interplay between $\mathcal{F}_{ij}(\tau/z)$ and $\omega_{q\bar{q}}$.

To arrive at a reliable prediction the leading terms of the form $(\alpha_s \ln^2 \ldots)^n$ thus have to be included. The results are based on alternatively momentum space or impact parameter techniques which were originally developed for the Drell Yan process and applied to top pair production in [64]. An additional complication arises from the blow up of the coupling constant associated with the radiation of soft gluons for $\hat{s} \to 4m_t^2$. This has been interpreted in [64] as a breakdown of perturbation theory. Different regulator prescriptions have been advertised. In [64] a cutoff $\hat{s} - 4m_t^2 > \mu_0^2$ with $\mu_0^2 \ll 4m_t^2$ was introduced to exclude a small fraction of the phase space.

The result is fairly stable for $q\bar{q}$ induced reactions with μ_0/m_t chosen between 0.05 and 0.2. The small contribution from gluon fusion, however, is sensitive towards μ_0 which had to be chosen in the range between $0.2m_t$ and $0.3m_t$, a consequence of the enhancement of radiation from gluons. A slightly different approach ("principal value resummation") has been advocated in [65] which circumvents the explicit μ_0 dependence of the result, but leads essentially to the same final answer (Table 2.1).

The result of the improved prediction (central value) is compared to the fixed order calculation (with $\mu^2 = 4m^2, m^2, m^2/4$) in fig. 2.12. Resummation evidently increases the cross section slightly above the previously considered range. The history of predictions is shown in table 2.2, with $\sqrt{s} = 1.8$ TeV and $m_t = 180$ GeV as reference values. The table demonstrates that the spread of predictions through a (fairly extreme) variation of structure functions (DFLM vs. ELHQ) and through a variation of the renormalisation and factorisation are comparable — typically around $\pm 10\%$. Leading log resummation increases the cross sections by $10 - 15\%$, with a sizeable sensitivity towards the cutoff prescription. A reduction in m_t by 5 GeV leads to an increase of σ by about 0.8 pb. Theory and experiment, with its present result of $7.6^{+2.4}_{-2.0}$ pb and 6.3 ± 2.2 pb from CDF and D0 respectively are thus well compatible (Fig. 2.13).

m_t [GeV]	175	180
μ_0 (min)	4.72	3.86
μ_0 (centr)	4.95	4.21
μ_0 (max)	5.65	4.78
"principal value"	5.6	4.8

Table 2.1: Top production cross section (in pb) for $\sqrt{s}$ =1800 GeV for different values of the cutoff μ_0 [64] (first three lines) and for the "principal value" prescription [65] (fourth line).

	σ [pb]
Altarelli et al. [59]	3.52 (DFLM) 4.10 (ELHQ)
Laenen et al. [64]	3.5 $(\mu^2 = 4m^2)$ 3.8 $(\mu^2 = m^2)$ 4.05 $(\mu^2 = m^2/4)$ } MRSD
Resummation	
Laenen et al. [64]	3.86 4.21 4.78 } vary μ_0
Berends et al. [66]	4.8 central value
Berger et al. [65]	4.8 "principal value res."

Table 2.2: History of predictions for the production cross section for $\sqrt{s}$ = 1.8 TeV and m_t = 180 GeV.

2.2.2 Threshold behaviour

Near the $t\bar{t}$ threshold the cross sections are affected by resonance production and Coulomb rescattering forces [43], [67], [68]. These corrections can be estimated in a simplified potential picture. The driving one-gluon exchange potential is attractive if the $t\bar{t}$ is in color-singlet state and repulsive in a color-octet state [68],

$$\begin{aligned} \sigma^{(1)}(gg\to t\bar{t}) &= \frac{2}{7}\,\sigma_B(gg\to t\bar{t})\,|\,\Psi^{(1)}\,|^2 \\ \sigma^{(8)}(gg\to t\bar{t}) &= \frac{5}{7}\,\sigma_B(gg\to t\bar{t})\,|\,\Psi^{(8)}\,|^2 \\ \sigma^{(8)}(q\bar{q}\to t\bar{t}) &= \sigma_B(q\bar{q}\to t\bar{t})\,|\,\Psi^{(8)}\,|^2 \end{aligned} \tag{2.16}$$

Figure 2.12: The NLO exact cross section as a function of the top quark mass for three choices of scale: $\mu = m/2$ (upper solid line), $\mu = m$ (central solid line) and $\mu = 2m$ (lower solid line), and the NLO exact cross section plus the $\mathcal{O}(\alpha_s^4)$ contribution at $\mu = m$ (dashed line) (from [64]).

with the correction factors (see fig. 2.14) given in NLO by

$$\beta|\psi|^2 = \beta\left(1 + \frac{\pi\alpha_s}{2\beta_t}\left\{\begin{array}{cc} 4/3 & \text{singlet} \\ -1/6 & \text{octet} \end{array}\right.\right) \tag{2.17}$$

The summation of the leading $\pi\alpha_s/\beta$ terms to all orders results in the familiar Sommerfeld correction factor

$$\beta|\psi|^2 = \beta\frac{x}{1-e^{-x}} \tag{2.18}$$

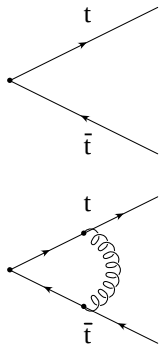
For $t\bar{t}$ in the singlet configuration $x = x^{(1)} \equiv \frac{4}{3}\frac{\pi\alpha_s}{\beta}$, for octet states $x = x^{(8)} \equiv -\frac{1}{6}\frac{\pi\alpha_s}{\beta}$.

The Coulombic attraction thus leads to a sharp rise of the cross section at the threshold in the singlet channel, even if no resonance can be formed anymore, since the phase space suppression of the Born term $\sigma_B \propto \beta_t$ is neutralized by the Coulomb enhancement of

Figure 2.13: Comparison of experimental results for the top quark mass and production cross section with the theoretical prediction.

the wave function $|\Psi_1|^2 \propto \alpha_s/\beta$. In the octet channel (dominant for $q\bar{q}$ annihilation), by contrast, the cross sections are strongly reduced by the Coulombic repulsion which leads effectively to an exponential fall-off of the cross sections $\sigma_8 \propto \exp[-\pi\alpha_s/6\beta_t]$ at the threshold [68]. Due to the averaging over parton luminosities the effects are less spectacular in $p\bar{p}$ or pp than in e^+e^- collisions.

The enhancement and suppression factors are compared to simple phase space $\sim \beta$ in Fig. 2.17. The dotted line corresponds to the phase space factor β, the dashed line to the perturbative NLO calculation (2.17), the solid line to the Coulomb enhancement given in eq. (2.18). The predictions for the singlet, octet ($\equiv q\bar{q}$), and properly weighted gluon fusion channel are displayed in Figs. 2.17 a, b, and c, respectively.



$$\sim \beta_t$$

$$\sim \beta_t \cdot \frac{\pi\alpha_s}{2\beta_t} \begin{cases} 4/3 & \text{singlet} \\ -1/6 & \text{octet} \end{cases}$$

Figure 2.14: Vertex corrections from gluon exchange in the threshold region.

2.2.3 Electroweak corrections

Another potentially important modification which is closely tied to the Coulomb enhancement originates from vertex corrections induced by light Higgs boson exchange. In a simplified treatment these are lumped into a Yukawa potential

$$V(r) = \frac{\sqrt{2}G_F m_t^2}{4\pi} \frac{e^{-m_H r}}{r} \tag{2.19}$$

resulting in a reduction δE of the apparent threshold, with $\delta E = -200$ MeV for $m_H = 100$ GeV as characteristic example. The change in the normalisation by $+10\%$ could become relevant for precision measurements. The situation is quite similar to the one discussed for e^+e^- colliders in section 3.2.

Genuine electroweak contributions of $\mathcal{O}(\alpha\alpha_s^2)$ have been calculated to both the $q\bar{q}$ and $gg \to t\bar{t}$ subprocesses [69]. The corrections include vertex corrections and box diagrams built-up by vector bosons and the Higgs boson (Fig. 2.16).

Except for a small region close to the production threshold, which is dominated by the Yukawa potential, the corrections are always negative; they can become sizeably large, in particular if the top is very heavy and if the energy of the subsystem exceeds 1 TeV, not uncommon for $t\bar{t}$ production at the LHC. In this case however, the large negative corrections are compensated by positive contributions from real radiation of W, Z, or H. The corrections for the $q\bar{q}$ and gg subprocess as functions of the parton energies are shown in Fig. 2.17. The sharp increase of the corrections close to threshold for a light Higgs is clearly visible and, similarly, the large negative correction for large parton energies. After convoluting the cross sections of the subprocesses with the parton distributions, a reduction of the Born cross section at a level of a few percent is observed (Fig. 2.18).

2.2.4 Gluon radiation

Up to this point the discussion has centered around the predictions for inclusive top quark production. Additional ingredients for the experimental analysis are the detailed

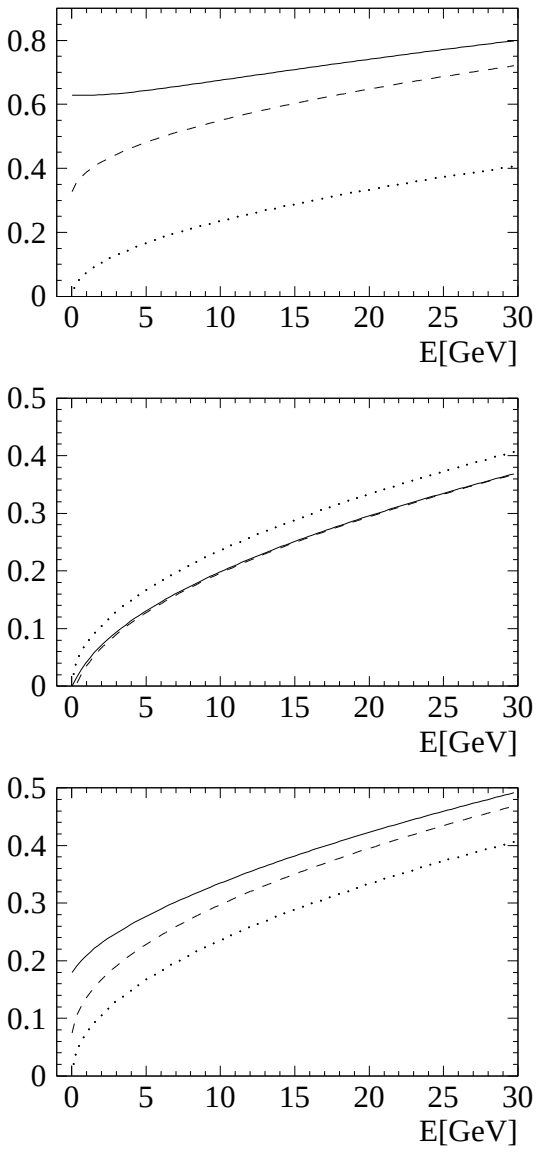


Figure 2.15: Threshold enhancement and suppression factors for singlet (upper), octet (middle), and properly weighted gluon fusion (lower figure) for $\alpha_s = 0.15$. Dotted line: Born approximation; dashed line: NLO approximation; solid line: Coulomb enhancement.

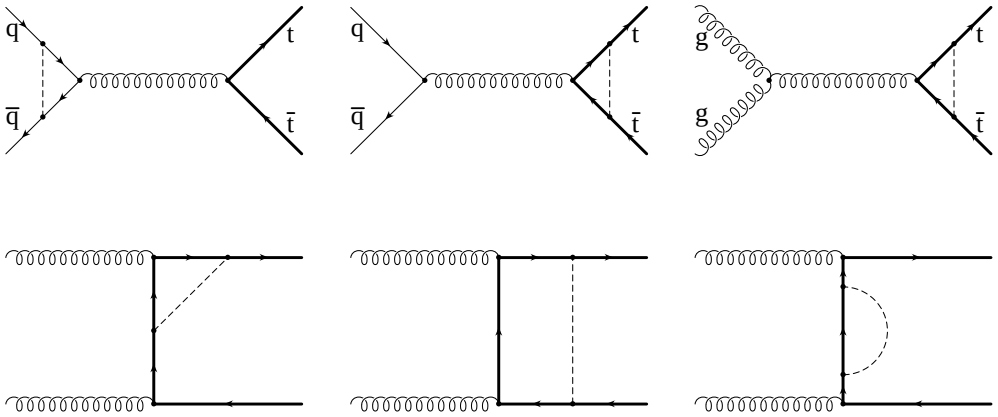


Figure 2.16: Feynman diagrams contributing to the electroweak radiative corrections. Dashed lines stand for H, Z, γ, or W.

topological structure of the signal, the number of jets, the characteristics of the underlying event, and, of course, predictions for the background. This information allows to adjust in an optimal way experimental cuts and to measure the top quark through a kinematic analysis of its decay products. As a typical example the impact of gluon radiation on the top mass determination has been analysed recently. An idealized study has been performed e.g. in [70]. Radiated gluons are merged with the b jet from top decay or with the quark jet from $W \to q\bar{q}$ if they are found within a cone of opening angle

$$R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < R_{cut} \tag{2.20}$$

with respect to b, q or $\bar{q}$ and if their rapidity is below $|\eta| < 2$. In this case the gluon is considered as top decay product and hence contributes to its invariant mass. If the gluon jet falls outside the cuts, it is assigned the rest of the event.

Gluon radiation associated with top quark production, if erroneously associated with top decay, will thus increase the apparent m_t. Radiation from top decay, if outside the forementioned cuts, will however, decrease the measured mass of the quark. The interplay between the two compensating effects is displayed in Fig. 2.20. For a realistic $R_{cut} \approx 0.6$ a reduction Δm of around 2 GeV is predicted.

2.3 Single top production

Virtual W bosons, originating from $u \to dW$ splitting, can merge with bottom quarks from gluon splitting $g \to b\bar{b}$ to produce single top quarks in association with fairly collinear d- and $\bar{b}$- jets (Fig. 2.21). The interaction radius in the QCD gg fusion process shrinks with rising energy so that the cross section $\sigma(gg \to t\bar{t}\ ) \sim \alpha_s^2/\hat{s}$ [mod. log's] vanishes asymptotically. By contrast, the interaction radius in the weak fusion process [56] is set by the Compton wave length of the W boson and therefore asymptotically non-zero,

Figure 2.17: Relative correction to the parton cross section for $m_t = 250$ GeV (upper figure: $q\bar{q} \to t\bar{t}$, lower figure: $gg \to t\bar{t}$) (from [69]).

Figure 2.18: Relative correction to the hadronic cross section for $s = (16\ \mathrm{TeV})^2$ (from [69])

$\sigma \to G_F^2 m_W^2/2\pi$. The subprocess has to be folded with the quark-gluon luminosities

$$\sigma(p\bar{p} \to t + X) = \int_\tau^1 \frac{\mathrm{d}z}{z} \mathcal{L}_{ug}(\tau/z)\sigma(ug \to t + \bar{b} + d) \tag{2.21}$$

plus a similar contribution from $\bar{d}g \to t + \bar{b} + \bar{u}$. The fall-off of the total cross section $\sigma(pp \to t\bar{b})$ is less steep than for the QCD fusion processes. As a result, the Wg fusion process would have dominanted for large top quark masses ≥ 250 GeV at the LHC (Fig. 2.23).

For $m_t = 180$ GeV, the case of practical interest, single top production is about a factor 5 below the QCD reaction. Nevertheless, as shown in Fig.2.23, about 10^6 top quarks will be produced at the LHC by this mechanism at an integrated luminosity of $\int \mathcal{L} = 10^4 pb^{-1}$. Also at the TEVATRON this process should be accessible with the anticipated luminosity.

A close inspection of the various contributions to the subprocess $ug \to t + \bar{b} + d$ reveals immediately that the by far dominant part of the cross section is due to b exchange, with the b quark being near its mass shell. Since the b quark is almost collinear to the incoming gluon, this cross section is logarithmically enhanced $\sim \ln(m_t^2/m_b^2)$ over other mechanisms. This naturally suggests to approximate the process by the subprocess $u + b \overset{\mathrm{W}}{\to} d + t$ with the b-quark distribution generated perturbatively by gluon splitting based on massless evolution equations. The weak cross sections can be presented in a compact form,

$$\sigma(ub \overset{\mathrm{W}}{\to} dt) \quad = \quad \frac{G_F^2 m_W^2}{2\pi} \frac{(\hat{s} - m_t^2)^2}{\hat{s}\ (\hat{s} + m_W^2 - m_t^2)} \qquad \to \frac{G_F^2 m_W^2}{2\pi}$$

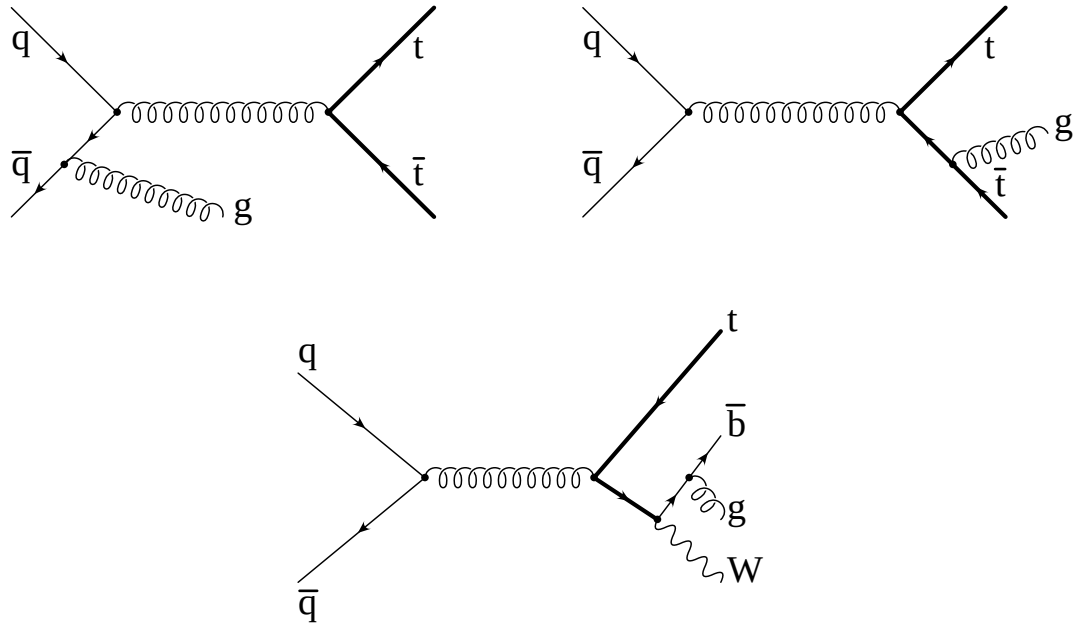


Figure 2.19: Gluon radiation from top production (upper) and decay (lower diagrams).

$$\sigma(d\bar{b} \stackrel{W}{\to} u\bar{t}) = \frac{G_F^2 m_W^2}{2\pi} \left[1 + \frac{\hat{s}\ (2m_W^2 - m_t^2) - 2m_W^2 m_t^2}{\hat{s}^{\ 2}} - \frac{(2\hat{s}\ + 2m_W^2 - m_t^2)m_t^2}{\hat{s}^{\ 2}} \log \frac{\hat{s}\ + m_W^2 - m_t^2}{m_W^2} \right] \to \frac{G_F^2 m_W^2}{2\pi}$$

and identically the same expressions for the $\mathcal{C}$-conjugate reactions.

Top quarks are created in $u + g$ collisions, anti-top quarks in $d + g$ collisions where the absorption of a W^- transforms a $\overline{b}$ quark to a $\overline{t}$ quark. The naïve expectation from valence quark counting for the ratio of $t/\overline{t}$ cross sections, $\sigma(u \to t) : \sigma(d \to \overline{t}) \sim 2 : 1$ is corroborated by a detailed analysis; in fact, the ratio turns out to be 2.1 for top quark masses of about 150 GeV.

The remaining possibilities for single top production are Compton scattering (Fig. 2.22(a))

$$g + b \to W + t \tag{2.22}$$

and the Drell Yan process (Fig. 2.22(b))

$$u + \bar{d} \to t + \bar{b} \tag{2.23}$$

The predicted cross sections are too small to be of practical interest. Single top quark production via the dominant mechanism (Fig. 2.21) offers a unique way for a measurement of the CKM matrix element V_{tb} and thus, indirectly of the top quark life time. As discussed in section 1.2.1 V_{tb} is strongly constrained to be very close to one for three generations — in a four generation model V_{tb} may be quite different from these expectations.

Figure 2.20: Shift in the apparent top mass as a function of R cut (from [70]).

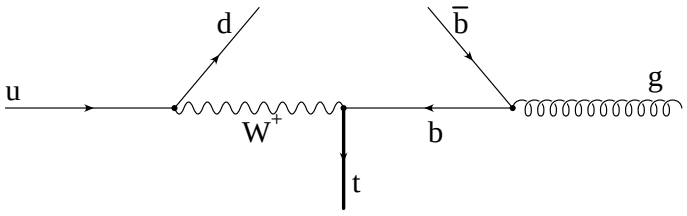


Figure 2.21: Characteristic diagram for single top production.

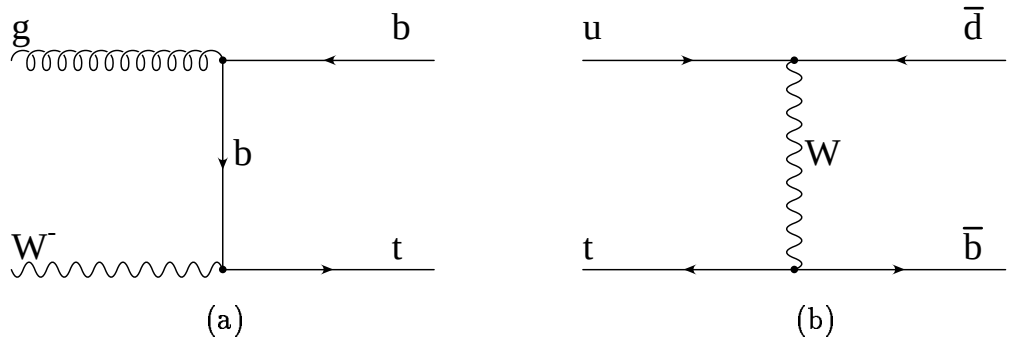


Figure 2.22: Characteristic diagrams for top production via Compton scattering (a) and Drell Yan process (b)

2.4 Quarkonium production

Both charm and bottom quarks have been discovered at hadron colliders in the form of quarkonium resonances J/ψ and Υ through their distinct signals in the $\mu^+\mu^-$ channel. The search for toponium at a hadron collider is, however, entirely useless. The broad ($\Gamma \sim$ 2 GeV) resonances decay with an overwhelming probability through single quark decay and are therefore indistinguishable from open top quarks produced close to threshold.

The situation could be different in extensions of the SM. Decays of a fourth generation b'

$$b' \to t + W \tag{2.24}$$

are suppressed by small mixing angles. Alternatively, if $m_{b'} < m_t + m_W$, the $b' \to c + W$ mode would have to compete with loop induced FCNC decays — leaving ample room for narrow quarkonium states. Another example would be the production of weak isosinglet quarks which are predicted in Grand Unified Theories. The decay of these quarks would again be inhibited by small mixing angles.

Of particular interest is the search for $\eta_{b'}$, the $^1S_0$ state composed of b' and $\bar{b}'$ [71, 72, 73]. It is produced with appreciable cross section. Its dominant decay mode

$$\eta_{b'} \to H + Z \tag{2.25}$$

is enhanced by the large Yukawa coupling, governing the coupling of the heavy quark to the Higgs and the longitudinal Z. For large $m_{b'}$ one obtains

$$\frac{\Gamma(\eta_{b'} \to H + Z)}{\Gamma(\eta_{b'} \to \gamma\gamma)} \sim \frac{m_{b'}^4}{M_W^4} \tag{2.26}$$

The branching ratios as functions of M_η are displayed in Fig. 2.24. The complete set of QCD corrections for leading and subleading annihilation decay modes can be found in [73]. They do not alter the picture significantly.

It should be emphasized that the decay $\eta_{b'} \to H + Z$ proceeds through the axial part of the neutral current coupling which, in turn, is proportional to the third component of the

Figure 2.23: Cross sections for several mechanisms of top quark production [44]. Various parametrizations and models of the parton densities [62] have been used.

weak isospin. Bound states of isosinglet quarks would, therefore, decay dominantly into two gluon jets.

The cross section for *open* $b\bar{b}'$ production at the LHC (with $m_{b'} = 300$ GeV) amounts to about 100 pb. The fraction of the phase space where bound states can be formed, i.e. for relative quark velocity $\beta < 0.1$, covers around 10^{-2} of the relevant region

$$\frac{\int_0^{0.1} \mathrm{d}\beta\beta}{\int_0^1 \mathrm{d}\beta\beta} = 10^{-2} \tag{2.27}$$

and indeed one predicts a production cross section somewhat less than 1 pb from a full calculation.

For a detailed calculation of the production cross section a proper treatment of the QCD potential is required to obtain a reliable prediction for the bound state wave function at the origin. The structure of the NLO corrections for the production cross section, in particular of the dominant terms, bears many similarities with the result for open production and for the Drell Yan process (eq. 2.8). For gluon fusion the partonic cross section is (in the $\overline{MS}$ scheme) given by [74, 75]

$$\hat{\sigma}^{gg} = \frac{1}{s}\frac{\pi^2}{3}\frac{R^2(0)}{M^3}\alpha_{\overline{\mathrm{MS}}}^2(M^2)\left\{\delta(1-z)\right.$$

Figure 2.24: Branching ratios of $\eta_{b'}$ for the dominant decay modes as functions of the bound state mass M.

$$+\frac{\alpha_{\overline{\rm MS}}(M^2)}{\pi}\left\langle\delta(1-z)\left(N_C\left(1+\frac{\pi^2}{12}\right)C_F\left(\frac{\pi^2}{4}-5\right)-\frac{4}{3}T_f\ln(2)\right)+N_C F(z)\right\rangle\Bigg\} \tag{2.28}$$

where

$$\begin{aligned} F(z) &= \Theta(1-z)\Bigg[\frac{11z^5+11z^4+13z^3+19z^2+6z-12}{6z(1+z)^2} \\ &\quad +4\left(\frac{1}{z}+z(1-z)-2\right)\ln(1-z)+4\left(\frac{\ln(1-z)}{1-z}\right)_+ \\ &\quad +\left(\frac{2(z^3-2z^2-3z-2)(z^3-z+2)z\ln(z)}{(1+z)^3(1-z)}-3\right)\frac{1}{1-z}\Bigg] \end{aligned} \tag{2.29}$$

and $z = M^2/\hat{s}$. Both Born term and the virtual correction are proportional to $\delta(1-z)$, the structure of the dominant term due to gluon splitting $\sim \ln(1-z)/(1-z)$ is again universal. Quark-gluon and quark-antiquark initiated subprocesses of order α_s^3 can be found in [74, 75]. It may be worth mentioning that the structure of QCD corrections to light Higgs production [76] is nearly identical to eq. 2.28. From Fig. 2.25 it is evident that $\eta_{b'}$ states with masses up to 1 TeV are produced at the LHC with sizeable rates. The fairly clean signature of the $Z + H$ decay mode might allow to discover these exotic quarkonia and the Higgs boson at the same time.

Figure 2.25: Production cross section for $\eta_{b'}$ including QCD corrections as function of the bound state mass M for $\sqrt{s} = 16$ TeV and 40 TeV.

Chapter 3

Top quarks in e^+e^- annihilation

A variety of reactions is conceivable for top quark production at an electron positron collider. Characteristic Feynman diagrams are shown in Fig.3.1. e^+e^- annihilation through the virtual photon and Z (Fig.3.1a) dominates and constitutes the reaction of interest for the currently envisaged energy region.

In addition one may also consider [77] a variety of gauge boson fusion reactions (Fig.3.1b-d) that are in close analogy to $\gamma\gamma$ fusion into hadrons at e^+e^- machines of lower energy. Specifically these are single top production,

$$e^+e^- \rightarrow \bar{\nu}e^-t\bar{b} \tag{3.1}$$

or its charge conjugate and top pair production through neutral or charged gauge boson fusion

$$\begin{aligned} e^+e^- &\rightarrow e^+e^-t\bar{t} \\ e^+e^- &\rightarrow \bar{\nu}\nu t\bar{t} \end{aligned} \tag{3.2}$$

The experimental observation of these reactions would allow to determine the coupling of top quarks to gauge bosons, in particular also to longitudinal W bosons and Z bosons, in the space-like region and eventually at large momentum transfers. This would constitute a nontrivial test of the mechanism of spontaneous symmetry breaking.

The various cross sections increase with energy in close analogy to $\gamma\gamma$ reactions, and eventually even exceed e^+e^- annihilation rates. However, at energies accessible in the foreseeable future these reactions are completely negligible: for an integrated luminosity of 10^{40} cm^{-2}, at $E_{cm} = 500\ GeV$ and for $m_t = 150\ GeV$ one expects about one $e^+e^-t\bar{t}$ event (still dominated by $\gamma\gamma$ fusion). At that same energy the cross sections for $e^+\nu\bar{t}b + c.c.$ and $\nu\bar{\nu}t\bar{t}$ final states are still one to two orders of magnitude smaller.

Another interesting class of reactions is e^+e^- annihilation into heavy quarks in association with gauge or Higgs bosons:

$$e^+e^- \rightarrow t\bar{t}Z \tag{3.3}$$

$$e^+e^- \rightarrow t\bar{b}W^- \tag{3.4}$$

Figure 3.1: Feynman diagrams for $t\bar{t}$ or $t\bar{b}$ production.

$$e^+e^- \rightarrow t\bar{t}H \tag{3.5}$$
$$e^+e^- \rightarrow t\bar{b}H^- \tag{3.6}$$

Two amplitudes contribute to the first reaction [78]: The $t\bar{t}$ system may be produced through a virtual Higgs boson which by itself was radiated from a Z (Fig.3.2). The corresponding amplitude dominates the rate and provides a direct measurement of the Yukawa coupling. The radiation of longitudinal Z's from the quark line in principle also carries information on the symmetry breaking mechanism of the theory. The transverse part of the $t\bar{t}Z$ coupling, i.e. the gauge part, can be measured directly through the cross

Figure 3.2: Amplitudes relevant for $e^+e^- \rightarrow t\bar{t}Z$ and for $e^+e^- \rightarrow t\bar{t}H$.

section or various asymmetries in $e^+e^- \to t\bar{t}$. The longitudinal part, however, could only be isolated with $t\bar{t}Z$ final states. For an integrated luminosity of 10^{40} cm^{-2} one expects only about 40 events (see sect. 3.1.7) and it is therefore not clear whether these can be filtered from the huge background and eventually used for a detailed analysis.

Light Higgs bosons may be produced in conjunction with $t\bar{t}$ [79]. They are radiated either from the virtual Z with an amplitude that is present also for massless fermions or directly from heavy quarks as a consequence of the large Yukawa coupling (Fig.3.2). The latter dominates by far and may therefore be tested specifically with heavy quark final states. The predictions for the rate will be discussed in sect. 3.1.7. Depending on the mass of the Higgs and the top quark, the reaction could perhaps be detected with an integrated luminosity of $10^{40} cm^{-2}$.

Top quark production in $\gamma\gamma$ collisions is conceivable at a "Compton collider". It requires special experimental provisions for the conversion of electron beams into well-focused beams of energetic photons through rescattering of laser light. A detailed discussion can be found in [80].

Chapter 3 will be entirely devoted to $t\bar{t}$ production in e^+e^- annihilation. Section 3.1 will be concerned with the energy region far above threshold — with electroweak aspects as well as with specific aspects of top hadronisation. The emphasis of section 3.2 will be on the threshold region which is governed by the interplay between bound state formation and the rapid top decay.

3.1 Top production above threshold

3.1.1 Born predictions

From the preceding discussion it is evident that the bulk of top studies at an e^+e^- collider will rely on quarks produced in e^+e^- annihilation through the virtual γ and Z, with a production cross section of the order of σ_{point}. For quarks tagged at an angle ϑ, the differential cross section in Born approximation is a binomial in $\cos\vartheta$

$$\frac{d\sigma}{d\cos\vartheta} = \frac{3}{8}\left(1+\cos^2\vartheta\right)\sigma_U + \frac{3}{4}\sin^2\vartheta\sigma_L + \frac{3}{4}\cos\vartheta\sigma_F \tag{3.7}$$

U and L denote the contributions of unpolarized and longitudinally polarized gauge bosons along the ϑ axis, and F denotes the difference between right and left polarizations. The total cross section is the sum of U and L,

$$\sigma = \sigma_U + \sigma_L \tag{3.8}$$

the forward/backward asymmetry is given by the ratio

$$A^{FB} = \frac{3}{4}\frac{\sigma_F}{\sigma} \tag{3.9}$$

The σ^i can be expressed in terms of the cross sections for the massless case in Born approximation,

$$\begin{aligned}
\sigma_B^U &= \beta\sigma^{VV} + \beta^3\sigma^{AA} \\
\sigma_B^L &= \frac{1}{2}\left(1-\beta^2\right)\beta\sigma^{VV} \\
\sigma_B^F &= \beta^2\sigma^{VA}
\end{aligned} \tag{3.10}$$

with

$$\begin{aligned}
\sigma^{VV} &= \frac{4\pi\alpha^2(s)e_e^2e_Q^2}{s} \\
&\quad + \frac{G_F(s)}{\sqrt{2}}e_ee_Q(v_e+\rho a_e)v_Q\frac{m_Z^2\left(s-m_Z^2\right)}{\left(s-m_Z^2\right)^2+\left(\frac{s}{m_Z}\Gamma_Z\right)^2} \\
&\quad + \frac{G_F^2}{32\pi}\left(v_e^2+a_e^2+2\rho v_ea_e\right)v_Q^2\frac{m_Z^4s}{\left(s-m_Z^2\right)^2+\left(\frac{s}{m_Z}\Gamma_Z\right)^2} \\
\sigma^{AA} &= \frac{G_F^2}{32\pi}\left(v_e^2+a_e^2+2\rho v_ea_e\right)a_Q^2\frac{m_Z^4s}{\left(s-m_Z^2\right)^2+\left(\frac{s}{m_Z}\Gamma_Z\right)^2} \\
\sigma^{VA} &= \frac{G_F\alpha(s)}{\sqrt{2}}e_e(\rho v_e+a_e)e_Qa_Q\frac{m_Z^2\left(s-m_Z^2\right)}{\left(s-m_Z^2\right)^2+\left(\frac{s}{m_Z}\Gamma_Z\right)^2} \\
&\quad + \frac{G_F^2}{16\pi}\left(2v_ea_e+\rho(v_e^2+a_e^2)\right)v_Qa_Q\frac{m_Z^4s}{\left(s-m_Z^2\right)^2+\left(\frac{s}{m_Z}\Gamma_Z\right)^2}
\end{aligned} \tag{3.11}$$

The fermion couplings are given by

$$v_f = 2I_3^f - 4e_f\sin^2\theta_w \quad , \qquad a_f = 2I_3^f \tag{3.12}$$

and the possibility of longitudinal electron polarization ($\rho = -1;+1;0$ for righthanded; lefthanded; unpolarized electrons) has been included. Alternatively one may replace $G_Fm_Z^2$ by

$$G_Fm_Z^2 = \frac{\pi\alpha(s)}{\sqrt{2}\sin^2\theta_W\cos^2\theta_W} \tag{3.13}$$

With $\sin^2\theta_W$ ($\approx$0.23) interpreted as $\sin^2\theta_{eff}$ [81] this formula accommodates the leading logarithms from the running coupling constant as well as the quadratic top mass terms in the threshold region.

3.1.2 Radiative corrections

QCD corrections to this formula are available for arbitrary m^2/s up to first order in α_s:

$$\begin{aligned}
\sigma &= \frac{(3-\beta^2)}{2}\beta\sigma^{VV}\left(1+\frac{4}{3}\frac{\alpha_s}{\pi}K_V\right) \\
&\quad + \beta^3\sigma^{AA}\left(1+\frac{4}{3}\frac{\alpha_s}{\pi}K_A\right)
\end{aligned} \tag{3.14}$$

The exact result [82] for $K_{V,A}$ can be found in [83]. These QCD enhancement factors are well approximated by [84]

$$\begin{aligned} K_V &= 1+\frac{4}{3}\alpha_s\left[\frac{\pi}{2\beta}-\frac{3+\beta}{4}\left(\frac{\pi}{2}-\frac{3}{4\pi}\right)\right] \\ K_A &= 1+\frac{4}{3}\alpha_s\left[\frac{\pi}{2\beta}-\left(\frac{19}{10}-\frac{22}{5}\beta+\frac{7}{2}\beta^2\right)\left(\frac{\pi}{2}-\frac{3}{4\pi}\right)\right] \\ \alpha_s &= \frac{12\pi}{25\log\left(4p_t^2/\Lambda^2\right)} \end{aligned} \tag{3.15}$$

The next to leading order corrections to K_V were calculated only recently [85, 86]. The scale in α_s chosen above was guessed on the basis of general arguments [67] which were confirmed by the forementioned complete calculations.

For small β these factors develop the familiar Couloumb enhancement $\sim \frac{2\pi\alpha_s}{3\beta}$, compensating the phase space $\sim \beta$. This leads to a nonvanishing cross section which smoothly joins the resonance region. Details of this transition will be treated in section 3.2.

To prepare this discussion, let us briefly study the limit of applicability of fixed order perturbation theory. The leading terms in the perturbative expansion close to threshold are obtained from Sommerfeld's rescattering formula $\left(x \equiv \frac{4}{3}\frac{\pi\alpha}{\beta}\right)$

$$\begin{aligned} K_V^{\rm Som} = \frac{x}{1-e^{-x}} &= 1+\frac{x}{2}+\frac{B_1x^2}{2!}-\frac{B_2x^4}{4!}+\frac{B_3x^6}{6!}\pm\ldots \\ &= 1+\frac{x}{2}+\frac{x^2}{12}-\frac{x^4}{720}+\frac{x^6}{5040}\pm\ldots \end{aligned} \tag{3.16}$$

with B_i being the Bernoulli numbers. At first glance one might require $x \leq 1$ for the perturbative expansion to be valid. However, significantly larger values of x are acceptable. The full Sommerfeld factor $K_V^{\rm Som}$ is remarkably well approximated by the first three terms of the series for surprisingly large x (only 6% deviation for $x = 4!$). For top quarks this corresponds to $\beta \approx 0.13 - 0.14$ and hence down to about 3 GeV above the nominal threshold. Upon closer inspection one also observes that the formula given in eq. (3.15) (a result of order α_s) coincides numerically well with the correction factor $K_V^{\rm Som}\left(1-\frac{16}{3}\frac{\alpha_s}{\pi}\right)$ which incorporates rescattering and hard gluon vertex corrections. The results presented in these lectures are based on the Born predictions plus $\mathcal{O}(\alpha_s)$ corrections.

Initial state radiation has an important influence on the magnitude of the cross section. $\sigma(s_{eff})$ is folded with the Bonneau Martin structure function, supplemented by the summation of large logarithms. A convenient formula for the non-singlet structure function in the leading logarithmic approximation has been obtained in [87], which is a natural extension of a formula proposed in [88]. This leads to a significant suppression by about a factor

$$\left(\frac{\delta W}{m_t}\right)^{\frac{2\alpha}{\pi}\left(\ln\frac{m_t^2}{m_e^2}-1\right)} \approx 0.5-0.6 \tag{3.17}$$

with $\delta W = 1-5\ GeV$ in the resonance and threshold region. The correction factor increases rapidly with energy, but stays below 0.9 in the full range under consideration (Fig.3.3).

Electroweak corrections to the production cross section in the continuum have been studied in [89]. Apart from a small region close to threshold they are negative. Relative to the G_F parametrized Born approximation they decrease the cross section by -6.3% to -9.3%, if m_t is varied between 100 and 200 GeV, m_H between 42 and 1000 GeV, and E_{cm} fixed at 500 GeV. QCD and electroweak corrections are thus of equal importance (Fig. 3.4).

Close to threshold and for relatively small Higgs boson masses a rapid increase of these corretions is observed (Fig. 3.5) which can be attributed to the attractive Yukawa potential induced by light Higgs boson exchange. Several GeV above threshold, and for m_H around or below 100 GeV it is more appropriate to split these corrections into hard and soft exchange and incorporate the latter in an instantaneous Yukawa potential [90].

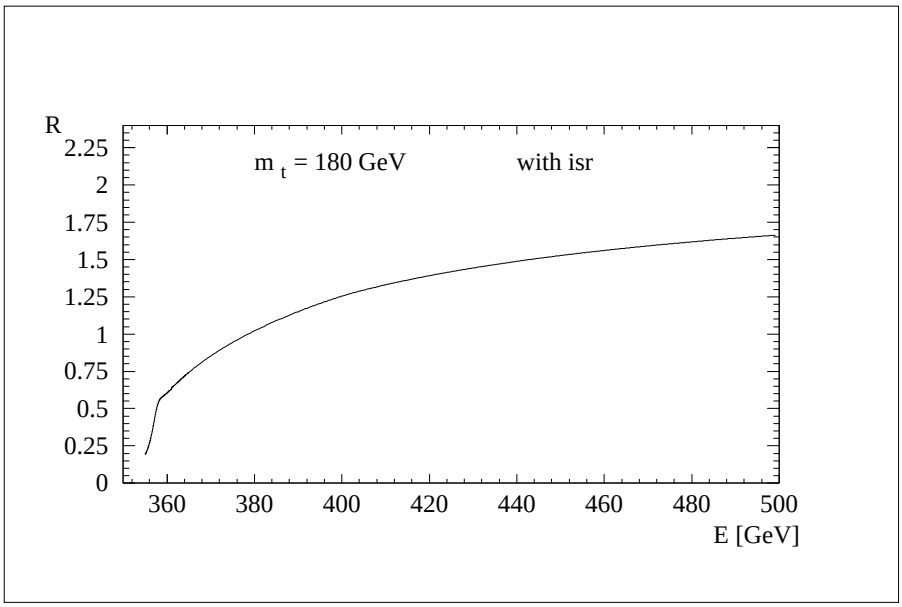


Figure 3.3: Cross section for $t\bar{t}$ production, including resonances, QCD corrections and initial state radiation in units of σ_{point}.

Longitudinal polarization

It should be mentioned that linear colliders might well operate to a large extent with polarized (electron) beams. The cross section for this case can be derived from (3.11). For top quarks the resulting right/left asymmetry

$$A_{LR} = (\sigma_L - \sigma_R)/(\sigma_L + \sigma_R) \tag{3.18}$$

is sizable (Fig.3.6) and amounts to about -0.4, reducing the production cross section with righthanded electrons. However, selection of righthanded electron beams decreases the W pair cross section even stronger, thereby enhancing the top quark signal even before cuts are applied. Electroweak corrections to A_{LR} in the threshold region have been calculated in [92].

Figure 3.4: Genuine electroweak corrections to top production in e^+e^- annihilation. From [89].

Figure 3.5: Relative size (in percent) of electroweak corrections in the threshold region for $m_t = 200$ GeV and different Higgs masses (from [91]).

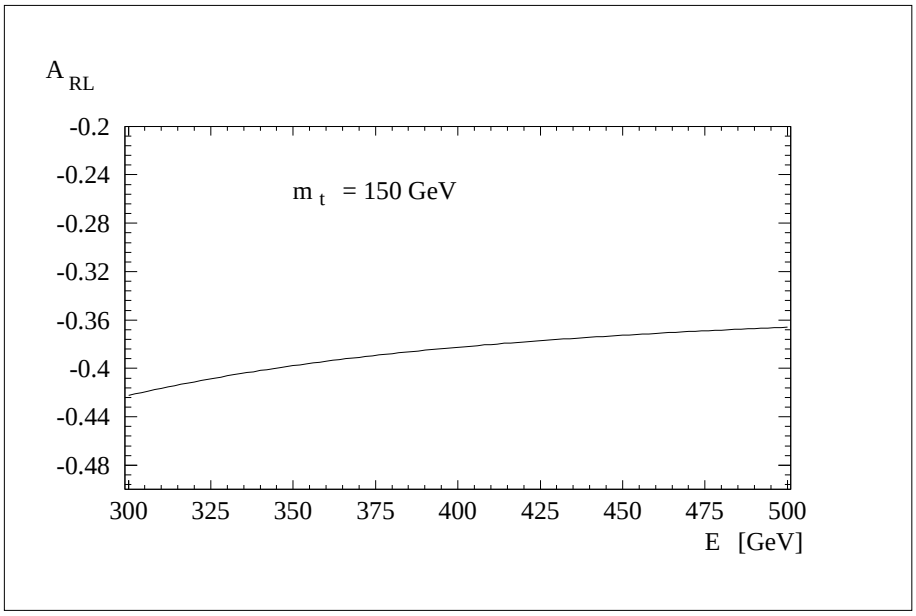


Figure 3.6: Right/left asymmetry as function of E_{cm} for $m_t = 150\ GeV$.

3.1.3 Top quark fragmentation

The experimental analysis of charm and bottom fragmentation functions has clearly demonstrated that heavy quark fragmentation is hard in contrast to the fragmentation of light quarks. This is a consequence of the inertia of heavy particles, the momentum of which is not altered much if a light quark $\bar{q}$ is attached to the heavy quark Q in the fragmentation process to form a bound state $(Q\bar{q})$, see e.g. [93]. At the same time soft infrared gluon radiation is damped if the color source is heavy.

For $m_t \geq 100$ GeV the strong fragmentation process and the weak decay mechanism are intimately intertwined [94]. The lifetime $\tau_* < \Lambda^{-1}$ becomes so short that the mesonic $(t\bar{q})$ and baryonic (tqq) bound states cannot be built–up anymore. Depending on the initial top quark energy, even remnants of the t quark jet may not form anymore [43]. Hadrons can be created in the string stretched between the t and the $\bar{t}$ only if the quarks are separated by about 1 fermi before they decay. If the flight path $\gamma\tau_*$ is less than 1/2 fm, the length of the $t - \bar{t}$ string is too short to form hadrons and jets cannot develop anymore along the flight direction of the top quarks. For $m_t \approx 180$ GeV top quark energies above 1 TeV are required to allow nonperturbative strings between t and $\bar{t}$. "Early" nonperturbative production of particles from the string between t and $\bar{t}$ is thus absent for all realistic experimental configurations. "Late" production from the b and $\bar{b}$ jets produced in top decays dominates. This is illustrated in Fig. 3.7. Early production dominates for $m_t = 90$ GeV, late production for $m_t = 120$ GeV and *a forteriori* for the actual value around 180 GeV.

The perturbative radiation of soft gluons, too, is interrupted by the t quark decay [95].

Figure 3.7: Rapidity distributions of "early" particles (full) and "late" ones (dashed) for three different top masses: 90 GeV in a) 100 GeV in b) and 120 GeV in c). (From [94].)

Figure 3.8: Sensitivity of the soft gluon distribution to the top width for back-to-backk t and b: $dN/d\cos(\theta_g)$ for $M = 140$ GeV, $\theta_b = 180°$, and $\Gamma = 0$, 0.7, 5, 20 GeV, and ∞ as marked. $W = 1000$ GeV. (From [96].)

The angular distribution (Θ) and the energy distribution (ω) of the radiated gluons is approximately given by

$$dP_g = \frac{4\alpha_s}{3\pi} \frac{\Theta^2 d\Theta^2}{\left[\Theta^2 + \frac{1}{\gamma^2}\right]^2 + \left[\frac{\Gamma}{\gamma\omega}\right]^2} \frac{d\omega}{\omega} \tag{3.19}$$

for a short–lived radiation source accelerated to $\gamma = E_t/m_t$. The gluons accumulate on the surface of a cone with half–aperture $\Theta_c \sim \gamma^{-1}$ for a long–lived t, but $\sim \gamma^{-1}\sqrt{\gamma\Gamma/\omega}$ if the particle decays quickly. The energy spectrum rises from zero to a maximum at $\omega \sim \gamma\Gamma$ before falling off $\sim 1/\omega$ for large ω, if the width is greater than the confinement scale Λ.

The impact of the finite width on the angular distribution of gluon radiation will be visible if $\omega \leq \Gamma_t E_t/m_t$. For a linear collider with c.m. energy of 2 TeV gluon jets with energies of 10 GeV and below would be affected [96]. The radiation pattern is shown in Fig. 3.8 for $m_t = 140$ GeV and $\sqrt{s} = 600$ GeV, with Γ_t tuned to different values in order to demonstrate the sensitivity of such a measurement. The impact of the conversion of gluons to hadrons has been ignored in this study. The picture is, further, complicated by the interference between radiation from top production and decay — a phenomen characteristic for unstable particles. These phenomena allow to probe the time evolution of hadronisation in a unique manner. Their understanding is a necessary prerequisite for any top mass measurement through a kinematic analysis of b-jet-W final states.

3.1.4 Static t parameters

Because of the large t mass, deviations from the Standard Model may manifest themselves in the top quark sector first. Examples in which the large mass is crucial are provided by multi–Higgs doublet models, models of dynamical symmetry breaking and compositeness. These effects can globally be described by form factors parametrizing the electroweak $t\bar{t}$ production current ($a = \gamma, Z$) and the weak (t, b) decay current ($a = -$) [97, 98],

$$j^a_\mu \sim F^a_{1L}\gamma_\mu P_L + F^a_{1R}\gamma_\mu P_R + \frac{i\sigma_{\mu\nu}q_\nu}{2m_t}[F^a_{2L}P_L + F^a_{2R}P_R] \tag{3.20}$$

[$P_{L,R}$ project on the left and right chirality components of the wave functions.] In the Standard Model, $F^-_{1L} = 1$ while all other F^-_i vanish; $F^\gamma_{1L} = F^\gamma_{1R} = 1$ and $F^\gamma_{2L} = F^\gamma_{2R} = 0$, analogously for the Z current. $\mathcal{CP}$ invariance requires $F^{\gamma,Z}_{2L} = F^{\gamma,Z}_{2R}$ in the $t\bar{t}$ production current, and equal phases for F^-_{1L} and F^-_{2R} *etc.* in the decay current. The static values of the form factors $F^{\gamma,Z}_2$ are the anomalous magnetic and electric dipole moments of the top quark.

The form factors are determined experimentally by measuring the angular distribution of the $t\bar{t}$ decay products, $e^+e^- \to t\bar{t}, t \to bW^+, W^+ \to f\bar{f}'$ *etc.* This requires the top quark to be treated as a free particle, the polarization of which not being affected by non–perturbative hadronic binding effects. This assumption is justified by the short lifetime of the top quark as discussed earlier. Details of the general helicity analysis can be found in the literature [97, 99].

3.1.5 Normal polarization of the top quarks

A non–zero component of the t polarization vector that is normal to the production plane can be generated only by the interference between complex helicity flip and non–flip amplitudes. Such relative phases can arise from $\mathcal{CP}$ violation but also from higher order loop corrections due to gluon exchange in the final state [100, 101, 97] or electroweak corrections involving Higgs and gauge bosons [97]. The QCD induced normal polarization is generally less than 5%, the electroweak normal polarization is smaller still. [By contrast, longitudinal and transverse polarization components within the $t\bar{t}$ production plane are generated already at the tree level of the electroweak interactions and they are large in general; see [100] for the discussion of details.]

3.1.6 Angular correlations of $t\bar{t}$ decay products

As stated in the previous chapter, top quarks are produced through the virtual photon and Z. In the threshold region they are polarized to a degree

$$P_t = A_{RL} \approx -0.4 \tag{3.21}$$

Assuming for the distribution of leptons from the decay of polarized top quarks

$$\frac{dN}{dx\, d\cos\theta} = f(x) + g(x)P_t \cos\theta \tag{3.22}$$

(with $g(x) = f(x)$ in the Standard Model, see eq.1.75) the angular distribution allows to test for the chirality of the tb current. Implicitly it was assumed that hadronization does not affect the top spin degrees of freedom [26, 102]. This assumption can be tested independently through the study of correlations between t and $\bar{t}$ decay products. In the threshold region the spins are correlated $\propto (1 + \frac{1}{3}\vec{s_+} \cdot \vec{s_-})$. This leads to the following correlated $\ell_+\ell_-$ distribution:

$$\frac{dN}{dx_+ dx_- d\cos\theta_{+-}} = f_+(x_+)f_-(x_-) + \frac{1}{3}g_+(x_+)g_-(x_-)\cos\theta_{+-} \tag{3.23}$$

where $f_+ = f_-$ and $g_+ = -g_-$. θ_{+-} denotes the angle between ℓ_+ and ℓ_-. After averaging the lepton energies,

$$\frac{dN}{d\cos\theta_{+-}} = 1 + \frac{1}{3}h_+h_-\cos\theta_{+-} \tag{3.24}$$

Note that the coefficient of the correlation term is $-h_+^2/3$ and hence always negative (assuming CP conservation). Since $|h_+| \leq 1$ it ranges between 0 and $-1/3$. This limiting value is assumed in the Standard Model. A detailed discussion with illustrative examples is given in [45].

3.1.7 Testing the Yukawa Coupling

With its relatively large Yukawa coupling $g_Y = \sqrt{2}m_t/v \approx 1$ the top quark is uniquely suited to test one of the basic ingredients of the Standard Model, the coupling between top quarks and the Higgs boson. The verification of this crucial prediction would confirm the mechanism for the generation of fermion masses and hence complete the measurement and analysis of basic couplings. Alternatively, any deviation would provide unambiguous proof for new physics.

Different strategies are at hand at an e^+e^-collider which are closely tied to the cms energy available and to the mass of the Higgs boson. For relatively light bosons a variety of possibilities appear to be promising: vertex corrections affect the cross section for $t\bar{t}$ production in the threshold region. For a collider in its early stage with an energy around 500 GeV it may well be the only option available and will be discussed more thoroughly in section 3.2. For higher energies, say around 1 TeV, a promising choice is the radiation of Higgs bosons from $t\bar{t}$ (see Fig. 3.2) with a cross section around 1 fb (Fig. 3.9) [79]. Alternatively one may analyse the top quark final states in conjunction with a Z boson. This reaction receives important contributions from the left one of the diagrams depicted in Fig. 3.2 if the Higgs mass happens to be relatively close to $2m_t$ [78]. Again, for simple kinematical reasons, high energies are crucial for the reaction to be accessible (Fig. 3.10).

3.2 Threshold behaviour

The previous section dealt with top quark production sufficiently far above threshold for the reaction to be well described by the Born cross section, modified slightly by QCD

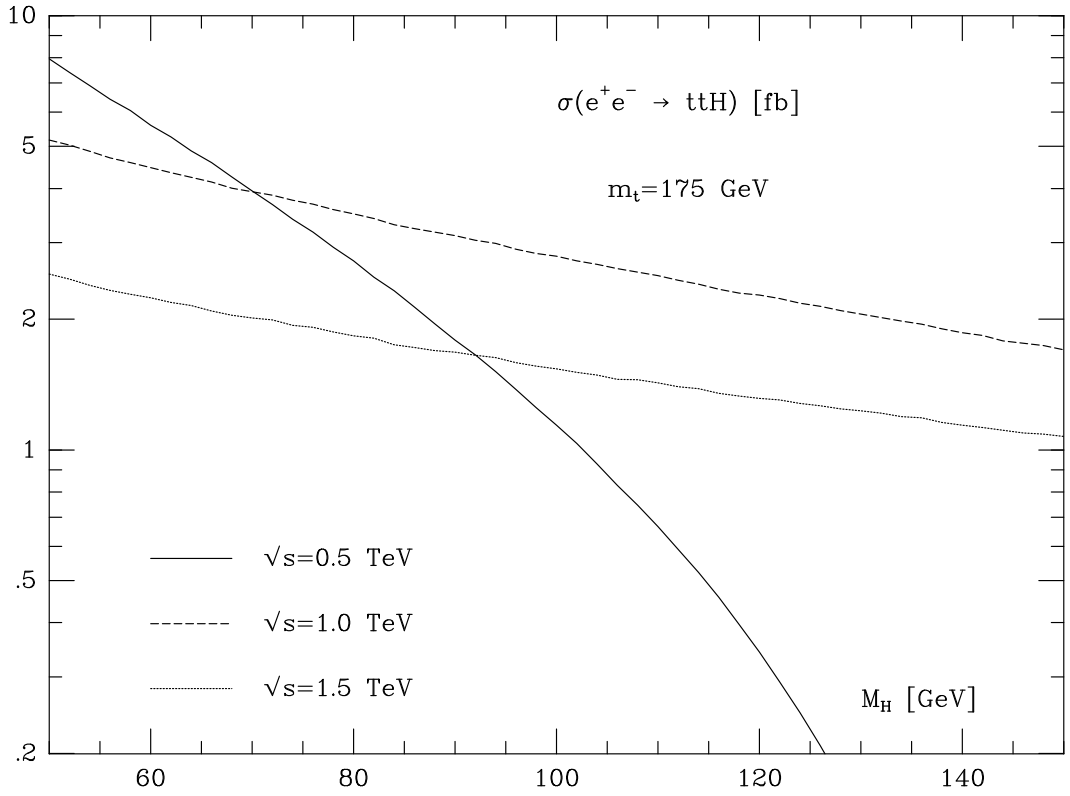


Figure 3.9: The cross section $\sigma(e^+e^- \to t\bar{t}H)$ (from [79]).

and electroweak corrections. This is in contrast to the situation in the threshold region, where QCD plays an important role and controls the cross section. Strong forces modify the Born prediction. They compensate the phase space suppression and enhance the production rate significantly, leading to a step function like behaviour at threshold. The large top decay rate also plays an important role. Quarkonium resonances cease to exist and merge into a structureless excitation curve which joins smoothly with the continuum prediction above the nominal threshold.

This sharply rising cross section allows to study top quarks in a particularly clean environment and with large rates. The following physics questions can be addressed:

- The QCD potential can be scrutinized at short distances, with the non perturbative tail cut off by the top decay. As a result Λ_{QCD} or α_s could be determined accurately.
- The top quark mass can be measured with a precision which is only limited by the theoretical understanding of the excitation curve, but in any case better than 500 MeV.
- Top quarks are strongly polarized (about 40%) even for unpolarized beams; and longitudinal beam polarisation will enhance this value even further. Detailed studies of top decays, in particular of the $V-A$ structure of the tbW coupling are therefore feasible.
- The interquark potential is — slightly — modified by the Yukawa potential induced by Higgs exchange. The excitation curve and the top quark momentum distribution

Figure 3.10: The Higgs mass (m_H) dependences of the total cross sections of $e^+e^- \to t\bar{t}Z$ for various top quark masses m_t. The c.m. energy $\sqrt{s}$ is set to be 500 GeV (a) and 1 TeV (b). (From [78].)

may therefore lead to an indirect measurement of the Yukawa coupling.

- The large number of top quarks in combination with the constrained kinematics at threshold, could facilitate the search for new decay modes expected in extensions of the SM.

With this motivation in mind the following points will be discussed: After a brief review of qualitative features of threshold production (section 3.2.1), the present status of our theoretical understanding of the total cross section will be presented in section 3.2.2.

The momentum distribution of top quarks and their decay products offers an alternative and complementary route to probe the interquark potential, as shown in section 3.2.3. Spin effects and angular distributions are sensitive towards the small P-wave contribution induced by the axial part of the neutral current. The theoretical framework and the resulting predictions are collected in section 3.2.4. Rescattering, relativistic corrections and other terms of order α_s^2 will be touched upon in section 3.2.5.

3.2.1 Introductory remarks

For a qualitative understanding it is illustrative to compare the different scales which govern top production close to threshold. The quarks are produced at a scale comparable to their Compton wave length

$$d_{\rm prod} \sim 1/m_t \tag{3.25}$$

Electroweak vertex corrections do not alter this behaviour significantly, since Z- or W-boson exchange proceeds at a distance $\approx 1/m_Z$ which is still short compared to scales characteristic for the bound state dynamics. For the QCD potential

$$V_{\rm QCD} = -\frac{4}{3}\frac{\alpha_s}{r} \tag{3.26}$$

one anticipates an effective coupling constant $\frac{4}{3}\alpha_s \approx 0.2$, if α_s is evaluated at the scale of the Bohr momentum

$$k_B \approx \frac{4}{3}\alpha_s\frac{m_t}{2} \approx 20GeV \tag{3.27}$$

The resulting Bohr radius

$$r_{\rm Bohr} = 1/k_B \tag{3.28}$$

is small compared to hadronic scales. The binding energy of the 1S level

$$E_B = \left(\frac{4}{3}\alpha_s\right)^2\frac{m_t}{4} \approx 2GeV \tag{3.29}$$

and, quite generally, the separation between different resonances, is smaller than the decay rate

$$2\Gamma_t \approx 3GeV \tag{3.30}$$

whence all resonances will merge and join smoothly with the continuum.

The coupling strength κ of the Yukawa potential

$$V_Y = -\frac{\kappa}{r} e^{-m_H r} \quad \text{with} \quad \kappa = \sqrt{2} G_F \frac{m_t^2}{4\pi} = 0.042 \tag{3.31}$$

is comparable to the QCD coupling $4/3\alpha_s = 0.2$ in magnitude. The exponential damping, however, with a cutoff $1/m_H \ll r_{\rm Bohr}$ and a lower limit $m_H > 65$ GeV, reduces the impact of the Yukawa potential quite drastically. (The situation may be different in multi-Higgs-models: the couplings could be enhanced and, even more important, the Higgs might be lighter!) Furthermore, the nonrelativistic treatment is no longer adequate and retardation effects must be taken into consideration.

The large top quark width plays a crucial role for the threshold behaviour, which is best understood in the framework of (nonrelativistic) Green's function techniques. The production of $t\bar{t}$ from a pointlike source (actually of extension $1/m_t$) at x' with frequency $\omega = E$ is characterized by the time dependent Green's function $G(\vec{r}, \vec{r}', t)$ which is a solution of the time dependent Schrödinger equation with a pointlike source term

$$(i\partial_t - H) G(\vec{r}, \vec{r}', t) = \delta(\vec{r} - \vec{r}') e^{-iEt} \tag{3.32}$$

with

$$H = \frac{p^2}{2m} + V(\vec{r}) \tag{3.33}$$

In the problem at hand $m = m_t/2$ is the reduced mass, $\vec{r}$ the relative distance between t and $\bar{t}$, and the width $\Gamma = 2\Gamma_t$. The location of the source is at the origin $\vec{r}' = 0$ by convention, and the second argument of G will be suppressed, $G(\vec{r}, t) \equiv G(\vec{r}, 0, t)$. For a qualitative discussion one may ignore the potential and obtains for a stable quark

$$\begin{aligned} G(\vec{r}, t) &= -\frac{m}{2\pi r} e^{i\sqrt{2mE}r} e^{-iEt} \\ \tilde{G}(\vec{p}, t) &= \frac{1}{E - \frac{p^2}{2m}} e^{-iEt} \end{aligned} \tag{3.34}$$

The corresponding current is flowing in radial direction from the source

$$\vec{j} = \frac{i}{2m} \left(G \vec{\nabla} G^* - G^* \vec{\nabla} G \right) = \vec{e}_r \frac{m^2 v}{4\pi^2 r^2} \tag{3.35}$$

with a constant flux through a sphere around the origin, reflecting the conservation of probability.

The width $\Gamma = 2\Gamma_t$ is introduced in the Schrödinger equation through the replacement

$$H \to H - i\Gamma/2 \tag{3.36}$$

and, consequently, through the substitution $E \to E + i\Gamma/2$ in eq. (3.34). The exponential damping of the flux in radial direction

$$\vec{j} = \vec{e}_r \frac{m^2 v}{4\pi^2 r^2} e^{i\left(\sqrt{2m(E+i\Gamma/2)} - \sqrt{2m(E-i\Gamma/2)}\right) r} \tag{3.37}$$

is most easily interpreted in two limiting cases. For $E \ll \Gamma$ the decrease

$$\vec{j} \sim \frac{\vec{e_r}}{r^2} e^{-\sqrt{2m\Gamma}r} \tag{3.38}$$

is solely driven by the large width, with a cutoff $x_c^{-1} = \sqrt{2m\Gamma} \approx 2.4$ GeV for realistic parameters. For $E \gg \Gamma$, on the other hand, the current decreases like

$$\vec{j} \sim \frac{\vec{e_r}}{r^2} e^{-\Gamma r/v} \tag{3.39}$$

In this latter case top quarks may travel appreciable distances, up to $x_{\rm cut} \approx v/\Gamma$. However, for realisitic beam energies they hardly propagate beyond the perturbative region.

Predictions for the Coulomb potential

The large top decay rate restricts the range of sensitivity to the short distance part of the potential

$$\tilde{V}(q) = -\frac{16\pi}{3}\frac{\alpha_V(q)}{q^2} \tag{3.40}$$

which is approximately Coulombic, with a logarithmic variation of α_V. Most of the qualitative features of top quark threshold physics can be understood even on the basis of the results for constant α_V, which are available in analytical form.

A remarkable feature of heavy quark production is the sharp rise of the cross section at threshold, a consequence of the attractive Coulomb force. The step function joins smoothly with the smeared resonances.

Let us try to quantify this aspect with the help of simple nonrelativistic quantum mechanics ($\alpha = \frac{4}{3}\alpha_s$, $m = m_t/2$). The narrow resonances below the nominal threshold (E=0) are located at $E_n = -E_{Ryd}/n^2$ with $E_{Ryd} = k_B^2/(2m) = \alpha^2 m/2$;

The production amplitude from a pointlike source with frequency $\omega = E$ is proportional to

$$\langle \vec{r} = 0 | \psi_n \rangle = \psi_n(0), \tag{3.41}$$

the rate correspondingly to

$$R \sim |\psi_n(0)|^2 \delta(E - E_n) \tag{3.42}$$

The wave function at the origin decreases with the third power of the radial quantum number,

$$|\psi_n(0)|^2 \sim 1/n^3 \tag{3.43}$$

their spacing becomes increasingly dense

$$\Delta E = E_{n+1} - E_n = E_{Ryd}\left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) \sim E_{Ryd}\frac{2}{n^3} \tag{3.44}$$

such that their average contribution to the cross section approaches a constant value (Fig. 3.11). Above threshold one has to project the state $\langle \vec{r} = 0|$ onto the Coulomb wave functions in the continuum $\psi_E(\vec{r})$. These replace the plane waves which are appropriate for the case where final state interaction is absent.

The production amplitude is thus governed by

$$\langle \vec{r} = 0 | \psi_E \rangle = \psi_E(0) \tag{3.45}$$

and the rate $\sim |\psi_E(0)|^2$. The threshold phase space factor v is thus compensated by the $1/v$ singularity in $|\psi_E(0)|^2$, and the cross section approaches a constant value for $E \to 0$.

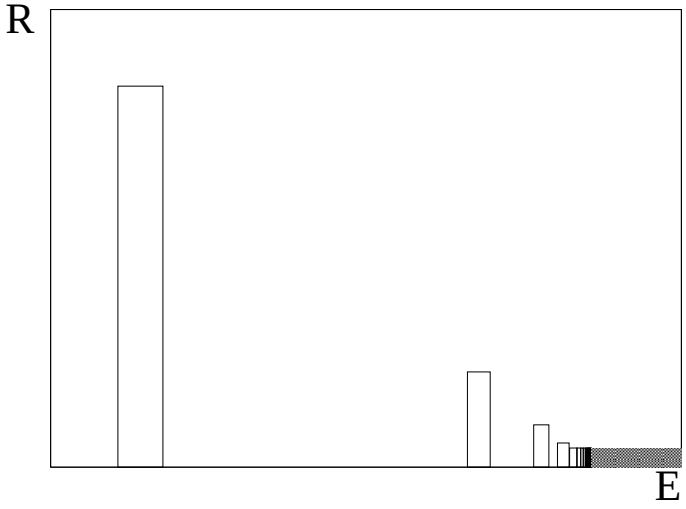


Figure 3.11: Schematic representation of resonances (wiht the area of the boxes adjusted to represent the weights of the respective delta functions) and the continuum cross section.

The explicit calculation yields

$$R(E) = \frac{9\pi}{2\alpha^2}\Gamma_e \delta(E - E_n) = 3Q_t^2 \frac{3}{2} \sum_n \frac{k_B^2}{m^2}\frac{1}{n^3}\pi\delta\left(E - E_{Ryd}\frac{1}{n^2}\right) \tag{3.46}$$

for energies below threshold, and

$$R(E) = 3Q_t^2 \frac{3}{2}\beta \frac{x}{1 - e^{-x}} \tag{3.47}$$

with $x = k_B/k = \pi\alpha/\beta$ for energies above threshold.[1]

The perturbative expansion breaks down in the limit $\beta \to 0$. The first term of this formal series

$$\beta \frac{x}{1 - e^{-x}} = \beta\left(\frac{x}{2} - \ldots\right) \tag{3.48}$$

[1]For a textbook discussion of Coulomb scattering states and a derivation of this result see e.g. [103, 104]

underestimates the exact result by a factor two.

Eq. (3.46) allows to connect the formalism based on narrow individual resonances with a formulation which is tailored to the situation at hand, namely wide overlapping resonances which merge into a smooth continuum.

Instead of summing the contributions from a large number of high radial excitations one may directly calculate the imaginary part of the Greens function for complex energy

$$\begin{aligned}\sigma(e^+e^- \to t\bar{t}) &= \frac{24\pi^2\alpha^2}{s}\frac{\rho_v(s)}{m_t^2}\left(1-\frac{16}{3}\frac{\alpha_s}{\pi}\right)\sum_n |\psi_n(0)|^2 \frac{\Gamma_t}{(E_n-E)^2+\Gamma_t^2} \\ &= \frac{24\pi^2\alpha^2}{s}\frac{\rho_v(s)}{m_t^2}\left(1-\frac{16}{3}\frac{\alpha_s}{\pi}\right)\sum_n Im\frac{\psi_n(0)\psi_n^*(0)}{E_n-E-i\Gamma_t} \\ &= -\frac{24\pi^2\alpha^2}{s}\frac{\rho_v(s)}{m_t^2}\left(1-\frac{16}{3}\frac{\alpha_s}{\pi}\right) ImG(0,0,E+i\Gamma) \qquad (3.49)\end{aligned}$$

The factor $\rho_v(s)$ incorporates the contributions from the intermediate photon and Z and is given by

$$\rho_v^{Born}(s) = \left|e_t e_e + \frac{1}{y^2}\frac{v_t v_e M_\theta^2}{s-M_Z^2+iM_Z\Gamma_Z}\right|^2 + \left|\frac{1}{y^2}\frac{v_t a_e M_G^2}{s-M_Z^2+iM_Z\Gamma_Z}\right| \qquad (3.50)$$

$$v_f = 2I_{3f} - 4e_f \qquad a_f = 2I_{3f} \qquad y = 16\sin^2\theta_W\cos^2\theta_W$$

($\alpha = \alpha_{\rm eff} = 1/128$ has been adopted in the numerical evaluation. Radiative corrections to this formula have been discussed in [92].)

The problem can be solved in closed analytical form for an exact Coulomb potential [105]

$$\begin{aligned}\mathrm{Im}G_{E+i\Gamma_t}(0,0) &= -\frac{m_t^2}{4\pi}\left[\frac{k_2}{m_t}+\frac{2k_B}{m_t}\arctan\frac{k_2}{k_1}\right. \\ &\quad \left.+\sum_{n=1}^{\infty}\frac{2k_B^2}{m_t^2\pi^4}\frac{\Gamma_t k_B n + k_2\left(n^2\sqrt{E^2+\Gamma_t^2}+\frac{k_B^2}{m_t}\right)}{\left(E+\frac{k_B^2}{m_t n^2}\right)^2+\Gamma_t^2}\right] \\ k_{1,2} &= \left[m_t\left(\sqrt{E^2+\Gamma_t^2}\mp E\right)/2\right]^{1/2}, \\ k_B &= \frac{2}{3}\alpha_s m_t. \qquad (3.51)\end{aligned}$$

To arrive at a realistic prediction of the total (and, in sect 3.2.3 of the differential) cross section the Coulomb potential must be replaced by the realistic QCD potential.

3.2.2 The QCD potential

On the basis of earlier conceptual work in [106, 107] the asymptotic behaviour of the static potential has been derived in [108, 109]. In momentum space the potential reads

in the $\overline{MS}$ subtraction scheme

$$V\left(Q^2, \alpha_{\overline{MS}}(Q^2)\right) = -\frac{16\pi}{3}\frac{\alpha_{\overline{MS}}(Q^2)}{Q^2}\left[1 + \left(\frac{31}{3} - \frac{10}{9}n_f\right)\frac{\alpha_{\overline{MS}}(Q^2)}{4\pi}\right] \tag{3.52}$$

The renormalization scale μ^2 has been chosen as Q^2, and n_f refers to the number of massless quarks. Employing standard arguments based on the renormalization group, the Q^2 expansion of $\alpha_{\overline{MS}}(Q^2)$ is given by

$$\begin{aligned} \frac{\alpha_{\overline{MS}}(Q^2)}{4\pi} &= \frac{1}{b_0 \log\left(Q^2/\Lambda^2_{\overline{MS}}\right)}\left[1 - \frac{b_1}{b_0^2}\frac{\log\log\left(Q^2/\Lambda^2_{\overline{MS}}\right)}{\log\left(Q^2/\Lambda^2_{\overline{MS}}\right)}\right] \\ b_0 &= 11 - \frac{2}{3}n_f, \qquad b_1 = 102 - \frac{38}{3}n_f \end{aligned} \tag{3.53}$$

The leading behaviour of the potential at small distances [$\gamma_E = 0.5772\cdots$]

$$\begin{aligned} V(r) &= \frac{16\pi}{(33-2n_f)r\log 1/(\Lambda_{\overline{MS}}r)^2} \\ &\quad \left[1 - \frac{b_1}{b_0^2}\frac{\log\log 1/(\Lambda_{\overline{MS}}r)^2}{\log 1/(\Lambda_{\overline{MS}}r)^2} + \frac{\left(\frac{31}{3} - \frac{10}{9}n_f\right)/b_0 + 2\gamma_E}{\log 1/(\Lambda_{\overline{MS}}r)^2} + \cdots\right] \end{aligned} \tag{3.54}$$

is thus directly given in terms of the QCD scale parameter Λ. The exploration of $V(r)$ for small distances could thus lead to a direct determination of Λ. For quark masses above 50-100 GeV the ground state properties become independent of the potential in the nonperturbative region. As discussed in the previous section the large decay rate acts as a cutoff and the predictions are fairly insensitive to the actual regularisation. However, an additional constant which can be traded against a shift in m_t must be carefully calibrated.

In practice one connects the theoretically predicted short distance part smoothly with the empirically determined potential above $\sim$0.1 fermi. The asymptotic form given in (3.52) is based on the assumption that n_f species of light quarks, taken as massless, contribute to the vacuum polarization, and heavier ones are ignored. The value of $\Lambda_{\overline{MS}}$ in (3.53) must be properly related to $\Lambda_{\overline{MS}}$ as determined from other experiments with a different number of effective light flavors [110, 111]. For the momentum range of around 15 GeV explored by the $t\bar{t}$ system n_f =5 seems adequate.

In the subsequent discussion the Green's function will be calculated in momentum space with the help of the Lippman-Schwinger integral equation [112]. The representation of the QCD potential in momentum space with the large Q^2 behaviour given by eqs. 3.52 and 3.53 will be employed. The intermediate and small momentum dependence will be based on Richardson's potential. This choice allows to vary $\alpha_{\overline{MS}}(M_Z^2)$ (or equivalently $\Lambda_{\overline{MS}}$) between 0.11 and 0.13, while maintaining a smooth Q^2 dependence of $\alpha_V(Q^2)$ (Fig. 3.12). An additive constant in coordinate space (corresponding to a δ-function in momentum space) is adjusted to fix $V(r = 1GeV^{-1}) = -1/4$ GeV for arbitrary $\alpha_{\overline{MS}}$. This constraint avoids the unmotivated and uncontroled variation of the long distance part of $V(r)$ with a change in $\alpha_{\overline{MS}}$. The potential in coordinate space is shown in Fig. 3.13.

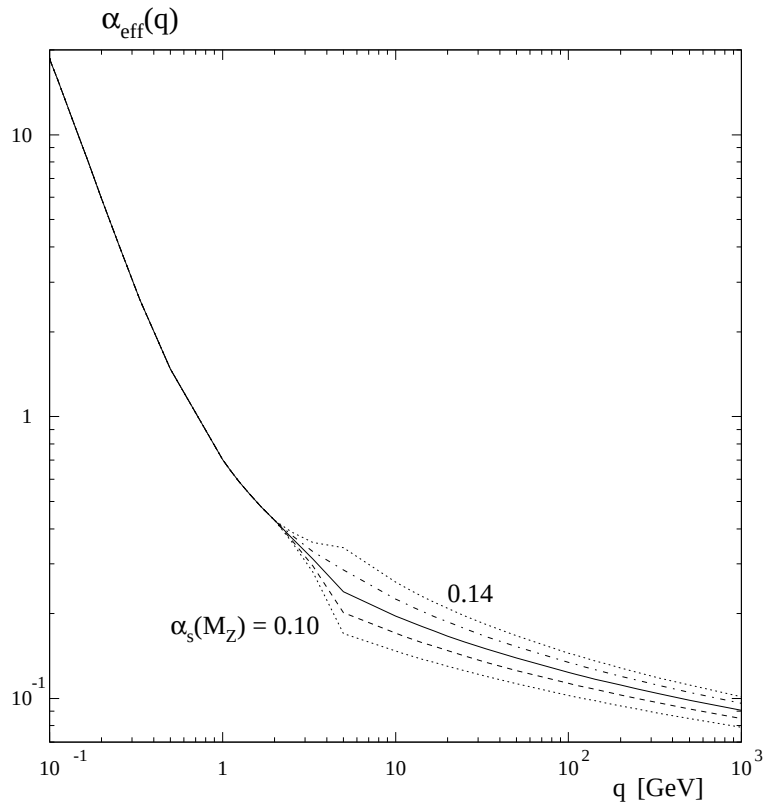


Figure 3.12: α_{eff} for different values of $\alpha_s(M_Z)$: solid: 0.12, dashed: 0.11, dashed-dotted: 0.13, dotted: 0.10 and 0.14.

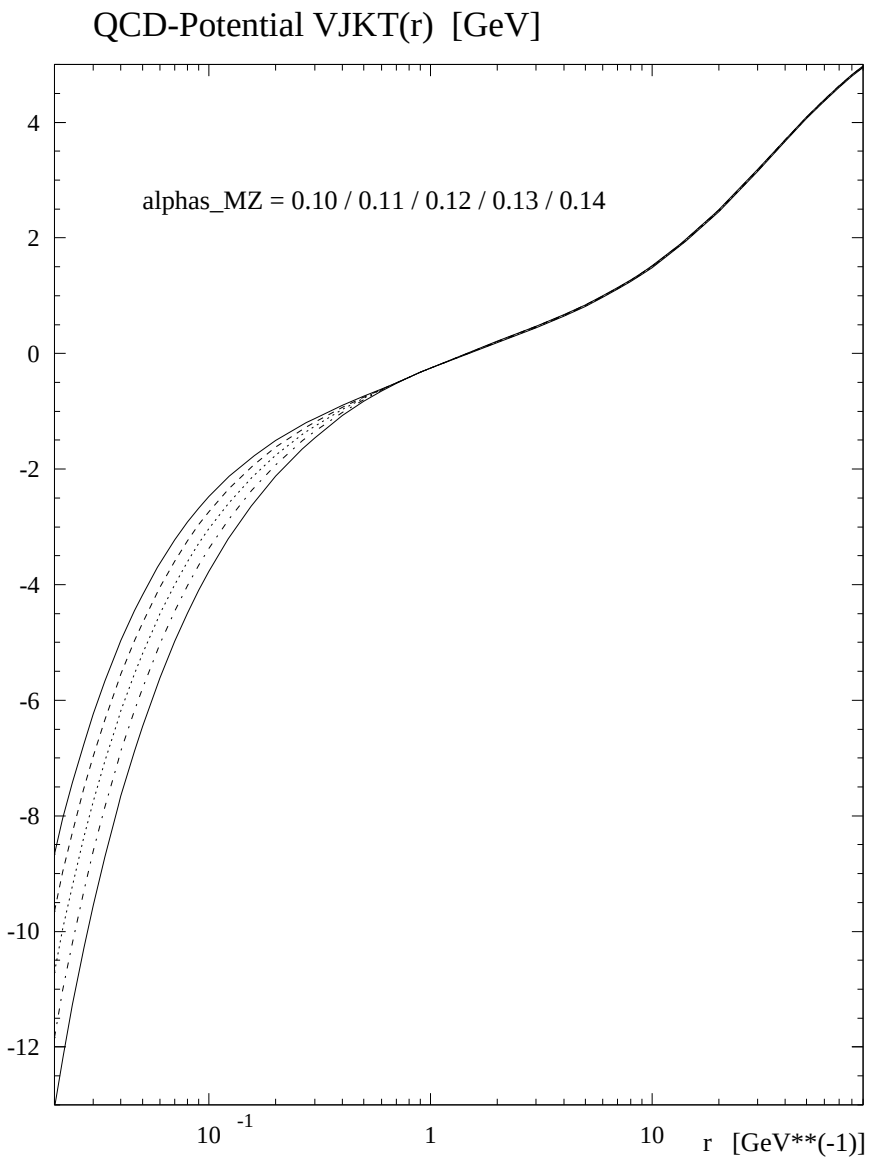


Figure 3.13: QCD potential in the position space V_{JKT} for different values of $\alpha_s(M_Z)$: solid: 0.12, dashed: 0.11, dashed-dotted: 0.13, dotted: 0.10 and 0.14.

3.2.3 Realistic predictions for $\sigma_{t\bar{t}}$.

For a realistic QCD potential the Green's function can only be calculated with numerical methods. An elegant algorithm for a solution in coordinate space has been suggested in [113]. As a consequence of the optical theorem (see also eqn. (3.49)) only the imaginary part of $G(\vec{r}=0,\vec{r}'=0,E+i\Gamma_t)$ is needed to predict the total cross section. The differential equation for the Green's function

$$\left[(E+i\Gamma_t)-\left(-\frac{\nabla^2}{2m}+V(\vec{r})\right)\right]G(\vec{r},\vec{r}'=0,E+i\Gamma_t)=\delta(\vec{r}) \tag{3.55}$$

is solved in a way which provides direct access to $\mathrm{Im}G(\vec{r}=0,\vec{r}'=0,E+i\Gamma_t)$ without the need to calculate the full $\vec{r}$ dependence. Alternatively, in [112, 114] the Green's function in momentum space was obtained from the Lippmann-Schwinger equation

$$\begin{aligned} G(\vec{p},E+i\Gamma_t) &= G_0(\vec{p},E+i\Gamma_t)+G_0(\vec{p},E+i\Gamma_t) \\ &\quad \times\int\frac{d\vec{q}}{(2\pi)^3}\tilde{V}(\vec{p}-\vec{q})G(\vec{q},E+i\Gamma_t) \\ G_0(\vec{p},E+i\Gamma_t) &= \frac{1}{E-p^2/2m+i\Gamma_t} \end{aligned} \tag{3.56}$$

The total cross section is in this case obtained from the integral over the differential distribution

$$\frac{d\sigma}{d^3p}=\frac{3\alpha^2}{\pi s m_t^2}\rho_v(s)\Gamma_t|G(\vec{p},E+i\Gamma_t)|^2 \tag{3.57}$$

This second formulation is particularly suited to introducing a momentum and energy dependent width $\Gamma(p,E)$ which allows to incorporate the phase space suppression and certain α_s^2 rescattering corrections to be discussed below in section 3.2.6.

It is well known that the coupling of the virtual photon to the quarkonium boundstate is modified by "hard" gluon exchange. The vertex correction to the vector current produces an additional factor $\left(1-\frac{16}{3}\frac{\alpha_s}{\pi}\right)$ for the quarkonium decay rate into e^+e^- through the virtual photon or Z. This factor can be calculated by separating the gluon exchange [115] correction to the vertex into the instanteneous potential piece and a remainder which is attributed to gluons with high virtualities of order m_t. A similar approach has been developed in [90] for Higgs exchange. The vertex correction is again decomposed into a part which is given by the instantaneous Yukawa potential

$$V_{Yuk}(r)=-\kappa\frac{e^{-m_H r}}{r} \tag{3.58}$$

with $\kappa=\sqrt{2}Gm_t^2/4\pi$ and a remainder which is dominated by highly virtual Higgs exchange. The rapid increase of the correction in the threshold region (cf. sect. 3.1.2) is driven by the potential; the remainder, the hard vertex correction, is fairly energy independent. The total cross section is thus sensitive to the top mass, the width (which in the SM is uniquely determined by m_t), the strong coupling constant α_s and the mass of the

Higgs boson. This dependence is illustrated in Figs. 3.14-3.17. Apart from the trivial shift of the threshold due to a change in m_t the shape of σ is affected by the rapidly increasing width of the top quark which amounts to 0.81 GeV, 1.57 GeV and 2.24 GeV for $m_t = 150$ GeV, 180 GeV, and 200 GeV respectively. This is demonstrated in Fig. 3.14: A fairly

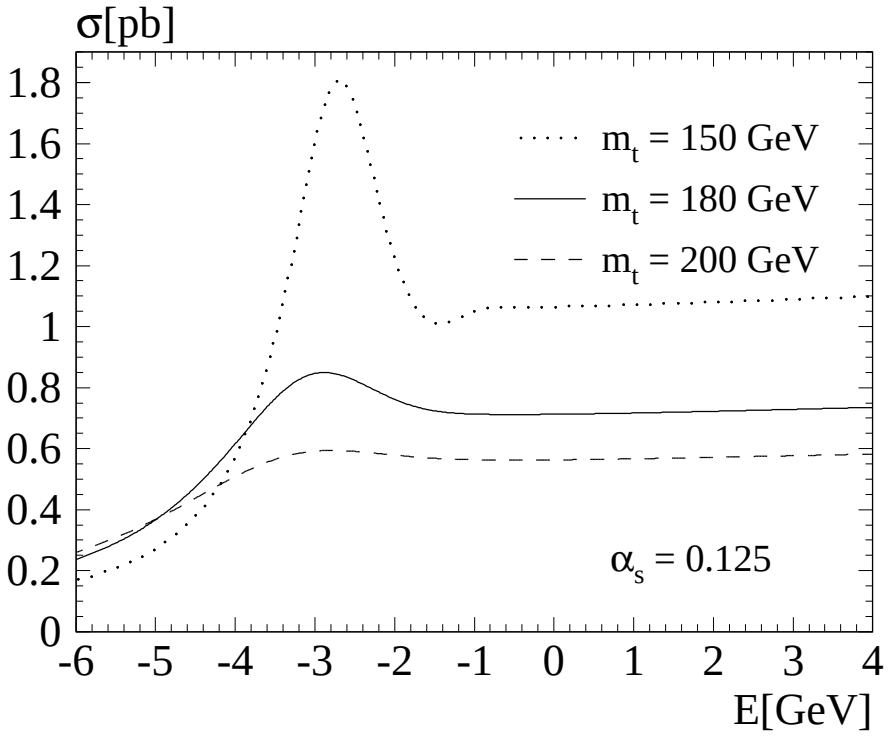


Figure 3.14: Total cross section as function of $E = \sqrt{s} - 2m_t$ for three values of the top quark mass.

pronounced 1S peak is still visible for $m_t = 150$ GeV, for $m_t = 200$ GeV, however, only a smooth shoulder is predicted. The behaviour is qualitatively very similar, if we keep m_t fixed say at 180 GeV and decrease or increase Γ_t be the corresponding amount. The shape of the cross section will therefore allow to determine the width of the top quark. A qualitatively very different response is observed towards a change in α_s (Fig. 3.15). The binding energy increases with α_s, the aparent threshold is thus lowered (This is the reason for the strong correlation between α_s and m_t in the experimental analysis based on σ_{tot} only [35, 116].) and the height of the "would-be resonance" is increased. Even several GeV above threshold one observes a slight increase of the cross section with α_s, a consequence of the enhanced attraction between t and $\bar{t}$ (cf. eq. 3.47). The impact of the running of α_s on the shape of the cross section is evident from Fig. 3.15. The full QCD prediction with running α_s (for $\alpha_{\overline{MS}}(M_Z^2) = 0.125$) is compared to the prediction for a Coulomb potential with α_s fixed. It is impossible to describe the height of the peak and the continuum above with the same value of α_s, even allowing for an arbitrary additive constant V_0. The influence of a variation in m_H is shown in Fig. 3.17. Cross section

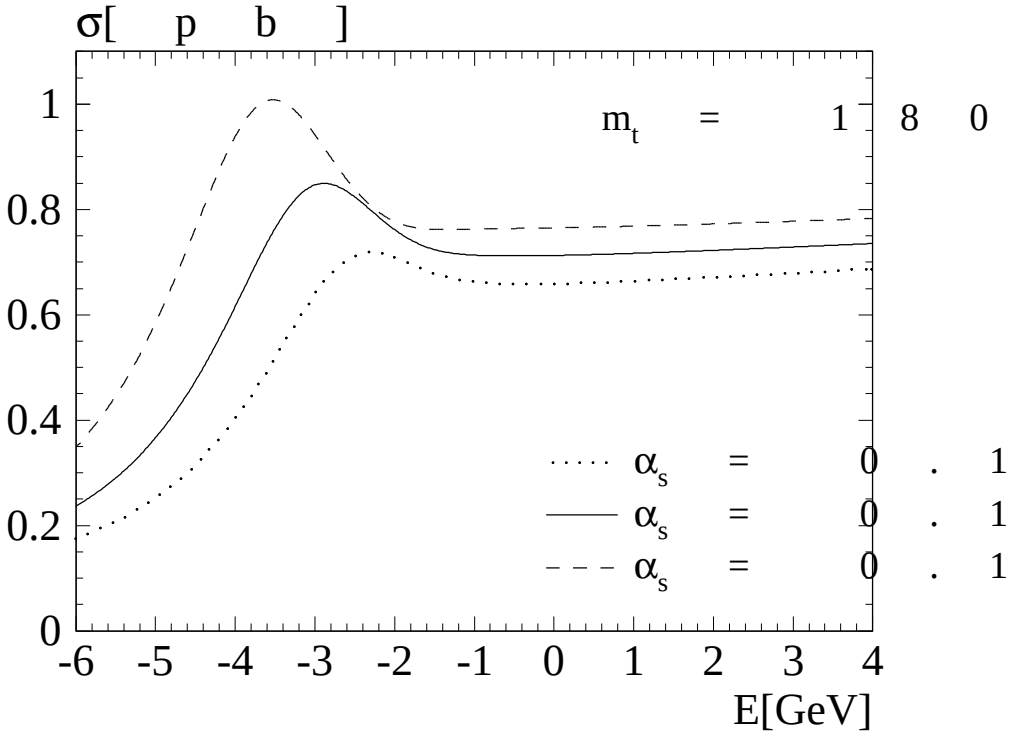


Figure 3.15: Total cross section as function of $E = \sqrt{s} - 2m_t$ for three values of α_s.

measurements with a precision better than 10% will become sensitive to the effect of a light Higgs boson.

Up to this point the amplitude induced by virtual Z and γ are included in Born approximation only. Electroweak corrections and initial state radiation are neglected. A detailed discussion of electroweak corrections to the cross section and the left right asymmetry in the context of the SM can be found in [92]. The corresponding discussion for the two-Higgs-doublet model is presented in [117]. In this model one might encounter enhanced Yukawa couplings which would amplify the effect under discussion.

Initial state radiation leads to a fairly drastic distortion of the shape of the cross section, in particular to a smearing of any pronounced structure. This is illustrated in Fig. 3.18 where the predictions with and without initial state radiation are compared for otherwise identical parameters.

Beamstrahlung and the energy spread of the beam lead to a further smearing of the apparent cross section. These accelerator dependent issues are treated in more detail in [35]. (For a related discussion see also [116].)

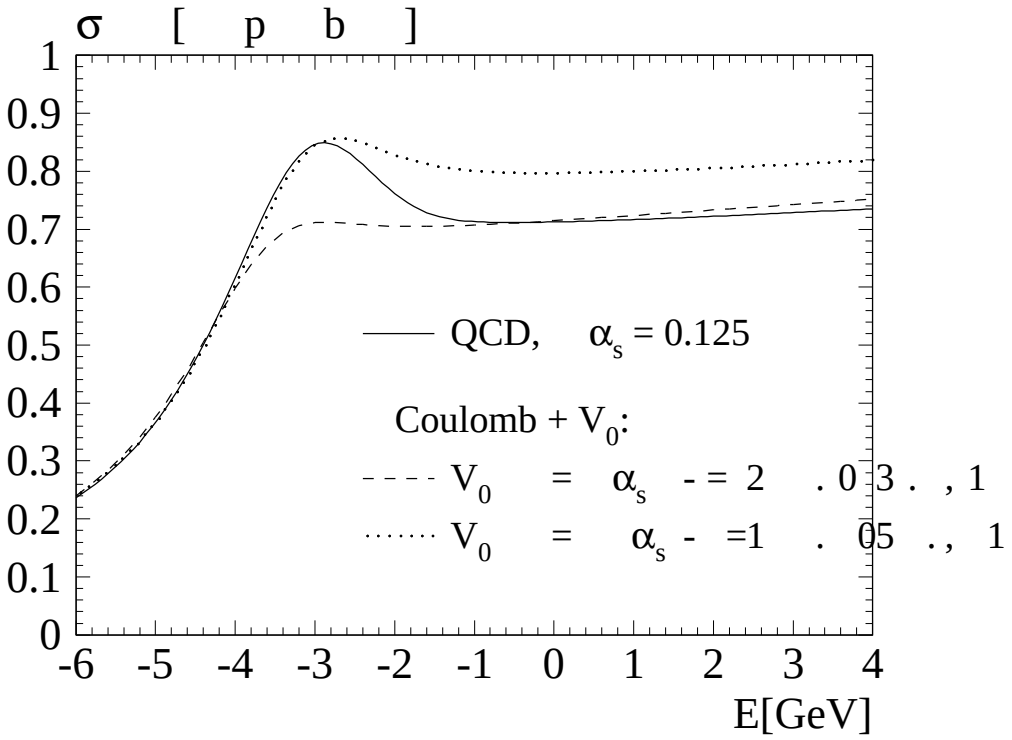


Figure 3.16: Comparison between the predicted cross section for constant (dashed and dotted lines) and running (solid lines) α_s.

3.2.4 Momentum distributions of top quarks

The Green's function in momentum space and the momentum distribution of top quarks (and thus their decay products) are intimately related. For a narrow quarkonium resonance with orbital quantum number n the quarks' momentum distribution is evidently given by the wave function in momentum space

$$\frac{dN}{d\vec{p}} = \frac{|\tilde{\psi}_n(\vec{p})|^2}{(2\pi)^3} \tag{3.59}$$

For J/ψ or Υ this distribution is not directly accessible to experiment since these states decay through $Q\bar{Q}$ annihilation only. For toponium, however, which is dominated by single quark decay, the decay products carry the information of their parent momentum and hence allow for the reconstruction of the original quark momentum distribution [118].

For one individual resonance this leads to the differential $t\bar{t}$ cross section (without Z contribution and transverse gluon correction).

$$\frac{d\sigma_n}{d\vec{p}}(\vec{p}, E) = \frac{3\alpha^2 Q_t^2}{\pi s m_t^2} |\psi_n(\vec{r}=0)|^2 \frac{\Gamma_t}{(E-E_n)^2 + \Gamma_t^2} |\tilde{\psi}_n(\vec{p})|^2 \tag{3.60}$$

Once Γ_t is sufficiently large, interferences between different radial excitations become important and the right-hand side of this equation has to be replaced by the square of

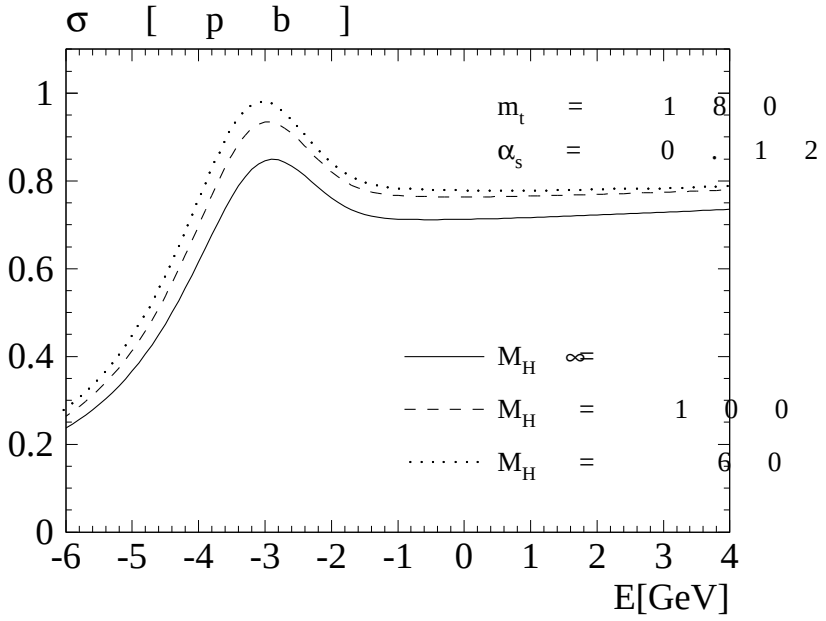


Figure 3.17: Total cross section as function of $E = \sqrt{s} - 2m_t$ for different values of the Higgs mass.

the Green's function [119, 112]

$$\frac{d\sigma_n}{d\vec{p}}(\vec{p}, E) = \frac{3\alpha^2 Q_t^2}{\pi s m_t^2}\Gamma_t \left|G(\vec{p}, E + i\Gamma_t)\right|^2 \tag{3.61}$$

with

$$G(\vec{p}, E + i\Gamma_t) = \int d\vec{r} e^{i\vec{p}\vec{r}} G(\vec{r}, \vec{r}' = 0, E + i\Gamma_t) \tag{3.62}$$

As discussed in sect. 3.2.2, the Green's function can be obtained in momentum space as a solution of the Lippmann-Schwinger equation. For an energy close to the 1S peak it exhibits a fairly smooth behaviour reminiscent of the 1S wave function in momentum space (Fig. 3.19). With increasing energy an oscillatory pattern of the amplitude is observed, and a shift towards larger momenta (Fig. 3.19). These results are intentionally displayed for $m_t = 120$ GeV, where the oscillations are still clearly visible, in contrast to m_t=180 GeV where all oscillations are smeared by the large width Γ_t. The corresponding predictions for the distributions at $m_t = 180$ GeV are displayed in Fig. 3.20. The transition from a wide distribution below the nominal threshold to a narrow one with the location of the peak determined by trivial kinematics is clearly visible. The impact on the energy distribution of the W's from top decay is shown in Fig. 3.21.

To characterize the momentum distribution by a single parameter, one may either choose its peak value, or the expectation value of the modulus of the momentum $\langle p \rangle$, the latter

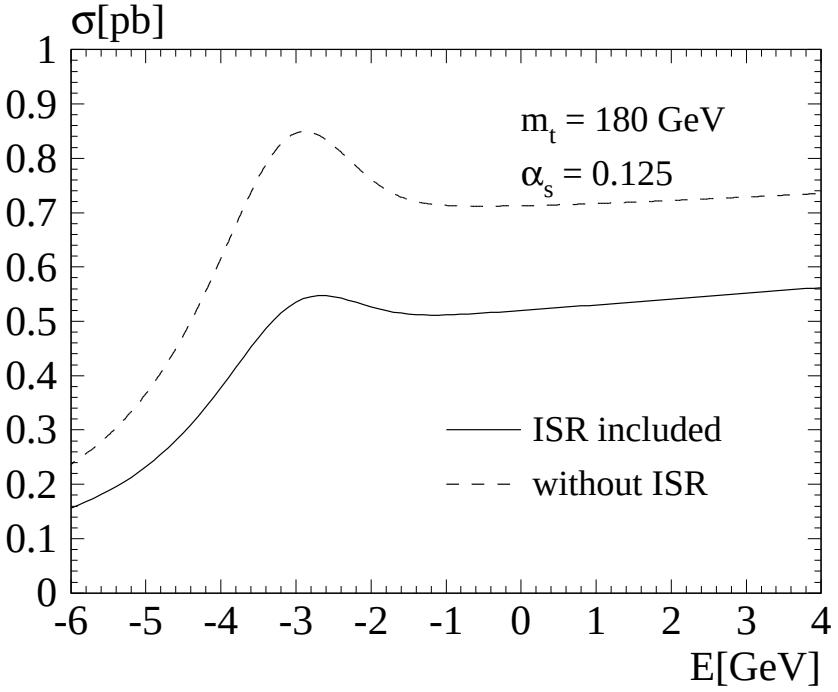


Figure 3.18: Comparison of the $t\bar{t}$ production cross section without (dashed) and with (solid line) initial state radiation.

being well adopted to the experimental analysis. In the situation at hand the definition of $\langle p \rangle$ has to be introduced with some care. The free Green's function G_0 (see eq. 3.56) drops $\sim p^{-2}$ for large momenta and this behaviour is recovered also in the presence of interaction. The expectation value $\langle p \rangle$ diverges logarithmically with the cutoff. In the narrow width approximation one finds for the leading terms

$$\begin{aligned} \langle p \rangle &= \frac{\int_0^{p_m} d\vec{p}\, p \left| G_0(\vec{p}, E + i\Gamma) \right|^2}{\int_0^{p_m} d\vec{p} \left| G_0(\vec{p}, E + i\Gamma) \right|^2} \\ &= \sqrt{m_t E} \left(1 + \frac{\Gamma}{E} \frac{\ln p_m^2 / \left(\sqrt{E^2 + \Gamma^2} m_t \right)}{\pi} \right) \end{aligned} \tag{3.63}$$

where a cutoff p_m has been introduced. As a consequence of the small numerical prefactor of the divergent term and its logarithmic cutoff dependence the result is fairly insensitive to the exact value of the cutoff for p_m of order m_t. Alternatively one may replace the phase space element $d\vec{p}/m_t$ by the relativistic version $d\vec{p}/E = d\vec{p}/\sqrt{m_t^2 + p^2}$ to obtain a convergent result. In future measurements the cutoff will presumably be provided by the experimental analysis.

In order to study the dependence of $\langle p \rangle$ on the strong coupling constant, consider for the moment the predictions for a stable quark. Some intuition and qualitative understanding can already be gained from the predictions based on a pure Coulomb potential [120].

For a stable top quark of fixed mass the "effective threshold" can be associated with the location of the $1S$ resonance $\sqrt{s_{thr}} = 2m_t + E_{1S}$ with $E_{1S} = -E_{Ryd} = -\alpha^2 m_t/4$ which

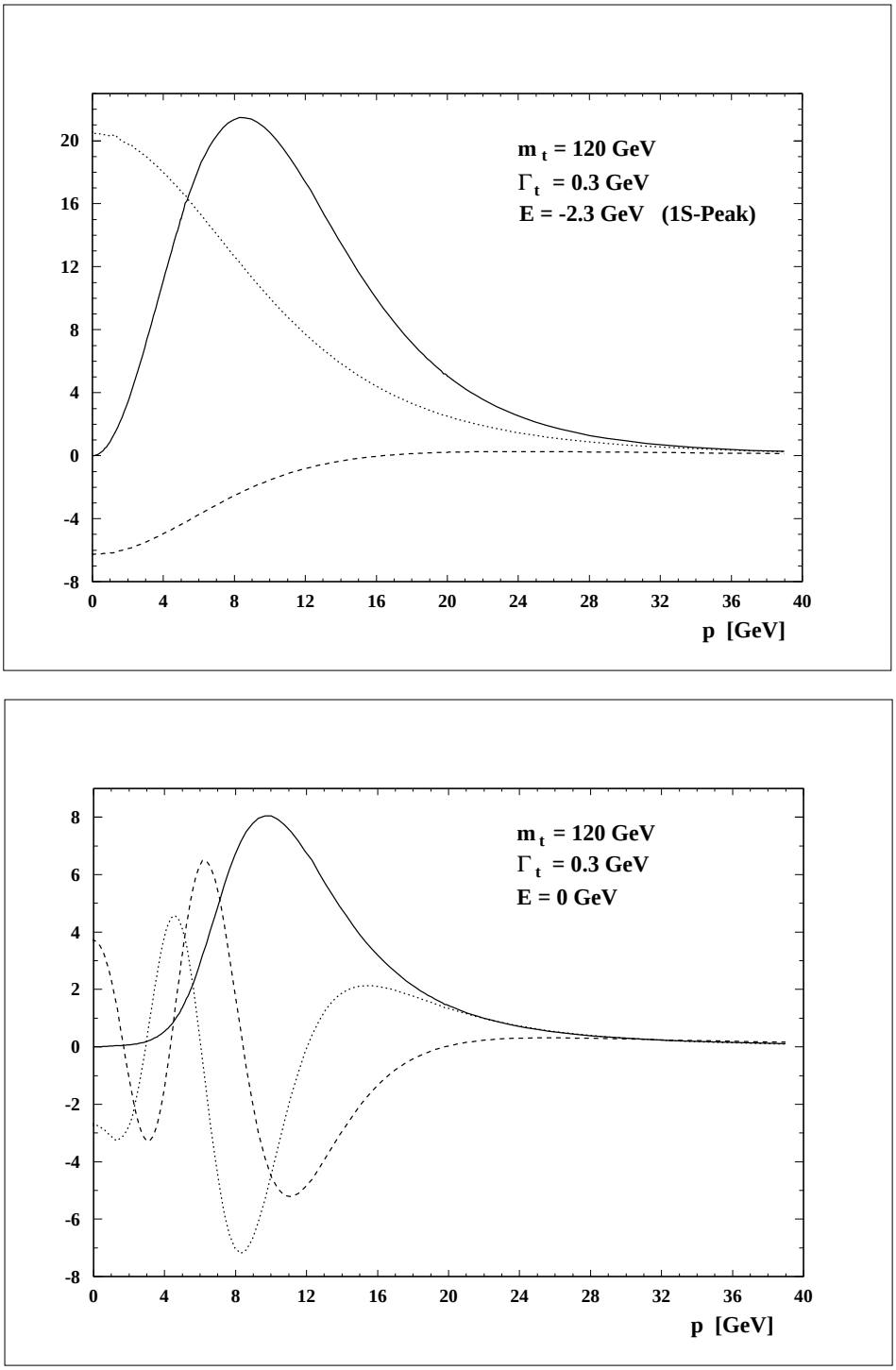


Figure 3.19: Real (dashed) and imaginary (dotted) parts of the Green's function for an energy corresponding to the 1S peak (upper figure) and for $E = 0$ (lower figure). Solid curve: $|pG(p)|^2 \cdot 0.002$.

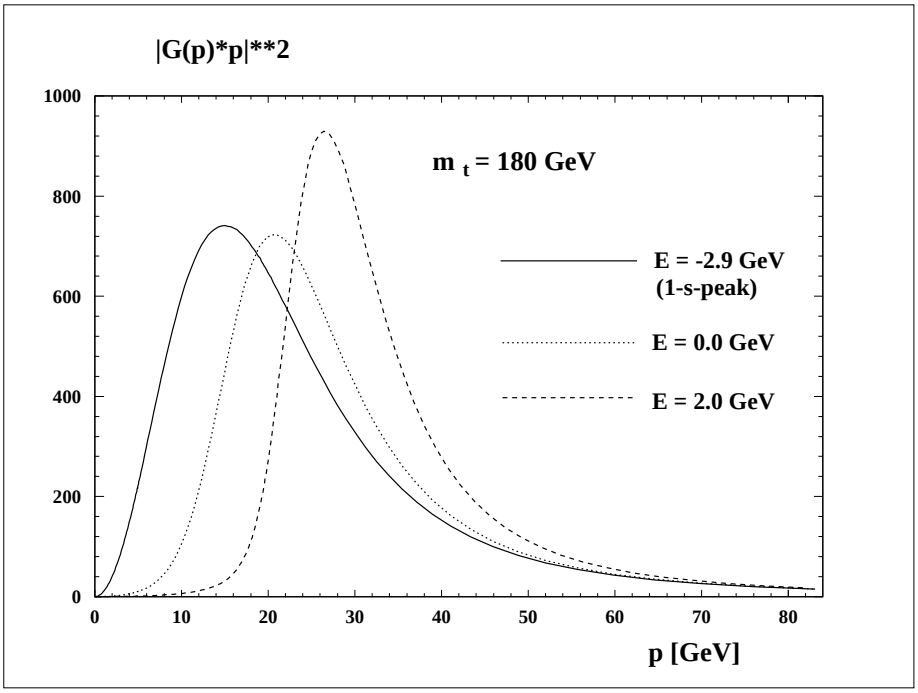


Figure 3.20: Momentum distribution of top quarks for three different cms energies.

decreases with increasing α. The height of the resonance cross section is proportional to the square of the wave function at the origin and hence proportional to α^3, as long as the resonances are reasonably well separated. In the limit of large Γ_t, i.e. far larger than E_{Ryd}, the overlapping $1S$, $2S$ ... resonances have to fill the gaps between the peaks. Since these gaps themselves increase proportional to α^2, one is left in the extreme case of large width with a cross section linear in α. Note that this corresponds to the behaviour of the cross section close to but slightly above the threshold which is also proportional to α.

For realistic top masses one thus observes a dependence of the peak cross section linear in α. Since the location of the peak itself depends on α, only the analysis of the full shape allows to extract the relevant information.

In a next step also the momentum distribution of top quarks has to be exploited to obtain further information. The discussion is again particularly simple for the Coulomb potential $V(r) = -\alpha/r$ and provides a nice exercise in nonrelativistic quantum mechanics. The average momentum, in units of the Bohr momentum $\alpha m_t/2$, can be written in terms of a function $f(\epsilon)$ which depends only on one variable $\epsilon = E/E_{Ryd}$ if the energy $E = \sqrt{s} - 2m_t$ is measured in terms of the Rydberg energy.

$$\langle p \rangle = \frac{\alpha m_t}{2} f(\epsilon) \tag{3.64}$$

For positive arguments the function f can be derived from obvious kinematical considerations.

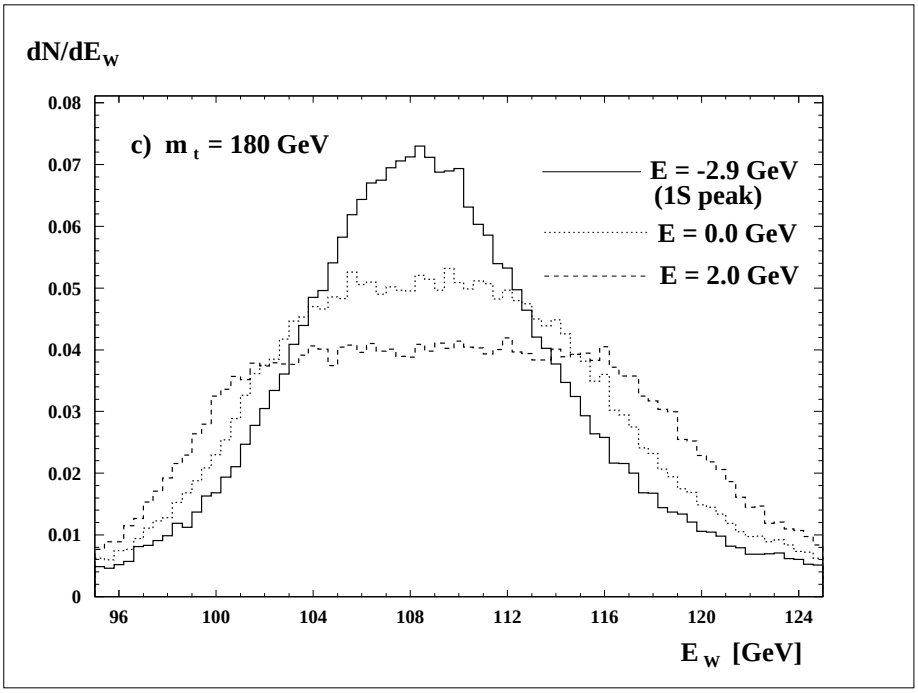


Figure 3.21: Energy distribution of W's from top quark decay for three different cms energies.

$$f(\epsilon) = \sqrt{\epsilon} \qquad \text{for} \qquad \epsilon \geq 0 \tag{3.65}$$

For the discrete negative arguments $\epsilon_n = -1/n^2$, corresponding to the locations of the bound states, the radial wave functions in momentum space are given in terms of the Gegenbauer polynomials C_n^m

$$\psi(\vec{p}) = \frac{16\pi n^{3/2}}{(1+n^2p^2)^2} C_{n-1}^1 \left(\frac{n^2p^2-1}{n^2p^2+1} \right) Y_0^0(\theta,\varphi) \tag{3.66}$$

with

$$\int \frac{d\vec{p}}{(2\pi)^3} |\psi(\vec{p})|^2 = 1 \ . \tag{3.67}$$

Using the explicit forms of C_n^m

$$C_0^1(z) = 1, \qquad C_1^1(z) = 2z, \qquad C_2^1(z) = 4z^2 - 1 \tag{3.68}$$

one obtains through straightforward calculation

$$f(-1) = \frac{8}{3\pi}, \qquad f(-1/4) = \frac{16}{15\pi}, \qquad f(-1/9) = \frac{24}{35\pi} \ . \tag{3.69}$$

For arbitrary n one derives the general result

$$f\left(-\frac{1}{n^2}\right) = \frac{8n}{(2n-1)(2n+1)\pi} \tag{3.70}$$

with the asymptotic behaviour

$$f\left(-\frac{1}{n^2}\right) \to \frac{2}{n\pi} \ . \tag{3.71}$$

This is in accord with the result expected from classical mechanics: For the average momentum of a particle on a closed orbit in the Coulomb potential one derives

$$\langle p^{2\beta} \rangle = \left(\frac{\alpha m_t}{2}\right)^{2\beta} \left(\frac{-E}{E_{Ryd}}\right)^{\beta} \frac{1}{2\pi} \int_0^{2\pi} d\xi \frac{(1-e^2\cos^2\xi)^{\beta}}{(1-e\cos\xi)^{2\beta-1}} \ . \tag{3.72}$$

Quantum mechanical orbits with angular momentum zero and high radial quantum numbers correspond to classical motions with excentricity $e = 1$ (i.e. straight lines). In this limiting case the classical expectation value is easily evaluated, and for $\beta = 1/2$ one finds agreement with the quantum mechanical result. For small negative energies one therefore obtains the behaviour $f(\epsilon) = 2\sqrt{-\epsilon}/\pi$. Significantly below threshold, however, the average momentum obtained from the Green's function increases more rapidly with decreasing energy and between the $1S$ and the $2S$ state one observes an approximately linear dependence on the energy.

From these considerations the dependence of the average momentum on α (with E fixed) is easily understandable, in particular the seemingly surprising observation that well below threshold $\langle p \rangle$ decreases with increasing α. From (3.64) one derives for a shift in α (keeping the energy E fixed) the following shift in $\langle p \rangle$

$$\frac{\delta\langle p \rangle}{\langle p \rangle} = \left(1 - 2\frac{f'(\epsilon)}{f(\epsilon)}\epsilon\right) \frac{\delta\alpha}{\alpha} \ . \tag{3.73}$$

Above threshold as well as close to but below threshold $f \propto \sqrt{|\epsilon|}$. Hence $\epsilon f'/f = 1/2$ and the average momentum remains unaffected. The location of the minimum is thus an ideal place to fix the mass of the top quark. Significantly below threshold, however, $\epsilon f'/f \approx 1$ and the factor in front of $\delta\alpha/\alpha$ becomes negative. This explains the decrease of $\langle p \rangle$ with increasing α.

These results are illustrated in Fig.3.22. In Fig.3.22b we demonstrate that $\langle p \rangle$ as evaluated with the program for the Green function (solid line) coincides perfectly well with the values calculated from the analytical formula on resonance, indicated by the triangles. The prediction from classical mechanics, namely $\langle p \rangle \propto \sqrt{|\epsilon|}$ is shown by the dotted line and agrees nicely for positive and negative energies. In Fig.3.22a α_s is increased from 0.20 to 0.24 and $\langle p \rangle$ changes in accord with the previous discussion.

For definiteness we have chosen $m = m_t/2 = 60$ GeV for the reduced mass and $\alpha = 4\alpha_s/3$ with α_s varying between 0.20 and 0.24. The curves demonstrate the decrease of $\langle p \rangle$ by about 10% for the corresponding increase in α. The triangles mark the locations of the resonances and the expectation values for the momentum as calculated from (3.70).

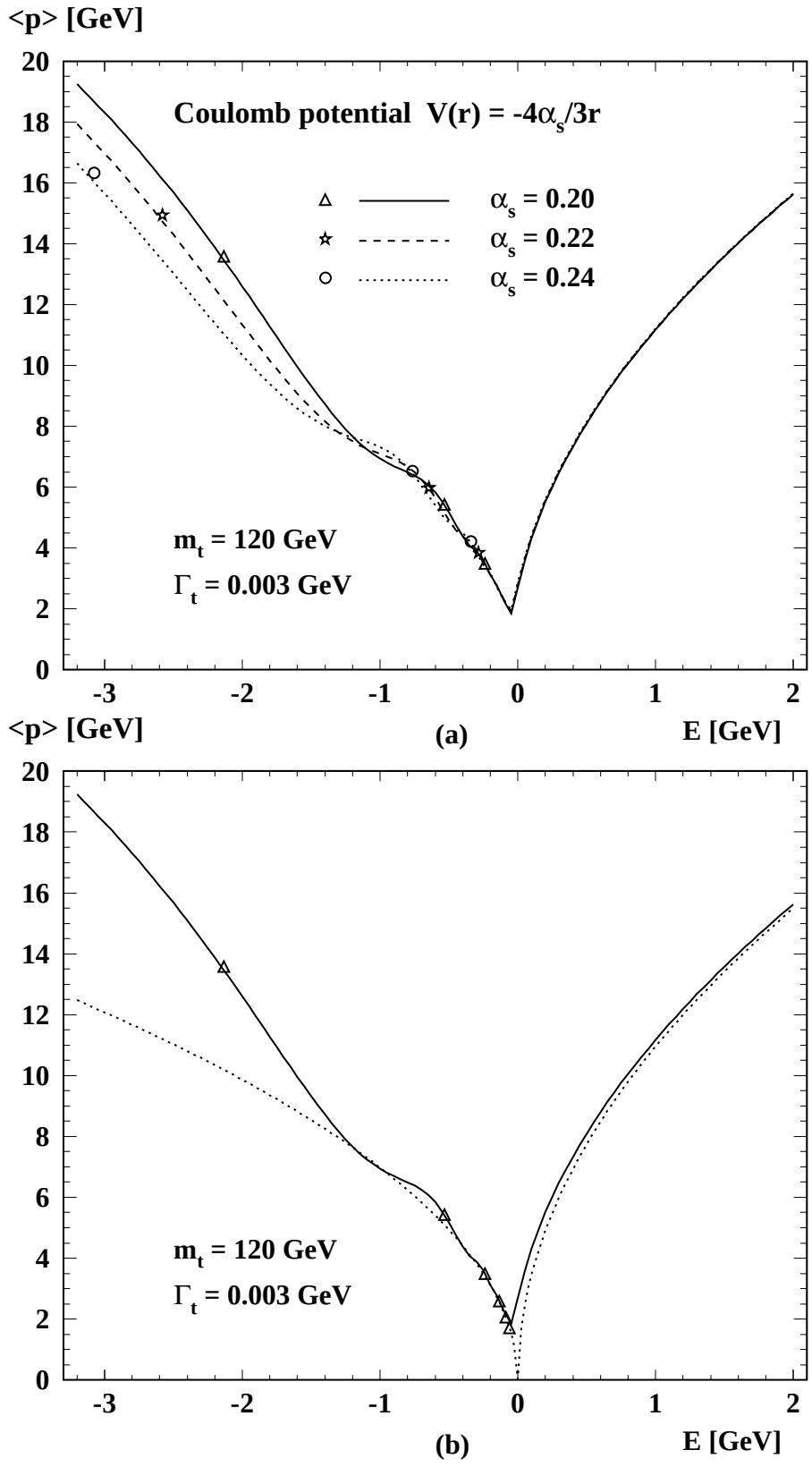


Figure 3.22: a) Average momentum as a function of E for different values of α_s. The markers show the results of the analytical calculation at $1S$, $2S$, $3S$ energies. b) Comparison with the analytical result for discrete energies and with the square–root dependence close to threshold.

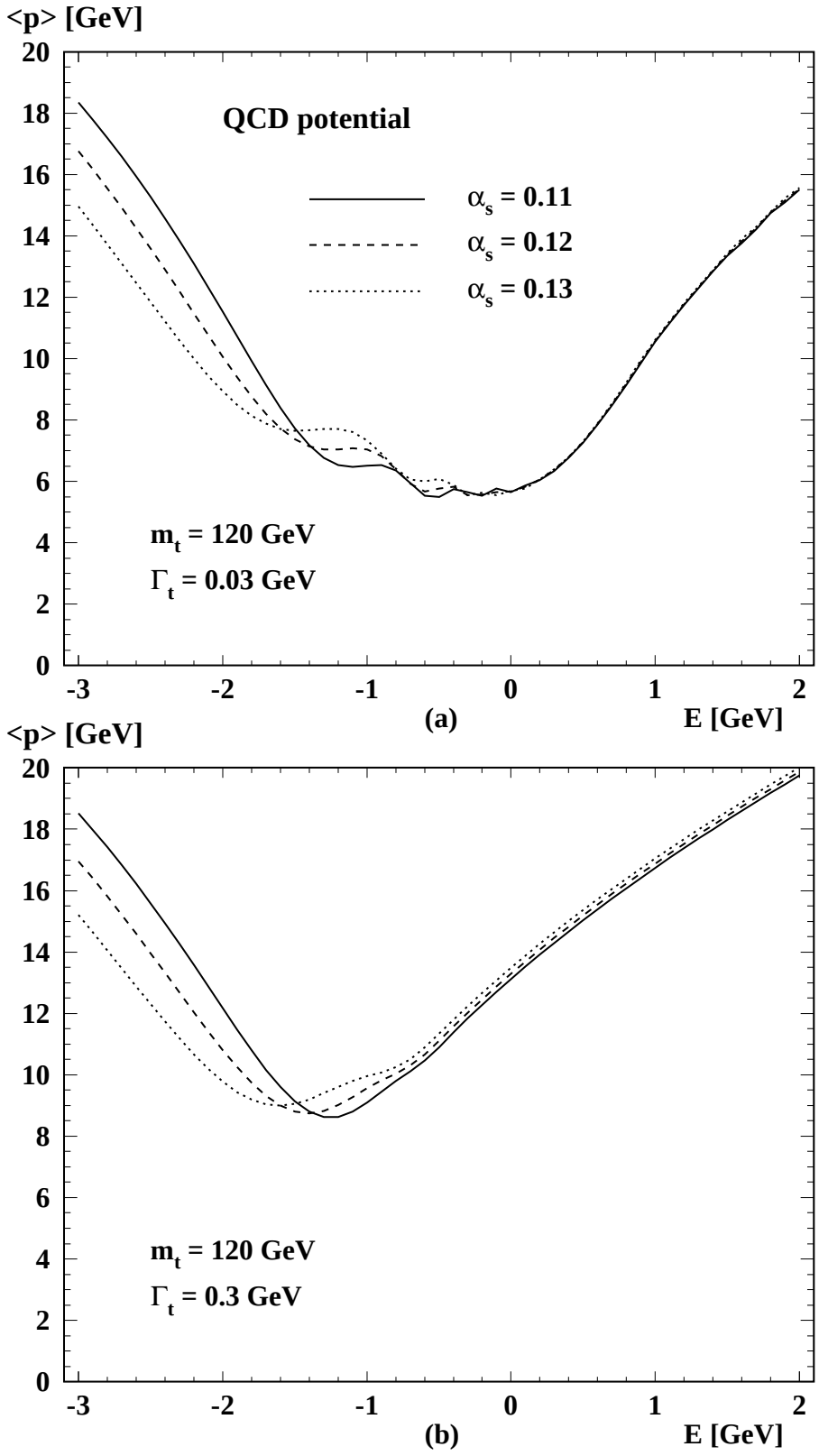


Figure 3.23: Energy dependence of $\langle p \rangle$, the average t quark momentum for $\alpha_s = 0.13$ (dotted) 0.12 (dashed) and 0.11 (solid) line for $m_t = 120$ GeV. a) $\Gamma_t = 0.03$ GeV and b) $\Gamma_t = 0.3$ GeV.

The qualitative behaviour remains unchanged for realistic QCD potentials corresponding to different values of $\alpha_s(M_Z)$. Qualitatively the same behaviour is observed as in Fig.3.22. In Fig.3.23a the top quark width has been set to an artificially small value of 0.03 GeV, in Fig.3.23b the realistic value of 0.3 GeV has been adopted. The finite width leads to an additional contribution to the momentum of order $\sqrt{\Gamma m}$.

An important feature is evident from Fig.3.23: The momentum calculated for positive energy is nearly independent from α_s and reflects merely the kinematic behaviour, just as in the case of the Coulomb potential. This is characteristic for the choice of a potential [114] where the large distance behaviour is fixed by phenomenology and decoupled from the short distance value of α_s.

The different assumptions on the long distance behaviour are reflected in differences between the predictions of [121, 119, 112] for the precise location of the $t\bar{t}$ threshold for identical values of α_s and m_t and in differences in the α_s dependence of the momentum distributions for fixed m_t and energy (see also [116]). All these differences can be attributed to the freedom in the additive constant discussed before. The same additive constant appears in $b\bar{b}$ spectroscopy, such that the mass difference between top and bottom is independent from these considerations.

In Fig. 3.24 the predictions for $\langle p \rangle$ vs. energy are presented for the case of a realistic QCD potential, assuming $m_t = 120/150/180$ GeV. The strong rise of $\langle p \rangle$ as a consequence of the strong increase of Γ_t is clearly visible.

3.2.5 Angular distributions and polarization

Close to threshold the production amplitude is dominantly S-wave which leads to an isotropic angular distribution. The spin of top quarks is alligned with the beam direction, with a degree of polarization determined by the electroweak couplings, the beam polarization and the mass of the top quark, but independent of the production dynamics, in particular of the potential.

Small, but nevertheless experimentally accessible corrections do arise from the small admixture of P-wave contributions and from rescattering of the top quark decay products. Let us concentrate for the moment on the first mechanism. P-wave amplitudes are proportional to the top quark momentum. For stable noninteracting particles the momentum vanishes at threshold. However, as discussed in the previous section the expectation value of the quark momentum is nonzero for all energies — a consequence of the large top decay rate and the uncertainty principle. Technically the P-wave contribution is calculated with the help of the Green's function technique. The generalization of the Lippman-Schwinger equation (3.56) from S- to P-waves reads as follows

$$\mathcal{F}(\mathrm{p}, E) \;=\; G_0(p, E) + G_0(p, E) \int \frac{d^3k}{(2\pi)^3} \frac{\mathbf{p}\cdot\mathbf{k}}{\mathbf{p}^2} V(\mathbf{p}-\mathbf{k}) \mathcal{F}(k, E) \qquad (3.74)$$

It is then straightforward to calculate the differential momentum distribution and the polarization of top quarks produced in electron positron annihilation. Let us recall the

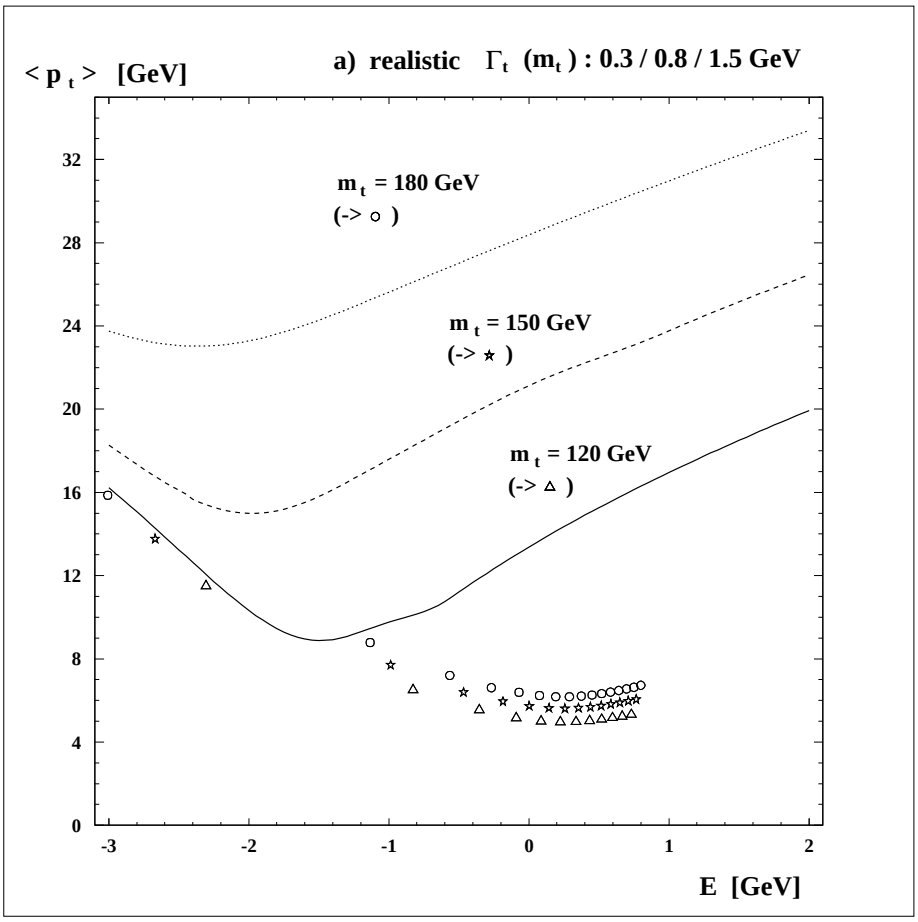


Figure 3.24: Energy dependence of the average top quark momentum for m_t = 120/150/180 GeV. Triangles, stars and circles correspond to $\langle p_t \rangle$ for S-states with $\Gamma_t = 0$.

following conventions for the fermion couplings

$$v_f = 2I_f^3 - 4q_f \sin^2\theta_{\rm W}, \qquad a_f = 2I_f^3. \tag{3.75}$$

$P_\pm$ denotes the longitudinal electron/positron polarization and $\chi = (P_+ - P_-)/(1 - P_+P_-)$ can be interpreted as effective longitudinal polarization of the virtual intermediate photon or Z boson. The following abbreviations will be useful below:

$$\begin{aligned} a_1 &= q_e^2 q_t^2 + (v_e^2 + a_e^2) v_t^2 d^2 + 2 q_e q_t v_e v_t d \\ a_2 &= 2 v_e a_e v_t^2 d^2 + 2 q_e q_t a_e v_t d \\ a_3 &= 4 v_e a_e v_t a_t d^2 + 2 q_e q_t a_e a_t d \\ a_4 &= 2(v_e^2 + a_e^2) v_t a_t d^2 + 2 q_e q_t v_e a_t d \\ d &= \frac{1}{16 \sin^2\theta_{\rm W} \cos^2\theta_{\rm W}} \frac{s}{s - M_Z^2}. \end{aligned} \tag{3.76}$$

The differential cross section, summed over polarizations of quarks and including S-wave and S-P–interference contributions, is thus given by

$$\begin{aligned} \frac{d^3\sigma}{dp^3} &= \frac{3\alpha^2\Gamma_t}{4\pi m_t^4}(1 - P_+P_-)\left[(a_1 + \chi a_2)\left(1 - \frac{16\alpha_{\rm s}}{3\pi}\right)|G(\mathrm{p}, E)|^2 + \right. \\ &\left. + (a_3 + \chi a_4)\left(1 - \frac{12\alpha_{\rm s}}{3\pi}\right)\frac{\mathrm{p}}{m_t}\mathrm{Re}\left(\, G(\mathrm{p}, E) F^*(\mathrm{p}, E)\,\right)\cos\vartheta\right]. \end{aligned} \tag{3.77}$$

The vertex corrections from hard gluon exchange for S-wave and P-wave amplitudes are included in this formula. It leads to the following forward-backward asymmetry [122, 123]

$$\mathcal{A}_{\rm FB}(\mathrm{p}, E) = C_{\rm FB}(\chi)\,\varphi_{\rm R}(\mathrm{p}, E), \tag{3.78}$$

with

$$C_{\rm FB}(\chi) = \frac{1}{2}\frac{a_3 + \chi a_4}{a_1 + \chi a_2}, \tag{3.79}$$

$\varphi_{\rm R} = \mathrm{Re}\varphi$, and

$$\varphi(\mathrm{p}, E) = \frac{(1 - 4\alpha_{\rm s}/3\pi)}{(1 - 8\alpha_{\rm s}/3\pi)}\frac{\mathrm{p}}{m_t}\frac{F^*(\mathrm{p}, E)}{G^*(\mathrm{p}, E)}. \tag{3.80}$$

This result is still differential in the top quark momentum. Replacing $\varphi(\mathrm{p}, E)$ by

$$\Phi(E) = \frac{(1 - 4\alpha_{\rm s}/3\pi)}{(1 - 8\alpha_{\rm s}/3\pi)}\frac{\int_0^{\mathrm{p}_m} d\mathrm{p}\,\frac{\mathrm{p}^3}{m_t} F^*(\mathrm{p}, E) G(\mathrm{p}, E)}{\int_0^{\mathrm{p}_m} d\mathrm{p}\,\mathrm{p}^2\,|G(\mathrm{p}, E)|^2}. \tag{3.81}$$

one obtains the integrated forward-backward asymmetry again. Again, the cutoff p_m must be introduced to eliminate the logarithmic divergence of the integral.

Polarization

To describe top quark polarization in the threshold region it is convenient to align the reference system with the beam direction (Fig. 3.25) and to define

$$\mathbf{s}_\parallel = \mathbf{n}_{e^-}, \quad \mathbf{s}_{\rm N} = \frac{\mathbf{n}_{e^-} \times \mathbf{n}_t}{|\mathbf{n}_{e^-} \times \mathbf{n}_t|}, \quad \mathbf{s}_\perp = \mathbf{s}_{\rm N} \times \mathbf{s}_\parallel. \tag{3.82}$$

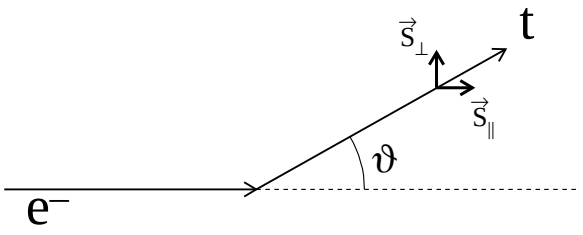


Figure 3.25: Definition of the spin directions. The normal component $\mathbf{s}_{\rm N}$ points out of the plane.

In the limit of small β the quark spin is essentially aligned with the beam direction apart from small corrections proportional to β, which depend on the production angle. A system of reference with $\mathbf{s}_{||}$ defined with respect to the top quark momentum [124] is convenient in the high energy limit but evidently becomes less convenient close to threshold.

Including the QCD potential one obtains for the three components of the polarization

$$\begin{aligned} \mathcal{P}_{||}(\mathbf{p},E,\chi) &= C^0_{||}(\chi) + C^1_{||}(\chi)\,\varphi_{\rm R}({\rm p},E)\,\cos\vartheta && (3.83)\\ \mathcal{P}_{\perp}(\mathbf{p},E,\chi) &= C_{\perp}(\chi)\,\varphi_{\rm R}({\rm p},E)\,\sin\vartheta && (3.84)\\ \mathcal{P}_{\rm N}(\mathbf{p},E,\chi) &= C_{\rm N}(\chi)\varphi_{\rm I}({\rm p},E)\sin\vartheta\,, && (3.85) \end{aligned}$$

$$\begin{aligned} C^0_{||}(\chi) &= -\frac{a_2+\chi a_1}{a_1+\chi a_2}, & C^1_{||}(\chi) &= \left(1-\chi^2\right)\frac{a_2a_3-a_1a_4}{(a_1+\chi a_2)^2},\\ C_{\perp}(\chi) &= -\frac{1}{2}\,\frac{a_4+\chi a_3}{a_1+\chi a_2}, & C_{\rm N}(\chi) &= -\frac{1}{2}\,\frac{a_3+\chi a_4}{a_1+\chi a_2} \;=\; -C_{\rm FB}(\chi), \end{aligned} \qquad (3.86)$$

with $\varphi_{\rm I} = {\rm Im}\,\varphi$, and $\varphi({\rm p},E)$ as defined in (3.80). The momentum integrated quantities are obtained by the replacement $\varphi({\rm p},E) \to \Phi(E)$. The case of non-interacting stable quarks is recovered by the replacement $\Phi \to \beta$, an obvious consequence of (3.81).

Let us emphasize the main qualitative features of the result:

- Top quarks in the threshold region are highly polarized. Even for unpolarized beams the longitudinal polarization amounts to about -0.41 and reaches ± 1 for fully polarized electron beams. This later feature is of purely kinematical origin and independent of the structure of top quark couplings. Precision studies of polarized top decays are therefore feasible.

- Corrections to this idealized picture arise from the small admixture of P-waves. The transverse and the normal components of the polarization are of order 10%. The angular dependent part of the parallel polarization is even more suppressed. Moreover, as a consequence of the angular dependence its contribution vanishes upon angular integration.

- The QCD dynamics is solely contained in the functions φ or Φ which is the same for the angular distribution and the various components of the polarization. (However,

this "universality" is affected by the rescattering corrections.) These functions which evidently depend on QCD dynamics can thus be studied in a variety of ways.

- The relative importance of P-waves increases with energy, $\Phi \sim \sqrt{E/m_t}$. This is expected from the close analogy between $\Phi_{\rm R} = {\rm Re}\,\Phi$ and β. In fact, the order of magnitude of the various components of the polarization above, but close to threshold, can be estimated by replacing $\Phi_{\rm R} \to {\rm p}/m_t$.

The C_i are displayed in Fig. 3.26 as functions of the variable χ ($\sin^2\theta_{\rm W} = 0.2317$, $m_t = 180$ GeV). As discussed before, $C_{\|}^0$ assumes its maximal value ± 1 for $\chi = \mp 1$ and the coefficient $C_{\|}^1$ is small throughout. The coefficient $C_\perp$ varies between $+0.7$ and -0.5 whereas $C_{\rm N}$ is typically around -0.5. The dynamical factors Φ are around 0.1 or larger, such that the P-wave induced effects should be observable experimentally.

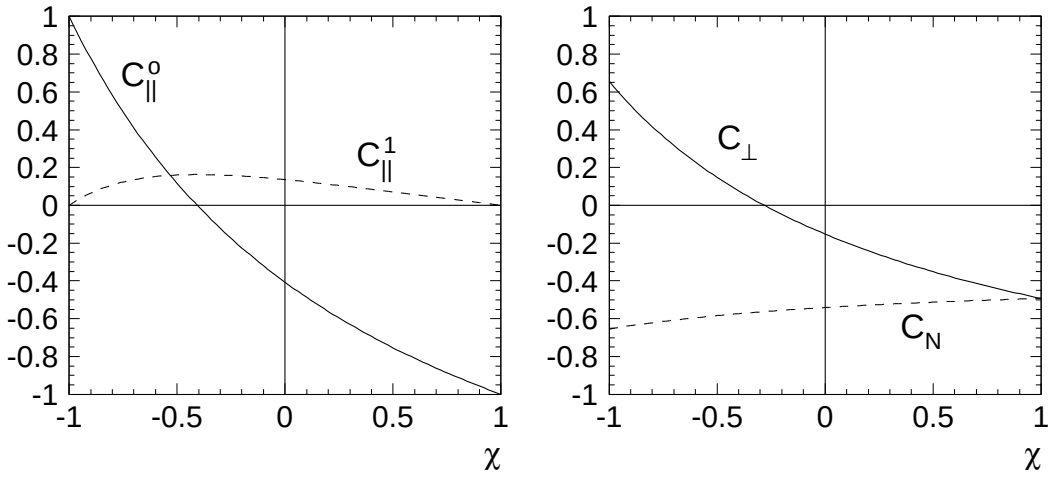


Figure 3.26: The coefficients (3.86) for $\sqrt{s}/2 = m_t = 180$ GeV.

The normal component of the polarization which is proportional to $\varphi_{\rm I}$ has been predicted for stable quarks in the framework of perturbative QCD [125, 124]. In the threshold region the phase can be traced to the $t\bar{t}$ rescattering by the QCD potential. For stable quarks, assuming a pure Coulomb potential $V = -4\alpha_{\rm s}/3r$, the nonrelativistic problem can be solved analytically [126] and one finds

$$\lim_{\substack{\Gamma_t \to 0 \\ E \to {\rm p}^2/m_t}} \left(E - \frac{\mathbf{p}^2}{m_t} + i\Gamma_t \right) G({\rm p}, E) \;=\; \exp\left(\frac{\pi \bar{k}}{2{\rm p}}\right) \Gamma(1 + i\bar{k}/{\rm p}) \tag{3.87}$$

$$\lim_{\substack{\Gamma_t \to 0 \\ E \to {\rm p}^2/m_t}} \left(E - \frac{\mathbf{p}^2}{m_t} + i\Gamma_t \right) F({\rm p}, E) \;=\; \left(1 - i\frac{\bar{k}}{\rm p}\right) \exp\left(\frac{\pi \bar{k}}{2{\rm p}}\right) \Gamma(1 + i\bar{k}/{\rm p}), \tag{3.88}$$

with $\bar{k} = 2m_t\alpha_s/3$ and hence

$$\varphi_{\rm I}({\rm p}, E) \;\to\; \frac{2}{3}\alpha_{\rm s}\frac{1 - 4\alpha_{\rm s}/3\pi}{1 - 8\alpha_{\rm s}/3\pi} \tag{3.89}$$

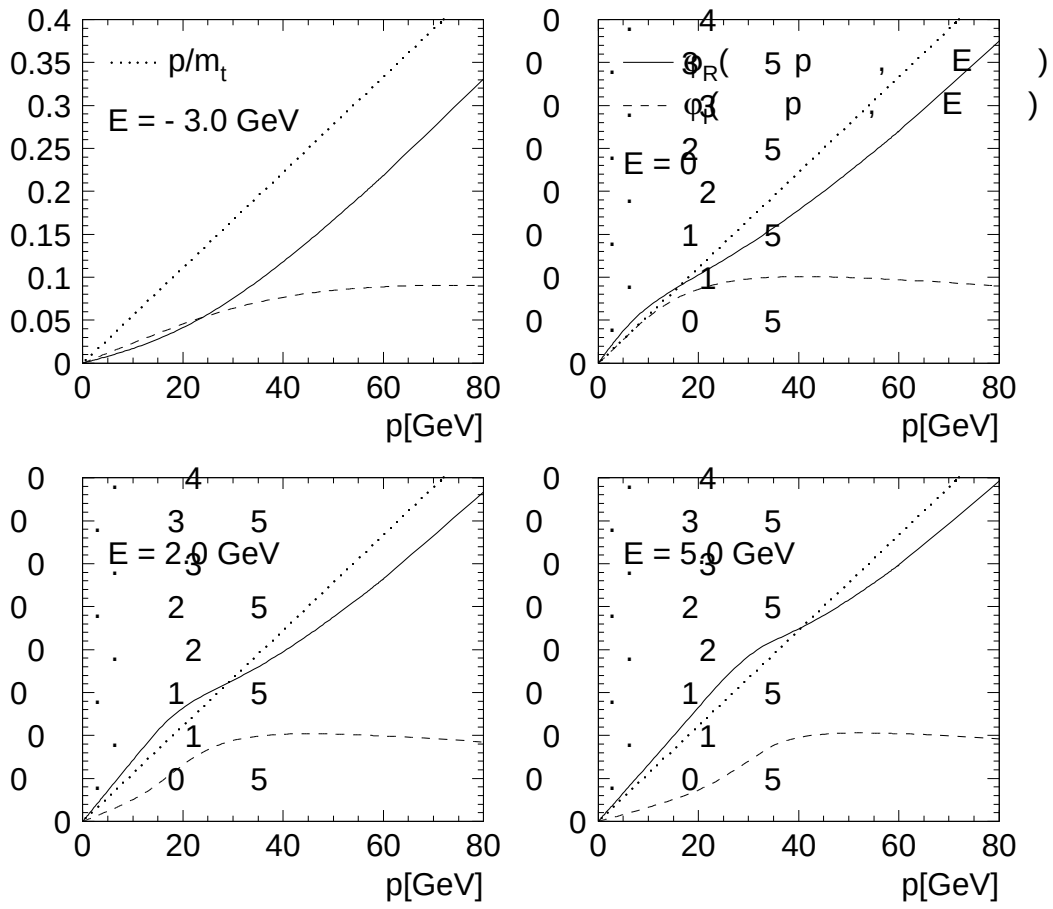


Figure 3.27: Real (solid) and imaginary (dashed) part of the function $\varphi(\mathrm{p}, E)$ for $m_t = 180$ GeV, $\alpha_\mathrm{s} = 0.125$ and four different energies. The dotted line shows the free particle case $\mathrm{Re}\,\varphi = \mathrm{p}/m_t$ (from [127]).

$$\Phi_\mathrm{I}(E) \;\rightarrow\; \frac{2}{3}\alpha_\mathrm{s}\frac{1-4\alpha_\mathrm{s}/3\pi}{1-8\alpha_\mathrm{s}/3\pi}. \tag{3.90}$$

The component of the polarization normal to the production plane is thus approximately independent of E and essentially measures the strong coupling constant. In fact one can argue that this is a unique way to get a handle on the scattering of heavy quarks through the QCD potential.

Predictions for real and imaginary parts of the function φ are displayed in Fig. 3.27 for four different energies.

The momentum integrated functions $\Phi(E)$ are shown in Fig. 3.28. From this figure, in combination with Fig. 3.26, it is clear that the contribution of P-wave amplitudes to the quark polarization will amount to 10% at most and by construction vanishes upon angular integration.

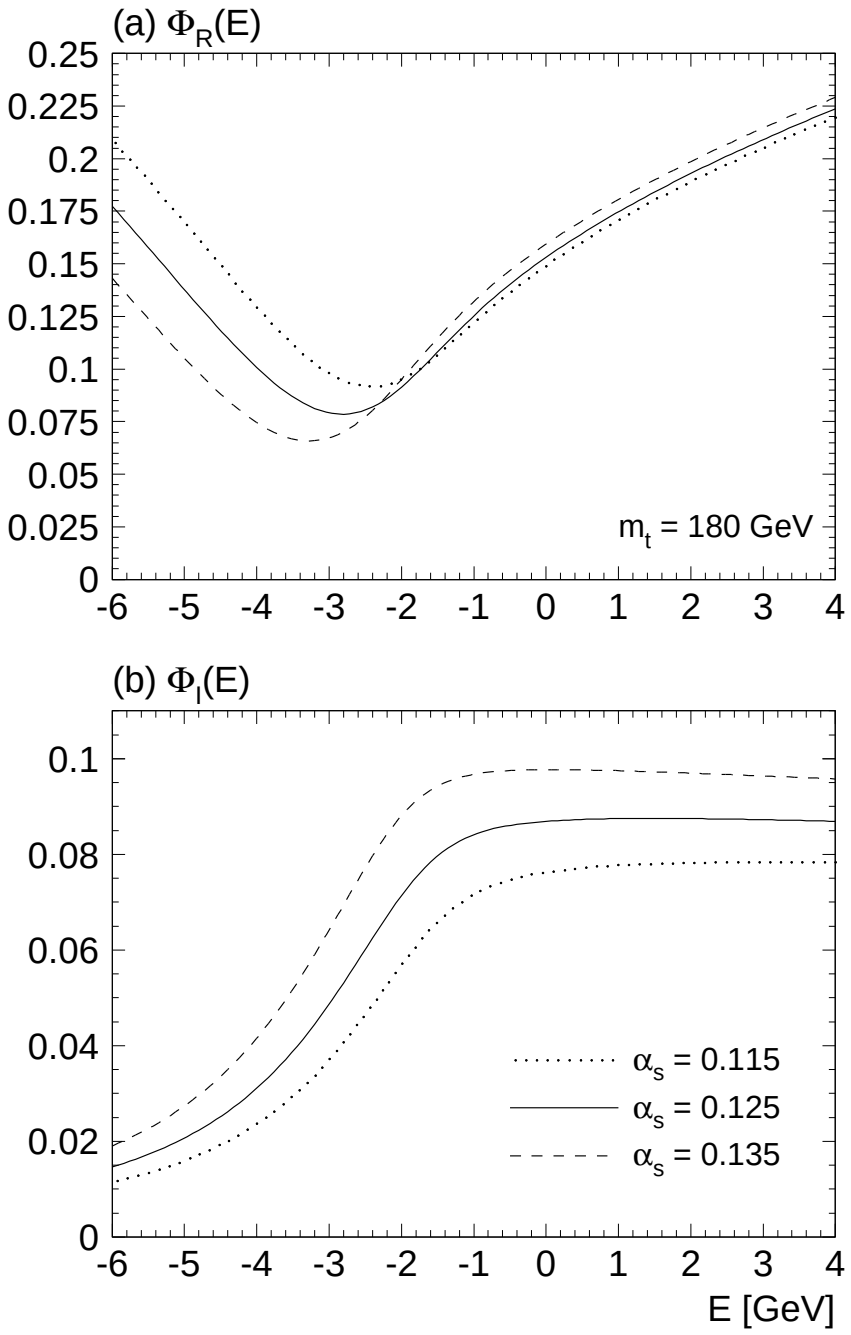


Figure 3.28: Real and imaginary part of $\Phi(E)$ for three different values of α_s (from [127]).

3.2.6 Rescattering

For a particle with a very small decay rate production and decay amplitudes can be clearly separated. This is fairly evident from the space-time picture of such a sequence. Prior to its decay the particle travels away from the production point and any coherence is lost between the two reactions. The situation is different for the case under discussion, an unstable top quark which decays within the range of interaction between t and $\bar{t}$. In such a situation the decay products from t are still affected by the force originating from $\bar{t}$ and vice versa (Fig. 3.29).

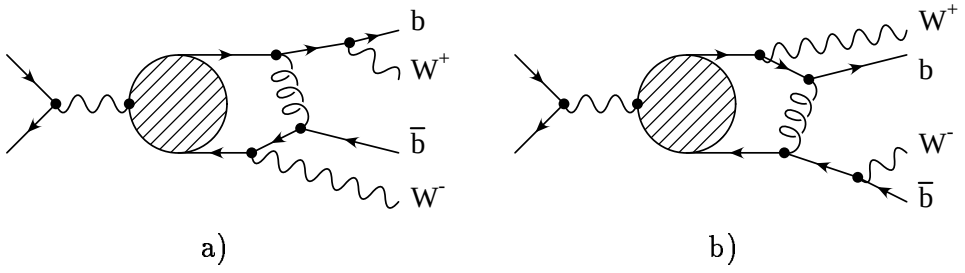


Figure 3.29: Lowest order rescattering diagrams.

In ref. [128, 129] it has been demonstrated that the total cross section remains unaffected by rescattering in order α_s. This result had been anticipated in [112] on the basis of earlier work which considered the decay rate of a muon bound in the strong field of a nucleus [130]. In contrast momentum and angular distributions [122, 129, 131] as well as the top quark polarization [127] are affected by rescattering. For example the momentum distribution has to be corrected by a factor $(1 + \psi_1(p, E))$ with

$$\psi_1(\mathrm{p}, E) = 2\,\mathrm{Im} \int \frac{d^3k}{(2\pi)^3} V(|\mathbf{k} - \mathbf{p}|) \frac{G(\mathrm{k}, E)}{G(\mathrm{p}, E)} \frac{\arctan \frac{|\mathbf{k}-\mathbf{p}|}{\Gamma_t}}{|\mathbf{k} - \mathbf{p}|} \tag{3.91}$$

$$\tag{3.92}$$

The distribution is shifted towards smaller momenta by about 5% (Fig. 3.30), an effect that could become relevant in precision experiments. The influence on the forward-backward asymmetry and the polarization is even more pronounced [127], as far as the S-P-wave interference terms are concerned which are thus intrinsically of order β. A detailed discussion of these effects is beyond the scope of these lectures and can be found in [127].

3.2.7 Relativistic corrections

In $\mathcal{O}(\alpha_s^2)$ one anticipates effects from relativistic corrections, from the reduction of the phase space through the binding energy and from the Coulomb wave function of the b quark. Individually these effects are large. For the sake of the argument, let us adopt

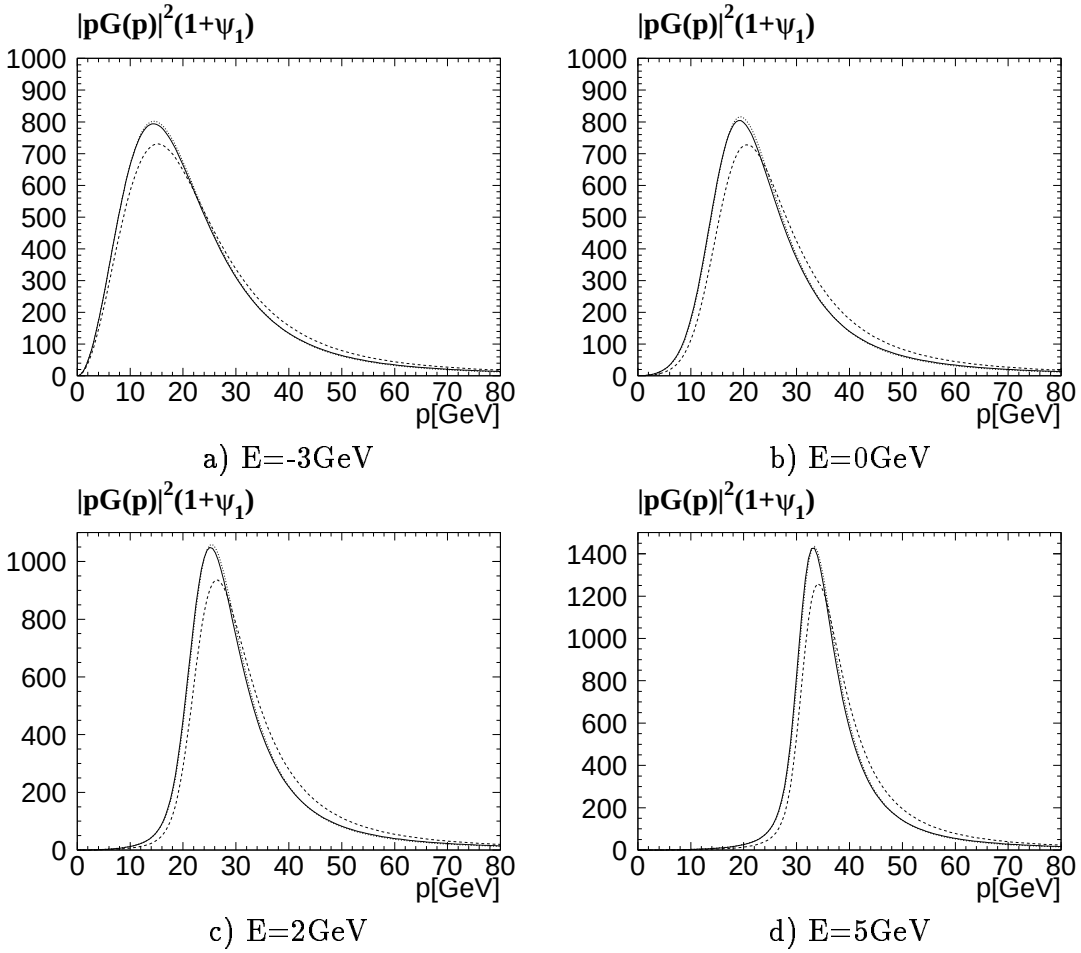


Figure 3.30: Modification of the momentum distribution through rescattering. Dashed line: no rescattering corrections included; Solid line: rescattering contribution with full potential included; dotted line: rescattering contribution with pure Coulomb potential and $\alpha_s = 0.187$ included (from [127]).

a pure Coulomb potential and a binding energy of -2.5 GeV. From the virial theorem one derives a potential energy of -5 GeV. The phase space of the quark decaying first is therefore reduced by this same amount. Assuming $m_t = 180$ GeV one would arrive at a reduction of Γ_t by about 10%. A full calculation of all $\mathcal{O}(\alpha_s^2)$ effects is not available at present and one has to resort to models and analogies [112, 119, 114]. For example, it has been shown [130, 132] that the decay rate of a muon bound in the field of a nucleus is given by

$$\Gamma = \Gamma_{\text{free}} \left[1 - 5(Z\alpha)^2\right] \left[1 + 5(Z\alpha)^2\right] \left[1 - \frac{(Z\alpha)^2}{2}\right], \tag{3.93}$$

where the first correction factor originates from the phase space suppression, the second from the Coulomb enhancement, and the third from time dilatation. Thus there is no first order correction to the total rate from rescattering in the nucleus potential, similar to the $t\bar{t}$ case discussed above. The second order contributions evidently compensate to a large extent. In a model calculation where these features are implemented [114] through a momentum dependent width, it is found that the total cross section as well as the momentum distribution are hardly affected. These considerations have recently been confirmed in a more formal approach [133].

Acknowledgments

I would like to thank A. Czarnecki, R. Harlander, A. Hoang, M. Jeżabek, E. Mirkes, M. Peter, T. Teubner, and P. Zerwas for pleasant collaborations and many helpful discussions and comments on the TOPics presented in this review. I am particularly grateful to A. Czarnecki. Without his TEXnical help this paper would never have been completed. The hospitality of the SLAC theory group during my sabbatical in the summer term 1995 is gratefully acknowledged.

Bibliography

[1] P. Sinerva, these proceedings;
CDF Collaboration (F. Abe et al.), *Phys. Rev. Lett.* **74** (1995) 2626;
D0 Collaboration (S. Abachi et al.), *Phys. Rev. Lett.* **74** (1995) 2632.

[2] D. Bardin et al., in Reports of the Working Group on Precision Calculations for the Z-Resonance, CERN 95-03, pg. 7.

[3] W. Hollik, talk at Int. Europhysics Conf. on HEP, Brussels, 1995, KA-TP-15-1995, hep-ph/9512232.

[4] M. Swartz , these proceedings, and references therein.

[5] For a recent compilation of theoretical results see B. Kniehl, in Reports of the Working Group on Precision Calculations for the Z-Resonance, CERN 95-03, pg. 299.

[6] A. Sirlin, ibid. pg. 299.

[7] G. Degrassi et al., ibid. pg. 163.

[8] W.A. Bardeen, C.T. Hill, M. Lindner, *Phys. Rev.* **D 41** (1990) 1647.

[9] G.L. Kane and M.E. Peskin, *Nucl. Phys.* **B 195** (1982) 29.

[10] C. Albajar et al. [UA1 Coll.], Phys. Lett. **B256** (1991) 121, E: ibid. **B262** (1991) 497.

[11] Particle Data Group, Phys. Rev. **D50**, 1171 (1994).

[12] A. Djouadi, J.H. Kühn and P.M. Zerwas, *Z. Phys.* **C 46** (1990) 411.

[13] Dong Su, private communication.

[14] TOPAZ Coll., E. Nakano et al., Phys. Lett. **B340** (1994) 135.
AMY Coll., F. Liu et al., Phys. Rev. **D49** (1994) 4339.
For a recent review see e.g. K. Miyabayashi, in *Recent electroweak results from TRISTAN* in XXXth Rencontres de Moriond.

[15] D. Schaile and P.M. Zerwas, Phys. Rev. **D45** (1992) 3262.

[16] M. Chanowitz, M. Furman and I. Hinchliffe, *Nucl. Phys.* **B 153** (1979) 402.

[17] N. Cabibbo, L. Maiani, G. Parisi and R. Petronzio, *Nucl. Phys.* **B 158** (1979) 296.

[18] M. Lindner, *Z. Phys.* **C 31** (1986) 295.

[19] G. Münster, Nucl. Phys. B (Proc. Suppl.) **34** (1994) 596; ibid., **42** (1995) 162.

[20] L. Hall, these proceedings.

[21] H. Albrecht et al. (ARGUS), *Phys. Lett.* **B 192** (1987) 245. M. Artuso et al. (CLEO), *Phys. Rev. Lett.* **62** (1989) 2233.

[22] M. Veltman, *Act. Phys. Pol.* **B 8** (1977) 475.

[23] M. Veltman, *Nucl. Phys.* **B 123** (1977) 89.

[24] L. Avdeev et al., *Phys. Lett.* **B 336** (1994) 560; Erratum ibid. **B349** (1995) 597.

[25] K.G. Chetyrkin, J.H. Kühn, and M. Steinhauser, *Phys. Lett.* **B 351** (1995) 331, *Phys. Rev. Lett.* **75** (1995) 3394.

[26] J.H. Kühn, *Act. Phys. Pol.* **B 12** (1981) 347.
J.H. Kühn, *Act. Phys. Austr.* **Suppl.XXIV** (1982) 203.

[27] M. Jeżabek and J. H. Kühn, *Phys. Lett.* **B 207** (1988) 91.

[28] B. H. Smith and M. B. Voloshin, *Phys. Lett.* **B 340** (1994) 176.

[29] A. Czarnecki, *Act. Phys. Pol.* **B 26** (1995) 845.

[30] S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, *Phys. Rev.* **D 28** (1983) 228.

[31] A. Denner and T. Sack, *Nucl. Phys.* **B 358** (1991) 46.

[32] R. Migneron G. Eilam, R.R. Mendel and A. Soni, *Phys. Rev. Lett.* **66** (1991) 3105.

[33] A. Denner and A.H. Hoang, Nucl. Phys. **B397** (1993) 483.

[34] M. Jeżabek and J. H. Kühn, *Phys. Rev.* **D 48** (1993) R1910; E: ibid. **D49** (1994) 4970.

[35] K. Fujii, these proceedings.

[36] J.H. Kühn and K.H. Streng, *Nucl. Phys.* **B 198** (1982) 71.

[37] M. Jeżabek and J. H. Kühn, *Nucl. Phys.* **B 320** (1989) 20.

[38] A. Czarnecki, M. Jeżabek, and J. H. Kühn, *Act. Phys. Pol.* **B 20** (1989) 961, E: ibid, **B23** (1992) 173.

[39] A. Czarnecki, M. Jeżabek, and J. H. Kühn, *Nucl. Phys.* **B 351** (1991) 70.

[40] A. Czarnecki, M. Jeżabek, J. G. Körner, and J. H. Kühn, *Phys. Rev. Lett.* **73** (1994) 384.

[41] A. Czarnecki and M. Jeżabek, *Nucl. Phys.* **B 427** (1994) 3.

[42] G.L. Kane, Proceedings, Madison Workshop 1979;
I. Bigi et al. in Ref.[43];
J.F. Gunion, H.E. Haber, G. Kane and S. Dawson, *The Higgs Hunter's Guide*, Addison–Wesley 1990;
C.S. Li and T.C. Yuan, *Phys. Rev. Lett.* **42** (1990) 3088.

[43] I. Bigi, Y.L. Dokshitzer, V.A. Khoze, J. Kühn and P. Zerwas, *Phys. Lett.* **B 181** (1986) 157.

[44] E. Reya, P.M. Zerwas [conv.] et al., in Proceedings ECFA LHC Workshop, Aachen 1990, CERN 90–10.

[45] P. Igo-Kemenes, J.H. Kühn [conv.], Proceedings, e^+e^- *Collisions at 500 GeV: The Physics Potential*, DESY 92-123, pg. 255-326, pg. 327-378.

[46] A. Czarnecki and S. Davidson, Phys. Rev. **D48** (1993) 4183; ibid., **D47** (1993) 3063.
A. Czarnecki and A.I. Davydychev, *Phys. Lett.* **B 325** (1994) 435.

[47] A. Czarnecki, *Phys. Rev.* **D 48** (1993) 5250.

[48] J. Gunion et al., Ref.[42];
A. Mendez and A. Pomarol, *Phys. Lett.* **B 252** (1990) 461;
C.S. Li and R.J. Oakes, *Phys. Rev.* **D 43** (1991) 855.

[49] K. Hikasa and M. Kobayashi, *Phys. Rev.* **D 36** (1987) 724;
H. Baer, M. Drees, R. Godbole, J.F. Gunion and X. Tata, Tallahassee Report FSU–HEP–910308 (1991).

[50] K. Hidaka, Y. Kizukuri and T. Kon, Gakugei Report TGU–12 (1991).

[51] J. Ellis and S. Rudaz, *Phys. Lett.* **B 128** (1983) 248;
A. Bouquet, J. Kaplan and C. Savoy, *Nucl. Phys.* **B 262** (1985) 299.

[52] G. Eilam, J.L. Hewett and A. Soni, *Phys. Rev.* **D 44** (1991) 1473;

[53] H. Fritzsch, *Phys. Lett.* **B 224** (1989) 423.

[54] W. Buchmüller and M. Gronau, *Phys. Lett.* **B 220** (1989) 641.

[55] M. Glück, J.F. Owens and E. Reya, *Phys. Rev.* **D 15** (1978) 2324;
B.L. Combridge, *Nucl. Phys.* **B 151** (1979) 429.

[56] S. Willenbrock and D.A. Dicus, *Phys. Rev.* **D 34** (1986) 155;
C.–P. Yuan, *Phys. Rev.* **D 41** (1990) 42.

[57] A.D. Martin, W.J. Stirling, and R.G. Roberts, hep-ph/9506423.

[58] W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler and J. Smith, Nucl. Phys. **B351** (1991) 507.

[59] P. Nason, S. Dawson and R. K. Ellis, *Nucl. Phys.* **B 303** (1988) 607; *ibid.* **B**327 (1989) 49; Erratum B335 (1990) 260.
G. Altarelli, M. Diemoz, G. Martinelli, and P. Nason, Nucl. Phys. **B308** (1988) 724.

[60] W. Beenakker, H. Kuijf, W. L. van Neerven and J. Smith, *Phys. Rev.* **D 40** (1989) 54.

[61] R. Meng, G. A. Schuler, J. Smith and W. L. van Neerven, *Nucl. Phys.* **B 339** (1990) 325.

[62] M. Diemoz, F. Ferrari, E. Longo and G. Martinelli, *Z. Phys.* **C 39** (1988) 21;
M. Glück, E. Reya and A. Vogt, *Z. Phys.* **C 48** (1990) 471 and Proceedings of the 'Small-x' Meeting, DESY, 1990.

[63] D. Appel, G. Sterman, and P. Mackenzie, *Nucl. Phys.* **B 309** (1988) 259.

[64] E. Laenen, J. Smith, and W.L. van Neerven, *Nucl. Phys.* **B 369** (1992) 543; *Phys. Lett.* **B 321** (1994) 254.

[65] E.L. Berger and H. Contopanagos, *Phys. Lett.* **B 361** (1995) 115.

[66] F.A. Berends, J.B. Tausk, and W.T. Giele, Phys. Rev. **D47** (1993) 2746.

[67] S. Güsken, J. H. Kühn and P. M. Zerwas, *Phys. Lett.* **B 155** (1985) 185.

[68] V. Fadin, V. Khoze, *JETP Lett.* **46** (1987) 417; *Sov. J. Nucl. Phys.* **48** (1988) 487;
V. Fadin, V. Khoze and T. Sjöstrand, *Z. Phys.* **C 48** (1990) 467.

[69] W. Hollik et al. in Ref.[44];
W. Beenakker, A. Denner, W. Hollik et al., R. Mertig, T. Sack, and D. Wackeroth, *Nucl. Phys.* **B 411** (1994) 343.

[70] B. Lampe, *Phys. Lett.* **B 348** (1995) 196.

[71] V. Barger et al., *Phys. Rev.* **D 35** (1987) 3366.

[72] J.H. Kühn and E. Mirkes, *Phys. Lett.* **B 311** (1993) 458.

[73] J.H. Kühn and M. Peter, *Z. Phys.* **C 67** (1995) 271.

[74] J.H. Kühn and E. Mirkes, *Phys. Lett.* **B 296** (1992) 425.

[75] J.H. Kühn and E. Mirkes, *Phys. Rev.* **D 48** (1993) 179.

[76] A. Djouadi, M. Spira, and P.M. Zerwas, *Phys. Lett.* **B 264** (1991) 440.

[77] O.J.P. Eboli, G.C. Marques, S.F. Novales and A.A. Natale, *Phys. Rev.* **D 34** (1986) 771,
S. Dawson and S. Willenbrock, *Nucl. Phys.* **B 284** (1987) 449,
C.P. Yuan, *Nucl. Phys.* **B 310** (1988) 1,
R.P. Kauffman, *Phys. Rev.* **D 41** (1990) 3343,
H. Veltman, *Phys. Rev.* **D 43** (1991) 2236.

[78] K. Hagiwara, H. Murayama and T. Watanabe, *Nucl. Phys.* **B 367** (1991) 257.

[79] A. Djouadi, J. Kalinowski and P.M. Zerwas, *Z. Phys.* **C 54** (1992) 255.

[80] J.H. Kühn, E. Mirkes, and J. Steegborn, *Z. Phys.* **C 57** (1993) 615..

[81] M. Consoli, W. Hollik and F. Jegerlehner, in "Z Physics at LEP 1", ed. G. Altarelli, CERN89-08.

[82] J. Jersak, E. Laermann and P.M. Zerwas, *Phys. Rev.* **D 25** (1982) 363,
J. Schwinger, "Particles, Sources and Fields", Addison-Wesley (1973),
L. Reinders, H. Rubinstein and S. Yazaki, *Phys. Reports* **C 127** (1985) 1.

[83] J.H. Kühn and P.M. Zerwas, in *Advanced Series on Directions in HEP*, Vol. 10, *Heavy Flavours*, Eds. A.J. Buras and M. Lindner, World Scientific, 1992, p. 434.

[84] J. Schwinger in [82], and ref. [67].

[85] A.H. Hoang, J.H. Kühn, and T. Teubner, *Nucl. Phys.* **B 452** (1995) 173.

[86] K. Chetyrkin, J.H. Kühn, and M. Steinhauser, *Phys. Lett.* **B 371** (1996) 93.

[87] M. Jeżabek, *Z. Phys.* **C 56** (1992) 285.

[88] S. Jadach and B.F.L. Ward, *Comp. Phys. Commun.* **56** (1990) 351.

[89] W. Beenakker, S.C. van der Marck and W. Hollik, *Nucl. Phys.* **B 365** (1991) 24.

[90] M. Jeżabek and J.H. Kühn, Phys. Lett. **B316** (1993) 360.

[91] W. Beenakker and W. Hollik, *Phys. Lett.* **B 269** (1991) 425.

[92] B. Grzadkowski, P. Krawczyk, J.H. Kühn, and R.G. Stuart, Nucl. Phys. **B281** (1987) 18.

[93] C. Peterson, D. Schlatter, I. Schmitt and P.M. Zerwas, *Phys. Reports* **27** (83) 105.

[94] T. Sjöstrand and P.M. Zerwas, in Proceedings, e^+e^- *Collisions at 500 GeV: The Physics Potential*, DESY 92-123, pg. 463-475.

[95] G. Jikia, *Phys. Lett.* **B 257** (1991) 196.

[96] V.A. Khoze, L.H. Orr, and J. Sterling, *Nucl. Phys.* **B 378** (1992) 413.

[97] G.L. Kane, G.A. Ladinsky and C.–P. Yuan, *Phys. Rev.* **D 45** (1992) 124.
C.–P. Yuan, *Phys. Rev.* **D 45** (1992) 782.

[98] W. Bernreuther and P. Overmann, Proc. of the Workshop e^+e^- Collisions at 500 GeV, the Physics Potential, Part C, P.M. Zerwas (ed.), DESY 93-123C, p. 287, and refs. therein.

[99] C. Schmidt and M.E. Peskin, *Phys. Rev. Lett.* **69** (1992) 410.

[100] J.H. Kühn, A. Reiter and P.M. Zerwas, *Nucl. Phys.* **B 272** (1986) 560.

[101] A. Devoto, J. Pumplin, W. Repko and G.L. Kane, *Phys. Rev. Lett.* **43** (79) 1062;

[102] J.H. Kühn, *Nucl. Phys.* **B 237** (1984) 77.

[103] L. Landau and E. Lifshitz, *Non-relativistic Quantum Mechanics*, sect. 36.

[104] A. Messiah, *Quantum Mechanics*, Chapter XI, sect. 7-11 and Appendix B, sect. 4.

[105] V.S. Fadin and V.A. Khoze, *JETP Lett.* **46** (1987) 525,

[106] T. Appelquist, M. Dine and I. Muzinich, *Phys. Rev.* **D 17** (1978) 2074.

[107] L. Susskind, in "Weak and Electromagnetic Interactions at High Energies", Les Houches (1977).

[108] W. Fischler, *Nucl. Phys.* **B 129** (1977) 157.

[109] A. Billoire, *Phys. Lett.* **B 92** (1980) 343.

[110] W. Bernreuther and W. Wetzel, *Nucl. Phys.* **B 197** (1982) 228.

[111] W.J. Marciano, *Phys. Rev.* **D 29** (1984) 580.

[112] M. Jeżabek, J.H. Kühn and T. Teubner, *Z. Phys.* **C 56** (1992) 653

[113] J.M. Strassler and M.E. Peskin, *Phys. Rev.* **D 43** (1991) 1500.

[114] M. Jeżabek and T. Teubner, *Z. Phys.* **C 59** (1993) 669

[115] For a pedagogical discussion see e.g. the appendix of W. Buchmüller and S.-H. H. Tye, Phys. Rev. **D24** (1981) 132.

[116] P. Igo-Kemenes, M. Martinez, R. Miguel, and S. Orteu, in Proceedings, e^+e^- *Collisions at 500 GeV: The Physics Potential*, Part C; ed. P.M. Zerwas, DESY 93-123C, pg. 319.

[117] R. Guth and J.H. Kühn, Nucl. Phys. **B368** (1992) 38.

[118] M. Jeżabek and J.H. Kühn, Phys. Lett. **B207** (1988) 91.

[119] Y. Sumino, K. Fujii, K. Hagiwara, H. Murayama, C.-K. Ng, *Phys. Rev.* **D 47** (1992) 56

[120] M. Jeżabek, J.H. Kühn and T. Teubner, in Proceedings, e^+e^- *Collisions at 500 GeV: The Physics Potential*, DESY 92-123C, pg. 303.

[121] J.M. Strassler and M.E. Peskin, *Phys. Rev.* **D 43** (1991) 1500

[122] H. Murayama and Y. Sumino, *Phys. Rev.* **D 47** (1993) 82.

[123] R. Harlander, M. Jeżabek, J.H. Kühn and T. Teubner, *Phys. Lett.* **B 346** (1995) 137.

[124] J.H. Kühn, A. Reiter and P.M. Zerwas, *Nucl. Phys.* **B 272** (1986) 560.

[125] A. Devoto, J. Pumplin, W. Repko and G.L. Kane, *Phys. Rev. Lett.* **43** (1979) 1062.

[126] V.S. Fadin, V.A. Khoze and M.I. Kotsky, *Z. Phys.* **C 64** (1994) 45.

[127] R. Harlander, M. Jeżabek, J.H. Kühn, and M. Peter, Karlsruhe preprint TTP 95-48, in print in Z. Phys. **C**.

[128] K. Melnikov and O. Yakovlev, *Phys. Lett.* **B 324** (1994) 217.

[129] Y. Sumino, PhD thesis, University of Tokyo, 1993 (unpublished).

[130] R.W. Huff, *Ann. Phys.* **16** (1961) 288.

[131] K. Fujii, T. Matsui and Y. Sumino, *Phys. Rev.* **D 50** (1994) 4341.

[132] H. Überall, *Phys. Rev.* **119** (1960) 365.

[133] W. Mödritsch and W. Kummer, *Nucl. Phys.* **B 430** (1994) 3.

TOP QUARK STUDIES AT HADRON COLLIDERS

Pekka K. Sinervo*
Department of Physics
University of Toronto, Toronto, Ontario, Canada M5S 1A7

ABSTRACT

The techniques used to study top quarks at hadron colliders are presented. The analyses that discovered the top quark are described, with emphasis on the techniques used to tag b quark jets in candidate events. The most recent measurements of top quark properties by the CDF and D$\emptyset$ collaborations are reviewed, including the top quark cross section, mass, branching fractions and production properties.

Future top quark studies at hadron colliders are discussed, and predictions for event yields and uncertainties in the measurements of top quark properties are presented.

*Supported by the Natural Sciences and Engineering Council of Canada.

1 Introduction

1.1 The Case for Top

The top quark and the Higgs boson are the heaviest elementary particles predicted by the standard model.[1] The four lightest quark flavours, the up, down, strange and charm quarks, were well-established by the mid-1970's. The discovery in 1977[2] of the Υ resonances, a new family of massive hadrons, required the introduction of the fifth quark flavour. Experimental and theoretical studies have indicated that this quark has a heavier partner, the top quark.

Indirect evidence for the top quark comes from a number of sources. The most compelling data come from the observed properties of the scattering process $e^+e^- \rightarrow b\bar{b}$, where the asymmetry in the scattering of the b quark relative to the incoming electron direction implies that the b quark has weak isospin of 0.5. The most precise measurement of this comes from the LEP collider, where this asymmetry has been found[3] to be in excellent agreement with the standard model expectation of 0.100 assuming that the b quark is a member of an $SU(2)$ doublet. The other member of that doublet would by definition be the top quark.

Additional indirect evidence comes from the study of b quark decays. It has been experimentally determined that the b quark does not decay via processes that yield zero net flavour in the final state (*e.g.*, $b \rightarrow \mu^+\mu^- X$), or where the decay results in only a quark of the same charge (*e.g.*, $b \rightarrow sX$ where X is a state with no net flavour quantum numbers).[4] The absence of these "flavour-changing neutral currents" in the standard model implies that the b quark is a member of an $SU(2)$ doublet.

Finally, evidence for the existence of a massive fermion that couples via the electroweak force to the b quark comes from detailed measurements of the Z° and W^+ bosons performed at LEP, SLC, the CERN S$p\bar{p}$S and the Fermilab Tevatron Collider. This body of data, and in particular the radiative mass shifts of the electroweak bosons, can only be described in the standard model by introducing a top quark. A recent compilation of data[5] indicates that the standard model top quark has a mass of

$$M_{top} = 169^{+16}_{-18}\ {}^{+17}_{-20}\ \mathrm{GeV/c^2}. \qquad (1)$$

The second uncertainty corresponds to variations of the unknown Higgs boson mass between 60 and 1000 GeV/c^2 (its nominal value is 300 GeV/c^2).

Taken together, these observations make a strong case for the top quark's existence. They also imply that our understanding of nature via the standard model would be profoundly shaken if the top quark was shown not to exist with its expected properties. The observation of the top quark is therefore of considerable significance.

1.2 Earlier Top Quark Searches

Direct searches for the top quark have been performed at virtually all of the high-energy collider facilities that have operated in the last twenty years.[6] The most model-independent searches have taken place at e^+e^- colliders, where one looks for the production and decay of a pair of massive fermions. Because of the relatively large mass of the top quark, its decay yields events that are quite spherical and are relatively easy to separate from the background of lighter quark production. The most stringent limits have been set by the LEP collaborations, which require that $M_{top} > 46$ GeV/c^2 at 95% confidence level (CL). These limits are insensitive to the decay modes of the top quark and the coupling of the top quark to the electroweak bosons.

Another relatively model-independent limit is set by measurements of the width of the W^+ boson. Direct and indirect measurements[7] of Γ_W indicate that the top quark is massive enough that the decay channel $W^+ \to t\bar{b}$ does not contribute to Γ_W. The limit set is $M_{top} > 62$ GeV/c^2 at 95% CL.

Direct searches for the top quark at hadron colliders have focused on two specific models for top quark decay: i) the minimal supersymmetric model (MSSM)[8] where the decay mode $t \to H^+b$ is also allowed (H^+ is the charged Higgs boson), and ii) the standard model where the top quark decays directly to $t \to W^+b$. The most stringent limit[9] assuming the MSSM requires that $M_{top} > 96$ GeV/c^2 at 95% CL for the case where $t \to H^+b$ always and $BR(H^+ \to \tau^+\nu_\tau) = 1.0$. This limit, however, depends on the overall width of the decay $t \to H^+b$, the Higgs boson branching fractions (H^+ is expected to preferentially decay to $c\bar{s}$ and $\tau\nu_\tau$ final states) and the H^+ detection efficiency. The DØ collaboration has published the most sensitive standard model search using a 15 pb^{-1} dataset, and has excluded a top quark with mass less than 131 GeV/c^2 at 95% CL.[10]

On the other hand, the CDF collaboration published a study of $\sim$ 20 pb^{-1} of data in April 1994 that claimed evidence for top quark production.[11] A total

of 12 events were observed in several decay modes above a predicted background of approximately 6 events. The probability that the observed event rate was consistent with a background fluctuation was estimated to be 0.25%. In addition, evidence was presented that the events in the sample were consistent with arising from the production and decay of a $t\bar{t}$ system and inconsistent with the properties expected of the dominant backgrounds. Although compelling, this observation was statistically limited and the possibility that it arose from a background fluctuation could not be ruled out.

In this report, I will focus on the latest results to come from the DØ and CDF top quark searches using data collected between 1992 and 1995. Both collaborations have acquired over three times more data, and have now reported conclusive evidence for top quark production.[12] I will describe the analyses performed by both collaborations and compare the two results.

I believe an extremely persuasive case has been made that the top quark has been found.

2 Production and Decay of Heavy Top

The production of heavy quarks in 1.8 TeV proton-antiproton ($p\bar{p}$) collisions is predicted to take place through the two leading-order quantum-chromodynamic (QCD) diagrams

$$q\bar{q} \rightarrow Q\bar{Q} \tag{2}$$

$$gg \rightarrow Q\bar{Q}, \tag{3}$$

with the relative rate of these two processes dictated largely by the mass of the heavy quark (Q), the parton distribution functions of the proton and phase space. Top quark pair-production is expected to dominate the production rate. The production of single top quarks through the creation of a virtual W^+ is smaller[13] (of order 10% of the $t\bar{t}$ rate) and expected to occur in a relatively small part of phase space. All heavy top quark searches have therefore ignored single top production.

The next-to-leading order corrections[14] to processes (2) and (3) are relatively small for heavy quark masses greater than $\sim$ 50 GeV/c^2. More recently, these estimates have been revised taking into account the effects of internal soft-gluon emission.[15,16] These cross sections are shown in Fig. 1 plotted as a function of

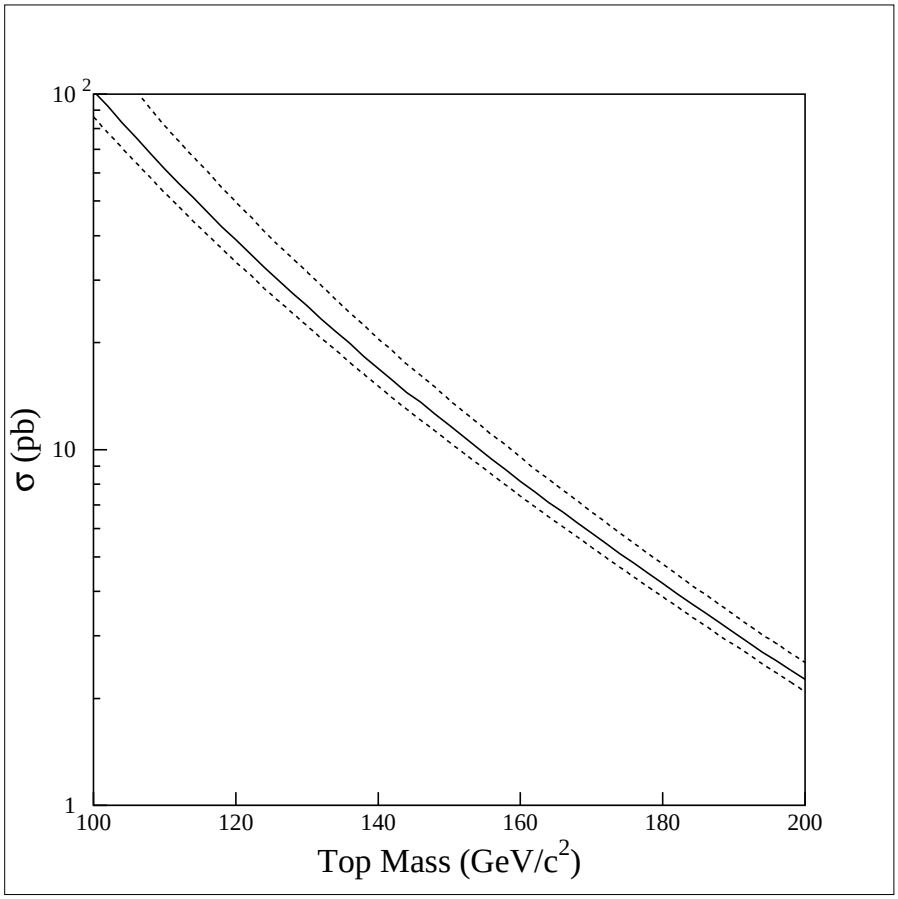


Figure 1: The total cross section for top quark production in 1.8 TeV $p\bar{p}$ collisions as estimated by E. Laenen *et al.*. The upper and lower curves are a measure of the theoretical uncertainties in the calculation.

the heavy quark mass. The uncertainty in these estimates reflects the theoretical uncertainty in this calculation, which is believed to be the choice of renormalisation scale. For top quark masses above 100 GeV/c^2, the primary contribution to the cross section comes from quark annihilation. This reduces the uncertainties arising from our lack of knowledge of the parton distribution functions of the proton, as these have been relatively accurately measured at large Feynman x, the kinematic region that would dominate very heavy quark production.

Top quark pair production will generate a top quark and anti-top quark that are recoiling against each other in the lab. The production diagrams favour configurations where the top quarks are produced isotropically in the lab frame. The relative motion of the $t\,\overline{t}$ system is expected to be small in comparison to the trans-

verse momentum* (P_T) distribution of the top quark itself.[17] The expected P_T distribution for a heavy top quark has a peak around half the top quark mass with a relatively long tail. The pseudorapidity distribution for top quarks is peaked at 0 and falls off rapidly so that most of the top quarks are produced in the "central" region with pseudorapidity $|\eta| < 2$. The combination of a relatively energetic heavy quark produced centrally is ideal from an experimental point of view. The top quark decay products are rather stiff and central, aiding their detection.

The standard model predicts that the top quark will decay almost always via $t \to W^+ b$. The W^+ decays approximately 2/3 of the time into $q\bar{q}'$ pairs ($u\bar{d}$ or $c\bar{s}$) and 1/3 of the time into one of the three lepton generations. This results in a decay topology consisting of 6 energetic partons that could either be charged leptons, neutral leptons or quark jets.

The decay channels involving τ^+ leptons are problematic given the difficulty of cleanly identifying these weakly decaying leptons in a hadron collider environment. They have therefore not been explicitly included in the searches I describe below. The final states involving 6 quark jets suffer an enormous background from QCD multijet production, with estimates of intrinsic signal-to-noise of $< 10^{-4}$. Because of these large backgrounds, this channel has not been the focus of most of the effort, and I will ignore it here also. However, recent work has demonstrated that a significant $t\,\overline{t}$ signal can be observed in these modes.[18]

With these considerations, there are five final states that are experimentally accessible:

$$\begin{aligned}
t\,\overline{t} &\to e^+\nu_e b\; e^-\bar{\nu}_e\bar{b} \quad (1/81) \\
t\,\overline{t} &\to \mu^+\nu_\mu b\; \mu^-\bar{\nu}_\mu\bar{b} \quad (1/81) \\
t\,\overline{t} &\to e^+\nu_e b\; \mu^-\bar{\nu}_\mu\bar{b} \quad (2/81) \\
t\,\overline{t} &\to e^+\nu_e b\; q\bar{q}'\bar{b} \quad (12/81) \\
t\,\overline{t} &\to \mu^+\nu_\mu b\; q\bar{q}'\bar{b} \quad (12/81),
\end{aligned} \tag{4}$$

where I have also listed the expected standard model branching fractions for each channel. In all cases where I refer to a specific charge state, the charge conjugate

*I will employ a coordinate system where the proton beam direction defines the $\hat{z}$ axis, and transverse variables such as transverse momentum (P_T) and transverse energy (E_T) are defined relative to this axis. The angle ϕ represents the azimuthal angle about the beam axis and the angle θ represents the polar angle relative to the beam axis. Pseudorapidity $\eta \equiv -\ln\tan(\theta/2)$ will often be employed instead of θ.

mode is implied. The first three dilepton channels turn out to be the cleanest final states, as the requirement of two energetic charged leptons and neutrinos virtually eliminates all backgrounds. They suffer from rather small branching fractions and are therefore the most statistically limited. The last two lepton+jets final states together correspond to approximately 30% of the $t\bar{t}$ branching fraction. However, these channels face the largest potential backgrounds.

3 Backgrounds to Top Quark Searches

Top quark production is an extremely rare process in $p\bar{p}$ collisions; its cross section of less than 100 pb can be compared with the total $p\bar{p}$ cross section of over 50 mb (almost nine orders of magnitude difference). Since the total cross section is dominated by "soft" QCD interactions, the top quark cross section can be more fairly compared with the cross section for other high Q^2 production processes, such as inclusive W^+ production (20 nb), Z° production (2 nb) and W^+W^- and W^+Z° production (10 and 5 pb, respectively). These processes are the sources of the most severe background to $t\bar{t}$ production.

It is necessary to control these backgrounds so that one can be sensitive to a top quark signal. All the channels listed in Eqs. (4) involve an energetic charged electron or muon, and one or more energetic neutrinos. The requirement of these two signatures in the final state using the DØ and CDF lepton identification systems are sufficient to adequately control the backgrounds associated with jets that might satisfy the lepton ID criteria. The remaining backgrounds are dominated by physics processes that generate real leptons in the final state.

In the case of the dielectron and dimuon modes, the single largest background comes from Drell-Yan production (including $Z^\circ \to e^+e^-$ and $Z^\circ \to \mu^+\mu^-$). This is controlled by requiring a neutrino signature as well as additional jet activity. The single largest physics background in the $e^+\mu^-$ final state comes from $Z^\circ \to \tau^+\tau^-$ decay, which can be similarily reduced by the requirement of a neutrino signature and additional jets.

The single largest physics background to lepton+jets final states come from inclusive W^+ production where additional jets are produced via initial and final state radiation.[19] The intrinsic rate for this background depends strongly on the multiplicity requirements placed on the jet candidates, as shown in Table 1 where the observed W+jet production cross section is presented as a function of jet

Jet Multiplicity	σB (pb)	$\sigma_T B$ (pb)
0	$1740 \pm 31 \pm 288$	$1753 \pm 26 \pm 123$
1	$336 \pm 14 \pm 63$	$287 \pm 4 \pm 21$
2	$76 \pm 12 \pm 18$	$59 \pm 2 \pm 5$
3	$14 \pm 3 \pm 3$	$11.0 \pm 0.3 \pm 1.0$
4	$4.0 \pm 1.6 \pm 1.2$	$2.0 \pm 0.1 \pm 0.3$

Table 1: The W+jet production cross section times the branching ratio for $W^+ \to l^+\nu_l$ as a function of jet multiplicity. The second column presents the observed cross sections for jets with corrected transverse energy > 15 GeV and $|\eta| < 2.4$. The third column shows the predicted QCD cross section based on a VECBOS Monte Carlo calculation.

multiplicity and compared with a QCD Monte Carlo prediction.[20] One can see from these rates that this background can overwhelm a $t\bar{t}$ signal. More stringent kinematic cuts can be applied to reject the W+jet events, taking advantage of the fact that the $t\bar{t}$ final states, on average, generate higher E_T W^+ bosons and additional jets. Alternatively, since the $t\bar{t}$ final state has two b quark jets in it, the requirement that one or more jets are consistent with arising from the fragmentation and decay of a b quark will preferentially reduce the W+jets background. Both of these techniques have been employed.

4 The Tevatron Collider

The Tevatron Collider is a 6 km circumference proton-antiproton storage ring that creates $p\bar{p}$ collisions at a centre-of-mass energy of 1.8 TeV. In its current configuration, the collider operates with six bunches of protons and six bunches of counter-rotating antiprotons that are brought into collision at two intersection points in the ring named B0 and D0. The B0 and D0 interaction regions house the CDF and D$\emptyset$ detectors, respectively.

The Tevatron embarked on a multi-year collider run starting in December 1992. The first stage of the run, known as Run IA, continued till August 1993, at which time approximately 30 pb^{-1} had been delivered to each interaction region. The second stage, Run IB, commenced in August 1994 and by February 1995

the collider had delivered an additional 80 pb^{-1} to each interaction region. The maximum luminosity of the Collider during this period was 1.7×10^{31} cm^{-2}s^{-1}, and has been steadily rising.

Run IB run ended in February 1996, with a total of ~ 150 pb^{-1} delivered to each interaction region.

5 The DØ and CDF Experiments

The DØ and CDF detectors have been designed to trigger and record the high P_T collisions that result when two partons in the $p\bar{p}$ system undergo a hard scatter. Both instruments detect electrons, muons, neutrinos and quark and gluon jets using a set of complementary subdetectors. However, they accomplish this common goal in rather different ways.

5.1 The DØ Detector

The DØ detector was designed with the philosophy that a uniform, hermetic, highly-segmented calorimeter should form the core of the detector.[21] A cut-away view of the detector is shown in Fig. 2. The DØ calorimeter employs a uranium absorber up to nine interaction lengths thick and a liquid argon readout system. This provides excellent hermeticity and uniformity, except perhaps in the transition region between the barrel and endcap cryostats. The overall resolution of the DØ calorimeter is

$$\frac{\sigma_E}{E} = \frac{0.15}{\sqrt{E}} \oplus 0.004 \text{ for electromagnetic showers} \quad (5)$$

$$\frac{\sigma_E}{E} = \frac{0.80}{\sqrt{E}} \text{ for hadrons,} \quad (6)$$

where E is measured in GeV.

A muon system consisting of charged particle detectors and 1.9 Tesla toroidal magnets located outside the calorimeter provides good muon identification. This system identifies muon candidates in the region $|\eta| < 3.3$ using sets of muon tracking chambers consisting of proportional drift tubes located interior and exterior to the large toroidal magnetic field. The deflection of the muon candidates in the magnetic field provides a momentum measurement with an accuracy of

$$\sigma\left(\frac{1}{p}\right) = \frac{0.18\,(p-2)}{p^2} \oplus 0.008, \quad (7)$$

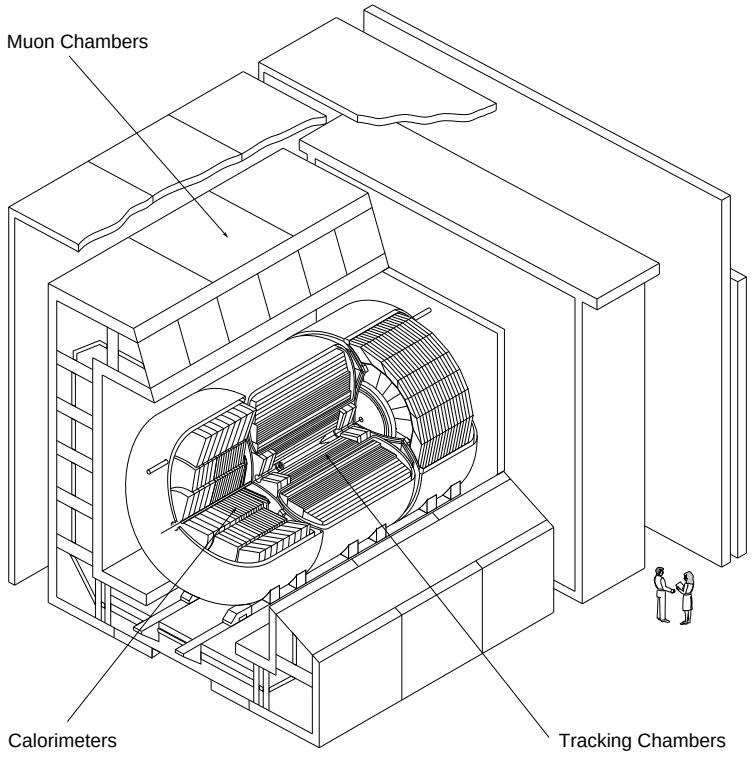


Figure 2: A cut-away view of the DØ detector. The inner tracking detectors are surrounded by the calorimeter cryostats, and both are situated inside the toroidal magnet. Planes of chambers outside the magnet provide muon identification and momentum measurement.

where p is the muon momentum measured in GeV/c.

Vertex, central and forward drift chambers provide charged particle detection in the interval $|\eta| < 3.2$. The tracking system does not incorporate a magnetic field, as the presence of a magnetic coil would degrade calorimeter performance.

5.2 The CDF Detector

The CDF detector[22] consists of a high-precision tracking system in a 1.4 T solenoid magnetic field, surrounded by a hermetic highly-segmented calorimeter, as shown in Fig. 3. The tracking system consists of three independent devices arranged coaxial to the beam line. A 4-layer silicon-strip detector (SVX) with inner and outer radii of 3.0 and 7.9 cm provides of order 40 μ precision on the impact

parameter of individual charged track trajectories extrapolated to the beam line. A set of time projection chambers (VTX) instrument the tracking region between 12 and 22 cm in radius, providing high-precision tracking in the r-z plane. An 84-layer drift chamber (CTC) detects charged particles in the region between 30 and 132 cm from the beamline. Together, these detectors measure particle transverse momentum to a precision σ_{p_T} given by

$$\frac{\sigma_{p_T}}{p_T} = 0.0009 p_T \oplus 0.0066, \tag{8}$$

for particles with $p_T \gtrsim 0.35$ GeV/c.

The central calorimeter (CEM and CHA) instruments the region $|\eta| < 1.1$, and is comprised of projective towers of size $\Delta\eta \times \Delta\phi = 0.1 \times 0.26$ radians. Each tower is made of a sandwich of Pb or Fe plates interleaved with scintillator. A Pb sandwich 25 radiation lengths thick is used to measure electromagnetic shower energies. An iron-scintillator sandwich approximately 5 interaction lengths thick is used to detect hadronic showers. Plug and Forward calorimeters (PEM, PHA, FEM and FHA) instrument the region $1.1 < |\eta| < 4.2$, and consist of similar absorber material. The showers in this region are detected with proportional wire chambers as they provide for a more radiation-resistant detector system. The presence of a solenoid magnet and a significant amount of material in front of the calorimeter leads to some compromise in calorimeter performance. The overall resolution of the CDF calorimeter is

$$\frac{\sigma_E}{E} = \frac{0.137}{\sqrt{E}} \oplus 0.02 \quad \text{(for electromagnetic showers)} \tag{9}$$

$$\frac{\sigma_E}{E} = \frac{0.50}{\sqrt{E}} \oplus 0.03 \quad \text{(for hadrons)}. \tag{10}$$

Planar drift chambers (CMU, CMP and CMX) located outside the calorimeter volume detect muons penetrating the calorimeter absorber, but precise muon momentum and direction come from the associated charged track detected in the inner tracking system. The central muon system is able to detect muons within the pseudorapidity interval $|\eta| < 1.0$. A forward muon system (FMU) consisting of large toriodal magnets surrounded by drift chambers and scintillator counters detect muons in the rapidity region $2.2 \leq |\eta| \leq 3.5$.

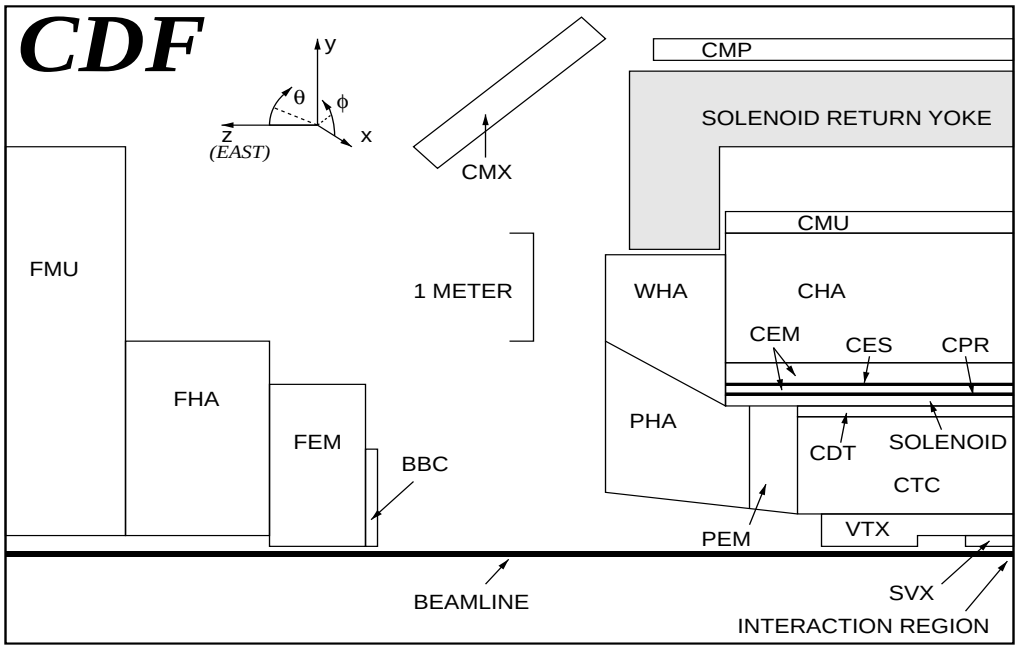


Figure 3: A schematic view of one quarter of the CDF detector. The interaction point is at the lower right corner of the figure.

5.3 Triggering and Data Acquisition

Pair production of standard model top quarks and their subsequent decay into either the dilepton or lepton+jets mode yields a signature that is relatively straightforward to trigger on. Both detectors employ multi-level trigger systems where at each level more information is brought together to form a decision. The trigger requirement of at least one energetic electron or muon is the primary tool used in identifying online a sample of top quark candidate events that are subsequently studied offline.

The requirement of at least one high P_T electron or muon in both CDF and DØ is imposed efficiently in the trigger. The production of leptons above a transverse energy of 15 GeV is dominated in both experiments by b and c quark production, and by inclusive W^+ boson production. For example, in CDF, the inclusive electron trigger is implemented with the following requirements:

1. The level 1 trigger demands that at least one calorimeter trigger cell with $\Delta\phi \times \Delta\eta = 0.26 \times 0.2$ has > 6 GeV of electromagnetic energy.
2. The level 2 trigger demands that there be a charged track candidate pointing

at an electromagnetic energy cluster, and requires that the cluster properties be consistent with those of an electromagnetic shower.

3. The level 3 trigger requires the presence of an electromagnetic cluster associated with a charged track reconstructed using the standard offline algorithms. Further quality cuts on the properties of the electromagnetic shower are also made.

These reduce the overall cross section of candidate events to approximately 50 nb, of which approximately 30% is comprised of real electrons. For comparison, the rate of $W^+ \to e^+\nu_e$ in this sample is of order 1 nb. The efficiency of this trigger for isolated electrons with $20 < E_T < 150$ GeV is $92.8 \pm 0.2\%$.

As another example, the D$\emptyset$ detector triggers on a sample of inclusive muon candidates by using a two level decision process:

1. The level 1 trigger demands the presence of a charged track stub in the muon toroidal spectrometer with a $p_T > 3$ GeV/c.

2. The level 2 trigger demands a high quality muon candidate consisting of a muon candidate in the muon system matched to a charged track observed in the central tracking system. The central track candidate must be reconstructed in all 3 dimensions, must be consistent with coming from the event interaction and must have P_T greater than 5 or 8 GeV/c, depending on the specific muon trigger.

The efficiency of this trigger is estimated to be $67 \pm 3\%$.

Both experiments employ inclusive electron and muon triggers, as well as triggers that identify smaller samples of events useful to the top search. Since the backgrounds to the dilepton sample are relatively small, it is convenient to identify the candidate events immediately in the trigger so that they can be analysed as soon as possible. A high-P_T dilepton trigger requiring at least two electron or muon candidates is therefore employed to flag these candidates immediately. The cross section for this trigger is only a few nb.

At a luminosity of 2×10^{31} cm^{-2}s^{-1}, a trigger cross section of 300 nb corresponds to an event rate of 6 Hz, which can be comfortably recorded and analyzed. Note, however, that even with a cross section of 10 nb, the total data sample for an integrated luminosity of 50 pb^{-1} will consist of 500 000 events, with each event comprised of order 200 kbytes of information.

5.4 The Run IA and IB Datasets

The Tevatron Collider started up after a three year shut-down in fall 1992, and continued running through the summer of 1993. As this was the DØ detector's first collider run, it was remarkable that the collaboration was able to successfully use 40-50% of the collisions for their physics studies. The CDF collaboration gathered 19.6 ± 0.7 pb^{-1} of data during this period.

From the start of Run IB in 1994 to February 1995, the Tevatron Collider had delivered over 100 pb^{-1} of collisions to each detector. The DØ and CDF collaborations had recorded and analysed ~ 45 pb^{-1} of this data by this date, giving the the two collaborations total Run I datasets of 50 and 67 pb^{-1}, respectively.

In between Run IA and IB, both collaborations made incremental improvements to their detectors. The DØ detector's muon trigger was improved and various detector subsystems were modified with the goal of improving overall robustness and efficiency. The CDF collaboration replaced the original 4-layer SVX detector with a mechanically identical device that used newer, radiation-hard silicon strip wafers, and employed an AC-coupled readout design. The new detector, known as the SVX', has much better signal-to-noise and is fundamentally better understood.

5.5 Event Reconstruction

A schematic of a $t\bar{t}$ event being produced in a $p\bar{p}$ collision and decaying into the final state partons is shown in Fig. 4 Given the large number of partons that arise from the decay of the $t\bar{t}$ system, each detector is required to reconstruct with good efficiency high energy electrons, muons and the jets resulting from the fragmentation of high energy quarks, and to tag the presence of one or more neutrinos by the imbalance of total transverse energy in the collision.

High energy electrons and muons are identified in both detectors by the charged track left in the central tracking systems, and by the behaviour of the leptons in the calorimeters and muon identification systems outside the calorimeters. Electrons will generate an electromagnetic shower in the calorimeter, with a lateral and longitudinal shower profile quite distinct from the shower intitiated by a charged hadron. Muons are readily identified as they generally pass unimpeded through the calorimeter and are detected outside the calorimeters as charged particles that point back to the particle trajectory in the central tracker. The CDF electron and

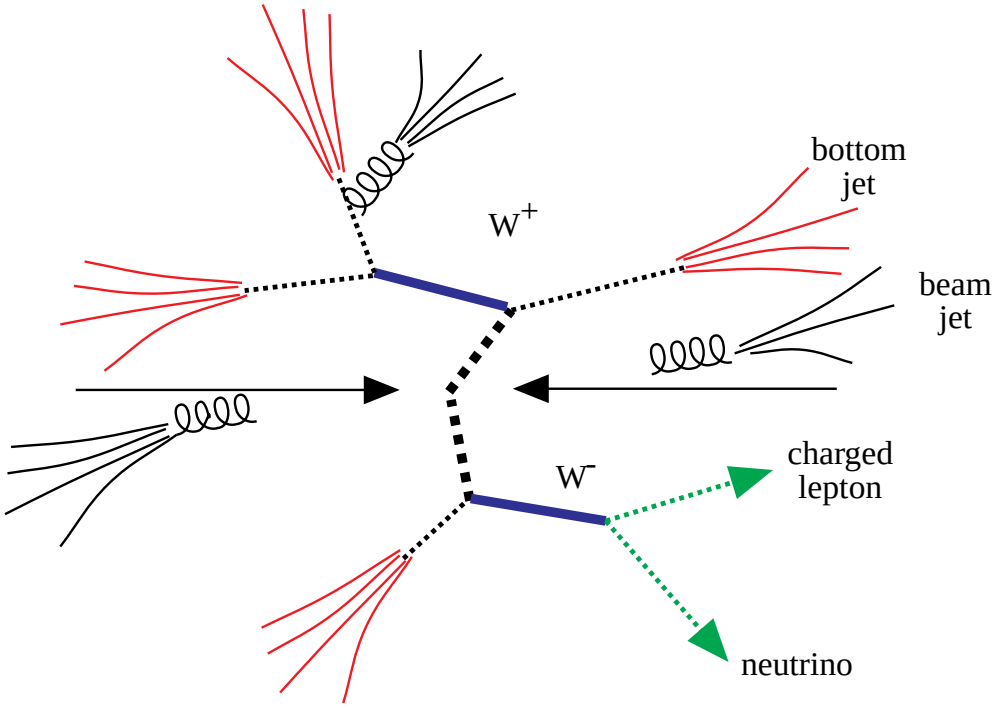


Figure 4: A schematic of a $t\bar{t}$ event produced at the Tevatron and decaying into a lepton+jets final state. In addition to the partons resulting from the decay, there are additional jets produced by initial and final state radiation.

muon reconstruction algorithms have efficiencies of $84 \pm 2\%$ and $90.6 \pm 1.4\%$ for leptons from W^+ boson decays. The D$\emptyset$ electron reconstruction has an efficiency of $72 \pm 3\%$. These efficiencies are quoted for electron and muon candidates that have already passed the trigger requirements discussed earlier.

Neutrinos can only be detected by requiring that they have sufficient transverse energy that the total measured energy flow sum to a value inconsistent with zero. In practical terms, this energy flow vector is known as missing transverse energy ($\not\!\!E_T$). Note that we cannot use the imbalance in energy flow along the beamline in this case as one can expect a significant imbalance due to the differing momentum of the partons in the proton and antiproton that collide to produce the $t\bar{t}$ system. The resolution in $\not\!\!E_T$ is driven by both the uniformity of the calorimeter and its inherent energy resolution. D$\emptyset$ has a missing transverse energy resolution in each

transverse coordinate of

$$\sigma_x = 1.08 + 0.019 \left(\sum E_T \right) \text{ GeV}, \tag{11}$$

where the summation gives the total scalar transverse energy observed in the calorimeter. CDF's transverse energy resolution is approximately 15-20% worse, which has a modest impact on its neutrino detection ability.

Jets are constructed in both detectors as clusters of transverse energy within a fixed cone defined in η-ϕ space.[23] The size of this cone is determined by the competing requirements of making it large enough to capture most of the energy associated with the fragmentation of a quark or gluon, and yet small enough that it doesn't include energy associated with nearby high energy partons or from the "underlying" event. The latter effect in itself contributes on average approximately 2 GeV per unit in η-ϕ space, and the fluctuations in this degrades the jet energy resolution (the size of this effect depends on the rate of multiple interactions). Monte Carlo (MC) calculations using a variety of models for quark fragmentation and underlying event assumptions, as well as studies of the underlying events have indicated that a jet cluster cone size substantially smaller than the traditional η-ϕ radii of 0.7 or 1.0 employed in QCD studies is required. The CDF analysis employs a cone radius of 0.4 in its top quark search, whereas the D$\emptyset$ collaboration has chosen to work with a cone radius of 0.5.

The requirement that most if not all daughters are reconstructed is not sufficient to reject all backgrounds to $t\,\overline{t}$ production. There are other kinematical variables that discriminate between $t\,\overline{t}$ and background events, most of them taking advantage of the fact that heavy top quark production will generate final state daughters that are on average quite energetic. This motivates the use of a variable called H_T defined as

$$H_T = \sum_{i=1}^{N_p} E_T^i, \tag{12}$$

where the sum is over all the jets and the leading electron cluster (in those channels where at least one electron is required). This variable is used by the D$\emptyset$ collaboration in both their dilepton and lepton+jets analysis, and its effectiveness in improving the signal-to-noise in the dilepton and lepton+jets channels is illustrated in Fig. 5. The CDF collaboration has recently reported the results of a top analysis using a similar variable.[24]

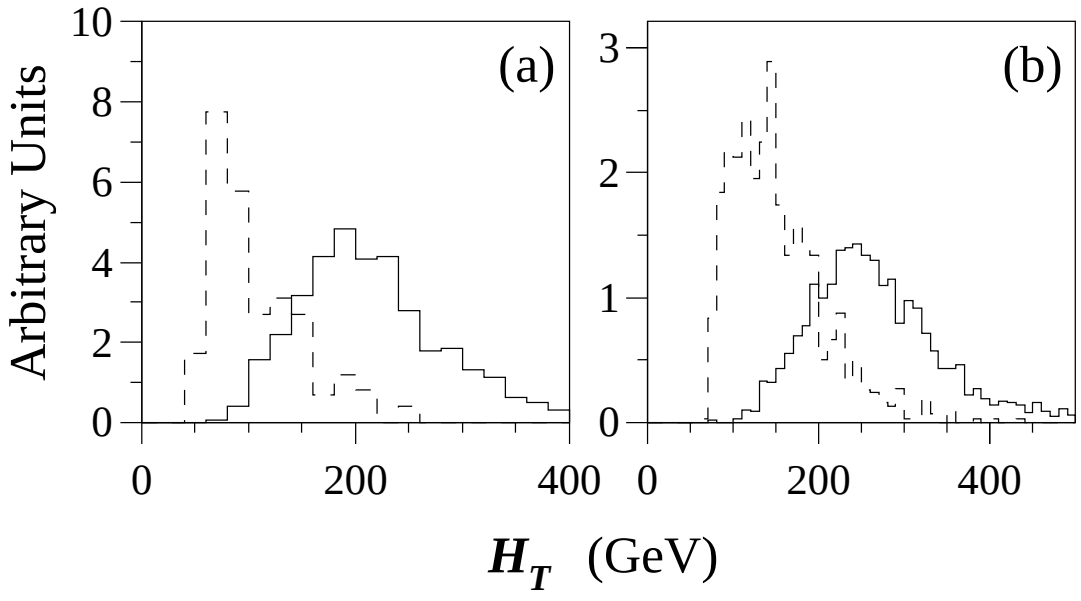


Figure 5: The H_T distributions for $e^+\mu^-$+jet events (a) and lepton+jet events (b). The solid histograms are the distributions expected from $t\bar{t}$ events for a top quark mass of 200 GeV/c^2. The dashed histograms are the expected distributions for the dominant backgrounds to $t\bar{t}$ production in both channels.

An additional kinematic variable known as aplanarity[25] ($\mathcal{A}$) has been employed by the D$\emptyset$ collaboration. This, as its name suggests, is a measure of how spherical a candidate event is: $t\bar{t}$ events are expected to have larger values of $\mathcal{A}$ than the corresponding physical backgrounds.

The final tool used in the reconstruction of $t\bar{t}$ events is the identification or "tagging" of jets that arise from the b quarks. There are two techniques employed by the collaborations. The first takes advantage of the fact that bottom hadrons decay semileptonically into electrons or muons about 20% of the time. D$\emptyset$ and CDF therefore search the interior of each jet cone for a muon candidate. CDF also searches for low-energy electron candidates that can be associated with the jet cluster. Because there are two b quarks in each $t\bar{t}$ decay, the efficiency of this soft lepton (SLT) tagging scheme ranges from 10-15%. The second technique is used exclusively by CDF and takes advantage of the long-lived nature of bottom hadrons and the SVX (or SVX') detector. A seach is performed for several charged tracks detected in the SVX that form a secondary vertex a significant distance from the primary interaction. The efficiency of this tagging scheme depends crucially on the performance of the SVX/SVX'. It is estimated that over 40% of all $t\bar{t}$

decays will have the presence of at least one SVX tag.

6 The Dilepton Top Quark Search

6.1 Dilepton Data Selection

The dilepton decay modes are the cleanest channel in which one would expect to observe a heavy top quark. They suffer from the relatively small total branching fraction of $t\bar{t}$ into these modes (a total of 4%), and from the presence of two neutrinos in the final state that are not individually observable.

The dilepton searches break down into three separate channels, the e^+e^-, $\mu^+\mu^-$ and $e^+\mu^-$ final states. The CDF analysis requires two isolated lepton candidates, each with $P_T > 20$ GeV/c and with $|\eta| < 1.0$. The candidates must satisfy standard lepton quality requirements that ensure high efficiency and high rejection from energetic, isolated charged hadrons. There are 2079 e^+e^- candidates, 2148 $\mu^+\mu^-$ candidates and 25 $e^+\mu^-$ candidates after these kinematical cuts. The large e^+e^- and $\mu^+\mu^-$ candidate samples are the result of Z° and Drell-Yan production, as can be seen by examining the invariant mass (M_{ll}) distribution of the dilepton system. This background is removed by rejecting those events with

$$75 < M_{ll} < 105 \text{ GeV/c}^2. \tag{13}$$

This leaves 215, 233 and 25 candidate events in the e^+e^-, $\mu^+\mu^-$ and $e^+\mu^-$ channels, respectively.

In addition, the events are required to have $\not\!\!E_T > 25$ GeV and at least two jet clusters with $E_T > 10$ GeV and $|\eta| < 2.0$, since $t\bar{t}$ events are expected to have two energetic neutrinos and a b quark and anti-quark in the final state. This still leaves a background in the e^+e^- and $\mu^+\mu^-$ sample from Drell-Yan production where the $\not\!\!E_T$ signal arises from an accompanying jet that is mismeasured. The distributions of the azimuthal opening angle between the missing transverse energy vector and the closest jet or charged lepton candidate in the event versus the missing transverse energy for each jet multiplicity are shown in Figs. 6 and 7 for the $\mu^+\mu^-$ and $e^+\mu^-$ channels, respectively. There is a clear cluster of events at small $\not\!\!E_T$-jet opening angles that extend to higher $\not\!\!E_T$ in the $\mu^+\mu^-$ (and e^+e^-) samples that results from the remnant Drell-Yan contamination in the samples. The same enhancement is not present in the $e^+\mu^-$ sample, which has no Drell-Yan

Mass (GeV/c^2)	D$\emptyset$	CDF
150	2.4	6.2
160	2.0	4.4
170	1.6	3.0
180	1.2	2.4

Table 2: The expected number of dilepton events arising from $t\bar{t}$ production for the D$\emptyset$ and CDF selections as a function of top quark mass. The uncertainties on these yields are of order 25-30%. The central value for the theoretical prediction for the $t\bar{t}$ cross section is assumed.

contamination. A stiffer $\not\!\!E_T$ cut requiring at least 50 GeV of missing transverse energy is imposed on those events that have $\not\!\!E_T$-jet opening angles less than 20°. The same region is occupied preferentially by backgrounds from $Z \rightarrow \tau^+\tau^-$ in the $e^+\mu^-$ sample so it is also removed.

This leaves a total of 7 candidate CDF events, 5 in the $e^+\mu^-$ channel and two in the $\mu^+\mu^-$ channel. No dielectron events survive the selection. One of the $\mu^+\mu^-$ events has an energetic photon candidate with a $\mu^+\mu^-\gamma$ invariant mass consistent with that of a Z° boson. Although the expected background from radiative $Z_\circ$ decay is only 0.04 events, the $\mu^+\mu^-\gamma$ candidate is removed from the sample in order to be conservative.

The D$\emptyset$ analysis requires two high P_T leptons; both leptons are required to have $P_T > 20$ GeV/c in the e^+e^- channel, $P_T > 15$ GeV/c in the $\mu^+\mu^-$ channel, and $P_T > 15(12)$ GeV/c for the electron (muon) in the $e^+\mu^-$ channel. A $\not\!\!E_T$ cut requiring at least 20 GeV and 25 GeV is placed on the $e^+\mu^-$ and e^+e^- channels, respectively (no $\not\!\!E_T$ requirement is placed on $\mu^+\mu^-$ candidate events). The selection requires at least two jets with corrected transverse energy > 15 with $|\eta| < 2.5$. Finally, e^+e^- and $e^+\mu^-$ candidate events are required to have $H_T > 120$ GeV and $\mu^+\mu^-$ events are required to have $H_T > 100$ GeV.

This leaves a total of 3 dilepton candidate events in the D$\emptyset$ dataset. There are 2 $e^+\mu^-$ events, no e^+e^- events, and 1 $\mu^+\mu^-$ event. The integrated luminosities corresponding to these three channels is 47.9 ± 5.7, 55.7 ± 6.7 and 44.2 ± 5.3 pb^{-1}, respectively. The number of observed events expected from $t\bar{t}$ production is shown in Table 2.

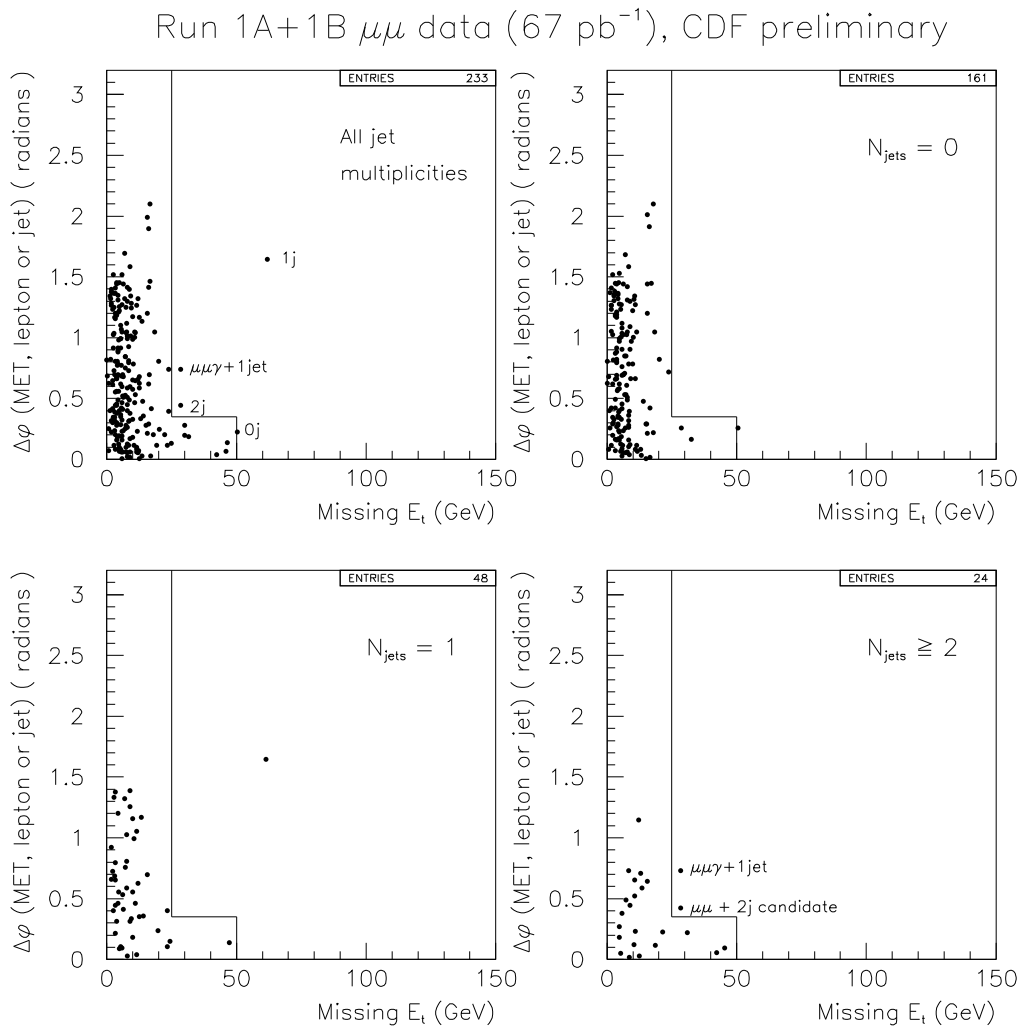


Figure 6: The distribution of the azimuthal opening angle between the missing E_T vector and the highest energy jet or lepton versus the $\not\!\!E_T$ is shown for all CDF candidate events, and for events with 0, 1 and ≥ 2 jets in the $\mu^+\mu^-$ channel. The boundary shows the cuts placed to reject the remaining Drell-Yan background.

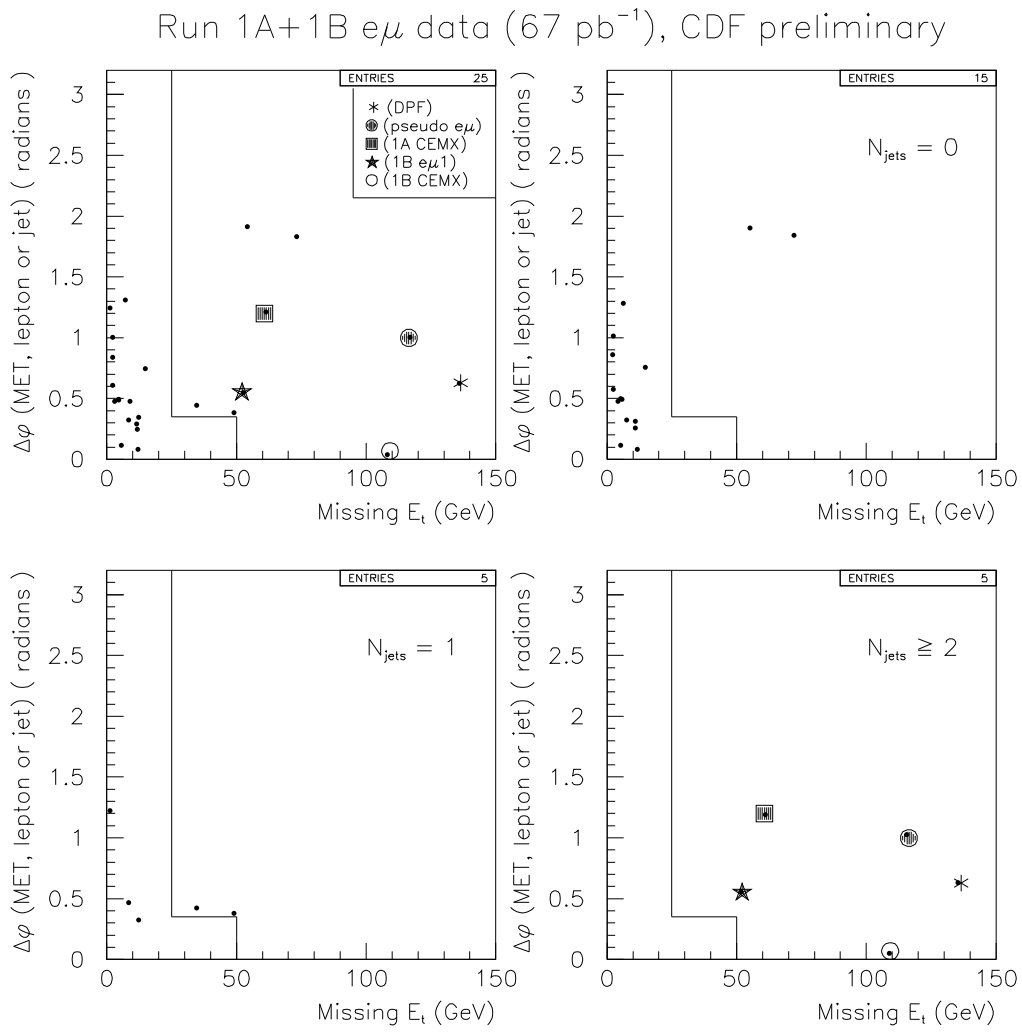


Figure 7: The distribution of the azimuthal opening angle between the missing E_T vector and the highest energy jet or lepton versus the events $\not\!\!E_T$ is shown for all CDF candidate events, and for events with 0, 1 and ≥ 2 jets in the $e^+\mu^-$ channel. The boundary shows the cuts placed to reject the $Z \to \tau^+\tau^-$ background.

6.2 Dilepton Backgrounds

The number of dilepton events observed by CDF and DØ is consistent with the rate expected from $t\bar{t}$ production for a top quark mass of order 140 to 150 GeV/c^2. It is necessary to accurately estimate the number of events expected from standard model background processes in order to interpret these event rates.

The most serious potential background comes from Z° boson production followed by the decay $Z^\circ \to \tau^+\tau^-$. The τ^+ leptons then decay leptonically leaving the dilepton signature and missing energy from the four neutrinos. The rate of this background surviving the selection criteria can be accurately estimated using the observed Z° boson kinematics in the dielectron and dimuon channels and simulating the decay of the τ^+ leptons. Other standard model sources of dileptons are divector boson production, $b\bar{b}$ and $c\bar{c}$ production and Drell-Yan production. Most of these are either very small (e.g., the backgrounds from W^+W^- and W^+Z° production) or can be estimated reliably from collider data (e.g. heavy quark production). Jets misidentified as leptons are a background source that also can be accurately estimated using the data. CDF uses the strong correlation between fake lepton candidates and the larger energy flow in proximity to the candidate. DØ employs similar techniques to estimate this background.

The estimated background rates in the three channels are listed in Table 3 and total to 1.3 ± 0.3 and 0.65 ± 0.15 for the CDF and DØ analyses, respectively. In both cases, there is an excess of observed candidate events above the expected backgrounds.

The significance of this observation can be quantified in a number of ways. One method is to ask how likely this observation is in the absence of $t\bar{t}$ production (the null hypothesis). The answer to this is an exercise in classical statistics,[26] where one convolutes the Poisson distribution of expected background events with the uncertainty in this expected rate. The significance of the CDF observation is then 3×10^{-3}; the significance of the DØ observation is 3×10^{-2}.

In themselves, each analysis cannot rule out the possibility that the observed events may be due to background sources. Taken together, however, they make the background-only hypothesis very unlikely.[†] The obvious next step is to seek

[†]One cannot simply multiply the two significances together. To combine these observations, one could define a single statistic (like the total number of observed events in both experiments) and then model the fluctuations of this variable in the case of the null hypothesis. This would give a larger probability of a background hypothesis than the product of the two probabilities.

Background	CDF	DØ
$Z \to \tau^+\tau^-$	0.38 ± 0.07	0.16 ± 0.09
Drell Yan	0.44 ± 0.28	0.26 ± 0.06
Fake $e^\pm$ or $\mu^\pm$	0.23 ± 0.15	0.16 ± 0.08
$W^+W^-/W^\pm Z^\circ$	0.38 ± 0.07	0.04 ± 0.03
Heavy quarks	0.03 ± 0.02	0.03 ± 0.03
Total	1.3 ± 0.3	0.65 ± 0.15

Table 3: The number of background events expected to survive the CDF and DØ dilepton analyses. Only the W^+W^- and heavy quark rates are estimated based on Monte Carlo calculations in the CDF analysis. The other estimates are derived from background rates obtained directly from data studies.

independent confirmation.

6.3 B Tagging in the Dilepton Sample

If the dilepton sample has a contribution from $t\bar{t}$ production, it is reasonable to search for evidence that two b quarks are being produced in association with the dilepton pair and neutrinos.

The CDF collaboration has examined these events for such indications using the b tagging algorithms described in detail in the following section. Three of the six events have a total of five tagged jets, three with SLT tags and two with SVX tags. CDF estimates that only 0.5 events with tags would be expected from non-$t\bar{t}$ standard model sources, whereas one would expect 3.6 tags if the events arose from the expected mixture of background and $t\bar{t}$ production. The data are certainly consistent with the $t\bar{t}$ hypothesis, and motivate a detailed study of the other potential channels.

7 The Lepton+Jets Top Quark Search

Both collaborations begin their lepton+jets analysis from a data sample dominated by inclusive W^+ production. They require events with significant $\not\!\!E_T$ and a well-identified, high transverse momentum electron or muon. DØ requires the

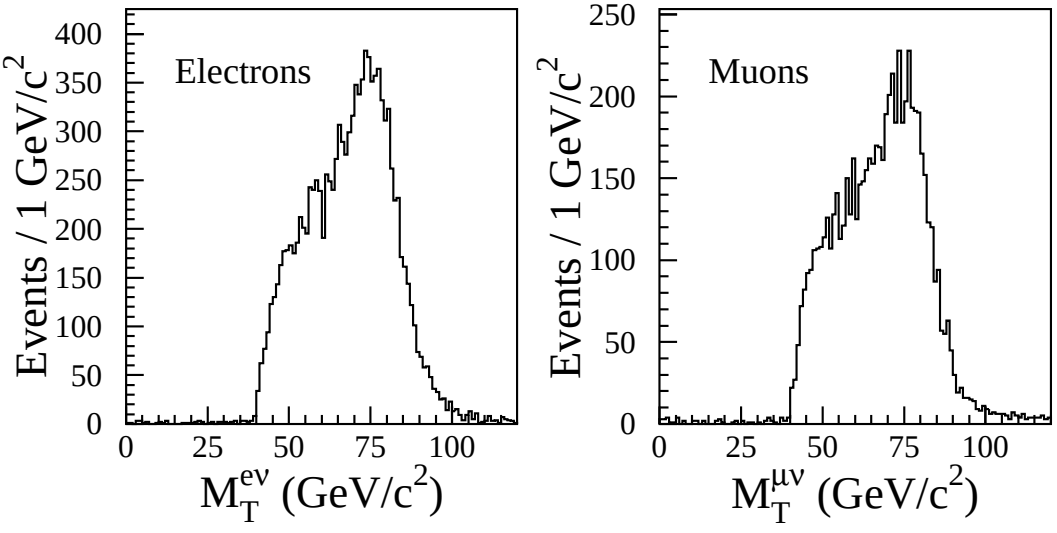


Figure 8: The transverse mass distribution for the CDF electron and muon samples after requiring a well-identified charged lepton and missing transverse energy > 20 GeV. These data are from Run IA only.

presence of an isolated electron with $E_T > 20$ GeV, and $\not\!\!E_T > 25$ GeV to identify an inclusive $W^+ \to e^+\nu_e$ sample, and an isolated muon with $p_T > 15$ GeV/c and $\not\!\!E_T > 20$ GeV to identify a $W^+ \to \mu^+\nu_\mu$ sample. CDF requires a candidate event to have $\not\!\!E_T > 20$ GeV and a charged lepton candidate in the central detector with $P_T > 20$ GeV/c and $|\eta| < 1.0$. The transverse mass for the resulting candidate events, defined as

$$M_T \equiv \sqrt{2E_T \not\!\!E_T(1 - \cos\phi_{l\nu})}, \tag{14}$$

where $\phi_{l\nu}$ is the azimuthal opening angle between the charged lepton and the $\not\!\!E_T$ vector, has a distribution with a clear Jacobian peak, as illustrated by the CDF data shown in Fig. 8.

7.1 The DØ Lepton+Jets Search

7.1.1 The DØ Kinematic Analysis

The production of W^+ bosons accompanied by additional jets form the largest single background in the lepton+jets search. However, there are significant differences in the kinematics of the partons in the $t\bar{t}$ and W+jets final state that can be used to differentiate between these processes. For example, the H_T distribution

is compared for the $t\bar{t}$ and W+jets final state in Fig. 5(b). One sees that this variable provides significant separation between signal and background with only a modest loss of signal.

The DØ collaboration defines a $t\bar{t}$ candidate sample by requiring that $H_T >$ 200 GeV, that there be at least four jets in the final state with $E_T > 15$ GeV and $|\eta| < 2.0$, and that the aplanarity of the event $\mathcal{A} > 0.05$. This leaves 5 e^++ jet events and 3 μ^++ jet events in the sample. They expect to observe 3.8 ± 0.6 events from $t\bar{t}$ production in this sample for a top quark mass of 180 GeV/c^2.

The backgrounds to $t\bar{t}$ production in this sample are dominated by the inclusive W+jets process. In order to estimate the size of this background, one can use the rate of observed events in the $W + 1$, $W + 2$, and $W + 3$ jet sample and extrapolate that to the number of events in the $W+ \geq 4$ jet sample. It is expected that the ratio of $W + n$ jet events to $W + (n - 1)$ jet events will be constant given the same jet requirements[19] when the H_T and aplanarity cuts are removed. This prediction can be tested using the $W + 1$ jets, $W + 2$ and $W + 3$ jet samples where one expects to see little $t\bar{t}$ contribution. The results of this test, shown in Fig. 9, confirm that this ratio remains constant.

The DØ collaboration then applies the H_T and aplanarity cuts and uses the relative efficiency of these cuts on $t\bar{t}$ signal and the W+jets background to extract the number of $t\bar{t}$ events in the sample and the number of background events that remain. The DØ collaboration estimates the size of the background in their $W + 4$ jet sample to be 1.9 ± 0.5 events. There is a clear excess of observed events above the predicted background.

7.1.2 B Tagging in the DØ Sample

DØ has performed a separate analysis requiring that one of the jets also be consistent with a b quark semileptonic decay. This study is complementary to the DØ kinematical analysis, and does not depend on the jet-scaling arguments to estimate the backgrounds.

DØ's excellent muon identification capability makes it possible to tag b hadrons by searching for the decay $b \rightarrow \mu^- \bar{\nu}_\mu X$. Because there are two b jets in each $t\bar{t}$ signal event, the fraction of tagged events will be twice the semileptonic branching fraction of b hadrons times the efficiency for identifying muons. DØ studies show that the use of standard muon identification requirements applied to candidates with $P_T >$ 4 GeV/c result in a tagging efficiency for $W+ \geq 3$ jet events of $\sim 20\%$.

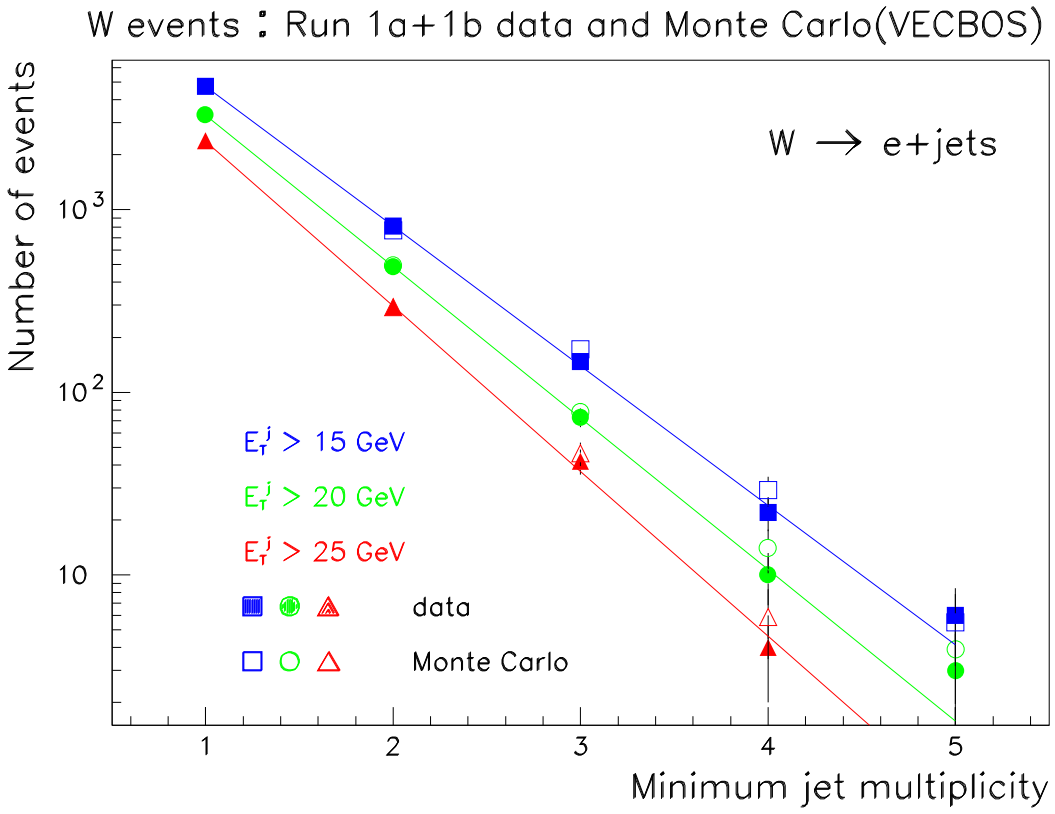


Figure 9: The rate of $W^+ \to e^+\nu_e$ events as a function of the minimum jet multiplicity and jet E_T requirements observed by the DØ collaboration (the charged conjugate mode is implied). These data are shown before the H_T or aplanarity cuts, and are compared to predictions from a QCD Monte Carlo calculation.

This is relatively insensitive to the actual top quark mass, rising slowly as a function of M_{top}.

"Fake" tags are expected to arise from real muons resulting from heavy quark (b, c) semileptonic decay and decays-in-flight of π and K mesons. This would imply that the fake rate per jet should be relatively independent of the number of jets in a given event, or the topology of the jets in the event. The DØ collaboration has measured the expected background rate for their tagging scheme using a large sample of events coming from their inclusive jet triggers. Since the jets in these events are expected to arise predominantly from light quarks and gluons, they form a good sample to estimate the probability of incorrectly b tagging a light quark or gluon jet. This leads to an over-estimate of the background from light quark jets, as some of the jets in this inclusive jet control sample will have c and b quarks

in them, albeit at a low rate. These studies show that the tag rate is between 0.005 and 0.010 per jet, and rises slowly with the E_T of the jet. Detailed Monte Carlo calculations using a full detector simulation verify this result. Based on this study, DØ expects that $\sim 2\%$ of the $W + 3$ and $W + 4$ jet background events will be tagged. With this fake rate, b tagging provides an order of magnitude improvement in signal-to-noise in this sample.

The DØ collaboration use a less stringent W+jets selection when also requiring a b quark tag in order to optimise the signal-to-noise of this analysis. The events are required to have $H_T > 140$ GeV, and the jet multiplicity requirement is relaxed to demand at least three jets with $E_T > 20$ GeV. In addition, the aplanarity cut is dropped altogether, and in the case of the electron + jets channel, the $\not\!\!E_T$ cut is relaxed to require $\not\!\!E_T > 20$ GeV. There are 3 events in the e+jet and μ+jet channels that survive these requirements, whereas only 0.85 ± 0.14 and 0.36 ± 0.08 events are expected from background sources, respectively. As in the dilepton and lepton + jets channels, a excess of candidate events over background is observed.

7.2 The CDF Counting Experiment

The CDF collaboration has performed an analysis of their lepton+jets data similar to that reported for the Run IA dataset.[11] The analysis avoids making stringent kinematical cuts that could result in large systematic uncertainties, and takes advantage of the presence of two b quarks in the signal events to control the expected backgrounds.

Starting from the inclusive W^+ boson sample, the CDF analysis requires at least three jets with $E_T > 15$ GeV and $|\eta| < 2.0$. This results in 203 events, with 164 and 39 events in the $W + 3$ and $W + \geq 4$ jet samples, respectively. The backgrounds estimated to make the largest contribution to this sample come from real W^+ boson production, from standard model sources of other isolated high E_T leptons (such as Z° boson production), from b and c quark semileptonic decays and from events where the lepton candidate has been misidentified. Most of the non-W^+ boson backgrounds have lower $\not\!\!E_T$, and are characterised by lepton candidates that are not well isolated from other particles in the event. The correlation between this additional energy flow and $\not\!\!E_T$ in the event allows one to directly measure this background fraction. This results in an estimate for the background from sources of non-isolated lepton candidates of $10 \pm 5\%$. The background rates

Background	Fraction of Sample (%)
WW, WZ Production	5.0 ± 2.3
$Z^\circ \to e^+e^-/\mu^+\mu^-$	5.2 ± 1.3
$Z^\circ \to \tau^+\tau^-$	3.3 ± 1.0
Fake Leptons, Conversions, $b\bar{b}$	10.0 ± 5.0
Total	23.5 ± 5.7

Table 4: The estimated fractions of events in the $W+ \geq 3$ jet sample arising from the different background sources to $t\,\overline{t}$ production. Only the requirement of at least three jets has been imposed.

from sources that produce isolated lepton candidates have been estimated using data and Monte Carlo calculations. These background estimates are summarised in Table 4.

7.2.1 Secondary Vertex Tagging

The CDF detector has the unique capability of detecting b quarks by reconstructing the location of the b quark's decay vertex using the SVX detector. A schematic of the decay topology for a bottom hadron is shown in Fig. 10. The charged particle trajectories are reconstructed in the CTC and then extrapolated into the SVX detector to identify the track's hits in the silicon strip detector.

The quality of the reconstructed SVX track is determined by the number of SVX coordinates found for the track and the accuracy of each coordinate. The algorithm to reconstruct secondary vertices considers all tracks above a transverse momentum of 1.5 GeV/c that have an impact parameter relative to the primary vertex $> 2\sigma$, where σ is the estimated uncertainty in the impact parameter measurement for the track. The algorithm first looks for vertices formed by three tracks, making relatively loose quality cuts on each of the tracks. A vertex is accepted if a χ^2 fit requiring the three tracks to come from a common point is acceptable. Any remaining high-quality tracks with large impact parameter are then paired up to look for two-track vertices. A jet containing a secondary vertex found in this way that has a positive decay length is considered SVX tagged (the sign of the decay length is taken from the dot product of the displacement vector between the primary and secondary vertices, shown as L_{xy} in Fig. 10, and the

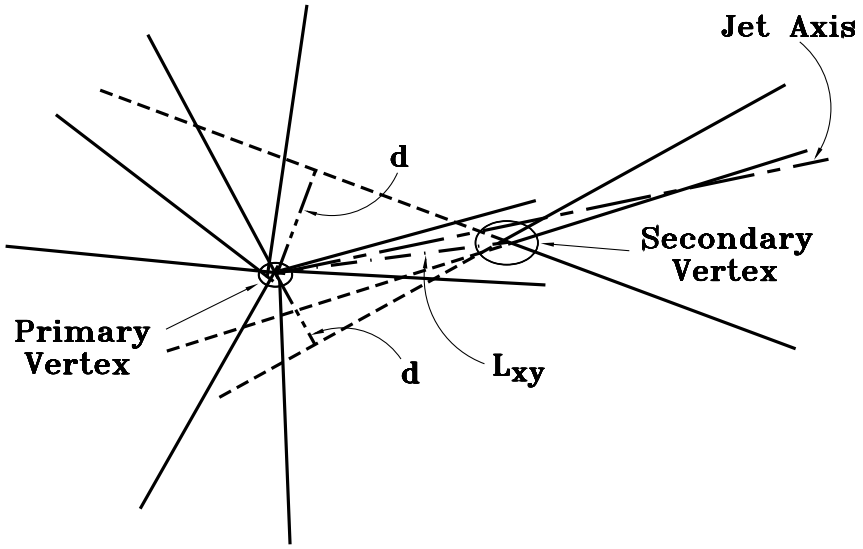


Figure 10: A schematic of the decay of a bottom quark, showing the primary and secondary vertices, and the charged tracks reconstructed in the CDF CTC and SVX detectors.

vector sum of the momenta of the daughter tracks).

The efficiency of this SVX tagging algorithm has been measured using a large sample of inclusive electron and $J/\psi \rightarrow \mu^+\mu^-$ candidates, where the heavy quark contents in these samples have been independently estimated. This efficiency agrees with that obtained using a full detector simulation; the ratio of the measured efficiency to the efficiency determined using the detector simulation is 0.96 ± 0.07.

The b quark SVX tags not arising from $t\bar{t}$ production come from track combinations that for some reason result in a fake secondary vertex (mistags) and from real sources of b and c quarks in $W+$ jet events. One way of estimating the mistag rate is to note that the rate of these fakes must be equal for those secondary vertices located on either side of the $p\bar{p}$ collision vertex as determined by comparing the displacement vector of the secondary vertex with the momentum vector of the tracks defining the secondary vertex (positive and negative tags, respectively). The rate of real b and c quarks not arising from $t\bar{t}$ production can be estimated using theoretical calculations and comparing these with observed rates in other channels.

The mistag probability has been measured using both samples of inclusive jets

and the inclusive electron and dimuon samples. The probability of mistagging as a function of the number of jets in the event and the transverse energy of the jet is shown in Fig. 11, based on the inclusive jet measurements where I have plotted both the negative and positive tag rates. The negative tag rate is perhaps the best estimate of the mistag rate, since one expects some number of real heavy quark decays in this sample to enhance the positive tag rate. The mistag rate per jet measured in this way is ~ 0.008, and is lower than the positive tag rate measured in the inclusive jet sample (~ 0.025), as expected from estimates of heavy quark production in the inclusive jet sample.

To account for all sources of background tags, the number of tagged events expected from sources of real heavy quark decays (primarily $W^+b\bar{b}$ and $W^+c\bar{c}$ final states) is determined using a Monte Carlo calculation and a full simulation of the detector. The sum of this "physics" tag rate and the mistag rate then gives an estimate of the total background to $t\bar{t}$ production. This estimate can be checked by using the positive tag rate in inclusive jet events as a measure of the total non-$t\bar{t}$ tag rate in the $W+$ jet events. This gives us a somewhat higher background rate, due primarily to the expected larger fraction of b and c quarks in the inclusive jet sample compared to the $W+$jet events.

The efficiency for finding at least one jet with an SVX tag in a $t\bar{t}$ signal event is calculated using the ISAJET Monte Carlo programme[27] to generate a $t\bar{t}$ event, and then applying the measured tagging efficiencies as a function of jet E_T to determine how many reconstructed b quark jets are tagged. The SVX tagging efficiency, *i.e.* the fraction of $W+ \geq 3$ jet $t\bar{t}$ events with at least one SVX-tagged jet, is found to be 0.42 ± 0.05, making this technique a powerful way of identifying $t\bar{t}$ candidate events.

7.2.2 Soft Lepton Tagging

The CDF collaboration developed the original lepton-tagging techniques to search for b quarks in $t\bar{t}$ production,[28] requiring the presence of a muon candidate in proximity to one of the jets. The collaboration has enhanced these techniques by extending the acceptance of the muon system and by searching for electron candidates associated with a jet cluster. In both cases, it is optimal to allow for relatively low energy leptons (down to P_T's as low as 2 GeV/c), so this technique has become known as "soft lepton tagging." A candidate jet cluster with a soft lepton candidate is considered to be SLT tagged.

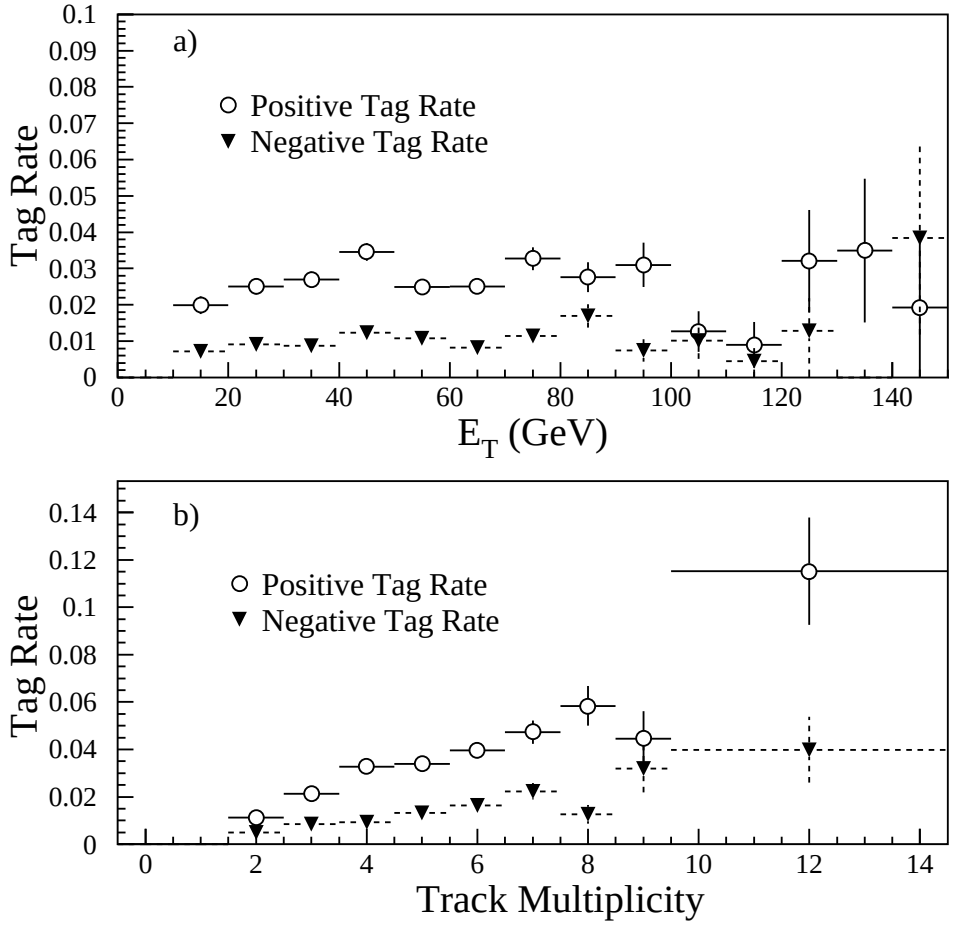


Figure 11: The rate of SVX tags as a function of the transverse energy of the jet and the charged track multiplicity in the jet, as measured using the inclusive jet sample. Tag rates for both positive and negative decay length vertices are shown.

The efficiency of this tagging technique depends on the ability to identify leptons in the presence of additional hadrons that come from the fragmentation of the b quark and the decay of the resulting c quark system. Muons are identified by requiring a charged track in the CTC that matches a muon track stub. Electron candidates are defined by an electromagnetic shower in the calorimeter with less than 10% additional energy in the hadronic calorimeter towers directly behind the shower, a well-reconstructed track in the CTC that matches the position of the shower and shower profiles consistent with those created by an electron. The overall efficiency for finding at least one SLT tag in a $t\bar{t}$ event is 0.22 ± 0.02, and is not a strong function of the top quark mass.

The rate at which this algorithm misidentifies light quark or gluon jets as

having a soft lepton is determined empirically by studying events collected by requiring the presence of at least one jet cluster. The mistag rate for muon tags varies between 0.005 and 0.01 per charged track, and rises slowly with the energy of the jet. The mistag rate for electrons also depends on the track momentum and how well isolated it is from other charged tracks; it typically is of order 0.005 per track. Fake SLT tags where there is no heavy flavour semileptonic decay is expected to be the dominant source of background tags in the $t\bar{t}$ sample, due to the larger SLT fake rates as compared to the SVX mistag rates.

7.2.3 Tagging Results in the CDF Lepton+Jets Sample

The SVX and SLT tagging techniques have been applied to the W+jet sample as a function of the number of jets in the event, and the expected number of mistags has been calculated for each sample. This provides a very strong consistency check, as the number of observed tags in the $W + 1$ jet and $W + 2$ jet samples should be dominated by background tags; the fraction in these two event classes expected from $t\bar{t}$ production is less than 10% of the total number of candidate events.

The number of candidate events and tags is shown in Table 5. There is good agreement between the expected number of background tags and the number of observed tags for the $W + 1$ jet and $W + 2$ jet samples. However, there is a clear excess of tags observed in the $W+ \geq 3$ jet sample, where we observe 27 and 23 SVX and SLT events, respectively, and expect only 6.7 ± 2.1 and 15.4 ± 2.3 SVX and SLT background tags. The excess of SVX tags is particularly significant, with the probability of at least this number of tags arising from background sources being 2×10^{-5}. The excess of SLT tags is less significant because of the larger expected background. The probability that at least 23 observed SLT tags would arise from background only is 6×10^{-2} and confirms the SVX observation.

It is interesting to note that if we attribute the excess number of SVX tags in the $W+ \geq 3$ jet sample to $t\bar{t}$ production, we would expect approximately 10 $W + 2$ jet tagged events resulting from $t\bar{t}$ production. This is in good agreement with the excess of observed tags (13 ± 7) in this sample, and corroborates the hypothesis that the excess in the $W+ \geq 3$ jet sample is due to the $t\bar{t}$ process.

A striking feature of the tagged sample is the number of events with two or more tagged jets. The 27 SVX tags are found in 21 events, so that there are 6 SVX double tags. There are also six SVX tagged events that have SLT tags.

Sample	SVX bkg	SVX tags	SLT bkg	SLT tags
W+1 jet	50 ± 12	40	159 ± 25	163
W+2 jet	21 ± 7	34	46 ± 7	55
W+$\geq$ 3 jet	6.7 ± 2.1	27	15.4 ± 2.3	23

Table 5: The expected number of background tags and the observed number of tags in the CDF lepton+jets sample as a function of the number of jets in event.

We would expect less than one SVX-SVX double tag and one SVX-SLT double tag in the absence of $t\,\overline{t}$ production, whereas we would expect four events in each category using the excess of SVX tags to estimate the $t\,\overline{t}$ production cross section. A schematic of one of the SVX double tagged events is shown in Fig. 12, where the tracks reconstructed in the SVX detector are displayed, along with the jets and lepton candidates they are associated with. These observations strengthen the $t\,\overline{t}$ interpretation of the CDF sample.

7.3 Summary of Counting Experiments

The results of the lepton+jets counting experiments performed by DØ and CDF are summarised in Table 6. Both collaborations observe an excess of events in all the channels in which one can reasonably expect evidence for the top quark. Many of the channels demonstrate correlated production of W^+ bosons with b quarks – exactly what we would expect from $t\,\overline{t}$ decay.

Taken together, this is overwhelming evidence that the two collaborations are observing phenomena that within the context of the standard model can only be attributed to pair production of top quarks.

8 Measurement of Top Quark Properties

In order to further test the interpretation that top quark production is responsible for the excess in the dilepton and lepton+jets channels, both collaborations have measured the rate of top quark production and identified a subset of their candidate lepton+jet events where it is possible to directly measure the mass of the top quark.

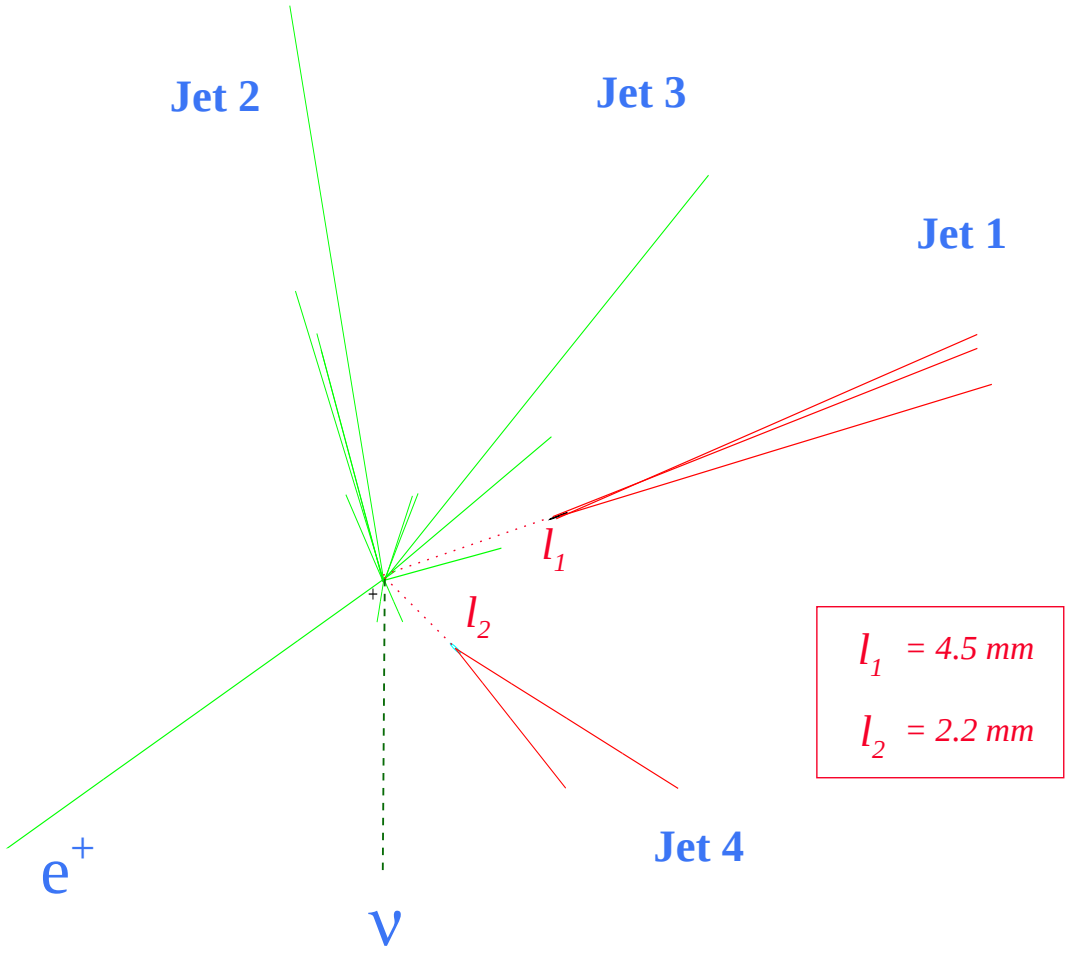


Figure 12: The schematic in the r-ϕ view of the SVX tracks reconstructed in one of the CDF lepton+jet events that has two SVX tagged b jets. The jets associated with the SVX tracks and the lepton candidates are identified. The decay lengths of each b candidate jet are noted in the figure. This event is fitted to a top quark mass of 170 ± 10 GeV/c^2, using the procedure discussed below.

Sample	Background	Observed
CDF Dileptons	1.3 ± 0.3	6
DØ Dileptons	0.65 ± 0.15	3
Lepton + Jets (DØ Kinematics)	0.93 ± 0.50	8
Lepton + Jets (DØ B Tagging)	1.21 ± 0.26	6
Lepton + Jets (CDF SVX tags)	6.7 ± 2.1	27
Lepton + Jets (CDF SLT tags)	15.4 ± 2.3	23

Table 6: The expected number of background events and the observed number of events in the different analyses. Note that some event samples and background uncertainties are correlated so it is not straightforward to combine these observations into a single statement of statistical significance.

These measurements allow us to test the standard model prediction for the cross section as a function of the top quark mass. The initial evidence for top quark production published by CDF[11] implied a top quark production cross section almost two standard deviations above the theoretically predicted value. Moreover, other standard model measurements, and in particular those performed at LEP, constrain the top quark mass. It is important to directly verify that these predictions agree with the top quark mass inferred from the Collider data.

The CDF and DØ Collaborations have also begun other studies of top quark properties that can be inferred from the Collider data. These include aspects of both top quark decay and production, and I discuss their status in the following subsections.

8.1 The Top Quark Cross Section

The acceptance of the DØ and CDF top quark searches depend on the top quark mass. We can therefore infer the $t\bar{t}$ production cross section as a function of the top quark mass given the number of observed events in each channel.

For a data sample with integrated luminosity $\mathcal{L}$, if we observe N_i^o candidate events in a particular channel i and we expect N_i^b background events, then the maximum likelihood solution for the cross section of the process combining all

Source	Uncertainty (%)
Lepton ID and Trigger	10
Initial State Radiation	7
Jet Energy Scale	6.5
b Tagging Efficiency	12

Table 7: The uncertainties in the acceptance calculation for the CDF cross section measurement using the SVX tagged sample.

channels is

$$\sigma = \frac{\sum_i \left(N_i^o - N_i^b\right)}{\mathcal{L}\left(\sum_i \epsilon_i\right)}, \tag{15}$$

where ϵ_i is the acceptance for the search. This assumes that the observed number of events has a Poisson distribution and that uncertainties on the acceptance can be ignored. The latter restriction can be relaxed by numerically solving for the maximum likelihood solution allowing for uncertainties in ϵ_i and N_i^b, and any correlations in the acceptances.

The CDF collaboration has performed a preliminary measurement of the $t\,\overline{t}$ cross section using the SVX tagged sample. This is the single most significant measurement and can be performed only knowing the SVX tagging efficiency and background rates. The addition of the SLT sample and the dileptons into the cross section measurement requires a knowledge of the efficiency correlations in the samples and is work in progress. The $t\,\overline{t}$ acceptance was determined using the ISAJET Monte Carlo programme, and found to be 0.034 ± 0.009. The uncertainties associated with this acceptance calculation are listed in Table 7. The expected background in the 21 tagged events is $N^b = 5.5 \pm 1.8$ events.[‡]

The resulting cross section determined from the SVX sample is $6.8^{+3.6}_{-2.4}$ pb for a nominal top quark mass of 175 GeV/c^2. This is approximately one standard deviation lower than the cross section determined in the Run IA CDF data. It is in good agreement with the theoretically predicted value of 4.9 ± 0.6 pb for the same top quark mass.

[‡]The previous estimate of the expected SVX background tags assumed that there was no contribution from $t\,\overline{t}$ production to the 203 events in the $W+ \geq 3$ jet sample prior to tagging.

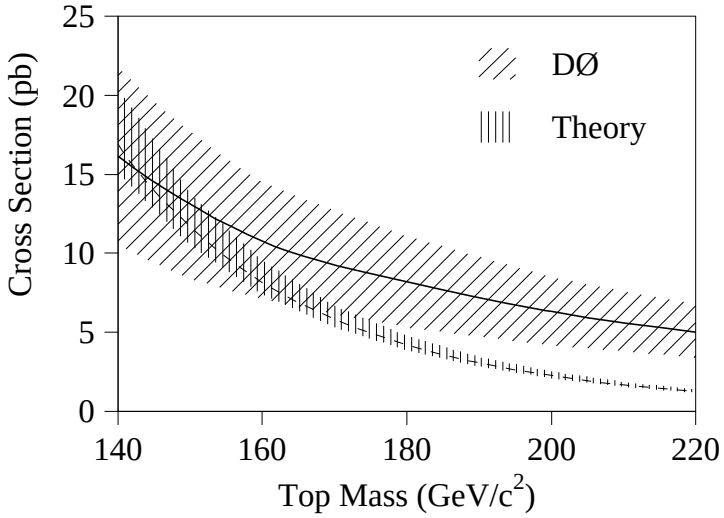


Figure 13: The top quark cross section determined by the DØ collaboration as a function of top quark mass. The QCD prediction for $t\,\overline{t}$ production is displayed as the heavier band.

The DØ collaboration estimates the $t\,\overline{t}$ cross section using the information from all the channels they have studied. They also perform a background subtraction and then correct for the acceptance, channel by channel. They determine $\sigma_{t\,\overline{t}} = 6.2 \pm 2.2$ pb, for a top quark mass of 200 GeV/c^2. This value doubles to ~ 12 pb if one assumes a top quark mass of 160 GeV/c^2. The top quark mass dependence of the DØ cross section is illustrated in Fig. 13.

The CDF and DØ estimates are in reasonable agreement with each other, although both have large uncertainties. A strong test of the lowest order calculation for $\sigma_{t\overline{t}}$ and next-to-leading order corrections will have to wait for substantially more statistics.

8.2 The Top Quark Mass

The top quark mass can be determined directly by correlating the kinematics of the observed partons in the final state. The sensitivity of this measurement depends on the amount of "missing" information in the events, and the inherent resolution of the detectors to jets and missing energy. The lepton + ≥ 4 jet events offer the possibility of fully reconstructing the $t\,\overline{t}$ system provided one assumes that the missing transverse energy arises from the undetected neutrino, and that four

of the jets come from the b and $\bar{b}$ quarks and the two quarks from the W^+ decay.

Perhaps the most serious complication to this procedure is the difficulty of associating final state jet clusters with the partons from the $t\,\overline{t}$ decay. The jets are only approximate measures of the initial state parton, and there is often not a 1-to-1 correspondence between partons resulting from the $t\,\overline{t}$ decay and observed jets. This is due to gluon radiation that can cause one parton to be observed as two jet clusters, and overlap of jet clusters, where two partons merge into a single jet cluster. To complicate matters further, additional partons are produced by initial and final state radiation, so the number of observed jet clusters may readily exceed four.

The number of combinatorial possibilities for assigning partons to jets in the case where only four jets are observed is twelve (we only have to identify the two jets associated with the W^+ decay and not have to permute these two). If we can identify one of the jets as arising from a bottom quark, the number of possible assignments reduces to six. Any technique that reconstructs the $t\,\overline{t}$ decay in this mode has to reduce the effect of these combinatorial backgrounds on the expected signal.

8.2.1 CDF Mass Analysis

The CDF collaboration measures the top quark mass by selecting a sample of lepton+jet events with at least four jets, and then making the parton-jet assignment that best satisfies a constrained kinematic fit. The fit inputs are the observed jet momentum vectors, the momentum vector for the charged lepton, the transverse energy vector for the neutrino and the vector sum of the momentum of the unassigned jets in the event. The uncertainties in these quantities are determined from the measured response of the detector. The fit assumes that the event arises from the process

$$\begin{aligned} p\bar{p} \;\rightarrow\; & t\bar{t}X, \\ & \quad \hookrightarrow q\bar{q}'\,\bar{b} \\ & \quad \hookrightarrow l^+\nu_l\, b \end{aligned} \tag{16}$$

The fit constrains the W^+ and W^- decay daughters to have an invariant mass equal to the W^+ mass and constrains the t and the $\bar{t}$ to have the same mass. The unknown recoil system X is observed in the detector as unassociated jets and the

"unclustered" energy in the calorimeter, *i.e.* the energy not associated with a jet. Only the four highest E_T jets are considered, reducing the possible combinations at the cost of some degradation in top quark mass resolution (in those cases where the $t\bar{t}$ daughter jets are not the four highest E_T jets in the event).

Formally, there are two degrees of freedom in the fit when we take into account the number of constraints and the number of unmeasured quantities. A χ^2 function including the uncertainties in the measurements is minimised subject to the kinematic constraints for each possible parton-jet assignment. The b-tagged jets in the event are only allowed to be assigned to the b or $\bar{b}$ quarks. Prior to the fit all jet energies are corrected in order to account for detector inhomogeneities and the effect of energy flow into and out of the jet clustering cone. The parton assignment that produces the lowest χ^2 is selected for the subsequent analysis. The event is rejected if the minimum χ^2 is greater than 10. Parton assignments that result in a top quark mass greater than 260 GeV/c^2 are also rejected as the experiment is not expected to have any sensitivity to top quark masses of that magnitude.

Monte Carlo studies have demonstrated that this procedure identifies the correct parton-jet assignment about 40% of the time. The top quark mass resulting from the fit in those cases is shown in Fig. 14 along with the mass distribution for all lepton + ≥ 4 jet events for a sample created assuming a top quark mass of 170 GeV/c^2. From a single event, one is able to measure the top quark mass to an accuracy of $\sim$ 10 GeV/c^2 when one makes the correct assignment. However, the full distribution shows that the fitting and parton assignment procedure retains much of this mass information even in those cases where the incorrect parton assignment has been made.

Starting with the 203 $W+ \geq 3$ jet events, the CDF collaboration selects a subset of events that have at least one additional jet with $E_T > 8$ GeV and $|\eta| < 2.4$. The requirements on the fourth jet are less stringent than the first three jets in order to enhance the efficiency for detecting all four jets from the $t\bar{t}$ decay. There are 99 such events in the CDF sample prior to requiring a b-tagged jet, and 88 of these pass the χ^2 cut on the best jet-parton assignment and kinematic fit. The additional requirement of at least one SVX or SLT-tagged jet leaves 19 events.

The background of non-$t\bar{t}$ events in this sample is estimated in the same manner used in the cross section analysis. One assumes that the 88 event sample is

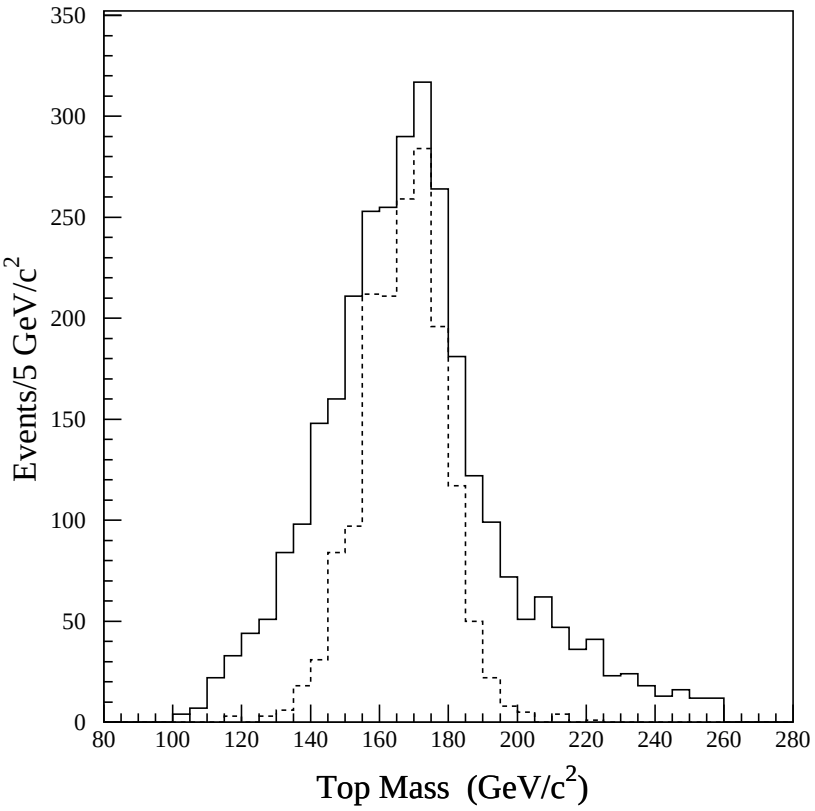


Figure 14: The fitted top quark mass in Monte Carlo events for those events in which the correct parton assignments have been made (dashed histogram) and for all events that pass the fit procedure (solid histogram). A top quark mass of 170 GeV/c^2 has been assumed.

a mixture of background and $t\bar{t}$ signal, and then applies the known background tag rates to determine how many of the non-$t\bar{t}$ events would be tagged. This results in a estimated background in the 19 events of $6.9^{+2.5}_{-1.9}$ events. This background is expected to be a combination of real W+jet events and events where an energetic hadron fakes the lepton signature. Studies of the Z+jet events, candidate events where the lepton is not well-isolated and W+jet Monte Carlo events show that the resulting top quark mass distribution for these different background events are all similar. The CDF collaboration therefore uses the W+jet Monte Carlo sample to estimate the background shape in the top quark mass distribution.

The resulting top quark mass distribution is shown in Fig. 15. One sees a clear peak around 170-180 GeV/c^2 with relatively long tails. The dotted distribu-

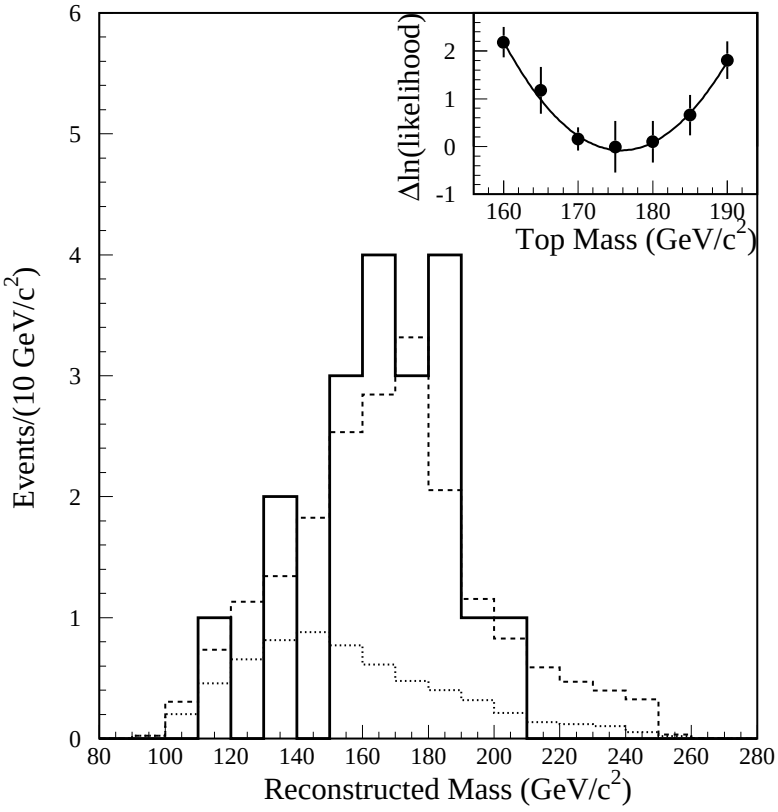


Figure 15: The fitted top quark mass for the 19 events in the CDF sample with four or more jets that satisfy the fit criteria. The dotted histogram reflects the shape and size of the estimated background. The dashed histogram is the result of a fit of the reconstructed mass distribution to a combination of $t\,\overline{t}$ signal and expected background. The inset distribution is the change in log-likelihood of this fit.

tion represents the shape of the non-$t\,\overline{t}$ backgrounds, normalised to the estimated background rate. The top quark mass is determined by performing a maximum likelihood fit of this distribution to a linear combination of the expected $t\,\overline{t}$ signal shape determined by Monte Carlo calculations for different top quark masses and the background. The background rate is constrained by the measured rate of non-$t\,\overline{t}$ events in the sample. The negative log-likelihood distribution for this fit is shown in the inset in Fig. 15. It results in a top quark mass of 176 ± 8 GeV/c^2.

Since the fit constrains the invariant mass of the jets assigned to be the W^+ boson daughters to the W^+ boson mass, one can only test the consistency of this

Source	Uncertainty (GeV/c^2)
Final State Gluon Radiation	7.7
Absolute Jet Energy Scale	3.1
Variations in Fit Procedures	2.5
Shifts Resulting from Tagging Biases	2.4
Monte Carlo Statistics	3.1
Non-$t\,\overline{t}$ Mass Distribution Shape	1.6
Miscellaneous Effects	2.0

Table 8: The systematic uncertainties associated with the CDF top quark mass measurement.

assignment by first relaxing this constraint and then examining the dijet invariant mass distribution. I show this in Fig. 16 for the $W+ \geq 4$ jet events that satisfy the selection criteria without the imposition of the dijet mass constraint. The comparison with the expected distribution from the combination of background events and $t\,\overline{t}$ signal is quite good. However, one should keep in mind the rather low statistics and the large expected mass resolution. This distribution will become a very important calibration tool when larger statistics samples become available.

The largest systematic uncertainties in this measurement arise from uncertainties in the modelling of gluon radiation in jets in the final state, absolute jet energy scale, variations in fitting procedures, and the shape of the non-$t\,\overline{t}$ background. A number of other potential sources of uncertainty have been studied, and have been found to contribute a total of ± 2.0 GeV/c^2 to the total systematic uncertainty. A summary of these uncertainties is given in Table 8, and total to ± 10 GeV/c^2.

One can quantify the significance of the shape of the mass distribution by performing an unbinned Kolmogorov-Smirnov test. The probability that the observed mass distribution could arise from purely background sources is 2×10^{-2}. This test is conservative in that it only compares the shape of the background with the observed data. Other measures of significance can be used. For example, one can define a relative likelihood for the top+background and background-only hypotheses and then ask how often a background-only hypothesis would result in a relative likelihood as significant as that observed. This test gives a probability for a background fluctuation of less than 10^{-3}. However, it is more model-dependent

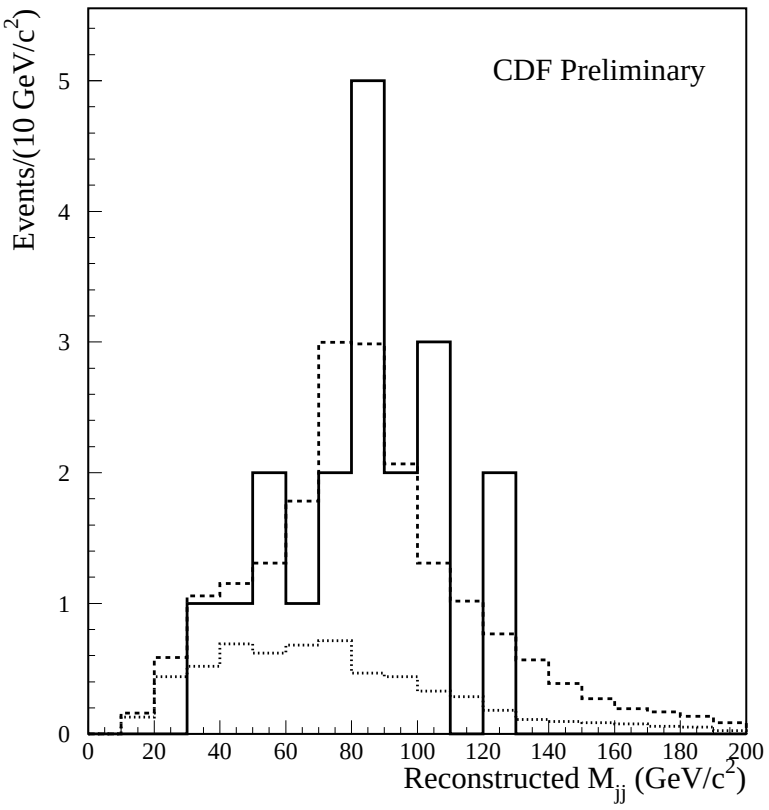


Figure 16: The solid histogram is the fitted dijet invariant mass distribution for the $W+ \geq 4$ jet events in the CDF sample that satisfy the fit criteria. In this case, the dijet invariant mass constraint has been relaxed and the lowest χ^2 solution has been plotted. The heavy dashed histogram is the expected distribution from a combination of $t\bar{t}$ signal and the non-$t\bar{t}$ background. The light dashed histogram is the background distribution normalised to the expected number of background events in this sample.

as it assumes a specific shape for the non-background hypothesis.

8.2.2 The DØ Mass Measurement

The DØ collaboration estimates the top quark mass using their sample of lepton + ≥ 4 jet events. In their analysis, they select 4-jet events by requiring that all jets have a corrected transverse energy > 15 GeV with $|\eta| < 2.4$. They also require the events to have $H_T > 200$ GeV and to have aplanarity > 0.05. They find 14 events that satisfy these requirements.

They then perform a χ^2 fit of the observed kinematics in each event to the $t\bar{t} \rightarrow W^+W^-b\bar{b}$ hypothesis, requiring that the mass of the assumed $t \rightarrow l^+\nu_l b$ system equal the mass of the $t \rightarrow q\bar{q}'b$ system making all possible parton-jet assignments in the final state. As in the CDF technique, they only consider the four highest E_T jets, and only fits with $\chi^2 < 7$ are considered acceptable. There are 11 events that have at least one configuration that gives an acceptable fit. For each event, they assign a top quark mass by averaging the top quark mass from the three best acceptable fits for that event, weighting the mass from each fit with the χ^2 probability from the fit. The resulting histogram of the invariant mass of the three-parton final state (the hypothesised top quark) is shown in Fig. 17(a). They performed the same analysis on a "looser" data sample of 27 events, where the H_T and aplanarity requirements were removed. This yields similar results, as shown in Fig. 17(b), although with significantly larger backgrounds. The mass distribution shows an enhancement at a three-parton invariant mass around 200 GeV/c^2, as expected from $t\,\overline{t}$ production (shown as the higher mass curve in both plots). The corresponding mass distribution expected from the QCD W+jet background is shown in Fig. 17(a)-(b) as the dashed curve at lower mass. It peaks at small values of three-parton invariant mass and together the combined background and signal hypothesis model the data well.

The mass distribution obtained using the looser selection is fit to a combination of $t\,\overline{t}$ signal and background, yielding a top quark mass of

$$M_{\rm top} = 199^{+19}_{-21} \pm 22 \text{ GeV/c}^2, \tag{17}$$

where the two uncertainties are statistical and systematic, respectively. A similar fit to the mass distribution using the 11 event sample results in a consistent result, but with larger statistical uncertainties. The negative log-likelihood distributions for the fits to the standard and loose selection are shown in Fig. 17(c) and (d),

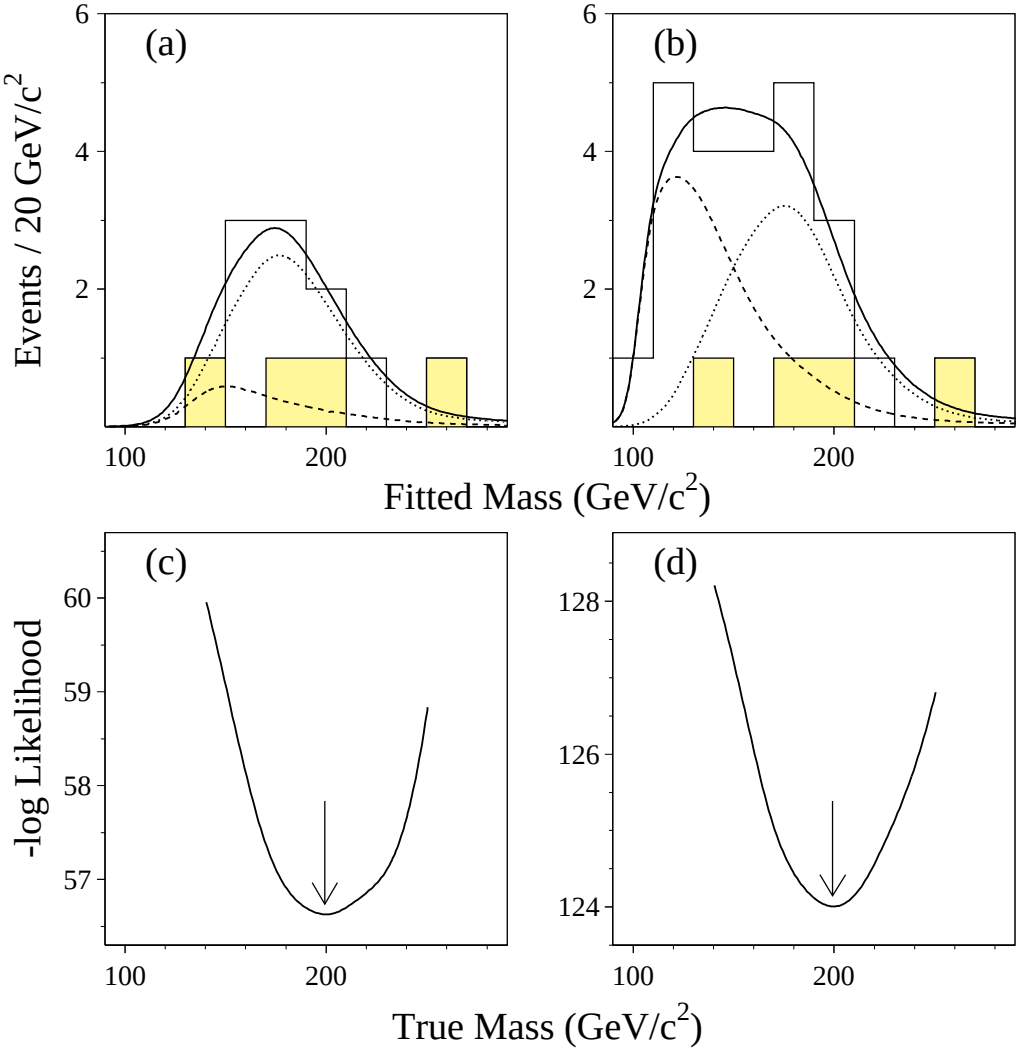


Figure 17: The distribution of the three-jet invariant mass versus the top quark mass obtained from the DØ lepton + 4 jet sample. Figures a) and b) show the results of the standard and "loose" selection, respectively. The dashed curves are the predicted background distributions, the dotted curves are the $t\bar{t}$ signal distributions and the solid curves are the sum of these. Figures c) and d) show the likelihood distribution for fits of the mass distributions to a combination of signal and background terms.

respectively. The systematic uncertainty is dominated by the sensitivity of this analysis to the DØ jet energy scale.

8.3 Top Quark Decays

The standard model predicts that the top quark will decay via a V-A interaction into the W^+b final state 100% of the time. It is important to confirm this prediction as various extensions to the standard model differ on the predicted phenomenology of top quark decays. There are effectively two separate predictions that should be tested:

1. The decay proceeds via the standard model charged current.
2. The top quark always decays to a b quark.

It is useful to address these two predictions separately as they involve different aspects of the standard model, namely the assumption that there is only one current involved in the top quark decay and on our understanding of the tW^+b vertex.

In the context of the standard model, the GIM mechanism is responsible for suppressing all flavour-changing neutral currents (FCNC). This has been experimentally tested in the strange and bottom quark sector, where limits on FCNC decays are quite stringent.[29] An extension to the top quark sector is therefore an important verification of this fundamental aspect of the electroweak interaction. The standard model does allow top quark charged current decays to either s or a d quarks, but only via the mixing of the quark mass eigenstates as parametrised by the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements V_{ts} or V_{td}. If we assume that there are only three generations and that the CKM matrix is unitary, then the 90% CL limits on these two elements are[30]

$$0.004 \leq |V_{td}| \leq 0.015 \qquad \text{and} \qquad 0.030 \leq |V_{ts}| \leq 0.048. \tag{18}$$

This predicts top quark branching fractions to s and d quarks of less than 0.3%. However, if we relax the condition of unitary and/or allow for a larger number of quark generations, then the strict limits on V_{ts} and V_{td} no longer apply, and the possibility exists for large top quark decay rates to these lighter quarks.

There are a number of standard model extensions that predict decay modes not involving a transition mediated by a W^+ boson.[31] The most obvious candidates are the flavour-changing neutral decays such as $t \to Z^\circ c$ or $t \to \gamma c$. Such models

therefore result in decays that violate both standard model predictions. There are also models that predict decay modes that always yield a b quark in the final state, but involve a transition mediated by something other than the W^+ boson. A popular example of this is the decay $t \to H^+ b$, where H^+ is a charged Higgs boson. Since the decay modes of the H^+ are in principle quite different from those of the W^+, this would result in a different rate of lepton+jet and dilepton final states coming from the $t\bar{t}$ system.

8.3.1 Top Quark Branching Fraction

The measurement of top quark branching fractions is currently limited by the rather small number of detected events, and by the large uncertainty in the top quark production cross section. The most sensitive measures of the top quark branching fraction $\mathcal{B}(t \to W^+ b)$ that do not depend on a knowledge of the $\sigma_{t\bar{t}}$ are the relative rate of single to double b quark tags in lepton+jet events, and the relative rates of zero, single and double b quark tags in dilepton events. The relative rate of zero b quark tags in lepton+jet events is not helpful in this case as this sample is contaminated with a large fraction of non-$t\bar{t}$ background.

These relative rates are sensitive to

$$\mathcal{R} \equiv \frac{BR(t \to W^+ b)}{BR(t \to W^+ q)} = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2}. \tag{19}$$

The fractions of zero, single and double tagged events can be related to $\mathcal{R}$ by the expressions

$$\begin{aligned} f_0 &= (1 - \mathcal{R}\epsilon)^2 \\ f_1 &= 2\mathcal{R}\epsilon(1 - \mathcal{R}\epsilon) \\ f_2 &= (\mathcal{R}\epsilon)^2, \end{aligned} \tag{20}$$

where ϵ is the b tagging efficiency. These can be solved for $\mathcal{R}$ to obtain the expressions

$$\mathcal{R} = \frac{2}{\epsilon(f_1/f_2 + 2)} \tag{21}$$

$$\mathcal{R} = \frac{1}{\epsilon(2f_0/f_1 + 1)}, \tag{22}$$

where the first expression is applicable to both the lepton+jets and dilepton event samples, and the second applies to the dilepton sample only.

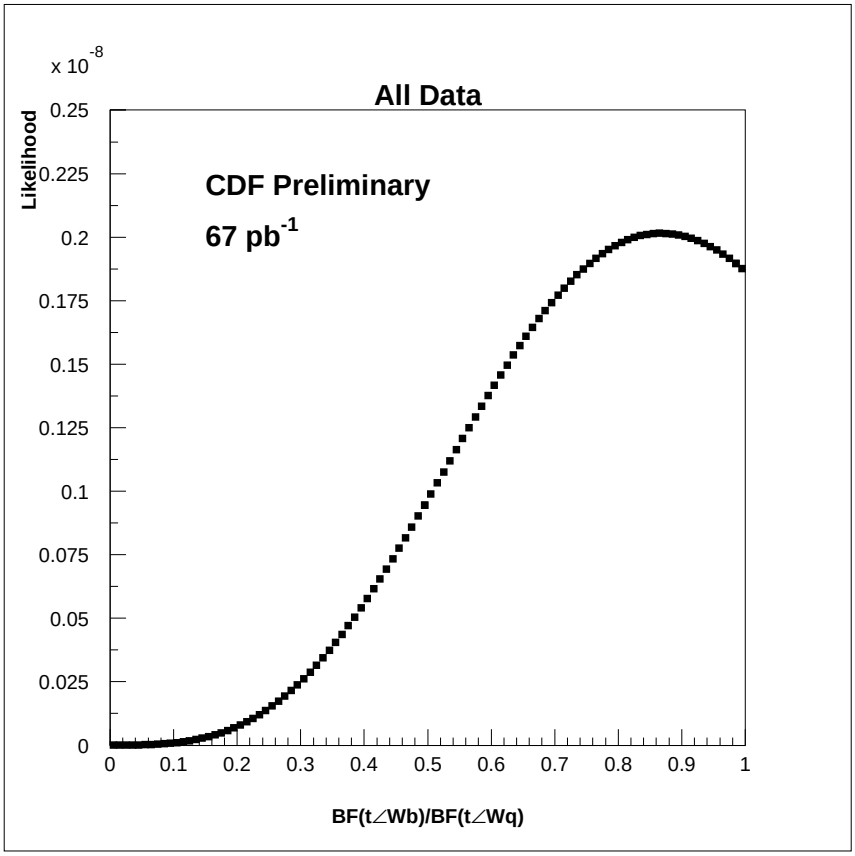


Figure 18: The likelihood function of $\mathcal{R}$ determined by using the relative rate of zero, single and doubled tagged events in the CDF dilepton data and the relative rates of single and double tagged events in the CDF lepton+jet data.

These relative rates of b tagged events are most efficiently combined by using a maximum likelihood technique to determine $\mathcal{R}$. The likelihood function that combines the CDF data from each channel is shown in Fig. 18 as a function of $\mathcal{R}$. The function peaks near unity, but has a large width that results from the limited statistics of the sample. From this distribution, one determines that

$$\mathcal{R} = 0.87^{+0.13}_{-0.30}(\text{stat})^{+0.13}_{-0.11}(\text{syst}), \tag{23}$$

where the systematic uncertainty is dominated by the uncertainty in b tagging efficiency.

Since $\mathcal{R}$ is a ratio involving three CKM matrix elements, we can convert this measurement into a statement about $|V_{tb}|$ by assuming, for example, the limits on

V_{td} and V_{ts} quoted in Eq. 18. This results in

$$|V_{tb}| = 0.11^{+0.89}_{-0.05}, \tag{24}$$

which is in agreement with the standard model expectation, albeit with large uncertainties. The result is most directly interpreted as implying $|V_{tb}| \gg |V_{ts}|$ or $|V_{td}|$.

8.3.2 Other Aspects of Top Quark Decays

The poor statistics of the DØ and CDF samples limit the detail with which one can study other aspects of top quark decays. However, I would like to mention two specific studies that are currently underway, though results are not yet available.

The V-A nature of the charged current results in the prediction that the decay $t \to W^+ b$ will result in W^+ bosons that are longitudinally polarised, that is, they will be produced with helicity aligned transverse to their momentum vector. This arises from the large top quark mass, as the fraction of longitudinal polarisation is given by

$$\frac{M^2_{top}/(2M^2_W)}{1 + M^2_{top}/(2M^2_W)}. \tag{25}$$

One will, with sufficient statistics, be able to extract this helicity information from the angular distribution of the charged or neutral lepton helicity angle measured in the lab frame that arises from the leptonic decay of the W^+ boson.[32]

One can also test for FCNC top decays by searching for evidence of Z° or γ bosons in final states such as

$$\begin{aligned} p\bar{p} &\to t\,\bar{t} \to Z^\circ c W^+ b \\ p\bar{p} &\to t\,\bar{t} \to \gamma c W^+ b \\ p\bar{p} &\to t\,\bar{t} \to Z^\circ c Z^\circ \bar{c}, \end{aligned} \tag{26}$$

which would arise if there was an appreciable FCNC top quark decay rate. These final states are essentially free of backgrounds,[33,34] so that the searches will be limited by the Z° branching ratios and the integrated luminosity.

8.4 Top Quark Production Properties

QCD calculations predict that top quarks should be produced with a relatively soft P_T distribution and in the central pseudorapidity region. Extensive theoretical studies have been done of heavy quark production, and the theoretical

uncertainties in the QCD predictions are quite modest. Although there has been some theoretical concern about the number and spectrum of additional jets arising from QCD radiation and higher-order processes, the general consensus is that these standard model uncertainties do not have a large effect on the production kinematics of top quarks.

However, there has been speculation that new physics beyond the standard model could have an influence on the production properties of the $t\,\overline{t}$ system.[35,36] There are in principle a large number of ways that such effects could be observed, which range from deviations from QCD in the $t\,\overline{t}$ production cross section to new particle resonances that couple strongly to the $t\,\overline{t}$ system and therefore influence the kinematics of the final state.

The statistics of the CDF and DØ samples limit our ability to exclude such anomolous effects, but one study illustrates how much we can learn from the Tevatron samples. A resonance coupling to the $t\,\overline{t}$ system (such as a heavy neutral gauge boson, or a Z') could result in an enhanced $t\,\overline{t}$ production cross section and be directly observed as an enhancement in the $t\,\overline{t}$ invariant mass distribution.[35] The observed $t\,\overline{t}$ invariant mass distribution from CDF is shown in Fig. 19, and is compared with what one would expect to observe if such a Z' boson does exist in Fig. 20. Note that this phenomena is predicted to strongly enhance the total $t\,\overline{t}$ production cross section for Z' boson masses of order 500 GeV/c^2 or less. These data have been used to exclude at the 95% CL the existence of a Z' with mass less than $\sim$ 470 GeV/c^2. This limit only takes into account statistical uncertainties; however, it is expected to be relatively insensitive to the systematic uncertainties that have not yet been fully characterised.

9 Future Top Quark Studies

9.1 Hadron Collider Development

Our current studies of the top quark system are based entirely on the top quark samples that have been collected at the Fermilab Tevatron Collider. With approximately 100 pb^{-1} of integrated luminosity, these samples are going to remain our only direct data on the top quark for the next three years.

The next Tevatron Collider run, known as Run II, is scheduled to begin in 1999 and will give us at least an order of magnitude improvement on the statistics

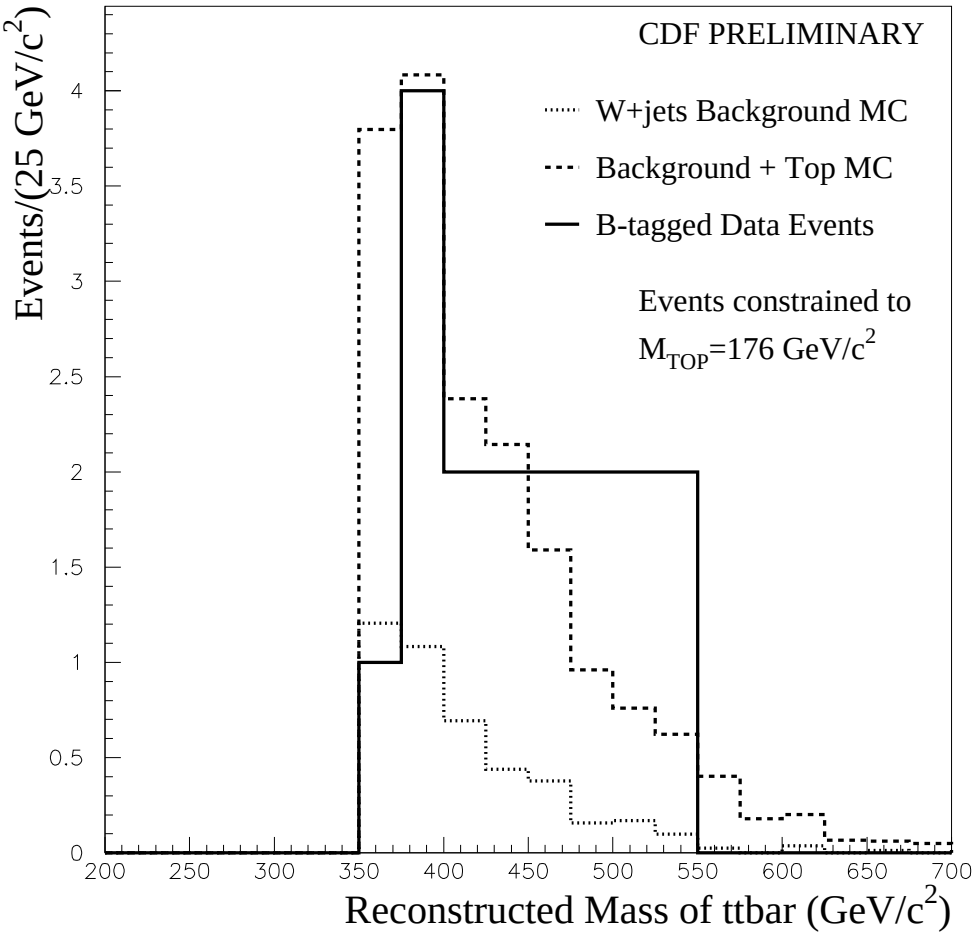


Figure 19: The $t\bar{t}$ invariant mass distribution of the CDF lepton+jet sample, using the fully-reconstructed lepton+≥ 4 jet events. The solid histogram is the data distribution, the heavy dashed histogram is the standard model prediction resulting from $t\bar{t}$ production and the estimated background, and the light dashed histogram is the mass distribution expected from the non-$t\bar{t}$ background. The top candidate events have been constrained to have a top quark mass equal to the CDF preliminary central value of 176 GeV/c^2.

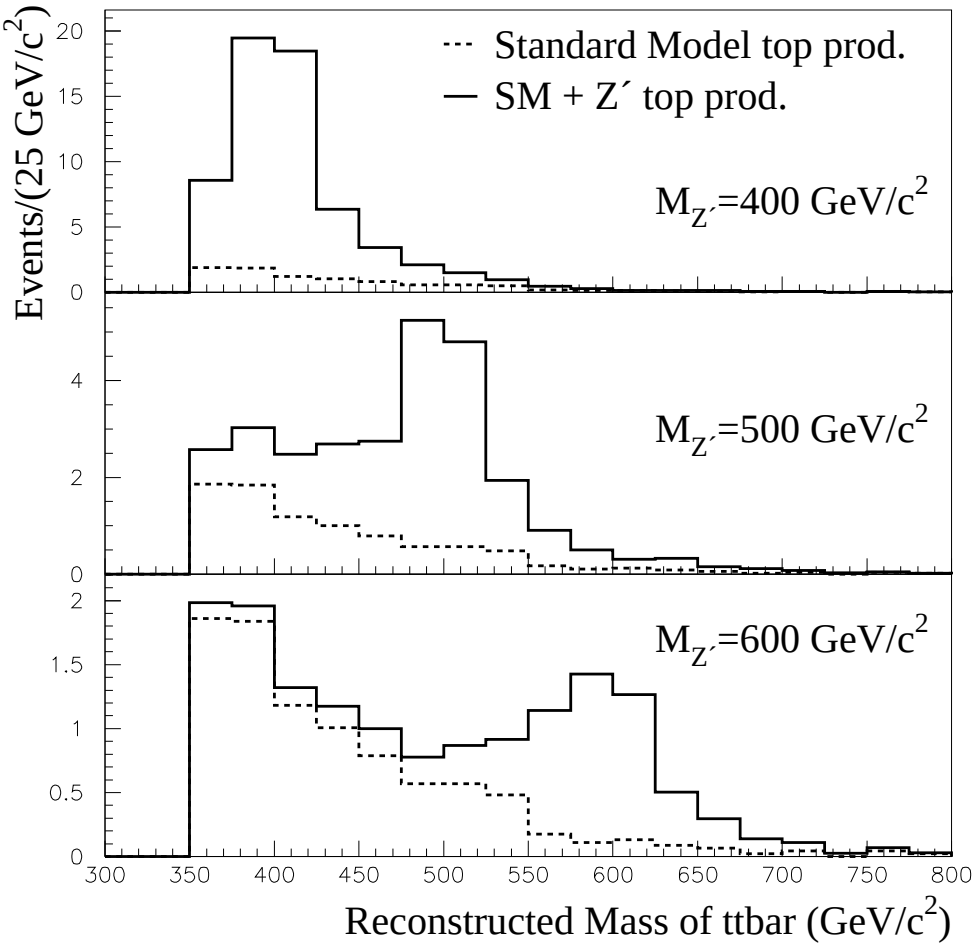


Figure 20: The $t\bar{t}$ invariant mass distribution that would be observed for different Z' boson masses. The theoretical predictions include the standard model QCD prediction combined with the Z' boson coupling to the $t\bar{t}$ system with Z' masses of 400, 500 and 600 GeV/c^2.

of Run I. This will be achieved with the construction of the Main Injector, a new synchotron that will replace the Tevatron's Main Ring as the accelerator and injector for the Collider, and the construction of a new $\bar{p}$ source. The Main Injector will allow significant increases in the maximum proton density that can be accomodated during acceleration and will provide a much larger acceptance of particles into the Tevatron Collider. In addition, the bunch spacing in the Tevatron Collider will be reduced from the current 3.0 μs to 396 ns and ultimately to 132 ns. The Tevatron maximum collision energy will also be increased by 10% to 2.0 TeV by improving the capability of the cryogenic systems.

These improvements will yield an instantaneous luminosity of 2×10^{32} cm^{-2}s^{-1}, an order of magnitude increase from Run I operating conditions. Over a period of four years, the facility is expected to provide each experiment with a data sample of 2 fb^{-1}, a factor of 20 increase in integrated luminosity over Run I. The increase in centre of mass energy results in a 30% increase in the $t\bar{t}$ yield, so an overall factor of 25 in produced top quark pairs is therefore expected.

The next step in top quark studies at hadron colliders will involve the Large Hadron Collider (LHC) currently under construction at CERN and scheduled for turn-on around 2004. The LHC, ultimately operating at $\sqrt{s} = 14$ TeV, will allow very high statistics studies due to the much larger $t\bar{t}$ production cross section and the much larger luminosity. The increased collision energy results in a $t\bar{t}$ production cross section of 1 nb, or a factor of 100 increase over the Run II production cross section. Even at relatively low initial luminosities of 10^{32} to 10^{33} cm^{-2}s^{-1}, the LHC will be producing top quarks at rates between 100 to 1000 times higher than the Tevatron during Run II. Although one has to take care in making direct comparisons due to the significantly more complex interactions that take place at the LHC, it is clear that this machine will have an enormous impact on what we will learn about the top quark.

I will briefly examine the top quark physics prospects of these two facilities in the following sections. A more detailed discussion of top quark physics prospects at the Tevatron is available.[37]

9.2 Tevatron Studies

The Run II top quark studies will benefit from both the much larger time-integrated luminosities made possible by the Main Injector and significant im-

provements in both the D$\emptyset$ and CDF detectors. Both collaborations are upgrading their charged particle detection systems by replacing all their subdetectors with new devices designed with the Run I experience in mind and optimised for Run II operating conditions. The D$\emptyset$ detector will now incorporate a superconducting solenoid magnet that will allow momentum analysis of charged particles, and both detectors will have enhanced silicon vertex tracking detectors that provide tracking coverage of virtually the entire luminous region. The collaborations are making other significant improvements in lepton identification systems, both for the detection of the high P_T leptons from the decay of W^+ bosons produced in $t\,\overline{t}$ events and the detection of the soft leptons from b quark decay.

9.2.1 Top Quark Event Yields

In order to estimate the expected number of reconstructed $t\,\overline{t}$ events, I have used the observed CDF yields of lepton+jet and dilepton events in Run I and taken into account the following effects:

- Run II will provide a factor of 20 increase in integrated luminosity.
- The SVX tagging efficiency will be improved by approximately a factor of 2 due to the increase in acceptance of the SVX subdetector to cover the entire luminous region at the interaction point.
- The soft lepton tagging efficiencies will be improved by of order 10% by extending the technique into the pseudorapidity region $1 \leq |\eta| \leq 2$.

With these assumptions, the expected yield of different categories of events are shown in Table 9. The uncertainties on these yields are relatively large and difficult to quantify. Although they are based on the observed Run I event yields, the expected improvement factors in tagging efficiency are based on extrapolations and detector simulations. However, they do form a relatively concrete basis on which to estimate the impact that the Run II data samples will have on top physics.

I have included in this table the predicted yields of the $Z+ \geq 4$ jet samples as well. With the given signal event yields, we are in a regime where the control of systematic uncertainties arising from detector effects and background uncertainties becomes essential to further improve the physics measurements. The $Z+$ jets data provides one of the key calibration samples as it constrains the theoretical models used to characterise the $W+$jets background to top production.

Channel	1 fb^{-1}	10 fb^{-1}
Tagged $W+ \geq 4$ jets	600	6000
Double tagged $W+ \geq 4$ jets	300	3000
Tagged Dilepton events	100	1000
$Z+ \geq 4$ jet events	200	2000

Table 9: Projected yields of observed events for 1 and 10 fb^{-1} of integrated luminosity for both the CDF and DØ experiments.

I will conservatively assume an integrated luminosity for Run II of 1 fb^{-1} for the following discussion, although many of the results will scale in a straight-forward manner with the assumed size of the data sample.

9.2.2 Run II Top Quark Cross Section

A more precise measurement of the top quark cross section is a good test of our understanding of perturbative QCD calculations. In addition, various extensions of the standard model predict that this cross section would be enhanced and therefore could be an indication of "new" physics.

The current uncertainties in $\sigma_{t\bar{t}}$ are dominated by the low statistics in the dilepton and lepton+jets signal samples. For Run II, these statistical uncertainties are expected to fall to of order 5% or better. The systematic uncertainties will therefore limit the measurement as these are currently at the level of 30-40%. However, it is possible to control most of these uncertainties as they arise from b tagging efficiencies, the background estimates and the integrated luminosity measurements. For example, the b tagging efficiencies can be obtained directly from the data using the rate of single to double-tagged lepton+jet events. I therefore expect these uncertainties to scale with the integrated luminosity.

I believe the systematic uncertainties will be limited, in fact, by how well we can measure the integrated luminosity in Run II. It is not clear that we will be able to determine this quantity to better than of order 3%, and I would therefore argue that this sets the "floor" on the systematic uncertainties on any absolute cross section measurements. If we expect that the other systematic uncertainies then scale with the number of observed candidate events, this implies an overall

systematic uncertainty of $\sim 7\%$.

With this assumption, the overall uncertainty in the cross section measurement could be of order 9%, which is considerably less than the current uncertainties of 15-20% on the standard model predictions.

9.2.3 Top Mass Measurement

We can conservatively estimate how well we can measure the top quark mass in Run II by extrapolating the uncertainties on the Run I mass measurements using the $W+ \geq 4$ jet sample.

Monte Carlo calculations have shown that the statistical uncertainty on M_{top} will scale as expected like $1/\sqrt{N}$, where N is the observed number of events in the sample. This assumes that the relative background rates will remain the same, a reasonable hypothesis since they are dominated by the intrinisic physics rates and not instrumentation effects. One therefore can expect a statistical uncertainty on M_{top} of ~ 2 GeV/c^2.

The control of the systematic uncertainties becomes the single most important aspect of this measurement. The largest source of systematic uncertainty relates to the measurement of the jet energies of the b quarks and quarks from the W^+ boson hadronic decays. Perhaps the most fundamental calibration tool is the observed W^+ signal in the dijet invariant mass distribution. However, independent calibrations can be performed by studying the balancing of observed energies in $Z+1$ jet and $\gamma+1$ jet events. With these studies, one can reasonably expect to reduce the systematic uncertainties arising from jet energy scales to of order 5 GeV in the Run I data set. Since this calibration is driven by the size of the Z+jet and γ+jet samples, one can assume that this uncertainty will scale statistically, resulting in a contribution to the systematic uncertainty of 1-2 GeV/c^2.

The other uncertainties that effect the current mass measurement together total 6-7 GeV/c^2 and should also scale statistically. Note that the largest contributions come form the understanding of the background shapes and the biases introduced by the tagging techniques. We would therefore predict that these would reduce to of 1.5-2.0 GeV/c^2 in a 1 fb^{-1} data sample. If we combine these together in quadrature, we arrive at a top quark mass systematic uncertainty of approximately 2.5 GeV/c^2, which is still larger than the expected statistical uncertainty. Further reductions in the systematic uncertainty are possible by, for example, using the double-tagged samples instead of just the single-tagged events.

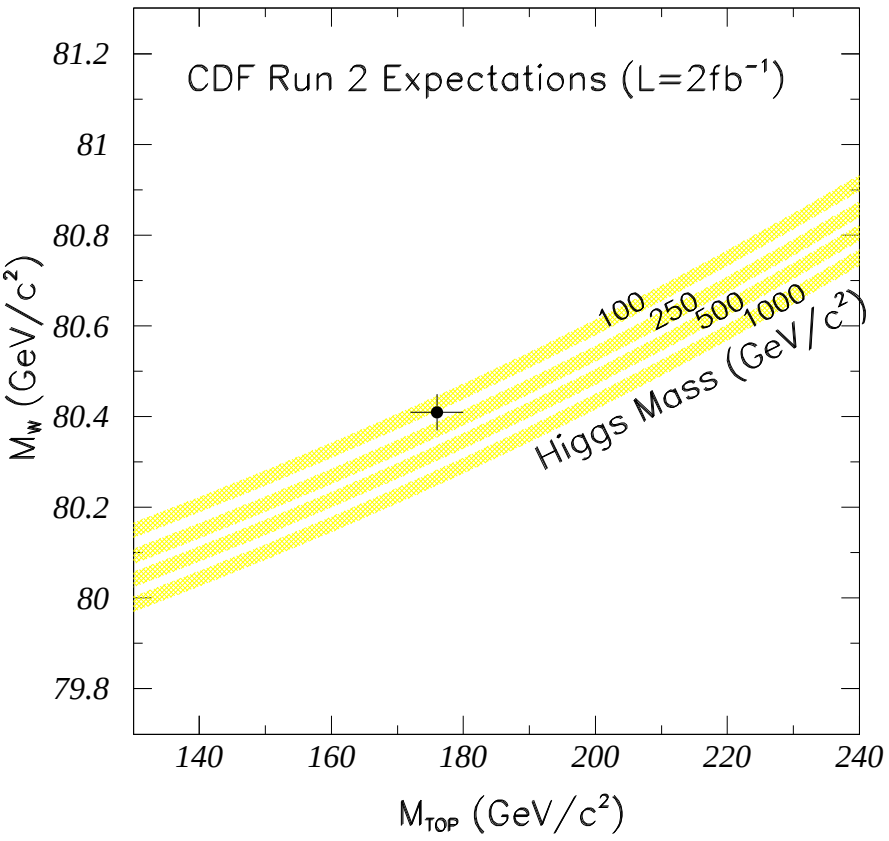


Figure 21: The expected precision of the top quark and W^+ boson mass measurements compared with the contours of standard model predictions for various Higgs mass assumptions. The central value represents the preliminary CDF Run I top quark and W^+ boson mass measurements. Note that the uncertainties assume an integrated luminosity of 2 fb^{-1}.

These data have an intrinsically better top quark mass resolution due to the reduced combinatorial background, and have a much smaller background due to the requirement of the second b tag.

Even without these expected improvements, the top quark mass uncertainty will be $\sim$ 3 GeV/c^2, when we combine both systematic and statistical uncertainties in quadrature. With the expected improvement in the W^+ boson mass measurement in Run II, we will have a very powerful test of the consistency of the standard model. This is illustrated in Fig. 21, where we plot the expected top quark mass versus the W^+ boson mass for various Higgs boson masses.

9.2.4 Top Quark Decays

The top quark branching fraction for the decay $t \to W^+b$ are most directly measured using the rates of tagged b quarks in both the lepton+jets and dilepton channels. The current statistical uncertainties on $\mathcal{B}(t \to W^+b)$ is set by the $\pm 20\%$ uncertainty on the rate of tagged W+jet events. This uncertainly will scale as $1/\sqrt{N}$, where N is the number of tagged events. Thus, given the extrapolated event yields, we can expect the statistical uncertainties on the tagging rates in the lepton+jets and dilepton samples to fall to of order $\pm 3\%$ and $\pm 4\%$, respectively.

The systematic uncertainties in these tagging rates are dominated by the uncertainty in the b tagging efficiency ϵ. In Run II, each experiment will have on the order of 10^7 B meson semileptonic decays that will provide a high statistics sample of relatively pure b decays that can be used to study the efficiencies of the various tagging techniques. With such large control samples, it is reasonable to expect that the systematic uncertainty on ϵ will scale with integrated luminosity.

With these assumptions, a simple Monte Carlo calculation predicts that one should be able to measure the branching fraction $\mathcal{B}(t \to W^+b)$ with a precision of $\pm 3\%$. As noted earlier, however, the constraint this places on V_{tb} depends on the values that V_{ts} and V_{td} can take on. If we assume the same range of values as given in Eq. 18, a Monte Carlo calculation combining both the lepton+jets and dilepton tagging fractions would allow us to constrain $V_{tb} \gtrsim 0.25$ at 90% CL. This constraint should also scale with luminosity so it will continue to improve with additional data. Although this limit is not as stringent as that obtained if one assumes unitarity of the CKM matrix, it is an important test of the assumption that only 3 quark generations couple to the electroweak force.

With the larger Run II statistics, it will also be possible to make more precise measurements of the detailed structure of the tW^+b vertex. For example, the V-A nature of the current involved in the decay predicts that the decay $t \to W^+b$ will result in W^+ bosons that are longitudinally polarised. One will be able to extract this helicity information from the angular distribution of the charged or neutral lepton helicity angle measured in the lab frame.[32] Monte Carlo studies[37] indicate that this fraction can be measured to of order 3% or better. This will make this a good test of the nature of the charged current decay. Any anomolous couplings are likely to become evident on the basis of this measurement.

Searches for anomolous top quark decays will also be possible. For example,

assuming that the γW^+ final state is not background limited, then a naïve calculation can be made assuming approximately 50% detection efficiency for the γ from the decay $t \to \gamma c$ or $t \to \gamma u$. The efficiency for detecting the γ+jet final state relative to the 3 jet final state resulting from the decay $t \to q\bar{q}'b$ would be ~ 0.5. With the expected lepton+jet event yields, we would be sensitive to $\mathcal{B}(t \to \gamma q)$ as small as 0.3%. Limits on decays mediated by Z° bosons would suffer by a factor of ~ 5 due to the necessary requirement of a dilepton decay of the Z° boson. These assume that the final states are not background limited at this sensitivity, an assumption that is difficult to test with the current data samples.

9.2.5 New Physics Searches

The search for new physics will continue at the Tevatron Collider during Run II, and the sensitivity of the $t\bar{t}$ system will only continue to improve with the increased event yields.

As one example of this, I show in Fig. 22 the expected $t\bar{t}$ invariant mass distribution after 1 fb^{-1} of running, assuming the existence of a Z' boson with a mass of 800 GeV/c^2. A clear signal is visible over the standard model prediction. One would be able to exclude the existence of such an object up to Z' masses of order 1 TeV/c^2 during Run II.

There are other speculations about new physics that will be addressed by studies of top production in Run II. The production of single top quarks via the process $q\bar{q}' \to W^* \to t\bar{b}$ is a direct way of measuring the partial width $\Gamma(t \to W^+ b)$ and searching for anomolous couplings between the top quark and the electroweak bosons.

These are only an example of the topics that will be addressed, but they demonstrate that the Tevatron during Run II will continue to be an exciting place to study top quark phenomenology.

9.3 LHC Studies

There have been many comprehensive studies performed of the potential for top quark physics at the much higher centre-of-mass energy afforded by the LHC. However, most of these studies are now dated as they were completed prior to the discovery of top. Not only does our current understanding of the properties of the top quark (most notably its mass) make many of these studies irrelevant, both the

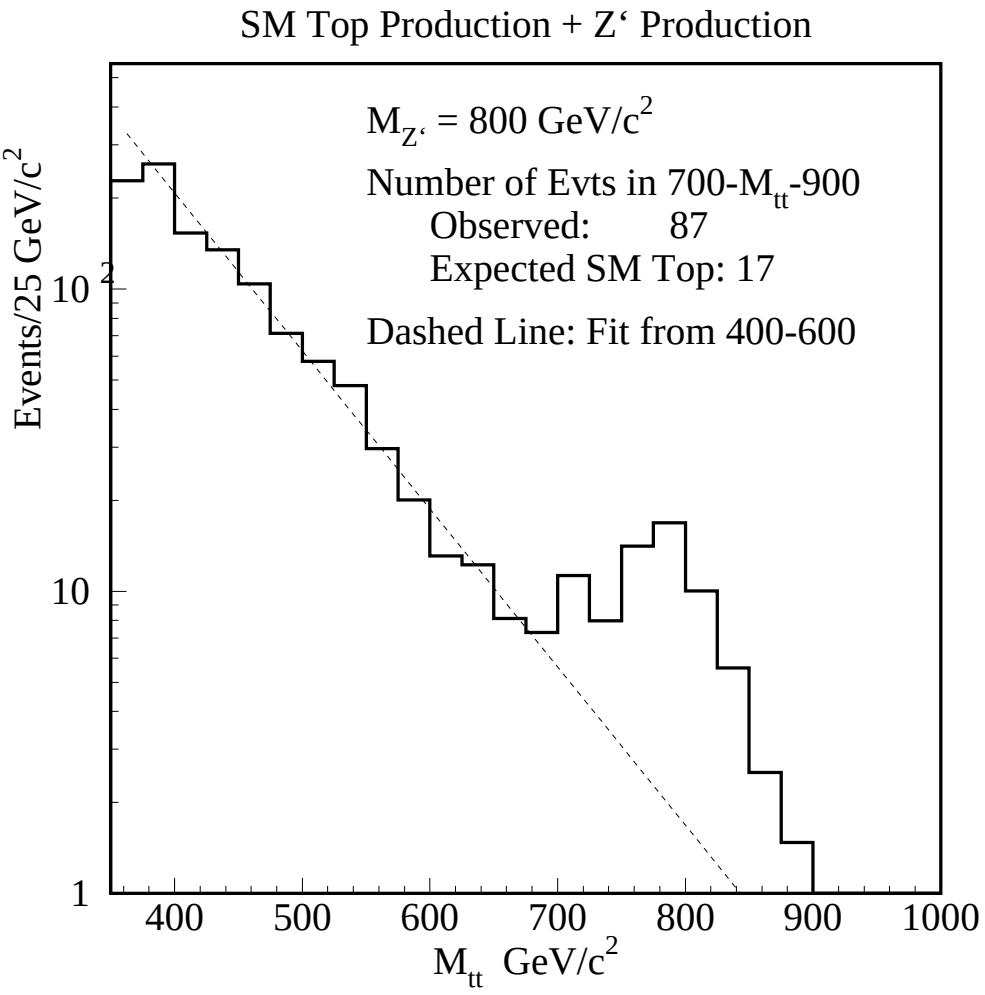


Figure 22: The expected $t\bar{t}$ invariant mass distribution assuming standard model production and the existence of a Z' boson that couples to the $t\bar{t}$ system.

D$\emptyset$ and CDF collaborations have taken enormous steps forward in understanding how to select and study $t\bar{t}$ candidate events in a hadron collider environment and these are not reflected in the previous studies.

For example, the earlier SSC and LHC studies[38] had concluded that a precise measurement of the top quark mass would be difficult given the large combinatorial backgrounds and the difficulty of performing a reliable jet energy calibration. These studies had concluded that top quark mass measurements with a precision of order 2-3 GeV/c^2 were possible with very large data samples. We now expect to acheive this level of precision at the Tevatron with the Run II data samples.

However, I note that the LHC will produce $t\bar{t}$ pairs at an enormous rate.

Even at a luminosity of 10^{33} cm^{-2}s^{-1}, the LHC will be producing of order 6000 $t\bar{t}$ pairs per day. Roughly speaking, an LHC experiment will be able to collect the same number of top events in one full day of running that would require a year's worth of data collection at the Tevatron. This will give an LHC experiment an enormous advantage in statistical power over the comparable Tevatron study. It is therefore reasonable to expect that most of the studies that I have discussed here will become very quickly systematics limited.

As a concrete example of this, the uncertainty in the top quark mass measurement will still be dominated by the systematic uncertainties in establishing the calorimeter energy scale. Although the *in situ* calibration of the calorimeter using the observed $W \to q\bar{q}'$ invariant mass distribution will provide a good calibration signal, the calibration of the b jet energy scales may become one of the limiting factors. Uncertainties arising from the additional "gluon" jets in the events will also remain, though they can be reduced by requiring, for example, two b tags and only considering lepton+4 jet events. The ultimate precision of an LHC mass measurement is difficult to quantify, but it is reasonable to expect that it can be reduced to of order 1 GeV/c^2 or perhaps less. At this level, the top quark mass is no longer expected to be the limiting factor in testing the consistency of the standard model.

The very large statistics samples available at the LHC make it possible to search for rare top quark decays. However, such a search will only be possible if the rare decay mode yields a sufficiently unique signature. For example, a signal for the rare decay $t \to Z^{\circ}c$ may ultimately be limited by the standard model process $p\bar{p} \to W^{+}Z^{\circ}X$ where the associated produced partons are b or c quark candidates. One can expect that the sensitivity of an LHC study will be at least an order of magnitude better than the corresponding Tevatron limit, but this is purely speculation as a detailed study taking into account potential backgrounds and signal efficiencies has not been performed.

10 Conclusions

The hadron collider environment has proved to be quite successful in discovering the top quark and beginning to elucidate its properties. However, these initial Tevatron studies of the top quark are currently statistics limited. Both the DØ and CDF collaborations have now completed data collection for Run I and have

event samples with sensitivities of approximately 100 pb^{-1}. With these data, both collaborations will be able to improve the statistical uncertainties on the top quark cross section and mass, and they are currently involved in additional studies that will reduce the systematic uncertainties in these measurements.

The CDF and DØ collaborations' preliminary estimates of the top quark mass, $176 \pm 10 \pm 13$ GeV/c^2 (CDF) and $199^{+19}_{-21} \pm 22$ GeV/c^2 (DØ), make it the heaviest known fermion in the standard model. The observed rate of $t\,\overline{t}$ events is consistent with standard model predictions, and make it the rarest phenomena observed in proton-antiproton annihilations. The very preliminary studies of top quark production and decay properties have yielded results that are consistent with the standard model predictions. However, additional analyses are underway and results from the full Run I data set will yield further insights on the properties of this unique fermion. Because of the massiveness of this fermion, it will be a unique probe into the physics of the standard model and what lies beyond this theory.

The Tevatron will continue to have a monopoly on direct $t\,\overline{t}$ studies for the next eight years. Run II, starting in 1999, will provide $t\,\overline{t}$ samples at least 20 times larger than those available in Run I, and will allow the first "high statistics" studies of the top quark. However, the LHC will be the ultimate hadron collider for top quark studies, as most of the standard model measurements will rapidly become systematics limited at this machine. In all, the future of top quark studies at hadron colliders looks very promising indeed.

Acknowlegements

I would like to thank the members of the DØ and CDF collaborations who kindly provided me with details of their respective analyses. I note that the research reported here would not have been successful without the excellent support of the staff of the Fermilab National Accelerator Laboratory and the institutions involved in this research. I would also like to acknowledge the organisers of the SLAC Summer Institute for their success in arranging such a stimulating and enjoyable meeting.

Support for this work from the National Sciences and Engineering Research Council of Canada is gratefully acknowledged.

References

[1] S. L. Glashow, Nucl. Phys. **22**, 579 (1961);
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, in *Elementary Particle Theory*, edited by N. Svartholm (Almquist and Wiksells, Stockholm, 1968), p. 367.

[2] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 352 (1977).

[3] D. Buskulic *et al.* (ALEPH Collaboration), Phys. Lett. **B335**, 99 (1994);
M. Acciari *et al.* (L3 Collaboration), Phys. Lett. **B335**, 542 (1994);
P. Abreu *et al.* (DELPHI Collaboration), Z. Phys. **C65**, 569 (1995);
G. Alexander *et al.* (OPAL Collaboration), CERN-PPE/95-179 (1995).

[4] R. Ammar *et al.* (CLEO Collaboration), Phys. Rev. **D49**, 5701 (1994).

[5] Review of Particle Properties, Phys. Rev. **D50**, 1304 (1994).

[6] D. Decamp *et al.* (ALEPH Collaboration), Phys. Lett. **B236**, 511 (1990);
O. Adriani *et al.* (L3 Collaboration), Phys. Lett. **B313**, 326 (1993);
P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. **B242**, 536 (1990);
M. Z. Akrawy *et al.* (OPAL Collaboration), Phys. Lett. **B236**, 364 (1990);
G. S. Abrams *et al.* (MARK II Collaboration), Phys. Rev. Lett. **63**, 2447 (1989);
I. Adachi *et al.* (TOPAZ Collaboration), Phys. Lett. **B229**, 427 (1989);
S. Eno *et al.* (AMY Collaboration), Phys. Rev. Lett. **63**, 1910 (1989);
H. Yoshida *et al.* (VENUS Collaboration), Phys. Lett. **B198**, 570 (1987).

[7] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 341 (1995);
F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 220 (1994);
C. Albajar *et al.* (UA1 Collaboration), Phys. Lett. **B253**, 503 (1991);
J. Alitti *et al.* (UA2 Collaboration), Phys. Lett. **B276**, 365 (1991).

[8] J. F. Gunion, H. E. Haber, G. Kane and S. Dawson, *The Higgs Hunter's Guide* (Addison-Wesley, New York, 1990); S. L. Glashow and E. E. Jenkins, Phys. Lett. **B196**, 233 (1987).

[9] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 2667 (1994).

[10] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **72**, 2138 (1994).

[11] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 225 (1994);
Phys. Rev. **D50**, 2966 (1994).

[12] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995);
S. Abachi *et al.* (DØ Collaboration), Phys.Rev. Lett. **74**, 2632 (1995).

[13] D. O. Carlson and C.-P. Yuan, Phys. Lett. **B306**, 386 (1993);
T. Stelzer and S. Willenbrock, "Single Top Quark Production via $q\bar{q} \to t\bar{b}$," DTP/95/40, ILL-(TH)-95-30 (1995).

[14] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988);
W. Beenakker, H. Kuijf and W. L. van Nerven, Phys. Rev. **D40**, 54 (1989);
G. Altarelli, M. Diemoz, G. Martinelli and P. Nason, Nucl. Phys. **B308**, 724 (1988).

[15] E. Laenen, J. Smith and W. L. van Neerven, Nucl. Phys. **B369**, 543 (1992);
E. Laenen, J. Smith and W. L. van Neerven, Phys. Lett. **B321**, 254 (1994).

[16] E. Berger and H. Contopanagos, Phys. Lett. **B361**, 115 (1995); S. Catani, M. L. Mangano, P. Nason and L. Trentadue, "The Top Cross Section in Hadronic Collisions," CERN-TH/96-21 (1996).

[17] S. Frixione, M. Mangano, P. Nason and G. Ridolfi, Phys. Lett. **B351**, 555 (1995).

[18] See, for example, P. Azzi, invited talk at XXXIst rencontres de Moriond (QCD session), March 23-30, 1996, CDF/PUB/TOP/PUBLIC/3679 (1996).

[19] F. A. Berends, W. T. Giele, H. Kuijf and B. Tausk, Nucl. Phys. **B357**, 32 (1991).

[20] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **70**, 4042 (1993).

[21] S. Abachi *et al.* (DØ Collaboration), Nucl. Instrum. Methods **A350**, 73 (1994).

[22] F. Abe *et al.* (CDF Collaboration), Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988).

[23] Details of the CDF and DØ jet cluster algorithms are provided in:
F. Abe *et al.* (CDF Collaboration), Phys. Rev. **D45**, 1448 (1992).
S. Abachi *et al.* (DØ Collaboration), FERMILAB-PUB-95-020-E, July 1995, submitted to Phys. Rev. **D**.

[24] F. Abe *et al.* (CDF Collaboration), FERMILAB-PUB-95-149-E, June 1995. Submitted to Phys. Rev. Lett.

[25] See, for example, V. D. Barger and R. J. N. Phillips, "Collider Physics," Addison-Wesley Publishing Co. (1987), p. 281.

[26] Review of Particle Properties, L. Montanet *et al.*, Phys. Rev. **D50**, 1278 (1994).

[27] F. Paige and S. D. Protopopescu, BNL Report No. 38034, 1986 (unpublished).

[28] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **68**, 447 (1992); Phys. Rev. **D45**, 3921 (1992).

[29] Review of Particle Properties, L. Montanet *et al.*, Phys. Rev. **D50**, 1229 (1994).

[30] Review of Particle Properties, L. Montanet *et al.*, Phys. Rev. **D50**, 1315 (1994).

[31] H. Fritzsch, Phys. Lett. **B224**, 423 (1989).

[32] G. Kane, C.-P. Yuan and G. Ladinsky, Phys. Rev. D **45**, 124 (1992); M. Jezabek and J. H. Kuhn, Phys. Lett. **B329**, 317 (1994).

[33] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 1936 (1995).

[34] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 1942 (1995).

[35] C. T. Hill, Phys. Lett. **B345**, 483 (1995).

[36] See, for example, B. Holdom and M. V. Ramana, Phys. Lett. **B353**, 295 (1995).

[37] D. Amidei *et al.*, "The TeV2000 Report: Top Physics at the Tevatron," CDF/DOC/TOP/PUBLIC/3265 and DØ Note 2653, 1996 (unpublished).

[38] See, for example, D. Froidevaux, "Top Quark Physics at LHC/SSC," CERN/PPE/93-148, published in *'93 Electroweak Interactions and Unified Theories*, Ed. J. Tran Thanh Van, Editions Frontieres (1993) 509-526;
W. W. Armstrong *et al.* (ATLAS Collaboration), "ATLAS Technical Proposal," CERN/LHCC/94-43 (1994);
D. Denegri *et al.* (CMS Collaboration), CERN-PPE-95-183 (1995);
E. L. Berger *et al.* (SDC Collaboration), "SDC Technical Design Report," SDC-92-201 (1992).

Top at Future Linear e^+e^- Colliders

K. Fujii

KEK, National Laboratory for High Energy Physics
1-1 Oho, Tsukuba-shi
Ibaraki-ken 305, JAPAN

No paper was received from the author in time for printing.

Review of Precision Electroweak Data

M. Swartz

Stanford Linear Accelerator Center
Stanford, CA 94309

No paper was received from the author in time for printing.

VERTEX DETECTORS: THE STATE OF THE ART AND FUTURE PROSPECTS

C. J. S. Damerell

Rutherford Appleton Laboratory

Chilton, Didcot, OX11 0QX, England

ABSTRACT

We review the current status of vertex detectors (tracking microscopes for the recognition of charm and bottom particle decays). The reasons why silicon has become the dominant detector medium are explained. Energy loss mechanisms are reviewed, as well as the physics and technology of semiconductor devices, emphasizing the areas of most relevance for detectors. The main design options (microstrips and pixel devices, both CCD's and APS's) are discussed, as well as the issue of radiation damage, which probably implies the need to change to detector media beyond silicon for some vertexing applications. Finally, the evolution of key performance parameters over the past 15 years is reviewed, and an attempt is made to extrapolate to the likely performance of detectors working at the energy frontier ten years from now.

© 1995 by C. J. S. Damerell.

TABLE OF CONTENTS

1 Introduction

2 Energy Loss of High-Energy Charged Particles in Silicon

2.1 Simplified Treatment
2.2 Improved Treatment
2.3 Implications for Tracking Detectors
2.4 Summary

3 Physics and Properties of Semiconductors

3.1 Conduction in Pure and Doped Silicon
3.2 The *pn* Junction
3.3 Charge Carrier Transport in Silicon Detectors

4 Microstrip Detectors

4.1 Introduction
4.2 The Generic Microstrip Detector
4.3 Microstrip Detectors; Detailed Issues
4.3.1 Design Optimization
4.3.2 Spatial Precision in Microstrip Detectors
4.3.3 Electronics for Microstrip Detectors
4.4 Physics Performance and Future Trends

5 Pixel-Based Detectors

5.1 Introduction
5.2 Charge-Coupled Devices (CCD's)
5.2.1 Structure and Basic Operation
5.2.2 CCD Charge Transfer and Readout; Detailed Issues
5.2.2.1 Charge Transfer Process
5.2.2.2 Charge Detection
5.2.2.3 Vertex Detector Readout Options
5.2.3 Physics Performance and Future Trends
5.3 Active Pixel Sensors (APS's)
5.3.1 Design Options
5.3.1.1 Monolithic Detectors
5.3.1.2 Hybrid Detectors
5.3.2 Performance and Future Trends

6 Radiation Damage in Silicon Detectors

6.1 Introduction
6.2 Ionizing Radiation
6.3 Displacement Damage
6.4 Detector-Specific Effects
6.4.1 Microstrip Detectors and APS Devices
6.4.2 CCD's
6.4.3 Local Electronics
6.5 Future Prospects

7 Beyond Silicon

7.1 Gallium Arsenide Detectors
7.2 CVD Diamond
7.3 Local Electronics

8 Conclusions

Acknowledgments

References

1 Introduction

There is for me a considerable sense of nostalgia in giving these lectures, since I previously gave such a series at the Summer Institute of 1984, which was especially noteworthy since it was coupled with the Pief-Fest to mark the retirement of Panofsky as Director of SLAC. Younger readers will be surprised to learn that the 1984 Institute, on the theme of the sixth quark, included evidence for the discovery of top with a mass of 40 ± 10 GeV.

In my 1984 lecture series, I suggested that these candidate top events really needed additional experimental evidence in order to be proved or disproved, and that this would best be provided by a precision vertex detector able to resolve the associated B decays. At the time, this suggestion was not taken particularly seriously. A lecture series relating to experimental methods of heavy quark detection at the same Institute made no mention of vertex detectors. Detectors with the required precision were only beginning to be used in the *fixed target* regime, and many of these were based on technologies such as bubble chambers that were manifestly not applicable to the collider environment. My own lectures made mention of techniques which have subsequently fallen into disuse for this reason. However, my main aim in those lectures was to establish a case for silicon vertex detectors in the collider environment. Our community was at that time in the early stages of planning the LEP and SLC detectors, and I focused particularly on Z^0 decays as the firm ground on which to base the case for these silicon vertex detectors. One was heavily dependent on Monte Carlo simulations of events with heavy flavor decays, where the possibilities for flavor tagging and some measure of topological vertexing could be demonstrated. Physicists at the time could be forgiven for not being wholly convinced by these simulations. Silicon detectors in those days were limited in size to a few square centimeters, were typically serviced by a huge amount of local electronics (easily accommodated in a fixed target experiment, but completely excluded in a collider), and detector reliability was a major problem. Here again, access for servicing which was easy in the fixed target environment would become much more difficult at the heart of a hermetic collider detector. In 1984, these Monte Carlo studies left on one side a host of technical problems which required many years of hard work to solve. Due to the loosely coupled R&D projects of many collaborations, the progress made since then has been immense. We now have a large variety of silicon vertex detectors in use in fixed target as well as collider experiments around the world. New designs are constantly being fabricated and tried out in test beams. The associated local electronics has shrunk spectacularly, and at the same time, become much faster and more powerful.

My task is thus made easier than 11 years ago; silicon vertex detectors have become well-established within the standard toolkit of high-energy experiments. I no longer need to rely on Monte Carlo studies to prove their usefulness; we can just look at the data. On the other hand, the array of detector types available has become somewhat bewildering, and I shall aim to provide some systematic guidance for nonexperts. Furthermore, despite the fact that the proponents of silicon detectors have been able to expand their horizons, even planning in some cases to displace gaseous tracking detectors with tens of square meters of strip detectors, they have begun to run into serious challenges in some vertex applications. In various hadron beam experiments, most spectacularly the LHC at its design luminosity, silicon detectors as we now know how to build them will fail after an unacceptably short time, when placed close to the interaction region. This has stimulated a major effort with other materials of greater radiation resistance, as we shall see towards the end of these lectures.

We are seeing the beginning of a technology division between e^+e^- colliders and hadron colliders, in regard to vertex detection at the energy frontier. Both are well-suited to the use of silicon at large radii, for general purpose tracking. But it is likely that at the luminosities needed for "discovery physics" at the TeV energy scale, silicon detectors will continue to be useful for high resolution vertex studies in the e^+e^- collider environment but not at LHC.

There are clearly great advantages in remaining with the silicon technology as far as possible. A major reason for its rapid growth as a material for tracking detectors is that the *planar process* for manufacturing silicon integrated circuits has been developed to an extremely fine art. These developments are continuing at a pace which reflects the billions of dollars annually invested, for purposes which have nothing to do with scientific research, let alone particle physics.

Before plunging into our rather specialized topic in fine detail, it is useful to take a brief look at the overall scene of silicon devices, particularly regarding their utility as radiation detectors. For, unlike some detection materials which are not widely used outside of our field (e.g., liquid argon), silicon finds applications in a vast range of scientific sensors. We in particle physics are concerned with its use for tracking microscopes that allow us to probe the smallest and shortest lived particles in nature. Silicon devices also provide the means to see the largest and oldest structures in the universe. Between these extremes, these sensors find a vast number of diverse applications, some of great importance to mankind (e.g., in medical imaging). Technically, all these areas are closely linked, so progress in one field may be significant to many others. All these scientific applications are dwarfed by the use of silicon sensors in the mass consumer markets, notably in video cameras but with applications now extending into other areas. What makes this field particularly dynamic is the flow of ideas from people with very different aims and agendas. The next major advance for HEP detectors may come from an astronomer concerned about radiation damage to his space-based telescope, or from the designer of an output circuit able to function at HDTV readout rates. Similarly, those designing devices for HEP use may dream up an advance that happens to be much more significant for some other field.

Why is silicon the preferred material for high-precision tracking detectors, as well as for such a wide range of radiation detectors?

Firstly, a *condensed medium* is essential if point measurement precision below about 10 μm is required. Gaseous tracking detectors are limited by diffusive spreading of the liberated electron cloud to precision of typically some tens of microns. Such detectors are entirely adequate for a host of particle tracking applications, but not for precision vertex detectors. Having established the need for a condensed medium, one should in principle consider liquids. There was some work done on high precision liquid xenon tracking detectors in the '70s [1] but there were many problems, not least of which was maintaining purity in conditions where the high mobilities of many contaminants rendered them particularly potent. In contrast, silicon wafers refined to phenomenal purity levels can then be sawn, exposed to the atmosphere, and assembled in complex geometries, with no degradation of their bulk electron lifetime characteristics. For these reasons, silicon and other solids are generally to be preferred, as opposed to liquids, for high-precision tracking purposes. There are, however, many possible solid state detection media, so why pick silicon?

Silicon has a band gap of 1.1 eV, *low* enough to ensure prolific production of liberated charge from a minimum-ionizing particle, hereafter referred to as a MIP (about 80 electron-hole pairs per micron of track length), but *high* enough to avoid very large dark current generation at room temperature (kT at room temperature = 0.026 eV). Being a low Z element of excellent mechanical properties (high modulus of elasticity) makes silicon an ideal material for use in tracking detectors where multiple scattering is of concern. This is nearly always the case in vertex detectors where tracks need to be extrapolated to the interaction region, and the dynamics of the fragmentation process ensures that even at the highest CM energies, many of the particles produced are of relatively low energy.

Besides these detector-related reasons, one has the vast IC technology developed specifically for this material. Silicon is currently unique in the combination of assets it brings with it; the growth of huge crystals of phenomenal purity, the possibility of *n*- and *p*-type doping, the possibility of selective growth of highly insulating layers (SiO_2 and Si_2N_3), and the possibility of doing all these using microlithographic techniques, allowing feature sizes of around one micron (and falling with time). A very readable account of the remarkable human stories associated with these amazing developments is to be found in George Gilder's book on the subject [2]. Very small feature sizes are, of course, precisely what one requires in order to construct detectors of precision below ten microns. Overall, the art of producing integrated circuits is probably by far the most sophisticated, fastest developing area of technological growth in the history of mankind. Without these developments, silicon as a detector of nuclear radiation would have remained a minor player, subject to arcane production procedures, of limited use for the spectroscopy of low-energy gamma rays, and wholly inappropriate for particle tracking purposes.

Though the scientific applications are of great importance, they are dwarfed by the use of silicon detectors for mass market consumer products and commercial interests.

Accurate figures are not readily available, but it seems that approximately $100M per year is spent on R&D of CCD's for domestic video and still cameras. These are interline transfer devices of no direct use for most scientific imaging applications. About $10M is spent on CCD development for medical and other scientific imaging applications (mostly X-rays). Silicon devices specifically aimed at particle tracking (microstrip detectors, CCD's, and active pixel sensors, hereafter referred to as APS devices) probably attract only $1M (order of magnitude) in R&D per year.

Even the consumer market for silicon sensors is dwarfed by the really hot commercial areas. For example, it was recently reported that NEC demonstrated a 1 Gbit DRAM. Production devices are expected to follow in three year's time, after the expenditure of *a further* $1.5B of R&D funding. Much of this will go in the development of submicron manufacturing capability, which ultimately will benefit the particle physics instrumentation community. We can eventually look forward to *submicron* tracking precision with *subnanosecond* timing information. However, the pace of such developments will be determined by the major players outside our own field, and there will inevitably be a time lag of several years between a technology being available for mass produced IC's and it being affordable for our purposes.

While the silicon processing infrastructure and R&D for a specific device can be enormously expensive, once production begins the costs can be modest. The ingredients of integrated circuits (sand, air, aluminum) are ridiculously cheap, and this benefit can be seen dramatically in large production runs. For example, SONY produces approximately five million CCD's per year for the domestic video camera market, at a production cost of only around $10, including the microlens and color filter system. This is a truly amazing achievement, as you can convince yourself by just looking through a microscope at one of these devices.

In summary, the match between silicon (and its attendant technologies) to the aspirations of the experimentalist wishing to construct tracking detectors of the highest possible precision, is evident. Were it not for the problems of radiation damage (which are most serious in the context of hadron colliders), there is little doubt that our field would by now have standardized completely on this material for vertex detection. Some time ago, test devices even surpassed photographic nuclear emulsions in precision, and with all the advantages of electronic readout. The challenge of hadron machines has stimulated some brave souls to undertake the monumental task of achieving similar technical performance using more radiation-resistant materials than silicon. They have, of course, to solve the problems not only of the detectors but also of the local electronics. We shall take a brief look at what they are doing in Sec. 7 of this paper. Other than that section, we shall devote ourselves exclusively to a discussion of silicon detectors and electronics.

2 Energy Loss of High-Energy Charged Particles in Silicon

High-energy charged particles traversing crystalline silicon can lose energy in two ways. Firstly, by ionization of the atomic electrons. This simple picture becomes rather more complex in regard to the valence electrons, as we shall see. The second energy loss mechanism (the so-called non-ionizing energy loss or NIEL) consists of displacement of silicon atoms from the crystal lattice, mostly by the process of Coulomb nuclear scattering. Only if the energy transfer to the nucleus exceeds approximately 25 eV can the atom be displaced from its lattice site. Below that, the energy is dissipated by harmless lattice vibrations. This implies an effective threshold energy for displacement damage with incident electrons (for example) of around 250 keV. Displacement of silicon atoms to *interstitial* positions (creating a *vacancy* in the lattice where the atom had previously been located) is one of the main radiation damage mechanisms. For a high-energy particle, the fraction of energy loss going into the NIEL mechanism is relatively small, but the cumulative effects on the detector performance can be severe.

A detector placed in a neutron flux experiences no signal from primary ionization, but the interactions can cause a high level of NIEL in view of the large neutron-silicon scattering cross section. For both charged hadrons and neutrons, other mechanisms of energy loss and radiation damage exist, notably neutron capture followed by nuclear decay, and inelastic nuclear scattering. The effects of non-ionizing energy loss on silicon detectors are considered in Sec. 6. In this section, we focus on the ionization energy loss only.

2.1 Simplified Treatment

Let us first imagine all the atomic electrons to be free, as if the crystal consisted of the silicon nuclei neutralized electrically by a homogeneous electron plasma. As a charged particle traverses the material, it loses energy by collisions (Coulomb scattering) with the electrons. Close collisions, while rare, will result in large energy transfers, while the much more probable distant collisions give small energy transfers. The process can be thought of classically in terms of the impulse generated by the attractive or repulsive Coulomb interaction between the projectile and the electron. The net impulse will be a kick transverse to the direction of travel of the projectile (see Fig. 1). The greater probability of remote collisions arises simply

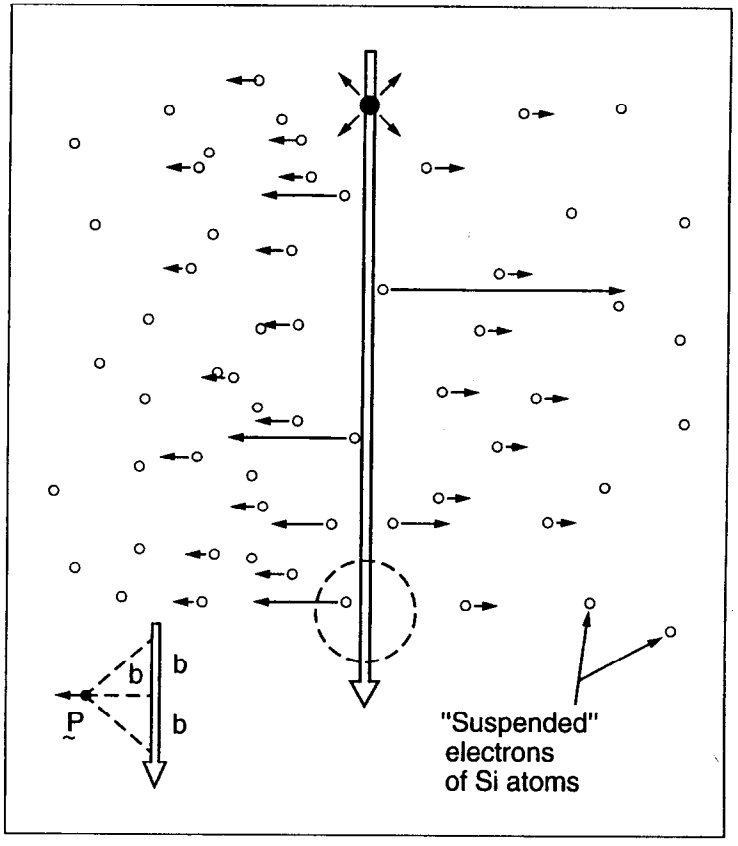


Fig. 1. Passage of charged particle through matter. Close collisions (electrons with small impact parameter b, shown by the inset) receive a powerful transverse impulse. Distant electrons receive a weak impulse.

from the greater volume of material available for collisions with a given impact parameter range, as the corresponding cylinder (of radius equal to the impact parameter) expands. In this simple case, the probability for a collision imparting energy E to an atomic electron is given by the Rutherford cross section

$$\frac{d\sigma_R}{dE} = \frac{2\pi {q_e}^4}{m_e c^2 \beta^2} \times \frac{1}{E^2}, \qquad (2.1)$$

where q_e and m_e are the charge and mass of the electron.

Note the mass of the struck particle in the denominator. This explains why scattering off the silicon nuclei, which are much more massive, causes very little energy loss, though these collisions do make the major contribution to the deviation in angle of the incident particle trajectory, via the process of multiple nuclear Coulomb scattering. Also, for sufficiently large momentum transfers, these collisions contribute to the NIEL referred to above.

We are interested in evaluating the *mean* energy loss and also the *fluctuations*, for traversal of a given thickness detector. An apparently simple approach would be to perform the integration over all E to obtain the mean energy loss, and to run a Monte Carlo calculation with multiple traversals to determine the energy loss distribution (straggling formula). However, we see that the integral diverges like $1/E$. The stopping power of this free-electron plasma would indeed be infinite, due to the long-range Coulomb interaction. In practice, the electrons are *bound*, and this prevents very low energy transfers to the vast number of electrons which are distant from the particle trajectory. This divergence is conventionally avoided by introducing a semi-empirical cutoff (binding energy) $E_{\min}$ which depends on the atomic number Z of the material. This is necessarily an approximate approach, since (for example) it ignores the fact that the outer electrons are bound differently in gaseous media than they are in solids. We shall need a more refined treatment to handle the cutoff in collisions with small energy transfer.

However, the Rutherford formula (with one small correction) is extremely useful as regards the close collisions, which are most important in defining the fluctuations in energy loss in "thick" samples (greater than approximately 50 μm of silicon, for example). The required correction is the upper cutoff $E_{\max}$ in energy transfer imposed by the relativistic kinematics of the collision process. If the projectile mass is much greater than m_e, we have $E_{\max} = 2m_e c^2 \beta^2 \gamma^2$. Due to the $1/E^2$ term in the Rutherford formula, we find that there is for each sample thickness, an energy transfer range in which the integrated probability of such transfers through the sample falls from almost unity to nearly zero. The Poisson statistics on energy transfers in this range gives rise to fluctuations on the overall energy loss for each traversal. Thus, the overall energy loss distribution consists of an approximately Gaussian core plus a high tail, populated by traversals for which a few close collisions occurred, each generating several times the mean energy loss. While the energy transfer region in which the probability function falls almost to zero is dependent on the sample thickness, this merely introduces an overall scale factor, so the *form* of the overall energy loss distribution is constant (the famous Landau distribution) over a wide range of detector thicknesses.

The rare close collisions with energy transfer greater than approximately 10 keV generate δ-electrons of significant range, which may be important in tracking detectors due to their potential for degrading the precision. For these close collisions, all atomic electrons behave as if they are free and the Rutherford formula may be used with confidence.

For thin samples, the energy loss fluctuations are not adequately handled by the Rutherford formula with cutoffs $E_{\min}$ and $E_{\max}$ In this case, the bulk of the energy loss arises from low-energy transfer collisions for which the binding of the atomic electrons must be handled in detail. We shall now consider the improved treatment of this case, specifically for crystalline silicon, though the same principles apply in general.

2.2 Improved Treatment

We note that energy loss is a discrete quantum mechanical process. We shall see that for very thin samples, a particle has even a finite probability of traversing the detector with no energy deposition at all.

For the low-probability close collisions, as noted above, it is valid to consider all atomic electrons as free, and the Rutherford formula applies. Ejected electrons of energies greater than approximately 10 keV will release further atomic electrons along their path. See Refs. [3] and [4] for a detailed treatment. For our purposes, it

is sufficient to note that the ultimate products that concern us are electrons, promoted into the conduction band of the material and holes (vacancies in the valence band), and that the generation of each electron-hole pair requires a mean creation energy W of approximately 3.6 eV. The precise value depends weakly on the temperature, see Fig. 2, and reflects the temperature dependence of the silicon band gap. Since this is around 1.1 eV, we note that electron-hole generation is a somewhat inefficient process; approximately 2/3 of the energy transferred from the primary (hot) electrons gives rise to phonon generation, eventually appearing as heat in the detector.

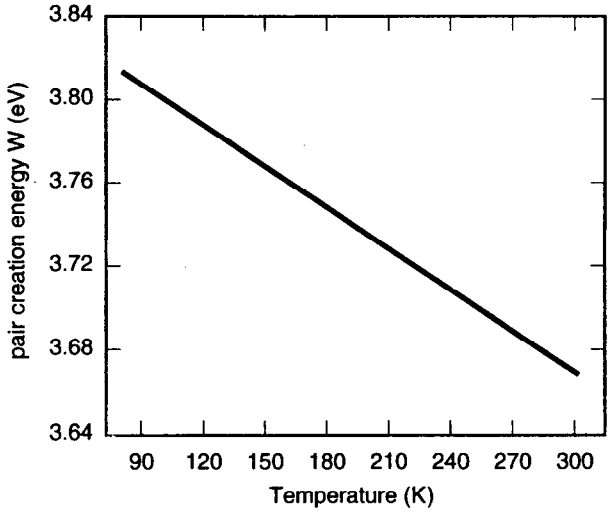


Fig. 2. Temperature dependence of the pair-creation energy W in silicon.

Beware, this has nothing to do with the non-ionizing energy loss (NIEL) referred to in the introduction to this section! Phonon generation (in contrast to NIEL) is a benign process which does not disrupt the crystal lattice and is usually ignored other than by enthusiasts for bolometric detectors. For our purposes, the δ-electrons ejected in close collisions can be considered to generate further electron-hole pairs at a mean rate of one per 3.6 eV of energy loss, *locally* on the track of the projectile, or *distributed* in the case that the δ-electron range is significant.

Qualitatively, the effect of the binding of the atomic electrons is to generate resonance-like enhancements in the energy loss cross section, above the values expected from the Rutherford formula. The K-shell electrons produce an enhancement in the 2 to 10 keV range, the L-shell in the 100 eV to 1 keV range, and the M-shell a resonance at around 20 eV. Below this resonance, the cross section rapidly falls to zero, in the region around 15 eV where the Rutherford formula would be cut off by the empirical ionization threshold energy.

The most satisfactory modern treatment proceeds from the energy-dependent photo-absorption cross section (a clean *point-like* process in the terminology of solid-state physics). This is, of course, closely linked to the energy loss process for charged particles, which fundamentally proceeds via the exchange of virtual photons. Combining photo-absorption and EELS (electron energy loss spectroscopy) data, Bichsel [5] has made a precise determination of the MIP energy loss cross section for silicon. The most subtle effects are connected with the valence (M-shell) electrons.

These valence electrons behave as a nearly homogeneous dense gas (plasma) embedded in a fixed positive-charge distribution. The real or virtual photons couple to this by generating longitudinal density oscillations, the quantum of which is called a *plasmon* and has a mean energy of 17 eV. The plasmons de-excite almost entirely by electron-hole pair creation. These somewhat energetic charge carriers are referred to as "hot carriers." Like the δ-electrons produced in the close collisions, they lose energy by thermal scattering, optical phonon scattering, and ionization. The topic of hot carriers is a major area of research, but for our purposes (as with the δ-electrons), we can ignore the details, since the end product that concerns us is again electron-hole pair creation at a rate of one per 3.6 eV of primary energy deposition. Figure 3 shows the photo-absorption cross section for silicon. The plasmon excitation is responsible for the extremely large cross section in the ultraviolet. It is by virtue of the low energy tail of this cross section in the visible that silicon has its optical sensing applications. The material becomes almost perfectly transparent once the photon energy falls below the 1.1 eV band gap energy.

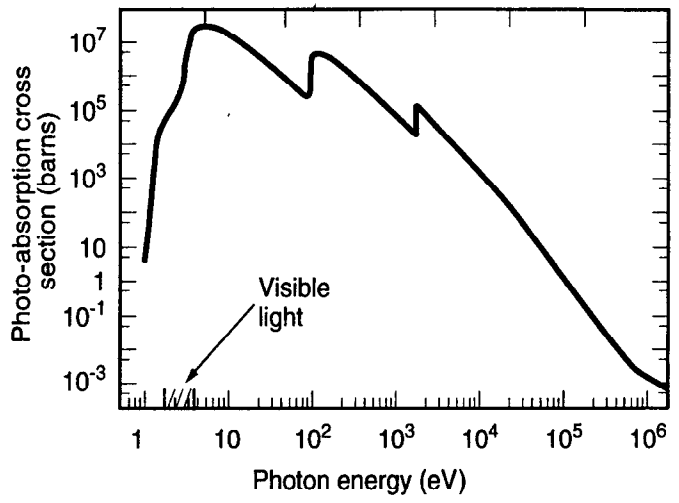


Fig. 3. Photo-absorption cross section of silicon versus photon energy.

The closely related energy loss cross section for a MIP is show in Fig. 4. Note that on average, it exceeds the Rutherford cross section by a factor of several in the energy range 10 eV to 5 keV. Above 10 keV, it is very close to the Rutherford value. By integrating the different components of this cross section, we can deduce the total mean collision rates associated with the different processes. These are as follows:

Electrons	Collision probability per micron
K (2)	8.8×10^{-3}
L (8)	0.63
M (4)	3.2

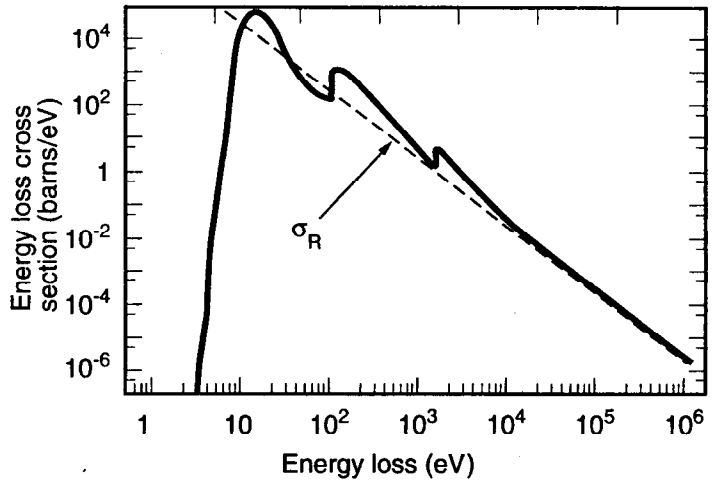


Fig. 4. Energy loss cross section for minimum-ionizing particles in silicon vs energy loss in primary collisions. The Rutherford cross section σ_R is also plotted.

Thus, despite the fact that on average a slice of silicon 1 μm in thickness will yield 80 electron-hole pairs, the Poisson statistics on the *primary* process (on average 3.8 collisions per micron) clearly implies a very broad distribution, with even a non-negligible probability of zero collisions, i.e., absolutely no signal. For thin samples, a correct statistical treatment of the primary process is essential if realistic energy loss (straggling) distributions are to be calculated. Their shapes are a strong function of the sample thickness, quite unlike the thickness-independent Landau distribution. The situation is depicted graphically in Fig. 5.

The area of each circle represents energy loss in a primary collision process. Those of smallest size correspond to plasmon excitation, while the larger ones represent the ionization of L-shell electrons. For these ten randomly selected tracks, the total energy deposition in the sample ranges from 37 eV to 390 eV.

2.3 Implications for Tracking Detectors

For high-precision tracking, there are clear advantages in keeping the silicon detector as thin as possible. A physically thin detector is optimal as regards multiple scat-

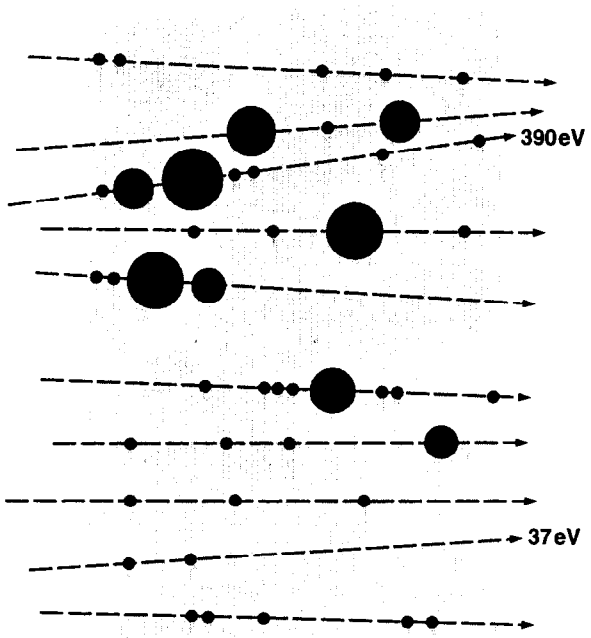


Fig. 5. Monte Carlo calculation of energy deposition in a 1 μm thick silicon detector. Area of a blob represents the energy deposited in each primary collision process.

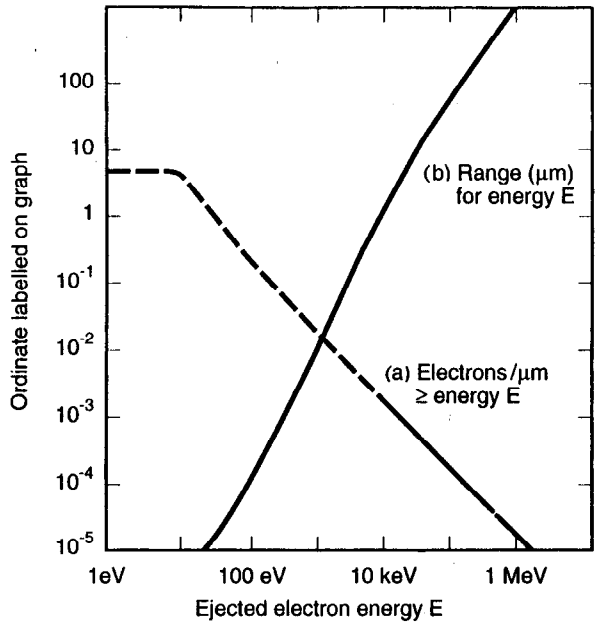


Fig. 6. (a) shows the number of electrons per micron of MIP track above a given energy, and (b) shows the range in silicon corresponding to that energy.

tering. A detector with the thinnest possible active region (which may be less than the physical thickness, as we shall see) is optimal as regards point measurement precision, for *two distinct reasons.*

For normal incidence tracks, the concern arises from δ-electrons of sufficient range to pull the centroid of the charge deposition significantly off the track. Figure 6(a) is an integral distribution of the number of primary electrons per micron of energy greater than a given value, and Fig. 6(b) shows the range of electrons of that energy in silicon. The range becomes significant for high-precision trackers for E greater than approximately 10 keV, for which the generation probability is less than 0.1% per micron. Thus, a detector of thickness 10 μm is much less likely to yield a "bad" co-ordinate than one of thickness 100 μm.

If the magnitude of the energy deposition in the detector is measured (by no means always possible), some of the bad co-ordinates will be apparent by the abnormally large associated energy. They could then be eliminated by a cut on the energy deposit, but this usually leads to unacceptable inefficiency and is rarely implemented. The situation is summarized in Fig. 7, which indicates the probabilities of the centroid for a track being pulled by more than a certain value (1 μm and 5 μm) as a function of detector thickness. The advantage of a thin active medium is apparent.

The second reason for preferring detectors to be as thin as possible applies to the case of angled tracks. In principle (and occasionally in practice), it may be possible to infer the position of such a track by measuring the entry and exit points in the detector, but more usually, the best one can do is to measure the centroid of the elongated charge distribution and take this to represent the track position as it traversed the detector mid-plane. In this case, large fluctuations in the energy loss (due to ejection of K- and L-shell electrons and δ-electrons) may be sufficient to cause serious track pulls for thick detectors. This is illustrated in Fig. 8. In the thin detector, there is a 10% probability of producing a δ-electron which, if it occurs near one end of the track, pulls the co-ordinate from its true position by 4 μm. In the thick detector, there is the same probability of producing a δ-electron which can pull the co-ordinate by 87 μm.

However, our enthusiasm for thin, active detector layers must be moderated by the primary requirement of any tracking system, namely a high efficiency per layer. Figure 9 (based on Ref. [5]) illustrates the problem we could already anticipate from Fig. 5. For very thin detectors (e.g., 1 μm Si), we see a very broad energy loss distribution with peaks corresponding to 0, 1, 2, ... plasmons excited, followed by a long tail extending to very large energy losses. An efficient tracking detector could never be built with such an active layer. Even at 10 μm silicon thickness, the true distribution is much broader than Landau and has a dangerous low tail. By 300 μm, the Landau distribution gives an adequate representation. Thus, while very thin detectors are ideal from the viewpoint of tracking precision, great care must be taken to assure that *system noise* allows a sufficiently *low threshold* to achieve the desired detector efficiency.

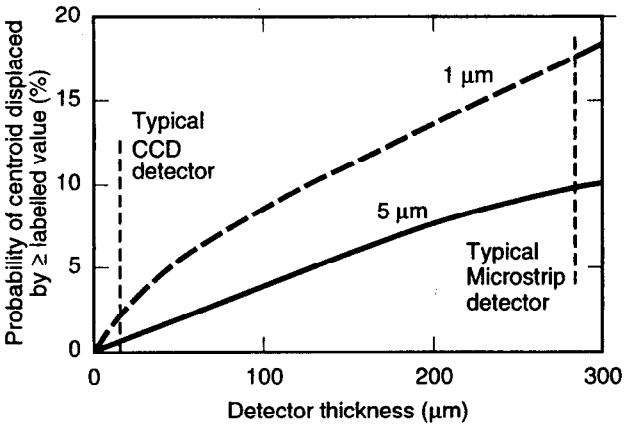


Fig. 7. Detector precision limitations from δ-electrons for tracks of normal incidence, as a function of detector thickness.

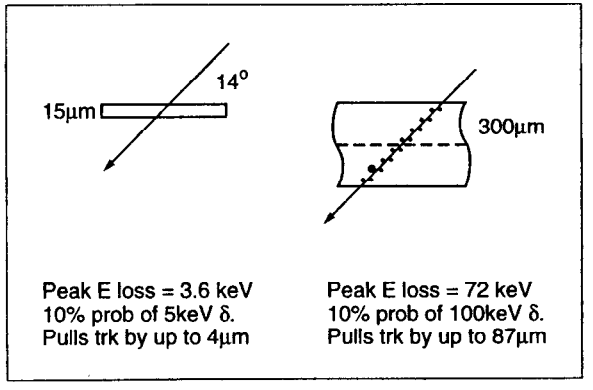


Fig. 8. Effect of energy loss fluctuations on detector precision for angled tracks.

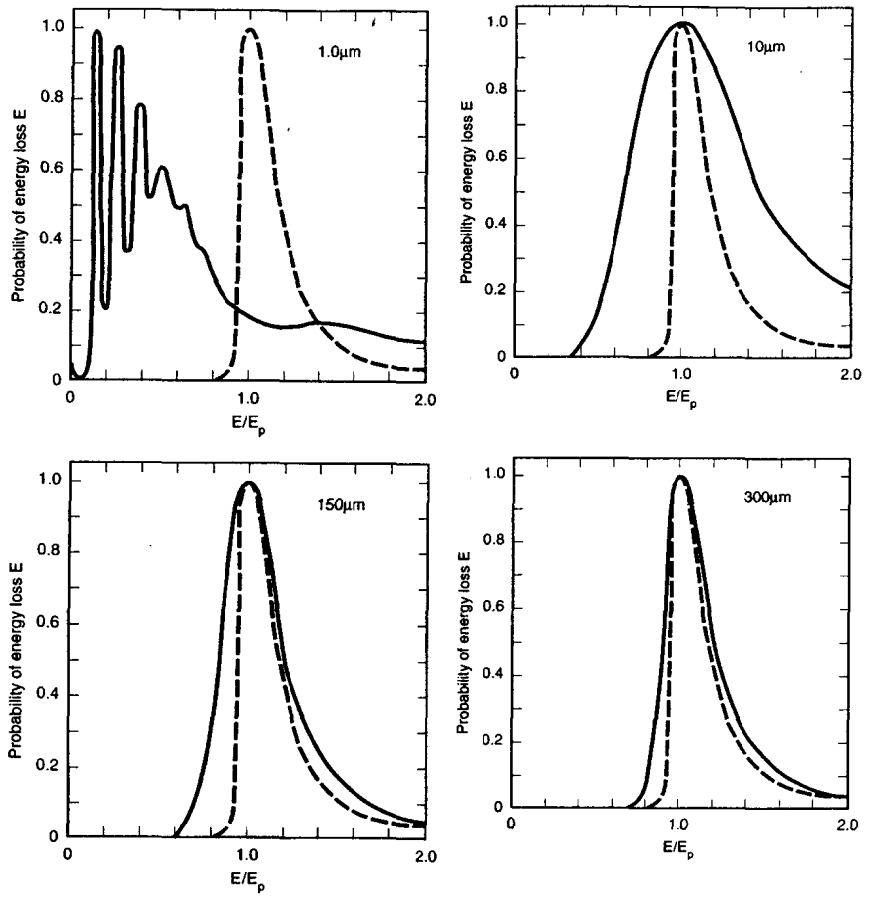


Fig. 9. Energy loss distributions for various silicon detector thicknesses, with (in each case) a Landau distribution for comparison. The separate peaks corresponding to 0, 1, 2 ... plasmon excitation are already merged by a thickness of 10 μm.

2.4 Summary

The valence electrons of silicon are very easily excited into plasmon oscillations from which they dislodge a small number (typically five) of electrons into the conduction band. A MIP thus creates a fine trail of electron-hole pairs along its track. The quantity W (energy needed to create an electron-hole pair) is approximately 3.6 eV, but depends on the band gap and hence (weakly) on the temperature. This energy loss process allows, in principle, unprecedented precision (much better than 1 μm) even compared to a nuclear emulsion (which needs typically a 400 eV δ-electron to blacken a grain). One does need to be prepared to exclude the measurements associated with large energy deposition, but these are rare in thin detectors.

How can this potential performance be achieved in practice? Standard IC processing (the planar technology) provides us with a host of suitable tools. This is after all one of the few areas of engineering in which submicron tolerances are now standard practice. In Secs. 4 and 5, we shall explore some types of detectors currently available. But first, we consider some of the basic properties of silicon which allow us in principle to collect and sense the signal charges we have been discussing in this section.

3 Physics and Properties of Semiconductors

Gaseous silicon has a typical structure of atomic energy levels (see Fig. 10). It has an ionization potential of 8.1 eV, i.e., it requires this much energy to release a valence electron, compared with 15.7 eV for argon. As silicon condenses to the crystalline form, the discrete energy levels of the individual atoms merge into a series of energy bands in which the individual states are so closely spaced as to be essentially continuous. The levels previously occupied by the valence electrons develop into the *valence band*, and those previously unoccupied become the *conduction band*. Due to the original energy level structure in gaseous silicon, it turns out that there is a gap between these two bands. In conductors, there is no such gap; in semiconductors, there is a small gap (1.1 eV in silicon, 0.7 eV in germanium), and in insulators, there is a large band gap. In particular, the band gap in silicon dioxide is 9 eV. This makes it an excellent insulator, and coupled with the ease with which the surface of silicon can be oxidized in a controlled manner, accounts partly for the pre-eminence of silicon in producing electronic devices.

We shall denote as E_v and E_c the energy levels of the top of the valence band and the bottom of the conduction band (relative to whatever zero we like to define). The energy needed to raise an electron from E_c to the vacuum E_0 is called the electron affinity. For crystalline silicon, this is 4.15 eV.

3.1 Conduction in Pure and Doped Silicon

To understand the conduction properties of pure silicon, the *liquid analogy* is helpful. This is illustrated in Fig. 11: (a) shows the energy levels in silicon under no applied voltage with the material at absolute zero temperature. All electrons are in the valence band, and under an applied voltage, (b) there is no change in the population of occupied states, and so no flow of current; the material acts like an insulator. At a high temperature, (c) a small fraction of the electrons are excited into the conduction band, leaving an equal number of vacant states in the valence band. Under an applied voltage, (d) the electrons in the conduction band can flow to the right and

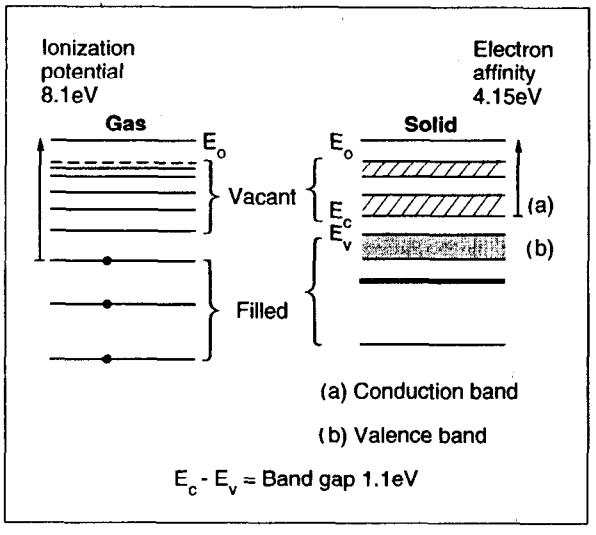


Fig. 10. Sketch of allowed energy levels in gaseous silicon which become energy bands in the solid material.

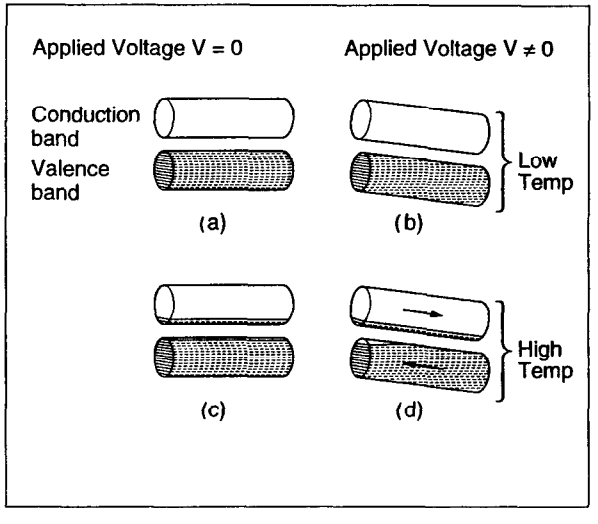


Fig. 11. Liquid analogy for a semiconductor.

there is a repopulation of states in the valance band which can be visualized as the leftward movement of a bubble (holes) in response to the applied voltage.

Now kT at room temperature is approximately 0.026 eV. This is small compared with the band gap of 1.1 eV, so the conductivity of pure silicon at room temperature is very low. To make a quantitative evaluation, we need to introduce the Fermi-Dirac distribution function $f_D(E)$ which expresses the probability that a state of energy E is filled by an electron. Figure 12(a) shows the form of this function

$$f_D(E) = \frac{1}{1+\exp\left(\frac{E-E_f}{kT}\right)}. \qquad (3.1)$$

Note that E_f, the Fermi level, is the energy level for which the occupation probability is 50%. Figure 12(b) sketches the density of states $g(E)$ in silicon. The concentration of electrons in the conduction band is given by the product $f_D g$, and the density of holes in the valence band by $(1-f_D)g$, as shown in Fig. 12(c). In pure silicon, the Fermi level is approximately at the midband gap, and the concentrations of electrons and holes are, of course, equal. These concentrations, due to the form of f_D, are much higher in a narrow band gap semiconductor, Fig. 12(d), than in a wide gap material, Fig. 12(e).

So far, we have been discussing pure (so-called *intrinsic*) semiconductors. Next, we have to consider the *doped* or extrinsic semiconductors. These allow us to achieve high concentrations of free electrons [n-type, Fig. 12(f)], or of holes [p-type, Fig. 12(g)], by moving the Fermi level very close to the conduction or valence band edge. The procedure for doing this is to replace a tiny proportion of the silicon atoms in the crystal lattice by dopant atoms with a different number of valence electrons.

Figure 13 shows the lattice structure characteristics of diamond, germanium, and silicon crystals. Silicon, with four valence electrons, forms a very stable crystal with covalent bonds at equal angles in space. It is possible (e.g., by ion implantation) to introduce a low level of (for example) pentavalent impurities such as phosphorus. By heating (*thermal activation* as it is called), the phosphorus atoms can be induced to take up lattice sites in the crystal. For each dopant atom, four of its electrons share in the covalent bonding with neighboring silicon atoms, but its fifth electron is extremely loosely bound. At room temperature, this electron would be free, and hence available for conduction in a sea of fixed positive charge (the phosphorus ions, present at precisely the same average density as the liberated electrons). At absolute zero, all valence electrons would be bound and the phosphorus-doped (n-type) silicon effectively an insulator. The mathematical description of the effect of doping in silicon is to retain the Fermi-Dirac distribution function, but to raise the Fermi level (50% occupation probability) close to the binding energy of the fifth electron, i.e., close to the conduction band edge. The population of those electrons within the conduction band is again given by the overlap of the Fermi-Dirac distribution function (now shifted in energy) and the density of states in the conduction band. Except at very low temperatures (where the Fermi-Dirac function is extremely sharp), the result is a high density of electrons (*majority carriers*) and a negligible density of holes (*minority carriers*) in the n-type material in equilibrium, as shown in Fig. 12(f).

Alternatively, silicon may be doped with trivalent impurities such as boron. In this case, three strong covalent bonds are formed, but the fourth bond is incomplete. This vacancy (hole) is easily filled by an adjacent electron. Thus, as in the intrinsic material, holes behave as reasonably mobile, positively charged carriers in a sea of fixed negative charge (the boron atoms with an additional electron embedded in the fourth covalent bond). The carrier concentrations (now with holes as majority carriers) are given by shifting the Fermi-Dirac distribution to within the hole binding energy, i.e., close to the valence band edge as shown in Fig. 12(g).

The general situation regarding doped silicon is sketched in Fig. 14, which indicates the energy levels corresponding to various commonly used dopant atoms. Pentavalent atoms are referred to as *donors* and trivalent atoms as *acceptors*. Note that the carriers are bound by only approximately 0.045 eV in the common n- and p-type dopants phosphorus and boron, compared to kT at a room temperature of 0.026 eV.

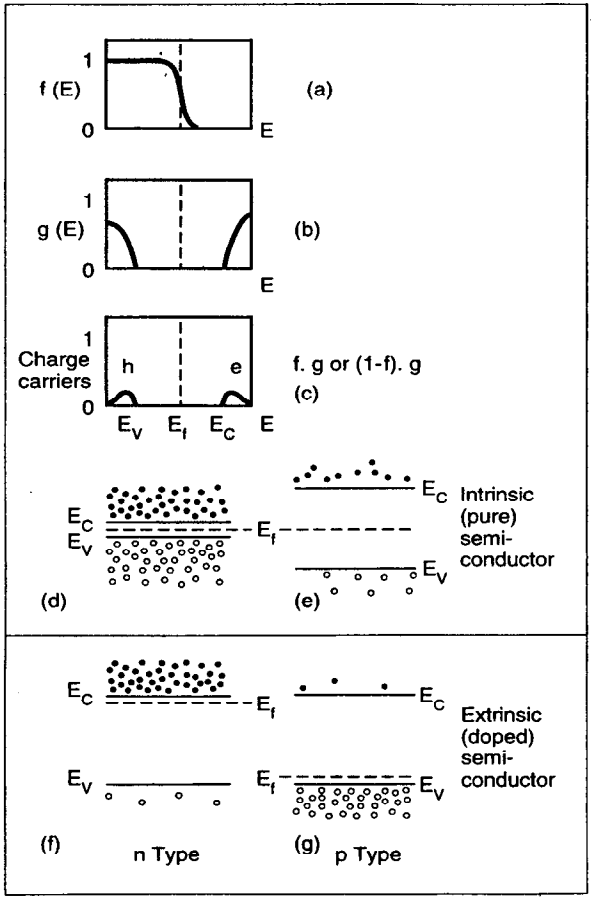


Fig. 12. (a) Fermi-Dirac distribution function. The slope increases as the temperature is reduced. (b) Density of states below and above forbidden band gap. (c) Concentration of charge carriers (electrons and holes) available for conduction. (d) and (e) Charge carrier distributions in narrow and wide band gap semiconductors. (f) and (g) Charge carrier distributions in *n*- and *p*-type semiconductors.

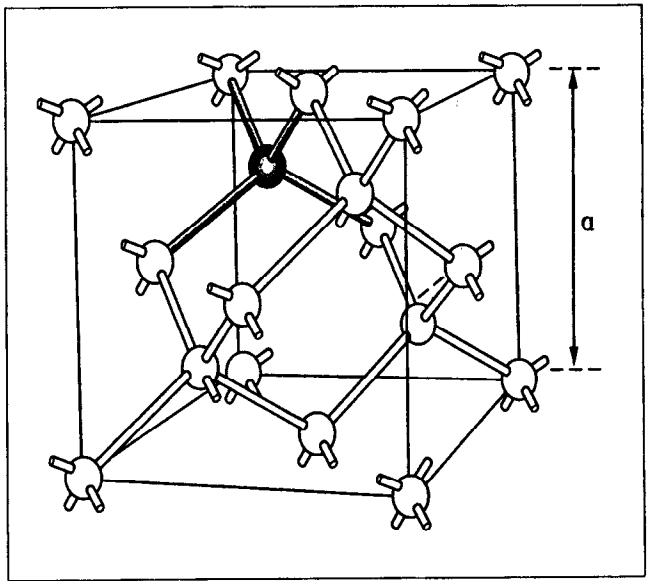


Fig. 13. Lattice structure of diamond, germanium, silicon, etc. where *a* is the lattice constant.

Figure 15 shows the concentration of electrons in n-type silicon (1.15 x 10^{16} arsenic dopant atoms per cm^3) as a function of temperature. Below about 100 K, one sees the phenomenon of *carrier freeze-out*, loss of conductivity due to the binding of the donor electrons. This is followed by a wide temperature range over which the electron concentration is constant, followed above 600 K by a further rise as the thermal energy becomes sufficient to add a substantial number of intrinsic electrons to those already provided by the dopant atoms. These will, of course, be accompanied by an equal concentration of mobile holes. The general behavior shown in Fig. 15 is typical of all doped semiconductors.

The *resistivity* ρ of the material depends not only on the concentration of free holes and electrons but also on their *mobilities*. As one would intuitively expect, the hole mobility is lower than that for electrons. Both depend on temperature and on the impurity concentration. At room temperature, in lightly doped silicon, we have

$$\text{electron mobility} \qquad \mu_n = 1350 \text{ cm}^2 \text{ (V s)}^{-1},$$

$$\text{hole mobility} \qquad \mu_p = 480 \text{ cm}^2 \text{ (V s)}^{-1},$$

and the resistivity is given by

$$\rho = \frac{1}{q_e(\mu_n \times n + \mu_p \times p)} \tag{3.2}$$

(n and p are the electron and hole concentrations).

For pure silicon at room temperature, $n_i = p_i = 1.45 \times 10^{10}$ cm^{-3} which gives $\rho_i = 235$ K Ω cm.

The carrier drift velocity (v_p for holes and v_n for electrons) is related to the mobility by $v_{p,n} = \mu_{p,n}\mathcal{E}$, where $\mathcal{E}$ is the electric field strength. This relationship applies only up to a maximum field, beyond which saturation effects come into play and one enters the realm of "hot carriers" which lose energy by impact ionization (creation of additional electron-hole pairs). Figure 16 shows the situation for silicon, as well as

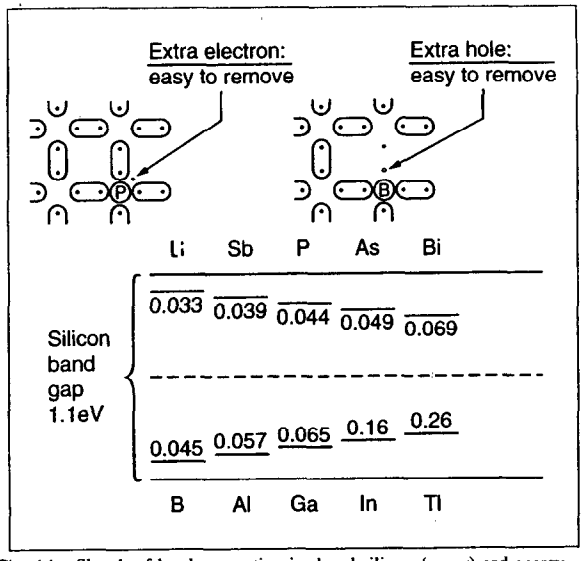


Fig. 14. Sketch of band occupation in doped silicon (upper) and energy levels within the band gap corresponding to various n- and p-type dopants (lower). Levels of acceptor atoms are conventionally measured from the top of the valence band, and levels of donor atoms are measured from the bottom up the conduction band.

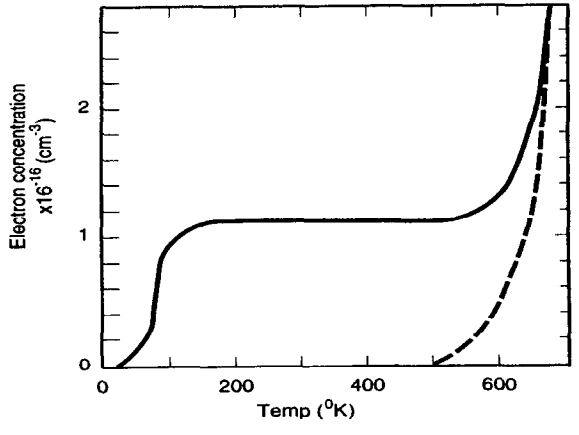


Fig. 15. Electron concentration versus temperature for n-type (arsenic doped) silicon. The dashed curve shows the concentration for intrinsic material.

the fact that nearly ten times higher electron drift velocities are achievable in gallium arsenide, which therefore has the potential for much faster electronic circuits.

The ionization rate is defined as the number of electron-hole pairs created per unit of distance travelled by an electron or hole. It depends primarily on the ratio $q_e\mathcal{E} / E_i$ where E_i is the effective ionization threshold energy, damped by terms reflecting the energy loss of carriers by thermal and optical phonon scattering, see Ref. [6]. For silicon, E_i is approximately equal to W (3.6 eV) for electrons and 5.0 eV for holes. The ionization rate becomes significant for electric fields in the range 10^5 to 10^6 V/cm in silicon, leading to the saturation of carrier drift velocity shown in Fig. 16.

The resistivity as a function of dopant concentration is shown in Fig. 17. For silicon detectors, as will be explained in the next section, we are often concerned with unusually high resistivity material, some tens of KΩ cm. From Fig. 17, one sees, for example, that 20 KΩ cm p-type material requires a dopant concentration of 5 x 10^{11} cm^{-3}. Remembering that crystalline silicon has 5 x 10^{22} atoms per cm^3, this implies an impurity level for the *predominant* impurities of 1 in 10^{11}. Even in the highly developed art of silicon crystal growing, this presents a major challenge. The resistivity noted above in connection with pure silicon (over 200 KΩ cm) is entirely unattainable in practice. Very high resistivity n-type silicon can be produced in the form of *compensated* material. The most uniformly doped material which can be grown is (for technical reasons) p-type, and this (with a resistivity of about 10 KΩ cm) is used to start with. It is then turned into n-type material by the procedure known as neutron doping. The crystal is irradiated with slow neutrons and by means of the reaction

$$Si^{30} + n \rightarrow Si^{31},$$

followed by

$$Si^{31} \rightarrow P^{31} + \beta^- + \overline{\nu}$$

is turned into n-type material. The resistivity is monitored and the irradiation ceases when this, having passed through a maximum, falls to the required value. In this way, material of resistivity as high as 100 KΩ cm can be made. Achieving

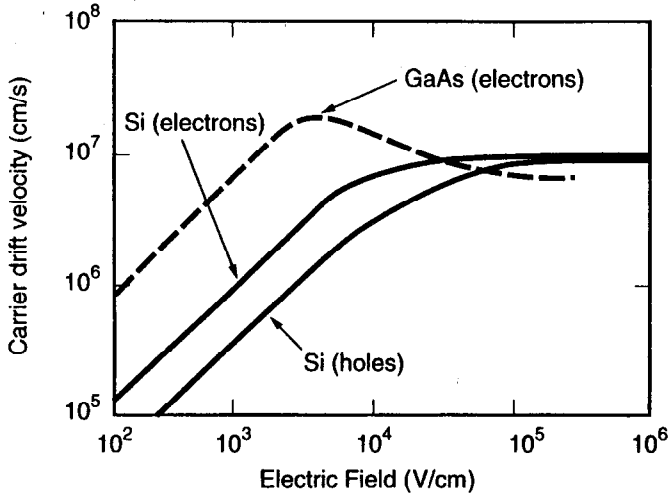


Fig. 16. Carrier drift velocity (electrons and holes) for silicon, and electron velocity for gallium arsenide as a function of electric field in the material.

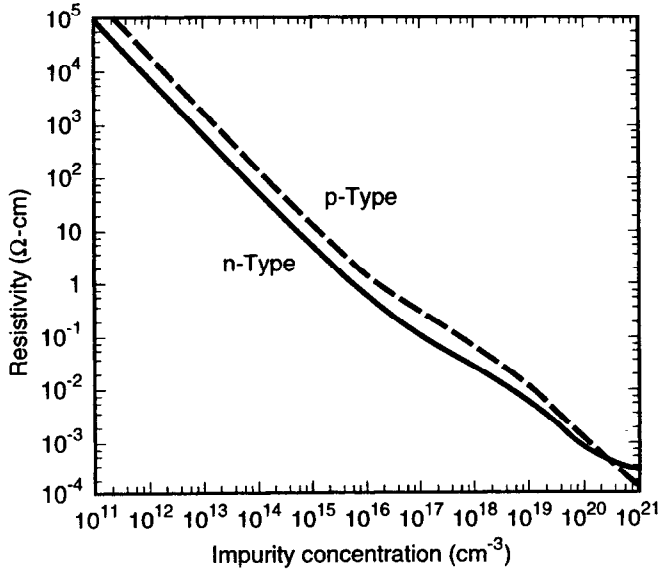


Fig. 17. Resistivity of silicon at room temperature as a function of acceptor or donor impurity concentration.

reasonable uniformity through the wafer of such a high resistivity is obviously extremely difficult.

We now consider more quantitatively the relationship between the carrier concentration and the Fermi level. The number of conduction band states occupied by electrons is given by

$$n = \int_{E_c}^{E_t} N(E) f_D(E) dE.$$

E_c and E_t are the energy at the bottom and top of the conduction band; $f_D(E)$ is the function (3.1); $N(E)$, the density of states, is given by the band theory of solids and is proportional to $(E - E_c)^{1/2}$. For the commonly encountered situation where Boltzmann statistics applies, for which the Fermi level is at least several times kT below E_c, the above integral can be approximately evaluated to yield

$$n = N_c \exp\left(-\frac{E_c - E_f}{kT}\right). \tag{3.3}$$

N_c is called the effective density of states. Its meaning is not as intuitively clear as the simple density of states $N(E)$; unlike $N(E)$, it is temperature dependent, being proportional to $T^{3/2}$.

The equivalent approximation for the hole concentration is

$$p = N_v \exp\left(-\frac{E_f - E_v}{kT}\right). \tag{3.4}$$

For *intrinsic* semiconductors, thermal agitation excites electrons from the valence band to the conduction band, leaving an equal number of holes in the valence band. In this case, $n = p = n_i$, where n_i is the intrinsic carrier density. There is a dynamic equilibrium between thermal generation on the one hand, and recombination of electrons in the conduction band with holes in the valence band, on the other. The neutrality condition obtained by equating Eqs. (3.3) and (3.4) leads to

$$E_f = E_i = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln\left(\frac{N_v}{N_c}\right). \tag{3.5}$$

Thus, the Fermi level of an intrinsic semiconductor lies very close to the middle of the band gap. The intrinsic carrier density is given from Eqs. (3.3) and (3.4) also:

$$pn = n_i^2 = N_c N_v \exp(-E_g / kT) \tag{3.6}$$

where $E_g = E_c - E_v$.

Note that

$$n_i = \sqrt{N_c N_v} \exp(-E_g / 2kT) \tag{3.7}$$

$$\propto T^{3/2} \exp(-E_g / 2kT).$$

Thus, n_i has a rapid temperature dependence, doubling for every 12°C rise for silicon around room temperature.

For doped silicon, e.g., n-type, the neutrality condition is between the ionized donors and the conduction band electrons created by the ionization process. For a dopant energy level E_d, the number of ionized donors is related to the Fermi level by the relation

$$N_d^+ = \frac{N_d}{1 + 2\exp\left(\frac{E_f - E_d}{kT}\right)}. \tag{3.8}$$

See Ref. [7]. From Eqs. (3.3) and (3.8), we have the neutrality condition

$$N_c \exp\left(-\frac{E_c - E_f}{kT}\right) = \frac{N_d}{1 + 2\exp\left(\frac{E_f - E_d}{kT}\right)}. \tag{3.9}$$

Figure 18 shows graphically the solution of Eq. (3.9) for two temperature values. At room temperature, the donor atoms are completely ionized and the carrier

concentration is essentially equal to N_d, with $E_f = E_{f1}$ a little below E_d. At the reduced temperature, $E_f = E_{f2}$ falls in the small energy range between E_d and E_c, and the carrier concentration plummets. Conversely, at very high temperatures, thermal excitation of valence band electrons would become dominant, causing the carrier concentration to rise rapidly, and the Fermi level to stabilize near the middle of the band gap, off-scale to the left in the figure. For p-type material, the number of ionized acceptors is given by

$$N_a^- = \frac{N_a}{1 + 4\exp\left(\frac{E_a - E_f}{kT}\right)}. \tag{3.10}$$

The difference in the factors in the denominator arises from the difference between the ground-state degeneracy for donor and acceptor levels.

In general, for doped material, we have

$$\left.\begin{aligned} n &= n_i \exp\left(\frac{E_f - E_i}{kT}\right) \\ p &= n_i \exp\left(\frac{E_i - E_f}{kT}\right) \end{aligned}\right\} \tag{3.11}$$

and $pn = n_i^2 = N_c N_v \exp(-E_g / kT)$ *just as for intrinsic material.* Thus, the deviation of a doped semiconductor from the intrinsic condition can be simply represented by a shift in the Fermi energy level with respect to the intrinsic level. The constancy of the pn product for different doping conditions is a particular example of the very important *law of mass action* which applies as much in semiconductor theory as it does in chemistry. In thermal equilibrium, the increase in electron concentration by donor doping causes a decrease in the concentration of mobile holes (by recombination) such that the pn product is constant. The ionized donors in this sense are passive bystanders, serving to preserve charge neutrality. It is generally valid to think of n-type material in equilibrium as containing only mobile electrons, and p-type material as containing only mobile holes, the majority carriers in each case.

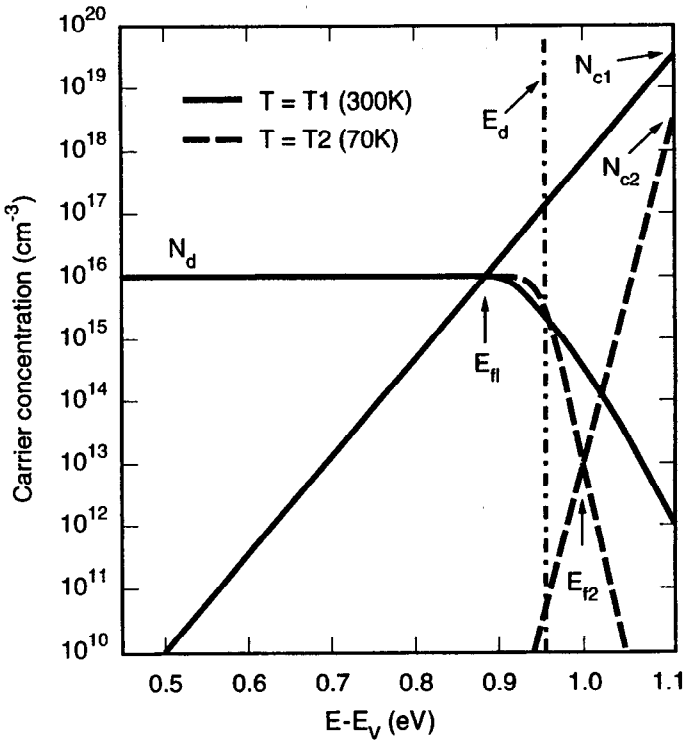


Fig. 18. Number of ionized acceptors and number of conduction band electrons versus the Fermi energy level E_f.

3.2 The *pn* Junction

We now need to introduce a most important fact related to conducting materials which are electrically in contact with one another and in thermal equilibrium; *they all must establish the same Fermi energy*. This applies to

metal/semiconductor systems
n-type/*p*-type systems, etc.

Charge flows from the high- to low-energy region for that carrier type until this condition is established. For example, at a *pn* junction, there develops a fixed space charge of ionized donors and acceptors, creating a field which opposes further drift of electrons and holes across the junction. The *depletion approximation* says that the semiconductor in this condition changes abruptly from being neutral to being fully depleted. This is far from obvious, and in fact, there is a finite length (the *Debye length*, typically $0.1\,\mu m$) over which the transition takes place. But the depletion approximation will be adequate for all the examples we need to consider. Let us look in some detail at the important case of the *pn* junction. Before contact [Fig. 19 (a)], the surface energy E_0 is equal in both samples; the *p*-type Fermi level is close to E_v and the sample is densely populated by holes; the *n*-type Fermi level is close to E_c and the sample is densely populated by electrons.

On contact, the electrons diffuse into the electron-free material to the left, and the holes diffuse to the right. In so doing, the electrons leave exposed donor ions (positively charged) over a thickness x_n in the *n*-type material, and the holes leave exposed acceptor ions (negatively charged) over a thickness x_p in the *p*-type material. This builds up an electric field which eventually just balances the tendency for current to flow by diffusion. Once this condition is reached [Fig. 19 (b)], the Fermi levels in the materials have become equal. The electrical potentials in the two samples (for example, the potential energy at the surface E_0 or at the conduction band edge E_c) are now unequal.

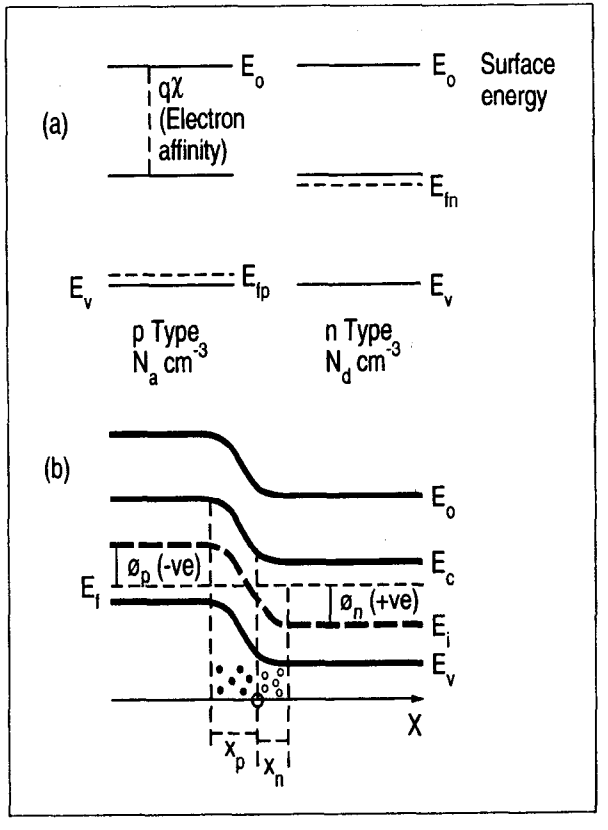


Fig. 19. (a) Energy levels in two silicon samples (of *p* and *n* type) when electrically isolated from one another. (b) When brought into contact, the Fermi level is constant throughout the material. The band edges bend in accordance with the space charge generated.

Intuitively, this can be understood as follows. Initially, the electrons at a particular level in the conduction band of the *n*-type material see energy levels in the *p*-type material at equal or lower energy which are unpopulated, so they diffuse into them. The developing space charge bends the energy bands up, so that these levels become inaccessible. Eventually, only very high-energy electrons in the *n*-type material see anything other than the absence of states of the band gap in the *p*-type material, and conversely for the holes in the *p*-type material.

Let us develop this quantitatively, adopting a coordinate system in which the *pn* junction of Fig. 19(b) is at position $x = 0$. E_0, E_c, E_i, and E_v all follow the same x dependence. The zero of the electric potential ϕ is arbitrary, so we define

$$\phi = -\frac{(E_i - E_f)}{q_e}. \tag{3.12}$$

Thus, ϕ is 0 for intrinsic material

positive for *n*-type

negative for *p*-type.

From Eq. (3.11), in the case of fully ionized donors and acceptors,

$$\phi_n = \frac{kT}{q_e}\ln\left(\frac{N_d}{n_i}\right)$$

$$\phi_p = -\frac{kT}{q_e}\ln\left(\frac{N_a}{n_i}\right).$$

The potential barrier

$$\phi_i = \phi_n - \phi_p = \frac{kT}{q_e}\ln\left(\frac{N_d N_a}{n_i^2}\right). \tag{3.13}$$

Notice that the potential barrier falls linearly with temperature since it is sustained by the thermal energy in the system. We may deduce the electric field strengths $\mathcal{E}(x)$ near the junction by using Poisson's equation

$$\frac{d^2\phi}{dx^2} = -\frac{d\mathcal{E}}{dx} = -\frac{q_e}{\varepsilon_s}\rho(x).$$

ε_s is the permittivity of silicon $= \varepsilon_r\varepsilon_0$.

ε_0 is the permittivity of space $= 8.85\times10^{-14}$ F cm^{-1}

$= 55.4$ e$^-$/V μm.

ε_r is the dielectric constant or

relative permittivity of silicon = 11.7.

For $x_n \geq x \geq 0$,

$$\left.\begin{array}{ll} \text{For } -x_p \leq x \leq 0, & \dfrac{d\mathcal{E}}{dx} = +\dfrac{q_e N_d}{\varepsilon_s} \qquad \therefore \mathcal{E}(x) = -\dfrac{q_e N_d}{\varepsilon_s}(x_n - x). \\ & \dfrac{d\mathcal{E}}{dx} = -\dfrac{q_e N_a}{\varepsilon_s} \qquad \therefore \mathcal{E}(x) = -\dfrac{q_e N_a}{\varepsilon_s}(x + x_p). \end{array}\right\} \tag{3.14}$$

The *undepleted* silicon on either side of the junction is *field free*. The depleted silicon close to the junction experiences an electric field whose strength is maximum at the junction and is directed always to the left, i.e., opposing the flow of holes to the right and opposing the flow of electrons to the left.

Requiring continuity of the field strength at $x = 0$ implies

$$N_a x_p = N_d x_n. \tag{3.15}$$

Thus, if one wants to make a deep depletion region on one side of the junction (important, as we shall see, for many detectors), we need to have a very low dopant concentration, i.e., very high resistivity material.

The electric field strength varies linearly with x; the electric potential, by integration of Eq. (3.14), varies quadratically.

For $x_n \geq x \geq 0$,

$$\left.\begin{aligned} &\phi(x) = \phi_n - \frac{q_e N_d}{2\varepsilon_s}(x_n - x).^2 \\ \text{For } -x_p \leq x \leq 0, \quad &\\ &\phi(x) = \phi_p + \frac{q_e N_a}{2\varepsilon_s}(x + x_p)^2. \end{aligned}\right\} \tag{3.16}$$

Requiring continuity of the potential at $x = 0$ implies

$$x_n + x_p = \left[\frac{2\varepsilon_s}{q_e}\phi_i\left(\frac{1}{N_a} + \frac{1}{N_d}\right)\right]^{1/2}. \tag{3.17}$$

From Eq. (3.13), ϕ_i depends only weakly on N_a and N_d.

If, for example, $N_a >> N_d$, we have $x_p \approx 0$ and Eq. (3.17) gives $x_n \propto N_d^{-1/2}$.

So a factor of two increase in resistivity leads to a factor of only $\sqrt{2}$ increase in depletion depth.

Figure 20 summarizes these results on the characteristics of an unbiased *pn* junction, with the inclusion of some typical numerical values based on $N_a = 10^{14}\,\text{cm}^{-3}$ and $N_d = 2 \times 10^{14}\,\text{cm}^{-3}$. The peak field in this case is about 3 kV/cm.

We now consider the effect of applying a voltage across the junction. Under equilibrium conditions, electron-hole pairs are continually generated by thermal excitation throughout the semiconductor. In the case of zero bias [Fig. 21(a)], the electrons and holes generated within the bulk of the semiconductor recombine. Those generated in the depletion region are swept into the undepleted silicon, holes to the left, electrons to the right. This effect would act to reduce the potential barrier and so is compensated by a small flow of *majority carriers* which find themselves with just sufficient energy to diffuse across the barrier in the opposite directions at just the rate needed to cancel the charge generation in the depleted material. The overall effect is of zero current flow, i.e., equilibrium.

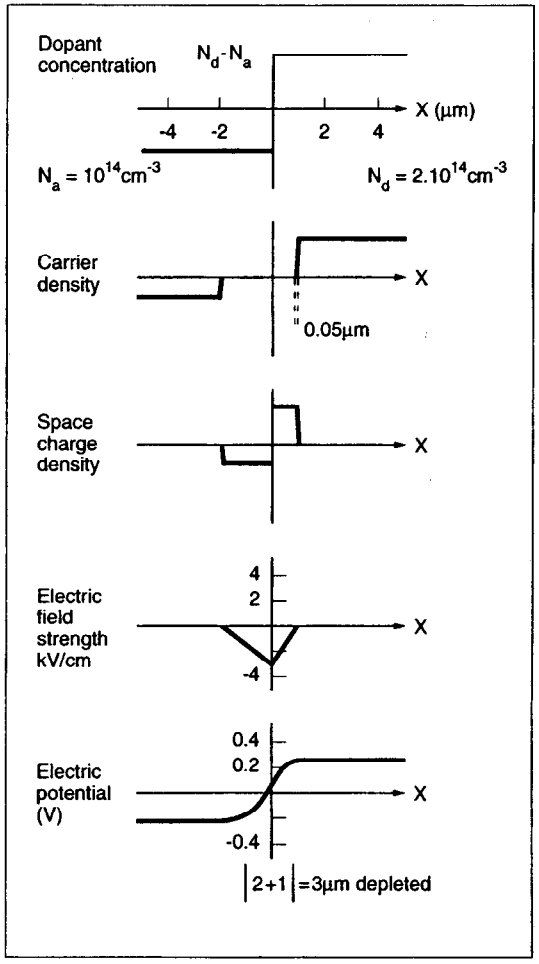


Fig. 20. Summary of various quantities across an unbiased *pn* junction.

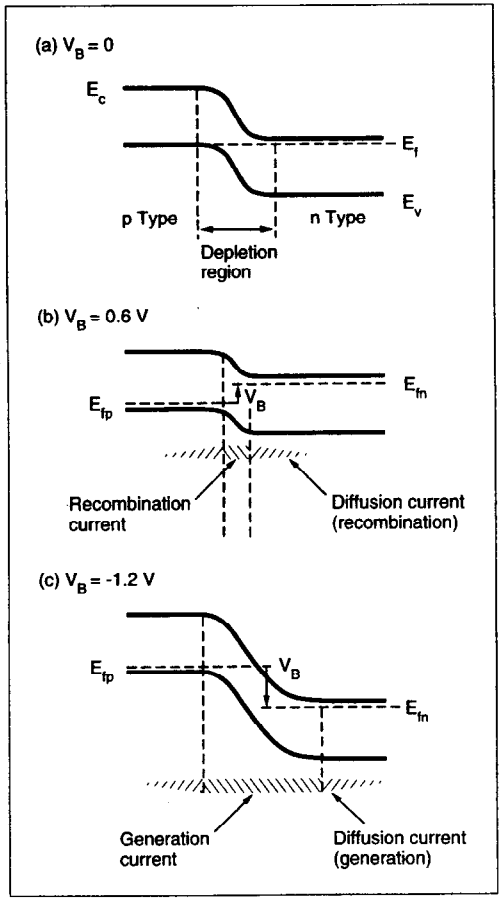


Fig. 21. Effect of an applied voltage across the semiconductor junction.

By applying a forward bias [Fig. 21(b)], we separate the previously equal Fermi levels by an amount equal to the bias voltage; the system is no longer in thermal equilibrium or this condition could not be maintained. Although there is still an electric field in the depletion region which is directed against the current flow, the depletion region is narrowed and the potential barrier is now inadequate to prevent majority carriers from flooding across it, holes from the left and electrons from the right. Many of these will recombine within the depletion region giving rise to the *recombination current.* Those which survive are absorbed within one or two diffusion lengths by recombination with the majority carriers on that side of the junction, giving rise to the *diffusion current.* Beyond these regions, there is just a steady flow of majority carriers supplied from the voltage source to keep the current flowing. Notice that in a forward-biased junction, the current flow results entirely in electron-hole *recombination.*

With a reverse bias, we have the situation shown in Fig. 21(c). The depletion region is now much wider and electron-hole pairs generated within it are efficiently swept into the undepleted silicon, electrons to the right and holes to the left, giving rise to the *generation current.*

Unlike the case of the unbiased junction, there is now no supply of majority carriers able to overcome the increased potential barrier across the junction. On the contrary, the thermal generation of *minority carriers* within one or two diffusion lengths of the depletion region leads to some holes generated in the *n*-region reaching this depletion region and then being briskly transported across it, and conversely for electrons generated in the *p*-region. This leads to the so-called *diffusion current.* In the case of the reverse-biased junction, the current flow is thus caused entirely by electron-hole *generation.* The current flow across reverse-biased junctions is of great importance in determining the noise limits in silicon detectors. An immediate observation is that since this current arises from *thermal* generation of electron-hole pairs, the operating temperature will be an important parameter.

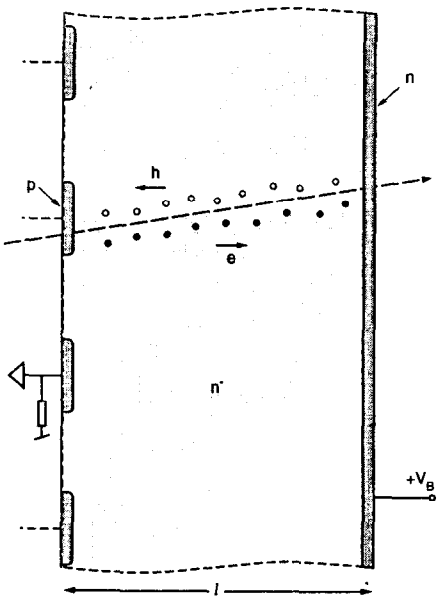


Fig. 22. Operating principles of a simple microstrip detector. It consists essentially of a reverse biased diode, operated at a voltage sufficient to fully deplete the thick, high-resistivity n layer.

Before continuing to discuss this point, it is worth noting that we have finally collected up enough information to calculate the characteristics of a typical particle detector, and it is instructive to do so. Referring to Fig. 22, we have a silicon detector of thickness ℓ made of good-quality, high-resistivity n-type silicon (ρ = 10 KΩ cm). On the front surface, we make shallow implants of acceptor atoms (the p strips), and on the back surface, we make a highly doped n-type implant to provide a good low-resistance ohmic contact. The terms n^+ and p^+ are conventionally used to represent high doping levels, n and p represent moderate levels, n^-, p^-, or π and ν represent low levels, and i is used for intrinsic or compensated material of the highest possible resistivity.

Now we apply a positive voltage V_B to the n-type surface with the aim of completely depleting the detector. In this way, we shall ensure complete collection of the electrons and holes generated by the passage of a charged particle; with incomplete depletion, we would lose signal by recombination. Equation (3.17) applies, with the difference that we replace ϕ_i by $V_B + \phi_i$ since the junction is biased in the direction which assists the previously existing depletion voltage.

We have

$$x_n + x_p \approx x_n = \left[\frac{2\varepsilon_s}{q_e}(V_B + \phi_i)\left(\frac{1}{N_a} + \frac{1}{N_d}\right)\right]^{1/2}$$

$$\approx \left[\frac{2\varepsilon_s}{q_e} \times \frac{V_B}{N_d}\right]^{1/2}.$$

From Fig. 17, we see that $N_d \approx \dfrac{6\times10^{15}}{\rho}$, and we require $x_p = \ell$

$$\therefore V_B = \frac{q_e}{2\varepsilon_s} \times \frac{6\times10^{15}}{\rho} \times \ell^2$$

$$= \frac{10^{-4}}{2\times55.4\times11.7} \times \frac{6\times10^{15}}{\rho} \ell^2 \times 10^{-8},$$

where ℓ is in μm and ρ in Ω cm

$$\therefore V_B = \frac{4.63\ell^2}{\rho}.$$

For the above example, V_B = 42 V is the potential needed to fully deplete the detector. We also find $\mathcal{E}_{max}$ = 2.8 kV/mm. This looks comfortable in terms of the breakdown field in silicon. However, in a real detector, it is important to pay

attention to the regions near the edges of the p strips, where the fields can be very much higher.

Returning to the general properties of the reverse biased junction, the most important parameter influencing the leakage current is the operating temperature. At high temperatures, above 100°C typically, the leakage current is dominated by thermal electron-hole generation within approximately one diffusion length of the depletion edge. The diffusion length for minority carriers is

$$L_D = \sqrt{D\tau_m}\,, \tag{3.18}$$

where D is the diffusion constant and is related to the mobility μ by the Einstein relation

$$D = \frac{kT}{q_e}\mu\,. \tag{3.19}$$

For electrons $D_n = 34.6\ \mathrm{cm^2 s^{-1}}$
For holes $D_p = 12.3\ \mathrm{cm^2 s^{-1}}$
} at room temperature.

τ_m is the minority carrier lifetime, and it can vary from about 100 ns to more than 1 ms depending on the quality of the silicon. This point will be discussed further. This leakage current (termed the diffusion current, as previously noted) depends only weakly on the reverse bias voltage but is highly temperature dependent due to its origin in the thermal generation of minority carriers.

At lower temperatures (less than about 100°C), the diffusion current becomes negligible and the generation current dominates. This continues to show a similarly fast temperature dependence, but is now also quite voltage dependent, since the depletion width is proportional to $V_B^{1/2}$.

The diffusion and generation currents depend on the rate of generation of electron-hole pairs, and the diffusion current depends also on the minority carrier lifetime. These quantities are, in fact, closely related. Direct thermal generation of an electron-hole pair is quite rare in silicon for reasons which depend on the details of the crystal structure. Most generation occurs by means of intermediate generation-recombination centers (impurities and lattice defects) near the band gap center. Thus, an electron-hole pair may be thermally created in a process where the hole is released into the valence band and the electron is captured by the trapping center in one step, to be subsequently emitted into the conduction band. These *bulk trapping states* vary enormously in their density and can be held down to a low level by suitable processing. It is precisely these states which determine the minority carrier lifetime already mentioned. Reducing the density of bulk trapping states does two things. It cuts down the thermal generation of charge carrier pairs in the material, so reducing the concentration of minority carriers available for the generation of current across a reverse-biased junction. It also increases the minority carrier lifetime and so the diffusion length (but only at $\tau^{1/2}$). The first effect vastly outweighs the second, so that a low density of bulk trapping states is highly advantageous in ensuring low leakage current. As we shall see later, even originally high-grade silicon can deteriorate due to the production of bulk trapping states by radiation damage. Mid-band gap impurities such as gold are a particularly serious source of bulk trapping centers. Even in low concentrations, gold atoms strongly reduce the carrier lifetimes and lead to greatly increased leakage current.

These effects are less serious in cases where one is collecting large signals promptly. But in cases of small signals and/or long storage times (such as in a silicon drift chamber, or CCD), particular care is needed. One important design criterion is to keep the stored charges well away from the surface of the silicon, since the silicon/silicon dioxide interface always has a high level of lattice defects. This criterion has led to the development of various forms of *buried channel* radiation detectors.

3.3 Charge Carrier Transport in Silicon Detectors

While the charge generated by an ionizing particle is being transported by the internal field in the detector, the process of diffusion spreads out the original very fine column of charge. In the case of very highly ionizing particles (such as alphas), the original density of electrons and holes can be so high that space-charge effects are important. In the case of MIP's, however, such effects are negligible and the time

development of the electron and hole charge distributions may be treated by simple diffusion theory.

Consider a local region of electron charge, for example, a short section of the particle track length within the silicon. Under the influence of the internal field, this will be drifted through the material, and at the same time, will diffuse radially as indicated in Fig. 23.

The RMS radius of the charge distribution increases as the square root of drift time t_d, as in Eq. (3.18), with standard deviation $\sigma = \sqrt{2Dt_d}$. Thus, 50% of the charge is contained within a radius of 0.95 $\sqrt{Dt_d}$. Assuming a "typical" drift field in depleted silicon of 1 kV/cm and using the fact that the drift velocity $V_d = \mu_n \mathcal{E}$, we obtain the following indication of the growth of a charge packet with time:

Drift Time	Charge Radius	Drift Distance
10 ns	6 μm	135 μm
1 μs	60 μm	14 mm.

Diffusive charge spreading is an attractive option for improving spatial precision beyond the limits of the detector granularity. For example, one might hope to achieve precision of one or two microns from a strip detector with 25 μm pitch, by centroid finding on the basis of measured charge collection in adjacent strips. This depends on achieving a charge radius of ≥ 30 μm which (from the above table) implies large drift distances and/or gentle drift fields. Ideas for improved precision by centroid finding may be limited by the available resistivity of silicon.

There is, however, an alternative approach that has so far been applied only to CCD detectors but which could be of more general interest. A wafer cut from a silicon crystal will normally have a rather uniform dopant concentration. It is possible subsequently to grow relatively thick (up to around 100 μm) *epitaxial layers* on the substrate wafer, of excellent crystalline quality and quite different (but also uniform) dopant concentration. For detector applications, a low-resistivity substrate with a high-resistivity epi layer is of particular interest. In the CCD case, as we shall see, the epi layer would be implanted with an n layer and biased so as to deplete only approximately 3 μm depth. The charge carrier transport associated with (for example) a charged particle track traversing such a structure is depicted in Fig. 24. Electrons within the thin depletion region are promptly collected into the buried channel, with no time for lateral diffusion. Electrons from the highly doped p^+ bulk are completely disposed of by recombination (very short minority carrier diffusion length in this material). However, electrons generated in the undepleted epitaxial layer find themselves able to diffuse homogeneously in all directions. Those which approach the p/p^+ junction experience a potential barrier as we have already discussed in the case of the unbiased pn junction, of magnitude

$$\phi_B = \frac{kT}{q_e} \ln\left(\frac{N_{p^+}}{N_p}\right).$$

For a 20 Ω cm epi layer on a highly doped 0.1 Ω cm substrate, we find

$$\phi_B = 180 \text{ mV compared with } \frac{kT}{q_e} = 26 \text{ mV}$$

at 300 K. The p/p^+ interface therefore acts as a *perfect mirror*, and the electrons continue diffusing until they happen to approach the pn depletion edge, at which point they are stored. Thus, a MIP leaves an electron charge cluster which is transversely spread by an amount related to the epi layer thickness. Such a detector made with partially undepleted thick epi material is in principle better for precision tracking by centroid finding than a fully depleted detector. To fully exploit this concept, one has to pay attention to the detector granularity, epi layer thickness, readout noise, etc. The most spectacular results in precision centroid finding in CCD's have been obtained not as yet with MIP's but with defocused star images in a satellite guidance system, where precision below 0.1 μm has been achieved using 20 μm pixels. This constitutes a very important demonstration of the inherent pixel-to-pixel homogeneity possible with high-quality silicon processing.

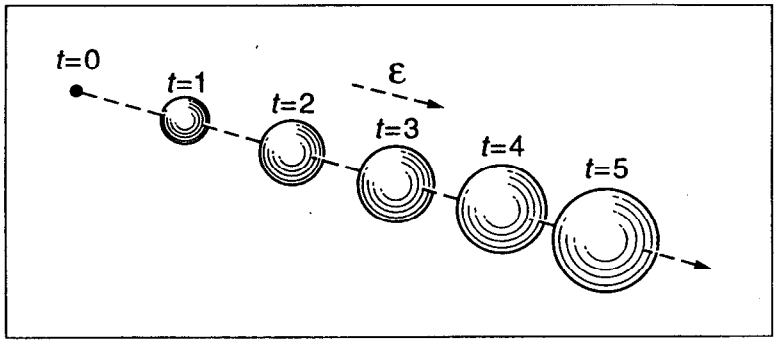


Fig. 23. Combined drift and diffusion of an initially compact charge cluster (electrons or holes) as a function of time over equal time intervals.

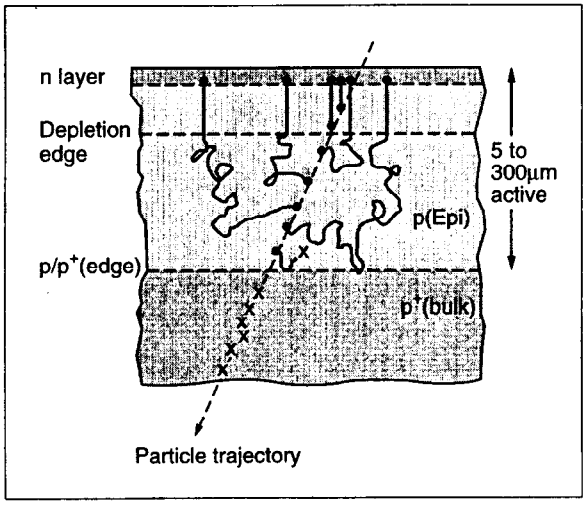


Fig. 24. Charge collection from a silicon structure as used in some pixel devices.

4 Microstrip Detectors

4.1 Introduction

Charged particles deposit a significant fraction of their energy by ionization in all types of materials, but only some are suitable as detector media. The conceptually most elementary detector types are insulators in which the signal is collected simply by applying a voltage to a pair of metal plates attached to the opposite faces of the detector layer, so creating an electric field within the material. The detection medium may be a gas (ionization chamber), a liquid (e.g., liquid argon calorimeter), or a solid (e.g., diamond detector). However, this principle cannot be applied to semiconductor detectors since even the highest purity material would generally have unacceptably low resistivity (i.e., excessive leakage current) except at extremely low temperature. As we have seen, it is possible to generate a region of internal electric field devoid of free charge carriers, and hence having greatly reduced leakage current, by creating a reverse-biased junction. Electron-hole pairs generated within the depletion region, for example, by thermal or optical excitation, or by the passage of a charged particle, are promptly swept to the surface for collection. This principle has been used for the detection of ionizing particles in silicon for over 40 years [8]. We have already noted some variations on this theme in connection with pixel devices (collection of minority carriers from undepleted material adjacent to depleted silicon), but the microstrip detector follows exactly this simple tradition.

The pioneering microstrip detectors of the early '80s (Ref. [9]) were based on the processes used for many years to manufacture nonsegmented semiconductor detectors for nuclear physics applications. The diodes were simply formed by the surface barrier between metal (aluminum) strips and the high-resistivity substrate. The strips were wire bonded to huge fanout boards which housed local pre-amplifiers connected to every Nth strip ($N \approx 5$). The principle of capacitive charge division was used to interpolate the track coordinates for signals collected on floating strips. The ratio of board area to detector area was almost 1000 to 1; this was tolerable in fixed-target experiments having unlimited space for local equipment outside the aperture of the forward spectrometer.

Closely following on these early developments, two revolutions took place which totally changed the technology of these detectors, opening up for them a much more powerful role in particle physics.

The first of these revolutions was to switch from surface barrier detectors to ion implantation, thus adopting the highly developed techniques used for processing integrated circuits. The microstrip detector becomes essentially a *p-i-n* diode structure, as we discussed in Sec. 3.2. The *p* strips (Fig. 22) were overlaid with metal (aluminum) to provide a low resistant path and connected to external electronics. This development [10] had been considered impossible by many semiconductor detector experts at the time. The high-resistivity material used almost uniquely by detector people was supposedly incompatible with the high-temperature processing required for the activation stage of ion implanted material. Kemmer showed that these experts were incorrect; it was problems of cleanliness in processing, rather than the high temperatures themselves, which led to the dreaded resistivity drops. The first result of this revolution was *more* robust detectors and hence the possibility of much larger areas. As important, the door was opened for the inclusion of a host of features already developed for IC's, such as techniques for isolating edge-related leakage currents (guard rings), for biasing with high dynamic resistance, and so on. Some of these will be discussed in Sec. 4.3.

The second revolution was the development of readout chips with high-density front-end amplifiers [11, 12]. Using integrated circuit technology, the front-end could be shrunk to a pitch of 50 μm, permitting the microstrip channels to be wire bonded directly to these compact IC's located along the edge of the detector. Furthermore, the readout chips embodied resettable storage of the analogue signals, and multiplexed readout. Thus, the number of cables needed for the detector readout was reduced by about a factor of 100. We shall in Sec. 4.3.3 record great ongoing progress in developing special readout IC's to suit a wide range of experimental conditions.

The combination of robust, sophisticated microstrip detectors and extremely compact electronics has led to their application in a host of experiments. With the SLC, Mark II, and LEP detectors, they crossed the barrier from fixed-target experiments into the collider environment, with excellent results in heavy flavor physics.

4.2 The Generic Microstrip Detector

Microstrip detectors come in a large variety of designs, each with its own strengths and weaknesses, each with a certain range of applications.

Due to the fact that high resistivity *n*-type material is more readily available, most detectors have used *n*-type wafers as starting material, though this may be changing in some application areas. The 111 crystal-orientation is conventionally used, but reasons why this too may be changing are discussed in the next section. As already mentioned, the pioneering detectors all used p^+ strips, collecting holes from the track of the ionizing particle. More recently, the back surface (n^+ implant) has also been subdivided into strips (which can as well be angled, perhaps at 90° to the *p* strips) giving us double-sided microstrip detectors.

Such a detector, and the associated internal electric field, is sketched in Fig. 25. The reverse bias is achieved by applying a positive voltage to the *n* strips, the *p* strips being grounded. In each case, series resistors (usually on-chip polysilicon) are used to create a high impedance path. The electric field (directed in the negative *Z* direction) would be uniform across the depleted n^- substrate, were it not for the finite resistivity and hence the presence of a low density of fixed positive charges. Due to this space charge, the magnitude of the field falls steadily from its peak value at the *pn* junction, towards the *n* side. The sketch shows an overdepleted detector. For the just-depleted case, the field would sink to zero at the surface of the *n* strips. Once we enter the heavily-doped *p*- or *n*-strip region, the field develops a large gradient, falling abruptly to zero.

The sketch indicates an AC coupled detector. The metal readout strips are isolated from the implanted strips by a thin layer of dielectric (silicon dioxide). Thus, the amplifier inputs sense the fast signal without also being obliged to sink the DC leakage current. Both AC and DC coupled microstrip detectors are common. In applications where radiation levels are low, and hence degradation in leakage current

is not a problem, the extra simplicity of DC coupled detectors may be advantageous. Early microstrip detectors were all DC coupled.

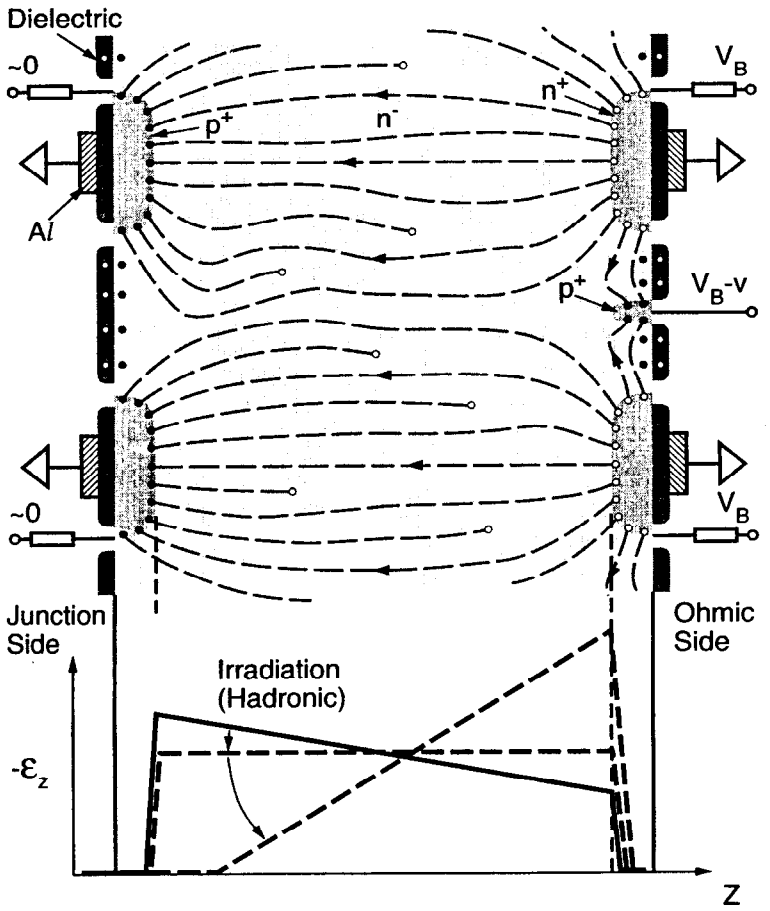


Fig. 25 Sketch of a cross section of a generic double-sided microstrip detector. Exposed fixed charges are shown by open circles (positive) and filled circles (negative). Also shown is the electric fields distribution in such a detector before and after radiation-induced displacement damage in the silicon.

Between neighboring charge collection strips on both sides is a passivation layer of silicon dioxide. Such oxide layers inevitably collect some positive charge (holes trapped as interface states) which is compensated by a very thin accumulation layer of mobile electrons in the bulk material. On the p side, these are repelled by the exposed negatively charged dopant atoms in the p strips. However, on the n side, they create a low-resistance interstrip leakage path. Signal electrons collected on one n strip will readily flow to neighboring strips; the strips are effectively shorted together. This problem can be overcome in a number of ways; Fig. 25 shows one of the cleanest solutions which is drawn straight from the textbooks of IC design. p^+ "channel stops" are implanted between the n strips. They are biased somewhat negatively relative to the strips, and hence, acquire a negatively charged depletion layer which repels the mobile electrons in the surface accumulation layer, so blocking the leakage path that would otherwise be present.

Before leaving this figure, there is one further point worthy of note, relating to the collection of signal charge. After the passage of an ionizing particle, holes begin to drift to the left, electrons to the right. Once the charges separate, the space-charge self-repulsion in principle leads to expansion of the charge cloud during the drift time. A localized charge distribution of N carriers (holes or electrons) will expand with time to a sphere of radius r_s, where

$$r_s = \left[4.5 \times 10^{-7} \frac{\mu N}{\varepsilon_s} t_d \right]^{1/3} \text{cm}.$$

ε_s is the permittivity of silicon, and t_d is the drift time in seconds. For collection of holes or electrons in a microstrip detector, r_s amounts to less than 1 μm and can be neglected (while the signal from an α particle can expand to $r_s \approx 10$ μm; see Ref. [9]). As we saw in Sec. 3.3, diffusive charge spreading can, on the other hand, be considerable. This is sensitively dependent on the type of charge carrier collected, on the detector resistivity, and on the biasing conditions.

For the conditions shown in Fig. 25, a strongly overdepleted detector, the electric field is reasonably uniform. For a just-depleted detector, the *holes* would all pass through the high-field region close to the *pn* junction, and those generated in that half of the detector would be entirely drifted through a fairly high field. For the *electrons*, on the contrary, all would pass through the low field region before reaching the *n* strips. Hence (even without the effect of the relative mobilities), the electron cloud will experience greater diffusive charge spreading than the hole cloud. In principle, this would give us higher precision (by centroid fitting) on the *n* side than on the *p* side. This question is discussed in more detail in the next section.

There are, however, several reasons why such fine tuning of detector parameters may not yield the desired improvement in precision.

Firstly, in a radiation environment, the effective dopant concentration varies with time. As depicted in Fig. 25 and discussed in detail in Sec. 6, hadronic irradiation causes the depleted material to become steadily more *p* type. Having passed through the compensated condition (when it could be depleted with a few volts), the resistivity falls steadily. After a certain dose (for fixed operating voltage), the detector would fail to deplete fully and the hole signal would be lost (no longer collected on an individual *p* strip). The electron signal would still be collected, but from a steadily decreasing thickness of detector. Thus, any precision advantage gained by fine tuning the depletion conditions could not be preserved through the life of the detector.

Secondly, due to their thickness, microstrip detectors have a significant probability of loss of precision due to δ-electrons, as discussed in Sec. 2.3. Results published from test beams often limit the signal charge to less than approximately 1.7 times the MIP mean value, in order to restrict the tails on the coordinate residuals. In tracking detectors with a limited number of points per track, one would not normally have the luxury of such a filter. For binary readout detectors, one would not even know which were the large signal clusters.

Thirdly, detector precision is seriously degraded for angled tracks, as we shall see in detail in the next section.

Finally, most tracking detectors in experiments operate in a magnetic field which (because of the Lorentz angle) degrades the measurement precision. In a conventional collider geometry with a solenoid magnet, the Z measurements are unaffected but the precision of the $R\phi$ measurement is degraded. For details, see the next section.

4.3 Microstrip Detectors: Detailed Issues

4.3.1 Design Optimization

All silicon microstrip detectors are of approximately 300 μm thickness. For much thinner detectors, the loss of signal charge, exacerbated by the reduction in signal voltage due to the increased capacitance from strip to substrate, results in a poor signal-to-noise performance. Even thicker detectors might be required, for example, in cases of modules having several long strips linked together and to a single readout chip. The capacitance to substrate is a particularly important issue in cases where capacitive charge division is used for the readout of floating strips. To avoid serious signal loss, it is essential that the geometry be chosen so that the interstrip capacitance greatly exceeds the strip-to-substrate capacitance, or one would suffer from serious loss of signal from floating strips. In some large systems currently under design (e.g., the ATLAS Silicon Central Tracker or SCT), the individual modules are 12 cm in length, with strip capacitances of around 18 pF (1–2 pF/cm is typical). Such large capacitances represent a considerable challenge for readout electronics, as we shall see in Sec. 4.3.3.

As already mentioned, a high-resistivity *n*-type substrate is conventionally used. High-resistivity *p*-type material is now available (both bulk and epitaxial), providing an interesting option for detector fabrication. Such detectors would have the advantage that under irradiation, they simply become steadily more *p* type. Thus, one would avoid the complications (e.g., in guard-ring structures) associated with the junction shifting over from the *p* side to the *n* side during the life of the detector.

The 111 crystal orientation is conventionally used in microstrip detectors, since it provides the densest surface, and hence the lowest probability of "spiking" (growth of aluminum deeply into the crystal in local regions, possibly shorting out the diode

structure). For IC manufacture (and also for MOS detector types such as CCD's), the 100 crystal orientation is preferred due to the lower density of dangling bonds at the silicon/silicon dioxide surface, and hence lower trapped charge at the interface. This may be particularly important in some microstrip detector applications, and for this reason some groups are doing exploratory work with 100 material. For AC coupled detectors, the area of metal in contact with silicon is reduced by many orders of magnitude compared to the early DC coupled devices. Also, metallization equipment is now extremely refined compared to 10 years ago, so the problem of spiking should be largely in the past.

For biasing microstrip detectors, the most commonly used method (also the simplest) is via on-chip polysilicon resistors. A problem with this approach is that as one has to allow for higher leakage current (due to radiation damage and/or longer strips), the resistance value needs to be reduced in order not to disturb the bias voltage excessively. This in turn can lead to loss of signal and worsening signal-to-noise ratio. The ideal solution would be a low DC resistance and a high dynamic resistance. Two approaches have been adopted, the reach-through structure [13] and the FOXFET biasing scheme [14]. This Field OXide FET structure, which employs a thick gate oxide, is vulnerable to radiation damage effects [15, 16]. The present situation appears to be that polysilicon biasing is the only safe solution for detectors to be used in a high-radiation environment.

For the *n*-strip isolation in detectors (one- or two-dimensional) where the electron signal is collected, two methods have been adopted. The channel stop approach [13] has been illustrated in Fig. 25. An alternative "field plate" method uses an MOS gate structure, in the form of "wings" attached to the aluminum readout strips in AC coupled detectors [17]. This is illustrated in Fig. 26.

For all these various microstrip detector structures, careful attention should be paid (by two-dimensional simulation) to the peak electric fields induced near the strip edges. Poorly understood leakage current has characterized many of the designs which at first glance looked quite reasonable. In a pioneering paper [18], Ohsugi and coauthors demonstrated the sensitivity to geometrical details in the specific case of AC coupled *p*-strip sensors. Breakdown was demonstrated in structures where the

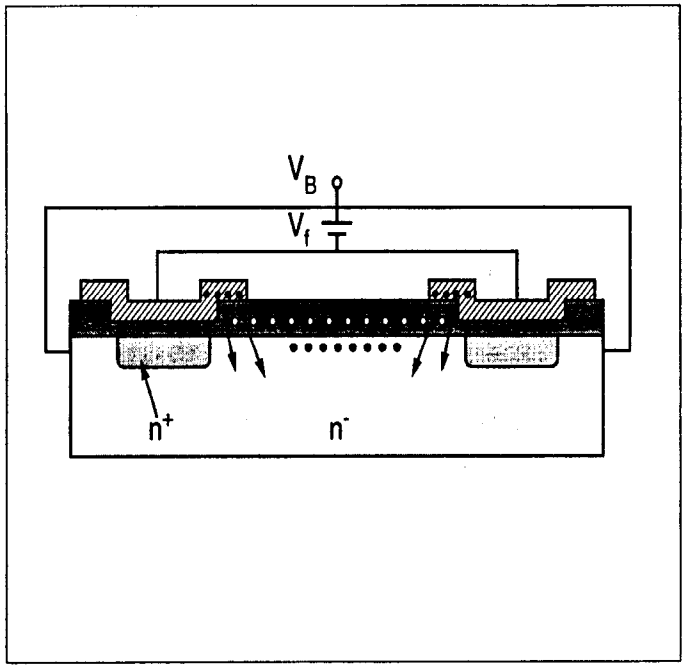


Fig. 26. The technique of *n*-stop isolation by field plate separation with extended AC coupled electrodes (one of several field plate approaches).

relative edges of the p^+ implant and the aluminium electrode led to peak fields at the edge of the implanted strips exceeding the breakdown field in silicon of 30 V/μm. While such problems can in principle be avoided by careful design, it is very easy to encounter some local variations, edge effects at the ends of the strips, etc., which can still cause problems. To this end, the diagnostic tool demonstrated in their paper is of enormous value. Using an infrared microscope equipped with a CCD camera, very small regions of avalanche breakdown can be seen clearly. This marvellous tool [19] is of value wherever anomalous leakage currents are encountered either due to design deficiencies or to process faults. One of the problems that has plagued manufacturers of large area microstrip detectors, particularly in the case of double-metal structures (see below), is that of pinholes in the dielectric, permitting unwanted leakage paths. An infrared microscope can be used to explore the positions of these defects, and possibly to suggest solutions (e.g., improved step coverage across gate edges). Similar problems have been encountered and solved in this way in the world of CCD detectors. For *n*-strip microstrip detectors, there is evidence (not surprisingly) that field plate devices are more susceptible to microdischarges than *p*-stop devices. However, much depends on the specific design details.

It is hardly surprising that another issue which still causes problems in microstrip detector design is that of uncontrolled oxide layers (e.g., interstrip, as depicted in sketch form in Fig. 25). In other detector types such as CCD's, care is taken to avoid even fine cracks between gate electrodes (by overlapping neighboring electrodes) since gate oxide inevitably contains trapped interface charge, the magnitude of which increases with irradiation. The electrical effects of such trapped charge can be minimized by the presence of a metal or polysilicon cover layer held at a well-defined potential. Microstrip detectors do not easily lend themselves to such design rules, but one may escape from trouble due to the accumulation layer of electrons already referred to. However, particularly if one is aiming for high efficiency for detection of (say) soft X-rays which deposit their signal near the surface, there are numerous examples of anomalous dead layers and other effects probably related to the uncontrolled oxide. This is an area for ongoing concern regarding the design of microstrip detectors.

The use of high resistivity silicon is driven by the desire to have a manageable operating voltage for full depletion; 150 V is commonly considered an upper limit. Under intense hadronic irradiation, this may set an uncomfortably short lifetime for the detector. It has been pointed out [20] that careful design of microstrip detectors (particularly as regards implant profiles, strip edges, guard rings, etc.) may enable operating voltages to be set even above 1 kV before microdischarges or breakdowns occur. Such a design would considerably extend the usefulness of microstrip detectors in high-radiation environments. Note that it is usually the breakdown voltage rather than the leakage current which shortens the lifetime of a detector in a radiation environment. The leakage current can always be reduced by cooling. There is long experience of this in the area of CCD detectors, and large systems of cooled microstrip detectors are now in the planning stages [21].

We have discussed briefly the availability of double-sided detectors, which are of interest in that they provide apparently two advantages over (for example) a pair of single-sided detectors: firstly, less material (of particular significance for vertex detectors), and secondly, some degree of resolution of the ambiguity problem for multihit events. Regarding the latter, the idea is that one can measure the signal charges in the *p*- and *n*-side clusters and use the correlation between them to rule out some of the associations (e.g., between a below-average cluster in one view and a multi-MIP cluster in the other view). In fact, this is not a very practicable idea, since the level of ambiguity is not greatly reduced.

Regarding extraneous material in the active volume, much depends on the angle between the strips on the two sides. If this is small (e.g., a few degrees), both sides of the module can be read out from the end without complications. If, however, one requires a large angle between the two strip planes (e.g., 90°), there are two options. Consider the case of a Z view as well as the conventional $R\phi$ view in a collider environment. The most obvious option, implemented in the pioneering double-sided ALEPH vertex detector [22], would be to place the Z readout chips along the long edge of the module. This results in a large amount of electronics in the active volume of the barrel detector system, which is not a good idea if precision vertexing is the goal. Later detectors have followed one of two different approaches. Both move the Z readout chips to the ends of the barrel, outside the active volume, in the same

general area as the $R\phi$ readout chips. The most ambitious approach is to integrate the linking traces onto the detector modules themselves, using a double-metal technology [23, 24]. A dielectric layer separates the Z-strips from the orthogonal readout strips, and metallized vias provide the connections between the two levels. Due to the larger number of Z-strips than readout strips in a typical module (a long rectangle), the Z-strips may be connected in a repeating pattern, resulting in some degree of ambiguity as to the spatial position (normally not a problem given the overall track-finding software). Alternatively, the Z-strip pitch may be made correspondingly coarser than the pitch of the $R\phi$ readout strips. There is one inevitable disadvantage to the double-metal approach, which is the increased capacitance of every strip; the detector strips and readout strips form a web of closely linked electrodes, separated only by the thickness of the dielectric layer. This, coupled with the fact that tracks at the ends of the polar angle range may deposit their charge over a number of Z-strips, can lead to a serious degradation in the signal-to-noise in the detector. The capacitance problem can be greatly reduced, with only a modest degradation in terms of material in the active volume, by routing the readout traces on separate thin substrates (e.g., copper traces on kapton) [25]. The Z-strips are wire bonded to diagonal readout strips at the edge of the detector, the signals being carried to the electronics in a zig-zag geometry, using additional Z-strips to link the diagonal readout strips. This idea is illustrated in Fig. 27. In this way, a low and acceptable ambiguity level as to which of a few widely separated strips was hit, is the price paid for accessing the data in an economical form with little additional material, and a generally acceptable overhead in capacitance.

There remains the choice between double-sided detectors and two back-to-back single-sided detectors, one for $R\phi$ and one for Z. As has been noted, the correlated cluster signal information is not often very useful, so the key issue is that of the additional material in the back-to-back approach. In vertexing applications, this is always important, though seldom decisive. There is necessarily additional material in the form of support structures, etc., so we are certainly not talking about a factor of two, and the multiple scattering is proportional to the square root of the thickness. If the double-sided option came free of additional costs, it would clearly be preferred. However, this is far from the case. Double-polished silicon wafers are available and

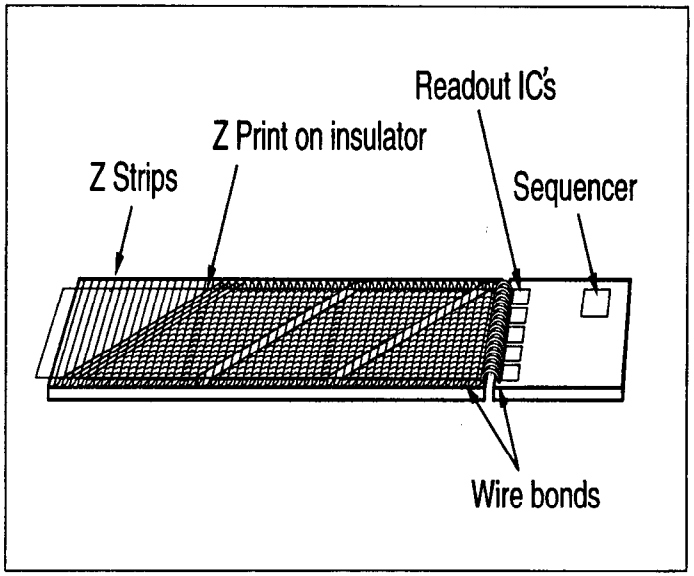


Fig. 27. A scheme for Z-strip readout using a separate metallized substrate (glass or kapton).

are not in themselves particularly expensive. However, for bulk production, it is desirable to use as far as possible the standard IC manufacturing equipment, which is all explicitly geared to single-sided processing. It has been claimed that the cost of double-sided relative to single-sided detectors is 3:1. This may be true for some small production runs, where it merely reflects the reduced yield of the double-sided devices. However, for large-volume production such as we are now seeing planned (e.g., for the LHC SCT's), it should be possible to greatly reduce the cost per unit area of detectors made with standard processing equipment. In this case, the cost ratio mentioned above is likely to become much more unfavorable. Time will tell.

4.3.2 Spatial Precision in Microstrip Detectors

Early microstrip detectors with very fine readout pitch (and huge fanout factors) had wonderful spatial precision but are now only of historical interest. We are at present effectively constrained by the readout pitch of all existing front-end electronics, namely 50 μm. This can be reduced by a factor of two by attaching readout IC's at each end of a module, and this has been done in environments of high track density. Also, one can include floating strips as has already been discussed. Spatial precision of approximately $\frac{25}{\sqrt{12}}\mu m = 7.2\mu m$ is thus in some ways natural for a silicon microstrip detector when read out with currently standard electronics. In large tracking systems, one has frequently to work very hard to achieve such levels of stability and systematic precision, for many reasons. Having said this, considerably better spatial precision has been achieved, mostly in test- beam situations.

Let us consider first the case of normal incidence tracks. As we saw in Sec. 4.2, the extra diffusive spreading would suggest that (for a given strip pitch) one might be able to achieve a higher precision in the charge collection on the n side (electrons) as opposed to the p side (holes). However, most experimental results to date have been obtained with detectors made with p strips on n-bulk silicon.

Using a single-sided detector with p strips on a 20 μm pitch and analogue readout on every strip, Belau *et al.* [26] were able to measure the *spatial distribution* of the hole charge collected. This varied from σ=2.5 μm to 1.9 μm as the operating voltage was raised from 120 V (just-depleted) to 200 V (overdepleted). From this, they *calculated* the precision achievable with a readout pitch of 20, 60, and 120 μm to be σ= 2.8, 3.6, and 5.9 μm, in the optimal case of the just-depleted detector. Measurements with 60 and 120 μm readout pitch [27] yielded precisions of 4.5 and 7.9 μm, a little worse than calculated, but better than $\frac{20}{\sqrt{12}}$ = 5.8 mm which would be the limit for a digital system with 20 μm readout pitch. Evidently, some degree of useful charge spreading is achieved with detectors having narrow strip pitch. For electron collection, the lower average electric field yields even better *calculated* precision, 0.8 μm to 3.6 μm, for the three cases mentioned above. In this case, they did not have data for comparison.

In all this, please remember the caveat about δ-electrons mentioned in Sec. 2.3. In these test beam studies, clusters with more than 1.7 times the mean MIP signal were discarded, with the consequential efficiency loss that could probably not be tolerated in a detector used for a physics experiment.

Results with a more typical arrangement of readout of every strip on a pitch of 50 μm have been reported for double-sided detectors [28]. For normal incidence, the precision achieved was 8.8 μm on the p side. This slightly worse figure is attributed to the higher electronic noise in that system. The signal-to-noise was 16 μm for the p side and degraded (for not completely clear reasons) to ten for the n side. The precision for the n-side signal was 11.6 μm, confirming the suggestion that the system noise played a large part in the measured spatial precision.

For normal incidence tracks, we may conclude that spatial precision in the region 5–10 μm is typical for strip pitch $\leq$ 50 μm, and with readout pitch $\leq$ 150 μm. The degradation in precision with increasing readout pitch is fairly modest. The usual reason for requiring a fine readout pitch (typically, equal to the strip pitch) is the need to preserve an optimal two-track resolution.

Once we permit angled tracks (which really only are of concern for the RZ view as opposed to the $R\phi$ view in colliders), the situation deteriorates fairly rapidly. The particle leaves a trail of charge carriers which are collected on a number of Z strips.

Taking the overall centroid is a bad approximation to the track position at the center plane of the detector, due to the energy-loss fluctuations along the track. The problem has been studied theoretically [29] and experimentally [30], as a result of which Hanai *et al.* have developed an algorithm ("convoluted Gaussian centroid") which leads to an experimental precision as a function of α, the track angle to the detector normal, varying from 8 μm ($\alpha = 0°$) to 40 μm ($\alpha = 75°$). These results were obtained using a single-sided p-strip detector with 25 μm strip pitch and 50 μm readout pitch.

A dangerous factor affecting spatial precision in microstrip detectors is the effect of magnetic fields. Empirical measurements have been reported in Ref. [26]; these agree well with calculations. For the p-strip signal in a just-depleted detector, a magnetic field of 1.7 T shifts the measured co-ordinate by about 10 μm and increases the width of the collected charge distribution from 5 to 12 μm. The relevant parameter determining these effects is the Hall mobility μ_n^H for electrons and μ_p^H for holes; see Shockley [7]. With the usual arrangement in collider barrel detectors (magnetic field $\mathcal{H}$ perpendicular to electric field), the charges drift at the Lorentz angle θ^L with respect to the electric field, where θ^L is almost independent of the magnitude of the electric field and is given by

$$tan\theta_{n,p}^L = \mu_{n,p}^H \times \mathcal{H} .$$

Now

$$\mu_p^H \approx 310 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$$

and

$$\mu_n^H = 1650 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}.$$

For a typical case of a magnetic field of 1.5 T and a 300 μm thick detector, the charge distribution of the holes shifts by $\approx 7\,\mu$m, while that for the electrons shifts by $\approx 37\,\mu$m [31]. Thus, collection of the *electron* signal in future collider experiments is liable to serious systematic effects, unless the n-strips are oriented at least approximately along the direction of drift induced by the magnetic field (the $R\phi$ direction in a barrel detector).

Finally, a reminder that in any silicon detector of thickness approximately 300 μm, the production of δ-electrons of significant range is quite a common occurrence, so the residual distributions will inevitably have a significant non-Gaussian tail, unless steps are taken to exclude large-signal clusters, with the attendant loss of efficiency.

4.3.3 Electronics for Microstrip Detectors

We have seen that silicon microstrip detectors have developed and diversified to an extraordinary degree, due partly to the ingenuity of those involved, and partly to the tools and devices provided for them by the integrated circuit industry. As regards the readout electronics, the progress has been at least as spectacular, for the same two reasons. The current picture is in fact one of somewhat bewildering complexity, since the diversity of options is so great. Part of this diversity reflects the variable detector applications, but even for one single application (e.g., the ATLAS SCT), there is not yet unanimity among the experts as to the optimal approach. The issues are quite subtle and the boundary conditions keep shifting. In this section, we shall aim to take a general look at the principles leading to these various options and make some remarks about the relevant areas of application. What is clear, however, is that the ASIC designer now has enormous power and flexibility at his disposal, so that a new application area is likely to lead to the very rapid evolution of one or more new readout schemes full of wonderful ideas to handle the peculiarities of that particular application.

Even from the very beginning of the ASIC initiative which opened the door for silicon microstrip detectors to find a home in collider detectors, there was not a unanimous approach. At that time, there *was* unanimity at the level of the functional requirements (amplifier, sample-and-hold, multiplexed analogue output) but two technological solutions; nMOS [12] and CMOS [11] were pushed by different groups. In the event, the "low and slow" CMOS solution proved superior, largely due to the much lower power dissipation (around 2 mW per channel compared with ten times that for nMOS). This pioneering CMOS chip, the first of a family of CAMEX chips, was joined by others, of which the most commonly used are the MX

(3-7) (Ref. [32]), SVX (1-3) (Ref. [33]), and AMPLEX [34] families. More recently, a bipolar chip for the front-end electronics has made its appearance [35].

Why is the user of silicon microstrip detectors faced with such a large array of readout options? Some part of the reason is sociocultural. There never was a "standard" drift chamber preamplifier; different laboratories like to do their own thing, and this competition is extremely healthy in encouraging new ideas. But mostly, these various approaches have been driven by the need to equip detectors to work in increasingly varied and hostile conditions. Beam-crossing intervals at SLC (8 ms) and LEP Phase 1 (22 μs) allowed very relaxed shaping times of 1 or 2 μs. The detector modules were small (strip lengths ≤ 6 cm) and the radiation environment almost nonexistent. Under these benign conditions, the ASIC designers were able to achieve spectacularly good signal-to-noise from a variety of single- and double-sided detectors. Moving to the Tevatron (originally 3.5 μs, upgrading to 396 ns and eventually 132 ns), HERA (96 ns), and in the future, the SLAC and KEK *B* factories (4 ns), and LHC (25 ns) represents a phenomenal challenge. Compounded with the escalating beam-crossing rate is the need to increase the module sizes (strip lengths of 12 cm will be used in the large ATLAS SCT, for example), plus the fact that the detectors at all hadron machines will encounter significant, if not fatal, radiation damage. Some relief is provided by cooling the detectors to reduce leakage current, but for the most part, it has been up to the chip designers to get the physicists out of a very uncomfortable situation. This is a rapidly evolving story, and it is far from clear where we shall end up. In the case of the LHC detectors, several critical decisions have to be taken over the next year, and these will be based on the results of much hard work going on in design labs and in test beams. Let us review in very general terms the main approaches, all of which are certainly appropriate, for some applications.

Firstly, the generic analogue chip comprises typically 128 channels, one of which is shown in its essentials in Fig. 28. The amplifier/shaper may include a CR-RC filter. It has been shown [36] that more sophisticated filtering schemes do not lead to a major improvement in noise performance. On receipt of a trigger, the signals are sampled and stored on capacitors C_s, which are read out (sequentially for each channel on the chip) via the analogue output, for remote digitization. Such a readout chip minimizes the logic local to the detector (and hence, is optimal from the viewpoint of power dissipation, which is usually an important issue), but it cannot be used in high-rate environments where even the first-level trigger appears after several beam crossings. The most obvious response to this situation is firstly to reduce the shaping time so as to retain an analogue signal which is unambiguously associated with its beam crossing. However, this causes inevitably a penalty in noise performance and may not be necessary. Given the sparsity of the tracks in the detector, each strip has a low probability of being hit on successive beam crossings. Then one may retain a longer shaping time and use a filtering approach [37] to recover the fast timing information by deconvoluting the sampled voltages of a shaped pulse, to retrieve the original impulse signal with high precision. This ingenious approach may extend the range of applicability of CMOS front-end electronics into the realm of LHC operating conditions, and has been adopted by the CMS Collaboration. Their analogue signal (50 ns shaping time) is sampled at the beam crossing rate of 40 MHz. The samples are stored in an analogue pipeline of 128 cells, and if a positive level-1 trigger signal is received, are deconvoluted by the analogue signal processor. All this happens in parallel for each channel.

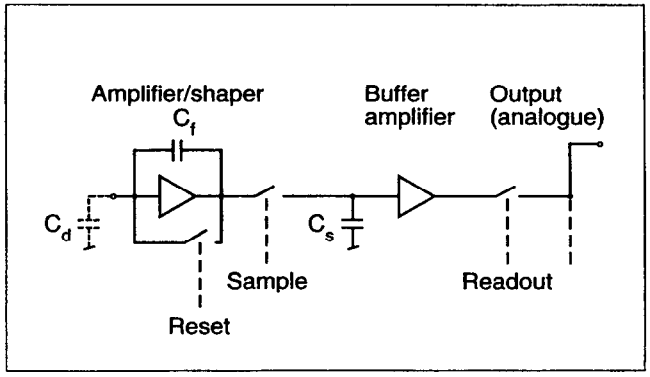


Fig. 28. Block diagram of one channel of a typical analogue readout chip.

The stored signals are read out at leisure via a multiplexer, connected off-chip to an electro-optical modulator. This consists of a multiquantum-well device which amounts essentially to a mirror of voltage-controlled reflectivity. Consuming almost no power, this device permits a change of reflectivity from 30% to 60% by changing the voltage across an InP/InGaAs sandwich [38]. The device is connected to an optical fiber, at the remote end of which is the drive laser, receiver module, flash ADC, and event builder memory. The beauty of such links is that they permit very high-speed data transmission with almost no power dissipation at the detector end. Used (as here) in analogue mode, they permit seven-bit resolution which is entirely adequate for microstrip detector applications.

The SVX family of readout chips has pioneered the digital approach. An example is shown in Fig. 29. Analogue signals are again put into a pipeline (one per channel). On receipt of a level-1 trigger, the relevant signal is transferred to a storage capacitor which serves as one input to a comparator used as a Wilkinson ADC circuit. The other comparator input is ramped at a fixed rate, and the time to reach equality of input is stored digitally as a measure of the signal amplitude. The digital data are then read out via a multiplexer.

Finally, we consider the bipolar option. Bipolar IC technology has been making great strides in recent years, and it has become possible to shrink amplifiers down to a pitch of 50 μm, as has been true for some time with CMOS systems. As a result, stray capacitances have been greatly reduced, and furthermore, very small transistors can be made with high bandwidth and low current. In short, the power dissipation has dropped to an extremely competitive level. At hadron colliders, even with cooled detectors, the problem of leakage current in long-strip modules after a few years of radiation damage will be considerable. The shot noise associated with the leakage current tends to favor short shaping times as opposed to the longer shaping time with deconvolution mentioned previously. The lower limit on the useful amplifier shaping time is given by the charge collection time of typically 20 ns. Below that, one encounters increasingly severe signal loss (the *ballistic deficit* effect). The superior transconductance of the bipolar transistor compared with CMOS (even if run in the weak inversion mode) suggests that to achieve adequate signal-to-noise performance

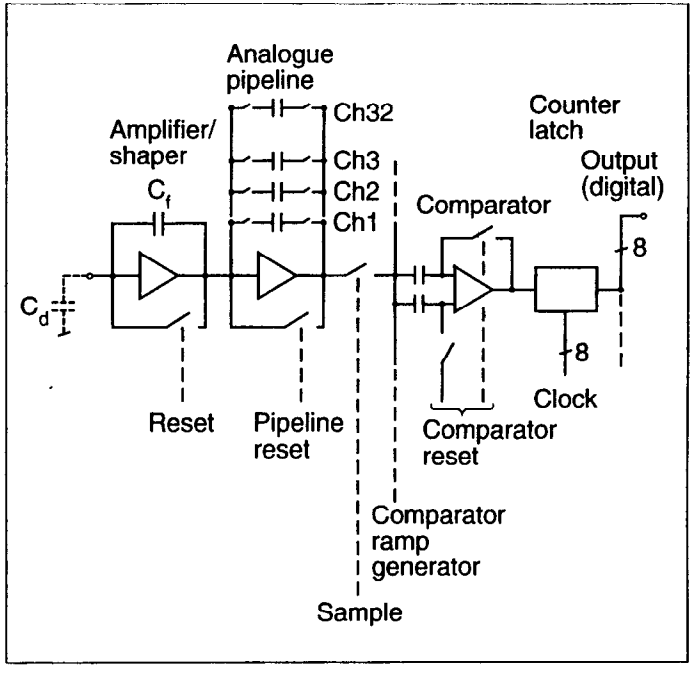


Fig. 29. Block diagram of one channel of a typical digital readout chip, of the SVX type.

for long microstrip modules in fast readout conditions, the bipolar option may be superior.

A disadvantage (possibly minor) of the bipolar approach is that (due to the near non-availability of rad-hard bi-CMOS) one necessarily has an analogue chip followed by a digital CMOS readout chip. Doubling the number of wire bonds in the system is not a major overhead, and there are advantages. For LHC applications, the size of the digital processing chip is such that yield is a significant consideration. Having the analogue front-end as a separate chip may be more economical overall.

This bipolar/CMOS combination has been used with excellent performance in the demanding environment of the ZEUS Leading Proton Spectrometer (LPS) at HERA [39, 40]. The basic system (Fig. 30) consists of a bipolar amplifier/comparator chip with 20 ns risetime, followed by a low-power digital pipeline. Not only does the front-end break with tradition in microstrip readout systems, but so does the digital system. The designers have adopted the simple "binary" approach of recording only the addresses of above-threshold strips, not the pulse heights. In fact, their system (which has been carefully designed to minimize common-mode noise) operates extremely stably with a constant threshold of 0.78 fC set for all channels.

Lack of pulse height information, of course, limits the spatial precision to $\frac{P}{\sqrt{12}}$, where p is the strip pitch, but as we have seen, this precision is in any case close to the limit achieved in nearly all systems. Furthermore, it is only in small radius vertex detectors that there are major physics advantages in pushing the point measurement precision to the highest achievable value.

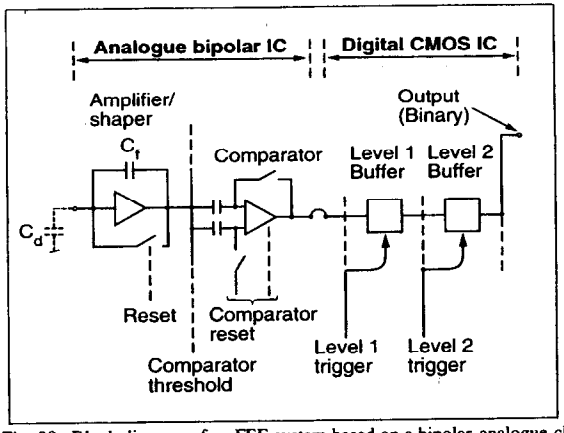


Fig. 30. Block diagram of an FEE system based on a bipolar analogue chip followed by a digital readout chip (binary readout).

The readout system takes advantage of the hierarchical trigger structure of ZEUS, which will also be followed in LHC. In the ZEUS application, they use a synchronous level-1 buffer of about 6 μs followed by an asynchronous level-2 buffer. Data are thus stored on-chip until a valid level-2 trigger arrives after about 1 ms.

All these considerations of readout options are complicated by another question, that of radiation damage. The move to hadron colliders has stimulated a major effort to develop rad-hard versions of the local detector electronics.

In the case of CMOS, a number of companies (Harris, UTMC, Honeywell, and DMILL) are involved with the chip designers already mentioned. For example, a 100 Mrad-hard version of the MX7 chip in 1.2 μm CMOS exists. These chips tend to somewhat exceed the 50 μm channel width, but for applications such as the LHC SCT's, this is acceptable. One cloud on the horizon is that, with the downturn in military spending, there is less funding for development of rad-hard electronics. However, as the industry moves into submicron processing, the devices have improved radiation resistance as a by-product (thinner oxide, etc.), so the trend may be to add a few steps to achieve adequate performance from a process not specifically developed for optimal radiation hardness.

For the bipolar IC's, the radiation damage situation is more favorable, due to the lack of sensitivity to oxide charge. The main cause of deterioration is bulk damage, which results in a reduction of the current gain β at high doses. Typically, an *npn* transistor suffers a β degradation of approximately a factor of two after 5 Mrads. The circuit designer can allow for such degradation, which makes these IC's usable at all but the smallest radii needed for vertex detectors at LHC. This region (as we shall see) is territory almost certainly out of bounds for silicon strip detectors due to the radiation damage in the detectors themselves.

Very recently, one company, DMILL (LETI), has produced some bi-CMOS chips using a rad-hard process. Whether they will find a sufficient market to sustain this initiative, and if so, whether these will offer a way to the future for HEP detectors, remains to be seen. At least for the time being, the combination of bipolar chips with rad-hard CMOS digital chips appears to be the safest means to satisfy our requirements.

Thus, in conclusion, both the CMOS and bipolar IC's we have discussed can, it appears, be designed to tolerate the worst radiation conditions likely to be encountered by silicon microstrip detectors. The inevitable noise degradation due to growth of leakage current in the detectors, plus other detector-related issues to be discussed in Sec. 6, are what finally limit the scope for these detectors. There is no possible cure in the electronics for these deficiencies, once they reach an unacceptable level in the detectors.

4.4 Physics Performance and Future Trends

Silicon microstrip detectors were originally developed as vertex detectors for charm physics at fixed-target experiments. Here, with the benefit of the high track momenta, they were able to achieve excellent impact parameter precision, and hence, clean reconstruction of a wide range of charm particle decays.

The move to e^+e^- colliders (initially SLC and LEP) called for much larger detector areas (and here the detector manufacturers were well able to oblige) and much more compact electronics (and, as we have seen, the ASIC designers solved these problems for us). Nevertheless, the physics capabilities of the detectors took a step backwards. Due to the lower particle momenta and the large radius beampipe (initially 10 cm at LEP, eventually reduced to 5.5 cm), the impact-parameter precision for tracks in hadronic jets was relatively poor. Nonspecialists were at first understandably ignorant of the situation, because all groups were proudly demonstrating beautiful miss-distance plots based on muon pairs. The situation for tracks in jets was, of course, much worse. The Holy Grail for vertex detectors is to present to the experimentalist a clear topology of the event, with high efficiency for associating all tracks uniquely with their true parent vertices. Fortunately for the detector builders, there is a host of lesser objectives which are still extremely useful for physics. The long lifetime of beauty particles means that b tagging is relatively straightforward. The cleanliness of the $\tau^+\tau^-$ signal at the Z^0 means that lifetime measurements (though imprecise at the individual event level) can be made accurately, given high event samples. Areas such as charm tagging and the separation between charged and neutral B states are much more challenging.

The one- and two-dimensional microstrip vertex detector systems at LEP have covered the range of radii typically 60 to 110 mm, and (in their finally upgraded forms) delivered precision in impact parameters as a function of momentum p GeV/c of:

$$\sigma^b_{XY} \approx 20 \oplus \frac{80}{p\sin^{3/2}\theta}\ \mu\text{m}$$

and

$$\sigma^b_{RZ} \approx 30 \oplus \frac{80}{p\sin^{3/2}\theta}\ \mu\text{m}.$$

With an average track momentum of about 5 GeV/c, this implies a mean impact parameter precision for normal incidence ($\theta = 90^\circ$) of around 30 μm. For reasonable topological vertexing (including charm), one would like to do five to ten times better than this. Another problem for the extraction of physics with microstrip detectors is that of poorly understood tails on residual distributions. These are presumably due to a combination of factors such as energy loss fluctuations, δ-electrons, cluster merging (by no means negligible in the core of jets), and so on. The general approach has been to artificially broaden the Monte Carlo residual

distributions to match the data. This is only a partial solution since it ignores the correlations that are surely present (e.g., a pair of tracks closely spaced in one view, giving poor coordinates on *all* planes due to partial merging of clusters).

The overall picture is one of impressively high efficiency and purity for b tagging, with flagging performance in the more challenging areas. For LEP2, the b-tagging requirement is considered to be of paramount importance (Higgs and SUSY searches). To do better as regards topological vertexing at the Z^0 would have required a smaller beampipe, giving a shorter extrapolation length to the interaction point (IP), and hence better impact parameter precision. But then, the track merging problem on the inner barrel would have been more severe. In any case, the time for such discussions is past.

The pioneering silicon microstrip vertex detector at hadron colliders has been the SVX family (same name as their readout chips) at the Tevatron. SVX1, the original detector, played a crucial role in the discovery of the top quark, again by performing the relatively simple task of b tagging. This is the first major discovery in particle physics in which a silicon vertex detector has been essential in achieving a convincing signal, and I am sure it will not be the last. After all the technical complications we have been discussing, it is somewhat comforting to note that the detector used for the top discovery was a single-sided, DC coupled, low signal-to-noise, radiation-soft detector. Such a detector would never survive the conditions after the Tevatron upgrades, and this vertex detector has already been upgraded at least once.

New microstrip vertex detectors are on the way. CLEO II has one (on a small-radius beam pipe, necessarily tackling the more challenging requirements of charm decay), and the SLAC and KEK B factories are building them, primarily to measure the longitudinal position of the B and $\overline{B}$ decay points with respect to the IP.

The ZEUS LPS set the trend for microstrip detectors to be used as momentum spectrometers in high-intensity conditions in which wire chambers could not survive. This trend continues with the D0 silicon tracker ($\approx 5\ m^2$) and the CMS and ATLAS SCT's ($40\ m^2$ for ATLAS). What has happened is that the high-energy, high-luminosity hadron collider environment has become too unfriendly for wire chambers on almost any radius. Therefore, silicon microstrips are taking over as detectors with tracking precision $< 100\ \mu m$ and are able to handle the hit rates and the integrated radiation doses. For such large detectors, spatial precision is less of an issue (it will, in fact, be a challenge to build them with few micron stability, so the intrinsic detector precision may not be the driving factor). This is one reason for the interest in (for example) binary readout.

However, these detectors clearly have their limitations. There is a nasty hole of radius ≈ 30 cm in ATLAS and CMS within which microstrip detectors dare not venture, due (as we shall see in Sec. 6) to problems of radiation damage. With the huge event multiplicities, track merging would also be very serious. In this region, silicon pixel detectors may find a home and (at the smallest radii) other detector options, as we shall discuss in Sec. 7. The overall result is that the main emphasis in the world of silicon microstrip detectors has shifted from aiming to achieve the ultimate in spatial precision with the minimal detector thickness (including pushing for double-sided detectors) to aiming to cover very large areas as economically as possible, with electronics having an extremely high rate capability. The pressure for the most economical solution may argue against double-sided detectors, particularly since the material associated with the additional silicon layers is not seriously detrimental to the momentum resolution of the tracks that are important for physics. Fortunately, the size of the collaborations has grown at least as fast as the areas to be covered, so there is every reason to believe that they will succeed in these challenging tasks.

To describe any advanced technology as mature is usually misleading. Silicon microstrip detectors and particularly the associated electronics will continue to evolve for many years. However, as the OPAL Collaboration demonstrated when they decided they needed a silicon microstrip vertex detector to retain LEP competitiveness, it is possible starting from scratch to build a sophisticated detector with this technology within a year, provided one does not try to invent a lot of new features.

For linear e^+e^- colliders, microstrip vertex detectors were never ideally suited, due to the high background associated with the single-pass collider operation. (As Witold Kozanecki puts it, "Backgrounds at SLC are similar to those at LEP, *during injection.*") This will also be true at small radii for the future high-energy linear collider. However, as at LHC, there is a good chance that silicon microstrip detectors may be the chosen technology for the outer tracking system at this machine.

5 Pixel-Based Detectors

5.1 Introduction

There are exceptions to every classification scheme. I was delighted to read a paper [41] contributed to the recent European Conference on Semiconductor Detectors which neatly bridged the gap between the one-dimensional microstrip detectors and the two-dimensional pixel-based detectors. How could this be? The authors were interested in detecting hard X-rays for which the attenuation length in silicon is rather long. To achieve a reasonable efficiency, they had the excellent idea to turn a microstrip detector *edge-on* to the X-ray direction, so that the strip length (several mm) became the effective detector thickness. In this way, they were able to achieve 80% efficiency for detecting 20 keV X-rays. By sweeping the detector slowly across the image, they were able to build up full two-dimensional images of excellent quality.

More usually, the term "pixel detector" is taken to mean a device equipped with a one- or two-dimensional array of pixels (picture elements). The two-dimensional variety, given the sensitivity of silicon for visible light, is the basis for the huge commercial market for camcorders and other electronic image-capture products. This marks the most important contrast with respect to the previously discussed microstrip detectors; while the strips can provide very precise position information, the fact that they are inherently one-dimensional precludes any application in which the desired output is some form of picture. Hence, pixel devices are of much greater interdisciplinary importance (both in terms of scientific sensors and in commercial terms) than microstrip detectors.

However, for tracking devices such as vertex detectors, how important is it to have this picture-taking capability? Figure 31 demonstrates that a few planes of pixel-based detectors give unambiguous track-finding capability, whereas the same number of planes of *double-sided* microstrip detectors do not produce an immediately recognizable pattern of tracks. There are in fact $N!$ patterns possible in the case of a jet of N particles. This is not too bad for the three-particle case shown (six-fold ambiguous), but for a high-energy jet of ten tracks, there are 3.6×10^6 possibilities!

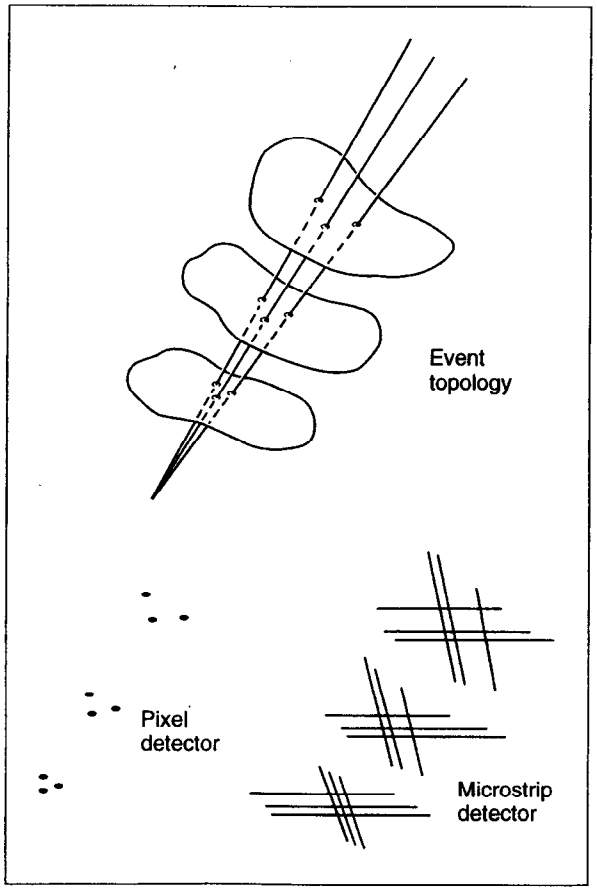


Fig. 31. Upper sketch: a few tracks traversing an unspecified set of three detector plates. Lower sketch: resulting information in case of one- and two-dimensional detector types.

What this means in practice is that such detectors would need to combine information from different planes having strips oriented differently (not necessarily a practical option in a collider detector) or (more usually) rely on the external detectors to perform the pattern recognition. Since there can be a lot going on between the IP and the outer tracking detectors (decays, γ-conversions, secondary interactions), a pixel-based vertex detector capable of stand-alone pattern recognition is manifestly a much more powerful tool for physics.

A second advantage is that of granularity. A single typical microstrip (e.g., of the DELPHI detector) covers 70 mm $\times$ 50 μm. This area (in a CCD detector) would be covered by 9000 pixels. These four orders of magnitude in granularity make for a huge advantage in tolerable hit density before the problems of cluster-merging start to make life difficult for the track reconstruction algorithm. One can for this reason position a pixel-based detector much closer to the IP, with obvious advantages for impact parameter precision (shorter extrapolation, just as a short focal-length lens makes for a more powerful microscope). Furthermore, there are physics environments where the density of background hits close to the IP is so high that a microstrip detector would be obliged to back away in order to reduce the occupancy to a tolerable level, whereas a pixel-based detector would be perfectly comfortable.

The third advantage is in terms of radiation hardness. We shall address this complex issue in Sec. 6, but in many cases, the limiting parameter is growth of leakage current, with associated shot noise which eventually can overwhelm the signal. In a microstrip detector, the signal on one strip has to be found against the noise background associated with the entire strip. If the "strip" length is reduced by a factor 10^4 (above example), the noise associated with the leakage current is correspondingly reduced. This can make the difference between a detector lifetime of one month and 2,000 years.

There are two other partly connected advantages. Most forms of pixel-based detectors have extremely low capacitance nodes for the charge collection, and hence need much smaller charge signals for satisfactory signal-to-noise. Excellent MIP detection efficiency is achievable with active layers 20 times thinner than microstrips. As we have seen, this has major advantages for tracking precision, both

for normal-incidence particles (minimizing the problem of δ-electrons) and for angled tracks (minimizing the effect of energy-loss fluctuations). Originally, it was customary, in using these devices with thin active layers, to leave them mechanically thick (say 300 μm), but more recently, techniques have been developed for mechanical lapping, chemical etching, and handling so that thinner devices can now be built into HEP detectors, with a further reduction in multiple Coulomb scattering.

Against these advantages, pixel-based detectors have disadvantages which make them impractical or uncompetitive in some situations. In order to appreciate these, however, we need to consider the two important classes of such detectors, for which the characteristics are extremely different and indeed complementary. These classes are the charge-coupled devices and the active pixel sensors.

5.2 Charge-Coupled Devices (CCD's)

An imaging CCD (Fig. 32) consists firstly of a square matrix of potential wells, so that signal charge generated below the silicon surface can be accumulated, building up an image. Secondly, by manipulating clock voltages in the *parallel register* (the I ϕ gates), charge can be transferred in parallel from one row to the next. Charge in the bottom row of the matrix is transferred into the adjacent *linear register*. The stored signals are then transferred one at a time (by manipulating the R ϕ gates) onto the output node, which is connected to the input of an on-chip charge-sensing amplifier. Also on chip is a reset FET to restore the output node to its nominal value, usually after reading the signal from each pixel. Thus, the CCD image is converted from a two-dimensional charge pattern to a serial train of pulses, well-suited to display on a video monitor. The CCD was invented in 1970 (Ref. [42]). Devices of this pioneering design (so-called surface channel CCD's, because the signal charge is stored at the silicon/silicon dioxide interface) are still used in video cameras. However, within two years, the invention of the more sophisticated buried-channel architecture was published [43]. Here, the signal charge is stored in the bulk of the silicon approximately 1 μm below the surface, suitably remote from the interface states that (as we shall see) can trap signal charge. For the small signals usually

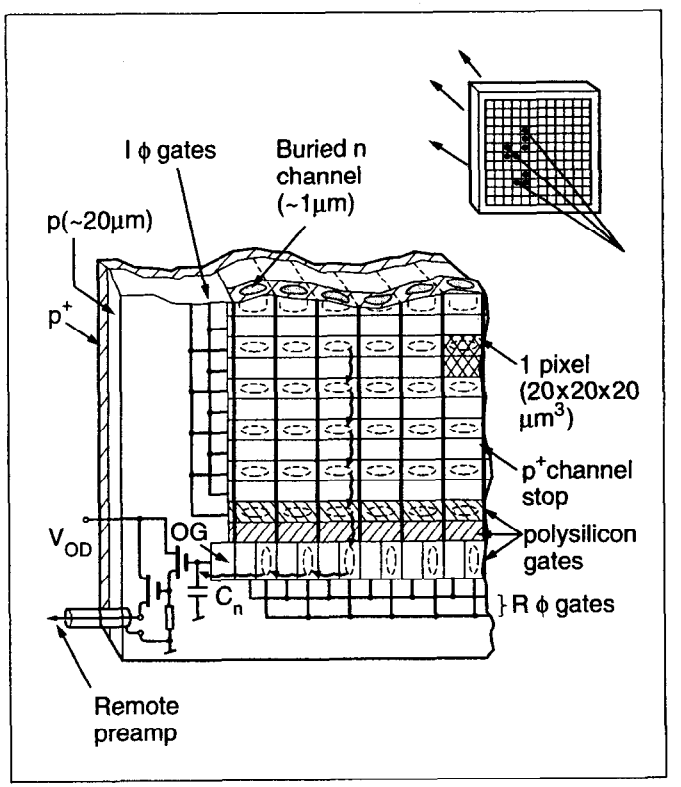


Fig. 32. Upper right: sketch of charge storage in a CCD detector traversed by a number of ionizing particles. Lower left: corner region of CCD showing the principal structural features.

sought in scientific applications, the buried-channel design is much more suitable, so we shall concentrate entirely on this.

Before diving into the details of scientific CCD's, let us take a brief look at the technology push being provided by industry. The largest CCD market is for camcorder sensors. The immediate aim in this market is to increase sensitivity so as to achieve good performance under typical indoor home lighting conditions. The next goal is CCD's for HDTV broadcast cameras (1920 x 1036 pixels, two readout channels, each running at 37 MHz) followed (in about 1998) by the HDTV camcorder. In addition, there is a big push for a high-quality electronic still photography camera, and eventually, an electronic motion picture camera. CCD design rules in the commercial sector are 0.5 μm (and reducing), and wafer sizes are 6" (and increasing). Both of these are currently beyond the reach of the manufacturer of scientific CCD's. The commercial devices use interline transfer and are typically only about 2 μm thick (active layer). This is excellent for sharp color images but makes them inapplicable for most radiation detector applications. The major commercial manufacturers are too busy chasing the frontiers associated with the mass market to be interested in the specialized needs of the scientific CCD users. Fortunately, there are several extremely high-quality manufacturers who serve this particular niche in the market. The possibility of using CCD's as high-precision detectors of MIP's was first evaluated theoretically about 15 years ago [44].

5.2.1 Structure and Basic Operation

We shall concentrate on the frame-transfer MOS CCD family since this is the most commonly used for scientific applications and the only one used to date for vertex detectors in HEP experiments.

Let us examine in some detail, with the aid of the general discussion of Sec. 3, how such a detector can be built. For more detailed information, there are some excellent books on CCD's [45, 46] as well as CCD conference proceedings and hundreds of published papers.

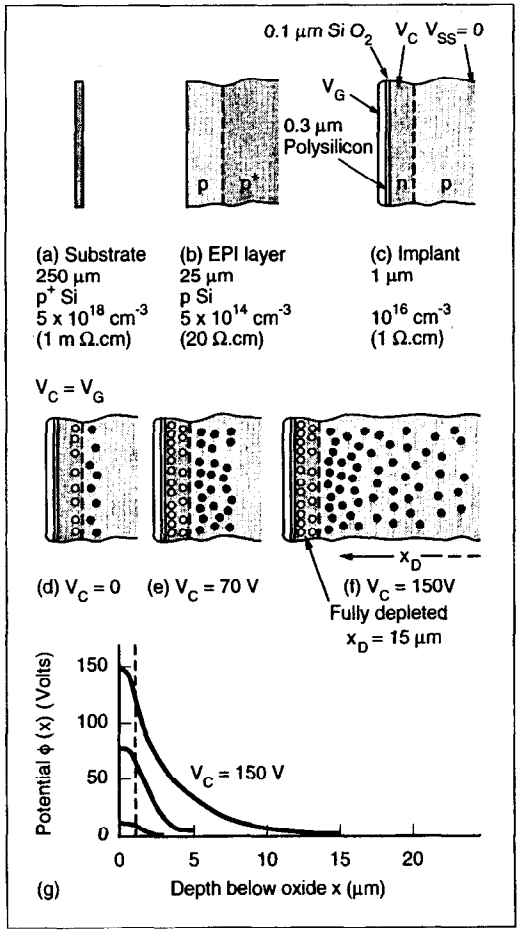


Fig. 33. (a)–(c) The successive stages in making a CCD-like structure (shown with increasing magnification). (d)–(f) The depletion process which would apply if V_c and V_G were increased together. (g) The corresponding potential distributions as a function of depth in the silicon.

Let us first consider the steps in making a device which would have some (but not yet all) of the features of a CCD. Starting with a low-resistivity, suitably inert substrate [Figs. 33(a)–(c)], we proceed to grow an epitaxial layer of higher resistivity silicon with a thickness adequate to contain all the necessary structures and associated field penetration. We next make a *pn* junction by the introduction of a shallow (approximately 1 μm) implant of *n*-type dopant. The surface is oxidized to make an insulating layer, and on top of this is deposited a thin conducting layer. The simplest would be aluminum, but for light detection, a high degree of transparency is important, and about 0.3 μm low resistivity "polysilicon" (amorphous silicon) would commonly be used. By analogy with FET's, the conducting surface layer is termed a "gate."

Let us now put some bias voltage onto the structure, as shown in Figs. 32(d)–(f). Grounding the substrate ($V_{ss} = 0$), we apply V_c to the *n* channel and V_G to the gate. Initially, assume $V_c = V_G$. Even with $V_c = 0$, as we learned in our discussion of the *pn* junction, there will be a thin depletion layer around the interface between the two types of silicon. By increasing V_c, we are able to deplete more of the material as the junction becomes more and more strongly reverse biased. With the parameters chosen in this example, a high voltage would be needed to achieve complete depletion of the *n* channel, at which point we should have depleted about 20 μm of the *p*-type substrate. The potential distributions for increasing values of V_c are shown in Fig. 32(g). For V_c = 150 V, such a device when traversed by particles would transport the generated electrons to the surface (silicon/silicon dioxide interface) and dump the holes into the undepleted substrate.

Now [looking at Figs. 34(a) and 34(b)], consider what happens if V_c is increased from zero while V_G is held at zero volts. Here, the situation is entirely different; the large capacitance between the *n* channel and the gate provides a further mechanism for depletion of the channel. The depletion around the *pn* junction proceeds as before, but the voltage across the oxide induces an increasing positive space-charge, starting from the silicon/silicon dioxide surface and growing into the body of the *n* channel. At a very low value of V_c (about eight volts), these depletion regions meet, causing the phenomenon known as *pinch-off*. The corresponding value of V_c

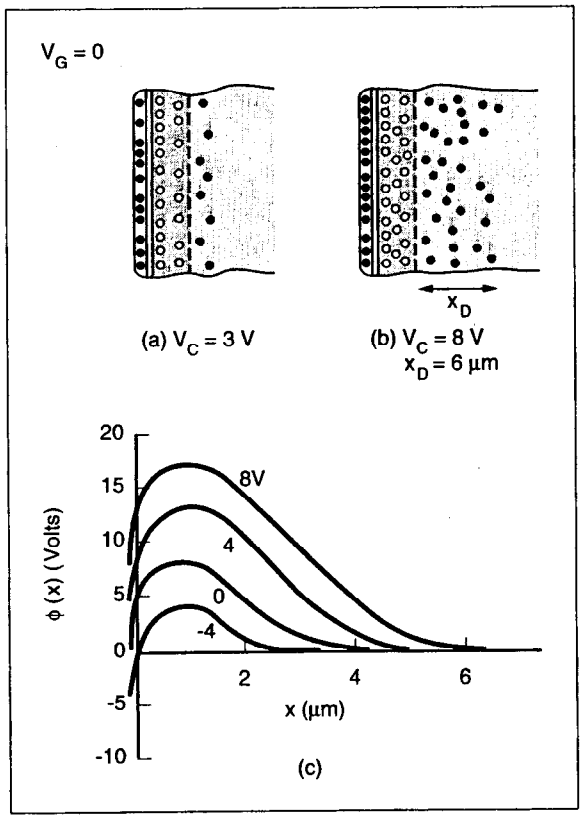


Fig. 34. (a) and (b) The depletion process in normally biased CCD operation with V_G negative with respect to V_c. (c) The corresponding potential distributions after channel pinch-off for various values of V_G.

is called the pinch-off voltage, and when it is reached, further increases of V_c (which can be controlled, say, by an edge connection) have no influence on the potential over the area of the detector. The depletion depth in the *p*-type material is only about 6 μm in this case. What is particularly interesting is the potential distribution in the silicon. This is shown in Fig. 34(c); look initially at the curve for $V_G = 0$. The quadratic form in both types of silicon is, of course, preserved (this is a consequence of Poisson's equation and uniform doping), but there is now a maximum in the electric potential just below the depth of the *pn* junction. This acts as a potential energy minimum for electrons, so (in contrast to the case $V_G = V_c$) the electrons liberated by the passage of a particle would accumulate approximately 1 μm below the silicon surface in the so-called *buried channel* of the device. This is a vital ingredient in the design of CCD's for our application. Tiny charges (< 10 e^-) can be safely stored and transported as long as they are held in the bulk of the silicon. Once they are allowed to make contact with the surface, they encounter numerous traps which cause serious loss of charge. Surface-channel CCD's, while quite commonly used, should be avoided for work with very low signal levels.

Notice that the situation depicted in Fig. 34(c) represents a non-equilibrium condition. Thermally generated electrons would accumulate in the potential energy minimum and drive more and more of the *n* channel out of depletion. CCD operation relies on some procedure for keeping the channel swept clean of electrons at an adequate rate.

Assuming that we avoid this accumulation of electrons, the effect of now varying the gate voltage V_G is to a first approximation simply to vary the depth (in volts) of the potential well, but hardly at all to change its depth (in microns) below the silicon. There is, in fact, a slow variation in the depletion depth with V_G, as can be seen from the figure. The quantitative calculation follows easily from what we have done in Sec. 3; see, for example, Ref. [46] for the details.

The device we have created has all the depth characteristics of an imaging CCD, but it still lacks two important features before it will have the necessary pixel structure over the surface. These are illustrated in Fig. 35. Firstly, at the required pixel granularity (say, 20 μm), p^+ implants are introduced of approximately 1 μm width and

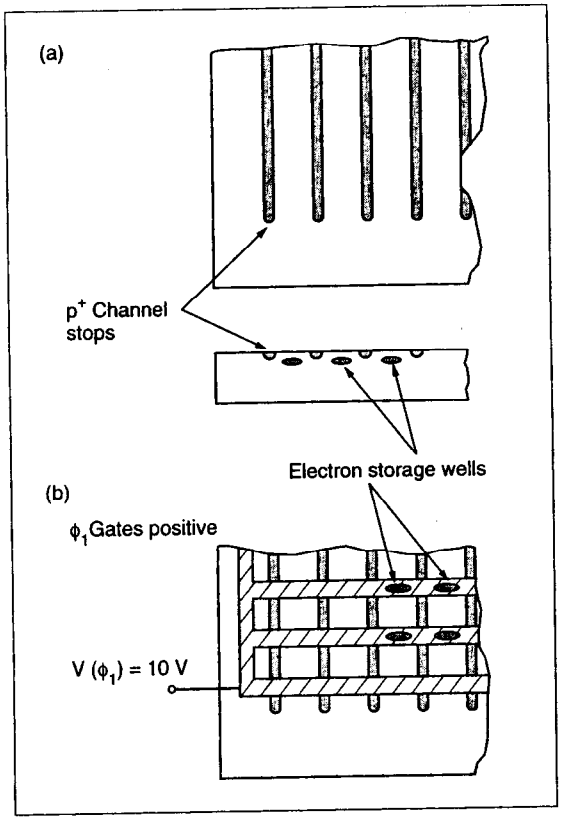


Fig. 35. Establishing the potential well structure: (a) Channel stops create potential barriers running vertically on the device. (b) Gates create horizontal potential barriers. The combined result is a matrix of localized wells, each of which constitutes a pixel.

1 μm depth. These become partly depleted as part of the overall biasing of the CCD, and so provide strips of negative space charge which effectively repel electrons. Thus, the electrons in the buried channel will now be confined to separate storage wells which run from top to bottom of the detector, in the view shown in Fig. 35(a). The typical doping level of the channel stops is $N_a = 10^{18}$ cm^{-3}.

Secondly, the charges are confined in the vertical direction by making a polysilicon gate structure which is not uniform across the surface but which consists of a series of horizontal bars. By biasing these positively [see Fig. 34(c) and Fig. 35(b)], we can achieve potential wells under each of the intersections between these gate electrodes and the regions midway between the channel stops. We now have a matrix of discrete potential wells which may exceed 10^6 in number on a typical CCD (800 channel stops x 2000 gate electrodes).

But still, we do not have a working CCD, since those potential wells are immobile. We can accumulate charge images but cannot read them out. To do this, we make a more complicated gate structure (Fig. 36). We arrange these gates in triplets (ϕ_1, ϕ_2, ϕ_3) in this so-called three-phase CCD structure. The static situation is for one phase (say, ϕ_1) to be high, so that the electrons are stored under this phase. Then by manipulating the voltages between ϕ_1 and ϕ_2 as shown in the figure, the electrons are moved to ϕ_2. Keeping ϕ_3 low throughout this operation ensures that the charges between adjacent pixels cannot be smeared together. The total physical width of $\phi_1 + \phi_2 + \phi_3$ electrodes together constitutes one pixel, e.g., $3 \times 7\ \mu\text{m} = 21\ \mu\text{m}$.

Now we have developed the capability to move all the stored charges down the device (for example) by one pixel at a time. Apart from three-phase CCD's, there exist other varieties (four-phase, two-phase, virtual phase, etc.).

At the bottom of the area array called the "imaging" or "I array" is a linear CCD, the output register or R register into which the charges stored in the bottom row of the I array can be shifted. Once in this register, charges in that row can be transferred sideways so that the charge contained in each pixel is sensed in turn by the on-chip circuit.

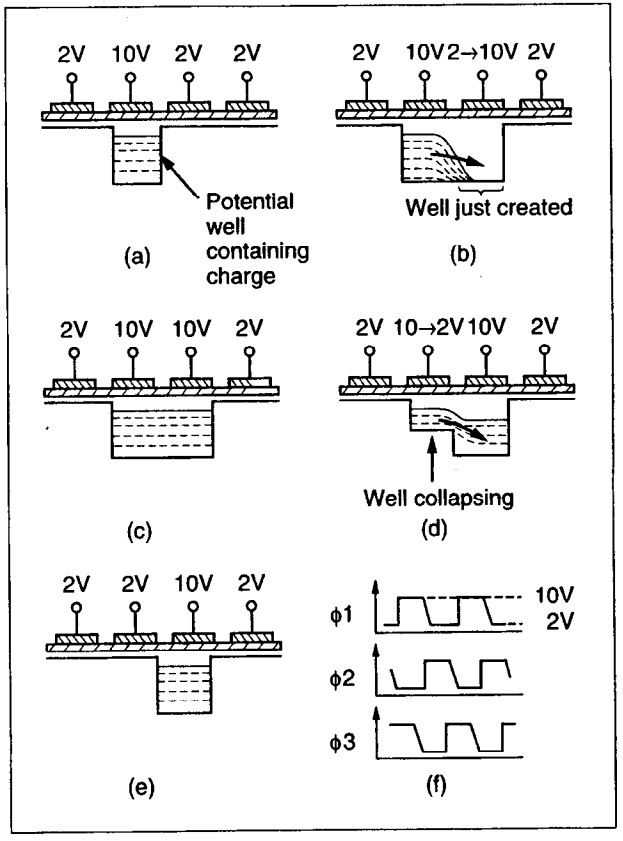


Fig. 36. After Ref. [46]. (a)–(e) Movement of potential well and associated charge packet by clocking of gate electrode voltages. (f) Clocking waveforms for a three-phase CCD.

Referring back to Fig. 32, which shows a two-phase CCD, note that each pixel (shaded area) covers the height of two I gates and is bounded by a pair of channel stops.

The CCD structure shown in this figure is sensitive to light or to particles over the full active area. It should be noted that this is not true of all imaging CCD's. Some, for example, have more complex channel stops, *pnp* structures which can be used for anti-blooming or for fast-clearing the CCD's. Such devices have dead bands between each pixel, a feature which makes them unacceptable for most applications as particle detectors.

In the spirit of Fig. 25 (simplified cross section of a generic microstrip detector), Fig. 37 shows the corresponding case of the MOS CCD. Note the buried channel, a region within the n^+ implant, not crossed by field lines, and the crossover in the electric field at that depth (lower plot). Note that the buried channel depth varies only slightly as the gate voltage is varied. Note also the intrinsic p/p^+ potential barrier created by the narrow depletion region at the back surface interface of the epitaxial silicon. We can correlate this with Fig. 24, which shows how the charge generated by a MIP along its track falls into a number of classes in such a structure. There is a region of typically 5 μm below the surface for which the charge is within the depletion depth and is fully collected into the "central" pixel, i.e., the one traversed by the particle. Next, the charge from the 15 μm of undepleted epitaxial silicon (which generally has a long diffusion length, maybe hundreds of microns) diffuses isotropically. About half of it diffuses into the depletion region and is caught in the central pixel *or* in neighboring ones; the other half gets there after being reflected off the p/p^+ potential barrier.

As has already been noted, the CCD potential wells represent a non-equilibrium condition. Thermal generation of electron-hole pairs in the material provides a source of electrons which accumulate. For TV imaging, these constitute a minor background, but for astronomy, the long integration times and low signal levels necessitate cooling, typically to liquid nitrogen temperatures. For particle detection, the requirements are less stringent and operating temperatures around 200 K may be

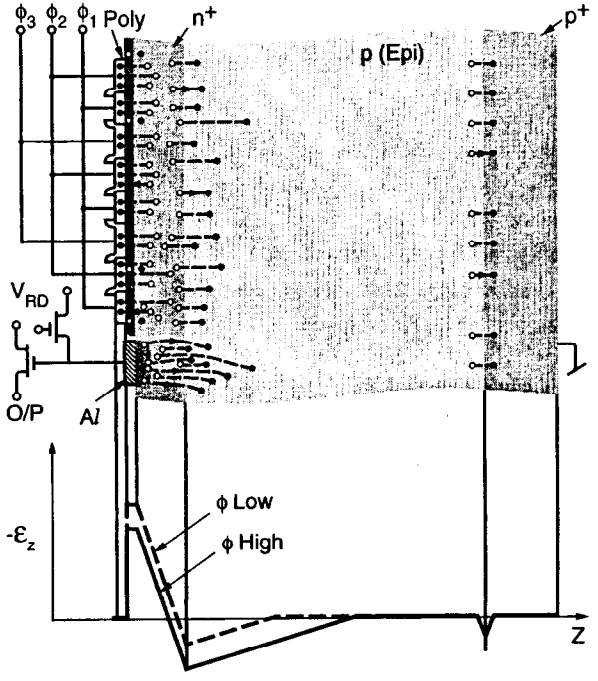


Fig. 37. Sketch of cross section of a generic three-phase MOS CCD. As in Fig. 25, exposed fixed charges are shown by open circles (positive) and filling circles (negative). Also shown is the electric field distribution in regions of high- and low-imaging gate voltage.

entirely adequate, but this depends strongly on the timing of the clearing and readout of the detectors.

It is worth noting that the scientific CCD has in recent years been extended by the development of the *pn* CCD [47]. This is (like many "innovations") far from new, having been developed, then forgotten, soon after the original CCD invention. At that time, it was known as the junction CCD [48]. As shown in Fig. 38, it is very like the MOS CCD except that negatively (reverse) biased p^+ implants are substituted for the MOS gate structure. *pn* CCD's are usually manufactured with a view to high X-ray efficiency, and hence are fabricated on high-resistivity silicon which is fully depleted, as in the microstrip detector. This case is shown in Fig. 38. For X-ray detection, there are recent papers reviewing the relative capabilities of both the MOS [49] and *pn* [50] CCD's. There is a considerable overlap as well as a degree of complementarity in their application areas [51]. For particle physics applications, MOS devices have been exclusively used to date, largely because of their ready availability at competitive prices from a number of manufacturers.

E. Fossum has written an excellent recent review of image sensor technologies (mostly CCD's) and of companies manufacturing these devices for scientific customers [52].

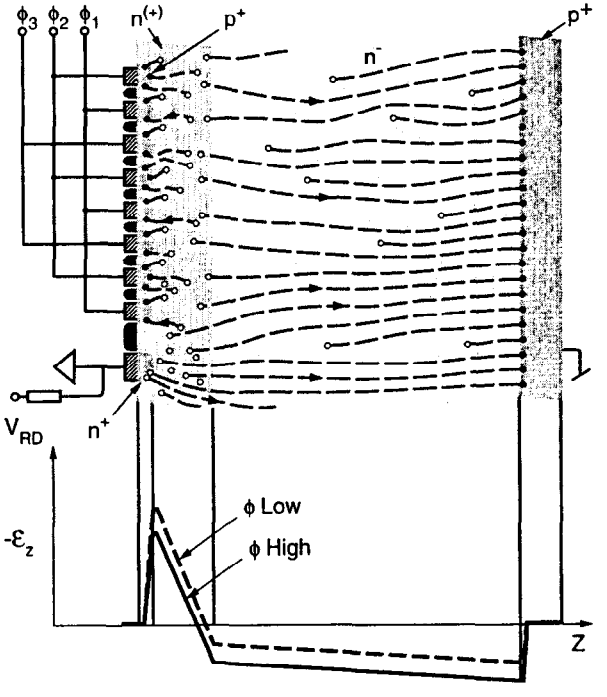


Fig. 38. Sketch of cross section of a generic *pn* CCD manufactured on high-resistivity silicon and depleted over the full thickness.

5.2.2 CCD Charge Transfer and Readout: Detailed Issues

5.2.2.1 Charge Transfer Process

As we have seen, the transfer of signal charge from pixel to pixel is accomplished by changing the voltage levels on the gate electrodes. Since the magnitude of the MIP signals is so small (approximately 2,000 e^- compared with about 10^5 e^- well capacity), one might imagine that very small drive pulse modulations would suffice to achieve good CTE. On the contrary, 5–10 V pulses are needed. Why is this? Firstly, in producing any IC, fixed positive charge is trapped at the silicon/silicon dioxide interface. This is dependent on the processing details, so one would never find perfect equality between (say) the three-phases of a register, which are obviously deposited in separate operations. The uncontrollable differences amount to

several tenths of a volt and result in effective charge storage for MIP signals even in the absence of any applied drive voltages. The idea of getting rid of early hits by letting the charge diffuse along the columns, as enunciated in my group's first paper [44] on the possible application of CCD's for MIP detection, simply does not work, as we found some years later. Small signals in CCD's cannot be eliminated like this; the electrons in the charge packet are, in fact, extremely cohesive.

The most significant factor that determines the minimal drive pulse voltages required for good CTE is the unavoidable presence of shallow traps which tend to pick up signal electrons at every stage of their long journey to the output node. Particularly in a sparse data situation such as one has in a particle detection system, such traps are dangerous. They may emit electrons with a long time constant, then sit empty until the arrival of a signal packet, at which point they capture electrons almost instantaneously. The signal packet moves on, with the trapped electrons being released only much later. As we shall discuss in Sec. 6, radiation damage can cause serious growth in the density of these bulk traps. The operating temperature is a very important parameter in minimizing this problem, since it profoundly affects the trap emission time constants. The problem of bulk traps affecting CTE in CCD's was first treated in a famous paper by Mohsen and Tompsett [53]. The topic has been revisited many times since; for a recent paper dealing specifically with CCD's operated at low temperature, see Ref. [54].

As well as the problem of traps of atomic dimensions, CCD's are also sensitive to more macroscopic potential wells (sometimes referred to as potential pockets) that can swallow part or all of the charge packet within one pixel. There are innumerable processing imperfections liable to cause such potential wells (minor variations in gate oxide thickness, tiny blemishes in gate polysilicon, minor crystal imperfections such as slip lines, and so on). Such manufacturing problems can be very difficult to diagnose; suffice it to say that less than 10% of large area devices made by a top-of-the-range CCD manufacturer are likely to suffer from such effects in more than 1% of the columns. As such, this is not a serious yield issue.

Both as regards atomic-scale bulk traps and as regards potential pockets, high-drive voltages can be extremely effective in releasing electrons from all but the deepest lying bulk traps, by virtue of the Poole-Frenkel effect [55] (lowering of a potential barrier by a potential gradient). Interestingly, the relevant strong electric fields arise not from the horizontal fringing fields, but from the fields developed along the vertical doping profile of the buried channel implant [54]. The device physics may be somewhat subtle, but the experimental observations are unambiguous: for good CTE, drive pulses in excess of 5 V and typically 10 V may be needed. What are the consequences of this?

As regards the parallel register, the capacitance to ground of each of the gates is pretty large, the polysilicon gate electrodes are somewhat resistive, so one may be limited to clock rates of around 100 kHz in order to achieve adequate voltage excursions at the center of a large CCD. The large currents induced in the CCD structure by the voltage excursions in the parallel register (which covers nearly all of the area of the device) generate massive feedthrough signals on the CCD output circuit. Neither the limited clocking frequency nor the feedthrough signals can normally cause any problems, since each parallel transfer is followed by typically 400 serial transfers as that row is read out, so the overall readout time is not seriously affected by the parallel transfer time constants.

For the serial register, equally large drive pulses are required. However, the associated capacitance is much smaller, and there is no problem to clock the serial register with good CTE in excess of 20 MHz. The theoretically maximum clocking rate is a very rapid function of the pixel size (length) [56], 60 MHz for 20 μm but only 4 MHz for 50 μm pixels. Experimentally, 20 μm pixels are easily clocked at 30 MHz.

In a vertex detector application, material in the active detector volume is to be minimized. In an optimized CCD design, the on-chip power dissipation associated with the drive pulses and readout amplifier are similar and extremely modest. A detector of some hundreds of mega-pixels can be cooled by a gentle flow of nitrogen gas. The cooling problems would become approximately a 100 times greater if the drive and readout electronics were contained within the low temperature enclosure. In practice, one always locates these outside the cryostat (using low mass striplines of approximately 30 cm length for the interconnections). Thus, the local electronics

can be run at room temperature, water cooled, and positioned in the small polar-angle region, beyond the coverage of the tracking detectors. Recent developments in electronics design have offered a remarkable opportunity for shrinking *all* the drive electronics into this small space where tracking is not required, at the heart of a collider detector. This allows the cleanest possible drive pulse generation, a major improvement on earlier systems for which these pulses had to be generated in modules on the periphery of the global detector, some tens of meters distant, and carried in on approximately 1000 fine coax cables.

As already noted, because of their low duty cycle, the parallel register drive pulses make only a minor contribution to the detector readout time. This readout time is effectively determined by the duration of the serial register clocks and the analogue signal-sensing electronics. In operating a CCD register, phases are always clocked in opposition, one coming down and another going up as the electron packet is passed on (see Fig. 36). The cleanest arrangement is the two-phase register, where an implant beneath each gate biases the charge packet to be always stored in the "downstream" half of the gate area. Balanced drive pulses to the two gates provide minimal disturbances to the CCD output circuit. But it is a very delicate business. The 10 V pulses are swinging around during the transfer of a MIP signal which (if one is lucky) may give a 1 mV step on the output node. The positive and negative edges of the drive pulses are unlikely to be balanced to better than a few percent.

Even if they were perfect, geometry layout differences on and off CCD (more importantly, the latter) can cause major feedthrough and ringing of the analogue signals by ten to 100 times the 1 mV level. For slow readout systems, one can wait for this to settle down. A major challenge in reading CCD's at 20 MHz or above with low noise (< 100 e^- RMS) is to achieve excellent isolation between the drive and analogue signals in compact systems. Use of miniature coax for the two critical drive lines between the local electronics and the detector is certainly helpful, but there are numerous possible feedthrough paths, all of which need to be extremely carefully controlled.

5.2.2.2 Charge Detection

The most commonly adopted CCD on-chip charge detection circuit is of the general form shown in Fig. 39. It consists of firstly an output diode, the very small n^+ implant seen in Fig. 37, linked to the serial register via the output gate (OG of Fig. 32). Thus, the CCD output node has its potential reset periodically to the reference voltage V_R via the reset transistor, which restores it to an appropriate voltage for collection of signal charge Q_s transferred by clocking from the buried channel of the serial register. This charge transfer causes the node potential to change by $\Delta V = Q_s / C_n$, where the node capacitance C_n is given by

$$C_n = C_d + C_g(1-G)$$

C_d is the node-substrate capacitance, and C_g is the gate-source capacitance of the transistor. G is the voltage gain of the source follower. For optimum signal to noise, these two capacitive components should be approximately matched. See Ref. [57] for a detailed discussion of the optimal transistor design parameters. This implies a small-sized transistor, which consequently has a relatively high impedance at its output source. For optimum noise performance, it is advantageous to use a depletion mode or buried channel MOSFET. This important discovery, made ten years ago [58], is understood in that the drain current in a surface channel FET experiences noise due to the continual random filling and emptying of interface states, which consequently modulate the channel characteristics. For a modern CCD [59], the advantage of a buried channel first-stage MOSFET is indicated in Fig. 40; the $1/f$ noise in the buried channel version is much reduced. There is a penalty in power dissipation in the buried channel device; for the same transconductance, a higher current is needed.

As already explained, for a vertex detector application, the off-chip amplifier and further processing should be external to the cryostat. Thus, the CCD amplifier needs to drive a capacitive load of some tens of picofarads. For slow readout, the first stage source follower alone is adequate, but for a high-speed system, the bandwidth

requirement implies a much larger transistor (lower g_m). Hence, the tendency in such cases will be to use a two- or three-stage output circuit, as shown in Fig. 39. With an optimized design, the noise performance is dominated by the first stage, even in the case where the later stage FET's are enhancement-mode devices.

A most important development in the early days of CCD signal processing was the invention of *correlated double sampling* or CDS [60], a technique which has since been adopted for charge detection circuits for microstrip detectors. The original aim was to reduce *reset noise* in CCD readout systems. The term "reset noise" refers to the unavoidable fluctuations in the node voltage (kTC_n) which arise from thermal fluctuations when the reset transistor is switched in and out of conduction. The procedure consists of sampling the node voltage twice after the reset, once before and once after the transfer of the signal charge. There are various options for filtering the signal preceding each sample; see Ref. [61] for a discussion. The optimal procedure consists of a signal integration for the same fixed period before and after sampling. In this case, the resultant total noise after sampling is given by

$$e_{nT}^2 = \int S_F(f)^2 e_n(f)^2 df ,$$

where $e_n(f)$ is the output circuit noise voltage per $Hz^{1/2}$ at frequency f, and $S_F(f)$ is the Fourier transform of the filter sampling function $S(t)$.

For the case of the dual integration for time τ,

$$S_F(f) = \frac{2\sin^2 \pi f\tau}{\pi f\tau} .$$

Note that this filter function falls to zero both at low and high frequency. Thus, CDS not only eliminates reset noise but also reduces the noise contribution from the output transistor, particularly in the low-frequency $1/f$ region, and in the high-frequency region (though the latter will normally be cut off by a suitable bandwidth limit to the main amplifier). The excellent noise performance of a modern CCD with the benefit of CDS is indicated in Fig. 40.

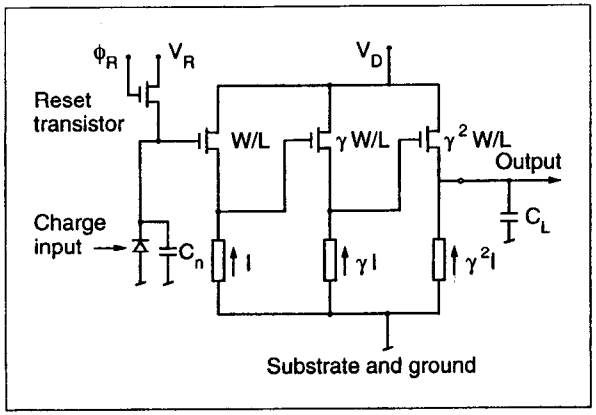


Fig. 39. Schematic diagram of a three-stage output circuit.

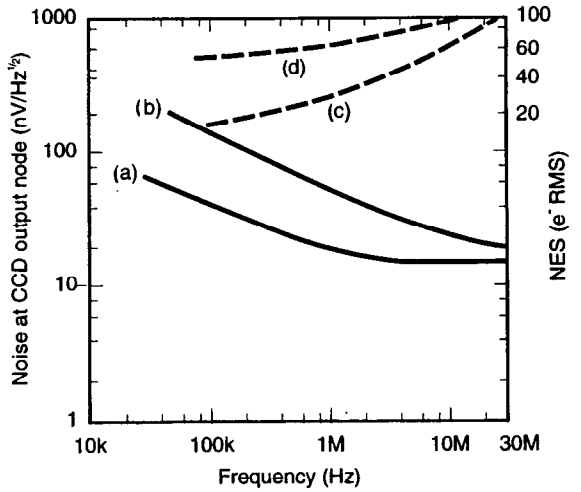


Fig. 40. Noise spectra of (a) buried channel and (b) surface channel first stage MOSFET's in a CCD (left-hand axis). (c) and (d) show the corresponding CDS noise equivalent signals in RMS e^- (right-hand axis).

The procedure normally followed in vertex detector readout, where readout speed is to be optimized, is to take advantage of the very small integrated charge to be expected in any row of the image, and hence, to reset the FET only at the end of each row. Thus, the signal charge of each successive pixel is just piled on top of its predecessors, and the CDS processing consists of simply taking successive differences of the filtered signal for pixel N, minus that previously sampled for pixel $(N-1)$. It is not necessary to wait for the clock feedthrough from the linear register to settle; as long as this is constant from one pixel to the next, it is eliminated by the CDS differencing procedure. There is clearly a limit to this, for example, if the feedthrough is so large as to push the amplifier beyond its linear range during the sampling period, or if the sampled signal is swinging too rapidly at the moment of ADC sampling. In a well-controlled system, the readout noise *clocked* will be little greater than the value measured with the CCD unclocked. But achieving this in a system running at 10 MHz or above can be a major challenge for the circuit designer.

5.2.2.3 Vertex Detector Readout Options

Given the many options for CCD architecture and external electronics, the vertex detector designer has the opportunity to adapt the system design to the needs of the experiment, within wide boundaries. This has become particularly apparent as the cost of fully customized CCD design has fallen to the level where it is appropriate to plan on a completely new design for any experiment.

In this discussion, we restrict ourselves to the general architecture of frame transfer CCD's. Interline transfer devices, which can offer (via the variant of gated anti-blooming drains) the option of fast clearing on the microsecond timescale, are not considered. Despite this convenience, such devices are unsuitable for high-precision tracking applications where high detection efficiency is also essential. One cannot afford, in a vertex detector where the overall thickness is critical, to have detector planes which are only 70% efficient; close to 100% MIP efficiency is essential.

As we saw in the previous section, the original idea of disposing of signals from out-of-time tracks by charge diffusion does not work; the only way to get rid of unwanted signals is to clock the charge out via the output node. During the pre-trigger conditions, this can normally be best achieved by running the detector in "fast clear" mode. By synchronously clocking the parallel and serial registers at the upper rate limit for the former (around 100 kHz), one can sweep unwanted signals out in a mean time interval of around 10 ms for a large-area CCD. In a fixed target or rapid-rate collider environment, this implies a certain density of background hits in the CCD at the time of the event trigger. If this density greatly exceeds $1/\mathrm{mm}^2$, one should consider carefully whether this is an appropriate environment for such a detector. But up to this density (occupancy only approximately 10^{-3} in a detector with 20 μm square pixels), it is no problem to filter out this background.

In a modern experiment, top-level trigger decisions may take a while to arrive, say, 1 ms. During this time, one would be continuing the fast clear operation, in ignorance of the wanted data in the detector. On receipt of the trigger, the clocking would change to readout mode. Valid data from a region of, say,

$$20\,\mu\mathrm{m} \times \frac{1\,\mathrm{ms}}{10\,\mu\mathrm{s}} = 2\ \mathrm{mm}$$

at the edge of each CCD would have been lost in the time interval between the event and the trigger. It is no problem to allow for this small reduction in the fiducial region, at the detector design stage.

At this point, one is presented with numerous options depending on conditions. Let us take three examples, a fixed-target experiment, and two e^+e^- linear collider scenarios. These are based on actual experience but should not be taken to mean that CCD vertex detectors are necessarily limited to these environments.

For a fixed-target experiment, there is normally extra space available outside the spectrometer aperture. Therefore, it makes sense to extend the CCD area by at least the size of the fiducial region and to continue fast clearing until the valid data are all in a storage area well away from the high flux beam region (Fig. 41). This was the procedure used in the NA32 experiment. The detector could then be read out at leisure. In fact, to keep conditions even cleaner, a small kicker magnet was used to dump the beam during readout, but this was barely necessary.

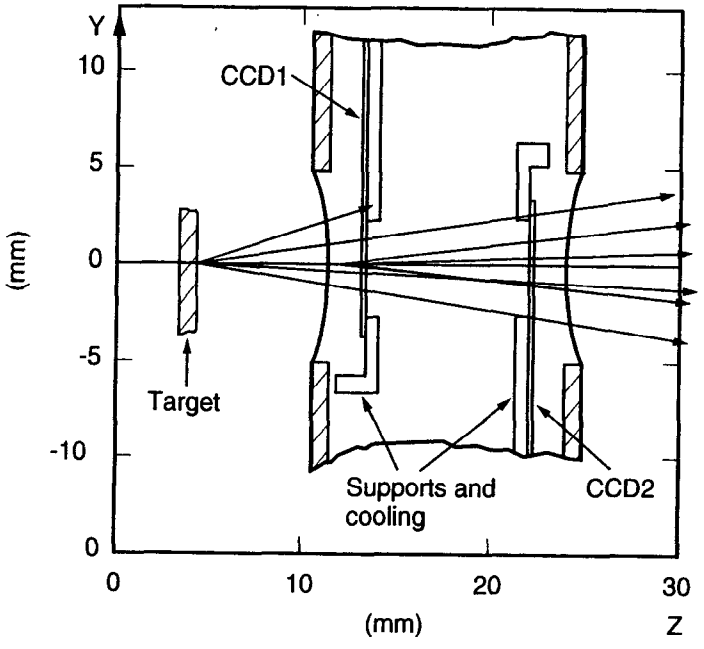


Fig. 41. CCD vertex detector for a fixed-target experiment (NA32). Data are fast-shifted into the quiet regions above and below the spectrometer aperture (CCD's 1 and 2, respectively) prior to readout.

For a linear collider environment such as SLC, the background comes mainly from synchrotron radiation and hence continues to accumulate throughout the readout period. Again, one has the possibility to inhibit this by dumping both beams. This has not been implemented in SLC because the backgrounds are quite tolerable. Furthermore, the trigger rate is sufficiently high that one would experience a significant deadtime loss from this. A CCD detector readout, though slow, can be made inherently deadtimeless; if a second trigger occurs during readout of one event, one just continues reading until data from the second event have been captured completely. Thus, if backgrounds permit, it is more efficient to avoid inhibiting collisions during the detector readout.

For the future linear collider, the SR background can be made negligible, and the small-radius background comes mainly from incoherent e^+e^- pair creation. Here, there are at least two extreme options. Firstly, to use a very small kicker magnet to move one of the beams by about 1 μm, out of collision, during readout. Secondly (if trigger rates are again high so that deadtime losses become an issue), to proceed as in SLC and live with the background. But in this case, one can take advantage of modern CCD design to use a multiport output register (up to 16 ports are commonly available, where in Fig. 32 we have illustrated just one in the corner). This increases the quantity of local readout electronics required, but one can then achieve full detector readout within the period of 5 ms between beam crossings. In practice, once the backgrounds and trigger rates for this environment have been quantitatively evaluated, one will be able to design a CCD vertex detector based on an optimized balance between these extremes.

The purpose of this section has not been to produce specific rules for the design of a CCD vertex detector readout system under specific experimental conditions; both of these are too variable for that. Instead, the hope is to encourage a flexibility of approach and to emphasize the opportunity presented to the vertex detector designer by fully customized CCD design.

5.2.3 Physics Performance and Future Trends

The major *attributes* of a CCD vertex detector are as follows:

1. Two-dimensional space point measurement, hence unsurpassed power for track reconstruction.

2. Two-track resolution. This is approximately 40 μm *in space* (see Fig. 42), compared with about 50 μm *in projection* for a strip detector, some 10^4 times worse.

3. Measurement precision about 3.5 μm for a MIP under typical readout conditions (RMS noise $\approx 50\ e^-$). Note that with less noisy readout (which at present means slower, but other improvements are possible) much higher precision can be achieved. For example, Ref. [62] demonstrates 0.9 μm precision for 15 keV X-rays in a CCD with 6.8 x 6.8 μm^2 pixels.

4. Thin active layer. This implies much lower conversion probability for X-ray background (e.g., synchrotron radiation) than in a thick microstrip detector.

5. Physically thin. Improved performance in terms of multiple scattering.

6. High granularity. Another factor leading to tolerance of high hit density (e.g., in particle jets close to the IP) and to high background. The former quality is demonstrated in Fig. 43, and the latter in Fig. 44.

A striking advantage of the high granularity is the almost total absence of merged clusters. This means that (in contrast to a microstrip-based vertex detector) it is straightforward to write a Monte Carlo program which accurately simulates the detector performance. This is demonstrated in the case of the SLD detector in Fig. 45, which shows the excellent agreement between data and Monte Carlo in the impact parameter distribution projected in orthogonal views. The Monte Carlo program has not needed to be fudged with any empirical smearing function in order to achieve this level of agreement.

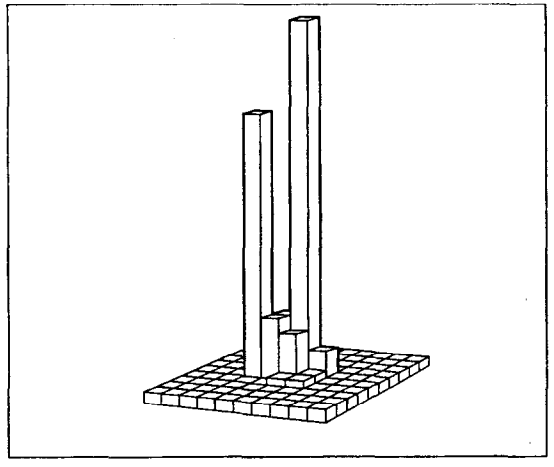


Fig. 42. Two MIP clusters separated in space by 40 μm, well resolved in a single CCD detector plane. Pixel size 20 x 20 μm^2.

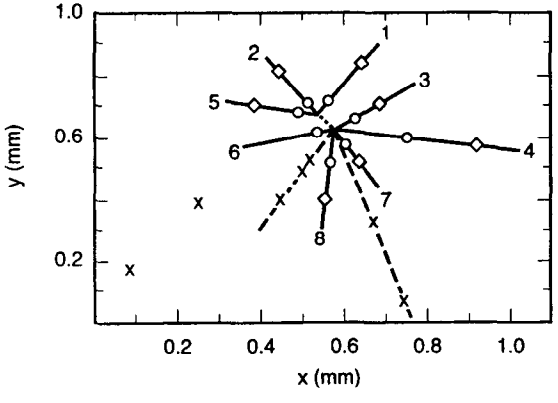


Fig. 43. Tracks from the IP and from a nearby charm decay in the NA32 vertex detector. Frame size 1 x 1 mm^2.

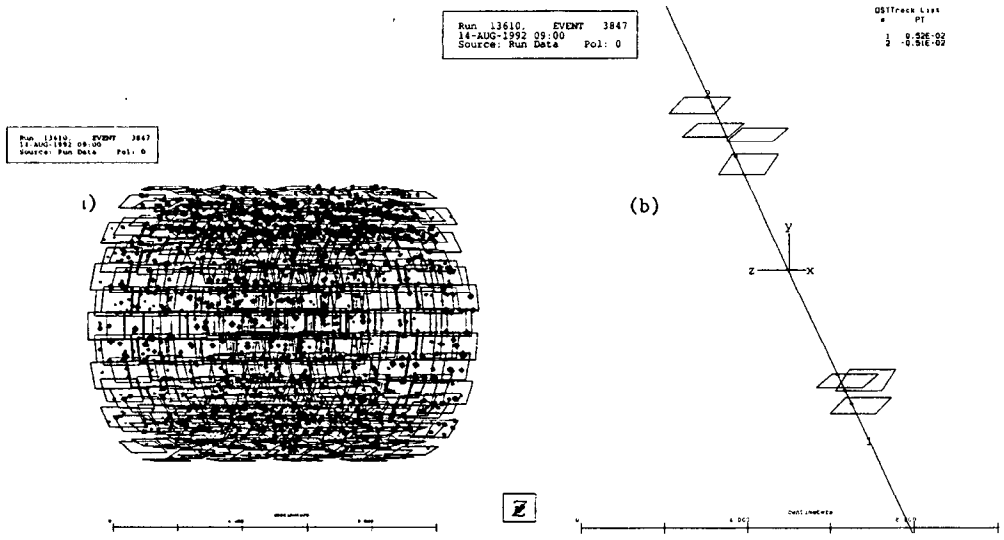


Fig. 44. (a) Raw data (mostly SR X-ray hits) in the SLD vertex detector. (b) The same event, with background filtered out by a drift chamber/vertex detector track-linking algorithm. This proved to be a $Z^0 \to \mu^+\mu^-$ event.

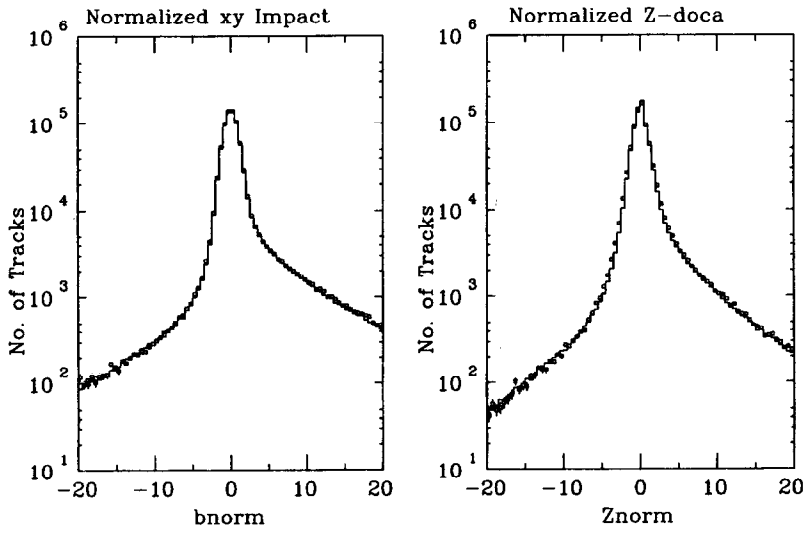


Fig. 45. Data (points) and Monte Carlo (histogram) distributions of impact parameter with respect to the IP in $Z^0 \to$ hadron decays (SLD experiment). The tails on the positive side are due to heavy flavor decays.

The major *deficiencies* of a CCD vertex detector are as follows:

1. Slow readout. This implies either beam suppression and hence a long dead-time associated with every top-level trigger, or a sufficiently benign background rate.

2. Radiation damage. See Sec. 6. In an environment of high hadronic flux, one either has to exchange CCD's fairly frequently (practicable in a fixed-target experiment) or avoid using them (e.g., at a hadron collider).

Both of these deficiencies can, to a great extent, be overcome with APS's (next section), but one then loses some of the previously listed attributes, as we shall see. Each detector type has its own niche.

The availability of fully customized large-area CCD's has opened the door for very exciting vertex detector developments. For example, Fig. 46 shows the CCD being used in the SLD upgrade detector. Adequate readout time is achieved with four outputs in this case. The devices have wire bonds at each end and are arranged end to end, one on either side of a beryllium motherboard, to build up two-CCD ladders out of which the detector (Figs. 47–49) is constructed. See Ref. [63] for a description of this 307 Mpixel detector.

For the future linear collider, one can be more adventurous. The CCD's can be thinned from 150 to 20 μm and attached to the same side of a beryllium stiffener (Fig. 50). By having outputs at one end only, the material in the active volume can be reduced from 0.35% RL per barrel (SLD upgrade) to 0.13% RL; see Ref. [64]. By a combination of larger and thinner CCD's, leading to higher precision-point measurements with more open geometries, one is seeing a steady evolution in the impact parameter precision achievable in the e^+e^- collider environment. The original SLD vertex detector yielded a measured precision of

$$\sigma_{XY}^{b} = 13 \oplus \frac{70}{p \sin^{3/2}\theta}\ \mu\text{m}$$

and

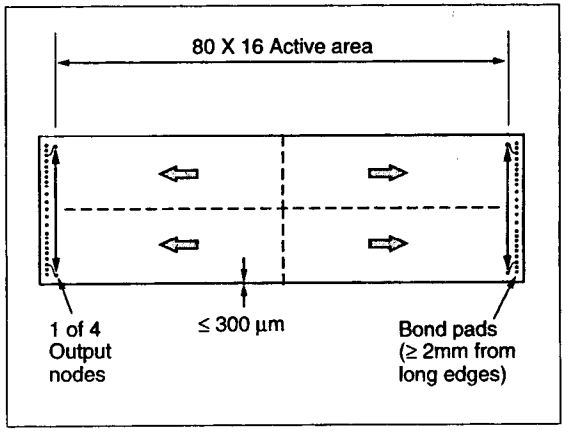


Fig. 46. Four-port CCD developed for the SLD upgrade vertex detector. Chip area = 13 cm^2.

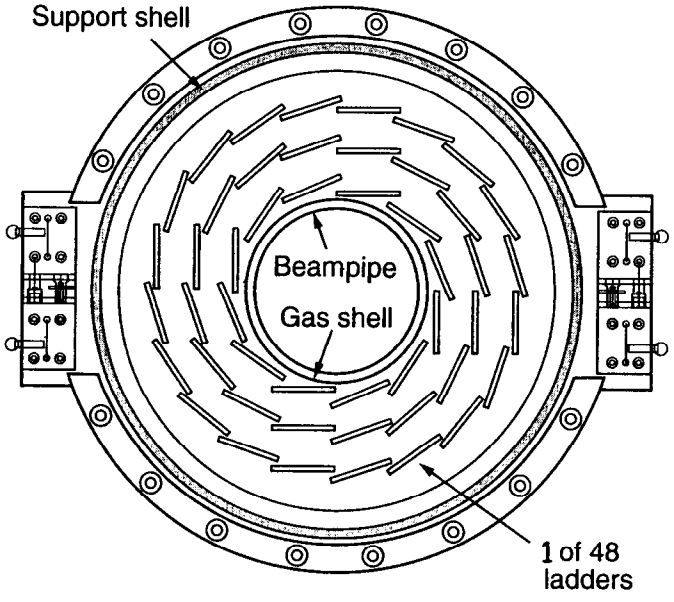


Fig. 47. Cross section (XY view) of SLD upgrade vertex detector.

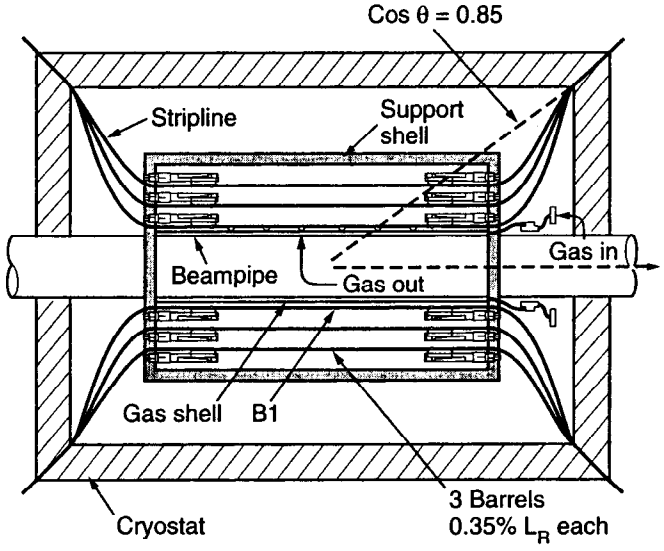


Fig. 48. Cross section (RZ view) of SLD upgrade vertex detector.

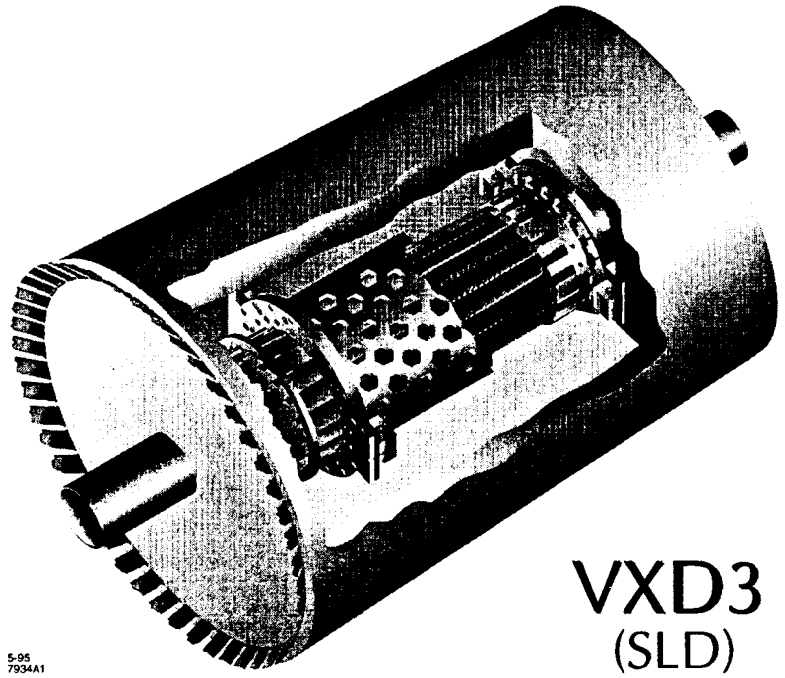


Fig. 49. Isometric drawing of SLD upgrade vertex detector.

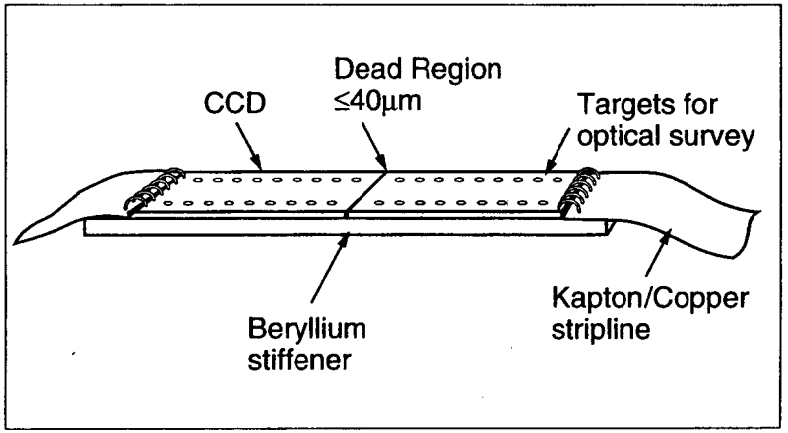


Fig. 50. Advanced two-CCD ladder design for a vertex detector for the future linear e^+e^- collider. Active length ≈ 16 cm.

$$\sigma_{RZ}^{b} = 35 \oplus \frac{70}{p \sin^{3/2}\theta}\ \mu\text{m}.$$

For the future LC detector, we anticipate

$$\sigma_{XY}^{b} = \sigma_{RZ}^{b} = 3 \oplus \frac{5.5}{p \sin^{3/2}\theta}\ \mu\text{m}.$$

Such a detector will be a tracking microscope of unprecedented power, having the capability to open numerous doors for exciting physics discoveries in the realm of Higgs and SUSY particles, as well as exploring the realm of the theoretically totally unexpected.

It should finally be emphasized that the low power dissipation in a well-designed CCD detector (approximately ten watts in the 307 Mpixel SLD upgrade detector) results in very low thermal management overheads. The detector can be cooled with a gentle flow of nitrogen gas, and the cryostat (see Figs. 48 and 49) consists of a low mass (< 1% RL) expanded foam enclosure. The operating temperature of around 200 K is chosen to minimize effects of radiation damage; see Sec. 6.

5.3 Active Pixel Sensors (APS's)

Both in the wider commercial world and in the area of scientific imaging, CCD's have established a dominant role, and as we have seen, are still in the midst of dynamic evolution. Yet they have limitations for vertex detectors, as has been emphasized. In addition, they have limitations for broader applications which have for many years stimulated studies, and more recently, actual devices, constructed according to a completely different architecture, the *active pixel sensor* or APS. The charge collection is as usual to one electrode of a reverse biased diode. But in the APS, these diodes form a discrete matrix over the device area, and each one is connected to its own signal processing circuit within the pixel. These circuits communicate to the outside world via some architecture, usually column-based. The essential point which has taken these devices into the real world has been the continuing shrinkage in feature sizes (and hence transistor sizes) available via the

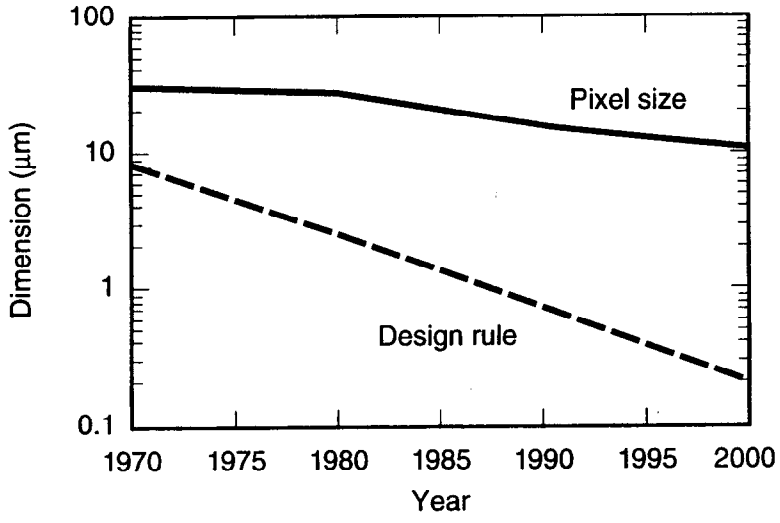


Fig. 51. Evolution of photolithographic feature size versus pixel size.

integrated circuit technology. Figure 51 (after Ref. [65]) illustrates this point. A recent review of developments in this field is to be found in Ref. [66]. Most of the commercial interest has been in the production of inexpensive CMOS chips combining low-quality imagers with processing electronics, for such applications as automatic chrominance control of automobile rear-view mirrors in response to headlights perceived in the field of view. One of the main factors limiting image quality is the separate processing of each channel; it is difficult to match these below 1%, and the eye is very sensitive to such blemishes. In terms of applications such as night vision systems, APS devices do have one interesting advantage over CCD's.

Since the readout can be nondestructive, one can watch on a monitor as the scene gains definition during the exposure time, possibly of advantage for some surveillance applications. However, commercial CMOS sensors made on low-resistivity material are typically limited to 1 or 2 μm detector active thickness, and hence are not useful for MIP detection. In addition, the growth in parasitic capacitance as the area is scaled up leads to escalating power requirements. Devices of area 100 x 100 pixels are relatively easy; beyond that, it becomes difficult. Finally, the spectacular evolution in design rule dimensions is generally associated with *smaller* IC's. Building sensors of area many square centimeters to such rules remains a distant dream. All of these factors do cause problems in the development of APS devices as vertex detectors.

For MIP detection, there are two main options. One of these is to take a high resistivity wafer and manufacture a single-sided microstrip-type detector, but with the strips cut into pads of the desired pixel dimensions, and to bump-bond this detector to a CMOS readout chip. This *hybrid* approach implies the less challenging route of keeping two technologies separate, rather than working to combine them. The second option, the *monolithic* approach, seeks to do the job on one chip. In both cases, the detector goals are similar and can be summarized as follows:

1. High-speed gating. In contrast to CCD's, the aim is to latch signals and associate them with specific beam cross-overs (BCO's) in environments such as LHC (BCO interval 25 ns), where the hit densities from each BCO are so high that one could not afford to integrate signals over more than one.

2. Time stamping. The idea is to transfer the hit information into a pipelined memory clocked at the BCO rate. On receipt of a level-1 trigger, those pixels that were hit at the corresponding BCO will be transferred to an on-chip buffer for readout, in the event that a level-2 or level-3 trigger is asserted.

3. Radiation hardness. Since (unlike the CCD) signal charge is not transported long distances through the silicon, the effects of bulk damage in terms of charge trapping are much reduced.

Leakage current impact is much reduced relative to microstrip detectors, due to the much smaller collection volume per detector element.

However, one does not escape the problems of type inversion and loss of charge collection efficiency (see Sec. 6). Furthermore, one has the same concerns regarding radiation effects in CMOS electronics (now in the active volume of the detector) as we noted in the microstrip environment.

As with microstrip detectors, there are three possible options for the readout electronics (binary, digital, and analogue), all of which are being actively pursued.

A major goal for physics is to be able to operate at relatively small radius (approximately 10 cm) for a reasonable lifetime in LHC at full luminosity. Several European and US groups are actively involved; for a recent review of the European work in this area, see Ref. [67].

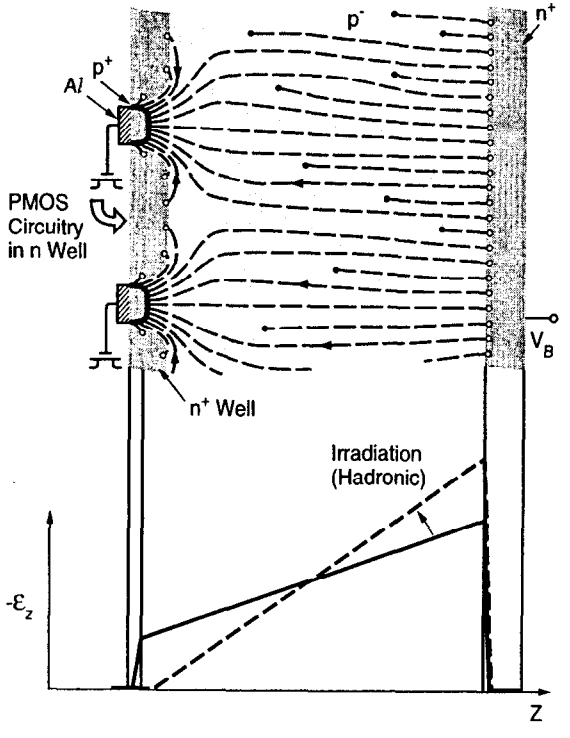


Fig. 52. Pixel structure of the generic monolithic APS. As with the microstrip detector (Fig. 25), hadronic irradiation tends to take the detector out of depletion, losing the signal.

5.3.1 Design Options

Let us consider in turn the two options available for MIP detection systems.

5.3.1.1 Monolithic Detectors

The generic monolithic detector pixel structure is sketched in Fig. 52. Full charge collection over the active area is achievable despite the fact that the p^+ collection implants occupy typically less than 10% of the surface area.

The main hurdle to overcome in moving from the commercial CMOS imager to a MIP-sensitive device was achieving compatibility between the high-temperature processing used for the CMOS activation steps and the preservation of high resistivity of the detector-grade silicon. This was demonstrated by Holland in a pioneering paper [68], in which the process of backside gettering is used for the removal of detrimental impurities from critical device regions. A similar process has been used since the mid-'80s in CCD manufacture, in which the heavily doped substrate is used to getter impurities from the epitaxial region from which the signal charge is collected.

To date, one prototype monolithic detector has been produced and demonstrated its capability for MIP detection [69]. This is an array of 10 x 30 pixels, pixel size 34 x 125 μm^2, overall area 1 mm^2. Ten percent of the chip area around two edges is taken up with CMOS circuitry. The analogue signals are read out sequentially at 1 MHz. Excellent MIP efficiency is achieved, with precision 2.0 μm x 22 μm in the two orthogonal directions. As with the commercial CMOS imagers, a considerable challenge is involved in scaling up the device size, but already a second generation detector of 96 x 128 pixels is under development [70]. European groups are also actively developing monolithic pixel detectors, aiming for the application to LHC vertex detectors.

5.3.1.2 Hybrid Detectors

Hybrid APS devices are being developed by several US and European groups for use in LHC detectors. The detector part consists of essentially a microstrip detector

structure, each strip being subdivided into a series of short strips which constitute the pixels. These are bump-bonded to the collection electrodes of a CMOS readout chip which would be similar in architecture to the monolithic versions. Hybrid detectors have the advantage of relative simplicity (no need to combine the detector and readout functions on one chip), but the complication of millions of interconnections and the disadvantage of extra material in the active volume. The thickness problem is exacerbated for both APS options by the high power dissipation (designers are aiming for about 0.5 W/cm^2, about 100 times higher than a CCD detector). Liquid-filled cooling tubes within the active volume are required.

Already one hybrid detector with 300 kpixels (of size 75 x 500 μm^2, too large for a vertex detector) is in use as a tracking detector in a high-track density, fixed-target environment [71]. This detector produces a binary output from each pixel at a readout rate of 2 MHz and has demonstrated excellent performance as a tracking detector. A second generation detector, shrinking the pixel size somewhat to 50 x 500 μm^2 while increasing the number of transistors per cell from 80 to 350 (using submicron technology), is in design. Zero suppression on-chip will greatly accelerate the speed of readout. These are vitally important steps en route to a viable LHC detector.

5.3.2 Performance and Future Trends

APS detectors for MIP detection are at a relatively early stage of development. They are demonstrating their capability in test beams and in fixed-target experiments as general tracking detectors. Their advancement to the level of an LHC vertex detector (see, for example, Fig. 53), with 100 Mpixels, depends on several challenging developments. Firstly, the functionality referred to earlier needs to be achieved in pixels of a reasonably small size, at least in one dimension (so that precise measurement in the $R\phi$ plane becomes possible). Secondly, the CMOS electronics needs to be sufficiently radiation hard, and finally, the detector needs also to demonstrate adequate radiation hardness. In fact, for the hybrid approach, one has in principle the option of going beyond silicon (see Sec. 7) for the detector, while retaining the rad-hard electronics for the readout. Overall, this is a very dynamic area of detector development, with an assembly of talented groups well-matched to the considerable challenges involved. Furthermore, even though the present prototypes are far from the eventual goals, ideas keep emerging and hold promise for ongoing important developments. An interesting new idea (Ref. [72]) involves the use of a p-channel JFET on a fully depleted high ohmic substrate (DEPJFET) for use as a unit cell for pixel detectors.

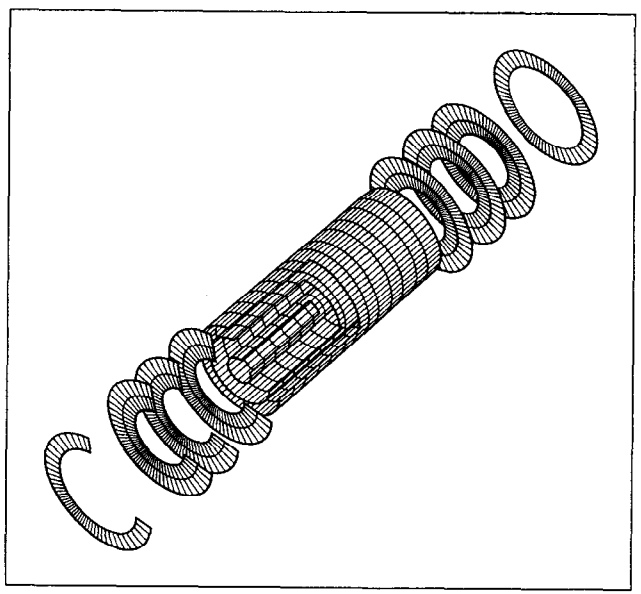


Fig. 53. Conceptual GEANT layout of a pixel vertex detector for ATLAS, consisting of three barrels plus endcaps. The innermost barrel (r = 4 cm) is not expected to survive for long at the full luminosity.

6 Radiation Damage in Silicon Detectors

6.1 Introduction

The subject of radiation damage in silicon devices has been studied intensively for decades, particularly in relation to the effects of nuclear reactors and weapons, both in the form of ionizing radiation and neutrons. References [73] and [74] are very useful books on the subject, Ref. [75] provides a valuable current review, and interesting historical reviews can be found in Refs. [76] and [77]. Yet, far from being exhausted, this is an extremely active area of study in connection with silicon tracking detectors. Why is this?

Firstly, silicon detectors are generally made from high resistivity material having long minority carrier lifetimes (order of magnitude milliseconds). Such material, unfamiliar to the field of electronic devices, behaves in unusual ways when irradiated; in general, it is more sensitive than electronic grade material to radiation effects. Secondly, there is an increasing number of important scientific applications (space-based equipment which spends time in radiation belts, detectors at small radius in LHC, etc.) for which the radiation environment is unusually hostile.

If we start by considering electromagnetic radiation of energy E_γ at long wavelengths (e.g., visible light), the effects in silicon devices (electron-hole pair generation) are entirely transient. Above about 10 eV, electron-hole pairs in silicon dioxide are generated. These nearly all recombine, but as E_γ is increased, the hot carriers have an increasing probability of becoming independent within the oxide layer, leading to some degree of *surface damage*. Once E_γ exceeds approximately 250 keV, the energy is sufficient to start dislodging silicon atoms from their lattice sites; we are entering the realm of *displacement damage*.

For massive charged particles, displacement damage sets in at much lower energy. Low-energy protons are extremely dangerous due to the large cross section for p Si Coulomb scattering.

These two mechanisms form the basis of all radiation damage effects that concern us in regard to silicon detectors and the local electronics supporting them. Yet the

possible range of consequences of these effects is rather diverse. Let us consider these in some detail.

6.2 Ionizing Radiation

The band gap in silicon dioxide is 8.8 eV, and on average, 18 eV is needed to release an electron-hole pair. Figure 54 shows the time development of the charge distribution in an irradiated MOS structure.

The radiation generates a charge Q_g in the oxide, where $Q_g \propto t_{ox}$. The magnitude of this charge is totally independent of the nature of the oxide, rad-hard or "soft." A fraction f_c of the charge is trapped at the interface (where f_c can vary from 2% for a hard oxide to 80% for standard oxide), giving a trapped charge $Q_{tr} = f_c Q_g$. This induces a flat-band voltage shift ΔV_{FB}, where

$$\Delta V_{FB} = Q_{tr} / C_{ox}.$$

Now $C_{ox} \propto 1/t_{ox}$, so

$$\Delta V_{FB} \propto t_{ox}^2.$$

Below 1200 Å, the dependence can be even faster, approximately as t_{ox}^3.

Note that this time development follows from the vastly different room temperature mobilities of electrons and holes in silicon dioxide, 2×10^5 cm^2/Vs and 20 cm^2/Vs, respectively.

As well as contributing a direct *interface charge*, the trapped holes can induce *interface states* in the case that they have been drifted towards the bulk silicon (as in Fig. 54). The interface state charge may be positive (for *n*-type substrates, i.e., *p*-channel MOS devices) or negative (for *p*-type substrates, i.e., *n*-channel MOS devices).

Note that at reduced temperature, the holes are effectively immobilized, so there is no performance difference between soft and hard oxide. This, however, is not a serious

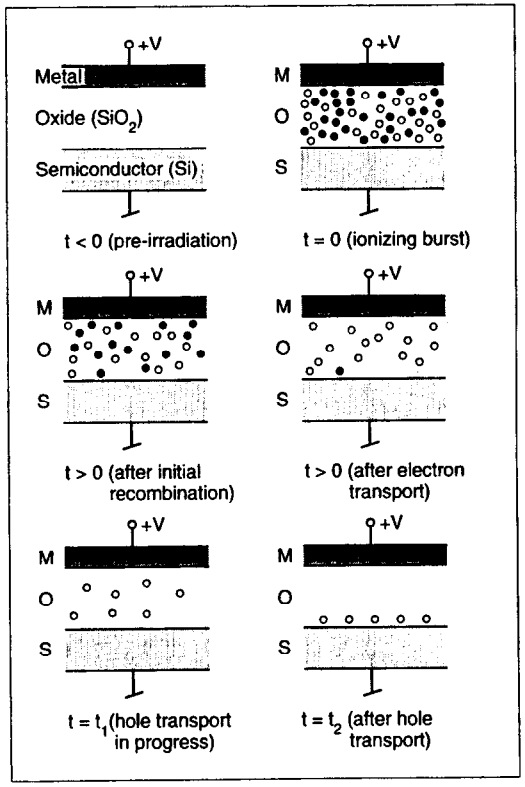


Fig. 54. Time development of charge distributions following a burst of ionizing radiation on a positively biased MOS structure.

concern for detector applications, since the detector can always be cycled up to room temperature for brief periods, restoring the holes to their normal room-temperature evolution.

The induced flat-band voltage shifts can cause various device and detector malfunctions. For nonhardened oxide, the effects are large; for example, 10 krads on a 700 Å oxide induces a 2 V shift. What can be done to reduce this?

> Firstly, the 100 substrate orientation is much preferred (minimal level of dangling bonds).
>
> Secondly, minimize t_{ox}, though not so far as to suffer a serious loss of device yield.
>
> Finally, observe special procedures in post-gate processing (most notably, keeping the temperature below 900 °C).

As well as the *gate oxide*, charge buildup in regions of *field oxide* on the device can be equally significant [78]. Huge voltage shifts are associated with the thick field oxide. In the case of *p* substrates, these induce inversion layers which can short all the *n* implants within the substrate. These effects are common to all device types (JFET's, bipolars, MOS devices, and detectors). Careful design practices (e.g., guard structures) are required to avoid them.

Recent developments may lead to a further breakthrough in the area of radiation hardening. It has been found that the conventional use of *hydrogen* to saturate dangling bonds may not be optimal. The Si-H bond is unstable with respect to X-radiation. To this end, a new process has been developed [79] based on semi-insulating polycrystalline silicon or SIPOS. Possible implications for radiation detectors are being evaluated.

6.3 Displacement Damage

Atomic collisions with high momentum transfer, as well as nuclear interactions, can permanently alter the properties of the bulk material. Such processes are grouped together as the source of *displacement damage*, in which silicon atoms are displaced from their normal lattice locations. These effects may be local single-atom displacements, in which case the damage is classified as a *point defect*; such defects commonly result from high-energy electromagnetic irradiation (X-rays or electrons). Displacement damage may also occur as *damage clusters* which consist of relatively large disturbed regions within the crystal; such defects commonly result from nuclear interactions of (for example) neutrons and protons. The most probable events of this type are elastic Coulomb scattering of silicon nuclei by the incident high-energy (charged) particle. As shown in Fig. 55 (based on Ref. [73]), a 50 keV recoil silicon nucleus can create clusters of damage (with knock-on and stopping of other nuclei) over a volume of several hundred Ångstroms typical dimensions.

The bulk damage due to the passage of high-energy particles can be described by the number of atomic (silicon) displacements per cm of track length. For protons traversing silicon, this rate falls from $\approx 10^4$/cm at 1 MeV to $\approx 10^2$/cm at 1 GeV. This nonionizing energy loss (NIEL) depends both on the particle type and energy, though at high energy (above approximately 1 GeV), it is nearly the same for all hadrons. See Refs. [80] and [81] for pioneering papers on this subject. The NIEL for various particle types is plotted in Fig. 56. To a good approximation, displacement damage effects depend only on the overall nonionizing dose received, except that the effects are much reduced for electromagnetic radiation. In this case, as well as the low specific NIEL value, all momentum transfers are so low as to liberate at most one atom (leading to point defects as opposed to cluster damage). Specifically for 5 MeV particles, an electron, proton, and neutron produce a primary knock-on atom (PKA) which on average generates in total 1.2, 4.2, and 8000 further displacements, respectively.

As far as the primary displacement damage is concerned, the generation of these clusters of vacancies (V) and interstitial silicon atoms (I) is the entire story. Even in low-resistivity material, the concentration of dopant atoms is so low that they play effectively no part in this process. However, the role of dopant and impurity atoms is crucial in understanding the electrical effects, because both vacancies and interstitials are mobile, and can combine stably with atoms other than silicon in the crystal structure.

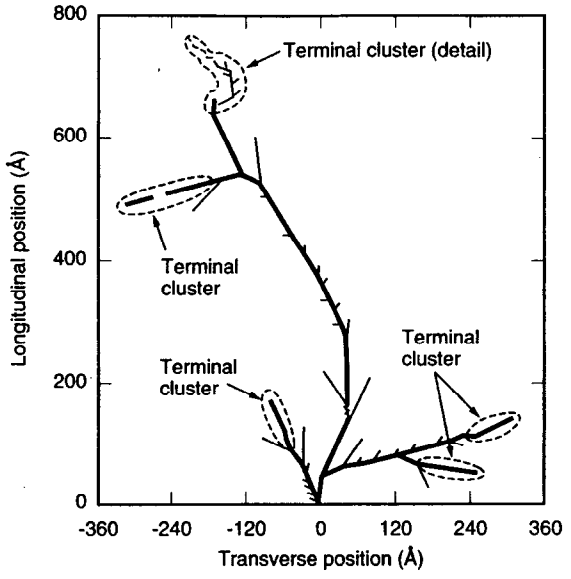


Fig. 55. Development of cluster damage due to a primary knock-on silicon atom of 50 keV, within the bulk material.

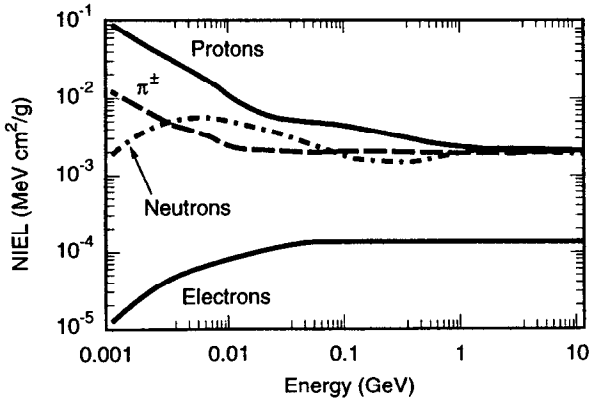


Fig. 56. NIEL for various particle types as a function of energy. A frequently used unit is the NIEL associated with a 1 MeV neutron.

Before considering this, we note that the practical effect is the development of a large number of energy levels within the band gap, some donor-like and some acceptor-like, some being capable of existing in more than two charge states. These levels, depending on their state of occupancy, can act as trapping centers and hence seriously degrade the minority carrier lifetime. In addition, these extraneous generation-recombination centers cause extra leakage current in depleted material and reduction in the carrier mobility. For electronic grade silicon, the description of displacement damage effects in terms of these macroscopic properties is sufficient.

For detector-grade material, the situation is more complex. It is rather like comparing the effects of an earthquake on a steel frame building as opposed to one made with bricks. The basic physics processes are the same, but the effects are very different. Detector-grade material (high resistivity, long minority carrier lifetime) is particularly sensitive to radiation-induced displacement damage. Let us start with an empirical description of what is observed, and then tackle the basic physics processes involved.

Measurements on *undepleted* detector-grade silicon reveal a monotonically increasing *rise* in resistivity with dose. This can be understood in that the disordered material generates a huge number of extra donor and acceptor states, populating the entire band gap. Statistically, the Fermi level drifts to approximately midgap, so the material becomes effectively *compensated.*

However, when one depletes the material, one finds a leakage current which grows linearly with dose (i.e., accumulated NIEL) but which anneals with more than one time constant. One is seeing the global effect of generation current from a number of intergap states which physically evolve with time. *Provided* the detector is designed for low-temperature operation, the leakage current is not a fundamental problem, since one can reduce it to an acceptable level by cooling.

Next, we consider the effective dopant concentration N_{eff}. From the resistivity measurements, we might have expected the material to change from n type to intrinsic, and to stabilize with a low value of N_{eff} as the Fermi level sits around mid-gap. On the contrary, as we saw in Figs. 25 and 52, the depleted material behaves

quite differently from the material in equilibrium. It becomes steadily more *p* type with fluence, going through *type inversion* at an equivalent fluence of approximately 5×10^{12} neutrons/cm^2, as shown in Fig. 57. As we saw in the case of the leakage current, the material shows a medium-term annealing behavior, which is extremely temperature dependent [82–84]. For highly irradiated samples (well beyond type inversion), N_{eff} falls back over a period of days (at room temperature) or years (at -20°C). However, this is by no means the end of the story. At room temperature, the material now enters a *reverse-annealing* phase; N_{eff} increases. The material becomes ever more *p* type; even after a year, the trend continues. This behavior can be entirely avoided by cooling. The data taken at -20°C show ongoing annealing to the end of the test period, with no tendency to flatten off; the material just becomes steadily more nearly intrinsic.

So what are the microscopic physics processes during this complex behavior pattern? One could even ask, why do we care? The answer to the second question is that there is a possibility that, once the details are understood, it may be possible by *defect engineering* to improve the radiation hardness of the material, e.g., by staving off the reverse annealing problem even at room temperature. This is a very active area of research. At a recent conference, contributions were varied and somewhat controversial [85]. DLTS measurements backed up by a semiconductor device model have enabled Matheson *et al.* [86] to produce a plausible explanation for some of the most striking of the above observations. Their results can be summarized as follows:

1. Based on photoluminescence and DLTS measurements on high resistivity *n*-type Wacker material, they find the following concentrations of expected and unexpected impurities:

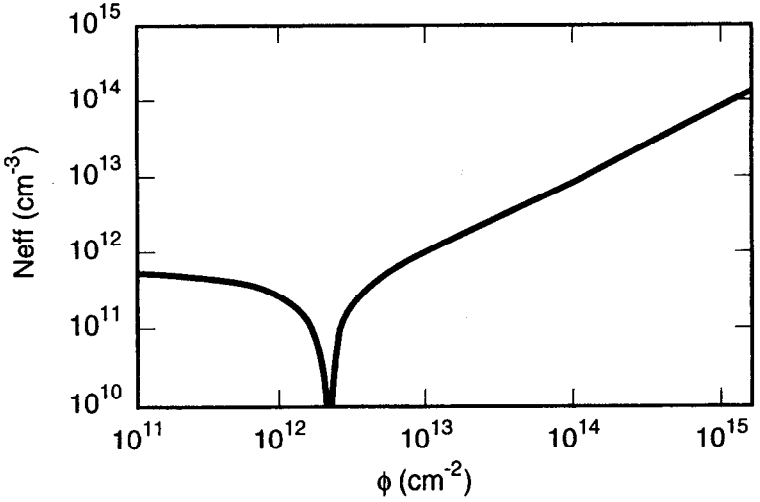


Fig. 57. Dependence of effective dopant concentration N_{eff} on fluence, at room temperature. The material, initially *n* type, goes through type inversion for $\phi \approx 5\times10^{12}$ neutrons / cm^2 equivalent dose.

[P] $\approx 10^{12}\,cm^{-3}$

[C_s] $1-5\times10^{13}\,cm^{-3}$ (substitutional carbon)

[O] $5\times10^{15}\,cm^{-3}$

[H] $10^{14}\,cm^{-3}$.

2. The mobile I and V centers diffuse away from the damage cluster and eventually mostly undergo one of the following reactions:

$I+V\rightarrow 0$ (repair)

$I+C_s\rightarrow C_i$ (interstitial carbon, transient)

$C_i+C_s\rightarrow CC$

$C_i+O\rightarrow CO$

$O+V\rightarrow VO$ $Si-A$ center (99.9%)

$P+V\rightarrow VP$ $Si-E$ center (0.11%)

$V+V\rightarrow VV$ (but most divacancies emerge from the primary cluster)

$VO+V\rightarrow V_2O$.

These observations rule out some of the almost-established folklore regarding the behavior of detector-grade material. The long-held belief that the resistivity rise was due to donor removal is excluded by the above figures. The phosphorus concentration is simply too low by several orders of magnitude.

3. The authors hypothesize that generation of some deep level acceptor is responsible for the reverse annealing. V_2O is a candidate, suggesting that a less oxygen-rich starting material might be free of this effect.

4. If such a deep-level acceptor is responsible, how does it become filled? The authors hypothesize that this is due to the bulk leakage current, and indeed demonstrate a suggestive correlation between the measured N_{eff} values during the annealing phase and the square root of the leakage current damage constant α. If this were the only effect involved, one would find simple proportionality between these. In fact, there is a nonzero offset, but it seems to me likely that this mechanism is a good part of what is a rather complex picture.

These pioneering studies have led to a concerted effort by LHC physicists to further understand the bulk radiation effects in detector-grade material, possibly leading to more radiation-resistant material in the longer term future.

The final empirical observation relevant to bulk damage effects in detectors is that of loss of *charge-collection efficiency*, CCE. For a 300 μm thick depleted detector, one finds approximately a 10% loss in CCE for a dose of $10^{14}\,n/cm^2$ equivalent. This is presumably related to the high density of trapping centers generated and probably implies a basic limit to the tolerable radiation dose for such thick detectors, at around the $10^{15}\,n/cm^2$ level.

6.4 Detector-Specific Effects

6.4.1 Microstrip Detectors and APS Devices

The major challenge which is driving much of the R&D discussed in the previous section is the LHC tracking detectors (vertex region and Central Tracker at larger radii). At small radius, the predominant background comes from pions of energy 100 MeV to 1 GeV, with albedo neutrons playing a relatively larger role at large radii [87]. The overall dose as a function of radius is listed in the following table, for a seven-year run at $\mathcal{L}=1.7\times10^{34}\,cm^{-2}s^{-1}$.

R (cm)	Ionizing Dose (Mrads)	Fluence (1 MeV $n/cm^2\times10^{-14}$)
4	117	40
6	55	18.8
10	22	7.4
11.5 L1	17	5.8
14.5 L2	11.4	3.9
30	3.8	1.3
52	1.8	0.6

If detector replacement during this period is assumed, one is entitled to divide by some factor, but there are reasons (beam-filling periods, etc.) to raise the estimate. Overall, these figures probably give a reasonable indication of the requirements.

Discounting, for the purposes of this discussion, the prospect of major progress through defect engineering, what do these figures imply for silicon tracking detectors in such an environment? (While we are discussing this in the context of LHC, the implications for other hadron beam or collider experiments follow directly.)

Within a radius R = 30 cm, one suffers increasingly serious CCE loss. This would be fatal for microstrip detectors. However, silicon pixel devices, with much smaller collection node capacitance, might be able to survive with a considerably smaller signal size, i.e., smaller depletion depth.

Beyond R = 30 cm, the detectors still go far beyond type inversion during their working life. This means one of two things. Either they are made on *p*-type substrates or they must be equipped with guard rings, etc., that allow the junction to move from the *p* side to the *n* side during operation. If one collects signals from the *p* strips (hole signal), one has to beware of loss of signal as the radiation dose increases (remember Fig. 25). This can be avoided by steadily increasing the operating voltage. Alternatively, one may collect the signal from *n* strips (electron signal), in which case the charge collection degrades more gracefully, as the devices fall below depletion. In either case, to prevent the global signal from falling too low, it is necessary to keep the devices at least almost fully depleted. This implies (for R $\geq$ 30 cm) high operating voltage (approximately 1 kV) at the end of the seven-year period, unless the detectors are cooled. Cooling to say -10°C can keep the depletion voltage down to approximately 150 V as well as providing the essential reduction in leakage current. However, if the detector is warmed up for a total of even one month during the seven-year period, the depletion voltage increases by a factor of two, due to rampant reverse annealing during that time.

In conclusion, environments such as LHC with high hadronic background provide a major challenge for silicon detectors. By switching from microstrips to pixels, one can hope to push below R = 30 cm, but within R = 10 cm, the region of interest for a general-purpose vertex detector with good impact parameter resolution, even these devices would not have a useful life expectancy at the full LHC luminosity. The most optimistic current expectation is for an inner layer of pixel detectors on R = 11.5 cm, with an active thickness of 150 μm and (at end-of-life) a depletion voltage of 350 V, 2 nA/pixel leakage current, and 30% ballistic deficit.

The hopes of being able to move into the heat below 10 cm have stimulated a considerable activity in devices made of material beyond silicon, as discussed in Sec. 7.

6.4.2 CCD's

For use as vertex detectors, CCD's have a role mainly in fixed-target experiments (where they are required to cover only a small area and hence can be changed frequently) and in e^+e^- collider experiments, where the hadronic backgrounds are low. Hence, our major concern is their tolerance of ionizing radiation. However, for other applications (notably space-based detectors that suffer from solar flares or spend time in the proton-radiation belts around the earth), the hadronic bulk damage effects can be serious.

Regarding ionizing radiation, the effect to be concerned with in CCD's is the slow shift in the potential of the parts of the device overlaid by gate oxide (the imaging area and output register), in relation to the potential of the output node (nominally fixed). Figure 58 (based on Ref. [88]) shows the flat-band voltage shift after irradiation of a CCD gate oxide at two extreme temperatures. For an *n*-channel CCD, the sign of the electric field is optimal (directed towards the gates, negative in the convention of Figure 58). Thus, at room temperature, the flat-band voltage shift ΔV_{FB} is negligible. However, the situation worsens as the temperature is reduced, and by 77 K, ΔV_{FB} is huge and equally bad for either polarity. Note that even at low temperature, ΔV_{FB} is negligible for an unbiased gate, so CCD's (and, in fact, any MOS devices) in radiation environments should be powered off when not in use. Furthermore, for devices operated cold, an occasional, brief warm-up to room temperature restores ΔV_{FB} to a much reduced level. One can, in addition, tune the output node voltage within limits. Modern standard production CCD's have

$\Delta V_{FB} \approx 20$ mV/krad and can be tuned for operation up to 100 krads. More advanced devices are now proven up to 1 Mrad of ionizing radiation.

In all this, it is extremely important that the polysilicon gate structure completely overlays the oxide layer. Figure 36 is an oversimplification; the actual CCD structure is sketched in Fig. 59.

Regarding bulk damage, we need to consider the effects on dark current, charge collection efficiency, and charge transfer efficiency. Even in heavily irradiated CCD's, the excess dark current can normally be dealt with by modest cooling. Given the thin epitaxial layer, the requirements made on minority carrier lifetime are not severe, and there is essentially no problem with CCE into the potential wells. However, once the electron charge packet starts its long journey to the output node, the situation is far more dangerous. The *n* channel being relatively highly doped, the generation of bulk defects is considerably simpler than was discussed for detector-grade material, being closely similar to that encountered in electronic devices. The mobile vacancies are predominantly captured by the phosphorus dopant atoms, giving an increasing density of Si-E centers (positively charged donor-like defects when empty; with an energy level E_{tr} of 0.44 eV below E_c). These defects have a high probability of capturing signal electrons which come within their electrical sphere of influence. Let us consider this case, a single type of bulk trap which uniformly populates the *n* channel. This situation is a restricted case of the general Shockley-Hall-Read theory of carrier capture and emission from traps, in which only capture and emission of electrons from/to the conduction band plays a part. Hole capture and emission are irrelevant since we are concerned with donor-like traps in depleted material. This situation has been considered by various authors [53, 54, 89, 90].

Let us firstly take a qualitative look at the situation. As the charge packet is transported from gate to gate (within a pixel or between neighboring pixels), *vacant* traps that lie within the storage volume of the charge packet will tend to capture electrons. If the traps are filled (either fortuitously, due to the passage of an earlier signal packet, or deliberately for this purpose by the injection of an earlier "sacrificial" charge packet), they will permit the signal electrons to pass undisturbed.

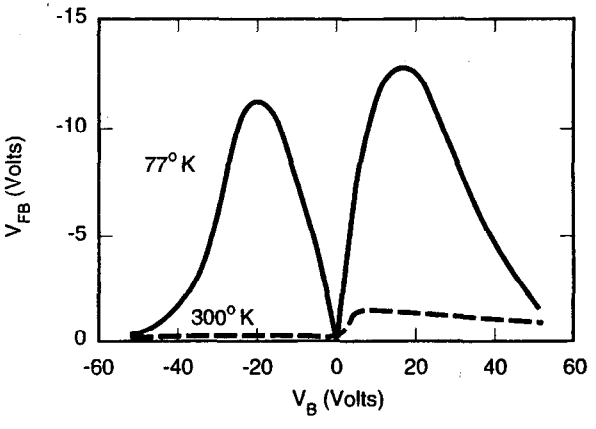


Fig. 58. Flat-band voltage shifts after 100 krads of ionizing radiation across a hardened gate oxide.

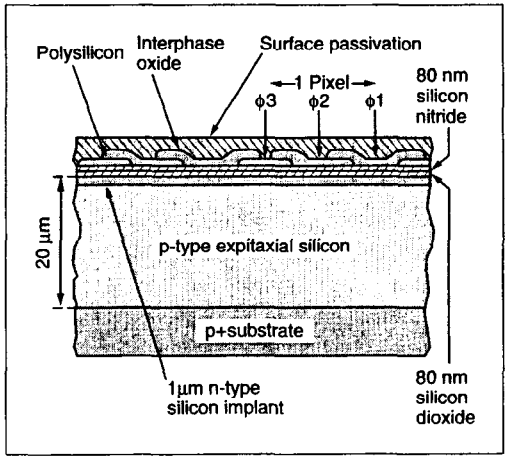


Fig. 59. Gate structure of a modern three-phase CCD register, designed to avoid potential wells due to radiation-induced charge build-up or other spurious charge in the oxide or surface passivation layers.

Also, if the signal packet is transported at a sufficiently high clock rate that the dwell time τ_g under any gate is small compared to the trapping time constant τ_c, the signal electrons will pass. Also, if the trap emission time constant τ_e is small compared with the clock pulse rise/fall time τ_r, the trapped electrons will be re-emitted in time to rejoin their parent charge packet. Only if electrons are *trapped and held long enough* to be redeposited in the next or later potential well, does the process contribute to a loss of CTE. This is evidently a multiparameter problem with some room for maneuver.

Let us now look at the process quantitatively. Assuming all traps are initially empty, the CTI is given by

$$\mathrm{CTI} = \sum_{j=1}^{N_F} F_j \times \frac{N_{tr}}{N_s}\left[1 - \exp\left(\frac{\tau_r}{\tau_e}\right)\right].$$

N_F is the number of phases per pixel (three for a three-phase structure).

F_j is the fill-factor for phase j, i.e., the probability that a trap in the charge packet storage volume will become filled during the dwell time.

$$F_j = 1 - \exp(-\tau_g / \tau_c).$$

For cases of practical interest, τ_c is of order of magnitude nanoseconds, and F_j may be taken to be unity. N_{tr} is the trap density. N_s, the signal charge density, is a function of the signal size but is effectively constant for charge packets larger than approximately 1000 e^- (Ref. [90]). For smaller charge packets, the effective signal density is reduced, and the CTI is correspondingly degraded. For very small charge packets of N_e electrons, one expects $N_s \propto 1/N_e$ since the signal electrons will occupy a constant volume determined by their thermal energy and the three-dimensional potential well in which they are stored. The volume of this potential well can be reduced (by techniques referred to as narrow channel or supplementary channel processing), so yielding a factor of up to four improvement in CTI, compared with standard channel devices [91].

Now

$$\tau_e = \frac{\exp[(E_c - E_t)/kT]}{\sigma_n X_n v_n N_c}.$$

The terms in the denominator are respectively the electron capture cross section for that trap type, an entropy factor, the electron thermal velocity, and the effective density of states in the conduction band. The numerator tells us that for shallow traps (or high temperature), τ_e is likely to be short, and conversely for deep traps and/or low temperatures, τ_e is likely to be long. In fact, for deep traps and appropriate clock times, by reducing the temperature, one can sweep the CTI through its full range from approximately zero (since the charge is re-emitted into the parent pixel during the drive-pulse risetime) to $3N_{tr}/N_s$ (for a three-phase CCD) and back to zero, as all traps are filled by some long preceding deliberate or accidental charge packets to have been clocked out of the device. Figure 60 (from Ref. [90]) nicely illustrates this point. This demonstrates the growth in CTI due to irradiation of a CCD with a high-energy electron source. The density of Si-E centers increases, but the effect on CTI can be minimized by operating at or below 190 K, where the trap emission time becomes adequately long. The degradation in CTI below 160 K is due to the emission time of a shallower trap becoming significantly long. Eventually (by about 70 K), the phosphorus donor ions can play a role (carrier freeze-out). This sets an effective lower limit to the useful operating temperature of *n*-channel CCD's.

For hadronic irradiation of CCD's, because of the much greater NIEL factor, the damage rates are greatly increased. The CTI effects are qualitatively similar [92], and it is believed that the Si-E center is responsible for 85% of the defects, with 15% due to the VV (divacancy) presumably generated in the initial damage clusters. There are possibly some further discrepancies with respect to the electromagnetic damage data; what is urgently needed are controlled experiments, involving both electromagnetic and hadronic irradiation of the same CCD types under similar clocking conditions, with well-defined injection of "sacrificial" charge packets to (as far as realistically possible) saturate the traps. One should also note the necessity to study the serial *and* parallel register in any test program. One might select a temperature low enough to

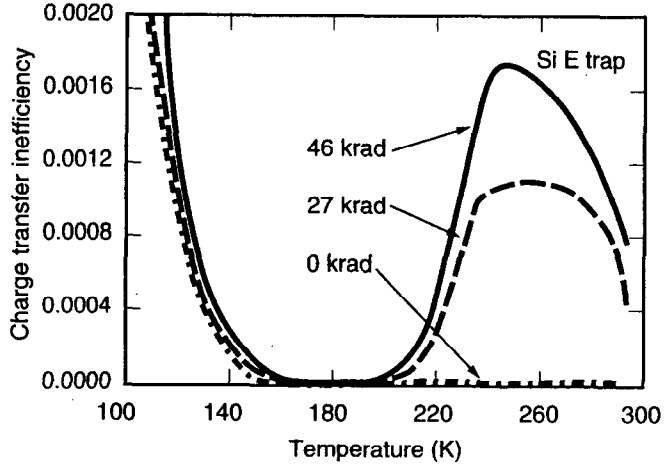


Fig. 60. From Ref. [90], effect of radiation damage on CTE in a CCD, as a function of operating temperature. Irradiated with a Sr^{90} β source.

have good parallel CTE against all traps, but find that this corresponds to long enough emission times for some intermediate depth trap to cause serious CTE loss in the serial register. There is no absolute rule that the serial register CTE exceeds that of the parallel register, though this is often the case.

6.4.3 Local Electronics

The issue of radiation hardness of local electronics for vertex detectors is extremely dependent on the detector type as well (of course) as on the nature of the experiment. In fixed-target experiments, it is no problem to keep the electronics out of the beam, so the issue does not arise. In collider experiments, it has already been mentioned that for CCD-based vertex detectors, it is desirable for thermal management reasons to keep the local electronics outside the cryostat, and due to the analogue multiplexing (by a factor of about 10^6) on the CCD, the number of connections required is small. This allows the electronics to be tucked away behind the tungsten mask used in the small angle region to shield the overall detector, providing a virtually radiation-free environment, even though the detector itself may accumulate as much as 1 Mrad during its working life.

The issue therefore really only arises in the case of nonmultiplexed detectors (microstrip and APS detectors particularly) where the electronics has to be connected by wire bonds or bump bonds, and is therefore inevitably in the same high radiation environment as the small-radius detectors. The worst example is LHC, for which the dose levels tabulated in Sec. 6.4.1 apply equally to the electronics. For the Central Trackers (reaching in to R $\approx$ 30 cm but not below), radiation resistance up to around 10 Mrads and 2×10^{14} n / cm^2 equivalent is required. This is achievable with "standard" rad-hard CMOS and bipolar IC processing. The commercial situation is somewhat unstable. Companies that previously worked closely with the defense industry in the USA and Europe are in some cases looking for new markets and are offering their services to ASIC designers in general, including those at HEP laboratories. Some of these companies, however, have decided that the nondefense markets are inadequate and have ceased to offer facilities for rad-hard electronics. As has been mentioned, the trend towards submicron processing lends itself incidentally to improved radiation resistivity, though care has still to be taken over such issues as

field oxide isolation. In general terms, the needs of the central tracker community for radiation-resistant microstrip electronics are well served; if anything, they have a wider choice than might absolutely be necessary.

For the vertex detector region ($R \leq 10$ cm), the situation is far more challenging (> 100 Mrads and $> 5\times10^{15} n / cm^2$ at $R = 4$ cm). Furthermore, hit densities and degradation in the detectors (noise related to leakage current, loss of charge collection efficiency) mandate pixel-based detectors. The precision requirements of a truly general vertex detector would imply precision of a few microns in both views (and hence small pixels). However, this high granularity should not be achieved at the expense of excessive power dissipation, or else the material introduced per layer (including cooling systems) becomes unacceptable. A general aim of not more than 1 W/cm^2 and 1% RL per layer (detector plus electronics) is generally considered reasonable, and the granularity (i.e., the physics capability) is adjusted to suit. This seems to me to be a very reasonable strategy; it has stimulated a huge and diverse effort, and as the technology advances, the physics requirements will become better met. The high particle fluxes at LHC (small radius) mandate a complex circuit for each pixel, and the requirement of radiation hardness, of course, increases the area of that circuit. This is a development area in which it will be necessary to take advantage of the latest developments into and beyond the time of LHC startup ten years from now. Fortunately, vertex detectors are compact and inexpensive in relation to their value for physics, and so can be rebuilt and upgraded pretty much in response to the technological advances.

6.5 Future Prospects

The radiation levels in space-based systems and accelerator environments such as LHC are generating new challenges. Those faced by the vertex detectors at hadron colliders are by far the most difficult. Detectors will necessarily be pixel-based, and the low-and-slow CCD pixel technology must be replaced by APS devices with as-yet unattainable performance. There is a temptation to abandon silicon as being inadequate for these radiation levels, both for the detectors and for the electronics. Yet it is clear that the essential limits to the radiation hardness of silicon, particularly as regards displacement damage in detector-grade material, are far from understood. The role of defects such as carbon and oxygen is only now beginning to be assessed. It therefore seems entirely appropriate to push hard on these developments, and the field of *defect engineering* is being applied to very good effect in elucidating this subject. If sufficient progress is made in radiation hardening, all the other attributes of silicon will give it a tremendous advantage over rival technologies. On the other hand, to have complete confidence that these enormous problems will be solved would be equally naive. It is therefore very important that some groups put their efforts into exploring alternatives, as discussed in the next section.

7 Beyond Silicon

Driven by the fierce radiation levels in future vertex detector environments (notably at LHC), it is natural to ask whether other detector media or IC technologies might be better suited to the task. Given the high probability that the pixel-based detectors to be used in these environments will necessarily be hybrid (as opposed to monolithic), it is even possible that the detector and electronics IC, bump-bonded together, may be made of different materials, either or both of which may be nonsilicon. There is a great deal of R&D under way in a number of technologies; space constraints permit only a glimpse at these in this paper.

7.1 Gallium Arsenide Detectors

Gallium arsenide has long been of interest for high-speed electronics and sensors, due to its high electron mobility (Fig. 16). In addition, the excellent radiation resistance of some heterojunction electronics devices based on gallium arsenide (see Sec. 7.3) has prompted research into its possible use as a detector medium in high-radiation environments. The essential concerns to date have been the lack of technological maturity by comparison with silicon devices, and the slow progress in overcoming these difficult problems.

The most basic material characteristics (high density, high Z, and high fragility), while advantageous for some applications such as X-ray detectors, are all going in the wrong direction for high-precision MIP tracking detectors, particularly vertex detectors. Nevertheless, the potential for high radiation tolerance is a major attraction.

The difficulties begin with the production of detector-grade material. The impressive work going on in this very complex area has been summarized in two excellent recent papers [93, 94]. Three methods of crystal growth and three methods of epitaxial layer deposition have been tried; of these, only one (liquid encapsulation, LEC) has yielded detector-grade material. Even here, resistivities are at present limited to around 100 Ω cm and electron lifetimes to around 10 ns.

The idea of using GaAs for high-speed (GHz) CCD's has great attractions [95], and considerable progress with test devices has been made. This work illustrates the need to extend basic designs with respect to those used with silicon. "Standard" capacitive gates imply processing complications that can be overcome by a resistive gate technology. This, however, gives large leakage current, which can in turn be overcome with a heterostructure design. The overall picture is one of considerable problems but enormous promise.

The use of pixellated GaAs detectors for hard X-rays, bump-bonded to silicon readout IC's, is being pursued by the Leicester University X-ray astronomy group [96, 97].

For tracking detectors in high-radiation environments, possibly including the most challenging vertex detector region, the RD-8 Collaboration at CERN is doing pioneering work [98]. MIP signals are not yet adequate for high-efficiency trackers, but progress in the quality of the starting crystals should improve that. For the present, compensated material (using iron or chromium doping) is used to achieve acceptable depletion depths. Reasonable resistance to neutron irradiation has been observed, but there are recent concerns (unpublished) as to the hardness with respect to protons. There is also the concern that as the carrier lifetime is increased as a result of improved crystal quality, the radiation tolerance may be correspondingly degraded. There is (to my knowledge) nothing to suggest that "detector-grade" GaAs (comparable in its properties to detector-grade silicon) would necessarily be more radiation resistant than silicon. All studies to date relate to material which can only be compared to silicon of resistivity around 100 Ω cm at best, with leakage currents approximately 1000 times greater than those of high grade silicon.

7.2 CVD Diamond

The availability of affordable diamonds grown by the chemical vapor deposition (CVD) process has opened up an exciting possibility for extremely radiation resistant tracking detectors, well-suited to the LHC vertexing environment. A comparison of some of the important parameters with respect to silicon and gallium arsenide is as follows:

Property	Silicon	GaAs	Diamond
Mass density g. cm^{-3}	2.33	5.32	3.5
Radiation length cm	9.4	2.3	12.0
Average e-h pairs per 100 μm	8900	13000	3600
Average e-h pairs per 0.1% RL	8400	3000	4500

Being in a class of its own as regards band gap for detector materials (see Fig. 61), there is no need to create a diode structure. Simply metallizing the insulator surfaces and applying a potential difference results in collection of the generated signal (up to the limit of the electron lifetime in the material) with negligible leakage current. A review paper of the CERN RD42 Collaboration on this subject reports excellent recent progress [99]. The method of crystal growth results in a defect density which diminishes as the thickness is increased (see Fig. 62). Carrier lifetimes have recently increased to the point that collection distances of 100 μm (adequate for an efficient MIP detector) have been achieved (Fig. 63). These properties have been stable with irradiation up to pion fluences of $6\times10^{13}\,cm^{-2}$. Of course, for the most challenging vertexing applications, they still need to be checked up to $10^{15}\,cm^{-2}$. Leakage currents are not a problem at any radiation dose.

This technology does appear to offer real hope for a reasonably low-mass detector sitting at the minimal radius ($\approx$ 4 cm) in an LHC experiment for a ten-year lifetime. Due to the track density, it would certainly need to be pixel-based (or very short strips!) so presumably, one is contemplating bump-bonding to appropriately robust electronics. This is the topic of the next section.

7.3 Local Electronics

For the high-radiation vertex detector environments where silicon-based IC's are (probably) ruled out, we are almost certainly in the world of pixels. The basic requirements for the front-end IC's include fast shaping times, low noise at low power, and excellent radiation hardness. The high electron mobility transistor

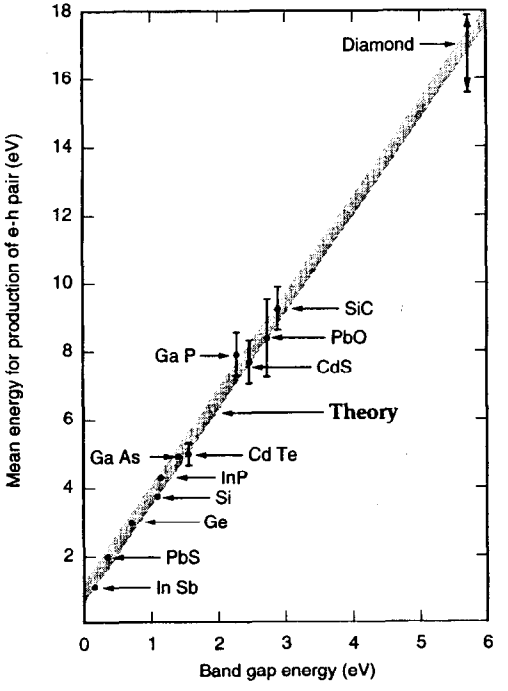


Fig. 61. Band gap and pair-creation energy, for various detector materials.

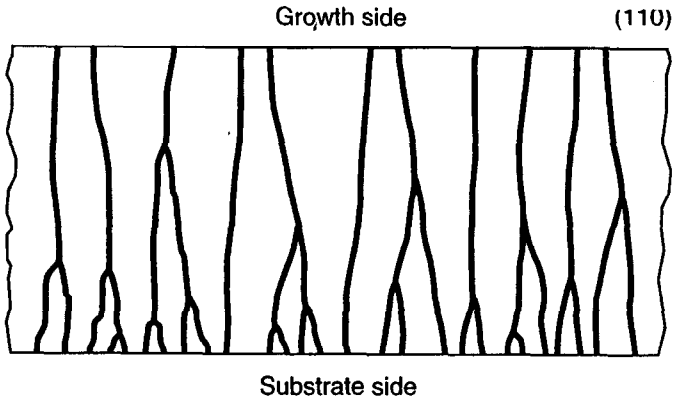


Fig. 62. Evolution of crystalline defects in CVD diamond as a function of thickness of the deposited layer.

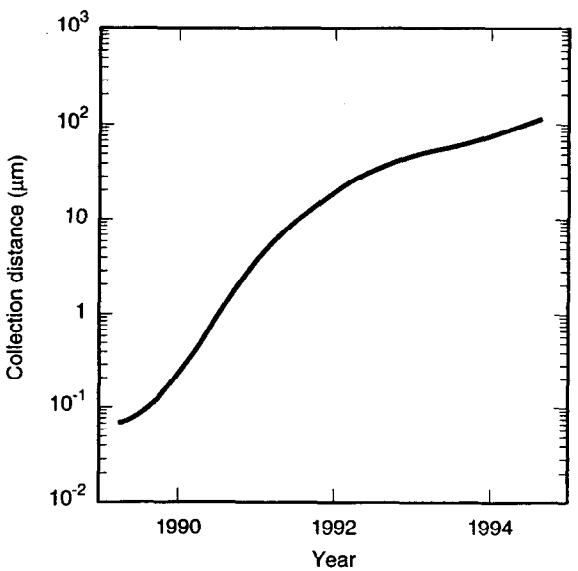


Fig. 63. Time development of collection distance in CVD diamond, from Ref. [99].

(HEMT) based on heterojunctions between different III-IV compounds offers considerable hope of satisfying these requirements. For a recent review paper, see Ref. [100]. The extraordinary radiation hardness of these devices, and indeed their availability as highly engineered structures, stems from the fact that electrons are transported in extremely thin layers (e.g., 10 nm thickness in the typical GaAs/AlGaAs heterostructure). Bulk damage effects are much less severe in such regions of high current density. The gain of both *n*- and *p*-type C-HFET's is stable after irradiation by 100 Mrad gammas and 10^{15} n / cm^2 (Ref. [101]), and these structures readily lend themselves to integrated electronics design (amplifiers, comparators, etc.) as required for APS readout electronics. The prospect of CVD diamond detectors bump-bonded to such readout IC's looks extremely promising. One is, however, still a long way short of demonstrating the LHC functionality at a reasonable pixel size and power dissipation. But there are no seemingly insurmountable obstacles in view.

8 Conclusions

Vertex detectors used in experiments up to the present time come in essentially two varieties, those providing one-dimensional information (microstrip detectors) and those being pixel-based and providing two-dimensional information (charge-coupled devices). The latter, though preferable in principle for several reasons, including superior track reconstruction capability, have restricted applicability in the HEP environment. Both of these detector types found their birthplace in the ACCMOR Collaboration in the early '80s, where they performed with unprecedented precision for charm reconstruction in a fixed-target experiment.

In the move to the collider environment, we experienced, in one sense, a step backward. Due to large beam pipes dictated by background levels at small radius, lower track momenta, and other factors, the enormous effort has been repaid by high-quality b tagging, but only limited charm capability. Fortunately for us, the physics rewards for these restricted technical achievements have been considerable, crowned recently by the discovery of top. The strength of the CDF analysis gained enormously from the b-tagging capability in that experiment.

For the future (B factories, LHC, and the e^+e^- linear collider, among others), the challenges will be still greater. Backgrounds and track densities in the event will in general increase at small radius, due mainly to the higher CM energies giving greater track multiplicities and to the increased luminosity needed to achieve the physics goals. Silicon microstrips, while of increasing value for general tracking, will tend to be pushed out of the small radius region where conditions are too hostile. Regarding the energy frontier (LHC and the future e^+e^- LC), we can expect to see a separation between the vertexing technologies.

For the LHC, one is looking for pixel-based detectors with high timing resolution and phenomenal radiation resistance. This probably leads to the realm beyond silicon, most probably hybrid detectors using GaAs or (more probably) CVD diamond, and hardened silicon or (more probably) heterojunction IC's for the front-end electronics. Some flexibility is gained by the general acceptance of the fact (demonstrated ten years ago in CCD detector systems) that the operating temperature should be considered a tunable parameter. By appropriate mechanical design, it is possible to make very low mass structures of micron-scale mechanical stability that can be repeatedly cycled between room temperature and the optimal cryogenic operating temperature. What is most important, as the overall LHC detectors enter their construction phase, is to preserve adequate funding for the R&D needed to surmount the great challenges associated with vertex detectors in that environment. R&D tends to be squeezed out under pressure of large construction projects, and it is important to remember that the LHC vertex detectors are on a significantly longer timescale than the rest of the system. The optimal detector designs may well continue to evolve through the physics life of the machine, leading to upgrade detectors on several occasions.

For the future e^+e^- Linear Collider, the picture seems to be rather clearer. The main challenge in sitting at small radius is to absorb a very high rate of background MIP hits from incoherent e^+e^- pair background. CCD detectors of unparalleled granularity have this capability, the 307 Mpixel SLD upgrade detector being a useful demonstration model. Ongoing CCD developments hold the promise of vertex detectors for this environment able to operate at R $\approx$ 10 mm with space-point precision of approximately 3 μm, and thickness less than 0.2% RL per layer. This combination is unachievable with any APS system conceived to date (thickness of 1% RL per layer is a reasonable goal for such detectors) and the poorer timing information from the CCD detector is not a serious drawback in this environment, given the long beam-crossing interval ($\leq$ 120 Hz bunch crossing frequency).

The *physics* requirements of these detectors operating at the energy frontier are, of course, difficult to define. Hopes of Higgs and SUSY particle decays via bottom provide a clear motivation. However, it is not impossible that even more exciting (i.e., unexpected) discoveries may result from clean recognition of charm jets, or indeed from clean operation in veto mode, recognizing jets which are devoid of heavy flavors. My personal inclination is to be wary of theoretical predictions and to aim to build a general purpose detector which is as powerful as possible within its measurement regime. For vertex detectors, this means aiming to see the full tree of sequential bottom and charm decays with high efficiency. History has taught us the danger of linking experiments too closely to theoretical ideas. One remembers

experiments at the CERN ISR where an intensive effort to discover the W boson was mounted. This search was, of course, doomed due to the machine energy being well below the W production threshold, but one could easily have discovered the J/ψ which was being prolifically produced in that environment, had experiments been provided with a modest two-muon detection system, rather than a highly sophisticated system geared up to single-muon detection. Such lessons have taught us that future detectors should be made as general as possible in their scope for physics discoveries. In the case of vertex detectors, achieving a good capability for identifying the heaviest long-lived quark of charge +2/3 (charm) as well as the heaviest quark of charge -1/3 (bottom) may pay unpredictable dividends for physics. In this regard, the present generation of collider vertex detectors, if given school grades, would attract comments such as "could do better," "a greater effort is needed in future," etc.

It is perhaps instructive to summarize the time development of the various types of vertex and tracking detectors with respect to some key parameters. Figure 64 shows the area coverage. Microstrip detectors have always been far ahead and seem well-placed to continue their prodigious expansion (to some tens of square meters) at LHC. CCD-based detectors may have peaked in area with the SLD upgrade. For the future LC, the smaller beam pipe leads to no greater an area coverage requirement than has currently been achieved. In this respect, smaller is better. APS systems, not yet used as vertex detectors, need to expand greatly for LHC, but the *performance* increase is a much greater challenge for them, as we have seen.

Figure 65 shows the corresponding situation as regards number of channels. At 300 Mpixels, the SLD vertex detector may have reached some sort of plateau, but the APS system for LHC will need to get close to this in order to meet the initial design specifications. This is an enormous extrapolation from where they are now.

Figure 66 shows a most important parameter, the multiple scattering term in the impact parameter resolution. Microstrip detectors have floated around the 30–60 μm region; however, this will become less significant as their role (at the energy frontier) evolves from vertex detector to general purpose tracker. The APS detector that will fill the vertexing hole at LHC aims for precision at the high end of this

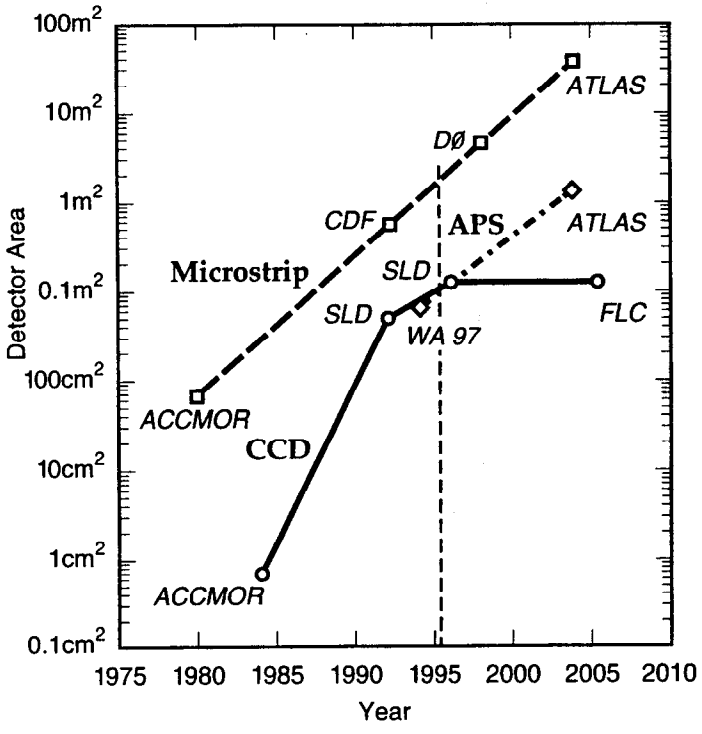


Fig. 64. Time development of area coverage of the leading-edge vertex and tracking detectors according to the main technologies (microstrips and CCD's). The APS is expected to enter the realm of vertex detectors in the LHC environment.

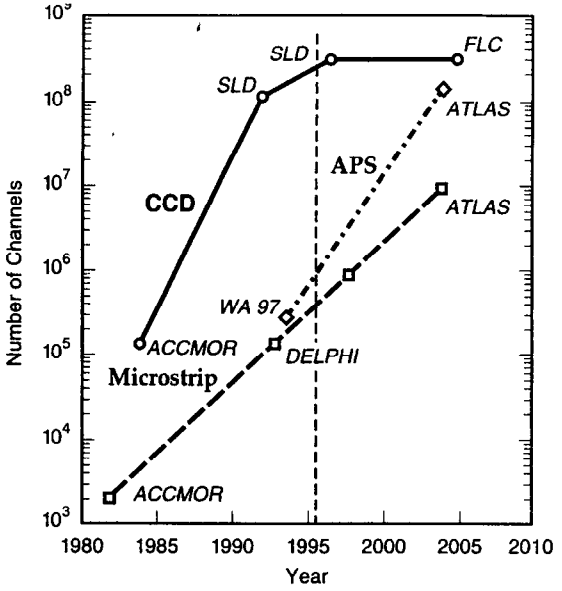


Fig. 65. As Fig. 64, but showing the number of channels in leading-edge detectors as a function of time.

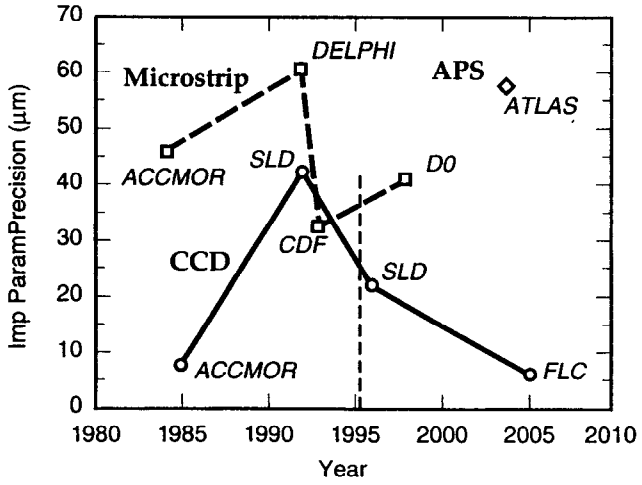


Fig. 66. As Fig. 64, but showing the multiple scattering term in the impact parameter precision as a function of time.

range, due to the fact that nobody yet dreams of going below a radius of 4 cm, and the detectors are intrinsically rather thick. CCD detectors started with marvellous performance in ACCMOR (resulting in some very clean charm physics), degraded badly in the collider environment, are gaining ground with the SLD upgrade detector, and hold promise of their original phenomenal performance (20 years later) at the future LC. The physics rewards on this second round of topologically excellent vertexing could (we hope) be enormous. Incidentally, the ongoing importance of this parameter stems from the increasing particle multiplicities in the events of interest. Despite the increased CM energies, the impact parameter precision for tracks in the 1 to 10 GeV range remains crucial for topological vertexing in the TeV collider regime.

Aside from their applications in particle physics, it is important to remember the very strong interdisciplinary aspects of these detectors. Their use in X-ray detection systems in pure and applied science is enormous, particularly for the pixel-based devices, since the ability to record an image is of rather general interest. Even if the highest aims for vertex detectors are slow in coming (sometimes because of the timescales of the new accelerators), the R&D is proving of great benefit to other areas.

Regarding the specific application to vertex detectors, there is an ongoing need for new ideas. Mostly these will come from young people. I would like to conclude these lectures with a special note of encouragement to these participants. If you get an idea, do not be put off by "the experts." I once attracted a considerable amount of negative expert comment (when I started to push CCD's for vertex detectors in 1980). The established community of experts on silicon radiation detectors was generally extremely skeptical. There were a few exceptions, such as Veljko Radeka and Emilio Gatti, who gave me greatly needed encouragement to carry on. So, if you get an idea, I advise you to pursue it and see where it leads without being too concerned as to the comments of critical bystanders. There is an ancient Chinese proverb that the one who thinks something to be impossible should not interrupt the one who is trying to do it. It would be better for science if some of us middle-aged physicists did more to remember this! I am sure there are wonderful ideas for novel

vertex detectors that nobody has yet thought of, and that some of the participants in this Institute may well discover them.

Acknowledgments

Having worked for 20 years in this field, I am conscious of the vast number of colleagues and collaborators from whom I have learned almost everything I know. In order to make a manageable list, I shall restrict my acknowledgments to those who have contributed most to my ability to prepare these lectures. These include David Burt, David Dorfan, Erik Heijne, Andrew Holland, Gerhard Lutz, Sherwood Parker, Peter Pool, Veljko Radeka, Hartmut Sadrozinski, Paul Seller, Steve Watts, and Peter Weilhammer. I would also like to thank Jacqueline Croft for the careful preparation of this report in record time, John Proch for excellent help with the figures, and David Sankey for digging me out of several PostScript-related traps. Finally, I would like to thank my wife Joan for her great support, and patience regarding the many lost weekends that went into the preparation of the lectures and this written report.

References

1. S. E. Derenzo *et al.*, LBL-1791 (1973).
2. G. Gilder, *Microcosm: The Quantum Revolution in Economics and Technology* (Touchstone, 1989).
3. G. W. Fraser *et al.*, Nucl. Instrum. Methods A **350**, 368 (1994).
4. P. Lechner and L. Strüder, Nucl. Instrum. Methods A **354**, 464 (1995).
5. H. Bichsel, Rev. Mod. Physics **60**, 663 (1988).
6. S. M. Sze, *Physics of Semiconductor Devices* (John Wiley, 1981).
7. W. Shockley, *Electrons and Holes in Semiconductors* (Van Nostrand, 1950).
8. G. K. McKay, Phys. Rev. **84**, 829 (1951).
9. J. B. A. England *et al.*, Nucl. Instrum. Methods **185**, 43 (1981).
10. J. Kemmer, Nucl. Instrum. Methods **169**, 499 (1980).
11. R. Hofmann *et al.*, Nucl. Instrum. Methods **225**, 601 (1984).
12. J. T. Walker *et al.*, Nucl. Instrum. Methods **226**, 200 (1984).
13. P. Holl *et al.*, IEEE Trans. Nucl. Sci., **NS-36**, 251 (1988).
14. P. Allport *et al.*, Nucl. Instrum. Methods A **310**, 155 (1991).
15. N. Bachetta *et al.*, Nucl. Instrum. Methods A **342**, 39 (1994).
16. T. I. Westgaard *et al.*, paper contributed to the 7th European Symposium on Semiconductor Detectors, 1995 (to be published).
17. B. S. Avset *et al.*, IEEE Trans. Nucl. Sci., **NS-37**, 1153 (1990).
18. T. Ohsugi *et al.*, Nucl. Instrum. Methods A **342**, 22 (1994).
19. *Hot electron analyzer, PHEMOS-50*, Hamamatsu Photonics.
20. G. Lutz, private communication.
21. ATLAS and CMS Silicon Central Trackers, technical proposals CERN/LHCC/94-43 (ATLAS) and CERN/LHCC/94-38 (CMS).
22. G. Batignani *et al.*, Nucl. Instrum. Methods A **236**, 183 (1993).
23. P. Weilhammer, Nucl. Instrum. Methods A **342**, 1 (1994).
24. R. Brenner, *Proc. 3rd International Workshop on Vertex Detectors*, Indiana University Report IUHEE-95-1 (1995).
25. E. Gross, *Proc. 3rd International Workshop on Vertex Detectors*, Indiana University Report IUHEE-95-1 (1995).
26. E. Belau *et al.*, Nucl. Instrum. Methods **214**, 253 (1983).
27. B. Hyams *et al.*, Nucl. Instrum. Methods **205**, 99 (1983).
28. L. Hubbeling *et al.*, Nucl. Instrum. Methods A **310**, 197 (1991).
29. H. Tajima *et al.*, Nucl. Instrum. Methods A **310**, 504 (1991).
30. H. Hanai *et al.*, Nucl. Instrum. Methods A **314**, 455 (1992).
31. A. S. Schwarz, Physics Reports **238**, 1 (1994).
32. P. Seller *et al.*, IEEE Trans. Nucl. Sci., **NS-35**, 176 (1988).

33. S. A. Kleinfelder *et al.*, IEEE Trans. Nucl. Sci., **NS-35**, 171 (1989).

34. E. Beauville *et al.*, Nucl. Instrum. Methods A **288**, 157 (1990).

35. D. E. Dorfan, Nucl. Instrum. Methods A **342**, 143 (1994).

36. Z. Y. Chang and W. Sansen, Nucl. Instrum. Methods A **305**, 553 (1991).

37. S. Gadomski *et al.*, Nucl. Instrum. Methods A **320**, 217 (1992).

38. C. Da Via *et al.*, Nucl. Instrum. Methods A **344**, 199 (1994).

39. E. Barberis *et al.*, UC Santa Cruz Preprint, SCIPP 94/29 (1994).

40. J. DeWitt *et al.*, UC Santa Cruz Preprint, SCIPP 94/34 (1994).

41. F. Arfelli *et al.*, paper contributed to the 7th European Symposium on Semiconductor Detectors, 1995 (to be published).

42. W. S. Boyle and G. E. Smith, Bell Syst. Tech. Journal **49**, 587 (1970).

43. R. H. Walden *et al.*, Bell Syst. Tech. Journal **51**, 1635 (1972).

44. C. J. S. Damerell *et al.*, Nucl. Instrum. Methods **185**, 33 (1981).

45. *Charge-Coupled Devices and Systems*, edited by M. J. Howes and D. V. Morgan (Wiley, 1979).

46. J. D. E. Beynon and D. R. Lamb, *Charge-Coupled Devices and Their Applications* (McGraw-Hill, 1982).

47. L. Strüder *et al.*, Nucl. Instrum. Methods A **253**, 386 (1987).

48. M. Kleefstra, *Solid State Electronics* **21**, 1005 (1978).

49. A. D. Holland, paper contributed to the 7th European Symposium on Semiconductor Detectors, 1995 (to be published).

50. H. Soltau, paper contributed to the 7th European Symposium on Semiconductor Detectors, 1995 (to be published).

51. C. J. S. Damerell, summary paper contributed to the 7th European Symposium on Semiconductor Detectors, 1995 (to be published).

52. E. R. Fossum, in *Proceedings of the SPIE International Conference* **2172**, 38 (1994).

53. A. M. Mohsen and M. F. Tompsett, IEEE Trans. Electronic Devices, **ED-21**, 701 (1974).

54. E .K. Bangart *et al.*, IEEE Trans. Electronic Devices, **ED-38**, 1162 (1991).

55. J. Frenkel, Phys. Rev. **54**, 647 (1938).

56. D. J. Burt, Nucl. Instrum. Methods A **305**, 564 (1991).

57. P. Centen, IEEE Trans. Electronic Devices, **ED-38**, 1206 (1991).

58. R. W. Brodersen and S. P. Emmons, IEEE Trans. Electronic Devices, **ED-23**, 215 (1976).

59. D. Burt, private communication.

60. M. H. White *et al.*, IEEE Journal of Solid State Circuits, **SC-9**, 1 (1974).

61. G. R. Hopkinson and D. H. Lumb, J. Phys. E. **15**, 1214 (1982).

62. K. H. Schmidt *et al.*, MPI-PhE/94-30 (1994).

63. SLD Vertex Detector Upgrade Group, SLAC-PUB-95-6950 (1995).

64. C. J. S. Damerell and D. Jackson, paper submitted to the Morioka Workshop on Physics at Future Linear e^+e^- Colliders, 1995 (to be published).

65. E. R. Fossum, transparencies of lecture given at the Univ. of Waterloo Pixel Device Conference, 1993.

66. S. K. Mendis *et al.*, *Proc. SPIE* **2172**, 19 (1994).

67. F. Antinori *et al.*, CERN DRDC/94-51 (1995).

68. S. Holland, Nucl. Instrum. Methods A **275**, 537 (1989).

69. C. J. Kenney *et al.*, Nucl. Instrum. Methods A **342**, 59 (1994).

70. S. Parker, private communication.

71 P. Middelkamp, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

72. G. Cesura *et al.*, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

73. V. A. J. van Lint *et al.*, *Mechanics of Radiation Effects in Electronic Materials* (Wiley, 1980).

74. G. C. Messenger and M. S. Ash, *The Effects of Radiation on Electronic Systems* (Van Nostrand Reinhold, 1986).

75. W. R. Dawes, Nucl. Instrum. Methods A **288**, 54 (1990).

76. V. A. J. van Lint, IEEE Trans. Nucl. Sci., **NS-41**, 2642 (1994).

77. E. E. Conrad, IEEE Trans. Nucl. Sci., **NS-41**, 2648 (1994).

78. J. R. Adams *et al.*, IEEE Trans. Nucl. Sci., **NS-24**, 2099 (1977).

79. W. Füssel, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

80. E. A. Burke, IEEE Trans. Nucl. Sci., **NS-33**, 1276 (1986).

81. G. P. Summers *et al.*, IEEE Trans. Nucl. Sci., **NS-33**, 1282 (1986).

82. H. J. Ziock *et al.*, Nucl. Instrum. Methods A **342**, 96 (1994).

83. E. Fretwurst *et al.*, Nucl. Instrum. Methods A **342**, 119 (1994).

84. F. Lemeilleur *et al.*, CERN-ECP/94-8 (1994).

85. Three papers contributed by G. Lutz, Z. Li, and S. Watts to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

86. J. Matheson *et al.*, CERN report RD20/TN/95/43.

87. M. Huhtinen and P. A. Aarnio, Nucl. Instrum. Methods **A 335**, 580 (1993).

88. J. M. Killiany, "Topics in applied physics," **38**, 147 (1980).

89. N. S. Saks, IEEE Trans. Nucl. Sci., **NS-24**, 2153 (1977).

90. M. S. Robbins, Ph.D. Thesis, Brunel University (1992) and RADECS 91, *IEEE Proceedings*, 368 (1992).

91. A. Holland *et al.*, IEEE Trans. Nucl. Sci., **NS-38**, 1663 (1991).

92. S. Watts *et al.*, ESTEC Report BRUCRD-ESA CCD-95-IR (1995).

93. M Schieber, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

94. E. Bauser, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

95. R. E. Colbeth *et al.*, in *Proceedings of the SPIE International Conference* **1071**, 108 (1989).

96. A. D. Holland *et al.*, Nucl. Instrum. Methods A **346**, 366 (1994).

97. A. D. Holland *et al.*, Leicester University preprint XRA 94/15 (1994).

98. RD-8 Collaboration Report, CERN/DRDC 94-32 (1994).

99. K. T. Knöpfle, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

100. G. Bertuccio *et al.*, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

101. W. Karpinski, paper contributed to the Seventh European Symposium on Semiconductor Detectors, 1995 (to be published).

THE ROLE OF TOP IN HEAVY FLAVOR PHYSICS

JoAnne L. Hewett*
Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

ABSTRACT

The implications of the massive top quark on heavy flavor transitions are explored. We review the generation of quark masses and mixings and the determination techniques, and present the status of the elements of the weak mixing matrix. Purely leptonic decays of heavy mesons are briefly summarized. We present a general introduction to flavor changing neutral currents and an extensive summary of radiative and other rare decay modes. The physics of neutral meson mixing is reviewed and applied to each meson system. We describe the phenomenology of CP violation and how it may be measured in meson decays. Standard Model predictions are given in each case and the effects of physics beyond the Standard Model are also discussed. Throughout, we contrast these transitions in the K and B meson systems to those in the D meson and top-quark sectors.

*Work supported by the Department of Energy, Contract DE-AC03-76SF00515.

©1995 by JoAnne L. Hewett.

1 Introduction

One of the outstanding problems in particle physics is the origin of the fermion mass and mixing spectrum. Despite the success of the Standard Model (SM) of particle physics, it does not provide any clues on the source of these parameters. In contrast to the case of electroweak symmetry breaking, we have no information about the relevant energy scale where these parameters originate; in fact, the relevant scale could lie anywhere from 1 TeV to the Planck scale. Other issues (besides quark mixing) related to the multifamilies of fermions are the suppression of FCNC effects and the CP-violation phases in fermion gauge couplings. Since the top quark has a mass at the electroweak symmetry breaking scale, it is believed that it may reveal some hints to these questions. In these lectures, we examine its role in heavy flavor transitions.

At present, the best approach in addressing these questions is to study the properties of all heavy fermions in detail. Detailed measurements of heavy quark systems are best realized at high precision, high luminosity machines, and several dedicated heavy flavor factories and experiments will be coming on line in the next decade. We will learn a wealth of new and precise information which will hopefully result in the development of a theory to explain the existence of families.

In these lectures, we review the generation of quark masses and mixings and the determination techniques, and present the status of the elements of the weak mixing matrix. Purely leptonic decays of heavy mesons are briefly summarized. We present a general introduction to flavor changing neutral currents and an extensive summary of radiative and other rare decay modes. The physics of neutral meson mixing is reviewed and applied to each meson system. We describe the phenomenology of CP violation and how it may be measured in meson decays. Standard Model predictions are given in each case and the effects of physics beyond the Standard Model are also discussed. Throughout, we contrast these transitions in the K and B meson systems to those in the D meson and top-quark sectors.

2 Quark Masses and Mixing

In the Standard Model (SM), a single complex scalar doublet is responsible for the spontaneous symmetry breaking $SU(2)_L \times U(1)_Y \to U(1)_{em}$. The fermions are massless before the symmetry breaking, with their masses being generated via Yukawa couplings after the spontaneous symmetry breaking occurs. Denoting the gauge (or weak or flavor) quark eigenstates as q_L^0 (q_R^0) for the left-handed doublet (right-handed singlet) quark fields, one can form the most general renormalizable quark-Higgs interaction

$$\mathcal{L}_{mass} \sim \frac{v}{\sqrt{2}} \left[\bar{u}^0_{L_i} h^u_{ij} u^0_{R_j} + \bar{d}^0_{L_i} h^d_{ij} d^0_{R_j} \right] + h.c. , \tag{1}$$

with v being the vacuum expectation value of the Higgs field, i, j are generation indices, and $h^{u,d}_{i,j}$ are 3×3 matrices of bare complex couplings which form the mass matrices for the up- and down-quarks

$$M_u = \frac{v}{\sqrt{2}} h^u , \qquad M_d = \frac{v}{\sqrt{2}} h^d . \tag{2}$$

The mass matrices are completely arbitrary and contain 36 unknown parameters! These matrices can be diagonalized by a bi-unitary transformation,

$$\begin{aligned} M_u^{diag} &= U_L^\dagger M_u U_R = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} , \\ M_d^{diag} &= D_L^\dagger M_d D_R = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} , \end{aligned} \tag{3}$$

where we have performed distinct rotations of the left- and right-handed fields, and the m_i represent the quark masses. Hence, six of the above 36 parameters become quark masses. The interaction Lagrangian can now be written as (where the generation indices have now been dropped)

$$\mathcal{L}_{mass} \sim \bar{u}^0_L U_L U_L^\dagger M_u U_R U_R^\dagger u^0_R + \bar{d}^0_L D_L D_L^\dagger M_d D_R D_R^\dagger d^0_R + h.c. , \tag{4}$$

which is just given by

$$\begin{aligned} \mathcal{L}_{mass} &\sim \bar{u}_L M_u^{diag} u_R + \bar{d}_L M_d^{diag} d_R + h.c. , \\ &\sim \sum_{i=1}^{6} m_i \bar{q}_i q_i + h.c. , \end{aligned} \tag{5}$$

in the mass (or physical) eigenstate basis with $u_L = U_L^\dagger u^0_L$, etc., being the physical quark fields. Note that the Higgs-quark Yukawa couplings are manifestly diagonal

in the physical basis; this is a consequence of the fact that there is only one Higgs doublet in the SM. The SM charged current interaction in the weak basis,

$$\mathcal{L}_{cc} \sim \frac{g}{\sqrt{2}} \bar{u}_L^0 \gamma_\mu d_L^0 W^\mu + h.c., \tag{6}$$

then becomes

$$\mathcal{L}_{cc} \sim \frac{g}{\sqrt{2}} \bar{u}_L \gamma_\mu U_L^\dagger D_L d_L W^\mu \tag{7}$$

in terms of the physical fields. The product $U_L^\dagger D_L$ is known as the Cabibbo-Maskawa-Kobayashi (CKM) weak mixing matrix V_{CKM}. Since there are no right-handed charged currents in the SM, there is no analogous right-handed weak mixing matrix. In extensions of the SM which enlarge the gauge group to $SU(2)_R \times SU(2)_L \times U(1)$, such as the left-right symmetric model,[1] the quantity $V_{CKM}^R \equiv U_R^\dagger D_R$ is similarly defined.

The CKM matrix contains information on all quark flavor transitions and is the source of CP violation within the SM. Writing the charged current interaction explicitly in matrix form yields

$$\mathcal{L}_{cc} \sim \frac{g}{\sqrt{2}} [u_L\, c_L\, t_L] \gamma_\mu \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W^\mu + h.c. \tag{8}$$

Note that by construction, the CKM matrix is unitary, *i.e.*, $\sum_i V_{ij} V_{ik}^* = \delta_{jk}$. Unitarity tests thus provide an excellent probe of the SM. In general, any unitary $N \times N$ matrix can be expressed by N^2 parameters, $N(N-1)/2$ of which are rotation angles, and $N(N+1)/2$ phase angles. Here, the phases are associated with the quark fields, and $2N-1$ of them may be arbitrarily redefined, leaving $(N-1)(N-2)/2$ independent phases. For three generations of quarks, this leaves three rotation angles and one independent phase. Extrapolation to four generations would then require six rotation angles and three phases. A common parameterization of the three generation CKM matrix,

$$V_{CKM} = \begin{pmatrix} c_1 & -s_1 c_3 & -s_1 s_3 \\ s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\ s_1 s_2 & c_1 s_2 c_3 + c_2 s_3 e^{i\delta} & c_1 s_2 s_3 - c_2 c_3 e^{i\delta} \end{pmatrix}, \tag{9}$$

is based on three Euler angles for the flavor rotations and was first given by Kobayashi and Maskawa[2] in 1973. Note that this was postulated *before* the discovery of the third generation (as well as charm), in order to introduce a potential source of CP violation. Here, $c_i = \cos\theta_i, s_i = \sin\theta_i$ with $0 \le \theta_i \le \pi/2$ and $-\pi \le \delta \le \pi$. An instructive parameterization, which is based on an expansion of the elements in powers of $V_{us} \equiv \lambda$, is given by Wolfenstein[3]

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta + i\eta\lambda^2/2) \\ -\lambda & 1 - \frac{\lambda^2}{2} - i\eta A^2\lambda^4 & A\lambda^2(1 + i\eta\lambda^2) \\ A\lambda^3(1 - \rho - i\eta) & A\lambda^2 & 1 \end{pmatrix} \tag{10}$$

to $\mathcal{O}(\lambda^3)$ in the Real terms and $\mathcal{O}(\lambda^5)$ in the Imaginary terms. This parameterization illustrates the approximate diagonal nature of the CKM matrix, and exhibits which elements (and hence measurements thereof) are most sensitive to the various parameters λ, A, ρ, and η.

We now review the status of the experimental determinations of the CKM matrix elements.[4,5] We stress that the values of the CKM elements are fundamental input parameters within the SM and precise knowledge of these parameters may provide some insight into their origin.

• V_{ud}: This element is determined from super-allowed $0^+ - 0^+$ nuclear β decays. These transitions have large radiative corrections as well as some nuclear Z dependence. Recent analyses[6] of the nuclear structure dependent radiative corrections are inconsistent with each other within the level of the estimated uncertainties. Taking an average value of these results yields the PDG value[4] $|V_{ud}| = 0.9736 \pm 0.0023$, where the error is dominated by the theoretical uncertainties. Neutron β decay is less dependent on these nuclear uncertainties; however, the present determination[5] of $|V_{ud}|$ from this process is larger than the above value by several sigma. Pion β decay, $\pi^+ \to \pi^0 e^+ \nu_e$, would in principle yield the cleanest measurement of V_{ud}, but the branching fraction is of order 10^{-8}, making a precision determination difficult.

• V_{us}: This element is cleanly determined from the K_{e3} decays $K^+ \to \pi^0 e^+ \nu_e$ and $K_L^0 \to \pi^\pm e^\mp \nu_e$, giving $|V_{us}| = 0.2196 \pm 0.0023$ (Ref. 7). Hyperon decays yield[7] a slightly larger value of $|V_{us}| = 0.222 \pm 0.003$, but are plagued with uncertainties from SU(3) breaking effects. The average of these two measurements result in the PDG value[4] $|V_{us}| = 0.2205 \pm 0.0018$.

• V_{ub}: The explanation of CP violation within the SM, *i.e.*, the phase in the CKM matrix, requires a nonvanishing value of V_{ub}. It can be measured at the $\Upsilon(4S)$ by examining the endpoint region of the lepton momentum spectrum in inclusive

$B \to X\ell\nu_\ell$ decays and counting the excess of leptons beyond the kinematic limit for $B \to X_c\ell\nu_\ell$. While data has established that V_{ub} is nonzero, converting the measured rate into a value of V_{ub} introduces substantial errors. This conversion is highly model dependent due to the small phase space available at the endpoint, and to details in the hadronization from the large uncertainties in the calculation of the rates for the resonant modes, and the relative size of the contributions of resonant and nonresonant modes in this region. The subtraction of background from the small data sample injects an additional large source of error. The present experimental error on the ratio $|V_{ub}|/|V_{cb}|$ is comparable to the theoretical uncertainty, yielding[4,8] $|V_{ub}/V_{cb}| = 0.08 \pm 0.02$, and thus new, less model-dependent techniques in extracting this CKM element are necessary.

Exclusive semileptonic decays, $B \to X_u\ell\nu_\ell$ where $X_u = \pi, \rho$, or ω, have recently been observed by CLEO.[9] Interpretation of these results in terms of V_{ub} relies on the evaluation of the shape of the contributing form factors, as well as uncertainties in the size of the contributions from nonresonant decays such as $B \to \pi\pi\ell\nu_\ell$. A fit to the data and averaging over the form factor models yields[9] $|V_{ub}| = (3.3 \pm 0.2^{+0.3}_{-0.4} \pm 0.7) \times 10^{-3}$, where the errors are due to statistics, systematics, and estimated model dependence. Reductions in the theoretical errors can be obtained via direct measurements of the form factor q^2 distributions in $c \to d$ transitions such as $D \to \pi\ell\nu_\ell$.

An alternative method[10] of extracting V_{ub} from semileptonic B decays is to measure the invariant mass spectrum of the final state hadrons *below* the charm hadron threshold, *i.e.*, $m_X < m_D$. More than 90% of the $B \to X_u\ell\nu_\ell$ decays lie within this region, yielding almost an order of magnitude increase in data sample over the endpoint region. The theoretical uncertainties associated with the determination of the total semileptonic spectrum are significantly smaller within this kinematic region, and are less than those associated with exclusive semileptonic decays which rely on form factor calculations. In addition, the $B \to X_u\ell\nu_\ell$ transitions are largely nonresonant and multiple jetlike final states dominate, making the inclusive decay theoretically well-understood throughout this kinematic region.

• V_{cd}: This element is evaluated by examining charm production in neutrino and antineutrino scattering off valence d-quarks and folding in the semileptonic branching fraction of charm weighted by the ratio of D^0/D^+ production in neutrino scattering. Averaging the experimental results and including the scale dependence from the NLO QCD corrections leads to the PDG value[4,11] $|V_{cd}| = 0.224 \pm 0.016$.

• V_{cs}: In principle, this element can also be determined in neutrino induced charm production. Here, the scattering of interest clearly takes place off of strange quarks and the results are quite dependent on the s-quark parton density distributions. The most conservative assumptions about the parton densities set[4] the constraint $|V_{cs}| > 0.59$, which is not very restrictive. A better determination can be obtained from D_{e3} decay, $D \to \bar{K}e^+\nu_e$, although this process is form factor dependent and hence contains theoretical uncertainties. Combining various form factor calculations with the measured decay rate gives the PDG value[4] $|V_{cs}| = 1.01 \pm 0.18$, which is still not very well-determined. Employing the three generation unitarity constraint on the CKM matrix results in the most precise evaluation of this element.

• V_{cb}: Considerable theoretical and experimental progress has been made recently on the extraction of V_{cb} from exclusive and inclusive decays. Exclusive semileptonic decays offer a reliable model-independent determination of V_{cb} within the framework of heavy quark effective theory as the heavy quark symmetry normalizes the q^2-dependent hadronic form factors with good precision at zero recoil for the charm hadron system. This technique is best suited for the process $B \to D^*\ell\nu_\ell$ as the leading corrections to the HQET result arise only at higher order, $1/m_Q^2$.

The inclusive semileptonic branching fraction B_{SL} can be determined from (i) measurement of the inclusive lepton momentum spectrum. This technique yields significant data samples, but the procedure used to fit the observed spectrum to the expected shape for primary and secondary leptons from B and charm decay, respectively, introduces a large model dependence. (ii) Charge and angular correlations in dilepton events. This offers less model dependence as the measured correlations can be used to separate the primary and secondary lepton spectra, instead of relying on theory. (iii) Separate measurement of B_{SL} for charged and neutral B meson decay. Here, one B in the event must be reconstructed in order to tag the charge of the other. Determination of V_{cb} from B_{SL} via technique (i) at CLEO and LEP is already saturated by the theoretical error, while methods (ii) and (iii) still offer room for improvement experimentally.

Combining the results[4,12] on the exclusive and inclusive semileptonic decays gives $|V_{cb}| = 0.040 \pm 0.003$.

• V_{td}, V_{ts}: These elements can be determined from flavor changing neutral current processes which contain one-loop top-quark contributions. The value of $|V_{td}|$ can be deduced from $B_d^0 - \bar{B}_d^0$ mixing and from the decay $K \to \pi\nu\bar{\nu}$, with $K \to \pi\nu\bar{\nu}$ ultimately offering the theoretically cleanest technique. The ratio $|V_{td}/V_{ts}|$ can be found from the ratio of B_d to B_s mass differences, as well as from the ratio of exclusive rates $B(B \to \rho\gamma)/B(B \to K^*\gamma)$, if the long-distance physics can be cleanly separated out. Each of these processes will be thoroughly discussed below; however, we note here they all depend on the assumption that there are no large contributions from new physics. At present, three-generation CKM unitarity constraints offer the best restrictions on these elements within the SM.

• V_{tb}: The b-tagged events observed in top-quark decays at the Tevatron[13] have afforded the first direct measurement of V_{tb}. CDF and DØ measure the ratio of events with zero, one, and two b-tags to extract the ratio $B(t \to Wb)/B(t \to WX)$, which has the advantage of being independent of the $t\bar{t}$ production cross section and the W boson branching fractions. Within the three-generation SM, this procedure yields $|V_{tb}| = 0.97 \pm 0.15 \pm 0.13$. The most precise determination of this element is obtained from employing unitarity together with the direct measurements of V_{ub} and V_{cb}.

Combining the above data with the constraint of three-generation unitarity, results in the 90% C.L. bounds on the full 3 × 3 CKM matrix[4]

$$V_{CKM} = \begin{pmatrix} 0.9745 - 0.9757 & 0.219 - 0.224 & 0.002 - 0.005 \\ 0.218 - 0.224 & 0.9736 - 0.9750 & 0.036 - 0.046 \\ 0.004 - 0.014 & 0.034 - 0.046 & 0.9989 - 0.9993 \end{pmatrix} . \quad (11)$$

These ranges differ slightly from those itemized above due to the inclusion of the unitarity constraint. However, it is important to note that the data does not preclude the existence of more generations.

3 Leptonic Decays

Pseudoscalar mesons can decay to a purely leptonic final state, $P^{\pm}(Q\bar{q}) \to \ell^{\pm}\bar{\nu}_{\ell}$, through the annihilation diagram. The matrix element for this process can be written as (for $m_P^2 \ll M_W^2$)

$$\begin{aligned} \mathcal{M} &= \langle \bar{\nu}_{\ell}\ell^- | \mathcal{L}_{eff} | P^- \rangle , \quad (12) \\ &= \frac{G_F}{\sqrt{2}} V_{Qq} \bar{u}(p_{\ell})\gamma^{\mu}(1-\gamma_5)v(p_{\bar{\nu}_{\ell}})\langle 0|\bar{v}(p_Q)\gamma_{\mu}(1-\gamma_5)u(p_{\bar{q}})|P^-\rangle . \end{aligned}$$

The hadronic matrix element must be of the form

$$\langle 0|\bar{Q}\gamma_{\mu}\gamma_5 q|P^-(p)\rangle = i f_P p_{\mu}^P , \quad (13)$$

since p_{μ}^P is the only four-vector associated with the initial state. The factor f_P is known as the pseudoscalar meson decay constant. Assuming massless neutrinos, the transition rate is then calculated to be

$$\Gamma(P^- \to \ell^-\bar{\nu}_{\ell}) = \frac{G_F^2 |V_{Qq}|^2}{8\pi} f_P^2 m_P m_{\ell}^2 \left(1 - \frac{m_{\ell}^2}{m_P^2}\right)^2 , \quad (14)$$

and is helicity suppressed due to the overall m_{ℓ}^2 factor. In the case of pion decay, the inclusion of the radiative corrections and a comparison with the experimental value for $\pi^- \to \mu^-\bar{\nu}_{\mu} + \mu^-\bar{\nu}_{\mu}\gamma$ yields the well-known value for the pion decay constant of $f_{\pi} = 131$ MeV. Similarly, the kaon decay constant is measured to be $f_K = 160$ MeV with a roughly 1% error due to the uncertainties associated with the value of V_{us}.

The leptonic decays of D and B mesons have not yet been observed (except for the case $D_s^- \to \mu^-\bar{\nu}_{\mu}$), and will be discussed further below. Assuming that the relevant CKM matrix elements for these heavier quark systems are well-known, these decays would provide important information on the value of their associated pseudoscalar decay constants, which in turn are essential for the study of $D^0 - \bar{D}^0$ and $B^0 - \bar{B}^0$ mixing, CP violation in the charm and bottom sector, and in non-leptonic decays.

3.1 Leptonic Decays of Charm Mesons

The SM transition rate for the purely leptonic decay of a pseudoscalar charm meson is given by Eq. (14) with the substitutions $P \to D_q$ and $Q \to c$. The resulting branching fractions are small due to the helicity suppression and are listed in Table 1 using the central values of the CKM parameters given in Ref. 4 and assuming $f_D = 200$ MeV and $f_{D_s} = 230$ MeV. The existing upper limit for f_D is $f_D < 290$ MeV, and is derived from the 90% C.L. bound[14] $B(D^+ \to$

Meson	$\mu^-\bar{\nu}_\mu$	$\tau^-\bar{\nu}_\tau$
D^-	3.52×10^{-4}	9.34×10^{-4}
D_s^-	4.21×10^{-3}	4.11×10^{-2}

Table 1: SM branching fractions for the leptonic decay modes, assuming $f_D = 200$ MeV and $f_{D_s} = 230$ MeV.

$\mu^+\nu_\mu) < 7.2 \times 10^{-4}$ from MARK III. One $D \to \mu\bar{\nu}_\mu$ event has been observed[15] by the BES Collaboration, leading to $f_D = 300^{+180}_{-150}\ ^{+80}_{-40}$ MeV. This is consistent with the MARK III upper bound. Several measurements of f_{D_s} have now been performed, and they are all consistent within the present level of accuracy. CLEO has observed[16] the process $D_s^{*+} \to D_s^+\gamma \to \mu\nu\gamma$ by examining the mass difference $\delta M \equiv M_{\mu\nu\gamma} - M_{\mu\nu}$ and have obtained

$$\frac{\Gamma(D_s^+ \to \mu^+\nu)}{\Gamma(D_s^+ \to \phi\pi^+)} = 0.245 \pm 0.052 \pm 0.074\,. \tag{15}$$

Using $\Gamma(D_s^+ \to \phi\pi^+) = 3.7 \pm 0.9\%$, they find $f_{D_s} = 344 \pm 37 \pm 52 \pm 42$ MeV where the last error reflects the uncertainty in the $\phi\pi^+$ branching fraction. Two emulsion experiments have measured[17] $f_{D_s} = 232 \pm 45 \pm 20 \pm 48$ MeV and $f_{D_s} = 194 \pm 35 \pm 20 \pm 14$ MeV, respectively. And, the BES Collaboration has reported[18] the observation of candidate events in $e^+e^- \to D_s^+D_s^-$ with the subsequent decay $D_s \to \mu\bar{\nu}_\mu$ yielding $f_{D_s} = 430^{+150}_{-130} \pm 40$ MeV. Here, the errors are expected to improve once more statistics are obtained. The current world average[19] value for the branching fraction is $B(D_s^- \to \mu^-\bar{\nu}_\mu) = (4.6 \pm 0.8 \pm 1.2) \times 10^{-3}$, corresponding to $f_{D_s} = (241 \pm 21 \pm 30)$ MeV.

L3 has recently reported[20] the observation of $D_s^- \to \tau^-\bar{\nu}_\tau$ with a branching fraction of $(7.4 \pm 2.8 \pm 1.6 \pm 1.8)\%$, allowing the determination $f_{D_s} = 309 \pm 58 \pm 33 \pm 38$ MeV. Folding this determination with the world average f_{D_s} obtained from the $\mu\bar{\nu}_\mu$ channel gives[19] $f_{D_s} = (255 \pm 20 \pm 31)$ MeV.

A variety of theoretical techniques have been employed to estimate the value of f_D and f_{D_s}. Lattice QCD studies[21] calculate these quantities in the quenched approximation through a procedure that interpolates between the Wilson fermion scheme and the static approximation. The nonrelativistic quark model is used to relate the decay constant to the meson wave function at the origin, $f_M = \sqrt{12/M_M}\,|\psi(0)|$, which is then inferred from isospin mass splitting of heavy mesons.[22] Other approaches employ the relativistic quark model[23] or QCD sum rules.[24] For

Decay Constant	Lattice	Quark Model Nonrelativistic	Quark Model Relativistic	Sum Rule
f_D	205 ± 15	207 ± 60	240 ± 20	170 – 235
f_{D_s}	235 ± 15	259 ± 74	290 ± 20	204 – 270
f_{D_s}/f_D	1.15 ± 0.05	1.25		1.21 ± 0.06

Table 2: Theoretical estimates of the weak decay constants in units of MeV (taking $m_c = 1.3$ GeV in the sum rule approach).

each of these calculational methods, the resulting ranges for the values of the pseudoscalar decay constants are presented in Table 2. A more complete collection of results is given in Ref. 23. Given the large errors, it is clear that these approaches are consistent. We also see that the theoretical predictions tend to be lower on average than the present experimental determinations. Once the experimental precision improves, discrimination between the various theoretical models should be possible, allowing for a better extrapolation to the B system. The theoretical uncertainties associated with the ratio f_{D_s}/f_D are much smaller, as this ratio should deviate from unity only in the presence of broken SU(3) flavor symmetry.

Non-SM contributions may affect the purely leptonic decays. Signatures for new physics include the measurement of non-SM values for the absolute branching ratios, or a deviation from the SM prediction

$$\frac{B(D_{(s)}^- \to \mu^-\bar{\nu}_\mu)}{B(D_{(s)}^- \to \tau^-\bar{\nu}_\tau)} = \frac{m_\mu^2\left(1 - m_\mu^2/m_{D_{(s)}}^2\right)^2}{m_\tau^2\left(1 - m_\tau^2/m_{D_{(s)}}^2\right)^2}\,. \tag{16}$$

This ratio is sensitive to violations of $\mu - \tau$ universality.

As an example, we consider the case where the SM Higgs sector is enlarged by an additional Higgs doublet. As we will see below, these models generate important contributions[25] to the decay $B^- \to \tau^-\bar{\nu}_\tau$, and it is instructive to examine their effects in the charm sector. Two such models, which naturally avoid tree-level flavor changing neutral currents, are Model I, where one doublet (ϕ_2) generates masses for all fermions and the second (ϕ_1) decouples from the fermion sector, and Model II, where ϕ_2 gives mass to the up-type quarks, while the down-type quarks and charged leptons receive their mass from ϕ_1. Each doublet receives a vacuum expectation value v_i, subject to the constraint that $v_1^2 + v_2^2 = v_{\rm SM}^2$. The charged Higgs boson present in these models will mediate the leptonic decay through an

effective four-Fermi interaction, similar to that of the SM W boson. The $H^\pm$ interactions with the fermion sector are governed by the Lagrangian

$$\mathcal{L} = \frac{g}{2\sqrt{2}M_W} H^\pm [V_{ij} m_{u_i} A_u \bar{u}_i (1-\gamma_5) d_j + V_{ij} m_{d_j} A_d \bar{u}_i (1+\gamma_5) d_j \\ m_\ell A_\ell \bar{\nu}_\ell (1+\gamma_5) \ell] + h.c., \quad (17)$$

with $A_u = \cot\beta$ in both models and $A_d = A_\ell = -\cot\beta(\tan\beta)$ in Model I(II), where $\tan\beta \equiv v_2/v_1$. In Models I and II, we obtain the result

$$B(D^- \to \ell^- \bar{\nu}_\ell) = B_{SM} \left(1 + \frac{m_D^2}{m_{H^\pm}^2}\right)^2, \quad (18)$$

where B_{SM} is the SM value of the leptonic branching fraction. In Model II, the D_s^- decay receives an additional modification

$$B(D_s^- \to \ell^- \bar{\nu}_\ell) = B_{SM} \left[1 + \frac{m_{D_s}^2}{m_{H^\pm}^2} \left(1 - \tan^2\beta \frac{m_s}{m_c}\right)\right]^2. \quad (19)$$

We see that the effect of the $H^\pm$ exchange is independent of the leptonic final state and the above prediction for the ratio in Eq. 16 is unchanged. This is because the $H^\pm$ contribution is proportional to the charged lepton mass, which is then a common factor with the SM helicity suppressed term. However, the absolute branching fractions can be modified; this effect is negligible in the decay $D^- \to \ell^- \bar{\nu}_\ell$, but could be of an order of a few percent in D_s^- decay if $\tan\beta$ is very large.

3.2 Leptonic Decays of B Mesons

The SM transition rate for the purely leptonic decays $B^- \to \ell^- \bar{\nu}_\ell$ is again given by Eq. (14), with appropriate substitutions. Here, the resulting SM branching fractions, shown in Table 3, are even smaller than in the case of charm mesons due to the value of V_{ub}. These SM predictions are somewhat imprecise due to the uncertainty in f_B and V_{ub}, and hence can vary over the range

$$B_{SM} \left(\frac{f_B}{180\,\text{MeV}}\right)^2 \left(\frac{V_{ub}}{0.0035}\right)^2, \quad (20)$$

where B_{SM} is the result listed in the table. We see from the table that the 90% C.L. experimental bounds[20,26] are roughly one to two orders of magnitude above the SM predictions for the cases of $B^- \to \mu^- \bar{\nu}_\mu, \tau^- \bar{\nu}_\tau$. The B factories presently

Mode	SM Prediction	Experimental Bound
$e^- \bar{\nu}_e$	6.9×10^{-12}	$< 1.5 \times 10^{-5}$ (CLEO)
$\mu^- \bar{\nu}_\mu$	2.9×10^{-7}	$< 2.1 \times 10^{-5}$ (CLEO)
$\tau^- \bar{\nu}_\tau$	6.6×10^{-5}	$< 5.7 \times 10^{-4}$ (L3)

Table 3: SM branching fractions for the B^- leptonic decay modes, assuming $f_B = 180$ MeV and taking the central value[4] of the CKM matrix element V_{ub}. The results of experimental searches[19,25] are also shown.

Decay Constant	Lattice	Quark Model Nonrelativistic	Quark Model Relativistic	Sum Rule
f_{B_d}	175 ± 25	140 ± 40	155 ± 15	155 – 185
f_{B_s}	200 ± 25	175 ± 50	210 ± 20	195 – 220
f_{B_s}/f_{B_d}	1.15 ± 0.05	1.25		1.22 ± 0.02

Table 4: Theoretical estimates of the weak decay constants in units of MeV (taking $m_b = 4.67$ GeV in the sum rule approach).

under construction at SLAC and KEK should be able to observe $B \to \tau\nu_\tau$ (and potentially the $\mu\bar{\nu}_\mu\gamma$ mode as well). This measurement will require the full (or partial) reconstruction of the other B's in the event as well as a large statistical sample. Theoretical estimates for $f_{B_{d,s}}$ are tabulated in Table 4 using the same approaches as in the cases discussed above in charm decays. See Ref. 23 for a more complete compilation.

Observation of these decays would, of course, provide a classic measurement of the decay constant f_B (assuming V_{ub} is known from other sources), but only if no new physics contributes to the decay. For example, in two-Higgs-doublet models (2HDM), tree-level charged Higgs exchange can again mediate this transition. In the 2HDM of Type II, the branching fraction is now modified by

$$B(B^- \to \ell^- \bar{\nu}_\ell) = B_{SM} \left(1 - \tan^2\beta \frac{m_B^2}{m_{H^\pm}^2}\right)^2. \quad (21)$$

Taking the SM and L3 bound on $B^- \to \tau^- \bar{\nu}_\tau$ listed in Table 3 implies $\tan\beta/m_{H^\pm} < 0.38$ GeV^{-1}, assuming $f_B = 190$ MeV and $|V_{ub}| = 0.0033 \pm 0.0008$. Once this decay is detected, tests for this type of scalar exchange can be performed by measuring the helicity of the final state τ. The measured branching fraction from

LEP for the decay $B \to X\tau\nu$ yields[25,26] a similar constraint of $\tan\beta/m_{H^\pm} < 0.52$ GeV^{-1}, which is independent of the uncertainties in f_B and V_{ub}.

4 Flavor Changing Neutral Current Decays

Flavor changing neutral current (FCNC) decays only occur at the loop level in the SM. Tree-level neutral currents are flavor diagonal due to the fact that all fermions with the same charge and helicity have identical transformation properties under the $SU(2)_L \times U(1)_Y$ gauge group,[27] so that the flavor to mass eigenstate rotation matrices commute with the neutral current operator. In fact, this property provided the original motivation for Glashow, Iliopoulos, and Maiani[28] (GIM) to introduce the charm quark in order to suppress phenomenologically unacceptable large values of FCNC processes in the kaon system. This allowed for the strange quark to have the same electroweak quantum number assignments as the down quark; hence eradicating the tree-level strangeness changing neutral current. The GIM mechanism thus achieves this tree-level cancellation without any artificial adjustments to the parameters of the theory and also provides additional suppression for FCNC that are induced at the loop level.

The one-loop processes which mediate FCNC's can generally be classified as electromagnetic, weak, or gluon penguin diagrams and box diagrams. Samples of these classes of diagrams are displayed in Fig. 1. The generic amplitude for a diagram of this type can be written as

$$\mathcal{A} \sim \sum_{i=1}^{3} V_{iQ} V_{iq}^* F(m_i^2/M_W^2)\,, \tag{22}$$

where the sum extends over the three generations of quarks of mass m_i contributing internally to the diagram, the V_{ij} are the relevant CKM elements appearing at the vertices, and the function F represents the result of the loop integrals for the diagrams. Using the unitarity property of the three-generation CKM matrix, $\sum_i V_{iQ} V_{iq}^* = 0$, allows one to rewrite the amplitude as

$$\mathcal{A} \sim V_{2Q} V_{2q}^* \left[F(m_2^2/M_W^2) - F(m_1^2/M_W^2)\right] + V_{3Q} V_{3q}^* \left[F(m_3^2/M_W^2) - F(m_1^2/M_W^2)\right]\,. \tag{23}$$

This clearly demonstrates that the amplitude would vanish if all the contributing internal quarks were degenerate! Hence, the magnitude of FCNC transitions is related to the size of the internal quark mass splittings. This point is illustrated in

Meson/Quark	Branching Fraction
K	$10^{-10} - 10^{-8}$
D	$10^{-18} - 10^{-10}$
B	$10^{-8} - 10^{-4}$
t	$10^{-12} - 10^{-9}$

Table 5: Typical values of FCNC branching fractions in the SM.

Table 5 which displays the typical SM range of FCNC branching fractions for each meson/quark system. As we would expect, the B meson system has the largest FCNC rates due to the large degree of mass splitting in the up-quark sector and due to the diagonal structure of the CKM matrix, whereas the charm mesons and top-quark rates are very suppressed by the efficiency of the GIM mechanism.

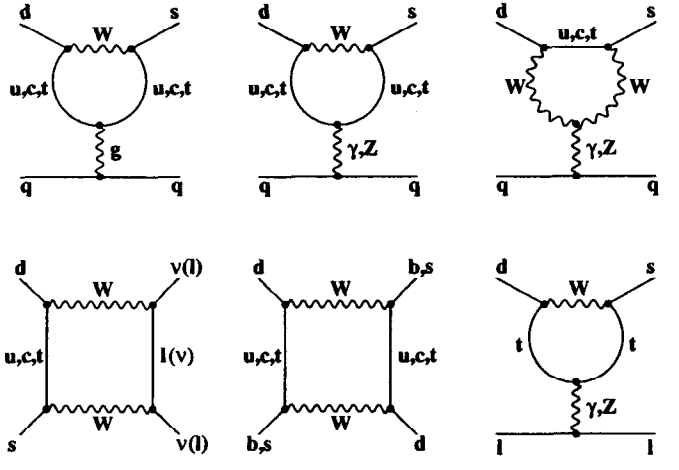


Figure 1: Feynman diagrams.

QCD corrections to these rare processes can be quite important. They are computed[29] via the Operator Product Expansion combined with renormalization group evolution. This procedure allows for an efficient separation of short-distance physics (corresponding to scales higher than μ) and long-distance contributions (scales lower then μ). Within this framework, the exclusive transition $M \to F$

can be expressed as

$$\mathcal{A}(M \to F) = \frac{G_F}{\sqrt{2}} V_{CKM} \sum_i C_i(\mu) \langle F|\mathcal{O}_i(\mu)|M\rangle \,, \tag{24}$$

where $\mathcal{O}_i$ represents the complete set of effective operators which govern the transition, the C_i are the Wilson coefficients which are related to the Inami-Lim functions at the scale $\mu = M_W$, and μ corresponds to the scale at which the transition takes place. The μ dependence of the Wilson coefficients is given by the renormalization group equations (RGE) and must be cancelled by the μ dependence contained in $\langle \mathcal{O}_i \rangle$. The use of the RGE allows for the summation of the large logarithms $\log(M_W/\mu)$ at a given order in perturbation theory. The long-distance, or nonperturbative, contributions are contained in the evaluation of the matrix elements of the operators.

4.1 Radiative Decays

We start our discussion of FCNC transitions with the study of radiative decays, $Q \to q'\gamma$. The most general Lorentz decomposition of the radiative amplitude is

$$\begin{aligned} \mathcal{A}(Q \to q'\gamma) &= \epsilon^\lambda(q) \langle q'|J^{em}_\lambda|Q\rangle \,, \\ &= \epsilon^\lambda(q) \bar{u}_{q'}(p-q) \left[iq^\nu \sigma_{\lambda\nu}(A + B\gamma_5) + \gamma_\lambda(C + D\gamma_5) \right. \\ &\qquad \left. + q_\lambda(E + F\gamma_5) \right] u_Q(p) \,, \end{aligned} \tag{25}$$

where ϵ^λ is the photon polarization, q represents the outgoing photon's momentum, and $A - F$ are the invariant amplitudes for each case. Electromagnetic gauge invariance, which dictates $\partial^\lambda J^{em}_\lambda = 0$, yields the condition

$$-m_{q'}(C + D\gamma_5) + m_Q(C - D\gamma_5) + q^2(E + F\gamma_5) = 0 \,. \tag{26}$$

For an on-shell photon ($q^2 = 0$), this gives $C = D = 0$. Folding in the property $\epsilon^\lambda \cdot q_\lambda = 0$, we are left with the magnetic dipole transition amplitude

$$\mathcal{A}(Q \to q'\gamma) = \epsilon^\lambda(q) \bar{u}_{q'}(p-q) \left[iq^\nu \sigma_{\lambda\nu}(A + B\gamma_5) \right] u_Q(p) \,. \tag{27}$$

This amplitude is represented by a gauge invariant set of loop diagrams (in this case, electromagnetic penguin diagrams) which sum to a finite result as there are

Quark	F_2	$\lvert V^*_{ci} V_{ui}\rvert F_2$	$\lvert V_{ib} V^*_{is}\rvert F_2$
d	1.57×10^{-9}	3.36×10^{-10}	
s	2.92×10^{-7}	6.26×10^{-8}	
b	3.31×10^{-4}	3.17×10^{-8}	
u	2.27×10^{-9}		1.29×10^{-12}
c	2.03×10^{-4}		7.34×10^{-6}
t	0.39		1.56×10^{-2}

Table 6: Contributions to $c \to u\gamma$ and $b \to s\gamma$.

no counterterms to absorb the infinities. The perturbative calculation of these diagrams yield the familiar result (neglecting the mass of the final state quark)

$$\begin{aligned} \Gamma(Q \to q'\gamma) &= \frac{G_F^2 \alpha m_Q^5}{128\pi^4} \left| V_{2Q} V^*_{2q} \left[F_2(m_2^2/M_W^2) - F_2(m_1^2/M_W^2) \right] \right. \\ &\qquad \left. + V_{3Q} V^*_{3q} \left[F_2(m_3^2/M_W^2) - F_2(m_1^2/M_W^2) \right] \right|^2 \,, \end{aligned} \tag{28}$$

where the function F_2 is given in Inami and Lim.[30] It is instructive to compare the magnitude of these functions for the decays $c \to u\gamma$ and $b \to s\gamma$, for the various internal quark states; this is presented in Table 6. Dominance of the t-quark intermediate state in $b \to s\gamma$ is evident, even upon including the CKM factors. Its effect is so large that the other intermediate states are numerically negligible and hence are typically omitted. The amplitudes for $c \to u\gamma$ differ from that of $b \to s\gamma$ in two important respects: (i) there is no single intermediate state which dominates, and (ii) the overall magnitude is much smaller. The effectiveness of the GIM mechanism is clearly demonstrated.

4.1.1 Radiative B Decays

Radiative B decays have become the benchmark FCNC process and provide one of the best testing grounds of the SM. The CLEO Collaboration has reported[31] the observation of the inclusive decay $B \to X_s\gamma$ with a branching fraction of $(2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$ and 95% C.L. bounds of $1 \times 10^{-4} < B(B \to X_s\gamma) < 4.2 \times 10^{-4}$, as well as an updated measurement[32] for the related exclusive process $B(B \to K^*\gamma) = (4.2 \pm 0.8 \pm 0.6) \times 10^{-5}$. This yields a value of 0.181 ± 0.068 for the ratio of exclusive to inclusive rates. On the theoretical side, the reliability of the calculation of the quark-level process $B \to X_s\gamma$ has improved with the

completion of the next-to-leading logarithmic QCD corrections. It has thus provided strong restrictions on the parameters of several theories beyond the SM.[33] This constitutes the first direct observation of a penguin mediated process(!) and demonstrates the fertile ground ahead for the detailed exploration of the SM in rare B transitions.

In the SM, the quark-level transition $B \to X_s\gamma$ is mediated by W-boson and t-quark exchange in an electromagnetic penguin diagram, as discussed above. To obtain the branching fraction, the inclusive rate is scaled to that of the semileptonic decay $B \to X\ell\bar{\nu}_\ell$. This procedure removes uncertainties from the overall factor of m_b^5 and reduces the ambiguities involved with the imprecisely determined CKM factors. The result is then rescaled by the experimental value for the semileptonic branching fraction,

$$B(B \to X_s\gamma) = \frac{\Gamma(B \to X_s\gamma)}{\Gamma(B \to X\ell\bar{\nu}_\ell)} B(B \to X\ell\bar{\nu}_\ell)\,. \tag{29}$$

The QCD corrections are calculated[34] via an operator product expansion based on the effective Hamiltonian

$$\mathcal{H}_{eff} = \frac{-4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_{i=1}^{8} C_i(\mu)\mathcal{O}_i(\mu)\,, \tag{30}$$

which is then evolved from the electroweak scale down to $\mu \sim m_b$ by the Renormalization Group Equations (RGE). The $\mathcal{O}_i$ are a complete set of renormalized operators of dimension six or less which mediate $b \to s$ transitions. They consist of the two current-current operators $O_{1,2}$, the four strong penguin operators O_{3-6}, and the electro- and chromo-magnetic dipole operators O_7 and O_8, respectively, and can be written as

$$\begin{aligned}
\mathcal{O}_1 &= (\bar{c}_\alpha\gamma_\mu P_L b_\beta)(\bar{s}_\beta\gamma^\mu P_L c_\alpha)\,,\\
\mathcal{O}_2 &= (\bar{c}_\alpha\gamma_\mu P_L b_\alpha)(\bar{s}_\beta\gamma^\mu P_L c_\beta)\,,\\
\mathcal{O}_3 &= (\bar{s}_\alpha\gamma_\mu P_L b_\alpha)\sum_q(\bar{q}_\beta\gamma^\mu P_L q_\beta)\,,\\
\mathcal{O}_4 &= (\bar{s}_\alpha\gamma_\mu P_L b_\beta)\sum_q(\bar{q}_\beta\gamma^\mu P_L q_\alpha)\,,\\
\mathcal{O}_5 &= (\bar{s}_\alpha\gamma_\mu P_L b_\alpha)\sum_q(\bar{q}_\beta\gamma^\mu P_R q_\beta)\,,\\
\mathcal{O}_6 &= (\bar{s}_\alpha\gamma_\mu P_L b_\beta)\sum_q(\bar{q}_\beta\gamma^\mu P_R q_\alpha)\,,
\end{aligned} \tag{31}$$

$$\begin{aligned}
\mathcal{O}_7 &= \frac{e}{16\pi^2} m_b(\bar{s}_\alpha\sigma_{\mu\nu}P_R b_\alpha)F^{\mu\nu}\,,\\
\mathcal{O}_8 &= \frac{g_s}{16\pi^2} m_b(\bar{s}_\alpha\sigma_{\mu\nu}T^a_{\alpha\beta}P_R b_\beta)G^{a\mu\nu}\,,
\end{aligned}$$

where the terms proportional to m_s in $O_{7,8}$ have been neglected. We note that the magnetic and chromomagnetic dipole operators, $\mathcal{O}_{7,8}$, contain explicit mass factors which must also be renormalized as shown below.

The C_i represent the corresponding Wilson coefficients which are evaluated perturbatively at the electroweak scale, where the matching conditions are imposed and then evolved down to the renormalization scale μ. The expressions for the coefficients at the W scale are

$$\begin{aligned}
C_{1,3-6}(M_W) = 0 \quad &, \qquad C_2(M_W) = 1\,,\\
C_7(M_W) = -\frac{1}{2}F_2(x_t) \quad &, \quad C_8(M_W) = -\tfrac{1}{2}D(x_t)\,,
\end{aligned} \tag{32}$$

with $x \equiv m_t^2/M_W^2$ and

$$\begin{aligned}
F_2(x) &= Q\left[\frac{x^3-5x^2-2x}{4(x-1)^3} + \frac{3x^2\ln x}{2(x-1)^4}\right] + \frac{2x^3+5x^2-x}{4(x-1)^3} - \frac{3x^3\ln x}{2(x-1)^4}\,,\\
D(x) &= \frac{x^3-5x^2-2x}{4(x-1)^3} + \frac{3x^2\ln x}{2(x-1)^4}\,,
\end{aligned} \tag{33}$$

where Q represents the charge of the internal quark.

The leading logarithmic QCD corrections to the decay width have been completely resummed, but lead to a sizable μ dependence of the branching fraction, and hence, it is essential to include the next-to-leading order corrections to reduce the theoretical uncertainty. In this case, the calculation of the perturbative QCD corrections involves several steps, requiring corrections to both the Wilson coefficients and the matrix elements of the operators in Eq. (30) in order to ensure a scheme-independent result. For the matrix elements, this includes the QCD bremsstrahlung corrections[35] $b \to s\gamma + g$, and the NLO virtual corrections which have recently been completed in both the naive dimensional regularization (NDR) and 't Hooft-Veltman schemes.[36] Summing these contributions to the matrix elements and expanding them around $\mu = m_b$, one arrives at the decay amplitude

$$\mathcal{M}(b \to s\gamma) = -\frac{4G_F V_{tb}V_{ts}^*}{\sqrt{2}} D\langle s\gamma|\mathcal{O}_7(m_b)|b\rangle_{tree}\,, \tag{34}$$

with

$$D = C_7^{eff}(\mu) + \frac{\alpha_s(m_b)}{4\pi}\left(C_i^{(0)eff}(\mu)\gamma_{i7}^{(0)eff}\log\frac{m_b}{\mu} + C_i^{(0)eff}r_i\right)\,. \tag{35}$$

Here, the quantities $\gamma_{i7}^{(0)eff}$ are the entries of the effective leading order anomalous dimension matrix, and the r_i are computed in Greub *et al.*,[36] for $i = 2, 7, 8$. The first term in Eq. (35), $C_7^{eff}(\mu)$, must be computed at NLO precision, while it is consistent to use the leading order values of the other coefficients. The explicit logarithms $\alpha_s(m_b)\log(m_b/\mu)$ in the equation are cancelled by the μ dependence of $C_7^{(0)eff}(\mu)$. This feature significantly reduces the scale dependence of the resulting branching fraction. The contribution to the inclusive width including these virtual corrections is then

$$\Gamma_{NLO}^{virt}(B \to X_s\gamma) = \frac{m_{b,pole}^5 G_F^2 \alpha_{em} |V_{tb}V_{ts}^*|^2}{32\pi^4} F|D|^2 , \qquad (36)$$

where the factor $F = m_b^2(m_b)/m_{b,pole}^2 = 1 - 8\alpha_s(m_b)/3\pi$ arises from the mass factor present in the magnetic dipole operator. This should be compared to the familiar leading order result (which omits the virtual corrections to $\langle \mathcal{O}_7 \rangle$)

$$\Gamma(B \to X_s\gamma) = \frac{m_{b,pole}^5 G_F^2 \alpha_{em}}{32\pi^4} |V_{tb}V_{ts}^*|^2 |C_7^{eff}(\mu)|^2 . \qquad (37)$$

For the Wilson coefficients, the NLO result entails the computation of the $\mathcal{O}(\alpha_s)$ terms in the matching conditions, and the renormalization group evolution of the $C_i(\mu)$ must be computed using the $\mathcal{O}(\alpha_s^2)$ anomalous dimension matrix. The former step has been computed in Ref. 37. The latter step is quite difficult, since some entries in the matrix have to be extracted from three-loop diagrams, and has recently been completed,[38] with the conclusion being that in the NDR scheme the NLO correction to $C_7^{eff}(\mu)$ is small.

The total inclusive width is then given by the sum of the virtual and bremsstrahlung corrections, $\Gamma(B \to X_s\gamma) = \Gamma^{virt} + \Gamma^{brems}$, where Γ^{brems} is given in Greub *et al.*,[35,36] and the branching fraction is calculated by scaling to the semileptonic decay rate. The leading order power corrections in the heavy quark expansion are identical for $B \to X_s\gamma$ and $B \to X\ell\bar{\nu}_\ell$, and hence cancel in the ratio.[39] This allows us to approximate $\Gamma(B \to X_s\gamma)$ with the perturbatively calculable free quark decay rate. For $m_t^{phys} = 175 \pm 6\,\text{GeV}$ (Ref. 13), $m_b/2 \le \mu \le 2m_b$, $\alpha_s = 0.118 \pm 0.003$ (Refs. 4, 40), $B_{sl} = (10.23 \pm 0.39)\%$ (Ref. 41), and $m_c/m_b = 0.29 \pm 0.02$, we find the branching fraction

$$B(B \to X_s\gamma) = (3.25 \pm 0.30 \pm 0.40) \times 10^{-4} , \qquad (38)$$

where the first error corresponds to the combined uncertainty associated with the value of m_t and μ, and the second error represents the uncertainty from the other parameters. This is well within the range observed by CLEO. In Fig. 2, the inclusive branching fraction is displayed as a function of the top mass from Ref. 42. The dashed lines indicate the error in the branching ratio if we fix $\mu = m_b$ and vary all the other parameters over their allowed ranges given above. The solid lines indicate the error for $m_b/2 < \mu < 2m_b$ with all other parameters fixed to their central values. This visually demonstrates that the error in the theoretical calculation of $B \to X_s\gamma$ is not overwhelmed by the scale uncertainty; other uncertainties are now comparable. Within the SM (and assuming $V_{tb} = 1$), comparison with the experimental result gives $|V_{ts}/V_{cb}| = 0.85 \pm 0.12(exp) \pm 0.10(th)$, which is consistent with unity.[43]

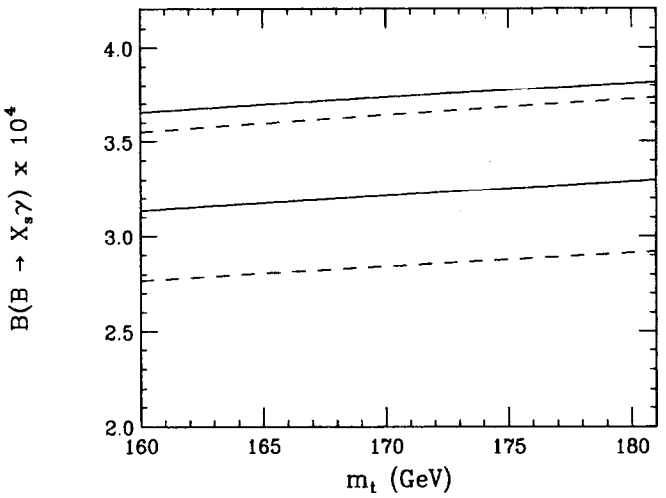


Figure 2: The branching ratio of $B \to X_s\gamma$ vs m_t. The dashed lines indicate the error in the branching ratio if we fix $\mu = m_b$ and vary all the other parameters over their allowed ranges: $\alpha_s(M_Z) = 0.118 \pm 0.003$, $B_{sl} = 10.23 \pm 0.39\%$, and $m_c/m_b = 0.29 \pm 0.02$. The solid lines indicate the error for $m_b/2 < \mu < 2m_b$ and all other parameters fixed to their central values.

Before discussing explicit models of new physics, we first investigate the constraints placed directly on the Wilson coefficients of the magnetic moment operators from the CLEO measurement of $B \to X_s\gamma$. Writing the coefficients at the matching scale in the form $C_i(M_W) = C_i^{SM}(M_W) + C_i^{new}(M_W)$, where $C_i^{new}(M_W)$

represents the contributions from new interactions. Due to operator mixing, the CLEO measurement of $B \to X_s\gamma$ then limits the possible values for $C_i^{new}(M_W)$ for $i = 7, 8$. These bounds are summarized in Fig. 3, where the allowed regions lie inside the diagonal bands.[42] We note that two bands occur due to the overall sign ambiguity in the determination of the coefficients. Here, the solid bands correspond to the constraints obtained from the current CLEO measurement, taking into account the variation of the renormalization scale $m_b/2 \leq \mu \leq 2m_b$, as well as the allowed ranges of the other input parameters. The dashed bands represent the constraints when the scale is fixed to $\mu = m_b$. We note that large values of $C_8^{new}(M_W)$ are allowed even in the region where $C_7^{new}(M_W) \simeq 0$. Experimental bounds on the decay $b \to sg$ are needed to constrain C_8.

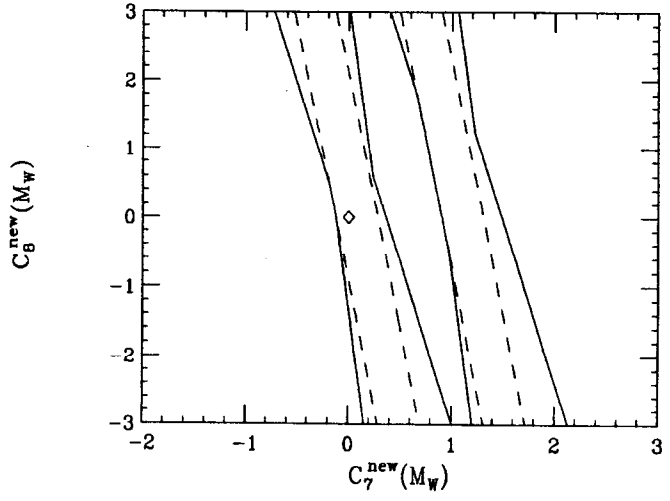


Figure 3: Bounds on the contributions from new physics to $C_{7,8}$. The region allowed by the CLEO data corresponds to the area inside the solid diagonal bands. The dashed bands represent the constraints when the renormalization scale is set to $\mu = m_b$. The diamond at the position (0,0) represents the Standard Model.

• Fourth Generation

In the case of four families, there is an additional contribution to $B \to X_s\gamma$ from the virtual exchange of the fourth generation up-quark t' (Ref. 44). The Wilson coefficients of the dipole operators are then modified by

$$C_{7,8}(M_W) = C_{7,8}^{SM}(m_t^2/M_W^2) + \frac{V_{t'b}V_{t's}^*}{V_{tb}V_{ts}^*} C_{7,8}^{SM}(m_{t'}^2/M_W^2)\,. \qquad (39)$$

V_{ij} represents the 4 x 4 CKM matrix which now contains nine parameters; six angles and three phases. The values of the elements of the 4 x 4 CKM matrix are much less restricted than their three-generation counterparts, as one can no longer apply the three-generation unitarity constraints.[4] Hence, even the overall CKM factor in the $B \to X_s\gamma$ branching ratio, $|V_{tb}V_{ts}^*/V_{cb}|$, can take on different values. Figure 4(a) displays the resulting branching fraction as a function of $m_{t'}$ for $m_t = 180$ GeV; here the vertical lines represent the range of possible values as the CKM elements are varied. These ranges were determined by generating 10^8 sets of the nine parameters in the 4 x 4 CKM matrix and demanding consistency with (i) four-generation unitarity and the extraction of the CKM elements from charged current measurements, (ii) the value of the ratio $|V_{ub}/V_{cb}|$, (iii) ϵ, and (iv) $B^0 - \bar{B}^0$ mixing. We see that there is little or no sensitivity to the t'-quark mass, and that the CLEO measurement places additional constraints on the 4 x 4 CKM matrix. In fact, we find that consistency with CLEO demands $0.20 \leq |V_{tb}V_{ts}| \leq 1.5 \times 10^{-2}$ and $0.23 \leq |V_{t'b}V_{t's}| \leq 1.1 \times 10^{-3}$.

• Two-Higgs-Doublet Models

In 2HDM, the $H^\pm$ contributes to $B \to X_s\gamma$ via virtual exchange together with the top quark. At the W scale, the coefficients of the dipole operators take the form (in Model II described above)

$$C_i(M_W) = C_i^{SM}(m_t^2/M_W^2) + A_{1_i}^{H^\pm}(m_t^2/m_{H^\pm}^2) + \frac{1}{\tan^2\beta} A_{2_i}^{H^\pm}(m_t^2/m_{H^\pm}^2)\,, \qquad (40)$$

where $i = 7, 8$. The analytic form of the functions A_{1_i}, A_{2_i} can be found in Refs. 45 and 46. In Model II, large enhancements appear for small values of $\tan\beta$, but more importantly, we see that $B(B \to X_s\gamma)$ is always larger than that of the SM, independent of the value of $\tan\beta$ due to the presence of the $A_{1_i}^{H^\pm}$ term. This leads to the familiar bound[31] $m_{H^\pm} > 260$ GeV obtained from the measurement of $B(B \to X_s\gamma)$ by CLEO. However, this constraint does not make use of the recent NLO calculation. We remind the reader that a full NLO calculation would also require the higher order matching conditions for the $H^\pm$ contributions. Nevertheless, we recall that the results on the NLO corrections to $C_7^{eff}(\mu)$ indicate they are small,[38] and a good approximation is obtained by employing the uncorrected $H^\pm$

matching conditions. Since the NLO corrections to the matrix element drastically reduces the μ dependence of the branching fraction, we would expect the resulting $H^\pm$ constraints to improve. Indeed, we find that[42] the CLEO bound excludes the region to the left and beneath the curves in Fig. 4(b). For $m_t^{phys} = 169$ GeV, we see that $m_{H^\pm} > 300$ GeV. This is calculated by using the same procedure that produced the previous charged Higgs mass bound by CLEO, *i.e.*, all the input parameters (*e.g.*, α_s, μ, m_c/m_b, and $B(B \to X\ell\bar{\nu}_\ell)$) are varied over their allowed ranges in order to ascertain the most conservative limit. This bound holds in the general two-Higgs-doublet-model II, and in supersymmetry if the superpartners are all significantly massive.

• **Supersymmetry**

There are several new classes of contributions to $B \to X_s\gamma$ in supersymmetry. The large $H^\pm$ contributions from Model II discussed above are present; however, the limits obtained in supersymmetric theories also depend on the size of the other super-particle contributions and are generally much more complex. In particular, it has been shown[47,48] that large contributions can arise from stop-squark and chargino exchange (due to the possibly large stop-squark mass splitting), as well as from the gluino and down-type squark loops (due to left-right mixing in the sbottom sector). The additional neutralino-down-squark contributions are expected to be small. Some regions of the parameter space can thus cancel the $H^\pm$ contributions resulting in predictions for the branching fraction at (or even below) the SM value, while other regions always enhance the amplitude. In minimal supergravity models with radiative breaking, the sign of the sparticle loop contributions is found to be correlated with the sign of the higgsino mass parameter μ (Refs. 48, 49). A scatter plot in the $R_7 - R_8$ plane is presented[42] in Fig. 4(c), where $R_i \equiv \frac{C_i^{susy}(M_W)}{C_i^{SM}(M_W)} - 1 \equiv \frac{C_i^{new}(M_W)}{C_i^{SM}(M_W)}$. Each point in the scatter plot is derived from the minimal supergravity model with different initial conditions, and is consistent with all collider bounds and is out of reach of LEPII. The first thing to note from the figure is that large values of R_7 and R_8 are generated, and the R_7 and R_8 values are very strongly correlated. The diagonal bands represent the bounds on the Wilson coefficients from the observation of $B \to X_s\gamma$ as determined previously. We see that the current CLEO data already places significant restrictions on the supersymmetric parameter space. Further constraints will be obtainable once a 10% measurement of $B(B \to X_s\gamma)$ is made, and the sign of C_7 is determined from a global fit described below. In this case, if no deviations from the SM are observed, the supersymmetric contributions will be restricted to lie in the dashed band. It is clear that these processes can explore vast regions of the supersymmetric parameter space. In fact, it is possible that spectacularly large deviations in rare B decays could be manifest at B factories, while collider experiments would not detect a hint of new physics.

• **Anomalous Top-Quark Couplings**

If the top quark has anomalous couplings to on-shell photons or gluons, the rate for $B \to X_s\gamma$ would be modified. The effect of an anomalous magnetic and/or electric dipole moment in the Lagrangian

$$\mathcal{L} = e\bar{t}\left[Q_t\gamma_\mu + \frac{1}{2m_t}\sigma_{\mu\nu}(\kappa_\gamma + i\tilde{\kappa}_\gamma\gamma_5)q^\nu\right]tA^\mu\,, \tag{41}$$

on the Wilson coefficients is

$$C_{7,8}(M_W) = C_{7,8}^{SM}(m_t^2/M_W^2) + \kappa_{\gamma,g}F_{1_{\gamma,g}}(m_t^2/M_W^2) + \tilde{\kappa}_{\gamma,g}F_{2_{\gamma,g}}(m_t^2/M_W^2)\,. \tag{42}$$

The functions $F_{1,2}$ can be found in Ref. 50. The effects of anomalous chromodipole moments arise from operator mixing. When the resulting branching fraction and the CLEO data are combined, the constraints (at leading order) in Fig. 4(d) are obtained[51] for $m_t = 180$ GeV. In this figure, the allowed region is given by the area inside the solid (dashed) semicircle when $\kappa_g, \tilde{\kappa}_g = 0 (= \kappa_\gamma, \tilde{\kappa}_\gamma)$. These bounds are considerably weaker than those obtainable from direct top-quark production at colliders.[52]

One of the goals of a high-luminosity B physics program is to extract the ratio of CKM elements $|V_{td}|/|V_{ts}|$ from a measurement of

$$\frac{B(B^- \to \rho^-\gamma)}{B(B^- \to K^{*-}\gamma)} = \frac{B(B^0 \to \rho^0\gamma) + B(B^0 \to \omega\gamma)}{B(B^0 \to K^{*0}\gamma)} = \left|\frac{V_{td}}{V_{ts}}\right|^2 \xi\Omega\,, \tag{43}$$

where ξ accounts for SU(3) symmetry breaking and Ω represents the phase space ratio. CLEO has recently placed[32] the bounds on the exclusive branching fractions, $B(B^0 \to \rho^0\gamma) < 3.9 \times 10^{-5}$, $B(B^- \to \rho^-\gamma) < 1.1 \times 10^{-5}$, and $B(B^0 \to \omega^0\gamma) < 1.3 \times 10^{-5}$. Combining this with their measurement of $B \to K^*\gamma$ and theoretical estimates[53] of the SU(3) breaking factor places the 90% C.L. constraint $|V_{td}|/|V_{ts}| < 0.45 - 0.56$. However, this technique of determining this ratio of

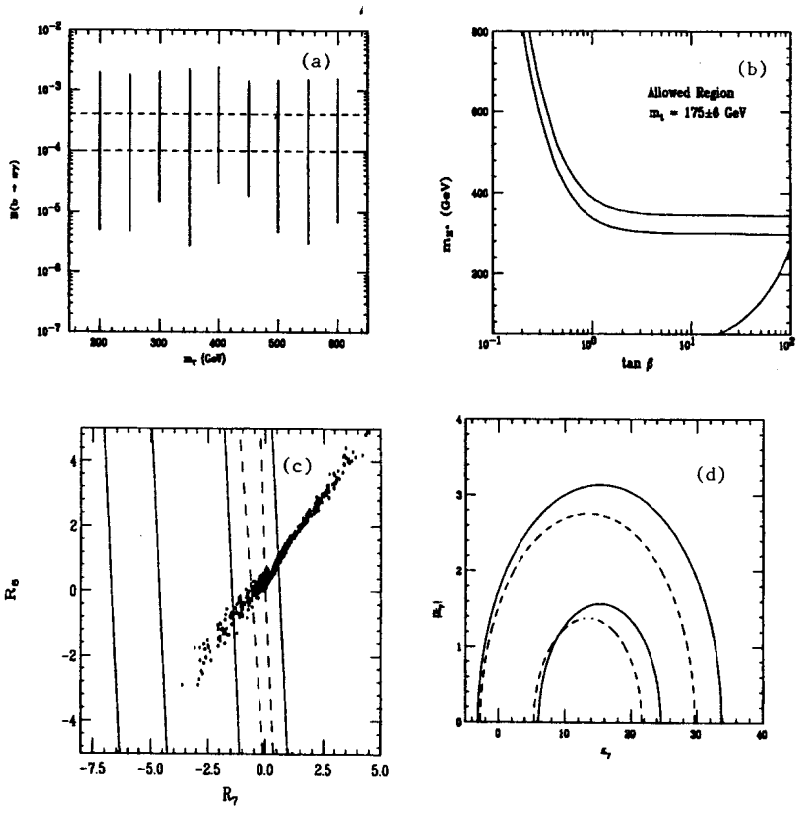


Figure 4: (a) The range of values for $B(B \to X_s\gamma)$ in the four-generation SM as a function of $m_{t'}$. (b) Constraints in the charged Higgs mass – $\tan\beta$ plane from the CLEO bound on $B(B \to X_s\gamma)$. The excluded region is that to the left and below the curves. The top line is for $m_t^{phys} = 181$ GeV and the bottom line is for $m_t^{phys} = 169$ GeV. We also display the restriction $\tan\beta/m_{H^\pm} > 0.52$ GeV^{-1} which arises from measurements of $B \to X\tau\nu$ as discussed in Ref. 25. (c) Parameter space scatter plot of R_7 vs R_8 in minimal supergravity model. The allowed region from CLEO data, as previously obtained, lies inside the two sets of solid diagonal bands. The dashed band represents the constraints from a potential 10% measurement of $B \to X_s\gamma$. (d) Bounds on anomalous top-quark photon couplings from $B \to X_s\gamma$. The solid and dashed curves correspond to the cases described in the text. In each case, the allowed regions lie inside the semicircles.

CKM elements depends critically on the assumption that these exclusive decay modes are dominated by short-distance penguin transitions. If this assumption is false, and the long-distance contributions to these decays were found to be large, this method would be invalidated. In fact, it has been pointed out by numerous authors[54] that long-distance contributions to $B \to X_d\gamma$ may be significant, and hence, these decays may not yield a good determination of the CKM element V_{td}. These long-distance effects originate from annihilation diagrams and from the inclusion of the light-quark intermediate states in the penguin amplitudes. However, separate measurements of charged and neutral B decays into $\rho\gamma$ and $\omega\gamma$ may be useful in sorting out the magnitude of the long-distance contributions. In contrast, long-distance effects in exclusive $B \to X_s\gamma$ decays are estimated to be small.[55]

4.1.2 Radiative Charm Decays

It is instructive to compare radiative decays in the charm system, $D \to X_u + \gamma$, with those of B mesons. While separation of the long- from the short-distance contributions is somewhat difficult in the B sector, radiative charm decays provide an excellent laboratory for the determination of the long-distance effects and would hence test the calculational models. In the charm case, it should be possible to directly determine the rate of the long-distance reactions which are expected to dominate. For example, the inclusive $c \to u\gamma$ penguin transitions do not contribute to $D^0 \to \bar{K}^{*0}\gamma$, and this mode would be a direct measurement of the nonperturbative effects. Before QCD corrections are applied, the short-distance inclusive rate is extremely small with $B(c \to u\gamma) \approx 1 \times 10^{-17}$; however, the QCD corrections *greatly* enhance this rate.[56] These corrections are calculated via an operator product expansion, where the effective Hamiltonian is evolved at leading logarithmic order from the electroweak scale down to the charm quark scale by the Renormalization Group Equations. This procedure mirrors that used for $b \to s\gamma$, except that two effective Hamiltonians must be introduced in order to correctly account for the evolution above and below the scale $\mu = m_b$. We thus have

$$\mathcal{H}_{eff}(M_W > \mu > m_B) = -\frac{4G_F}{\sqrt{2}} \sum_{q=d,s,b} V_{cq}^* V_{uq} [C_1(\mu)\mathcal{O}_1^q + C_2(\mu)\mathcal{O}_2^q], \quad (44)$$

$$\mathcal{H}_{eff}(m_b > \mu > m_c) = -\frac{4G_F}{\sqrt{2}} \sum_{q=d,s,b} V_{cq}^* V_{uq} [C_1(\mu)\mathcal{O}_1^q + C_2(\mu)\mathcal{O}_2^q + \sum_{i=3}^{8} C_i(\mu)\mathcal{O}_i],$$

where the operators are as defined in Eq. (32) with the appropriate substitutions. This procedure results[57] in the inclusive branching fraction $B(D \to X_u\gamma) \simeq (2-5) \times 10^{-8}$, with the range corresponding to the difference between neutral and charged D decay. (We note that these radiative branching fractions have also been scaled to semileptonic charm decay in order to reduce the CKM and m_c uncertainties.) We see that in this case, the rate is given entirely by operator mixing! The penguin contributions to the exclusive channels would then be typically of order 10^{-9}, which is still significantly smaller than the long-distance estimates presented in the following text in Table 8. We note that for radiative charm decays, the predicted values of the exclusive branching fractions from long-distance effects are within reach of B factories.

4.2 Other Rare B Decays

As discussed above, FCNC processes in the B sector are not as suppressed as in the other meson systems and can occur at reasonable rates in the SM. This is due to a sizable loop-level contribution from the top quark, which results from the combination of the large top mass (giving a big GIM splitting) and the diagonal nature of the CKM matrix. Long-distance effects are expected to play less of a role due to the heavy B mass, and hence rare processes are essentially short-distance dominated. Many classes of new models can also give significant and testable contributions to rare B transitions.

Other FCNC decays of B mesons include $B_{d,s}^0 \to \ell^+\ell^-, \gamma\gamma$, $B \to X_{s,d} + \ell^+\ell^-, X_{s,d}\nu\bar{\nu}$, with $\ell = e\mu\tau$. In the SM, they are mediated by appropriate combinations of electromagnetic and weak penguins as well as box diagrams, and generally have larger rates, as discussed above, due to the heavy top quark and the diagonal nature of the CKM matrix. The SM predictions and current experimental situation[4,58,59] for these decays are summarized in Table 7, taking $m_t = 180$ GeV. The purely leptonic decays, $B^0 \to \ell^+\ell^-$, can be enhanced by contributions from new physics at both the loop-level (for example, in Extended Technicolor models[60]

Decay Mode	Experimental Limit	B_{SM}
$B_d^0 \to e^+e^-$	$< 5.9 \times 10^{-6}$ (CLEO)	2.6×10^{-15}
$B_d^0 \to \mu^+\mu^-$	$< 1.6 \times 10^{-6}$ (CDF)	1.1×10^{-10}
$B_d^0 \to \tau^+\tau^-$	—	2.1×10^{-8}
$B_s^0 \to e^+e^-$	—	5.3×10^{-14}
$B_s^0 \to \mu^+\mu^-$	$< 8.4 \times 10^{-6}$ (CDF)	$(3.6 \pm 1.8) \times 10^{-9}$
$B_s^0 \to \tau^+\tau^-$	—	5.1×10^{-7}
$B^0 \to \mu^+\mu^-$	$< 8.0 \times 10^{-6}$ (DØ)	
$B^0 \to e^\pm\mu^\mp$	$< 5.9 \times 10^{-6}$ (CLEO)	0
$B^0 \to e^\pm\tau^\mp$	$< 5.3 \times 10^{-4}$ (CLEO)	0
$B^0 \to \mu^\pm\tau^\mp$	$< 8.3 \times 10^{-4}$ (CLEO)	0
$B_d^0 \to \gamma\gamma$	$< 3.9 \times 10^{-5}$ (L3)	1.0×10^{-8}
$B_s^0 \to \gamma\gamma$	$< 1.5 \times 10^{-4}$ (L3)	3×10^{-7}
$B \to X_s + \gamma$	$(2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$ (CLEO)	$(3.25 \pm 0.30 \pm 0.40) \times 10^{-4}$
$B \to K^*\gamma$	$(4.2 \pm 0.8 \pm 0.6) \times 10^{-5}$ (CLEO)	$(4.0 \pm 2.0) \times 10^{-5}$
$B^0 \to \rho^0\gamma$	$< 3.9 \times 10^{-5}$ (CLEO)	$(0.85 \pm 0.65) \times 10^{-6}$
$B^0 \to \omega^0\gamma$	$< 1.3 \times 10^{-5}$ (CLEO)	$(0.85 \pm 0.65) \times 10^{-6}$
$B^- \to \rho^-\gamma$	1.1×10^{-5} (CLEO)	$(1.9 \pm 1.6) \times 10^{-6}$
$B \to X_s + e^+e^-$	—	$(6.25^{+1.04}_{-0.93}) \times 10^{-6}$
$B \to X_s + \mu^+\mu^-$	$< 3.6 \times 10^{-5}$ (DØ)	$(5.73^{+0.75}_{-0.78}) \times 10^{-6}$
$B \to X_s + \tau^+\tau^-$	—	$(3.24^{+0.44}_{-0.54}) \times 10^{-7}$
$B^0 \to K^0 ee/\mu\mu$	$< 1.5/2.6 \times 10^{-4}$ (CLEO)	$(5.0 \pm 3.0)/(3.0 \pm 1.8) \times 10^{-7}$
$B^- \to K^- ee/\mu\mu$	$< 1.2/0.9 \times 10^{-5}$ (CLEO)	$(5.0 \pm 3.0)/(3.0 \pm 1.8) \times 10^{-7}$
$\bar{B}^0 \to \bar{K}^{*0} ee/\mu\mu$	$< 1.6/2.5 \times 10^{-5}$ (CLEO/CDF)	$(2.0 \pm 1.0)/(1.25 \pm 0.62) \times 10^{-6}$
$\bar{B}^- \to \bar{K}^{*-} ee/\mu\mu$	$< 6.9/11 \times 10^{-4}$ (CLEO)	$(2.0 \pm 1.0)/(1.25 \pm 0.62) \times 10^{-6}$
$B^+ \to K^+ e^\pm\mu^\mp$	$< 1.2 \times 10^{-5}$ (CLEO)	0
$\bar{B}^0 \to \bar{K}^{*0} e^\pm\mu^\mp$	$< 2.7 \times 10^{-5}$ (CLEO)	0
$B \to X_s + \nu\bar{\nu}$	$< 7.7 \times 10^{-4}$ (ALEPH)	$(3.8 \pm 0.8) \times 10^{-5}$

Table 7: Standard Model predictions for the branching fractions for various rare B meson decays with $f_{B_d} = 180$ MeV. Also shown are the current experimental limits.[4,57,58]

or by virtual $H^\pm$ exchange[61] in 2HDM), and at tree-level, *e.g.*, with leptoquark exchange.[62] However, as can be seen from the table, the experimental probes of these purely leptonic decays are orders of magnitude above the expected rates, and hence only potentially large tree-level contributions can currently be tested. Indeed, the most stringent constraints on tree-level leptoquark contributions in B decays are obtained from the exclusive reaction $B \to Ke\mu$ (Ref. 62). However, in this case, there exist large uncertainties associated with the hadronic matrix elements, yielding some sloppiness in the resulting bounds.

The transition $b \to s\ell^+\ell^-$ merits further attention as it offers an excellent opportunity to search for new physics. For example, it has been found[63] that Extended Technicolor models with a GIM mechanism already violate(!) the experimental upper bound on $B \to X_s\mu\mu$, but more traditional ETC models yield a rate which is close to the SM prediction. The decay proceeds via electromagnetic and Z penguin as well as by W box diagrams, and hence can probe different coupling structures than the pure electromagnetic process $b \to s\gamma$. For $B \to X_s\ell^+\ell^-$ the Hamiltonian of the effective field theory [see Eq. (30)] is expanded to include two additional operators, $\mathcal{O}_{9,10}$. This formalism leads to the physical decay amplitude (neglecting the strange quark mass)

$$\mathcal{M}(B \to X_s\ell^+\ell^-) = \frac{\sqrt{2}G_F\alpha}{\pi} V_{tb}V_{ts}^* \Big[C_9^{eff}\, \bar{s}_L\gamma_\mu b_L \bar{\ell}\gamma^\mu\ell + C_{10}\bar{s}_L\gamma_\mu b_L\bar{\ell}\gamma^\mu\gamma_5\ell - 2C_7^{eff} m_b \bar{s}_L i\sigma_{\mu\nu}\frac{q^\nu}{q^2} b_R \bar{\ell}\gamma^\mu\ell \Big] , \quad (45)$$

where q^2 represents the momentum transferred to the lepton pair. The expressions for $C_i(M_W)$ are given by the Inami-Lim functions.[30] A NLO analysis for this decay has recently been performed,[64] where it is stressed that a scheme-independent result can only be obtained by including the leading and next-to-leading logarithmic corrections to $C_9(\mu)$ while retaining only the leading logarithms in the remaining Wilson coefficients. The residual leading μ dependence in $C_9(\mu)$ is cancelled by that contained in the matrix element of $\mathcal{O}_9$. The combination yields an effective value of C_9 given by

$$C_9^{eff}(\hat{s}) = C_9(\mu)\eta(\hat{s}) + Y(\hat{s}) , \quad (46)$$

with $Y(\hat{s})$ being the one-loop matrix element of $\mathcal{O}_9$, $\eta(\hat{s})$ represents the single gluon corrections to this matrix element, and $\hat{s} \equiv q^2/m_b^2$ is the scaled momentum transferred to the lepton pair. The effective value for $C_7^{eff}(\mu)$ refers to the leading order scheme-independent result obtained by Buras *et al.* [34] The corresponding formulae for $C_i(\mu)$, $Y(\hat{s})$, and $\eta(\hat{s})$ are collected in Refs. 43 and 64. The operator $\mathcal{O}_{10}$ does not renormalize, and hence its corresponding coefficient does not depend on the value of μ (except for the μ dependence associated with the definition of the top-quark mass). The numerical estimates [in the naive dimensional regularization (NDR) scheme] for these coefficients are then (taking $m_b^{pole} = 4.87$ GeV, $m_t^{phys} = 175$ GeV, and $\alpha_s(M_Z) = 0.118$)

$$C_7^{eff}(\mu = m_b {}^{-m_b/2}_{+m_b}) = -0.312^{-0.059}_{+0.034} ,$$
$$C_9(\mu = m_b {}^{-m_b/2}_{+m_b}) = 4.21^{+0.31}_{-0.40} , \quad (47)$$

and

$$C_{10}(\mu) = -4.55 . \quad (48)$$

The reduced scale dependence of the NLO versus the LO corrected coefficients is reflected in the deviations $\Delta C_9(\mu) \lesssim \pm 10\%$ and $\Delta C_7^{eff}(\mu) \approx \pm 20\%$ as μ is varied in the range $m_b/2 \le \mu \le 2m_b$. We find that the coefficients are much less sensitive to the values of the remaining input parameters, with $\Delta C_9(m_b), \Delta C_7^{eff}(m_b) \lesssim 3\%$, varying $\alpha_s(M_Z) = 0.118 \pm 0.003$ (Refs. 4 and 41), and $m_t^{phys} = 175 \pm 6$ GeV (Ref. 13) corresponding to $m_t(m_t) = 166 \pm 6$ GeV. The resulting inclusive branching fractions (which are computed by scaling the width for $B \to X_s\ell^+\ell^-$ to that for B semileptonic decay) are found to be $(6.25^{+1.04}_{-0.93}) \times 10^{-6}$, $(5.73^{+0.75}_{-0.78}) \times 10^{-6}$, and $(3.24^{+0.44}_{-0.54}) \times 10^{-7}$ for $\ell = e, \mu$, and τ, respectively, taking into account the above input parameter ranges, as well as $B_{sl} \equiv B(B \to X\ell\nu) = (10.23 \pm 0.39)\%$ (Ref. 19), and $m_c/m_b = 0.29 \pm 0.02$ (Refs. 36 and 4). There are also long-distance resonance contributions to $B \to X_s\ell^+\ell^-$, arising from $B \to K^{(*)}\psi^{(\prime)} \to K^{(*)}\ell^+\ell^-$. These appear as an effective $(\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma_\mu\ell)$ interaction and are incorporated into C_9^{eff} via the modification $Y(\hat{s}) \to Y'(\hat{s}) \equiv Y(\hat{s}) + Y_{res}(\hat{s})$, where $Y_{res}(\hat{s})$ is given in Ref. 65. These pole contributions lead to a significant interference between the dispersive part of the resonance and the short-distance contributions. However, suitable cuts on the lepton pair mass spectrum can cleanly separate the short-distance physics from the resonance contributions.

Various kinematic distributions associated with the final state lepton pair render $B \to X_s\ell^+\ell^-$ an excellent SM testing ground. These distributions include the lepton pair invariant mass distribution,[66] the lepton pair forward-backward asymmetry,[67] and the tau polarization asymmetry[68] in the case $\ell = \tau$. They

are presented in Fig. 5, with and without the resonance contributions. Note that both asymmetries are large. As an example of how new physics can affect these distributions, we display in Fig. 5(d) the tau polarization asymmetry for various changes of sign of the contributing Wilson coefficients. Measurement of all three kinematic distributions as well as the rate for $B \to X_s\gamma$ would allow for the determination of the sign and magnitude of all the Wilson coefficients for the contributing operators and thus provide a completely model-independent analysis. A 95% C.L. Monte Carlo fit to these coefficients has been performed[42] in order to ascertain how much quantitative information will be obtainable at future B factories. In this fit, "data" has been generated assuming the SM is realized, and by dividing the lepton pair invariant mass spectrum into bins, where six of the bins are taken to be in the low dilepton invariant mass region below the J/ψ resonance, and three of the bins above the ψ' pole. The "data" has been statistically fluctuated using a normalized Gaussian distributed random number procedure. The errors in each bin are expected to be statistics dominated. However, for $B \to X_s\gamma$, the statistical precision will eclipse the possible systematic and theoretical accuracy, and a flat 10% error in the determination of the branching fraction is thus assumed. A three-dimensional χ^2 fit to the coefficients $C_{7,9,10}(\mu)$ is performed for three values of the integrated luminosity, 3×10^7, 10^8, and 5×10^8 $B\bar{B}$ pairs, corresponding to one year at e^+e^- B factory design luminosity, one year at an upgraded accelerator, and the total accumulated luminosity at the end of the program. Hadron colliders will, of course, also contribute to this program, but it is more difficult to assess their potential systematic and statistical weights without further study.

The 95% C.L. allowed regions as projected onto the $C_9(\mu) - C_{10}(\mu)$ and $C_7^{eff}(\mu) - C_{10}(\mu)$ planes are depicted in Figs. 6(a) and 6(b), where the diamond represents the expectations in the SM. We see that the determinations are relatively poor for 3×10^7 $B\bar{B}$ pairs and that higher statistics are required in order to focus on regions centered around the SM. Clearly, C_9 and C_{10} are highly correlated, whereas C_7^{eff} and C_{10} are not. We see that the sign, as well as the magnitude, of all the coefficients including C_7^{eff} can now be determined.

Supersymmetric contributions to $B \to X_s\ell^+\ell^-$ have recently been analyzed in Refs. 42 and 69. In Fig. 7, the correlation between R_9 and R_{10} (recall $R_i \equiv \frac{C_i^{susy}(M_W)}{C_i^{SM}(M_W)} - 1$) is displayed using the same supersymmetric parameter space as in

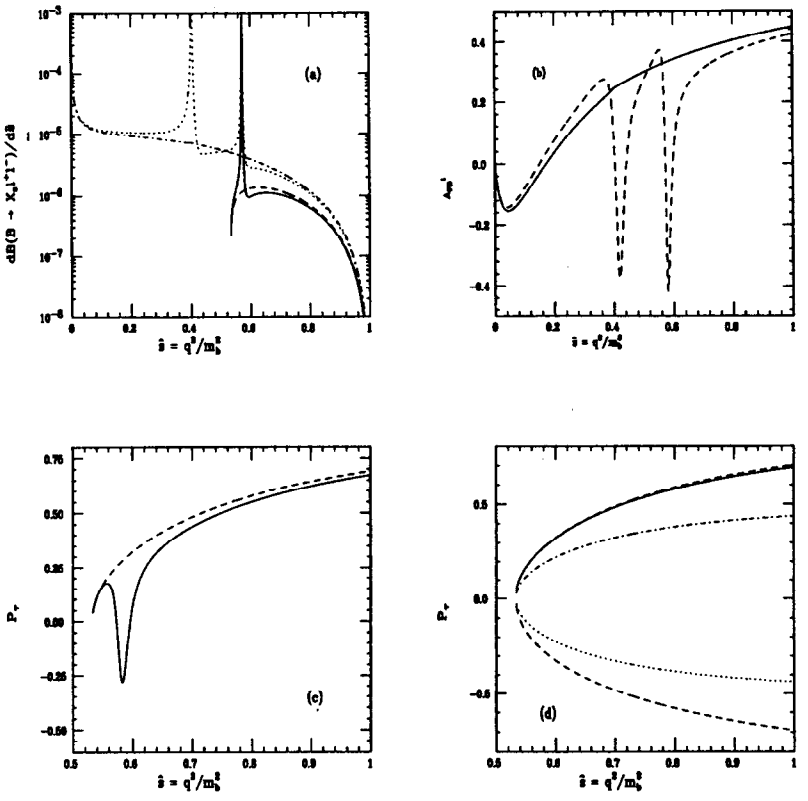


Figure 5: (a) Differential branching fraction, (b) lepton pair forward-backward asymmetry, and (c) tau polarization asymmetry as a function of $\hat{s}$ for $\ell = \tau$ (solid and dashed curves) and $\ell = e$ (dotted and dash-dotted curves), with and without the long-distance contributions. (d) Tau polarization asymmetry with changes in sign of the Wilson coefficients at the electroweak scale, corresponding to $C_{10}, C_9, C_{9,10}, SM$, and $C_{7,8}$ from bottom to top.

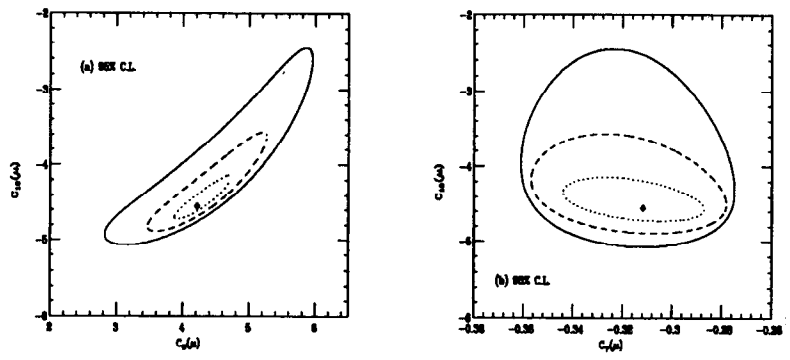


Figure 6: The 95% C.L. projections in the (a) $C_9 - C_{10}$ and (b) $C_7^{eff} - C_{10}$ planes, where the allowed regions lie inside of the contours. The solid, dashed, and dotted contours correspond to 3×10^7, 10^8, and 5×10^8 $B\bar{B}$ pairs. The SM prediction is labeled by the diamond.

Fig. 4(c). We see that R_9 is always positive since the charged Higgs boson and chargino contributions always add constructively. We see that the values of R_9 and R_{10} are bounded by about 0.04, a small number compared to the range found for R_7 in Fig. 4(c), and rendering the minimal supergravity contributions to $R_{9,10}$ essentially unobservable. The solid lines in this figure correspond to the 95% C.L. bounds obtainable with very high integrated luminosity (5×10^8 $B\bar{B}$ pairs) at B factories from the global fit shown above.

4.3 Other Rare D Decays

While investigations of the K and B systems have and will continue to play a central role in our quest to understand flavor physics, in-depth examinations of the charm-quark sector have yet to be performed, leaving a gap in our knowledge. Since charm is the only heavy charged +2/3 quark presently accessible to experiment in copious amounts, it provides the sole window of opportunity to examine flavor physics in this sector. In addition, charm allows a complementary probe of SM physics (and beyond) to that attainable from the down-quark sector.

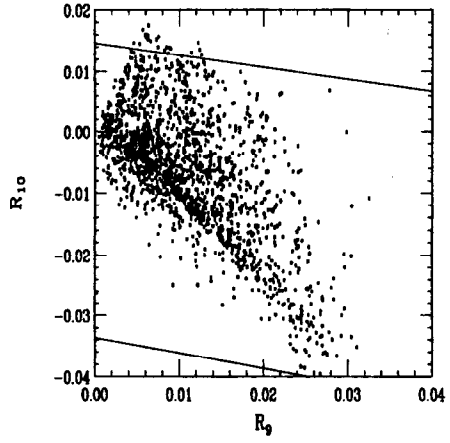


Figure 7: Parameter space scatter plot of R_9 vs R_{10} in minimal supergravity model. The global fit to the coefficients obtained in Fig. 6 with 5×10^8 $B\bar{B}$ pairs corresponds to the region inside the diagonal bands.

Due to the effectiveness of the GIM mechanism, short-distance SM contributions to rare charm processes are very small. Most reactions are thus dominated by long-range effects which are difficult to reliably calculate. However, for some interactions, there exists a window for the potential observation of new physics. In fact, it is precisely because the SM flavor changing neutral current rates are so small that charm provides an untapped opportunity to discover new effects and offers a detailed test of the SM in the up-quark sector.

FCNC decays of the D meson include the processes $D^0 \to \ell^+\ell^-, \gamma\gamma$, and $D \to X + \ell^+\ell^-, X + \nu\bar{\nu}, X + \gamma$, with $\ell = e, \mu$, and with the radiative decays being discussed above. The calculation of the SM short-distance rates for these processes is straightforward and the transition amplitudes and standard loop integrals, which are categorized in Ref. 30 for rare K decays, are easily converted to the D system. The loop integrals relevant for $D^0 \to \gamma\gamma$ may be found in Ref. 70. Employing the GIM mechanism results in a general expression for the loop integrals which can be written as

$$A = V_{cs}V_{us}^*[F(x_s) - F(x_d)] + V_{cb}V_{ub}^*[F(x_b) - F(x_d)], \qquad (49)$$

with $x_i \equiv m_i^2/M_W^2$ and $F(x_d)$ usually being neglected (except in the 2γ case). The s- and b-quark contributions are roughly equal as the larger CKM factors compensate for the small strange quark mass. The values of the resulting inclusive short-distance branching fractions are shown in Table 8, along with the current experimental bounds.[4,71] The corresponding exclusive rates are typically an order of magnitude less than the inclusive case. We note that the transition $D^0 \to \ell^+\ell^-$ is helicity suppressed and hence has the smallest branching fraction. The range given for this branching fraction, $(1 - 20) \times 10^{-19}$, indicates the effect of varying the parameters in the ranges $f_D = 0.15 - 0.25$ GeV and $m_s = 0.15 - 0.40$ GeV. It is clear that the typical branching fraction is indeed much smaller than that in the B meson system, illustrating the effectiveness of the GIM mechanism when there is no heavy top quark contributing inside the loop.

The calculation of the long-distance branching fractions are plagued with the usual hadronic uncertainties, and the estimates listed in the table convey an upper limit on the size of these effects rather than an actual value. These estimates have been computed by considering various intermediate particle states (*e.g.*, $\pi, K, \bar{K}, \eta, \eta', \pi\pi$, or $K\bar{K}$) and inserting the known rates for the decay of the intermediate particles into the final state of interest. In all cases, we see

Decay Mode	Experimental Limit	$B_{S.D.}$	$B_{L.D.}$
$D^0 \to \mu^+\mu^-$	$< 7.6 \times 10^{-6}$ (WA92)	$(1 - 20) \times 10^{-19}$	$< 3 \times 10^{-15}$
$D^0 \to e^+e^-$	$< 1.3 \times 10^{-5}$ (CLEO)		
$D^0 \to \mu^+e^\mp$	$< 1.9 \times 10^{-5}$ (CLEO)	0	0
$D^0 \to \gamma\gamma$	—	10^{-16}	$< 3 \times 10^{-9}$
$D \to X_u + \gamma$		1.4×10^{-17}	
$D^0 \to \rho^0\gamma$	$< 1.4 \times 10^{-4}$ (CLEO-prelim.)		$< 2 \times 10^{-5}$
$D^0 \to \phi^0\gamma$	$< 2.0 \times 10^{-4}$ (CLEO-prelim.)		$< 10^{-4}$
$D^+ \to \rho^+\gamma$	—		$< 2 \times 10^{-4}$
$D^+ \to \bar{K}^{*+}\gamma$	—		3×10^{-7}
$D^0 \to \bar{K}^{*0}\gamma$	—		1.6×10^{-4}
$D \to X_u + \ell^+\ell^-$		4×10^{-9}	
$D^0 \to \pi^0 ee/\mu\mu$	$< 4.5/18 \times 10^{-5}$ (CLEO/E653)		
$D^0 \to \bar{K}^0 ee/\mu\mu$	$< 1.1/2.6 \times 10^{-4}$ (CLEO/E653)		$< 2 \times 10^{-15}$
$D^0 \to \rho^0 ee/\mu\mu$	$< 1.0/2.3 \times 10^{-4}$ (CLEO/E653)		
$D^0 \to \eta ee/\mu\mu$	$< 1.1/5.3 \times 10^{-4}$ (CLEO)		
$D^+ \to \pi^+ ee/\mu\mu$	$< 6.6/1.8 \times 10^{-5}$ (CLEO/E653)	few$\times 10^{-10}$	$< 10^{-8}$
$D^+ \to K^+ ee/\mu\mu$	$< 480/3.2 \times 10^{-5}$ (MRK2/E653)		$< 10^{-15}$
$D^+ \to \rho^+\mu\mu$	$< 5.6 \times 10^{-4}$ (E653)		
$D^0 \to X_u + \nu\bar{\nu}$		2.0×10^{-15}	
$D^0 \to \pi^0\nu\bar{\nu}$	—	4.9×10^{-16}	$< 6 \times 10^{-16}$
$D^0 \to \bar{K}^0\nu\bar{\nu}$	—		$< 10^{-12}$
$D^+ \to X_u + \nu\bar{\nu}$	—	4.5×10^{-15}	
$D^+ \to \pi^+\nu\bar{\nu}$	—	3.9×10^{-16}	$< 8 \times 10^{-16}$
$D^+ \to K^+\nu\bar{\nu}$	—		$< 10^{-14}$

Table 8: Standard Model predictions for the branching fractions due to short- and long-distance contributions for various rare D meson decays. Also shown are the 90% C.L. current experimental limits.[71]

that the long-distance contributions overwhelm those from the SM short-distance physics.

Lepton flavor violating decays, *e.g.*, $D^0 \to \mu^\pm e^\mp$ and $D \to X + \mu^\pm e^\mp$, are strictly forbidden in the SM with massless neutrinos. In a model with massive nondegenerate neutrinos and nonvanishing neutrino mixings, such as in four-generation models, $D^0 \to \mu^\pm e^\mp$ would be mediated by box diagrams with the massive neutrinos being exchanged internally. LEP data restricts[72] heavy neutrino mixing with e and μ to be $|U_{Ne}U^*_{N\mu}|^2 < 7 \times 10^{-6}$ for a neutrino with mass $m_N > 45$ GeV. Consistency with this bound constrains the branching fraction to be $B(D^0 \to \mu^\pm e^\mp) < 6 \times 10^{-22}$. This same result also holds for a heavy singlet neutrino which is not accompanied by a charged lepton. The observation of this decay at a larger rate than the above bound would be a clear signal for the existence of a different class of models with new physics.

Examining Table 7, we see that there is a large window of opportunity to discover the existence of new physics in rare charm decays. Although the SM short-distance contributions are completely dominated by the long-distance effects, there are some modes where the size of the two contributions are not that far apart. The observation of any of these modes at a larger rate than what is predicted from long-distance interactions would provide a clear signal for new physics.

4.4 Rare Decays in the Kaon System

The SM level for the theoretically clean decay $K^+ \to \pi^+\nu\bar{\nu}$ should be reached in the next decade, with the present bound[73] on the branching fraction being $B(K^+ \to \pi^+\nu\bar{\nu}) < 2.4 \times 10^{-9}$ from E787 at Brookhaven. This transition is theoretically clean as it is short-distance dominated;[74] the relevant hadronic operator is extracted from $K^+ \to \pi^0 e^+ \nu$, and the next-to-leading order QCD corrections are fully known.[75] The SM processes responsible for this decay are Z-mediated penguin graphs and W box diagrams with both charm and top quarks contributing internally. The full NLO expressions for this decay are given in Ref. 75. The impact of the NLO corrections are to reduce the scale uncertainties from $\pm 22\%$ to $\pm 7\%$. Here we present the approximate result recently given by Buras,[12]

$$\begin{aligned} B(K^+ \to \pi^+\nu\bar{\nu}) &= 0.7 \times 10^{-10} \left[\left(\frac{|V_{td}|}{0.01} \right)^2 \left(\frac{|V_{cb}|}{0.04} \right)^2 \left(\frac{m_t(m_t)}{170\,\mathrm{GeV}} \right)^{2.3} + cc + tc \right], \\ &= (9.1 \pm 3.2) \times 10^{-11}, \end{aligned} \tag{50}$$

where the cc and tc terms represent the pure charm and charm-top contributions, respectively. Measurement of this rare decay would provide a sensitive and direct determination of V_{td}. The theoretical error[12,75] on an evaluation of V_{td} from this channel is at the $\pm 4\%$ level. Hence, this mode represents the most promising technique of determining V_{td}.

An enhancement over the SM rate would clearly signal new physics although such enhancements are not expected in most minimal extensions of the SM once the constraints from $B - \bar{B}$ mixing, ϵ_K, and $b \to s\gamma$ are taken into account.[76] These processes are to a large extent governed by the same parameters, limiting the impact of new physics in this case. A possible exception concerns the MSSM with SUSY particles in the 100 GeV range where there can be some enhancement.[77] There remains the possibility of large enhancements in SUSY models with broken R-parity, models with family symmetry producing a new type of neutrino, as well as certain leptoquark models.[76] Typically, these models are more weakly constrained overall and could also lead to nonstandard signals in other rare processes (for example, B or D decays). The Three-Higgs-Doublet model also can lead to a moderate enhancement (by a factor of three) of the standard rate for this decay.[78] This is to be contrasted with the 2HDM where the existing constraints preclude any significant effect in future kaon decay measurements.[46]

The process $K_L \to \mu^+\mu^-$ shares several features of the preceding one as far as sensitivity to new physics is concerned. However, the bounds obtained are not as reliable due to large and uncertain long-distance contributions. One interesting aspect of this process is the sensitivity to other sources of CP violation in the measurement of the longitudinal polarization of the muon, which is expected to be $P_L \approx 2 \times 10^{-3}$ in the SM.

Extensive discussions of other rare K decay modes can be found in Ref. 79.

4.5 Rare Decays of the Top Quark

Loop-induced flavor changing top quark decays are small in the SM, as in the charm-quark system, due to the effectiveness of the GIM mechanism and the

small masses of the $Q = -1/3$ quarks. However, these transitions are anticipated to be theoretically clean as long-distance effects are expected to be negligible. The SM rates for $t \to c\gamma, cZ, cg$ are given by 4.9×10^{-13}, 1.4×10^{-13}, 4.4×10^{-11}, respectively, for $m_t = 180$ GeV (Ref. 80). The branching fraction for $t \to ch$ as a function of the Higgs mass is represented by the solid curve in Figs. 8(a) and 8(b). We see that this rate is also tiny, being in the 10^{-13} range over the entire kinematically allowed region for the Higgs mass. Loop contributions from new physics have been studied in 2HDM[80,81] and in SUSY,[82] and generally can enhance these transition rates by three to four orders of magnitude for some regions of the parameter space. The effects of virtual $H^{\pm}$ exchange in 2HDM of Type II on the reactions $t \to cV$, $V = \gamma, Z, g$, are displayed in Fig. 8(c) for $m_t = 180$ GeV. We see that, indeed, enhancements are present for large values of $\tan\beta$. We also examine the decays $t \to ch, cH$ in Model II, where h and H respectively represent the lightest and heaviest physical neutral scalars present in 2HDM. The resulting rates are depicted in Figs. 8(a) and 8(b) for the demonstrative case of $m_{H^{\pm}} = 600$ GeV and $\tan\beta = 2(30)$, corresponding to the dashed (solid) curves. Here, we have made use of the SUSY Higgs mass relationships in order to reduce the number of free parameters. We note that the effects of super-partner virtual exchange should also be included (with, of course, a corresponding increase in the number of parameters!). We have also studied these modes in Model I and found similar rate increases for regions of the parameter space. Even if new physics were to produce such enhancements, the resulting branching fractions would still lie below the observable level in future experiments at an upgraded Tevatron, the LHC, or the NLC.

On the other hand, if these FCNC decays were to be detected, they would provide an indisputable signal for new physics. Hence, a model-independent approach in probing anomalous FCNC top-quark couplings has recently been taken by a number of authors.[83] By parameterizing the general tcV vertex in a manner similar to that presented in Eq. (6), and performing a Monte Carlo study of the signal rate versus potential backgrounds, Han *et al.*[83] have found that such anomalous couplings can be probed down to the level of $\kappa_{\gamma,Z} \equiv \sqrt{g_L^2 + g_R^2}|_{\gamma,Z} \simeq 0.1(0.01)$ at the Tevatron (LHC). This corresponds to values of the branching fractions for $t \to cZ$, $c\gamma$ at the level of few$\times 10^{-3}$ for the Tevatron bounds and 10^{-4} for the LHC. CDF has, in fact, already performed a search for these FCNC decays from their present top sample, and has placed the bounds[84] $B(t \to c\gamma + u\gamma) < 2.9\%$ and $B(t \to cZ + uZ) < 90\%$ at 95% and 90% C.L., respectively.

Potential non-SM tree-level decays of the top quark could feasibly occur at measurable rates in future colliders. Examples of these possible transitions are: (i) the decay of top into a charged Higgs, $t \to bH^+$ in multi-Higgs models,[85] (ii) the tree-level flavor-changing decay $t \to ch$, which can occur, if kinematically accessible, in multi-Higgs models without natural flavor conservation,[86,87] (iii) $t \to \tilde{t}\tilde{\chi}^0$ which can take place in supersymmetry if the top-squark is sufficiently light[88] (this possibility is related to the large value of the top Yukawa coupling, and is thus special to the top system), and (iv) $t \to \tilde{\ell}^+ d$ in SUSY models with R-parity violation.[89] For favorable values of the parameters, each of these modes could be competitive with the SM decay $t \to bW^+$. The observation of the top quark by CDF and DØ, which relies heavily on the expected signal from SM top decay,[13] can thus restrict the values of the branching fractions for these potential new modes. The possible constraints that could be obtained on the models which would allow the decays (i) $t \to bH^+$ and (ii) $t \to ch$ to occur, if these collaborations were to make the statement that the observed $t\bar{t}$ production rate is 50–90% of that expected in the SM, are given in Fig. 9. We have examined the case of the decay into a $H^{\pm}$ in Model II, taking $m_t = 180$ GeV, and find that the potentially excluded regions lie below the curves. Clearly, large regions of the parameter space have the potential to be ruled out. In the case of $t \to ch$ decay, we have parameterized the tree-level tch coupling as $(\sqrt{2}G_F)^{1/2} m_t(\alpha - \beta\gamma_5)$ and displayed the restrictions in the $k \equiv \sqrt{\alpha^2 + \beta^2} - m_h$ plane. The region above the curves would be excluded.

5 Neutral Meson Mixing

For a neutral meson system, the mass and lifetime eigenstates, P_L, P_H, with masses $m_{L,H}$ and widths $\Gamma_{L,H}$, are conventionally defined as mixtures of the two weak CP-conjugate eigenstates $P^0, \bar{P}^0$ as

$$\begin{aligned} |P_L\rangle &= p|P^0\rangle + q|\bar{P}^0\rangle, \\ |P_H\rangle &= p|P^0\rangle - q|\bar{P}^0\rangle, \end{aligned} \tag{51}$$

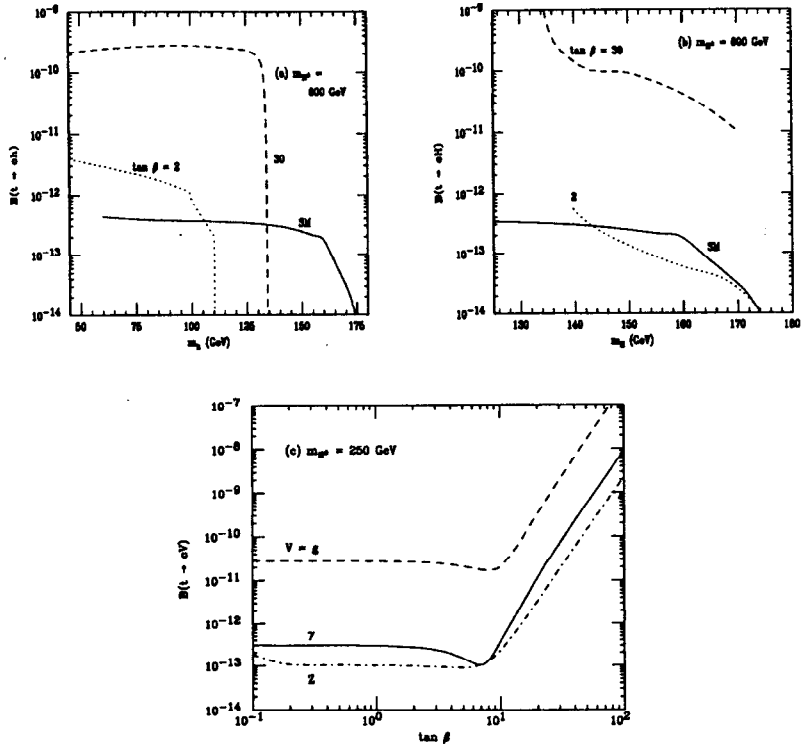


Figure 8: Branching fractions for (a) $t \to ch$ and (b) $t \to cH$ as a function of the neutral Higgs mass in 2HDM of Type II. The SM rate is represented by the solid curve. (c) $B(t \to cV)$ where $V = g, \gamma, Z$ as a function of $\tan\beta$ in Model II. In all cases, the top-quark mass is taken to be 180 GeV.

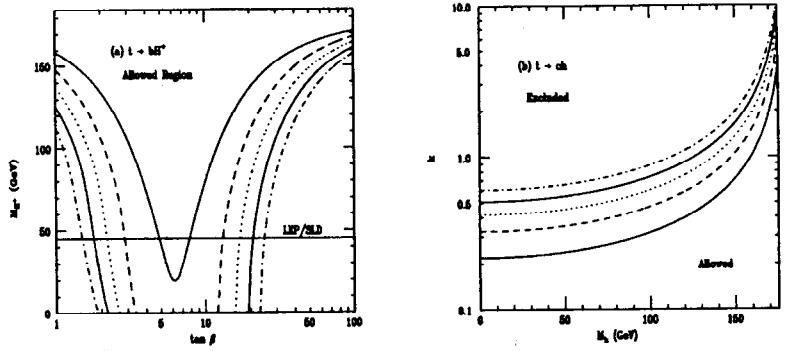


Figure 9: Constraints placed on the nonstandard decays (a) $t \to bH^+$ and (b) $t \to ch$ from demanding that the observed event rate for top-quark pair production is at least 50, 60, 70, 80, and 90% of that expected in the SM, corresponding to the dashed-dot, solid, dotted, dashed, and solid curves. $m_t = 180$ GeV is assumed.

with the normalization $|p|^2 + |q|^2 = 1$, and the subscripts L and H denoting the light and heavy states, respectively. Here, P^0 generically represents the pseudoscalar neutral meson systems K^0, D^0, B_d^0, and B_s^0. Note that there is no top-meson mixing as the top-quark decays too rapidly to form neutral meson bound states. There is also the equivalent definition

$$\begin{aligned} |P_L\rangle &= [(1+\epsilon_P)|P^0\rangle + (1-\epsilon_P)|\bar{P}^0\rangle]/\sqrt{2(1+|\epsilon_P|^2)}\,, \\ |P_H\rangle &= [(1+\epsilon_P)|P^0\rangle - (1-\epsilon_P)|\bar{P}^0\rangle]/\sqrt{2(1+|\epsilon_P|^2)}\,. \end{aligned} \tag{52}$$

The mixing parameters are related by $q/p = (1-\epsilon_P)/(1+\epsilon_P)$. In the limit of CP invariance, $|p|^2 = |q|^2$, $\mathcal{R}e\,\epsilon_P = 0$, and a phase convention can be found such that $\mathcal{I}m\,\epsilon_P = 0$ and $p = q = 1/\sqrt{2}$. Throughout our discussion, we will assume CPT invariance.

The Hamiltonian eigenvalue equation

$$\begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \begin{pmatrix} p \\ \pm q \end{pmatrix} = \left(M_{L,H} - \frac{i}{2}\Gamma_{L,H}\right) \begin{pmatrix} p \\ \pm q \end{pmatrix} \tag{53}$$

	K^0	D^0	B_d^0	B_s^0
x	0.476	< 0.083	0.72 ± 0.03	> 13.8

Table 9: Experimental measurements and constraints[4,12] on the parameter $x \equiv \Delta M/\Gamma$ for the various meson systems.

has the solution

$$\frac{q}{p} = \left[\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right]^{\frac{1}{2}} = \frac{-2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta M - \frac{i}{2}\Delta\Gamma}, \tag{54}$$

where $\Delta M = M_H - M_L = 2|M_{12}|$ and $\Delta\Gamma = \Gamma_H - \Gamma_L = 2|\Gamma_{12}|$. M_{12} describes transitions between P^0 and $\bar{P}^0$ via virtual states, and Γ_{12} represents contributions to decay channels which are common to both P^0 and $\bar{P}^0$. The parameter $x \equiv \Delta M/\Gamma$ is often used to describe the competition between the $P^0 - \bar{P}^0$ mixing and decay. The experimental measurements and constraints on x for the various meson systems are listed in Table 9.

The proper time evolution of an initially pure P^0 or $\bar{P}^0$ state is

$$\begin{aligned} |P^0(t)_{phys}\rangle &= e^{-\Gamma t/2}e^{-iMt}\left[\cos\left(\frac{\Delta Mt}{2}\right)|P^0\rangle \right. \\ &\quad \left. + i\frac{q}{p}\sin\left(\frac{\Delta Mt}{2}\right)|\bar{P}^0\rangle\right], \\ |\bar{P}^0(t)_{phys}\rangle &= e^{-\Gamma t/2}e^{-iMt}\left[i\frac{p}{q}\sin\left(\frac{\Delta Mt}{2}\right)|P^0\rangle \right. \\ &\quad \left. + \cos\left(\frac{\Delta Mt}{2}\right)|\bar{P}^0\rangle\right], \end{aligned} \tag{55}$$

where M is defined as $M \equiv (M_H + M_L)/2$. In systems where $\Delta\Gamma$ can be neglected, the probability of mixing can then be written as

$$\mathcal{P}(t) = \frac{1}{2}e^{-\Gamma t}\left[1 - \cos\left(\Delta Mt\right)\right]. \tag{56}$$

Time-dependent measurements of mixing in the B_d system have only recently been performed at SLD, LEP, and the Tevatron with new vertexing technology,[12] and provide a direct determination of ΔM. Previous results relied on the time-integrated mixing parameter

$$\chi = \int_0^\infty \mathcal{P}(t)dt = \frac{x^2}{2(1+x^2)}, \tag{57}$$

which is bounded to be $\chi \leq 0.5$, and determines the parameter x. $\Upsilon(4S)$ experiments measure the pure time integrated χ_d parameter of the B_d system, while high-energy experiments off the $\Upsilon(4S)$ measure the mixture

$$\chi_B = f_d\chi_d + f_s\chi_s, \tag{58}$$

where $f_{d,s}$ represent the fractions of produced b-quark hadrons that are B_d^0 and B_s^0, respectively. The values of these hadronization fractions are not precisely known, they are approximately $f_d \sim 0.39$ and $f_s \sim 0.12$ at SLC/LEPI energies, and hence introduce a source of uncertainty to the time-integrated mixing measurements. If χ approaches its upper value of 0.5, as is expected for the B_s system, it clearly does not provide a good determination of x, and one must then rely on the time-dependent approach.

As in the case of rare decays, both short- and long-distance physics processes contribute to meson mixing within the SM. The short-distance contributions are mediated by box diagrams with internal quark and W-boson exchange and are calculated via the operator product expansion in Eq. 24. The $\Delta Q = 2$ effective Lagrangian for a pseudoscalar meson $P = Q\bar{q}$ is

$$\mathcal{L}_{eff}(\Delta Q = 2) = \frac{G_F^2 M_W^2}{16\pi^2}\sum_{i,j} V_{q_i'Q}V_{q_i'q}^* V_{q_j'Q}V_{q_j'q}^* \eta_{ij} S\left(\frac{m_{q_i'}^2}{M_W^2}, \frac{m_{q_j'}^2}{M_W^2}\right)\mathcal{O}_{LL}, \tag{59}$$

where the sum extends over the two contributing internal quarks $q_{i,j}'$, η summarizes the QCD corrections, and S represents the Inami-Lim functions[30] from the evaluation of the box diagrams. Note that the GIM mechanism may be employed here as well, and hence we would expect sizable short-distance mixing in cases where the top quark contributes internally, such as in the K^0, B_d^0, and B_s^0 systems. The matrix element of the $\Delta Q = 2$ operator can be evaluated as

$$\begin{aligned} \langle\bar{P}^0|\mathcal{O}_{LL}|P^0\rangle &= \langle\bar{P}^0|\bar{q}\gamma_\mu(1-\gamma_5)Q\bar{q}\gamma^\mu(1-\gamma_5)Q|P^0\rangle, \\ &= \frac{4}{3}f_P^2 B_P m_P, \end{aligned} \tag{60}$$

with f_P (m_P) being the pseudoscalar meson decay constant (mass), which is measured in the purely leptonic decays as discussed above, and B_P being the so-called bag factor which represents the nonperturbative factors associated with the hadronic matrix element and comprises the major source of theoretical uncertainty in the calculation of meson mixing. These bag factors are estimated using

nonperturbative techniques such as lattice gauge theory, QCD sum rules, $1/N$ expansion, or chiral perturbation theory, with the lattice gauge results giving the most accurate evaluations.[21]

The long-distance contributions may be generally represented as the sum of common intermediate states I, which interact with the pseudoscalar mesons via an effective weak Hamiltonian $\mathcal{H}_{eff}$

$$\langle \bar{P}^0|\mathcal{A}|P^0\rangle = \sum_I \frac{\langle \bar{P}^0|\mathcal{H}_{eff}|I\rangle\langle I|\mathcal{H}_{eff}|P^0\rangle}{m_P^2 - m_I^2 + i\epsilon}. \tag{61}$$

The dominant classes of contributions of this type arise from (i) single particle intermediate states, called pole contributions, and (ii) two particle intermediate states, denoted as dispersive contributions. Due to the effectiveness of the GIM mechanism in reducing the size of the short-distance effects in the charm system, one expects long-distance processes to dominate $D^0 - \bar{D}^0$ mixing.

5.1 $K^0 - \bar{K}^0$ Mixing

The neutral kaon system provides a special laboratory for the study of mixing. The dominant CP conserving decays of the two physical states are $K_L \to 3\pi$ and $K_S \to 2\pi$. Due to the strong phase space suppression for the K_L decay, K_L and K_S have very different lifetimes, providing a clean separation of these two eigenmodes in the laboratory. The SM short-distance contributions to the $K_L - K_S$ mass difference arises from top- and charm-quark contributions to the W box diagram, giving

$$\begin{aligned} M_{12}(\Delta S = 2) &= \frac{G_F^2 M_W^2 f_K^2 B_K m_K}{12\pi^2}\left[(V_{cs}V_{cd}^*)^2\eta_c S(x_t) + (V_{ts}V_{td}^*)^2\eta_t S(x_t)\right. \\ &\left. +2V_{cs}V_{cd}^*V_{ts}V_{td}^*\eta_{ct}S(x_c,x_t)\right]. \end{aligned} \tag{62}$$

Here, $S(x)$ represents the Inami-Lim functions,[30] $x_i \equiv m_i^2/M_W^2$, and the η_i correspond to the QCD correction factors which have been computed[90,91] to NLO for each contribution, with their numerical values being $\eta_c = 1.38 \pm 0.20$, $\eta_{c,t} = 0.47 \pm 0.04$, and $\eta_t = 0.57 \pm 0.01$. The hadronic matrix element (or bag factor), B_K, represents a large uncertainty in the computation of M_{12}, with the results from various approaches being summarized in Table 10. Buras[12] advocates use of the value $B_K = 0.75 \pm 0.15$. The mass difference is then $\Delta M_K = 2|M_{12}|$.

Approach	B_K
Lattice	0.90 ± 0.06
$1/N$	0.70 ± 0.10
QCD Sum Rules	~ 0.60
Chiral Quark Model	0.87 ± 0.25
QCD Hadronic Duality	0.39 ± 0.10
SU(3) Symmetry	1/3

Table 10: Compilation of various determinations[21,92] of B_K.

The calculation of ΔM_K is unfortunately plagued with uncertainties from the potentially sizable long-distance contributions.[93] Even so, the $K_L - K_S$ mass difference has played a strong and historical role in constraining new physics. For example, the strongest bound[94] (albeit assumption dependent) on the mass of a right-handed W boson in the Left-Right Symmetric Model of $M_{W_R} \gtrsim 1.6$ TeV, the requirement of near degeneracy of the squark masses in supersymmetry,[95] and severe constraints on technicolor model building[96] such as the introduction of the Techni-GIM mechanism, are all obtained from $K^0 - \bar{K}^0$ mixing.

5.2 $B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$ Mixing

The quark level process which is dominantly responsible for $B^0 - \bar{B}^0$ mixing in the SM is that of top-quark exchange in a W box diagram. The mass difference for B_d meson mixing is then given by

$$\Delta M_d = \frac{G_F^2 M_W^2 m_B}{6\pi^2} f_{B_d}^2 B_{B_d}\eta_{B_d}|V_{tb}V_{td}^*|^2 F(m_t^2/M_W^2), \tag{63}$$

with $\eta_{B_d} = 0.55 \pm 0.01$ being the QCD correction factor which is calculated to NLO,[91] and F(x) being the usual Inami-Lim function.[30] For consistency with the NLO QCD calculations, the running top-quark mass evaluated at m_t should be used. An equivalent expression for B_s mixing is obtained with $d \to s$. This yields the SM values of $\Delta M_d = (3.0^{+9.0}_{-2.7}) \times 10^{-13}$ GeV and $\Delta M_s = (7.4^{+8.6}_{-4.3}) \times 10^{-12}$ GeV, where the ranges correspond to taking $m_t^{phys} = 175 \pm 6$ GeV, $|V_{td}| = 0.009 \pm 0.005$ and $|V_{ts}| = 0.040 \pm 0.006$ as given in Ref. 4, and $f_{B_d} = 175 \pm 25$ MeV, $B_{B_d} = 1.31 \pm 0.03$, (the combined quantity is quoted to be $f_{B_d}\sqrt{B_{B_d}} = 207 \pm 30$ MeV) $f_{B_s} = 200 \pm 25$ MeV, and $B_{B_s} = (1.01 \pm 0.01)B_{B_d}$ as suggested by a global summary

of lattice gauge theory results.[21] This agrees well with the experimental bounds[12] of $\Delta M_d = (0.464 \pm 0.012 \pm 0.013)$ ps^{-1} and $\Delta M_s > 9.2$ ps^{-1}, corresponding to the x parameter values in Table 9. This situation is summarized in Fig. 10.

$B_d^0 - \bar{B}_d^0$ mixing is measured with impressive accuracy and can be used to determine the value of V_{td}, giving[12] (in the Wolfenstein CKM parameterization)

$$\begin{aligned} |V_{td}| &= A\lambda^3[(1-\rho)^2+\eta^2]^{1/2}, \qquad (64) \\ &= 8.54\times 10^{-3}\left[\frac{207\ MeV}{f_{B_d}\sqrt{B_{B_d}}}\right]\left[\frac{170\ GeV}{\bar{m}_t(m_t)}\right]^{0.76}\left[\frac{\Delta M_d}{0.464\ ps^{-1}}\right]\sqrt{\frac{0.55}{\eta}}. \end{aligned}$$

Setting the input parameters at their 1σ values gives the range $|V_{td}| = 0.007 - 0.010$. Unfortunately, this evaluation of V_{td} is dominated by the uncertainties associated with the hadronic matrix elements and assumes that new physics does not contribute to $B_d^0 - \bar{B}_d^0$ mixing.

A measurement of $B_s^0 - \bar{B}_s^0$ mixing could also yield a value for the ratio of CKM elements $|V_{td}/V_{ts}|$ via

$$\frac{\Delta M_d}{\Delta M_s} = \frac{f_{B_d}^2 B_{B_d}\eta_{B_d} m_{B_d}|V_{td}|^2}{f_{B_s}^2 B_{B_s}\eta_{B_s} m_{B_s}|V_{ts}|^2} = \xi^2\lambda^2[(1-\rho)^2+\eta^2]. \qquad (65)$$

The factor which multiplies the ratio of CKM elements, ξ, measures the amount of SU(3) breaking effects. The ratio of hadronic matrix elements, $f_{B_d}\sqrt{B_{B_d}}/f_{B_s}\sqrt{B_{B_s}}$, is more accurately determined in lattice gauge theory than the individual quantities with the current global value[21] being 1.15 ± 0.05 in quenched calculations. However, unquenching is expected to increase this result by 10%. The LEP bound on ΔM_s then yields[12] the 95% C.L. constraint

$$\left|\frac{V_{td}}{V_{ts}}\right| < 0.29. \qquad (66)$$

We note that if V_{ts} is relatively large, a sensitive technique[97] of extracting $|V_{td}/V_{ts}|$ could be obtained from a measurement of $\Delta\Gamma/\Gamma$ for the B_s meson.

Remarkably, the above technique for extracting $|V_{td}/V_{ts}|$ remains valid in many scenarios beyond the SM. In this class of models, the virtual exchange of new particles alters the Inami-Lim function in Eq. (63) above, but not the factors in front of the function. The effects of the new physics then cancels in the ratio $\Delta M_d/\Delta M_s$. Models of this type include Two-Higgs-Doublet models and supersymmetry in the super-CKM basis. Notable exceptions to this feature can be found in models which (i) change the structure of the CKM matrix, such as the addition of a fourth generation, or extra singlet quarks, and in Left-Right Symmetric models, (ii) have couplings proportional to fermion masses, such as flavor changing Higgs models, or (iii) have generational dependent couplings, *e.g.*, leptoquarks or supersymmetry with R-parity violation.

It is difficult to use ΔM_d alone to restrict new physics due to the errors on the theoretical predictions for this quantity from the imprecisely determined CKM factors and B hadronic matrix elements. (This is unfortunate as ΔM_d is so precisely measured!) In most cases, the restrictions obtained from $B \to X_s\gamma$ surpass those from $B^0 - \bar{B}^0$ mixing.

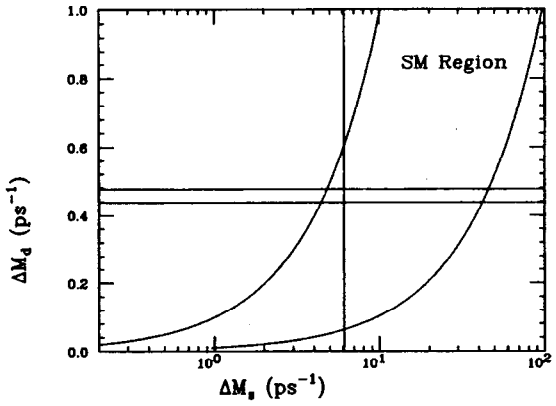


Figure 10: The SM expectation for the $\Delta M_d - \Delta M_s$ plane, where the predicted region lies inside the solid curves. The experimental bounds lie in between the solid horizontal lines and to the right of the solid vertical line.

5.3 $D^0 - \bar{D}^0$ Mixing

Currently, the limits[98] on $D^0 - \bar{D}^0$ mixing are from fixed target experiments, with $x_D \equiv \Delta m_D/\Gamma < 0.083$, implying $\Delta m_D < 1.3\times 10^{-13}$ GeV, from an analysis which assumes there is no interference between doubly-Cabibbo suppressed decays and the mixing amplitude. A more recent result,[99] which takes these interference

effects into account, obtains the bound on the ratio of wrong-sign to right-sign final states of $r_D \equiv \Gamma(D^0 \to \ell^- X)/\Gamma(D^0 \to \ell^+ X) < 0.50\%$ at 90% C.L., where

$$r_D \approx \frac{1}{2}\left|\frac{q}{p}\right|^2 \left[\left(\frac{\Delta m_D}{\Gamma}\right)^2 + \left(\frac{\Delta\Gamma}{2\Gamma}\right)^2\right] . \tag{67}$$

This gives $\Delta M_D < 1.58 \times 10^{-13}$ GeV, assuming $|q/p|^2 = 1$ and $\Delta\Gamma \approx 0$.

The short-distance SM contributions to Δm_D proceed through a W box diagram with internal d, s, b quarks. In this case, the external momentum, which is of order m_c, is communicated to the light quarks in the loop and cannot be neglected. The effective Hamiltonian becomes

$$\mathcal{H}_{eff}^{\Delta c=2} = \frac{G_F \alpha}{8\sqrt{2}\pi x_w}\left[|V_{cs}V_{us}^*|^2\left(I_1^s \mathcal{O}_{LL} - m_c^2 I_2^s \mathcal{O}_{RR}\right) + |V_{cb}V_{ub}^*|^2\left(I_3^b \mathcal{O}_{LL} - m_c^2 I_4^b \mathcal{O}_{RR}\right)\right] , \tag{68}$$

where the I_j^q represent integrals[100] that are functions of m_q^2/M_W^2 and m_q^2/m_c^2, and $\mathcal{O}_{LL} = [\bar{u}\gamma_\mu(1-\gamma_5)c]^2$ is the usual mixing operator while $\mathcal{O}_{RR} = [\bar{u}(1+\gamma_5)c]^2$ arises in the case of nonvanishing external momentum. The numerical value of the short-distance contribution is $\Delta m_D \sim 5 \times 10^{-18}$ GeV (taking $f_D =$ 200 MeV). The long-distance contributions have been computed via two different techniques: (i) the intermediate particle dispersive approach (using current data on the intermediate states) yields $\Delta m_D \sim 10^{-4}\Gamma \simeq 10^{-16}$ GeV (Ref. 101), and (ii) heavy quark effective theory which results in $\Delta m_D \sim (1-2)\times 10^{-5}\Gamma \simeq 10^{-17}$ GeV (Ref. 102). Clearly, the long-distance contributions overwhelm those from short-distance SM physics in $D^0 - \bar{D}^0$ mixing, and both contributions lie far below the present experimental sensitivity.

One reason the SM expectations for $D^0 - \bar{D}^0$ mixing are so small is that there are no heavy particles participating in the box diagram to enhance the rate. Hence, the first extension to the SM that we consider is the addition[103] of a heavy $Q = -1/3$ quark. We can now neglect the external momentum, and Δm_D is given by the usual expression[30]

$$\Delta m_D = \frac{G_F^2 M_W^2 m_D}{6\pi^2} f_D^2 B_D |V_{cb'}V_{ub'}^*|^2 F(m_{b'}^2/M_W^2) . \tag{69}$$

The value of Δm_D is displayed in this model in Fig. 11(a) as a function of the overall CKM mixing factor for various values of the heavy quark mass. We see that Δm_D approaches the experimental bound for large values of the mixing factor.

Another simple extension of the SM is to enlarge the Higgs sector by an additional doublet. First, we examine two-Higgs-doublet models which avoid tree-level FCNC by introducing a global symmetry; such models are discussed above in the sections on leptonic and radiative decays. The expression for Δm_D in this case can be found in Ref. 46. From the Lagrangian in Eq. (17), it is clear that Model I will only modify the SM result for Δm_D for very small values of $\tan\beta$, and this region is already excluded from existing data on $B \to X_s\gamma$. However, enhancements can occur in Model II for large values of $\tan\beta$, as demonstrated in Fig. 11(b).

Next, we consider the case of extended Higgs sectors without natural flavor conservation. In these models, the above requirement of a global symmetry which restricts each fermion type to receive mass from only one doublet is replaced[104] by approximate flavor symmetries which act on the fermion sector. The Yukawa couplings can then possess a structure which reflects the observed fermion mass and mixing hierarchy. This allows the low-energy FCNC limits to be evaded as the flavor changing couplings to the light fermions are small. We employ the Cheng-Sher ansatz,[86] where the flavor changing couplings of the neutral Higgs are $\lambda_{h^0 f_i f_j} \approx (\sqrt{2}G_F)^{1/2}\sqrt{m_i m_j}\Delta_{ij}$, with the $m_{i(j)}$ being the relevant fermion masses and Δ_{ij} representing a combination of mixing angles. h^0 can now contribute to Δm_D through tree-level exchange as well as mediating $D^0 - \bar{D}^0$ mixing by h^0 and t-quark virtual exchange in a box diagram. These latter contributions only compete with those from the tree-level process for large values of Δ_{ij}. In Figs. 11(c) and 11(d), we show the constraints placed on the parameters of this model from the present experimental bound on Δm_D for both the tree-level and box diagram contributions.

The last contribution to $D^0 - \bar{D}^0$ mixing that we consider here is that of scalar leptoquark bosons. They participate in Δm_D via virtual exchange inside a box diagram,[62] together with a charged lepton or neutrino. Assuming that there is no leptoquark-GIM mechanism, and taking both exchanged leptons to be the same type, we obtain the restriction

$$\frac{F_{tc}F_{tu}}{m_{LQ}^2} < \frac{196\pi^2 \Delta m_D}{(4\pi\alpha f_D)^2 m_D} , \tag{70}$$

where F_{tq} parameterize the *a priori* unknown leptoquark Yukawa couplings as $\lambda_{tq}^2/4\pi = F_{tq}\alpha$. The resulting bounds in the leptoquark coupling-mass plane are presented in Fig. 11(e).

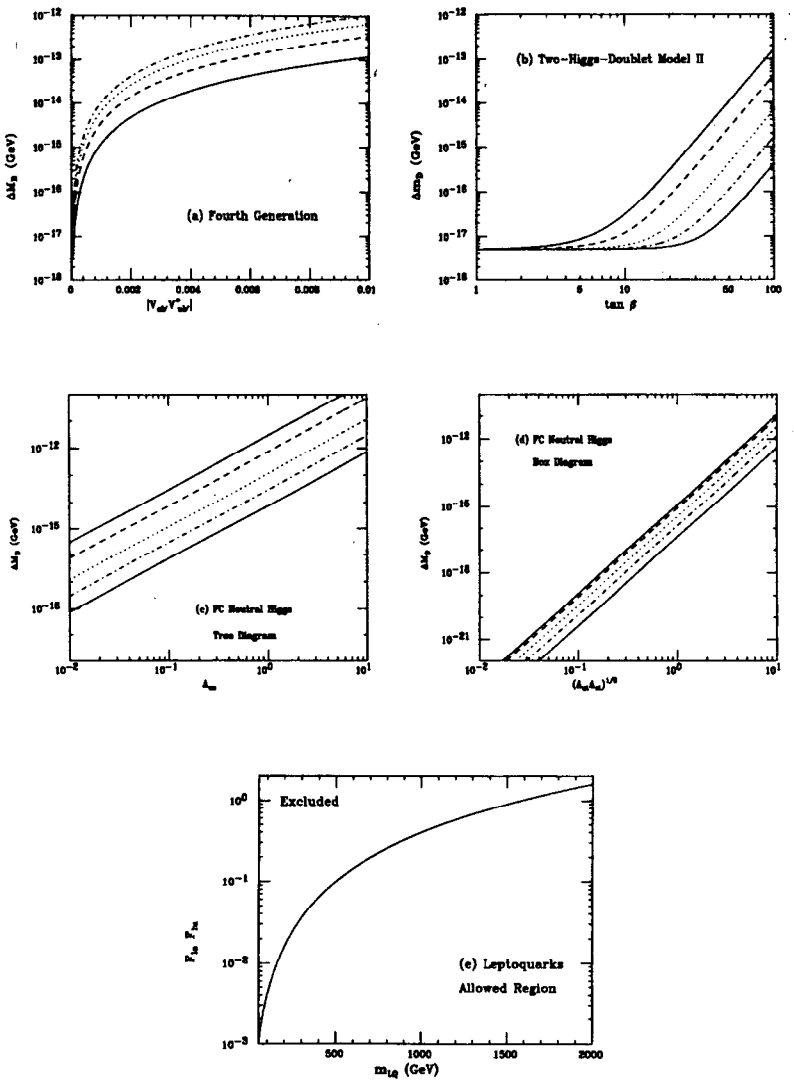


Figure 11: Δm_D in (a) the four-generation SM with the solid, dashed, dotted, dash-dotted curve corresponding to fourth generation quark masses $M_{b'} = 100, 200, 300$, and 400 GeV, respectively. (b) The two-Higgs-doublet Model II as a function of $\tan\beta$ with, from top to bottom, the solid, dashed, dotted, dash-dotted, solid curve representing $m_{H^\pm} = 50, 100, 250, 500$, and 1000 GeV. (c) Tree-level and (d) box diagram contributions to Δm_D in the flavor changing Higgs model described in the text as a function of the mixing factor for $m_h = 50, 100, 250, 500$, and 1000 GeV corresponding to the solid, dashed, dotted, dash-dotted, and solid curves from top to bottom. (e) Constraints in the leptoquark coupling-mass plane from Δm_D.

6 CP Violation

The symmetries C, charge conjugation (which describes particle-antiparticle interchange), P, parity (which relates left- to right-handed particles), and T, time reversal (which correlates a process with its time-reversed state), are all preserved under the strong and electromagnetic interactions. Weak processes, however, are known to violate each of these symmetries separately, while conserving the product CPT, which is an exact symmetry of the equations of motion. Weak decays violate C and P at a fairly large level, while the product CP has been observed to be violated at a much smaller rate.

CP violation arises in the SM from the existence of the phase in the three-generation CKM matrix as first postulated by Kobayashi and Maskawa.[2] Unitarity of the CKM matrix can be represented geometrically in terms of triangles in the complex plane. For example, the relation $V_{tb}V_{td}^* + V_{cb}V_{cd}^* + V_{ub}V_{ud}^* = 0$, can be depicted as the triangle displayed in Fig. 12. The figure depicts the rescaled triangle, where the length of all sides are scaled to $|V_{cd}V_{cb}^*|$, and hence the length of the base is unity. In this case, it can be shown that the apex of the triangle is located at the point (ρ, η) in the complex plane, where ρ and η are the Wolfenstein parameters describing the CKM matrix. Here, the unitarity angles α, β, and γ are related to the magnitudes of the sides of the triangle by

$$\alpha = arg\left(\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right),\ \beta = arg\left(\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right),\ \gamma = arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{71}$$

The values of these angles are rather poorly constrained at present, as will be discussed below. The area of the triangle represents the amount of CP violation in the SM, and can be described by the Jarlskog[105] parameter

$$\begin{aligned} J &= 2 \times \text{area of triangle}, \\ &= |V_{ud}||V_{us}||V_{ub}||V_{cb}|\sin\delta = A^2\lambda^6\eta \simeq \mathcal{O}(10^{-5}). \end{aligned} \tag{72}$$

Similar unitarity triangles, representing other orthogonality relations of the CKM matrix, may also be drawn. All such triangles clearly have the same area in the SM; however, the remaining triangles involve one side which is much shorter than the other two, and consequently one of their unitarity angles is extremely small. This is in contrast to the above triangle, where all three sides are of comparable magnitude, $\mathcal{O}(\lambda^3)$, and hence all three angles are naturally large.

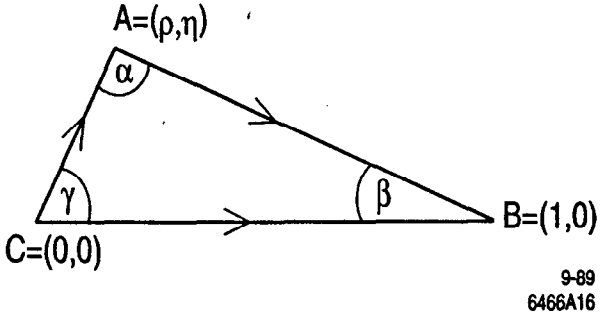


Figure 12: The rescaled unitarity triangle.

This explains why CP asymmetries are predicted to be large in neutral B_d decays. For example, the triangle representing the neutral K meson system, which is built from the relation $\sum_q V_{qd}V_{qs}^* = 0$, has two long sides of length $\mathcal{O}(\lambda)$ and a third side of length $\mathcal{O}(\lambda^5)$. Hence, CP asymmetries in this system are related to the small angle of this unitarity triangle and are of order 10^{-3}.

There are many additional sources of CP violation in theories beyond the SM, such as multi-Higgs-Doublet models, supersymmetry, and Left-Right Symmetric models.[106] It is worth noting that the observed matter-antimatter asymmetry of the universe may require additional sources of CP violation beyond the CKM phase.[107]

There is a vast literature on CP violation[108] to which we refer the reader for a more detailed discussion. Here, we now describe the three manifestations of CP violation and how they are observed in the various meson systems. In all cases, one should keep in mind the experimentally relevant number for measurement of a CP asymmetry at the $n\sigma$ level,

$$N = \frac{n^2}{Ba^2}\,, \tag{73}$$

where N is the number of identified P mesons required for observation of the asymmetry (not including efficiency reductions), B represents the branching fraction of the decay mode, and a is the value of the CP violating asymmetry.

6.1 CP Violation in Decays

CP violating effects may be observed directly in the decays of charged or neutral mesons. This is referred to as Direct CP Violation, or CP Violation in Decays. CPT symmetry assures us that the total width of a particle and its antiparticle are identical, *i.e.*,

$$\Gamma_{\text{total}} = \bar{\Gamma}_{\text{total}}\,. \tag{74}$$

If CP were conserved, this would also hold true for the partial decay width of a meson to a particular final state, $P \to f$, versus the time-reversed process, $\bar{P} \to \bar{f}$. For CP to be violated, these two partial widths must be different, *i.e.*,

$$\Gamma(P \to f) \neq \Gamma(\bar{P} \to \bar{f})\,. \tag{75}$$

In order for direct CP violation to occur, the decay amplitudes must have contributions from (at least) two different weak phases, and two separate strong phases. This can easily be seen as follows. Let us assume that the decay amplitude to the final state f has the form

$$A_f = A_1 e^{i\delta_1} + A_2 e^{i\delta_2}\,, \tag{76}$$

with $A_{1,2}$ being the two weak amplitudes after the strong phases $\delta_{1,2}$ have been factored out. In the SM, all tree-level contributions to a given transition enter with the same weak phase, whereas penguin diagrams can contribute with a different phase. Here, we identify A_1 as the tree-level amplitude, while A_2 represents the penguin transition. For the CP conjugate amplitude, the weak phases are conjugated, $A_{1,2} \to A_{1,2}^*$, but the strong phases are not. The CP asymmetry is then given by

$$\frac{|A_f|^2 - |\bar{A}_{\bar{f}}|^2}{|A_f|^2 + |\bar{A}_{\bar{f}}|^2} = \frac{2\mathcal{I}m\,(A_1^* A_2)\sin(\delta_1 - \delta_2)}{|A_1|^2 + |A_2|^2 + 2\mathcal{R}e\,(A_1^* A_2)\cos(\delta_1 - \delta_2)}\,, \tag{77}$$

which clearly vanishes if $A_{1,2}$ contain the same weak phase and if $\delta_1 = \delta_2$. Hence, direct CP violation arises from the product of the weak phase difference, which is odd under CP, and the strong phase difference from final state interactions, which is even under CP. Since this results from the interference between the tree-level and penguin transitions, the magnitude of direct CP violation is related to the size of the penguin contributions. Unfortunately, there is at present no unambiguous experimental signal for direct CP violation.

6.2 CP Violation in Mixing

Indirect CP violation, or CP violation due to mixing, is a consequence of the fact that the mass eigenstates $P_{L,H}$ are not CP eigenstates and is represented by the potential deviation of $|q/p|$ from unity. Clearly, these effects only arise in neutral meson decays. This process is theoretically clean as it is independent of the strong phases and thus provides a direct measurement of the CKM phase[109] (at least within the SM).

The CP violating observable that can be defined in this case is given by

$$\frac{1-|q/p|^2}{1+|q/p|^2} = \frac{2\mathcal{R}e\,\epsilon_P}{1+|\epsilon_P|^2} = \frac{2\mathcal{I}m\,(M_{12}^*\Gamma_{12}/2)}{|\Gamma_{12}/2|^2+|M_{12}|^2+[(\Delta m_P)^2+(\Delta\Gamma_P/2)^2]/4}, \quad (78)$$

where we give the expression in terms of M_{12} and Γ_{12}, which are defined in the previous sections. This quantity is independent of the phase convention and is directly observable. As noted by Ma *et al.*,[110] this observable demonstrates that CP nonconservation is determined by the relative phase between M_{12} and Γ_{12}. Defining $\Delta\Gamma_P/\Delta M_P \equiv a$ and taking the approximation (which is valid in the SM only) that $\Delta M \simeq 2\mathcal{R}e\,M_{12}$ and $\Delta\Gamma \simeq 2\mathcal{R}e\,\Gamma_{12}$, the above expression can be written in the more convenient form

$$\frac{2\mathcal{R}e\,\epsilon_P}{1+|\epsilon_P|^2} = \frac{a}{2(1+a^2/4)}\left[\frac{\mathcal{I}m\,\Gamma_{12}}{\mathcal{R}e\,\Gamma_{12}} - \frac{\mathcal{I}m\,M_{12}}{\mathcal{R}e\,M_{12}}\right]. \quad (79)$$

6.3 CP Violation in the Interference Between Mixing and Decay

Additional CP violating effects can arise from the interference of a pseudoscalar meson P^0 decaying to a final state f at time t, with a P^0 which mixes into a $\bar{P}^0$ state which then decays to f at time t. We define the phase convention independent quantity

$$r_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}. \quad (80)$$

When CP is conserved, $|q/p| = 1\,, |\bar{A}_f/A_f| = 1$, and the relative phase between these two quantities vanishes. If any one of these three conditions are not met, then $r_f \neq 1$ and CP is violated. As discussed above, if the first condition (the deviation of $|q/p|$ from unity) doesn't hold, then CP violation occurs through mixing, while if the magnitudes of the amplitudes differ, then CP violation occurs in decay. However, even if these first two conditions hold, it is possible that the relative phase between these two quantities is nonzero resulting in $\mathcal{I}m\,r_f \neq 0$ while $|r_f| = 1$. It is this case that we call CP violation in the interference between mixing and decay. This case is also independent of hadronic uncertainties and hence is theoretically clean and can be directly related to the CKM matrix elements. We will discuss the significance and potential measurements of this third type of CP violation in the kaon and B meson systems separately.

6.4 CP Violation in the Kaon System

The kaon system has provided our only experimental observation of CP violation. The charge asymmetry in semileptonic decay, $K_L^0 \to \ell\nu_\ell + X$,

$$a_{s\ell} = \frac{\Gamma(K_L \to \ell^+\nu_\ell X) - \Gamma(K_L \to \ell^-\bar{\nu}_\ell X)}{\Gamma(K_L \to \ell^+\nu_\ell X) + \Gamma(K_L \to \ell^-\bar{\nu}_\ell X)}, \quad (81)$$

has been measured[4] to have the value $a_{s\ell} = (3.27 \pm 0.12) \times 10^{-3}$. Since we can relate $\langle \ell^+\nu_\ell X|H|K_L\rangle = p\mathcal{A}$, while $\langle \ell^-\bar{\nu}_\ell X|H|K_L\rangle = q\mathcal{A}^*$, the charge asymmetry can be identified as a determination of $|q/p| \neq 1$ with

$$a_{s\ell} = \frac{2\mathcal{R}e\,\epsilon_K}{1+|\epsilon_K|^2}. \quad (82)$$

CP violation has also been measured in the decay of $K_L \to \pi\pi$. The amplitudes in the CP eigenstate basis can be written as

$$\begin{aligned} \mathcal{A}(K^0 \to \pi\pi(I)) &= A_I e^{i\delta_I}, \\ \mathcal{A}(\bar{K}^0 \to \pi\pi(I)) &= A_I^* e^{i\delta_I}, \end{aligned} \quad (83)$$

where I denotes the isospin of the $\pi\pi$ final state, δ_I is the final state phase shift, and A_I would be real if CP were conserved. It is interesting to note that experimentally, $|\mathcal{A}_0/\mathcal{A}_2| = 20$. The following ratios of CP violating to CP conserving amplitudes have been measured

$$\begin{aligned} \eta_{+-} &= \frac{\mathcal{A}(K_L^0 \to \pi^+\pi^-)}{\mathcal{A}(K_S^0 \to \pi^+\pi^-)}, \\ \eta_{00} &= \frac{\mathcal{A}(K_L^0 \to \pi^0\pi^0)}{\mathcal{A}(K_S^0 \to \pi^0\pi^0)}. \end{aligned} \quad (84)$$

These differences in these quantities can be parameterized as

$$\begin{aligned} \eta_{+-} &= \epsilon + \epsilon', \\ \eta_{00} &= \epsilon - 2\epsilon', \end{aligned} \quad (85)$$

where ϵ and ϵ' are defined by

$$\begin{aligned} \epsilon &= \epsilon_K + i\frac{\mathcal{I}m\,A_0}{\mathcal{R}e\,A_0}, \\ |\epsilon'| &= \frac{1}{\sqrt{2}}\frac{\mathcal{R}e\,A_2}{\mathcal{R}e\,A_0}\left[\frac{\mathcal{I}m\,A_2}{\mathcal{R}e\,A_2} - \frac{\mathcal{I}m\,A_0}{\mathcal{R}e\,A_0}\right]. \end{aligned} \tag{86}$$

These parameters, ϵ and ϵ', are defined so that the potentially direct CP violating effects are isolated and affect only ϵ'. In superweak models,[111] CP violation is confined to the mass matrix and ϵ' is predicted to vanish. As shown by Wu and Yang,[112] it is possible to adopt a phase convention such that $\mathcal{I}m\,A_0 = 0$. In this case, we then have

$$\begin{aligned} \epsilon &= \epsilon_K, \\ |\epsilon'| &= \frac{1}{\sqrt{2}}\frac{\mathcal{R}e\,A_2}{\mathcal{R}e\,A_0}\frac{\mathcal{I}m\,A_2}{\mathcal{R}e\,A - 2}. \end{aligned} \tag{87}$$

ϵ is then given by the CP violation effects due to mixing and can be calculated as

$$\begin{aligned} \epsilon &= \epsilon_K = \frac{e^{i\pi/4}}{\sqrt{2}\Delta M_K}\mathcal{I}m\,M_{12}, \\ &= \frac{G_F^2 M_W^2 f_K^2 B_K m_K}{6\sqrt{s}\pi^2\Delta m_K} A^2\lambda^6\eta\left[\eta_{ct}S(x_c,x_t) - \eta_c S(x_c) + \eta_t A^2\lambda^4(1-\rho)S(x_t)\right], \end{aligned} \tag{88}$$

in the Wolfenstein parameterization. We see that the measurement of ϵ_K guarantees that $\eta \neq 0$! The uncertainties in the calculation of ϵ_K are equivalent to those outlined above in the case of $K_L - K_S$ mixing.

Returning to direct CP violation in K decays, we see that ϵ'/ϵ can be expressed in terms of the operator product expansion[113] by relating $\mathcal{R}e\,A_{0,2}$ and $\mathcal{I}m\,A_{0,2}$ to the appropriate Wilson coefficients and hadronic matrix elements. The effective Hamiltonian for this process can then be written as

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}}V_{us}V_{ud}^*\sum_{i=1}^{10}\left[Z_i(\mu) - \frac{V_{ts}V_{td}^*}{V_{us}V_{ud}^*}y_i(\mu)\right]\mathcal{O}_i(\mu), \tag{89}$$

where the sum extends over the set of operators given by the current-current operators $\mathcal{O}_{1,2}$, the QCD penguin operators $\mathcal{O}_{3-6}$, and the electroweak operators $\mathcal{O}_{7-10}$. The functions z_i and y_i are related to the Wilson coefficients; their forms are given explicitly in Ref. 113. In this formalism, it is easy to see that ϵ'/ϵ is governed by both QCD and electroweak penguin transitions. In fact, due to the large value of the top quark mass, the electroweak penguin amplitudes play an important role[114] and enter ϵ'/ϵ with the opposite sign of the QCD penguin contributions. This serves to suppress the prediction of ϵ'/ϵ within the SM. In fact, for $m_t = 200$ GeV, the SM prediction for ϵ'/ϵ is zero! Due to this strong cancellation for large values of m_t and the uncertainties associated with the hadronic matrix elements, a precise SM prediction for ϵ'/ϵ is very difficult. However, a simplified analytic expression which highlights these uncertainties may be written as[12]

$$\frac{\epsilon'}{\epsilon} = 11\times 10^{-4}\left[\frac{\eta\lambda^5A^2}{1.3\times 10^{-4}}\right]\left[\frac{140\text{ MeV}}{m_2(2\text{ GeV})}\right]\left[\Lambda_{\overline{MS}}^{(4)}\right][B_6 - Z(x_t)B_8], \tag{90}$$

where $Z(x_t) \approx 0.18(m_t/M_W)^{1.86}$, and $B_{6,8}$ represent the hadronic matrix elements corresponding to operators $\mathcal{O}_{6,8}$. The most recent analysis,[115] which incorporates the latest determinations of all the input parameters, predicts the range

$$-1.2\times 10^{-4} \leq \epsilon'/\epsilon \leq 16.0\times 10^{-4}. \tag{91}$$

This prediction may be altered, however, if the value of the strange quark mass is as low as presently calculated in lattice gauge theories.[21]

The importance of the measurement of ϵ'/ϵ to understand more about the mechanism of CP violation cannot be overemphasized, although, due to presently conflicting experimental results, constraints on new physics from ϵ'/ϵ will not be taken into account here. The next round of experiments, which will reach a precision of 10^{-4}, might settle the issue of whether or not $\epsilon'/\epsilon \neq 0$. As shown above, the SM prediction allows for a wide range of values for ϵ'/ϵ. Ultimately, one wants to establish whether CP violation is milliweak ($\Delta S = 1$) as in the SM and/or superweak ($\Delta S = 2$). The latter occurs in multi-Higgs doublet models through scalar interactions, in SUSY models,[116] or in the LRM to give a few examples.[117]

We note briefly that the decay $K_L \to \pi^0\nu\bar{\nu}$, which is related to $K^+ \to \pi^+\nu\bar{\nu}$ discussed above, proceeds almost exclusively through direct CP violation and would provide a clean laboratory to measure this phenomenon. Unfortunately, the branching fraction is extremely small in the SM at $B(K_L \to \pi^0\nu\bar{\nu}) \approx 2.8\times 10^{-11}$, and the present experimental limits[4] lie above this prediction by roughly six orders of magnitude.

6.5 CP Violation in B Decays

CP violation in the B system will be examined[118] during the next decade at dedicated e^+e^- B factories and at hadron colliders. A theoretically clean technique

is offered[109] by the measurement of the time-dependent CP asymmetry, which involves the CP violating effects from the interference of mixing and decay. Given the proper time evolution of a neutral pseudoscalar meson state of Eq. (56) and the definition of r_f in Eq. (80), we can write the time-dependent rate for the decay of initially pure B^0 or $\bar{B}^0$ states into a CP eigenstate as

$$\begin{aligned}\Gamma(B^0_{phys} \to f_{CP}) &= |A|^2 e^{-\Gamma t}\left[\frac{1+|r_f|^2}{2} + \frac{1-|r_f|^2}{2}\cos(\Delta M t) - \mathcal{I}m\, r_f \sin(\Delta M t)\right], \\ \Gamma(\bar{B}^0_{phys} \to f_{CP}) &= |A|^2 e^{-\Gamma t}\left[\frac{1+|r_f|^2}{2} - \frac{1-|r_f|^2}{2}\cos(\Delta M t) + \mathcal{I}m\, r_f \sin(\Delta M t)\right].\end{aligned} \tag{92}$$

The time-dependent CP asymmetry can then be expressed as

$$\begin{aligned}a_f(t) &= \frac{\Gamma(B^0_{phys} \to f_{CP}) - \Gamma(\bar{B}^0_{phys} \to f_{CP})}{\Gamma(B^0_{phys} \to f_{CP}) + \Gamma(\bar{B}^0_{phys} \to f_{CP})}, \\ &= \frac{(1-|r_f|^2)\cos(\Delta M t) - 2\mathcal{I}m\, r_f \sin(\Delta M t)}{1+|r_f|2},\end{aligned} \tag{93}$$

and hence is directly related to $\mathcal{I}m\, r_f$. In fact, for decay modes which have $|r_f| = 1$, the time-dependent asymmetry reduces to

$$a_f(t) = -\mathcal{I}m\, r_f \sin(\Delta M t). \tag{94}$$

Recall that when there is no direct CP violation in a channel, all amplitudes that contribute to the decay mode have the same CKM phase, denoted generically here as ϕ_D, and hence $|\bar{A}_f/A_f| = |e^{-2i\phi_D}| = 1$. In this case, r_f can be completely expressed in terms of the CKM matrix elements as $r_f = \pm \exp -2i(\phi_D + \phi_M)$, where ϕ_M represents the mixing phase from $q/p = \sqrt{M^*_{12}/M_{12}} = e^{-2i\phi_M}$ (for $\Gamma_{12} \ll M_{12}$), and the overall sign is determined by the CP eigenvalue of the final state f. Clearly, the asymmetry is simply

$$a_f(t) = \pm \sin(2(\phi_D + \phi_M)) \sin(\Delta M t). \tag{95}$$

In order to relate the time-dependent CP asymmetry to the CKM parameters, one needs to examine the CKM dependence of mixing and of the amplitudes of the relevant decay channels. An extensive summary of these relations for various decay modes is given in the Particle Data Book.[4] Here, we briefly discuss two important cases, $B_d \to J/\psi K_s$ and $B_d \to \pi\pi$. In the first case, the quark subprocess responsible for the decay is $b \to c\bar{c}s$, which is dominated by a tree-level diagram mediated by W-boson exchange. There are small penguin contributions as well; however, the penguin weak phase, $\arg(V_{tb}V^*_{ts})$, is similar (modulo π) to the weak phase of the tree-level contribution. We thus have contributions to the weak phase from the CKM structure of the decay diagram, from $B^0_d - \bar{B}^0_d$ mixing, and from $K^0 - \bar{K}^0$ mixing in the final state. This gives

$$\begin{aligned}r_f &= -\left(\frac{q}{p}\right)_{B_d} \frac{\bar{A}_{J/\psi K_S}}{A_{J/\psi K_S}} \left(\frac{q}{p}\right)_K, \\ &= \frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}, \frac{V_{cs}V^*_{cb}}{V^*_{cs}V_{cb}}, \frac{V^*_{cs}V_{cd}}{V_{cs}V^*_{cd}},\end{aligned} \tag{96}$$

where the minus sign arises since $J/\psi K_S$ is CP-odd. Comparing this with Eq. (71) yields $\mathcal{I}m\, r_f = -\sin 2\beta$. This gives the theoretically cleanest determination of a unitarity angle! In the latter example, $B_d \to \pi\pi$, the quark subprocess is $b \to u\bar{u}d$, which is again dominated by tree-level W exchange. In this case, we have

$$\begin{aligned}r_f &= \left(\frac{q}{p}\right)_{B_d} \frac{\bar{A}_{\pi\pi}}{A_{\pi\pi}}, \\ &= \frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}, \frac{V_{ub}V^*_{ud}}{V_{ud}V^*_{ub}},\end{aligned} \tag{97}$$

which then gives $\mathcal{I}m\, r_f = \sin 2\alpha$. Unfortunately, this process is not as clean as $B_d \to J/\psi K_S$, as both the gluonic and electroweak penguin contributions enter with a different phase at an unknown size. $B_d \to \pi\pi$ thus suffers from what is called penguin contamination. The amount of this contamination needs to be separately determined.[119]

The present status of the unitarity triangle in the $\rho - \eta$ plane is summarized in Fig. 13(a), where the shaded area is that allowed in the SM. This region is determined by measurements of the quantities (i) $|V_{ub}|$ and $|V_{cb}|$, (ii) ϵ_K, and (iii) the rates for $B^0_d - \bar{B}^0_d$ and $B^0_s - \bar{B}^0_s$ mixing, as discussed above, together with theoretical estimates for the parameters which relate these measurements to the underlying theory, such as B_K, f_B, and B_B. The value of $\overline{m_t}(m_t)$ is taken to be consistent with the physical range 175 ± 6 GeV. Here we have employed the scanning technique, where both the experimental measurements and theoretical input parameters are scanned independently within their 1σ errors. This method yields the SM ranges for the angles of the unitarity triangle: $-0.89 \le \sin 2\alpha \le 1.00$, $0.18 \le \sin 2\beta \le 0.81$, and $-1.00 \le \sin 2\gamma \le 1.00$. Since

the ratio $\Delta M_{B_s}/\Delta M_{B_d}$ is more accurately related to the theoretical predictions than the separate quantities, we see that a measurement of $B_s^0 - \bar{B}_s^0$ mixing would be an invaluable tool in determining the angles of this triangle.

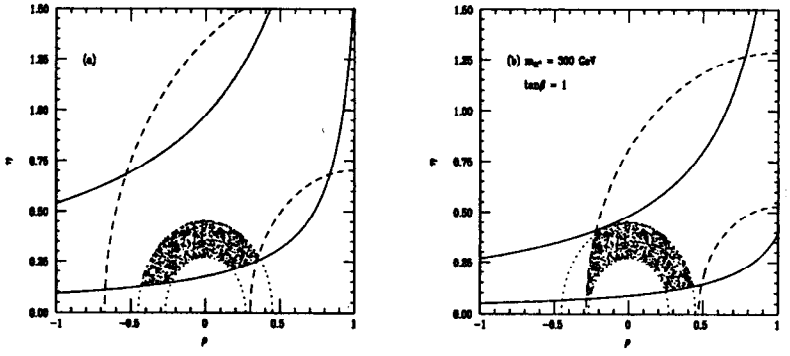


Figure 13: Constraints in the (a) SM and (b) two-Higgs-doublet Model II in the $\rho - \eta$ plane from $|V_{ub}|/|V_{cb}|$ (dotted circles), $B_d^0 - \bar{B}_d^0$ mixing (dashed circles), and ϵ (solid hyperbolas). The shaded area corresponds to that allowed for the apex of the unitarity triangle.

It is important to remember that this picture can be dramatically altered if new physics is present, even if there are no new sources of CP violation. Figure 13(b) displays the constraints in the $\rho - \eta$ plane in the two-Higgs-doublet Model II. In this case, the presence of the extra Higgs doublet is felt by the virtual exchange of the $H^\pm$ boson in the box diagram which mediates $B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$ mixing and governs the value of ϵ_K. For this $\rho - \eta$ region, the allowed ranges of the angles of the unitarity triangle become $-1.00 \leq \sin 2\alpha \leq 1.00$, $0.12 \leq \sin 2\beta \leq 0.81$, and $-1.00 \leq \sin 2\gamma \leq 1.00$. In fact, this opens up a new allowed region in the $\sin 2\alpha - \sin 2\beta$ plane, as shown in Fig. 14 from Ref. 120. Similar effects have also been pointed out in supersymmetric models.[121] Clearly, caution must be exercised when relating the results of future CP violation experiments to the $\rho - \eta$ plane.

The B factories presently under construction should be able to discern whether new physics contributes to CP violation. Signals for new sources of CP violation include (i) nonclosure of the three-generation unitarity triangle, (ii) new contributions to $B^0 - \bar{B}^0$ mixing which yield a nonvanishing phase for this process, (iii) nonvanishing CP asymmetries for the channels $B_d^0 \to \phi\pi^0, K_S^0 K_S^0$, (iv) inconsistency of separate measurements of the angles of the unitarity triangle, and (v_i) a deviation of CP rates from SM predictions. Models which contain additional CP phases include nonminimal Supersymmetry, Multi-Higgs Doublets, Left-Right Symmetric Models, and the Superweak Model. A concise review of the effects of these models on CP violating observables is given by *Grossman et al.*[106] We present here, as an example, the case of multi-Higgs models with three or more Higgs doublets. In this scenario, $B^0 - \bar{B}^0$ mixing receives additional contributions from the $H_{1,2}^\pm$ exchange which depends on the phase in the charged scalar mixing matrix. Interference between these contributions and the SM yield an overall nonzero phase in ΔM_{B_d}. Denoting this phase as θ_H, the unitarity angles measured by CP asymmetries in B decays are thus shifted by

$$a_{CP}(B \to J/\psi K_S) = -\sin(2\beta + \theta_H),\ a_{CP}(B \to \pi\pi) = \sin(2\alpha + \theta_H)\,. \qquad (98)$$

The magnitude of this effect depends on the size of θ_H, which has recently[120] been constrained by $B \to X_s\gamma$. Another interesting example is provided in models with an extra iso-singlet down quark; in this scenario, it has been found[122] that measurements of the unitarity angles α and β alone are not enough to distinguish and bound the new contributions, and that observation of both the third angle γ and B_s mixing are also needed. In summary, the large data sample which will become available will provide a series of unique consistency tests of the quark sector and will challenge the SM in a new and quantitatively precise manner.

6.6 CP Violation in the D Meson System

CP violation in the $Q = 2/3$ quark sector is complementary to that of the K and B systems, but has yet to be explored. In the SM, the CKM phase is responsible for generating CP violation, and in the charm system, the resulting rates are small. However, new sources of CP violating phases could greatly enhance the rates, thus rendering CP violation in the charm system a sensitive probe for physics beyond the SM.

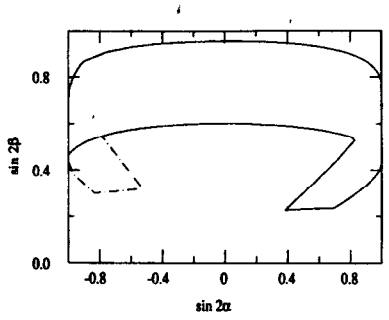


Figure 14: The allowed region in the $\sin 2\alpha - \sin 2\beta$ plane in the SM (solid) and in 2HDM (dot-dashed). From Ref. 120.

• Indirect CP Violation

However, since Δm_D is extremely small in the SM, the induced CP violation is negligible. If new physics were to enhance $D^0 - \bar{D}^0$ mixing, as seen to occur in the previous section for some models, then this mechanism could yield sizable CP violating effects. This interaction between mixing and CP violation in the D meson system has recently received attention in the literature.[123,124]

• Direct CP Violation

Before estimating the typical size of this asymmetry in the SM, we first note that in contrast to B decays, the branching fractions for the relevant modes, *i.e.*, $\pi^+\pi^-, K^+K^-$, etc., are rather sizable in the charm system, and for once, the large effects of final state interactions are welcomed! The size of the CP asymmetry in the SM is estimated[125] to be at most a few $\times 10^{-3}$. The present experimental sensitivity for various modes is in the vicinity of 10% (Ref. 126).

An interesting example of the potential size of CP violating effects from new physics is that of left-right symmetric models.[127] In this case, reasonably large values for CP asymmetries can be obtained for the Cabibbo allowed decay modes. This occurs due to the existence of an additional amplitude from the W_R exchange, which carries a different weak phase from that of the W_L mediated decay. The estimated values of the CP asymmetries in these models is of an order of a few $\times 10^{-2}$. CP asymmetries at the percent level are expected[128] in some nonminimal SUSY models for the decays $D^0 \to K_S^0\pi^0, K_S^0\phi$.

6.7 CP Violation in the Top-Quark Decays

CP violation in top-quark production and decay is expected to be very small in the SM;[129] however, numerous models with new interactions, such as multi-Higgs models and supersymmetry, can give rise to CP violation in the top system at interesting levels. Since the top-quark decays before it has time to hadronize, it provides a particularly good laboratory for the study of such effects. Searches for CP violating effects can be carried out by studying CP-odd spin-momentum correlations in the top-quark decay products. e^+e^- colliders, with polarized beams, are especially suited to carry out such investigations. Numerous studies of CP symmetry tests can be found in Refs. 130 and 131.

7 Conclusion

Rare processes in the kaon sector will be investigated with more precision with the large data sample which will be collected at DAPHNE. In particular, the CP violating parameter ϵ'/ϵ will be explored at the 10^{-4} level. Future runs of the AGS at Brookhaven could increase their total integrated luminosity by a factor of three to six, and hence, finally place the SM prediction for the long sought-after decay $K^+ \to \pi^+\nu\bar{\nu}$ within experimental reach.

A large amount of data on the B-meson system has been and will continue to be acquired during the next decade at CESR, the Tevatron, HERA, the SLAC and KEK B factories, as well as the LHC, and promises to yield exciting new tests of the SM. FCNC processes in the B sector are not as suppressed as in the other meson systems and can occur at reasonable rates in the SM. This is due to a sizable loop-level contribution from the top quark, which results from the combination of the large top mass (giving a big GIM splitting) and the diagonal nature of the CKM matrix. Long-distance effects are expected to play less of a role due to the heavy B mass, and hence rare processes are essentially short-distance dominated. Many classes of new models can also give significant and testable contributions to rare B transitions. The benchmark process for this type of new physics search is the inclusive decay $B \to X_s\gamma$ (and the related exclusive process $B \to K^*\gamma$) which has been recently observed by CLEO.[31] It has since provided strong restrictions on the parameters of several theories beyond the SM. This constitutes the first direct observation of a penguin mediated process and

demonstrates the fertile ground ahead for the detailed exploration of the SM in rare B transitions.

FCNC in the $Q = 2/3$ quark systems will also be explored at a deeper level within the next decade. Increased statistics in the D meson sector will be collected at the e^+e^- B factories and in a possible fixed target run of the Tevatron main injector or at a possible new dedicated heavy flavor experiment for the Tevatron collider. While it is not expected that the data sample will be large enough to reach the miniscule SM rates for the D meson FCNC transitions, important restrictions on new physics can be placed.

And, lastly, the physics of the top quark is just beginning to be explored. In the near future, the Tevatron main injector will produce roughly 7×10^3 $t\bar{t}$ pairs with 1 fb^{-1} of integrated luminosity, while in the longer term, the LHC and NLC will be top-quark factories. Since the top quark is the heaviest SM fermion with a mass at the electroweak symmetry breaking scale, it might provide a unique window to new physics.

In summary, we look forward to an exciting future in heavy flavor physics!

References

[1] For a review and original references, see R. N. Mohapatra, *Unification and Supersymmetry* (Springer, New York, 1986).

[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[3] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[4] R. M. Barnett *et al.* (Particle Data Group), Phys. Rev. D **54**, 1 (1996).

[5] W. J. Marciano in *Proceedings of the 5th International Symposium on Heavy Flavor Physics*, Montreal, Canada, edited by D. I. Britton (Editions Frontieres, France, 1994).

[6] F. C. Barker, Nucl. Phys. A **579**, 62 (1994), and A **540**, 501 (1992); I. S. Towner, Nucl. Phys. A **540**, 478 (1992); E. Hagberg *et al.*, Nucl. Phys. A **509**, 429 (1990); B. A. Brown and W. E. Ormand, Phys. Rev. Lett **62**, 866 (1989).

[7] H. Leutwyler and M. Roos, Z. Phys. C **25**, 91 (1984); M. Bourquin *et al.*, Z. Phys. C **21**, 27 (1983).

[8] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **255**, 297 (1991); J. Bartelt *et al.* (CLEO Collaboration), Phys. Rev. Lett. **71**, 4111 (1993).

[9] J. P. Alexander *et al.* (CLEO Collaboration), Phys. Rev. Lett. **77**, 5000 (1996).

[10] V. Barger, C. S. Kim, and R. J. N. Phillips, Phys. Lett. B **251**, 629 (1990); C. Greub and S.-J. Rey, SLAC-PUB-7245, hep-ph/9608247.

[11] H. Abramowicz *et al.*, Z. Phys. C **15**, 19 (1982); S. A. Rabinowitz *et al.*, Phys. Rev. Lett. **70**, 134 (1993); A. O. Bazarko *et al.*, Z. Phys. C **65**, 189 (1995).

[12] L. Gibbons, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996; A. Buras, *ibid.*

[13] P. Tipton, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996.

[14] J. Adler *et al.* (MARK III Collaboration), Phys. Rev. Lett. **60**, 1375 (1988).

[15] J. Z. Bai *et al.* (BES Collaboration), SLAC-PUB-7147 (1996).

[16] D. Acosta *et al.* (CLEO Collaboration), Phys. Rev. D **49**, 5690 (1993).

[17] A. Aoki *et al.*, Prog. Theor. Phys. **89**, 131 (1993); K. Kodama *et al.* (FNAL E635 Collaboration), Phys. Lett. B **382**, 299 (1996).

[18] J. Z. Bai *et al.* (BES Collaboration), Phys. Rev. Lett. **74**, 4599 (1995).

[19] J. Richman, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996.

[20] L3 Collaboration, CERN-PPE/96-198 (1996).

[21] J. M. Flynn, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996, hep-lat/9611016.

[22] J. Amundson *et al.*, Phys. Rev. D **47**, 3059 (1993); J. Rosner, Phys. Rev. D **42**, 3732 (1990).

[23] S. Capstick and S. Godfrey, Phys. Rev. D **41**, 2856 (1990).

[24] C. A. Dominguez and N. Paver, Phys. Lett. B **318**, 629 (1993).

[25] Y. Grossman, H. Haber, and Y. Nir, Phys. Lett. B **357**, 630 (1995); W.-S. Hou, Phys. Rev. D **48**, 2342 (1993); P. Krawczyk and S. Pokorski, Phys. Rev. Lett. **60**, 182 (1988); J. L. Hewett, in *Proceedings of the 1992 DPF Meeting*, Fermilab, edited by J. Yoh (World Scientific, 1993).

[26] M. Artuso *et al.* (CLEO Collaboration), Phys. Rev. Lett. **75**, 785 (1995); D. Buskulic *et al.* (ALEPH Collaboration), Phys. Lett. B **343**, 444 (1995).

[27] S. L. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977); E. Pachos, Phys. Rev. D **15**, 1966 (1977).

[28] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[29] See, for example, the recent review, G. Buchalla, A. J. Buras, and M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996); as well as F. J. Gilman and M. Wise, Phys. Rev. D **20**, 2392 (1979); *ibid.*, D **27**, 1128 (1983).

[30] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981).

[31] M. S. Alam *et al.*, CLEO Collaboration, Phys. Rev. Lett. **74**, 2885 (1995); R. Ammar *et al.*, CLEO Collaboration, Phys. Rev. Lett. **71**, 674 (1993); K. W. Edwards *et al.*, CLEO Collaboration, presented at the *1995 Meeting of the European Physical Society Conference on High Energy Physics*, Brussels, Belgium, July 1995, CLEO-CONF95-6.

[32] R. Ammar *et al.*, (CLEO Collaboration), presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996, CLEO-CONF-96-05.

[33] For a review of implications of non-SM physics in $B \to X_s\gamma$, see J. L. Hewett, in *The 21st SLAC Summer Institute on Particle Physics*, Stanford, CA, July 1993, hep-ph/9406302.

[34] N. G. Deshpande, P. Lo, J. Trampetic, G. Eilam, and P. Singer, Phys. Rev. Lett. **59**, 183 (1987); S. Bertolini, F. Borzumati, and A. Masiero, Phys. Rev. Lett. **59**, 180 (1987); B. Grinstein, R. Springer, and M. Wise, Phys. Lett. B **202**, 138 (1988); Nucl. Phys. B **339**, 269 (1988); A. Buras, M. Misiak, M. Münz, and S. Pokorski, Nucl. Phys. B **424**, 374 (1994).

[35] A. Ali and C. Greub, Z. Phys. C **49**, 431 (1991); Phys. Lett. B **259**, 182 (1991); Phys. Lett. B **361**, 146 (1995); N. Pott, Phys. Rev. D **54**, 938 (1996).

[36] C. Greub, T. Hurth, and D. Wyler, Phys. Lett. B **380**, 385 (1996); Phys. Rev. D **54**, 3350 (1996); C. Greub and T. Hurth, presented at DPF96, Minneapolis, MN, August 1996, hep-ph/9608449.

[37] K. Adel and Y.-P. Yao, Phys. Rev. D **49**, 4945 (1994).

[38] K. G. Chetyrkin, M. Misiak, and M. Münz, hep-ph/9612313; M. Misiak, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996.

[39] J. Chay, H. Georgi, and B. Grinstein, Phys. Lett. B **247**, 399 (1990); I. Bigi, N. Uraltsev, and A. Vainshtein, Phys. Lett. B **243**, 430 (1992); B. Blok and M. Shifman, Nucl. Phys. B **399**, 441 (1993); I. Bigi *et al.*, Phys. Rev. Lett. **71**, 496 (1993).

[40] P. N. Burrows, presented at the *3rd International Symposium on Radiative Corrections*, Cracow, Poland, August 1996, hep-ex/9612007.

[41] M. Schmelling, presented at the *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996.

[42] J. L. Hewett and J. D. Wells, Phys. Rev. D (in press), hep-ph/9610323.

[43] A. Ali, in *The 20th International Nathiagali Summer College on Physics and Contemporary Needs*, Bhurban, Pakistan, 1995, hep-ph/9606324.

[44] J. L. Hewett, Phys. Lett. B **193**, 327 (1987); W.-S. Hou, A. Soni, and H. Steger, Phys. Lett. B **192**, 441 (1987).

[45] T. G. Rizzo, Phys. Rev. D **38**, 820 (1988); X. G. He, T. D. Nguyen, and R. R. Volkas, Phys. Rev. D **38**, 814 (1988); W.-S. Hou and R. S. Willey, Phys. Lett. B **202**, 591 (1988); C. Q. Geng and J. N. Ng, Phys. Rev. D **38**, 2858 (1988); B. Grinstein, R. Springer, and M. Wise, Nucl. Phys. B **339**, 269 (1990); J. L. Hewett, Phys. Rev. Lett. **70**, 1045 (1993); V. Barger, M. Berger, and R. J. N. Phillips, Phys. Rev. Lett. **70**, 1368 (1993).

[46] V. Barger, J. L. Hewett, and R. J. N. Phillips, Phys. Rev. D **41**, 3421 (1990); A. Buras *et al.*, Nucl. Phys. B **337**, 284 (1990); J. F. Gunion and B. Grzadkowski, Phys. Lett. B **243**, 301 (1990).

[47] S. Bertolini *et al.*, Nucl. Phys. B **294**, 321 (1987), and Nucl. Phys. B **353**, 591 (1991); R. Barbieri and G. F. Giudice, Phys. Lett. B **309**, 86 (1993); R. Garisto and J. N. Ng, Phys. Lett. B **315**, 119 (1993); M. A. Diaz, Phys. Lett. B **322**, 207 (1994); F. M. Borzumati, Z. Phys. C **63**, 291 (1994); S. Bertolini and F. Vissani, Z. Phys. C **67**, 513 (1995); J. L. Lopez *et al.*, Phys. Rev. D **48**, 974 (1993); H. Anlauf, Nucl. Phys. B **430**, 245 (1994); H. Baer and M. Brhlik, hep-ph/9610224.

[48] T. Goto and Y. Okada, Prog. Theor. Phys. **94**, 407 (1995).

[49] G. L. Kane, C. Kolda, L. Roszkowski, and J. D. Wells, Phys. Rev. D **49**, 6173 (1994).

[50] J. L. Hewett and T. G. Rizzo, Phys. Rev. D **49**, 319 (1994).

[51] J. L. Hewett, T. Takeuchi, and S. Thomas, to appear in *Electroweak Symmetry Breaking and Beyond the Standard Model*, edited by T. Barklow *et al.*, SLAC-PUB-7088, hep-ph/9603391.

[52] For a review, see R. Frey, D. Gerdes, and J. Jaros, to appear in *Proceedings of the 1996 DPF/DPB Summer Study on Future Directions for High Energy Physics*, Snowmass, CO, July 1996.

[53] A. Ali, V. M. Braun, and H. Simma, Z. Phys. C **63**, 437 (1994); S. Narrison, Phys. Lett. B **327**, 354 (1994); J. M. Soares, Phys. Rev. D **49**, 283 (1994).

[54] See, for example, D. Atwood, B. Blok, and A. Soni, Int. J. Mod. Phys. A **11**, 3743 (1996); N. G. Deshpande, X.-G. He, and J. Trampetic, Phys. Lett. B **367**, 362 (1996); H.-Y. Cheng, Report IP-ASTP-23-94; J. Milana, Phys. Rev. D **53**, 1403 (1996); G. Ricciardi, Phys. Lett. B **358**, 129 (1995); A. Ali and V. M. Braun, Phys. Lett. B **359**, 223 (1995).

[55] E. Golowich and S. Pakvasa, Phys. Rev. D **51**, 3518 (1995); Z. Ligeti, L. Randall, and M. B. Wise, hep-ph/9702322; A. Khodjamirian *et al.*, hep-ph/9702318.

[56] G. Burdman, E. Golowich, J. L. Hewett, and S. Pakvasa, Phys. Rev. D **52**, 6883 (1995).

[57] C. Greub, T. Hurth, M. Misiak, and D. Wyler, Phys. Lett. B **382**, 415 (1996).

[58] For a review of present experiment results on rare B physics, see, T. E. Browder, K. Honscheid, and S. Playfer, in *B Decays*, 2nd edition, edited by S. Stone (World Scientific, 1995).

[59] S. Abachi *et al.* (DØ Collaboration), DØnote-2023 (1996); F. Abe *et al.* (CDF Collaboration), Fermilab-PUB-96/1040-E; P. Perrodo *et al.* (ALEPH Collaboration), submitted to *28th International Conference on High Energy Physics*, Warsaw, Poland, July 1996; Acciani *et al.* (L3 Collaboration), Phys. Lett. B **363**, 137 (1995); R. Ammar *et al.* (CLEO Collaboration), Phys. Rev. D **49**, 5701 (1994); E. Thorndike *et al.* (CLEO Collaboration), in the *Proceedings of the 27th International Conference on High Energy Physics*, Glasgow, Scotland, July 1994, edited by P. J. Bussey and I. G. Knowles (Institute of Physics, Bristol, 1995).

[60] L. Randall and R. Sundrum, Phys. Lett. B **312**, 148 (1993).

[61] J. L. Hewett, S. Nandi, and T. G. Rizzo, Phys. Rev. D **39**, 250 (1989); M. Savage, Phys. Lett. B **266**, 135 (1991).

[62] S. Davidson, D. Bailey, and B. A. Campbell, Z. Phys. C **61**, 613 (1994); M. Leurer, Phys. Rev. D **50**, 536 (1994), and D **49**, 333 (1994).

[63] B. Grinstein, Y. Nir, and J. M. Soares, Phys. Rev. D **48**, 3960 (1993).

[64] A. J. Buras and M. Münz, Phys. Rev. D **52**, 186 (1995).

[65] N. G. Deshpande, J. Trampetic, and K. Panrose, Phys. Rev. D **39**, 1461 (1989); C. S. Lim, T. Morozumi, and A. I. Sanda, Phys. Lett. B **218**, 343 (1989); Z. Ligeti and M. Wise, Phys. Rev. D **53**, 4937 (1996).

[66] N. G. Deshpande and J. Trampetic, Phys. Rev. Lett. **60**, 2583 (1988); B. Grinstein, M. J. Savage, and M. B. Wise, Nucl. Phys. B **319**, 271 (1989).

[67] A. Ali, G. F. Giudice, and T. Mannel, Z. Phys. C **67**, 417 (1995); A. Ali, T. Mannel, and T. Morozumi, Phys. Lett. B **273**, 505 (1991).

[68] J. L. Hewett, Phys. Rev. D **53**, 4964 (1996).

[69] T. Goto, Y. Okada, Y. Shimizu, and M. Tanaka, hep-ph/9609512.

[70] E. Ma and A. Pramudita, Phys. Rev. D **24**, 2476 (1981).

[71] A. Fryberger *et al.* (CLEO Collaboration), Phys. Rev. Lett. **76**, 3065 (1996); E. M. Aitala *et al.* (E791 Collaboration), Phys. Rev. Lett. **76**, 369 (1996); M. Selen, CLEO Collaboration, presented at *APS Spring Meeting*, Washington D.C., April 1994; T. Alexopoulos *et al.*, E771 Collaboration, Fermilab Report Fermilab-Pub-95-286-E (1995); K. Kodama *et al.* (E653 Collaboration), Phys. Lett. B **345**, 85 (1995); M. Adamovich *et al.* (WA92 Collaboration), Phys. Lett. B **353**, 563 (1995); A. Weir *et al.* (MARK II Collaboration), Phys. Rev. D **41**, 1384 (1990).

[72] A. Acker and S. Pakvasa, Mod. Phys. Lett. A **7**, 1219 (1992).

[73] S. Kettell (E787 Collaboration), hep-ex/9701003.

[74] C. Q. Geng, I. J. Hsu, and Y. C. Lin, Phys. Lett. B **355**, 569 (1995); S. Fajfer, Helsinki University Report HU-SEFT-R-1996-05, hep-ph/9602322.

[75] G. Buchalla and A. Buras, Nucl. Phys. B **412**, 106 (1994), and Phys. Rev. D **54**, 6782 (1996).

[76] I. Bigi and F. Gabbiani, Nucl. Phys. B **367**, 3 (1991).

[77] J. S. Hagelin and L. S. Littenberg, Prog. Part. Nucl. Phys. **23**, 1 (1989).

[78] S. Weinberg, Phys. Rev. Lett. **37**, 657 (1976); P. Krawczyk and S. Pokorski, Nucl. Phys. B **364**, 10 (1991); Y. Grossman and Y. Nir, Phys. Lett. B **313**, 126 (1993); Y. Grossman, Nucl. Phys. B **426**, 355 (1994).

[79] For recent reviews and original references, see L. Littenberg and G. Valencia, Ann. Rev. Nucl. Part. Sci. **43**, 729 (1993); J. L. Ritchie and S. G. Wojcicki, Rev. Mod. Phys. **65**, 1149 (1993); R. Battiston *et al.*, Phys. Rep. **214**, 293 (1992).

[80] G. Eilam, J. L. Hewett, and A. Soni, Phys. Rev. D **44**, 1473 (1991).

[81] B. Grzadkowski, J. F. Gunion, and P. Krawczyk, Phys. Lett. B **286**, 106 (1991); N. G. Deshpande, B. Margolis, and H. Trottier, Phys. Rev. D **45**, 178 (1992); M. Luke and M. Savage, Phys. Lett. B **307**, 387 (1993).

[82] C. S. Li, R. J. Oakes, and J. M. Yang, Phys. Rev. D **49**, 293 (1994); G. Couture, C. Hamzaoui, and H. Konig, Phys. Rev. D **52**, 1713 (1995).

[83] T. Han, R. D. Peccei, and X. Zhang, Nucl. Phys. B **454**, 527 (1995); T. Han *et al.*, Univ. California Davis Report UCD-96-07, hep-ph/9603247; E. Malkawi and T. Tait, Phys. Rev. D **54**, 5758 (1996).

[84] T. J. LeCompte, CDF Collaboration, presented at the *2nd Rencontres du Vietnam*, Ho Chi Minh City, Vietnam, October 1995, Fermilab-CONF-96/021-E.

[85] I. I. Bigi *et al.*, Phys. Lett. B **181**, 157 (1986); V. Barger and R. J. N. Phillips, Phys. Rev. D **41**, 884 (1990), *ibid.* D **40**, 2875 (1990), and Phys. Lett. B **201**, 553 (1988); C. S. Li and T. C. Yuan, Phys. Rev. D **42**, 3088 (1990).

[86] S. Pakvasa and H. Sugawara, Phys. Lett. B **73**, 61 (1978); S. Pakvasa, H. Sugawara, and Y. Yamanaka, Phys. Rev. D **25**, 1895 (1982); T. P. Cheng and M. Sher, Phys. Rev. D **35**, 3484 (1987); L. J. Hall and S. Weinberg, Phys. Rev. D **48**, 979 (1993); A. Antaramian, L. J. Hall, and A. Rasin, Phys. Rev. Lett. **69**, 1871 (1992).

[87] T. P. Cheng and M. Sher, Phys. Rev. D **35**, 3483 (1987); B. Mukhopadhyaya and S. Nandi, Phys. Rev. Lett. **66**, 285 (1991); T. P. Cheng and L. F. Li, Phys. Rev. D **45**, 1708 (1992); W.-S. Hou, Phys. Lett. B **296**, 179 (1992); L. J. Hall and S. Weinberg, Phys. Rev. D **48**, R979 (1993); W.-S. Hou and H. C. Huang, Phys. Rev. D **51**, 5285 (1995).

[88] K. Hikasa and M. Kobayashi, Phys. Rev. D **36**, 724 (1987); H. Baer *et al.*, Phys. Rev. D **44**, 725 (1991); J. Sender, hep-ph/9602354.

[89] K. Agashe and M. Graesser, LBNL Report LBNL-37823 (1995), hep-ph/9510439.

[90] S. Herrlich and U. Nierste, Nucl. Phys. B **419**, 292 (1994); *ibid.* B **476**, 27 (1996), and Phys. Rev. D **52**, 6506 (1995).

[91] A. J. Buras, M. Jamin, and P. H. Weisz, Nucl. Phys. B **347**, 491 (1990).

[92] W. A. Bardeen, A. J. Buras, and J. M. Gerard, Phys. Lett. B **211**, 343 (1988); J. M. Gerard, Acta Phys. Pol. B **21**, 257 (1990); J. Bijnens and J. Prades, Nucl. Phys. B **444**, 523 (1995); A. Pich and E. de Rafael, Phys. Lett. B **158**, 477 (1985); J. Prades *et al.*, Z. Phys. C **51**, 287 (1991); J. F. Donoghue, E. Golowich, and B. R. Holstein, Phys. Lett. B **119**, 412 (1993); R. Decker, *Nucl. Phys.* B Proc. Suppl. A **7**, 190 (1989); S. Narrison, Phys. Lett. B **351**, 369 (1995); V. Antonelli *et al.*, hep-ph/9610230.

[93] For a review, see J. M. Gerard, in *Proceedings of the 12th Warsaw Symposium on Elementary Particle Physics*, Kazimierz, Poland, 1989.

[94] G. Beall, M. Bander, and A. Soni, Phys. Rev. Lett. **48**, 848 (1982).

[95] See, for example, J. Ellis and D. V. Nanopoulos, Phys. Lett. B **110**, 44 (1982).

[96] S. Dimopoulos and J. Ellis, Nucl. Phys. B **182**, 505 (1981); S. Dimopoulos, H. Georgi, and S. Raby, Phys. Lett. B **127**, 101 (1983); R. S. Chivukula and H. Georgi, Phys. Lett. B **188**, 99 (1987); R. S. Chivukula, H. Georgi, and L. Randall, Nucl. Phys. B **292**, 93 (1987); L. Randall, Nucl. Phys. B **403**, 122 (1993).

[97] T. E. Browder and S. Pakvasa, Phys. Rev. D **52**, 3123 (1995).

[98] J. C. Anjos *et al.* (FNAL E691 Collaboration), Phys. Rev. D **60**, 1239 (1988).

[99] E. M. Aitala *et al.* (FNAL E791 Collaboration), Phys. Rev. Lett. **77**, 2384 (1996); G. Blaylock, A. Seiden, and Y. Nir, Phys. Lett. B **355**, 555 (1995).

[100] A. Datta, Z. Phys. C **27**, 515 (1985).

[101] For further details, see G. Burdman, E. Golowich, J. Hewett, and S. Pakvasa, SLAC-PUB-7136; G. Burdman in *CHARM2000 Workshop*, Fermilab, June 1994; J. Donoghue *et al.*, Phys. Rev. D **33**, 179 (1986).

[102] H. Georgi, Phys. Lett. B **297**, 353 (1992); T. Ohl *et al.*, Nucl. Phys. B **403**, 605 (1993).

[103] K. S. Babu *et al.*, Phys. Lett. B **205**, 540 (1988); T. G. Rizzo, Int. J. Mod. Phys. A **4**, 5401 (1989).

[104] S. Pakvasa and H. Sugawara, Phys. Lett. B **73**, 61 (1978); T. P. Cheng and M. Sher, Phys. Rev. D **35**, 3484 (1987); L. Hall and S. Weinberg, Phys. Rev. D **48**, 979 (1993).

[105] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).

[106] Y. Grossman, Z. Ligeti, and Y. Nir, hep-ph/9701231.

[107] V. Kuzmin, V. Rubakov, and M. Shaposhnikov, Phys. Lett. B **155**, 36 (1985); G. Farrar and M. Shaposhnikov, Phys. Rev. Lett. **70**, 2833 (1993), and Phys. Rev. D **50**, 774 (1994); P. Huet and E. Sather, Phys. Rev. D **51**, 379 (1995); A. Cohen, D. Kaplan, and A. Nelson, Ann. Rev. Nucl. Part. Sci. **43**, 27 (1993).

[108] For reviews, see *CP Violation*, edited by C. Jarlskog (World Scientific, Singapore, 1989); Y. Nir and H. Quinn, Ann. Rev. Nucl. Part. Sci. **42**, 211 (1992).

[109] A. B. Carter and A. I. Sanda, Phys. Rev. Lett. **45**, 952 (1980), and Phys. Rev. D **23**, 1567 (1981); I. I. Bigi and A. I. Sanda, Nucl. Phys. B **193**, 85 (1981), and *ibid.* B **281**, 41 (1987).

[110] E. Ma, W. A. Simmons, and S. F. Tuan, Phys. Rev. D **20**, 2888 (1979); L.-L. Chau, W.-Y. Keung, and M. D. Tran, Phys. Rev. D **27**, 2145 (1983).

[111] See, for example, L. Wolfenstein, Comments Nucl. Part. Phys. **21**, 275 (1994).

[112] T. T. Wu and C. N. Yang, Phys. Rev. Lett. **13**, 380 (1964).

[113] G. Buchalla, A. J. Buras, and M. K. Harlander, Nucl. Phys. B **337**, 313 (1990); A. J. Buras, M. Jamin, M. E. Lautenbacher, and P. H. Weisz, Nucl. Phys. B **370**, 69 (1992); addendum, *ibid.*, B **375**, 501 (1992); A. J. Buras, M. Jamin, and M. E. Lautenbacher, Nucl. Phys. B **408**, 209 (1993); M. Ciuchini, E. Franco, G. Martinelli, and L. Reina, Nucl. Phys. B **415**, 403 (1994).

[114] J. M. Flynn and L. Randall, Phys. Lett. B **224**, 221 (1989); Erratum, *ibid.*, B **235**, 412 (1990).

[115] A. J. Buras, M. Jamin, and M. E. Lautenbacher, hep-ph/9608365.

[116] E. Gabrielli, A. Masiero, and L. Silvestrini (1995) hep-ph/9509379 and hep/ph/9510215; J. Hagelin, S. Kelley, and T. Tanaka, Nucl. Phys. B **415**, 293 (1994), and Mod. Phys. Lett. A **8**, 2737 (1993).

[117] B. Winstein and L. Wolfenstein, Rev. Mod. Phys. **65**, 1113 (1993).

[118] D. Boutigny *et al.*, BABAR Collaboration, SLAC Report SLAC-0443 (1994); M. T. Cheng *et al.*, BELLE Collaboration, KEK Report KEK-94-02 (1994); The CLEO III Detector Proposal, Cornell Report CLNS-94/1277 (1994); The CDF Collaboration, CDF-DOC-ADVISORY-PUBLIC/2436; T. Lohse *et al.*, HERA-B Collaboration, DESY Report DESY-PRC-93-04 (1994); K. Kirsebom *et al.*, LHC-B Collaboration, CERN Report CERN/LHCC 95-5 (1995).

[119] M. Gronau and D. London, Phys. Rev. Lett. **65**, 3381 (1990); Y. Nir and H. Quinn, Phys. Rev. Lett. **67**, 541 (1991); N. G. Deshpande and X.-G. He, Phys. Rev. Lett. **74**, 26 (1995), Erratum *ibid.*, **74**, 4099 (1995).

[120] Y. Grossman and Y. Nir, Phys. Lett. B **313**, 126 (1993).

[121] G. C. Branco, G. C. Cho, Y. Kizukuri, and N. Oshimo, Phys. Lett. B **337**, 316 (1994), and Nucl. Phys. B **449**, 483 (1995); M. Wohar Phys. Rev. D **54**, 2198 (1996).

[122] D. Silverman, Int. J. Mod. Phys. A **11**, 2253 (1996).

[123] T. E. Browder and S. Pakvasa, Univ. of Hawaii Report UH-511-828-95 (1995); L. Wolfenstein, Carnegie Mellon Univ. Report CMU-HEP-95-04 (1995); G. Blaylock, A. Seiden, and Y. Nir, Univ. of Santa Cruz Report SCIPP-95-16 (1995).

[124] G. Burdman, E. Golowich, J. L. Hewett, and S. Pakvasa, SLAC Report SLAC-PUB-7136.

[125] F. Bucella *et al.*, Phys. Lett. B **302**, 319 (1993); G. Burdman in *Proceedings of the CHARM2000 Workshop*, Fermilab 1994, edited by D. Kaplan.

[126] P. L. Frabetti *et al.*, Phys. Rev. D **50**, 2953 (1994); J. Bartelt *et al.*, CLEO Collaboration, Phys. Rev. D **52**, 4860 (1995).

[127] A. Le Yaouanc *et al.*, Phys. Lett. B **292**, 353 (1992); S. Pakvasa, in *Proceedings of CHARM2000 Workshop*, Fermilab, June 1994, edited by D. Kaplan; M. Gronau, and S. Wakaizumi, Phys. Rev. Lett. **68**, 1814 (1992).

[128] I. I. Bigi, F. Gabbiani, and A. Masiero, Z. Phys. C **48**, 633 (1990).

[129] See, for example, G. Eilam, J. L. Hewett, and A. Soni, Phys. Rev. Lett. **67**, 1979 (1991), *ibid.*, **68**, 2103 (1992).

[130] For an overview, see R. Frey, talk presented at the *Future e^+e^- Collider Physics Study Group Meeting*, Estes Park, Colorado, June 21-23, 1995. See also, G. Kane, G. A. Ladinsky, and C. P. Yuan, Phys. Rev. D **45**, 124 (1992); C. P. Yuan, Phys. Rev. D **45**, 782 (1992); D. Atwood, A. Aeppli, and A. Soni, Phys. Rev. Lett. **69**, 2754 (1992); D. Atwood, A. Kagan, and T. G. Rizzo, Phys. Rev. D **52**, 6264 (1995); M. Peskin, in *Proceedings of the Second International Workshop on Physics and Experiments at Linear e^+e^- Colliders*, Waikoloa, HI, April 1993, edited by F. A. Harris *et al.* (World Scientific, 1994); M. Peskin and C. R. Schmidt, in *Proceedings of the First Workshop on Linear Colliders*, Saariselkä, Finland, September 1991 (World Scientific, 1992); P. Zerwas, *ibid.*; W. Bernreuther *et al.*, in *Proceedings of the Workshop on e^+e^- Collisions at 500 GeV, The Physics Potential*, (DESY, Hamburg) edited by by P. Igo-Kemenes and J. H. Kühn, 1992; A. Djouadi, ENSLAPP-A-365-92 (1992); T. G. Rizzo, Phys. Rev. D **50**, 4478 (1994); M. Frigeni

and R. Rattazzi, Phys. Lett. B **269**, 412 (1991); R. D. Peccei, S. Persis, and X. Zhang, Nucl. Phys. B **349**, 305 (1991); D. O. Carlson, E. Malkawi, and C.-P. Yuan, Phys. Lett. B **337**, 145 (1994); T. G. Rizzo, Phys. Rev. D **51**, 3811 (1995), and SLAC-PUB-95-6914, 1995.

[131] J. F. Donoghue and G. Valencia, Phys. Rev. Lett. **58**, 451 (1987), and E **60**, 243 (1988); C. A. Nelson, Phys. Rev. D **43**, 1465 (1991); C. R. Schmidt and M. E. Peskin, Phys. Rev. Lett. **69**, 410 (1992); W. Bernreuther *et al.*, Nucl. Phys. B **388**, 53 (1992), and E B **406**, 516 (1993); A. Brandenburg and J. P. Ma, Phys. Lett. B **298**, 211 (1993), and Z. Phys. C **56**, 97 (1992); B. Grzadkowski, Phys. Lett. B **305**, 384 (1993); B. Grzadkowski and W.-Y. Keung, Phys. Lett. B **319**, 526 (1993), and *ibid.* B **316**, 137 (1993); D. Atwood *et al.*, Phys. Rev. Lett. **70**, 1364 (1993); T. Arens and L. M. Sehgal, Nucl. Phys. B **393**, 46 (1993), Phys. Lett. B **302**, 501 (1993), and Phys. Rev. D **50**, 4372 (1994); W. Bernreuther and A. Brandenburg, Phys. Lett. B **314**, 104 (1993), and Phys. Rev. D **49**, 4481 (1994); D. Chang, W.-Y. Keung, and I. Phillips, Nucl. Phys. B **408**, 286 (1993), and B **429**, 255(E) (1994); W. Bernreuther and P. Overmann, PITHA 95/30 (1995).

Techniques for Particle Identification

Tom Ypsilantis
Collége de France
Lab. de Physique Corpusculaire
11 Place Marcelin-Berthelot
F-75231
Paris CEDEX 05
FRANCE

Lecture One: Introduction

Material not available. Please refer to the Summer Institute Prelecture Notes.

Lecture Two: A Long Baseline RICH with a 27-Kiloton Water Target and Radiator for Detection of Neutrino Oscillations

T. Ypsilantis,[(a)*] J. Séguinot,[(a)] and A. Zichichi[(b)]

(a) Collége de France

(b) Università di Bologna

*Lecturer, tel: (004122) 767 7577, fax: (004122) 785 0207, e-mail: thomas.ypsilantis@cern.ch.

©1995 by Tom Ypsilantis.

LECTURE TWO:

A Long Baseline RICH with a 27-Kiloton Water Target and Radiator for Detection of Neutrino Oscillations

T. Ypsilantis,[a]* J. Séguinot,[a] and A. Zichichi[b]

(a) College de France
(b) Università di Bologna

ABSTRACT

A 27 kt water volume is investigated as a target for a long baseline neutrino beam from CERN to Gran Sasso. Charged secondaries from the neutrino interactions produce Cherenkov photons in water which are imaged as rings by a spherical mirror.

The photon detector elements are 14 400 photomultipliers (PM's) of 127 mm diameter or 3600 HPD's of 250 mm diameter with single photon sensitivity. A coincidence signal of about 300 pixel elements in time with the SPS beam burst starts readout in bins of 1 ns over a period of 128 ns.

Momentum, direction, and velocity of hadrons and muons are determined from the width, center, and radius of the rings, respectively. Momentum is measured if multiple scattering dominates the ring width, as is the case for most of the particles of interest.

Momentum resolutions of 1–10%, mass resolutions of 5–50 MeV, and direction resolutions of < 1 mrad are achievable. Thresholds in water for muons, pions, kaons, and protons are 0.12, 0.16, 0.55, and 1.05 GeV/c, respectively.

Electrons and gammas can be measured with energy resolution $\sigma_E/E \approx 8.5\%/\sqrt{E(\text{GeV})}$ and with direction resolution ≈ 1 mrad.

The detector can be sited either inside a Gran Sasso tunnel or above ground because it is directional and the SPS beam is pulsed; thus the rejection of cosmic ray background is excellent.

1. Introduction

The lectures covered our recent work on TEA-Fast RICH counters [1], on CsI–Fast RICH counters [2], and new results on HPD photodetectors for RICH [3]. Since these works are now published, we refer the interested reader to the above references and concentrate here on the final seminar on long baseline RICH (LBL-RICH) [4].

1.1 Long Baseline RICH

The question of neutrino masses and mixing remains one of the most important unsolved problems of particle physics. Experiments in this field use either accelerator neutrinos, solar neutrinos, or atmospheric neutrinos, each sensitive to a different range of neutrino masses and mixing angles. Italy and CERN are now considering a neutrino beam traversing 732 km of earth to arrive at the Laboratorio Nazionale Gran Sasso (LNGS), where long baseline experiments will be installed. The possibility for such experiments was already among the physics goals of the Gran Sasso Project and special care was taken to build the experimental halls aligned towards CERN [5]. The advantage of long baseline neutrino experiments is, of course, their increased sensitivity to small mass differences.

For this purpose, the large water radiator and RICH detector, shown schematically in Fig. 1, was proposed as an experiment at the Gran Sasso laboratory [6]. The water is cheap and safe, and serves both

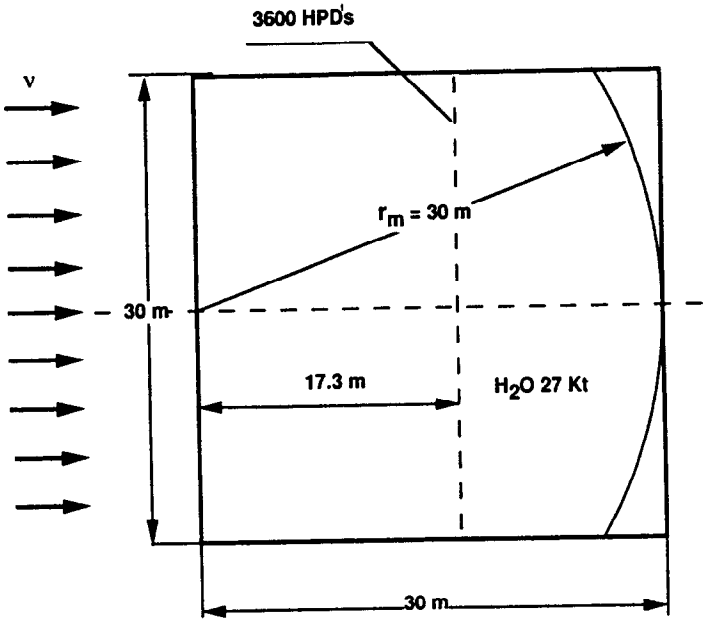


Fig. 1. The layout of the 27 kt water target and radiator between z = 0 to z = 30 m with x = ±15 m, and y = ±15 m. A mirror of curvature r_m = 30 m is at position z = 30 m.

as the neutrino target and the radiating medium. Cherenkov photons, produced by charged particles from neutrino interactions in the water, are detected with visible light photomultipliers (PM's). All hardware elements of this experiment (i.e., clean water, PM's, mirrors, and a swimming pool of Olympic volume) are completely proven and do not require additional R&D.

Of course, most of these elements were (are) present in the pioneering IMB, Kamiokande (and the new 50 kt Super-Kamiokande) detectors which investigate(d) solar neutrinos, atmospheric neutrino interactions, and proton decay.

Our technique, however, differs in one essential aspect, namely, its use of a mirror to give focused images which allow momentum, velocity, and mass determination from Cherenkov rings which are multiple scattering dominated (MSD). Without the mirror, the ring width is determined by track length, which itself is of little direct physical interest but prohibits the observation of multiple scattering which can determine momentum.

1.2 Some Basic Cherenkov Relations

The Cherenkov emission angle θ relative to the particle direction is given by Cherenkov's equation

$$\cos\theta = 1/n\beta, \qquad (1)$$

where n is the radiator refractive index and β the particle velocity. The number of detected photoelectrons N is given by the integral of the Frank-Tamm relation

$$N = N_0 Z^2 L \sin^2\theta, \qquad (2)$$

where L is the particle pathlength in the medium, Ze the particle charge, and N_0 is the detector response parameter defined as

$$N_0 = (\alpha/\hbar c)\int (QTR)\,dE = \left(370\ eV^{-1}cm^{-1}\right) TR q_{int}, \qquad (3)$$

and α is the fine structure constant, E the photon energy, $q_{int} \equiv \int Q dE$ the energy integral of quantum efficiency, T the radiator transmission, and R the mirror reflectivity. A glass window, visible light PM has Q varying from 4 to 28% for E from 2 to 3.5 eV. Integration gives q_{int} = 0.32 eV and for R = 0.95 and T = 1, we find N_0 = 112/cm for full PM coverage and N_0 = 22/cm for 20% coverage. Thus, in water, we expect one detected photoelectron per mm of pathlength [from Eq. (2) for n = 1.34, Z = β = 1, $\sin^2\theta \approx 0.44$, thus N/L $\approx$ 1 /mm].

1.3 Momentum from RICH

It is well-known that a ring image determines particle direction (i.e., the polar and azimuthal angles θ_p, ϕ_p) from the ring center and

particle velocity β from the ring radius [7]. Using $p = m\beta\gamma$ as the defining relation for momentum, we obtain the error

$$\frac{\sigma_p}{p} = \sqrt{\left(\frac{\sigma_m}{m}\right)^2 + \left(\gamma^2 \frac{\sigma_\beta}{\beta}\right)^2}, \qquad (4)$$

where m is assumed to be measured with error σ_m. The velocity resolution obtainable from RICH is $\sigma_\beta/\beta = \tan\theta\sigma_\theta/\sqrt{N}$ with θ the image radius and σ_θ the image width (per photon) [7], thus

$$\frac{\sigma_p}{p} = \sqrt{\left(\frac{\sigma_m}{m}\right)^2 + \left(\gamma^2 \frac{\tan\theta\sigma_\theta}{\sqrt{N}}\right)^2} \quad . \qquad (5)$$

Obviously, m must be measured in order to determine p; moreover, Eq. (5) shows that the error from β (the second term) degrades as γ^2 thus is rapidly limited at high momenta. Neither p or σ_p/p can be found without m; however, to measure m, we must use the same defining relation $m = p/\beta\gamma$ and by the same argument obtain Eq. (5) in a different form, i.e.,

$$\frac{\sigma_m}{m} = \sqrt{\left(\frac{\sigma_p}{p}\right)^2 + \left(\gamma^2 \frac{\tan\theta\sigma_\theta}{\sqrt{N}}\right)^2} \quad . \qquad (6)$$

Clearly, we need an independent measure of p along with β from RICH.

1.4 Multiple Scattering Dominance

When multiple scattering is the dominant angular error (MSD limit), then momentum can be determined from the ring image width [8]. Since this method is new (or sufficiently old that it has been forgotten), we review the technique and its capabilities.

Historically, some early measurements of pion and muon masses in nuclear emulsions were based on this effect. In the experiment of Goldschmidt-Clermont *et al.* [9], secondary particles produced by primary cosmic-ray protons in emulsion were tracked (by human scanners looking through microscopes) and range was measured to find the particle kinetic energy via the Bethe-Bloch relation. They also measured the rms angular deflection of the track to obtain the momentum. Similar results were obtained by Camerini *et al.* [10] from multiple scattering and grain counting (recall that $dE/dx \propto 1/\beta^2$ at low energies).

If the angular width of the ring is multiple scattering (ms) dominated, the quadratic γ dependence of Eqs. (5) and (6) is reduced to first order since

$$\sigma_\theta = \sigma_\theta(\mathrm{ms}) = \frac{k_{ms}}{p\beta}\sqrt{\frac{L}{X_0}} = \frac{k_{ms}}{m\gamma\beta^2}\sqrt{\frac{L}{X_0}}, \qquad (7)$$

where $k_{ms} = (13.6/\sqrt{2})$ MeV = 9.6 MeV and X_0 is the radiation length of the radiator medium [6]. Note that σ_θ is proportional to $1/p$; formally,

this is similar to particle bending in a magnetic field where the bending angle $\propto 1/p$. The strict similarity disappears when we insert the θ dependence of β and the (θ, N) dependence of L.

Combining Eqs. (2), (5), and (7), we obtain the momentum error of an MSD-RICH due to the β error [i.e., the second term of Eq. (5)]

$$\left(\frac{\sigma_p}{p}\right)_\beta = \frac{nK\cos^2\theta}{m\sqrt{\left(\sin^2\theta_n - \sin^2\theta\right)}} , \qquad (8)$$

where $\cos\theta_n = 1/n$ and $K \equiv nk_{ms}/\sqrt{(N_0X_0)}$. Thus, for MSD and m known, Eq. (8) gives the momentum error due to the β error.

1.5 Momentum (without a Magnet) from Ring Radius and Width if L Is Known

In general, momentum is determined from multiple scattering by solving Eq. (7) for $p = [k_{ms}/\beta\sigma_\theta(ms)]\sqrt{(L/X_0)}$. Expressed in terms of the measured variables (θ, σ_θ), we find

$$p = \frac{K'\cos\theta}{\sqrt{\left(\sigma_\theta{}^2 - \sigma_{\theta_0}{}^2\right)}}, \qquad (9)$$

where $K' = nk_{ms}\sqrt{(L/X_0)}$ and we have deduced $\sigma_\theta(ms) \equiv \sqrt{(\sigma_\theta{}^2-\sigma_{\theta_0}{}^2)}$, since σ_θ is the total measured ring width and σ_{θ_0} is the width from all momentum-independent sources. These include chromatic (E), pixel (xyz), emission point (u_e), and impact parameter (v_e, w_e) errors but not multiple scattering (ms) or slowing (sl).

Here K' is required to be a known quantity; thus L must be known (or measured). This is the case for most RICH detectors where L is known from external tracking detectors, but for the LBL-RICH, L is not (and cannot be) directly measured. This case is treated in Sec. 1.6.

The mass defining kinematical relation $m = p/\beta\gamma$ may now be written in terms of the measured variables (θ, σ_θ) as

$$m = nK'\cos\theta\sqrt{\frac{\sin^2\theta_n - \sin^2\theta}{\sigma_\theta{}^2 - \sigma_{\theta_0}{}^2}} . \qquad (10)$$

From Eqs. (9) and (10), we evaluate the momentum and mass errors as

$$\frac{\sigma_p}{p} = \sqrt{\frac{\frac{\varepsilon^2}{2} + (\tan\theta\sigma_\theta)^2}{N}}, \qquad (11)$$

$$\frac{\sigma_m}{m} = \sqrt{\frac{\frac{\varepsilon^2}{2} + \left[\left(\gamma^2+1\right)\tan\theta\sigma_\theta\right]^2}{N}} , \qquad (12)$$

where $\varepsilon \equiv \sigma_\theta{}^2/(\sigma_\theta{}^2-\sigma_{\theta_0}{}^2)$. Note that $\varepsilon \approx 1$ for MSD while $\varepsilon >> 1$ for non-MSD. These derivations use the estimate of the width error [11], i.e.,

$$\sigma_{\sigma_\theta} = \frac{\sigma_\theta}{\sqrt{2N}}. \qquad (13)$$

1.6 Momentum (without a Magnet) from Ring Radius and Width if N (not L) Is Known

Here, we consider the LBL-RICH case when L is not measured but only inferred from Eq. (2) as $L = N/(\varepsilon_g \varepsilon_a N_0 \sin^2\theta)$, where ε_g is the geometrical efficiency for imaging the photons and ε_a is the absorption loss (see Sec. 3.5). The momentum defining relation, Eq. (7), may now be written in terms of the measured variables (θ, σ_θ, N) as

$$p = \left(\frac{K}{\tan\theta}\right)\sqrt{\frac{N}{(\sigma_\theta^2 - \sigma_{\theta 0}^2)\varepsilon_g \varepsilon_a}}, \qquad (14)$$

where again $K = nk_{ms}/\sqrt{(N_0 X_0)}$. The mass equation $m = p/\beta\gamma$ now expressed in terms of the same variables (θ, σ_θ, N) is

$$m = \left(\frac{nK}{\tan\theta}\right)\sqrt{\left(\frac{\sin^2\theta_n - \sin^2\theta}{\sigma_\theta^2 - \sigma_{\theta 0}^2}\right)\frac{N}{\varepsilon_g \varepsilon_a}}; \qquad (15)$$

thus from Eqs. (14) and (15), we obtain for the momentum and mass errors

$$\frac{\sigma_p}{p} = \sqrt{\frac{\frac{1}{4} + \frac{\varepsilon^2}{2} + \left(\frac{\tan\theta\sigma_\theta}{\sin^2\theta}\right)^2}{N}}, \qquad (16)$$

$$\frac{\sigma_m}{m} = \sqrt{\frac{\frac{1}{4} + \frac{\varepsilon^2}{2} + \left[\left(\gamma^2 + \frac{1}{\sin^2\theta}\right)\tan\theta\sigma_\theta\right]^2}{N}}, \qquad (17)$$

where negligible errors are assumed for ε_g and ε_a (see Sec. 3.5). Note that the resolutions of Eqs. (11) and (12) are only marginally better than Eqs. (16) and (17); thus, little is lost by not having tracking detectors inside the water. In fact, we have not found a reasonable way to implement a tracker without seriously compromising the RICH imagery. Luckily, Eqs. (16) and (17) show that the impossible is also unnecessary. Note that for MSD, the first two terms in the numerator of Eq. (16) are 3/4 and the third term is negligible, so that $\sigma_p/p \approx 0.87/\sqrt{N}$. With $N \approx 850$ image points (or even half that many), the $1/\sqrt{N}$ term is small, i.e., 3–5%, thus indicating that good momentum resolution is possible, in principle.

2. Experimental Layout

The layout of Fig. 1 shows the radiator, mirror, and detector array. The mirror center of curvature C fixes the origin (0, 0, 0) of the ZXY coordinate system. The water volume starts at $z = 0$ and extends to $z = 30$ m and transversely to $x = \pm 15$ m, and $y = \pm 15$ m. A spherical mirror of curvature $r_m = 30$ m is placed at the far end of the cube, i.e., at $z = 30$ m.

The optimal detector sphere for tracks uniformly distributed in the water volume ($0 \le z \le r_m$) is at $z = (r_m/2)\sqrt{[1+(3\sin^2\theta/4)]} = 17.3$ m; thus the PM's could be arrayed on a spherical surface 17.3 m into the water target, hence 12.7 m upstream of the mirror. In fact, the PM's will be arrayed on a flat plane at z = 17.3 m because, in the case of a spherical array, too many photons are lost to the side walls. The PM's should cover 20% of the 30 x 30 m^2 surface, i.e., 180 m^2 with a pixel size of 125 mm, i.e., 14 400 PM's on a grid of 250 mm pitch. A coincidence of $\ge$ 100 PM hits in a 128 ns gate during the 6 μs SPS burst window (see Sec. 5.1) will signal an interesting event and start readout. The PM's will be read out with seven-bit TDC's or FADC's for a period of 128 ns in bins of 1 ns, thus increasing the detector granularity to 1.8 Mpixels, quite enough to image events of maximum size $N \le 2 \times 10^4$.

In order to reduce the cost and increase the number of pixels, we are designing 254 mm diameter HPD's [12, 13], each with 36 pads of size 36 x 36 mm^2 at the photocathode surface (and 9 x 9 mm^2 at the silicon wafer plane). A total of 3600 of these HPD's will be needed containing 129 600 pixels. Experiments have shown that $\le$ 1 ns timing can be obtained from HPD's if the pad signals are independently brought out of the vacuum envelope and treated by conventional fast electronics [14].

The mirror radius is chosen relatively small (r_m = 30 m, $f = r_m/2$ = 15 m) so that the image radius [$r_{image} \approx f\theta \approx$ (15 m) (0.72) = 10.8 m] is mostly contained (with geometric efficiency ε_g) inside the $\pm$15 m PM (or HPD) array. Because the detector array is 80% transparent, the water volume can be extended into the good optics region 17.3 m upstream of the detector surface. Photons from this region will be detected with 16% effective coverage rather than 20%.

A hadron track of pathlength 1λ (one absorption length in water is 850 mm) will make an image with $N \approx 850\varepsilon_g\varepsilon_a$ hit points. The latter factor $\varepsilon_a = e^{-\mu\ell}$ is the efficiency for photon transmission in pathlength ℓ of water ($\langle\ell\rangle \approx$ 30 m in the LBL-RICH). Water transparency with $\mu^{-1} >$ 100 m has been attained for 3.9 eV photons [15], whereas the LBL-RICH photon detection range is only from 2.5 to 3.5 eV.

The momentum range for MSD extends up to about 4.5 GeV/c. This range includes almost all hadrons produced by 1 to 20 GeV neutrinos via quasi-elastic (QEL) and deep inelastic scattering (DIS) via charged and neutral current interactions. The threshold momentum for Cherenkov radiation in water is p = 1.12m, hence 0.12, 0.16, 0.55, and 1.05 GeV/c for muons, pions, kaons, and protons, respectively. About 25% of the protons from quasi-elastic interactions are above the proton threshold. Generally, all above threshold hadrons will have their direction, momentum, velocity, mass, and $(Ze)^2$ measured in the LBL-RICH.

Electrons and gammas can also be measured because EM shower electrons in water (X_0 = 36 cm) radiate if p > 0.57 MeV/c. A Cherenkov sensitive shower is therefore somewhat shorter than a dE/dx sensitive shower and is less affected by low-energy fluctuations. Since it is fully contained in about 5 m, we take 25 m as the fiducial target length thus

defining the LBL–RICH fiducial mass as 22.5 kt. The shower produces a more diffuse, but still identifiable, ring ($\sigma_\theta \approx 50$ mrad) than a hadron ring ($\sigma_\theta \leq 15$ mrad) but with many more points, i.e., $N \approx 3000\, \varepsilon_g \varepsilon_a E_e$ (GeV). The water acts as the showering medium and radiator of a homogeneous Cherenkov calorimeter with resolution $\sigma_{Ee}/E_e \approx 8.5\%/\sqrt{E_e}$ (GeV). The direction of the particle initiating the EM shower is accurately determined by the ring center to better than 1 mrad.

Muon identification is obtained for $p \leq 1.5$ GeV/c by direct measurement of β and p (from MSD) which determines m with error σ_m from 3 to 6 MeV, sufficient to distinguish muons and pions. Above 1.5 GeV/c, σ_m becomes larger than 10 MeV, and muon identification by this method is not possible, but then the muon range is so long that the image "lights up like a muon sign." For example, a 1.1 GeV/c muon with 5 m range in water will make a ring with 5000 $\varepsilon_g \varepsilon_a$ image points compared to 800 $\varepsilon_g \varepsilon_a$ for a 1λ pathlength pion. The measurement of p for muons is also good, i.e., $\sigma_p/p \leq 6\%$ for $p \leq 4.5$ GeV/c. It becomes limited by emission point errors $\sigma_\theta(z_e)$ due to the long muon pathlength in water. Extension to higher momentum (i.e., $\sigma_p/p \approx 10\%$ for p = 15 GeV/c) can be attained by time slicing the track into a series of shorter segments.

The capability of the LBL-RICH is such that it can explore values of L/E_ν between 50 and 700 km/GeV with a broad band neutrino beam of energy E_ν between 1 and 15 GeV at Gran Sasso (L = 732 km). In neutrino disappearance ($\mu \rightarrow \tau$) and neutrino appearance ($\mu \rightarrow e$) experiments with broad band beams, it is essential to accurately determine E_ν. This is possible in LBL-RICH because the direction and energy of muons and electrons are well-measured as, indeed, are the hadrons.

2.1 Particle Momentum Resolution

The contributions to the angular error σ_θ vs impact parameter ρ_e (relative to C) for a 1 GeV/c pion with an 850 mm pathlength in water are shown in Fig. 2 for the detector layout of Fig. 1. Note that the dominant contribution is from multiple scattering $\sigma_\theta(ms)$ = 15 mrad, while chromatic $\sigma_\theta(E)$ = 3.6 mrad, pixel $\sigma_\theta(xyz)$ = 1.9 mrad, and slowing $\sigma_\theta(sl)$ = 0.4 mrad are less important. Only the impact parameter errors $\sigma_\theta(v_e)$, $\sigma_\theta(w_e)$, and emission point error $\sigma_\theta(u_e)$ vary with $\rho_e = \sqrt{(v_e^2+w_e^2)}$, but they are not significant even for ρ_e as large as 15 m.

We have evaluated the resolutions σ_p/p and σ_m for $\theta_p = 0$ tracks and pixel sizes $\Delta x = \Delta y = 125$ mm, $\Delta z = 1$ mm, $\Delta u_e = 850$ mm, $\Delta v_e = \Delta w_e = 100$ mm, and $\Delta\theta_p = \Delta\phi_p = 1$ mrad. The refractive index and dispersion n(E) of water were obtained from Ref. [16].

Figure 3 shows the resolution σ_p/p vs p for 15 m pathlength muons or for 0.85 m pathlength hadrons (π, K, P). The solid curves are from multiple scattering [Eq. (16)] while the dot-dash curves are from velocity β when m is known [(Eq. (8)]. Note that the solid curves are everywhere excellent, i.e., $1 < \sigma_p/p < 6\%$ for $p \leq 5$ GeV/c. For K's and P's, the dot-dash curves are everywhere < 1% and better than the solid

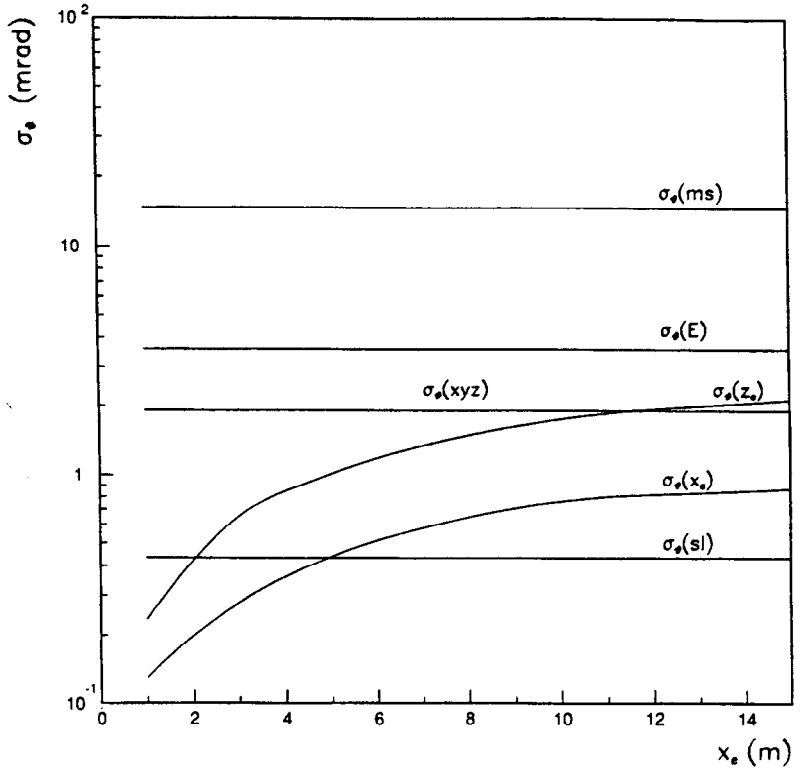


Fig. 2. The Cherenkov angle width vs impact parameter ρ_e for a 1 GeV/c pion track 85 cm long in water. The contributions shown are σ_θ(ms) from multiple scattering, σ_θ(E) chromatic, σ_θ(xyz) from pixel size, σ_θ(sl) from energy loss, $\sigma_\theta(z_e)$ from tracklength, and $\sigma_\theta(x_e)$ from impact parameter.

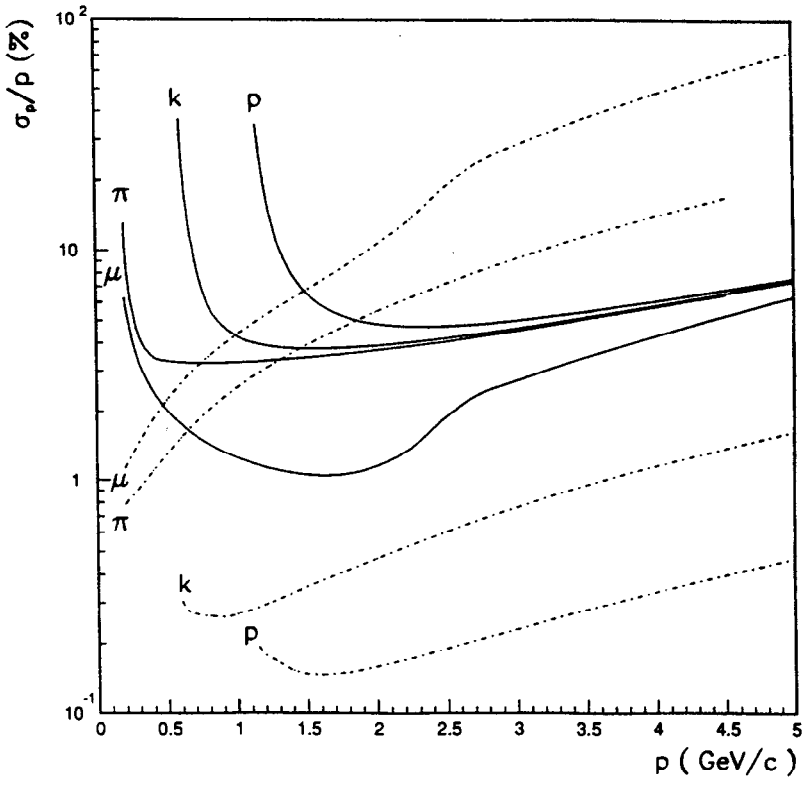


Fig. 3. The resolution σ_p/p vs p for (μ, π, K, P) in water and the geometry of Fig. 1. The solid curves are from multiple scattering [Eq. (16)], whereas the dot-dash curves are from the β measurement [Eq. (8)] assuming mass is known.

curves, whereas for π's and μ's, they are only better for $p < 1.3$ and 0.5 GeV/c, respectively.

Figure 4 shows the mass resolution σ_m/m from combined σ_θ and β measurements, i.e., Eq. (17). For $p \leq 1.25$ GeV/c, the resolution $\sigma_m \approx 5$–7 MeV is sufficient for μ/π identification. Above 1 GeV/c, the muon pathlength becomes so long that the muon is identified by its large N (i.e., $N > 5000\,\varepsilon_g\varepsilon_a$). The K mass resolution, $\sigma_m \approx 20$–30 MeV for $p < 5$ GeV/c, is sufficient for π/K identification and the P mass resolution, $\sigma_m \approx 50$–60 MeV for $p < 5$ GeV/c, is also sufficient for K/P identification. Therefore, the combined σ_θ and β measurements determine the identification of all stable particles and allows us to choose the best resolution curves of Fig. 3 (solid or dot-dash).

2.2 Particle Direction Determination

The polar angles (θ_p, ϕ_p) of a particle producing a ring image are determined with high precision from the ring center, i.e., $\sigma_{\theta p} \approx \sigma_{\phi p} \approx \sigma_\theta/\sqrt{N}$. For a 1 GeV/c pion track in water with Cherenkov pathlength of 85 cm, we have $\sigma_\theta \approx 15$ mrad and for $N \approx 400$, then $\sigma_{\theta p} \approx \sigma_{\phi p} \approx 0.75$ mrad. The direction error for electrons, gammas, and muons should be at least as good because N is considerably larger.

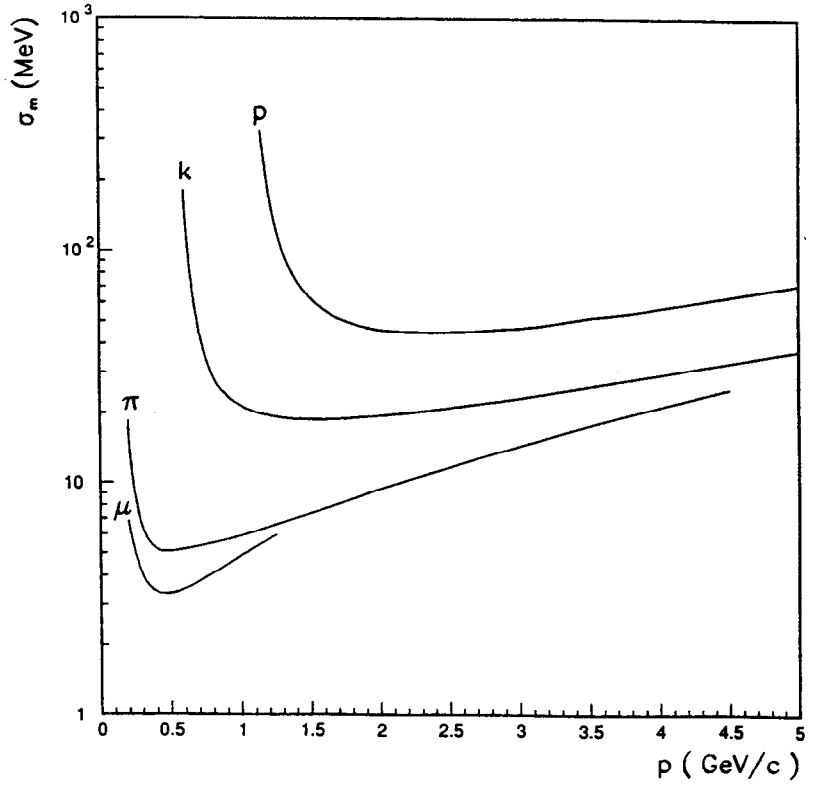


Fig. 4. The mass resolution σ_m vs p for (μ, π, K, P) in water and geometry of Fig. 1. The solid curves [Eq. (17)] are from combined measurements of multiple scattering and β.

2.3 Quasi-Elastic Neutrino Interactions

Here, we show events due to 12 GeV/c neutrinos interacting quasi-elastically to produce leptons (ℓ) via the process ($\nu_\ell + N \rightarrow \ell^- + P$) at random points in the water volume (see Fig. 1). The events were obtained from a PYTHIA 5.7 simulation with all fast decays allowed. They were subsequently introduced into GEANT to simulate the tracks with multiple scattering, energy loss, secondary interactions, and Cherenkov light emission. The PM hit points are labeled as muons, electrons (or gammas), protons, and pions. All images contain only 10% of the expected photoelectron hit points because of computer memory limitations.

- In Figs. 5–7, we show three successive events of the type $\nu_\mu + N \rightarrow \mu^- + P$.
- Figures 8–10 show three successive events of the type $\nu_e + N \rightarrow e^- + P$. The electrons were allowed to interact; thus the images shown are due to showers.
- Figures 11–14 show four successive events of the type $\nu_\tau + N \rightarrow \tau^- + P$. The τ^-'s were allowed to decay naturally via the dominant e^-, μ^-, ρ^-, or π^- branching modes.

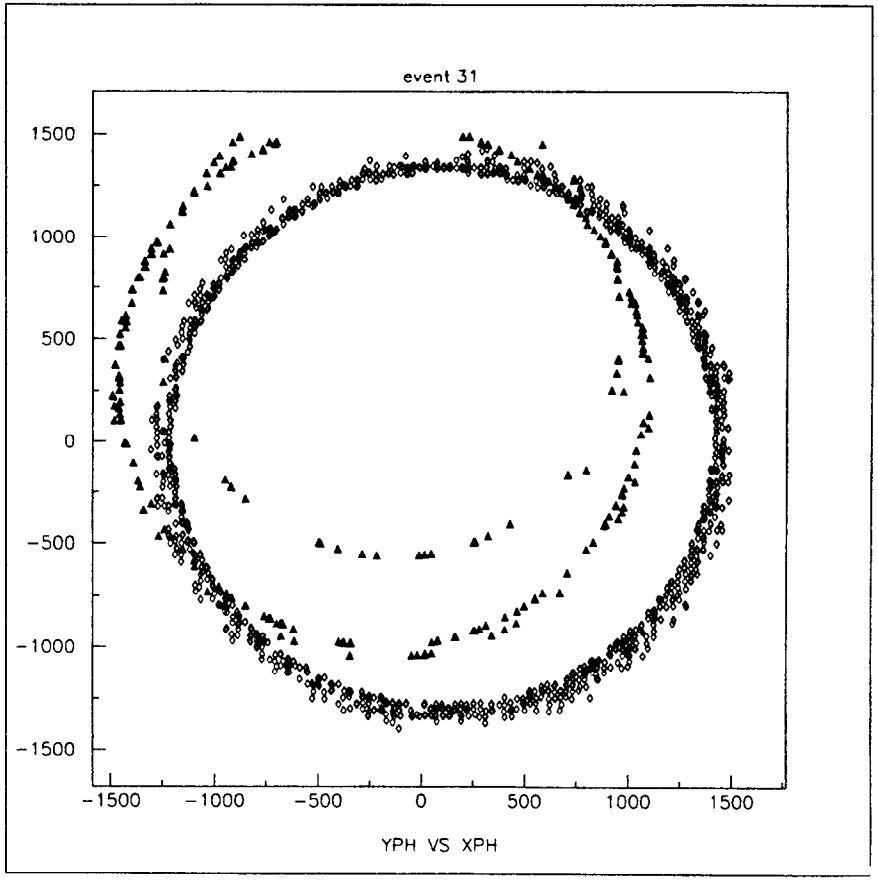


Fig. 5. A Monte Carlo simulation of a quasi-elastic event (#31) $\nu_\mu + N \rightarrow \mu^- + P$ for $E_{\nu\mu} = 12$ GeV. It has two proton rings (black triangles) (the smaller one is due to a scatter) and one very dense muon ring (open diamonds). Muon identification here is obvious. The diffuseness of the image is due to the long muon pathlength; thus emission point errors dominate. This effect can be removed by time slicing the image (thus breaking the track up into a series of shorter segments) and reconstructing each segment.

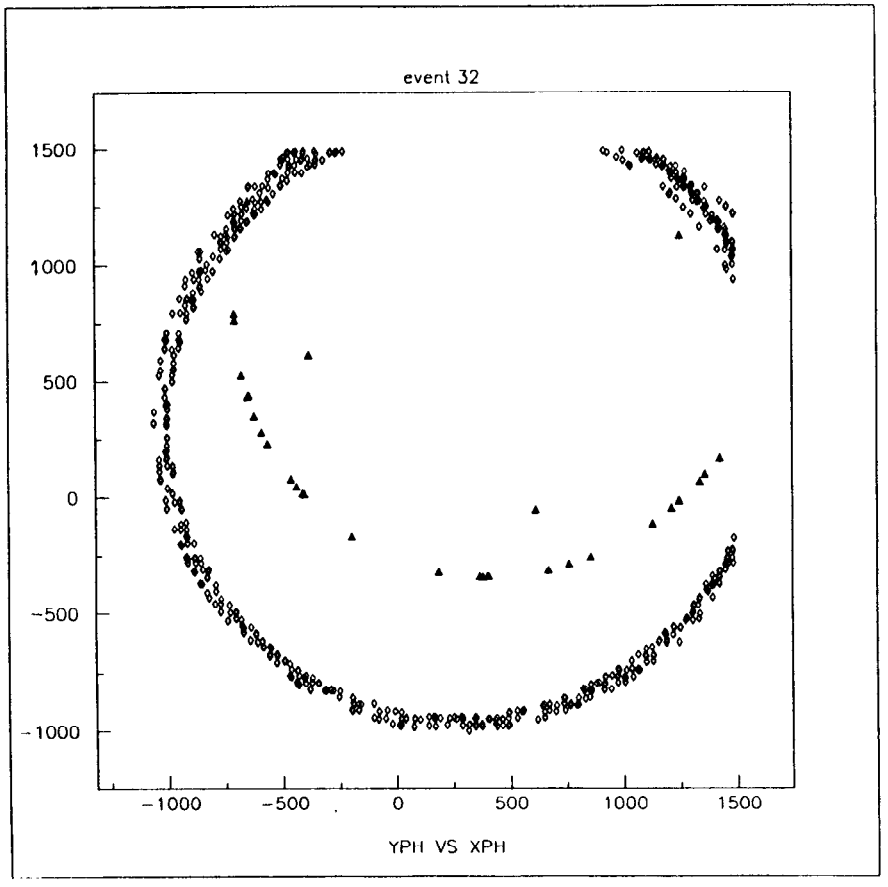


Fig. 6. A Monte Carlo simulation of a quasi-elastic event (#32) ν_μ + N $\rightarrow \mu^-$ + P for $E_{\nu\mu}$ = 12 GeV. It shows one proton ring (black triangles) with a hint of a second (it is evident if all N hits are plotted) along with a self-evident muon ring (open diamonds).

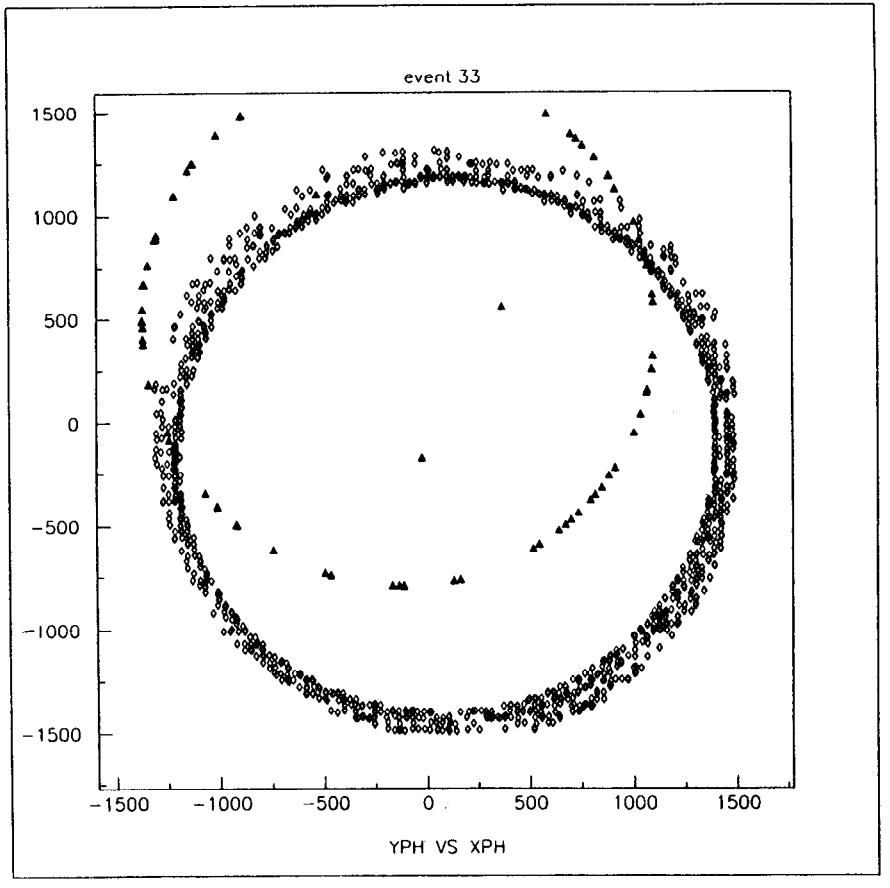


Fig. 7. A Monte Carlo simulation of a quasi-elastic event (#33) ν_μ + N $\rightarrow \mu^-$ + P for $E_{\nu\mu}$ = 12 GeV. It shows one proton ring (black triangles) with some extra hits (a second proton ring due to a scatter) and one muon ring (open diamonds). The rings are reasonably easy to identify (by eye) and so the pattern recognition algorithm will surely work.

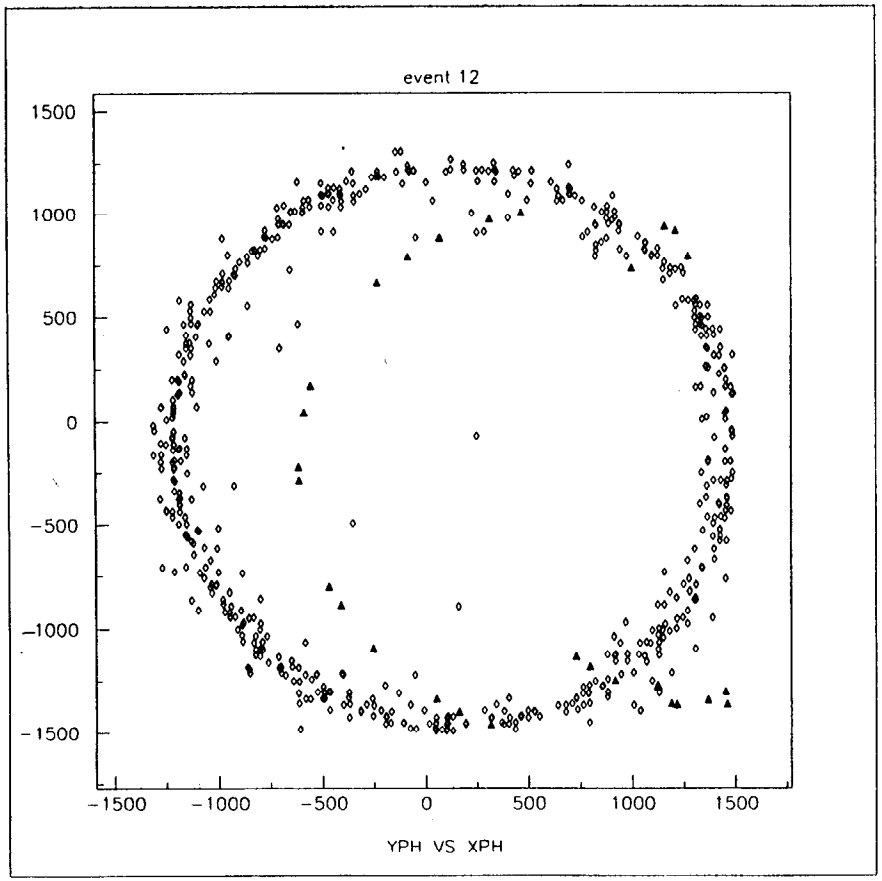


Fig. 8. A Monte Carlo simulation of a quasi-elastic event (#12) $\nu_e + N \rightarrow e^- + P$ with $E_{\nu e} = 12$ GeV. It has one proton ring (black triangles) (plus a scatter which may be resolvable) and one dense electron ring (open diamonds). Electron identification and energy measurement should be good.

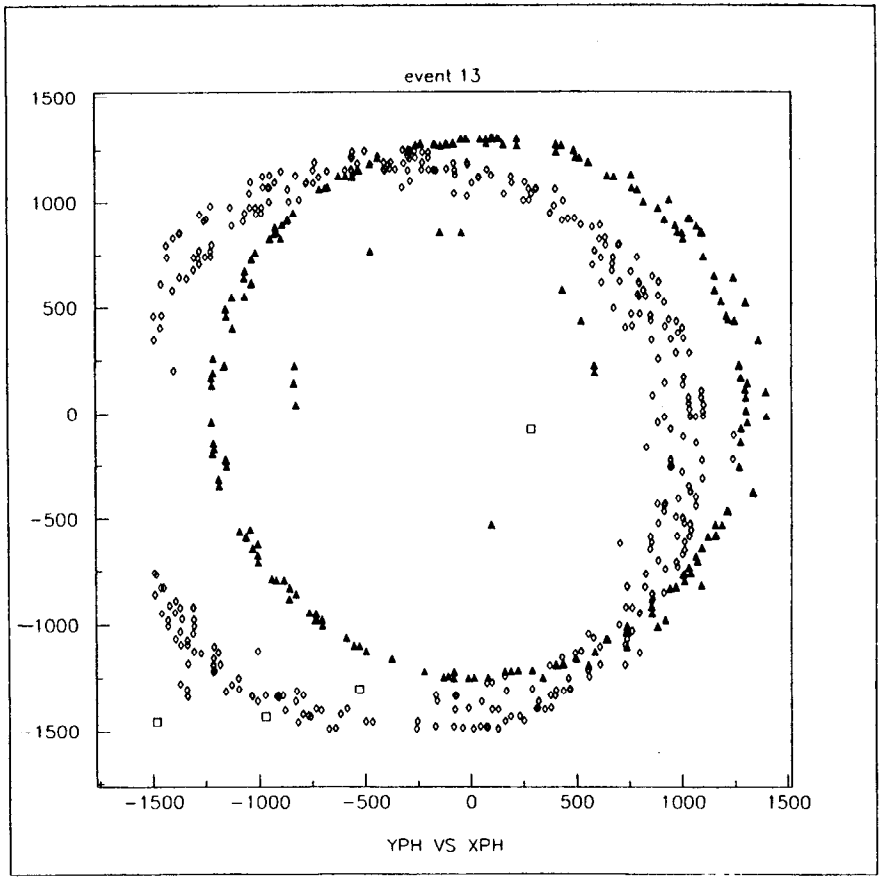


Fig. 9. A Monte Carlo simulation of a quasi-elastic event (#13) $\nu_e + P \rightarrow e^- + \Delta^{++}$ and $\Delta^{++} \rightarrow P + \pi^+$ for $E_{\nu e} = 12$ GeV. It shows two proton rings (black triangles) and one electron ring (open diamonds). There is a hint of a pion ring (open squares) which is evident if all N hits are plotted.

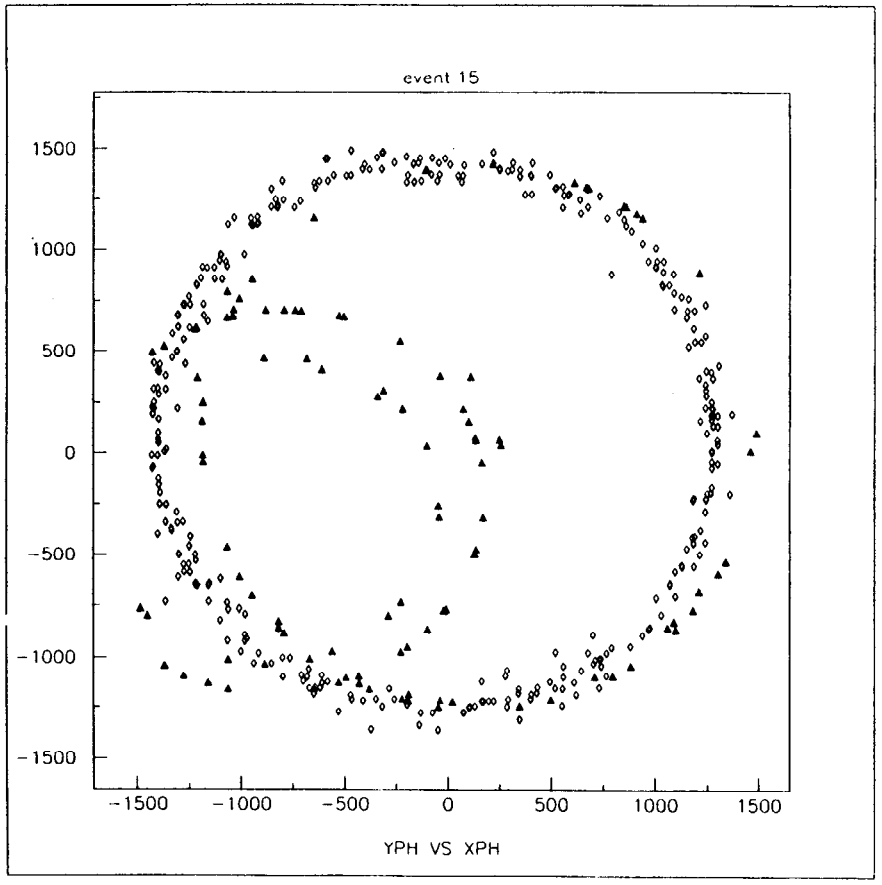


Fig. 10. A Monte Carlo simulation of a quasi-elastic event (#15) $\nu_e + N \rightarrow e^- + P$ for $E_{\nu e} = 12$ GeV. It has three proton rings (black triangles) (the smaller ones are due to scatterings) and one electron ring (open diamonds).

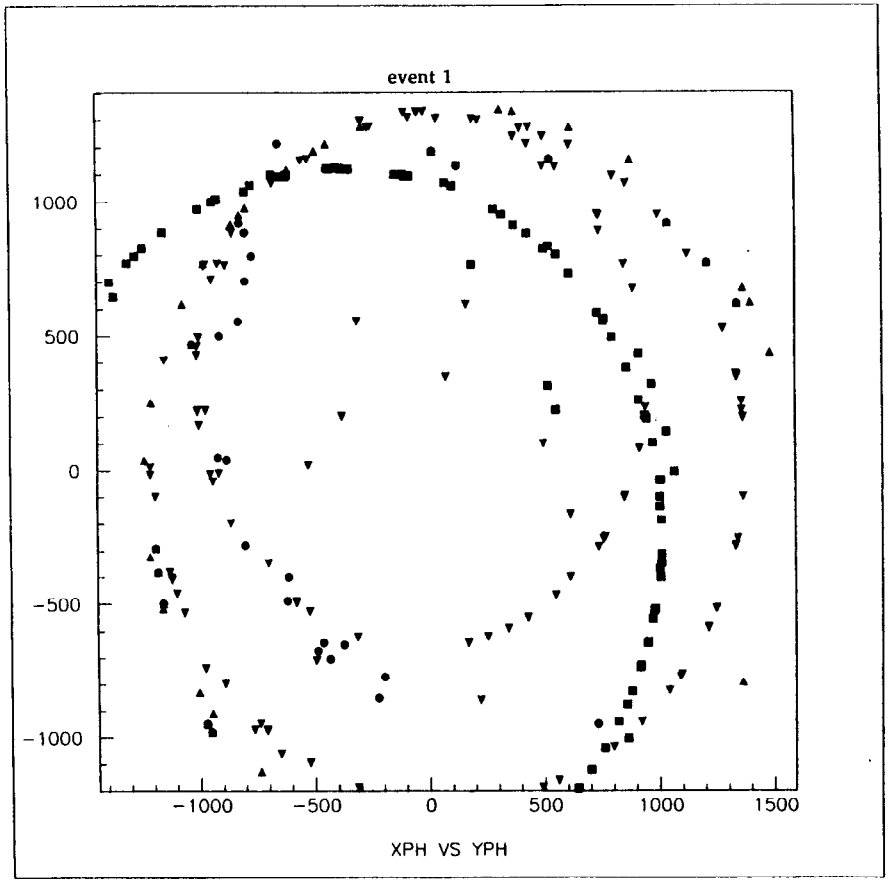


Fig. 11. A Monte Carlo simulation of a quasi-elastic event (#1) $\nu_\tau + N \rightarrow \tau^- + P$ for $E_{\nu\tau} = 12$ GeV. It shows two proton rings (black triangles) and one electron ring (black circles), and one or two pion rings (black squares). Clearly, this event would be challenging.

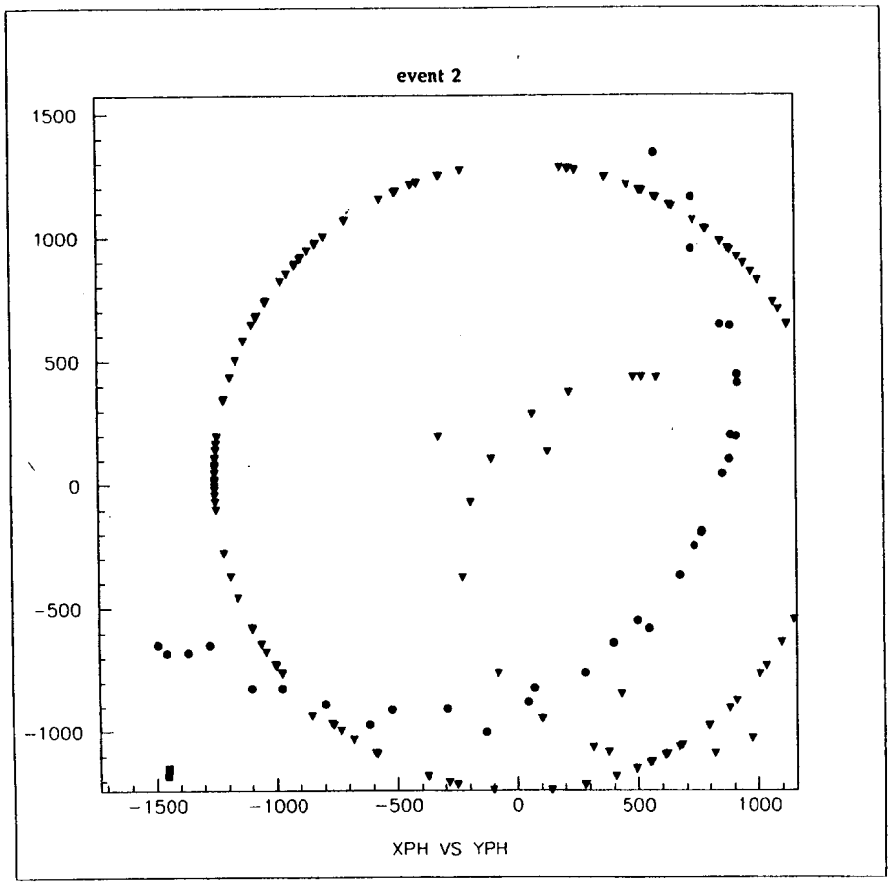


Fig. 12. A Monte Carlo simulation of a quasi-elastic event (#2) $\nu_\tau + N \rightarrow \tau^- + P$ for $E_{\nu\tau} = 12$ GeV. It shows two proton rings (black triangles) and one electron ring (black circles).

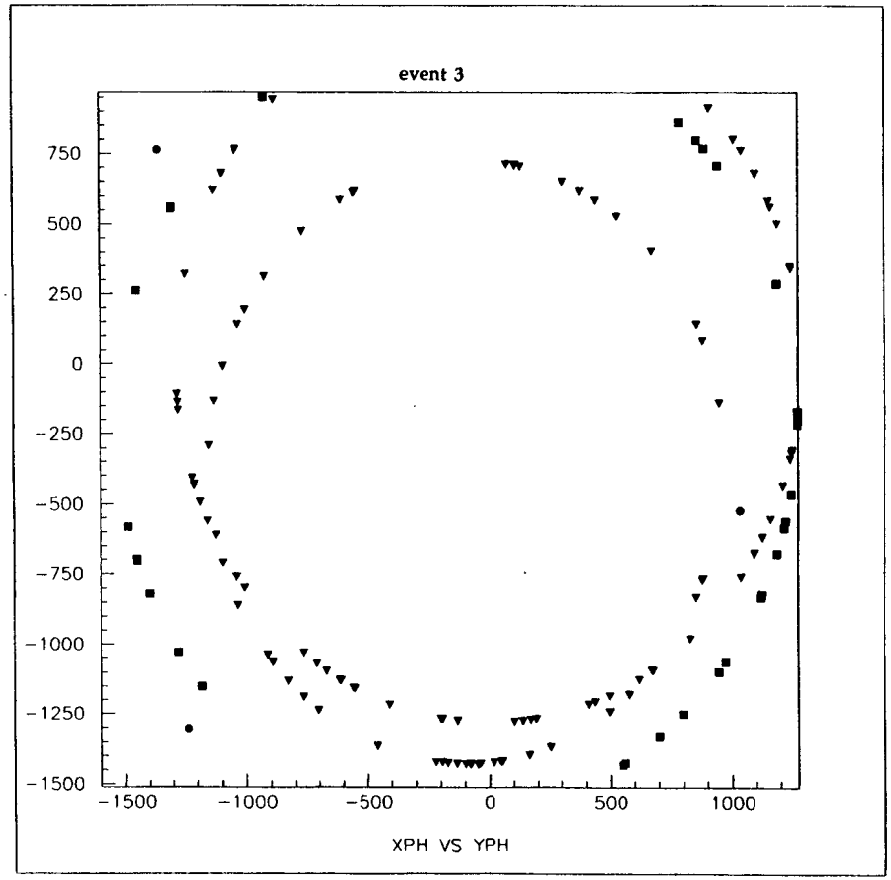


Fig. 13. A Monte Carlo simulation of a quasi-elastic event (#3) $\nu_\tau + N \rightarrow \tau^- + P$ for $E_{\nu\tau} = 12$ GeV. It shows two proton rings (black triangles) and one pion ring (black squares) with a hint of an electron ring (black circles).

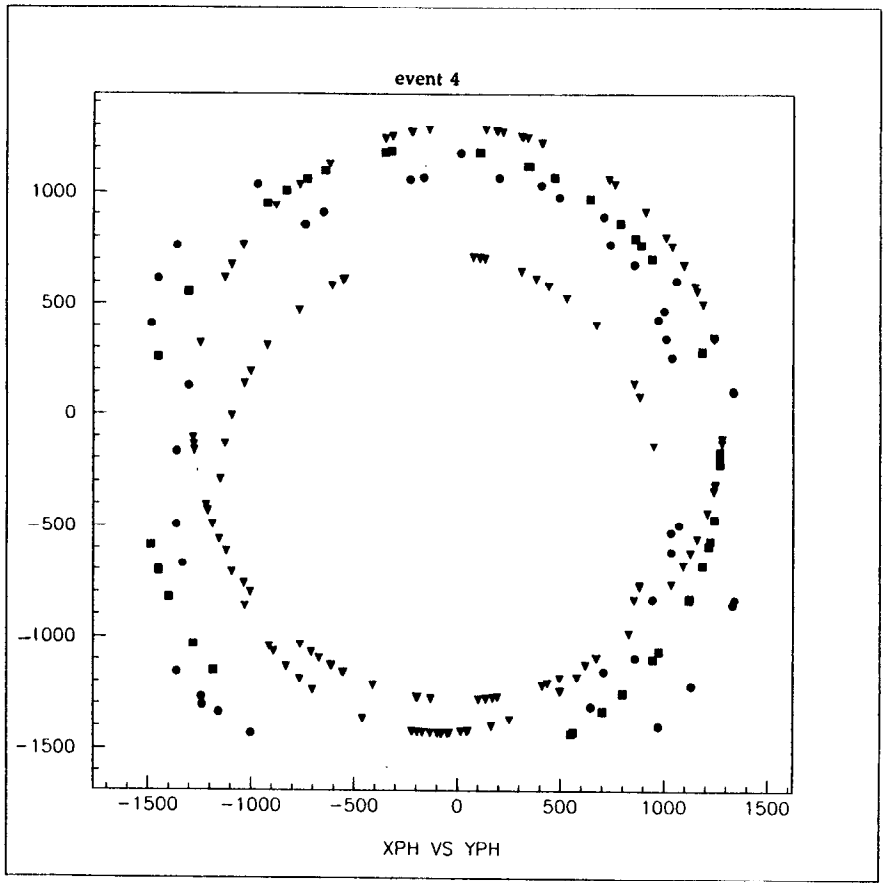


Fig. 14. A Monte Carlo simulation of a quasi-elastic event (#4) $\nu_\tau + N \rightarrow \tau^- + P$ for $E_{\nu\tau} = 12$ GeV. It shows two proton rings (black triangles) and two electron rings (black circles), and one pion ring (black squares). Clearly, this event would be challenging.

3. Properties of Ring Images

Every charged particle above Cherenkov threshold makes a ring image. Neutrals which decay into charged pairs also form images. Electrons and gammas will shower and produce somewhat more diffuse images.

Because Cherenkov light rays form parallel bundles in all planes containing the track and because spherical mirrors focus parallel bundles to a point on the focal surface, hence the contributions from all planes combine to form a ring. This means that parallel tracks form the same ring and that the ring center determines the particle direction (θ_p, φ_p) [7].

3.1 Parameters of the Image

The ring image is characterized by nine parameters: three detected photon coordinates (z, x, y); and five track parameters, i.e., the photon emission point (u_e, v_e, w_e) and the particle direction (θ_p, φ_p), and a single parameter for photon energy E. The photon emission point u_e is measured along the particle track, and the impact parameter $\rho_e = \sqrt{(v_e^2+w_e^2)}$ is the perpendicular distance to the track from the mirror center of curvature C.

We define two different coordinate systems (see appendix Fig. A1), the ZXY system, fixed relative to the mirror and the water tank (Fig. 1), with unit vectors $(\vec{k}, \vec{i}, \vec{j})$ and the PQR system, fixed to each track and

defined from C such that P is parallel, Q and R are normal to the track, with unit vectors $(\bar{a}_p, \bar{b}_p, \bar{c}_p)$. The photon emission point is (z_e, x_e, y_e) in the ZXY system and (u_e, v_e, w_e) in the PQR system. The photon detection point is (z, x, y) in the ZXY system and (u, v, w) in the PQR system. A particle tracker, if available, would measure (u_e, v_e, w_e), whereas the photon detector will measure (z, x, y). Note that (z, x, y) and (z_e, x_e, y_e) are independent of (θ_p, φ_p) while (u, v, w) and (u_e, v_e, w_e) depend on (θ_p, φ_p).

Seven of the nine variables, i.e., $\zeta_i = (z, x, y, v_e, w_e, \theta_p, \varphi_p)$, can be determined with arbitrary precision. The group (z, x, y) depend on the accuracy of the photon detector while $(v_e, w_e, \theta_p, \varphi_p)$ depend on the accuracy of the presumed tracker. In case of the LBL-RICH, these are found by the procedure developed below. The error in θ, due to errors in the variables ζ_i, may be expressed as

$$\sigma_\theta(\zeta_i) = \left(\frac{\partial\theta}{\partial\zeta_i}\right)\sigma_{\zeta_i}, \tag{18}$$

where $\partial\theta/\partial\zeta_i$ is calculated from the reconstruction relation $\theta = \theta(\zeta_i)$. Analytic forms for $\sigma_\theta(\zeta_i)$ evaluated from $\theta = \theta(\zeta_i)$ may be found in Ref. [7] and more generally in Appendix A.

In contrast, the photon emission point u_e along the track has an intrinsic error

$$\sigma_{u_e} = \frac{\Delta u_e}{\sqrt{12}} \tag{19}$$

limited by the radiator pathlength Δu_e; however, in a focused system $\partial\theta/\partial u_e = 0$ (or is small even for large values of ρ_e); thus $\sigma_\theta(u_e)$ is never dominant. Similarly, the photon energy error for a square detector response is

$$\sigma_E = \frac{\Delta E}{\sqrt{12}} \quad , \tag{20}$$

which can be reduced only by reducing the detector energy bandwidth ΔE. The corresponding Cherenkov angle error is

$$\sigma_\theta(E) = \left(\frac{\partial\theta}{\partial n}\right)\left(\frac{dn}{dE}\right)\sigma_E, \tag{21}$$

where n(E) is the radiator dispersion function. These errors define the limits of the RICH resolution.

An important advantage of RICH is that the Cherenkov angle distributions are Gaussian, without the Landau tail which characterizes dE/dx (energy loss) detectors.

3.2 Impact Parameter, Vertex, and Emission Point Vectors and Cherenkov Angle

The unit vector $\bar{a}_p$ along the track, parallel to P (see Fig. A1), has ZXY components

$$\begin{aligned}
&\vec{a}_p = a_{pz}\vec{k} + a_{px}\vec{i} + a_{py}\vec{j},\\
&a_{pz} = \cos\theta_p,\\
&a_{px} = \sin\theta_p \cos\phi_p,\\
&a_{py} = \sin\theta_p \sin\phi_p.
\end{aligned} \tag{22}$$

The unit vector $\vec{b}_p$ along Q has components in ZXY

$$\begin{aligned}
&\vec{b}_p = b_{pz}\vec{k} + b_{px}\vec{i} + b_{py}\vec{j},\\
&b_{pz} = -a_{px},\\
&b_{px} = \sin^2\phi_p + \cos\theta_p \cos^2\phi_p \equiv \varepsilon_x,\\
&b_{py} = (\cos\theta_p - 1)\cos\phi_p \sin\phi_p \equiv \eta.
\end{aligned} \tag{23}$$

The unit vector $\vec{c}_p = \vec{a}_p \times \vec{b}_p$ along R is normal to $\vec{a}_p$ and $\vec{b}_p$ with ZXY components

$$\begin{aligned}
&\vec{c}_p = c_{pz}\vec{k} + c_{px}\vec{i} + c_{py}\vec{j},\\
&c_{pz} = -a_{py},\\
&c_{px} = \eta,\\
&c_{py} = \cos^2\phi_p + \cos\theta_p \sin^2\phi_p \equiv \varepsilon_y.
\end{aligned} \tag{24}$$

Thus, $(\vec{k}, \vec{i}, \vec{j})$ and $(\vec{a}_p, \vec{b}_p, \vec{c}_p)$ are the unit vectors of the ZXY and PQR coordinate systems, respectively. The specific choice of QR axes is made such that $\vec{a}_p \to \vec{k}$, $\vec{b}_p \to \vec{i}$, $\vec{c}_p \to \vec{j}$ as $\theta_p \to 0$.

The particle production (neutrino interaction) vertex $\vec{r}_v$ has ZXY components (z_v, x_v, y_v) and PQR components (u_{va}, v_e, w_e), where

$$\begin{aligned}
&\vec{r}_v = z_v\vec{k} + x_v\vec{i} + y_v\vec{j} = u_{va}\vec{a}_p + v_e\vec{b}_p + w_e\vec{c}_p,\\
&u_{va} = \vec{r}_v \cdot \vec{a}_p,\\
&v_e = \vec{r}_v \cdot \vec{b}_p,\\
&w_e = \vec{r}_v \cdot \vec{c}_p.
\end{aligned} \tag{25}$$

The impact parameter $\vec{r}_{ip}$ is that vector normal to $\vec{a}_p$ which when added to a vector proportional to $\vec{a}_p$ gives the vertex vector $\vec{r}_v$; thus

$$\begin{aligned}
&\vec{r}_{ip} = \vec{r}_v - (\vec{r}_v \cdot \vec{a}_p)\vec{a}_p = v_e\vec{b}_p + w_e\vec{c}_p = z_{ip}\vec{k} + x_{ip}\vec{i} + y_{ip}\vec{j};\\
&z_{ip} = \sin\theta_p\left[z_v \sin\theta_p - (x_v \cos\phi_p + y_v \sin\phi_p)\cos\theta_p\right];\\
&x_{ip} = -z_v \sin\theta_p \cos\theta_p \cos\phi_p + x_v\left(1 - \sin^2\theta_p \cos^2\phi_p\right) - y_v \sin^2\theta_p \sin\phi_p \cos\phi_p;\\
&y_{ip} = -z_v \sin\theta_p \cos\theta_p \sin\phi_p - x_v \sin^2\theta_p \sin\phi_p \cos\phi_p + y_v\left(1 - \sin^2\theta_p \sin^2\phi_p\right).
\end{aligned} \tag{26}$$

where the second equality of the top equation of Eq. (25) identifies v_e and w_e as impact parameters. In Ref. [7], ϕ_p was explicitly set to zero by a rotation, and only v_e was considered (they are called x_e). This is equivalent to setting $y_v = 0$ as is shown by considering the ZXY components of $\vec{r}_{ip}$ in the limit $\phi_p \to 0$, i.e., $z_{ip} \to A\sin\theta_p$, $x_{ip} \to -A\cos\theta_p$ (here $A \equiv z_v\sin\theta_p - x_v\cos\theta_p$), and $y_{ip} \to y_v$; thus $y_{ip} = 0$ only if $y_v = 0$.

Therefore, the geometry of Ref. [7] is not completely general because it implicitly assumes $y_v = 0$ and uses only one impact parameter v_e; whereas in general, $y_v \neq 0$ and two impact parameters v_e and w_e are required. The formulation and reconstruction method given below is, however, quite general. The derivatives found in

Ref. [7] apply only for $y_v = 0$ and should be replaced by the more general ones given in Appendix A.

In ZXY, the photon emission point is (z_e, x_e, y_e); thus $\vec{r}_e = z_e\vec{k} + x_e\vec{i} + y_e\vec{j}$ with $r_e = \sqrt{(z_e^2 + x_e^2 + y_e^2)}$. In PQR, the emission point is at distance u_{ve} along $\vec{a}_p$ from $\vec{r}_v$, i.e.,

$$\vec{r}_e = \vec{r}_v + u_{ve}\vec{a}_p = u_e\vec{a}_p + v_e\vec{b}_p + w_e\vec{c}_p \qquad (27)$$

with $r_e = \sqrt{(u_e^2+v_e^2+w_e^2)}$ and the second equality stemming from Eq. (25) with $u_e \equiv u_{va} + u_{ve}$. Since the (z_e, x_e, y_e) coordinates are defined in ZXY, they are obviously independent of (θ_p, ϕ_p). The PQR components (u_e, v_e, w_e), expressed in terms of (z_e, x_e, y_e) and (θ_p, ϕ_p), are

$$\begin{aligned} u_e &= z_e a_{pz} + x_e a_{px} + y_e a_{py}, \\ v_e &= z_e b_{pz} + x_e b_{px} + y_e b_{py}, \\ w_e &= z_e c_{pz} + x_e c_{px} + y_e c_{py}. \end{aligned} \qquad (28)$$

and

$$\begin{aligned} z_e &= u_e a_{pz} + v_e b_{pz} + w_e c_{pz}, \\ x_e &= u_e a_{px} + v_e b_{px} + w_e c_{px}, \\ y_e &= u_e a_{py} + v_e b_{py} + w_e c_{py}. \end{aligned} \qquad (29)$$

In matrix notation, $u_e = Az_e$ and $z_e = Bu_e$ where u_e and z_e are column vectors $(u_e, v_e, w_e)^T$, $(z_e, x_e, y_e)^T$. Here A is the matrix with $a_{11} = a_{pz}$, $a_{12} = a_{px}$, $a_{13} = a_{py}$, $a_{21} = b_{pz}$, $a_{22} = b_{px}$, $a_{23} = b_{py}$, $a_{31} = c_{pz}$, $a_{32} = c_{px}$, $a_{33} = c_{py}$, and B is the matrix with $b_{11} = a_{pz}$, $b_{12} = b_{pz}$, $b_{13} = c_{pz}$, $b_{21} = a_{px}$, $b_{22} = b_{px}$, $b_{23} = c_{px}$, $b_{31} = a_{py}$, $b_{32} = b_{py}$, $b_{33} = c_{py}$; thus $AB = 1$, hence $B = A^{-1}$. Note also that $B = A^T$, where T indicates transpose, i.e., $(A^T)_{ij} = a_{ji}$.

The detection point vector $\vec{r} = r\vec{n}$ has components (z, x, y) in ZXY which are independent of (θ_p, ϕ_p) while (u, v, w) in the PQR system depend on (θ_p, ϕ_p) as

$$\vec{r} = z\vec{k} + x\vec{i} + y\vec{j} = u\vec{a}_p + v\vec{b}_p + w\vec{c}_p \qquad (30)$$

with $r = \sqrt{(z^2+x^2+y^2)} = \sqrt{(u^2+v^2+w^2)}$. The matrix relations between the column vectors $u = (u, v, w)^T$ and $z = (z, x, y)^T$ are $u = Az$ and $z = Bu$.

Finally, the photon direction unit vector $\vec{a}$, defined by the Cherenkov polar and azimuthal angles (θ, ϕ) in PQR, has its emission components ($a_e = \cos\theta$, $b_e = \sin\theta\cos\phi$, and $c_e = \sin\theta\sin\phi$) independent of (θ_p, ϕ_p), while its ZXY components (a_z, a_x, a_y) depend on (θ_p, ϕ_p) as

$$\vec{a} = a_e\vec{a}_p + b_e\vec{b}_p + c_e\vec{c}_p = a_z\vec{k} + a_x\vec{i} + a_y\vec{j} \quad . \qquad (31)$$

The matrix relations between the column vectors $a = (a_z, a_x, a_y)^T$ and $a_e = (a_e, b_e, c_e)^T$ are $a_e = Aa$ and $a = Ba_e$ with $a^2 = a_e^2 = 1$.

Because the emitted photon plane (containing $\vec{n}_e$ and $\vec{a}$) also contains C and because the mirror normal (at the reflection point) is in this same plane, therefore the reflected photon will also be in this plane. This is expressed by the vector equation

$$\vec{n} = \mu\vec{n}_e + \nu\vec{a} \qquad (32)$$

with μ and ν to be determined. We define the angles (Ω, Ω_e, Ω'), shown in Fig. A1, by the dot products:

$$\begin{aligned}\vec{n}\cdot\vec{n}_e &= \cos\Omega;\\ \vec{a}\cdot\vec{n}_e &= \cos\Omega_e;\\ \vec{a}\cdot\vec{n} &= \cos\Omega'.\end{aligned}$$

Taking the dot product of Eq. (32) with each of ($\vec{a}$,$\vec{n}_e$,$\vec{n}$) and solving the first two resultant equations for μ and ν gives

$$\begin{aligned}\mu &= \frac{\sin\Omega'}{\sin\Omega_e},\\ \nu &= \frac{\sin\Omega}{\sin\Omega_e}.\end{aligned} \qquad (33)$$

The third equation is satisfied if $\Omega' = \Omega_e - \Omega$; however, this is always true since the three vectors form a closed triangle. Note that Ω is found directly from experiment as the dot product of the emission point and detection point unit vectors, i.e.,

$$\cos\Omega = \vec{n}\cdot\vec{n}_e = \frac{zz_e + xx_e + yy_e}{rr_e}; \qquad (34)$$

thus, Ω does not depend on (θ_p, ϕ_p) because both (z, x, y) and (z_e, x_e, y_e) are measured in ZXY. From the geometry of Fig. A1, we find $\Omega = \Omega_e + \Lambda - 2\theta_m$ which, with the sine law relations $r_e\sin\Omega_e = r_m\sin\theta_m = r\sin\Lambda$, becomes

$$\Omega = \Omega_e + \arcsin\left(\frac{r_e\sin\Omega_e}{r}\right) - 2\arcsin\left(\frac{r_e\sin\Omega_e}{r_m}\right). \qquad (35)$$

Thus, we obtain Ω_e (and $\Omega' = \Omega_e - \Omega$) from Ω by numerical inversion of Eq. (35) and find μ and ν from Eq. (33).

An equivalent and computationally faster way of obtaining Ω_e has been recently developed [17]. Consider the two triangles of Fig. A1 containing the $\Omega_a = \Omega_e - \theta_m$ and $\Omega_b = \Lambda - \theta_m$. Straightforward geometry allows us to obtain a quartic equation in $s \equiv \sin\Omega_a$, i.e.,

$$s^4 + a_3s^3 + a_2s^2 + a_1s + a_0 = 0, \qquad (36)$$

where $a_3 = -2\eta\sin\Omega$, $a_2 = \eta(a\eta - 2\rho)$, $a_1 = \eta(2\eta - \rho\cos\Omega)$, and $a_0 = (1 - 4\rho^2\eta^2)/4$ with $\eta \equiv r_m/2r$, $\rho \equiv r/r_e$, and $a \equiv 1 + \rho^2 + 2\rho\cos\Omega$. Solution of the quartic equation gives two real and two complex roots. Of the two real roots, we choose the root which has a mirror hit point $z_m > 0$. The other real root has $z_m < 0$ corresponding to a light ray reflected from the spherical mirror surface upstream of C (where, in fact, no mirror physically exists). From the two triangles, we find the additional relations

$$\begin{aligned}\tan\theta_m &= \frac{s}{2\eta\rho - c},\\ \Omega_e &= \Omega_a + \theta_m,\\ \tan\Omega_e &= \frac{s}{c - \frac{1}{2\eta\rho}}.\end{aligned} \qquad (37)$$

where $c \equiv \cos\Omega_a$. Equations (36) and (37) are equivalent to Eq. (35) but are computationally faster.

We then proceed to solve Eq. (31) for $\bar{a}$ as

$$\bar{a} = \left(\frac{1}{\nu}\right)\bar{n} - \left(\frac{\mu}{\nu}\right)\bar{n}_e = \left(\frac{\sin\Omega_e}{\sin\Omega}\right)\bar{n} - \left[\frac{\sin(\Omega_e - \Omega)}{\sin\Omega}\right]\bar{n}_e. \qquad (38)$$

Since the components of $\bar{a}$ along $(\bar{a}_p, \bar{b}_p, \bar{c}_p)$ are ($a_e = \cos\theta$, $b_e = \sin\theta\cos\phi$, and $c_e = \sin\theta\sin\phi$), we find for the Cherenkov emission angles (θ, ϕ), the explicit relations

$$\begin{aligned}
\cos\theta &= \left(\frac{\sin\Omega_e}{\sin\Omega}\right)\left(\frac{u}{r}\right) - \left[\frac{\sin(\Omega_e - \Omega)}{\sin\Omega}\right]\left(\frac{u_e}{r_e}\right), \\
\sin\theta\cos\phi &= \left(\frac{\sin\Omega_e}{\sin\Omega}\right)\left(\frac{v}{r}\right) - \left[\frac{\sin(\Omega_e - \Omega)}{\sin\Omega}\right]\left(\frac{v_e}{r_e}\right), \\
\sin\theta\sin\phi &= \left(\frac{\sin\Omega_e}{\sin\Omega}\right)\left(\frac{w}{r}\right) - \left[\frac{\sin(\Omega_e - \Omega)}{\sin\Omega}\right]\left(\frac{w_e}{r_e}\right).
\end{aligned} \qquad (39)$$

Thus, θ and ϕ are obtained from (Ω, Ω_e), (u_e, v_e, w_e), and (u, v, w). The (θ_p, ϕ_p) dependence enters through (u_e, v_e, w_e) and (u, v, w), i.e., via the matrix A of Eq. (28). This reconstruction is quite general (it works for any detector surface, e.g., flat) and removes all but intrinsic aberrations of the image.

Clearly, the solution for Ω_e (and Λ and θ_m) is independent of (θ_p, ϕ_p), since physically, it represents a light ray propagating from $\bar{r}_e$ to $\bar{r}$. Similarly, the timing equation (Sec. 3.4) is also independent of (θ_p, ϕ_p).

This completes the summary of the Cherenkov angle reconstruction method [7].

3.3 Approximate Determination of Particle Direction

The geometry of the LBL–RICH makes it difficult (or impossible) to implement a tracker inside the water volume; however, (θ_p, φ_p) can be found in good approximation from the center of the ring image. Since the ring images can be identified (see Figs. 7–14), we assume that an array of image points (z_i, x_i, y_i) and arrival times t_i are known (i = 1 to N). From these points, we can approximately determine the particle direction (even if $\bar{r}_e$ is not known) by assuming that the array (z_i, x_i, y_i) satisfies the equation of a circle on a sphere. Intersection of a cone (cone angle θ, cone direction θ_p, ϕ_p) with a sphere of radius r gives the equation of a circle on a sphere as

$$\begin{aligned}
&z_i + \xi x_i + \psi y_i - \lambda r = 0, \\
&\xi = \tan\theta_p \cos\varphi_p, \\
&\psi = \tan\theta_p \sin\varphi_p, \\
&\lambda = \cos\theta/\cos\theta_p.
\end{aligned} \qquad (40)$$

Minimizing the function $\chi^2 = \sum_1^N (z_i + \xi x_i + \psi y_i - \lambda r)^2$ with respect to the unknown coefficients (ξ, ψ, λ) leads to the equations

$$\xi = \frac{\sigma_{xy}\sigma_{yz} - \sigma_{xz}\sigma_{yy}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{xy}},$$
$$\psi = \frac{\sigma_{xy}\sigma_{xz} - \sigma_{yz}\sigma_{xx}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{xy}}, \qquad (41)$$
$$\lambda r = \langle z \rangle + \xi \langle x \rangle + \psi \langle y \rangle.$$

where $\langle x \rangle = \frac{1}{N}\sum x_i$, etc., $\langle xy \rangle = \frac{1}{N}\sum x_i y_i$, etc., $\sigma_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle$, etc. Thus,

$$\cos\theta_p = \frac{1}{\sqrt{1+\xi^2+\psi^2}},$$
$$\cos\phi_p = \frac{\xi}{\sqrt{\xi^2+\psi^2}},$$
$$\sin\phi_p = \frac{\psi}{\sqrt{\xi^2+\psi^2}}, \qquad (42)$$
$$\cos\theta = \frac{\langle z \rangle + \xi \langle x \rangle + \psi \langle y \rangle}{r\sqrt{1+\xi^2+\psi^2}}.$$

This zero impact parameter ($u_e = v_e = 0$), spherical detector surface approximation provides an initial estimate of (θ_p, φ_p) but not of θ; however, this is not a problem because the reconstruction algorithm [Eq. (39)] gives a precise estimate for (θ, ϕ).

3.4 Timing Measurements and Determination of the Track Parameters

The track parameters z_e, x_e, and y_e can be determined from the measurement of time. The ring image pattern defines an array (z_i, x_i, y_i) of image points (i = 1 to N) and an associated photon arrival time $t_i \pm \sigma_t$ at the PM. The photon pathlengths are ℓ_1 from the emission point $\vec{r}_e$ to the mirror hit point $\vec{r}_m$, and ℓ_2 from $\vec{r}_m$ to the photon detection point $\vec{r}$. These are found from Fig. A1, using the sine law relations $r_e \sin\Omega_e = r_m \sin\theta_m = r \sin\Lambda \equiv r_a$, as

$$\ell_1 = r_m \cos\theta_m - r_e \cos\Omega_e;$$
$$\ell_2 = r_m \cos\theta_m - r\cos\Lambda. \qquad (43)$$

The total pathlength $\ell = \ell_1 + \ell_2$ is obtained from the detection point $r = \sqrt{(z^2+x^2+y^2)}$ and the assumed emission point $r_e = \sqrt{(z_e^2+x_e^2+y_e^2)}$ as

$$\ell = \ell_1 + \ell_2 = \frac{ct}{n} = 2\sqrt{r_m^2 - r_e^2 + r_b^2} - \sqrt{r^2 - r_e^2 + r_b^2} - r_b, \qquad (44)$$

where Ω_e is from Eq. (35) [or Eqs. (36) and (37)], $r_b \equiv r_e \cos\Omega_e$, and $t = n\ell/c$ is the time from photon emission to detection. As expected, Eq. (44) is independent of particle direction (θ_p, ϕ_p).

3.5 Reconstruction of (θ, ϕ) and Determination of Track Parameters $(z_e, x_e, y_e, \theta_p, \phi_p)$

The reconstruction and track finding strategy can now be stated:

(1) an image is visually identified with measured hit points and time (z_i, x_i, y_i, t_i);

(2) the initial emission point and emission time (z_e, x_e, y_e, t_e) are chosen randomly in the water volume and during the beam spill;

(3) Ω is calculated from Eq. (34);

(4) Ω_e is determined by numerically inverting Eq. (35) or solving Eqs. (36) and (37);

(5) an initial estimate of particle direction (θ_p, ϕ_p) is found from Eq. (42);

(6) values for (θ, ϕ) are then determined from Eq. (39) using Eqs. (28) and (29);

(7) the time-detector $\chi^2 = \chi_{td}^2$ function is constructed from measured t_i and t_i^{th} from Eq. (44) as

$$\chi_{td}^2 = \sum_1^N \left(\frac{t_i - t_i^{th}}{\sigma_t} \right)^2 \qquad (45)$$

and minimized by varying (z_e, x_e, y_e, t_e). Since absolute time of the primary interaction cannot be determined (or known from the beam spill because 20 beam bunches will be inside the water radiator at any given time), the emission time t_e may be determined from the ring. However, because χ_{td}^2 depends only on time differences, these can be referenced to any convenient zero which in our case is the onset of the SPS beam spill cycle. We have shown for the geometry of Fig. 1 with 1 ns time bins, 125 mm x, y pixels (and with 1λ emission point variation, chromatic, multiple scattering, and energy loss aberrations) that the χ_{td}^2 function near its minimum varies by about a factor of two for variations of 100 mm of the average emission point. The problem of finding a good starting point (z_e, x_e, y_e, t_e) within the large radiator volume is considered in Sec. 4.

(8) A finer determination of the track parameters $(z_e, x_e, y_e, \theta_p, \phi_p)$ and (θ, ϕ) is obtained by minimizing the width of the θ distribution, i.e.,

$$\sigma_\theta^2 = \frac{1}{N}\sum_1^N \theta_i^2 - \left(\frac{1}{N}\sum_1^N \theta_i \right)^2, \qquad (46)$$

by varying the track parameters near the minimum of the χ_{td}^2 function [Eq. (45)].

3.6 Vertex Point Determination

The distance along the track from vertex to emission point u_{ve} varies from photon to photon with $\langle u_{ve} \rangle = \lambda/2 = 425$ mm for the average hadron. A better estimate is obtained from the number of imaged points N, i.e.,

$$u_{ve} = \frac{N}{\varepsilon_g \varepsilon_a N_0 \sin^2\theta} \qquad (47)$$

corrected for geometrical efficiency ε_g, i.e., the fraction of photons which hit the mirror and are imaged (without hitting the side walls). This is obtained by simulation from the assumed emission point (z_e, x_e, y_e). In this simulation, (θ_p, ϕ_p) and θ are needed and ϕ is varied to find the efficiency. The absorption term $\varepsilon_a = e^{-\mu\ell}$ is evaluated from ℓ [Eq. (44)] using the measured water absorption coefficient $\mu = \mu(E)$ (Ref. [15]).

If two ring images (1 and 2) are fit by the preceding algorithm (Sec. 3.4) and if their emission points $\bar{r}_{e1}$ and $\bar{r}_{e2}$ are near each other (i.e., $\Delta r_{12} = |\bar{r}_{e1} - \bar{r}_{e2}| \approx \lambda$), then they are candidates to have a common vertex. The vector equation for the vertex point is

$$\bar{r}_V = u_{va1}\bar{a}_{p1} + v_{e1}\bar{b}_{p1} + w_{e1}\bar{c}_{p1} = u_{va2}\bar{a}_{p2} + v_{e2}\bar{b}_{p2} + w_{e2}\bar{c}_{p2}, \qquad (48)$$

where $u_{va1} = u_{e1} \pm u_{ve1}$ and $u_{va2} = u_{e2} \pm u_{ve2}$ [u_{ve1} and u_{ve2} are found from Eq. (47)]. Both signs are negative for a primary vertex, whereas one sign is negative and the other positive for a scatter or decay vertex. The three components of Eq. (48) used in a χ^2 minimization will strongly limit the vertex point with C = 3M constraints (M is the number of charged vertex tracks). Thus, even a two track vertex will provide six equations on the vertex point. We estimate that the vertex point can be found with cm-like accuracy although this has not yet been verified by simulation.

4. Photon Detection with PM's at the Mirror Surface

It is essential to find the photon emission point four-vector $(\bar{r}_e, t_e)$ in order that the Cherenkov angle reconstruction can be implemented (step 2 of Sec. 3.5). We have shown that random start points converge to the correct minimum if the point is within a 3 m radius four sphere about the true emission point. To explore the space inside a $(30\ m)^3$ volume would require choosing 10^3 random start points. In the time coordinate, the radiator is 1350 m long (i.e., a 6 μs beam spill with c/n = 0.225 m/ns), hence 450 segments of ± 1.5 m length. Combining these 450 points with the 1000 volume start points implies about 450 k random start points, which appear excessive.

For this reason, we have investigated the effect of replacing 4% of the reflecting mirror surface area, i.e., 36 m^2 with 2880 PM's of 127 mm diameter so as to directly detect Cherenkov photons (a la IMB,

Kamiokande, and Super-Kamiokande). This will result in a substantial sample of mirror PM (mPM) hits, since the mPM cover is 1/5 of the detector PM (dPM) cover, and thus, provides 160 mPM hits compared with 800 dPM hits.

4. 1 Determination of the Photon Emission Point and Time

We use a four-vector formulation of the problem due to G. Fiorentini [18] and denote by $\xi_\alpha \equiv (\vec{r}, ct/n)$ the four-vector components of the ith, mPM hit (subscripts i and m are dropped). Denoting the emission point vector components $\psi_\alpha \equiv (\vec{r}_e, ct_e/n)$, then the equation for the photon vector components is just the difference between hit and emission points, i.e., $a_\alpha = \xi_\alpha - \psi_\alpha$. We define the coordinate vector $\eta_\alpha \equiv \xi_\alpha - \langle\xi_\alpha\rangle$ so that its average over the i hits is zero (i.e., $\langle\eta_\alpha\rangle = 0$) and the vector $\delta_\alpha \equiv \psi_\alpha - \langle\xi_\alpha\rangle$, thus $a_\alpha = \eta_\alpha - \delta_\alpha$. The condition for a_α to be a photon four vector is then

$$a_\alpha a_\alpha = (\eta_\alpha - \delta_\alpha)^2 = \eta_\alpha\eta_\alpha - 2\eta_\alpha\delta_\alpha + \delta_\alpha\delta_\alpha = 0; \qquad (49)$$

averaging over the i photon hits and recalling that $\langle\eta_\alpha\rangle = 0$ gives

$$\begin{aligned} &\langle\eta_\alpha\eta_\alpha\rangle \equiv \langle s^2\rangle = -\delta_\alpha\delta_\alpha \equiv -s_e^2; \\ &\langle s^2\rangle = \langle z^2\rangle + \langle x^2\rangle + \langle y^2\rangle - \langle t^2\rangle; \\ &s_e^2 = z_e^2 + x_e^2 + y_e^2 - t_e^2, \end{aligned} \qquad (50)$$

where (z, x, y, t) are the hit mPM coordinates with time converted to space by the velocity factor c/n. Equation (50) thus gives a quadratic constraint on the unknowns (z_e, x_e, y_e, t_e) in the form $s_e^2 = -\langle s^2\rangle$, i.e., as averages over known mirror hit points $\langle s^2\rangle$. This relation constrains but does determine (z_e, x_e, y_e, t_e). For this purpose, we multiply Eq. (50) by η_β and again average over hit points to obtain

$$\begin{aligned} &\langle\eta^2\eta_\beta\rangle - 2\delta_\alpha\langle\eta_\alpha\eta_\beta\rangle = 0; \\ &\delta_\alpha T_{\alpha\beta} = V_\beta; \\ &T_{\alpha\beta} = \langle\eta_\alpha\eta_\beta\rangle; \\ &V_\beta = \langle\eta^2\eta_\beta\rangle/2, \end{aligned} \qquad (51)$$

where the third term of Eq. (49) drops out because $\langle\eta_\beta\rangle = 0$. This gives a set of four linear equations for $\delta_\alpha = (z_e, x_e, y_e, t_e)$ in terms of the tensor $T_{\alpha\beta}$ and the vector V_β. These are respectively, quadratic and cubic moments averaged over hit points, i.e., $T_{11} = \langle z^2\rangle$, $T_{12} = \langle zx\rangle$, $T_{13} = \langle zy\rangle$, $T_{14} = -\langle zt\rangle$, $T_{21} = \langle zx\rangle$, $T_{22} = \langle x^2\rangle$, $T_{23} = \langle xy\rangle$, $T_{24} = -\langle xt\rangle$, $T_{31} = \langle zy\rangle$, $T_{32} = \langle xy\rangle$, $T_{33} = \langle y^2\rangle$, $T_{34} = -\langle yt\rangle$, $T_{41} = \langle zt\rangle$, $T_{42} = \langle xt\rangle$, $T_{43} = \langle yt\rangle$, $T_{44} = -\langle t^2\rangle$, and $2V_1 = \langle zs^2\rangle = \langle z^3\rangle + \langle zx^2\rangle + \langle zy^2\rangle - \langle zt^2\rangle$, $2V_2 = \langle xs^2\rangle$, $2V_3 = \langle ys^2\rangle$, $2V_4 = \langle ts^2\rangle$. In order to obtain stable and accurate (meaningful) solutions to these linear equations, it was necessary to use the method of Gaussian elimination with partial pivoting [19].

Unfortunately, the constraint [Eq. (50)] is not contained in the linear relations for δ_α of Eq. (51). We have simulated many random events with full errors and aberrations and plotted the distance between the

solution values (z_{es}, x_{es}, y_{es}, t_{es}) to input values (z_e, x_e, y_e, t_e) in the form of the four-radius $r_4 = \sqrt{[(z_{es}-z_e)^2+(x_{es}-x_e)^2+(y_{es}-y_e)^2+(t_{es}-t_e)^2]}$ versus $C = [s_e^2+\langle s^2\rangle]/\langle s^2\rangle$, the normalized constraint. It was found that for $-1 < C < 1$, the four radius is within 3 m of the true emission point; hence, it is sufficient to guarantee that the χ_{td}^2 algorithm of Sec. 3.4 converges to the correct minima. Moreover, we found that about 90% of the random events have $|C| \le 1$.

The above results may be compared to the standard χ^2 formulation of the problem. Since the ith hit point is caused by a photon, it obeys the spherical wave relation

$$f_i = (z_i - z_e)^2 + (x_i - x_e)^2 + (y_i - y_e)^2 - (t_i - t_e)^2 = 0. \qquad (52)$$

Averaging over the i hits gives

$$\langle f\rangle = \langle z\rangle^2 + \langle x\rangle^2 + \langle y\rangle^2 - \langle t\rangle^2 + z_e^2 + x_e^2 + y_e^2 - t_e^2 = \langle s\rangle^2 + s_e^2 = 0; \qquad (53)$$

thus, we recover the constraint of Eq. (50). Here we have used, as before, the hit points with their average subtracted so that $\langle z\rangle = \langle x\rangle = \langle y\rangle = \langle t\rangle = 0$. We now define the spherical wave to mirror χ^2 as

$$\chi_{swm}^2 = \sum_i^N \left(\frac{f_i}{\sigma_{f_i}}\right)^2, \qquad (54)$$
$$\sigma_{f_i} = 2\sqrt{(z_i - z_e)^2\sigma_z^2 + (x_i - x_e)^2\sigma_x^2 + (y_i - y_e)^2\sigma_y^2 + (t_i - t_e)^2\sigma_t^2}.$$

Requiring $\partial\chi^2/\partial z_e = 0$, $\partial\chi^2/\partial x_e = 0$, $\partial\chi^2/\partial y_e = 0$, $\partial\chi^2/\partial t_e = 0$ results in the same formulas as Eq. (51) except the diagonal elements T_{ii} have the additional term ε, i.e., $T_{11} = \langle z^2\rangle + \varepsilon$, $T_{22} = \langle x^2\rangle + \varepsilon$, $T_{33} = \langle y^2\rangle + \varepsilon$, $T_{44} = -\langle t^2\rangle + \varepsilon$ (where $2\varepsilon = C\langle s^2\rangle = \langle s^2\rangle + z_e^2 + x_e^2 + y_e^2 - t_e^2$); thus the equations are no longer linear (i.e., cubic) in (z_e, x_e, y_e, and t_e) unless the constraint $C = 0$ is satisfied. An attempt to find an iterative solution (the diagonal elements were increased by the ε found in the preceeding iteration) failed since the solution did not converge. Possibly, this constraint may be imposed by the method of Lagrangian multipliers. In general, and for any start point, the χ_{swm}^2 function must be minimized to find the best emission point vector (z_e, x_e, y_e, and t_e).

4. 2 Determination of the Particle Direction

An approximate algorithm to determine particle direction using dPM hits has already been given in Sec. 3.3, but a more precise algorithm using mPM hits would be advantageous for choosing a better start point for the χ^2 minimizations (Step 2 of Sec. 3.5).

We present here an algorithm due to G. Fiorentini [19] using again the notation of the preceeding section. The three-vector η_i represents the kth mirror hit point (average subtracted, subscripts k and m dropped), thus the average over the k hits $\langle\eta_i\rangle = 0$. The emission point vector in the same coordinate system is δ_i, hence the photon vector (of

length ρ) is $\rho_i = \rho a_i = \eta_i - \delta_i$ and its dot product with particle direction is $\rho a_i a_{pi} = \rho\cos\theta = (\eta_i - \delta_i)a_{pi}$. The length ρ is simply the time difference $\eta_4 - \delta_4$, and hence the basic equation is $(\eta_i - \delta_i)s_i = (\eta_4 - \delta_4)$ in the variables $s_i \equiv a_{pi}/\cos\theta$. Averaging over hits gives the constraint $\delta_i s_i = \delta_4$. Multiplying the basic equation by η_m and averaging gives a set of linear equations $T_{mi}s_i = V_m$, where the tensor $T_{mi} = \langle\eta_m\eta_i\rangle$ and the vector $V_m = \langle\eta_m\eta_4\rangle$. Expressed in terms of the hit points (z, x, y, t), we get the same tensor as given in Sec. 4.1 (i.e., $T_{11} = \langle z^2\rangle$, $T_{12} = \langle zx\rangle$, . . .) whereas the vector components are $V_1 = \langle tz\rangle$, $V_2 = \langle tx\rangle$, $V_3 = \langle ty\rangle$. Note that the three solutions s_i just suffice to determine θ, θ_p, and ϕ_p. Written out in full, the constraint has the form $z_e s_1 + x_e s_2 + y_e s_3 = t_e$ where (z_e, x_e, y_e, t_e) are found from Eq. (51).

The problem then is to solve a set of three linear equations with a linear constraint. The solution of the linear equations by Gaussian elimination with partial pivoting [19] is straightforward, but we have not yet been able to include the constraint. The solutions with emission point, pixel, and timing errors included are excellent if the hit (data) points are generated without chromatic aberrations, multiple scattering, or energy loss; however, once these are included, the solutions become unstable and useless. For this reason, we looked for another less elegant algorithm which can provide some sensitivity to particle direction from mirror hit points.

The vector equation of the photon hit point $\vec{r} = \vec{r}_e + \rho\vec{a}$, where again ρ is the distance between the emission and hit points, and $\vec{a}$ the photon direction unit vector [Eq. (31)]. In the PQR coordinate system, the components are $u_i - u_{ei} = \rho_i\cos\theta$, $v_i - v_e = \rho_i\sin\theta\cos\phi$, and $w_i - w_e = \rho_i\sin\theta\sin\phi$ where ρ_i is the distance between ith emission and hit points. Recall that impact parameter coordinates (v_e and w_e) do not depend on i because they are constant anywhere along the track, whereas u_e varies with the point of emission along the track. Eliminating ρ_i and ϕ, we obtain the equation of a cone $g_i = 0$ and then solve for u_{ei}, i.e.,

$$g_i = (v_i - v_e)^2 + (w_i - w_e)^2 - \tan^2\theta(u_i - u_{ei})^2 = 0,$$
$$u_{ei} = u_i - \frac{\sqrt{\left[(v_i - v_e)^2 + (w_i - w_e)^2\right]}}{\tan\theta}. \qquad (55)$$

The emission point u_{ei} along the track is determined if θ, θ_p, ϕ_p are assumed and the impact parameters v_e and w_e are known. For this purpose, we use the solution of Eq. (51) which gives (z_e, x_e, y_e, t_e) from the mirror hit points (z_i, x_i, y_i, t_i). The assumed (θ_p, ϕ_p) direction defines the matrices A and B; thus $u_e = Az_e$ can be calculated and impact parameters v_e and w_e extracted. Since these are constant along the track, we use these in Eq. (55) along with the transformed hit point $u_i = Az_i$ to find the column vector $u_{ei} = (u_{ei}, v_e, w_e)^T$. Transforming back via $z_{ei} = Bu_{ei}$ gives a column vector (z_{ei}, x_{ei}, y_{ei}) for each hit point which inserted into Eq. (52) gives the cone-spherical wave-mirror (cswm) direction dependent constraint

$$\chi_{cswm}{}^2 = \sum_i^N \left(\frac{g_i}{\sigma_{g_i}} \right)^2,$$

$$\sigma_{g_i} = 2\sqrt{(Z\sigma_z)^2 + (X\sigma_x)^2 + (Y\sigma_y)^2},$$

$$Z = (v - v_e)b_{pz} + (w - w_e)c_{pz} - \tan^2\theta(u - u_{ei})a_{pz}, \qquad (56)$$

$$X = (v - v_e)b_{px} + (w - w_e)c_{px} - \tan^2\theta(u - u_{ei})a_{px},$$

$$Y = (v - v_e)b_{py} + (w - w_e)c_{py} - \tan^2\theta(u - u_{ei})a_{py}.$$

In summary, we now have three χ^2 functions to minimize $\chi_{td}{}^2$ [Eq. (45)] and $\chi_{swm}{}^2$ [Eq. (54)], which depend only on the emission point vector and the $\chi_{cswm}{}^2$ function [Eq. (56)] which depends both on emission point and particle direction (θ_p, ϕ_p). In addition, the width σ_θ of the reconstructed Cherenkov angle [Eq. (46)] depends both on emission point and particle direction.

We calculated the values of $\chi_{td}{}^2$, $\chi_{swm}{}^2$, $\chi_{cswm}{}^2$, θ, and σ_θ vs $\Delta(t_e, z_e, x_e, y_e, \theta, \theta_p, \phi_p)$ for representative events with realistic errors (Δ here indicates the difference between the true value of the variable and its assumed value). Each event was generated at a random point and time in the water volume (z_e, x_e, y_e, t_e) with a random direction (θ_p, ϕ_p), random pathlength (ℓ), and an emission point random between ($0 < u_e < \ell$). The ring images included chromatic aberrations, multiple scattering, and energy loss appropriate to the water radiating medium. The results showed that track variables (t_e, z_e, x_e, y_e) were determined to about 100 mm while the particle direction (θ_p, ϕ_p) was found within 5 mrad. More important, the values of θ and σ_θ were found always to be equal to the generated values thus allowing velocity determination and, in addition, momentum and mass determination over the range of MSD.

5. Beams and Sites

5.1 The CERN-SPS Extracted Beam

The CERN-SPS beam operates at 200 MHz, thus with a 5 ns periodicity. This means that succeeding RF bunches are separated in the water target by only 1.5 m; hence absolute timing cannot determine the interaction vertex point (this would be possible if the RF bunches were separated by > 100 ns). In other words, at any given time within an SPS burst, there will be 20 RF bunches inside the water target.

One, two, or three SPS beam bursts can be extracted every SPS cycle of 14 sec. If one burst is extracted, it will be 23 μs long and contain 1.3×10^{13} p. If two bursts are extracted, they will be 10 μs long separated by 50 ms each with 1.3×10^{13} p giving 2.6×10^{13} p/cycle. If three bursts are extracted, they will be 6 μs long separated by 50 ms, hence 3.9×10^{13} p/cycle. For a 44% SPS duty factor, there will be 10^6 cycles/y and so $\approx 4 \times 10^{19}$ p/y (Ref. [20]).

5.2 Neutrino Beam and Event Rates

The designed broad band neutrino beam ($1 < E_\nu < 20$ GeV) has a flux of 500 events/kt-10^{19} p; thus for 4×10^{19} p/y, the rate will be 2000 events/kt-y (Ref. [21]). Since the LBL-RICH contains 22.5 fiducial kt of water, it should amass up to 45 k events/y (for no oscillations, i.e., $\Delta m^2 = 0$).

5.3 Where to Site the LBL-RICH

5.3.1 Outside the Gran Sasso Tunnel

As is clear from Fig. 1, the Gran Sasso tunnel would have to be 43 m in diameter to contain a $(30\ m)^3$ cube. This is twice the diameter of the present and future Gran Sasso tunnels, so we have investigated operation of the LBL–RICH above ground. The possibility is to use an existing Gran Sasso tunnel of about 19 m diameter, but extending the length of the radiator to 100 m will be considered in the next section.

We assume the full unshielded cosmic ray flux of $180/m^2$-s (Ref. [22]). For the 900 m^2 surface area of LBL-RICH, the rate will be 0.16 MHz; thus, during a beam burst of 6 μs, we expect one muon to traverse the LBL–RICH. Since 3×10^6 SPS bursts result in 45 k events, the specific event rate is 0.015 events/burst or one signal event (S) per 67 bursts. During this particular burst, we expect one background muon (B), thus $S/B \approx 1$. The other 66 bursts will contain only an obvious B event which cannot be confused with S because it lies in another burst. At least four methods are available to reduce B to a negligible level. They are:

(1) By optically shielding the PM's so they only view the mirrors.

(2) By timing. Recall that the PM hits are binned in buckets of 1 ns width over a period of 128 ns. The B event will arrive randomly over the 6 μs burst gate since, because of its directionality, it cannot initiate the 100 PM trigger. The true S event arrives (during the 6 μs burst) with ≤ 128 ns dispersion relative to the 100 PM trigger signal, thus allowing a B reduction factor of 6000/128 = 47.

(3) By pattern. Since the B events are mostly vertical, they do not form good images, whereas the S events are mostly longitudinal and do form good images.

(4) By massive shielding. Even though the LBL–RICH will be above ground, it should be placed behind a mountain (when viewed from CERN), thus screening out the more horizontal muon tracks.

The cost estimate with mHPD's and dHPD's is 25 MSF.

5.3. 2 Inside a Gran Sasso Tunnel

A possible layout of the LBL-RICH inside a Gran Sasso tunnel of 19 m diameter is shown in Fig. 15. Using five sections of 20 m length and 18.6 m diameter, we can obtain the same 27 kt water mass. The HPD image plane is located at 11.5 m from the mirror's center of curvature, and the image radius is about 8.4 m.

This solution has several advantages; namely, it is shielded both from cosmic rays and from ecologists. Another advantage is that the muons will be extremely well measured in several of the five sections. The device becomes a Cherenkov total energy calorimeter with 1000 pe/m = 625 pe/GeV.

The disadvantages are that it has a smaller electron shower fiducial mass since the last 5 m of each section is needed to contain the shower (20 kt compared to 22.5 kt) and requires more HPD surface area [1360 compared to 900 m^2, thus 5 x 1100 = 5500 HPD's of 250 mm in diameter compared to 3600 for the $(30\ m)^3$ radiator].

The total cost estimate here is 40.6 MSF compared to the outside option of 25 MSF; hence costs scale for the same water mass approximately as the inverse ratio of mirror focal length (i.e., 30/20 = 1.5, whereas the inside option is actually 1.62 times more expensive).

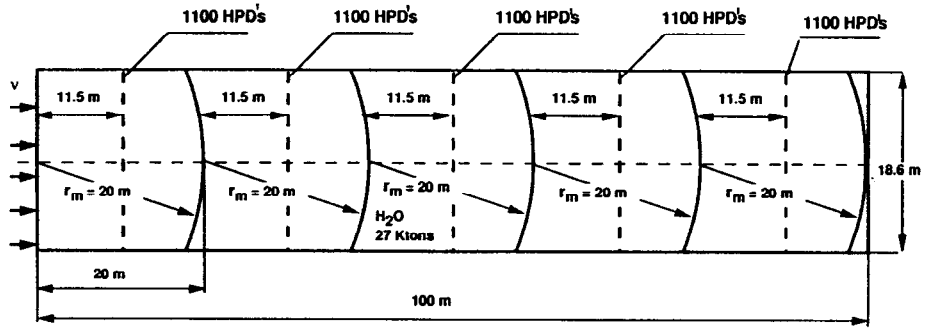


Fig. 15. The layout of the 27 kt water target and radiator filling a Gran Sasso tunnel of 18.6 m diameter. Five equivalent sections of 20 m length have reflecting mirrors of 20 m curvature placed at the end of each section. The detector HPD plane array (20% coverage) is placed 11.5 m downstream from the mirror center of curvature.

6. The LBL-RICH Physics Reach

The maximal mixing model of Harrison *et al.* [23] (which fits all existing neutrino data) predicts a large muon disappearance effect (44%) and a large electron appearance effect (22%) in the region $50 < L/E_\nu < 700$, thus for $1 < E_\nu < 15$ at L = 732 km. This region is easily accessable with the broad band beam in the LBL-RICH. The muon rate will be large without (and measurably less with) oscillations, whereas the electron rate will be small without (and measurably more with) oscillations. Since the LBL-RICH has good energy resolution for muons, electrons, and hadrons, it can fine-bin the muon or electron type events vs L/E_ν so as to observe the maximum oscillation swings.

In the two-neutrino mixing analysis, we expect to reach $\Delta m^2 \approx 9.2 \times 10^{-4}$ eV^2/sin(2θ) for $\nu_\mu \rightarrow \nu_\tau$ oscillations via the reaction $\nu_\tau + n \rightarrow \tau^- + p$ with decay $\tau^- \rightarrow e^- + \nu_e + \nu_\tau$ by cuts on p_t and E_e. Neutrino interactions in this low-energy, wide-band beam ($1 < E_\nu < 20$ GeV) are ≈ 95% deep inelastic (DIS) (67% CC, 33% NC) and ≈ 5% quasi-elastic (QEL). The beam flux is 2000 events/kt-y, hence the LBL–RICH (22.5 kt fiducial) will detect 2.3×10^5 events in five years. The number of ν_μ quasi-elastic events is $N_\mu(QEL) \approx 1.2 \times 10^4$ with a ν_e quasi-elastic background of $N_e(QEL) \approx 60$ (since $\nu_e/\nu_\mu = 0.5\%$) but is reduced to 0.4 by the kinematical cuts (6×10^{-3}) (Ref. [24]). The number of NC-DIS events is much larger (7.7×10^4) but drops to 15 by the same kinematical cuts (2×10^{-4}) (Ref. [24]). Further purification must be obtained from the ring patterns. If no $\tau \rightarrow e$ candidates are observed in this event sample ($N_{\tau e} = 2.3$ at the 90% confidence level), we obtain the oscillation probability

$$P = \frac{N_{\tau e}}{(N_\mu)(BR)(eff)\left(\frac{\sigma_{\nu\tau}}{\sigma_{\nu\mu}}\right)} = \frac{2.3}{(1.2\times10^4)(0.18)(0.21)(0.7)} = 0.0072 , \quad (57)$$

where BR is the $\tau \rightarrow e\nu\nu$ branching ratio, eff = signal efficiency = 0.21 (0.23 for p above threshold, 0.9 for $p_t > 0.3$ GeV), and 0.7 is the ratio of the ν_τ and ν_μ cross sections. However, for small oscillations, we can write

$$\sqrt{P} = \sin(2\theta)\frac{1.27\Delta m^2 L}{E_\nu} = 0.085 ; \quad (58)$$

hence for Gran Sasso at L = 732 km and with neutrino beam energy $E_\nu = 10$ GeV, we find

$$\Delta m^2 = \frac{9.2\times10^{-4}}{\sin(2\theta)}. \quad (59)$$

7. Summary

We have shown how a RICH counter can measure momentum and have applied this method to investigate long baseline neutrino oscillation experiments. This method allows large mass targets, but with measurement of momentum, direction, velocity, mass, and absolute charge for hadrons and muons. In addition, electrons and gammas can be measured by calorimetry in water with good energy and excellent direction resolution.

References

[1] J. L. Guyonnet *et al.*, Nucl. Instrum. Methods A **343**, 178 (1994) and J. Séguinot *et al.*, Nucl. Instrum. Methods A **314**, 465 (1992).

[2] J. Séguinot *et al.*, Nucl. Instrum. Methods A **371**, 64 (1996).

[3] T. Ypsilantis and J. Séguinot, Nucl. Instrum. Methods A **368**, 229 (1995).

[4] T. Ypsilantis *et al.*, Nucl. Instrum. Methods A **371**, 330 (1996).

[5] A. Zichichi, "Perspectives of Underground Physics: The Gran Sasso Project," CERN-EP/88-28.

[6] Letter of Intent for the Long Baseline RICH submitted on 1/1/96 to Laboratorio Nazionale del Gran Sasso: "A 27 kton Liquid Water Target and RICH Photon Counter to Detect Oscillations in the CERN-Gran Sasso Neutrino Beam," T. Ypsilantis, spokesman.

[7] T. Ypsilantis and J. Séguinot, Nucl. Instrum. Methods A **343**, 30 (1994).

[8] D. Hatzifotiadou *et al.*, College de France, LPC/94-29.

[9] Y. Goldschmidt-Clermont *et al.*, Proc. Phys. Soc. **61**, 138 (1948).

[10] U. Camerini *et al.*, Phil. Mag. **41**, 413 (1950).

[11] R. Barlow, *Statistics* (John Wiley and Sons), Sec. 5.2.3.

[12] R. De Salvo, CLNS 87-92, Cornell University, Ithaca, NY. See also R. De Salvo *et. al.*, Nucl. Instrum. Methods A **315**, 375 (1992); Nucl. Instrum. Methods A **365**, 76 (1995).

[13] E. Chesi *et al.*, Nucl. Instrum. Methods A **387**, 222 (1997).

[14] R. De Salvo (private communication).

[15] J. Quickendon *et al.*, J. Chem. Phys. **72**, 4416 (1980).

[16] I. Thormahlen, J. Straub, and U. Grigull, J. Phys. Chem. Ref. Data **14**, 933 (1985).

[17] R. Forty, Nucl. Instrum. Methods A **384**, 167 (1996).

[18] G. Fiorentini (private communication).

[19] G. E. Forsythe and C. B. Moler, *Computer Solution of Linear Algebraic Systems* (Prentice-Hall Inc.), pp. 68-70.

[20] E. Weisse (private communication).

[21] A. Ball *et al.*, Nucl. Instrum. Methods, in press.

[22] Particle Data Group, Phys. Rev. D **50**, 1173 (1994). See p. 1269.

[23] P. F. Harrison, D. H. Perkins, and W. G. Scott, Phys. Lett. B **349**, 137 (1995).

[24] A. Rubbia, Report to the CERN Medium/Long Baseline Working Group (1996).

Appendix A

We consider the more general geometry of a ring image. We define two different coordinate systems as shown in Fig. A1, the ZXY system, fixed relative to the mirror, with unit vectors $(\vec{k},\vec{i},\vec{j})$ and the PQR system, fixed to each track and defined from C such that P is parallel to the track and Q and R are normal to the track with unit vectors $(\vec{a}_p,\vec{b}_p,\vec{c}_p)$. The photon emission point is (z_e, x_e, y_e) in the ZXY system and (u_e, v_e, w_e) in the PQR system. The photon detection point is (z, x, y) in the ZXY system and (u, v, w) in the PQR system.

An external particle tracker would directly measure (u_e, v_e, w_e) and the photon detector measures (z, x, y). The other coordinates (z_e, x_e, y_e) and (u, v, w) must then be determined from knowledge of (θ_p, φ_p) via Eqs. (29) and (30).

In the LBL-RICH case, the mirror tracker measures (z_e, x_e, y_e, t_e) and the photon detector (z, x, y). Thus, (u_e, v_e, w_e) and (u, v, w) are functions of (θ_p, φ_p) via Eqs. (28) and (30).

The emission point vector $\vec{r}_e$, the photon vector $\vec{a}$, and the mirror hit vector $\vec{r}_m$ define one triangle with included angles $\pi-\Omega_e$, θ_m, and Ω_a as shown in Fig. A1. The mirror hit vector $\vec{r}_m$ with the reflected photon vector $\vec{a}_r$ and the detected point vector $\vec{r}$ define a second triangle with angles θ_m, $\pi-\Lambda$, and Ω_b; hence since $\Omega = \Omega_a + \Omega_b = \Omega_e + \Lambda - 2\theta_m$ and $r_e\sin\Omega_e = r_m\sin\theta_m = r\sin\Lambda$ ($\equiv r_a$), then Eq. (35) is proven.

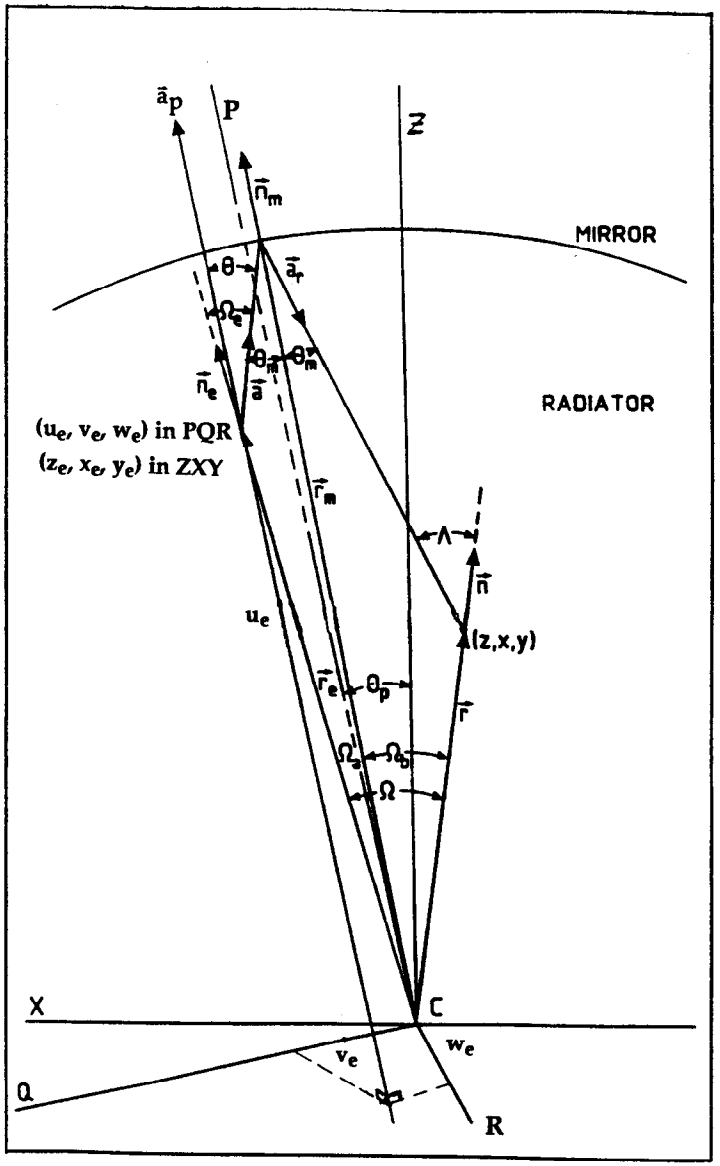


Fig. A1. The geometry of a mirror-focused ring image defined by the mirror center of curvature C, the Z axis along the neutrino beam, the X axis horizontal, and the Y axis vertical. The PQR axes, centered on C, are defined so that P is parallel to the particle direction $\vec{a}_p$ and Q and R are normal to $\vec{a}_p$ (i.e., parallel to $\vec{b}_p,\vec{c}_p$). The emission point coordinates in the PQR system are u_e along and v_e and w_e normal to the track.

Three relations are needed to allow calculation of the derivatives $\partial\theta/\partial\zeta_i$ for use in evaluation of the errors in Eq. (18); they are Eqs. (34), (35), and (39) (the first). Given these relations, it is straightforward to obtain the relation $\frac{\partial\theta}{\partial\zeta_i} = \frac{FK_i + GL_i + M_i}{\sin\theta}$ where

$$K_i = \frac{r_a}{\sqrt{r^2 - r_a{}^2}}\left(\frac{\partial \ln r}{\partial\zeta_i}\right) - Jr_a\left(\frac{\partial \ln r_e}{\partial\zeta_i}\right);$$

$$J = \frac{1}{\sqrt{r^2 - r_a{}^2}} - \frac{2}{\sqrt{r_m{}^2 - r_a{}^2}};$$

$$L_i = \frac{1}{\tan\Omega}\left[\frac{\partial \ln(rr_e)}{\partial\zeta_i}\right] - \frac{1}{rr_e\sin\Omega}\frac{\partial}{\partial\zeta_i}(zz_e + xx_e + yy_e); \qquad \text{(A1)}$$

$$M_i = \frac{\sin\Omega'}{\sin\Omega}\frac{\partial}{\partial\zeta_i}\left(\frac{u_e}{r_e}\right) - \frac{\sin\Omega_e}{\sin\Omega}\frac{\partial}{\partial\zeta_i}\left(\frac{u}{r}\right);$$

with

$$F = \frac{ru_e\cos\Omega' - ru_e\cos\Omega_e}{rr_e\sin\Omega(1 + Jr_b)}, \qquad \text{(A2)}$$

$$G = F + \frac{\sin\Omega_e}{\sin^2\Omega}\left[\frac{u\cos\Omega}{r} - \frac{u_e}{r_e}\right];$$

these are explicitly evaluated in Table A1 for both mirror and normal trackers.

Table A1

The coefficients K_i, L_i, M_i (versus ζ_i) which determine $\sin\theta(\partial\theta/\partial\zeta_i) = FK_i + GL_i + M_i$ [see Eqs. (A1) and (A2)] with $r_a \equiv r_e\sin\Omega_e = r_m\sin\theta_m = r\sin\Lambda$, and $r_b \equiv r_e\cos\Omega_e$. The primes in rows $\zeta_i = \theta_p$, or φ_p, indicate a partial derivative with respect to this variable. Quantities found in these rows (mirror tracker) are defined as $s_z = rz_e\sin\Omega - zr_b$, $s_x = rx_e\sin\Omega - xr_b$, $s_y = ry_e\sin\Omega - yr_b$.

ζ_i	K_i	L_i	M_i
z	$\frac{zr_a}{r^3\cos\Lambda}$	$\frac{zr_b - rz_e}{r^2 r_e\sin\Omega}$	$\left(\frac{uz}{r^2} - a_{pz}\right)\frac{\sin\Omega_e}{r\sin\Omega}$
x	$\frac{xr_a}{r^3\cos\Lambda}$	$\frac{xr_b - rx_e}{r^2 r_e\sin\Omega}$	$\left(\frac{ux}{r^2} - a_{pz}\right)\frac{\sin\Omega_e}{r\sin\Omega}$
y	$\frac{yr_a}{r^3\cos\Lambda}$	$\frac{yr_b - ry_e}{r^2 r_e\sin\Omega}$	$\left(\frac{uy}{r^2} - a_{pz}\right)\frac{\sin\Omega_e}{r\sin\Omega}$
Mirror tracker			
z_e	$\frac{-Jz_e\sin\Omega_e}{r_e}$	$\frac{rz_e\cos\Omega - zr_e}{rr_e{}^2\sin\Omega}$	$\left(a_{py} - \frac{z_e u_e}{r_e{}^2}\right)\frac{\sin\Omega'}{r_e}$
x_e	$\frac{-Jx_e\sin\Omega_e}{r_e}$	$\frac{rx_e\cos\Omega - xr_e}{rr_e{}^2\sin\Omega}$	$\left(a_{py} - \frac{x_e u_e}{r_e{}^2}\right)\frac{\sin\Omega'}{r_e}$
y_e	$\frac{-Jy_e\sin\Omega_e}{r_e}$	$\frac{ry_e\cos\Omega - yr_e}{rr_e{}^2\sin\Omega}$	$\left(a_{py} - \frac{y_e u_e}{r_e{}^2}\right)\frac{\sin\Omega'}{r_e}$
θ_p	0	0	$\frac{s_z a_{pz}' + s_x a_{px}' + s_y a_{py}'}{rr_e\sin\Omega}$
φ_p	0	0	$\frac{s_z a_{pz}' + s_x a_{px}' + s_y a_{py}'}{rr_e\sin\Omega}$
Normal tracker			
u_e	$\frac{-Ju_e\sin\Omega_e}{r_e}$	$\frac{ru_e\cos\Omega - (\bar{r}\cdot\bar{a}_p)r_e}{rr_e{}^2\sin\Omega}$	$\frac{(r_e{}^2 - u_e{}^2)\sin\Omega'}{r_e{}^3\sin\Omega}$
v_e	$\frac{-Jv_e\sin\Omega_e}{r_e}$	$\frac{rv_e\cos\Omega - (\bar{r}\cdot\bar{b}_p)r_e}{rr_e{}^2\sin\Omega}$	$-\frac{u_e v_e\sin\Omega'}{r_e{}^3\sin\Omega}$
w_e	$\frac{-Jw_e\sin\Omega_e}{r_e}$	$\frac{rw_e\cos\Omega - (\bar{r}\cdot\bar{c}_p)r_e}{rr_e{}^2\sin\Omega}$	$-\frac{u_e w_e\sin\Omega'}{r_e{}^3\sin\Omega}$
θ_p	0	$\frac{zz_e' + xx_e' + yy_e'}{-rr_e\sin\Omega}$	$\frac{za_{pz}' + xa_{px}' + ya_{py}'}{-r\sin\Omega/\sin\Omega_e}$
φ_p	0	$\frac{zz_e' + xx_e' + yy_e'}{-rr_e\sin\Omega}$	$\frac{za_{pz}' + xa_{px}' + ya_{py}'}{-r\sin\Omega/\sin\Omega_e}$

The Heavy Top Quark And Supersymmetry*

Lawrence J. Hall
Theoretical Physics Group, Lawrence Berkeley Laboratory
and Department of Physics, University of California, Berkeley, CA 94720
E-mail: hall_lj@lbl.gov

ABSTRACT

Three aspects of supersymmetric theories are discussed: electroweak symmetry breaking, the issues of flavor, and gauge unification. The heavy top quark plays an important, sometimes dominant, role in each case. Additional symmetries lead to extensions of the Standard Model which can provide an understanding for many of the outstanding problems of particle physics. A broken supersymmetric extension of spacetime allows electroweak symmetry breaking to follow from the dynamics of the heavy top quark; an extension of isospin provides a constrained framework for understanding the pattern of quark and lepton masses; and a grand unified extension of the Standard Model gauge group provides an elegant understanding of the gauge quantum numbers of the components of a generation. Experimental signatures for each of these additional symmetries are discussed.

*This work was supported in part by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under contract DE-AC03-76SF00098, and in part by the National Science Foundation under grant PHY-95-14797.

©1995 by Lawrence J. Hall.

Table of Contents

I. Symmetries and Symmetry Breaking

1. Symmetries
2. Flavor Symmetries
3. The Major Problems of the High-Energy Frontier
4. Supersymmetry
5. Summary

II. $SU(2) \times U(1)$ Breaking and the Weak Scale

1. A Symmetry Description
2. Matter vs. Higgs
3. A Heavy Top Quark Effect

III. Flavor in Supersymmetric Theories

1. The Fermion Mass and Flavor-Changing Problems
2. Approximate Flavor Symmetries
3. The Flavor-Changing Constraints
4. The Maximal Flavor Symmetry
5. A Brief Introduction to the Literature
6. The Minimal $U(2)$ Theory of Flavor
7. The Suppression of Baryon and Lepton Number Violation
8. Conclusions
9. Appendix A

IV. Supersymmetric Grand Unification

1. Introduction
2. Flavor Signals Compared
3. Flavor-Changing and CP-Violating Signals
4. The Top Quark Origin of New Flavor and CP Violation
5. Summary

V. The High-Energy Frontier

I. Symmetries and Symmetry Breaking

I.1 Symmetries

Much progress in particle physics has been made possible by understanding phenomena in terms of symmetries, which can be divided into four types: global or local action in spacetime or in an internal space. A symmetry of any of these types can be further classified as exact or broken, according to whether any breaking has been measured in experiments, as illustrated by well-known examples in Table 1. In these lectures, I discuss three of the four symmetry types, leaving out the gauging of spacetime symmetries which is expected to occur at the Planck scale.

An interesting feature of Table 1 is that of the six entries, only five have been discovered in nature: there is no experimental evidence for a broken, global symmetry of spacetime, hence the blank entry.

Table 1. Symmetries

	Exact	**Broken**
Local Internal	$SU(3)_{QCD}$ $U(1)_{EM}$	$SU(2) \times U(1)_Y$
Global Internal	Baryon number: B Individual lepton numbers: L_i	Isospin: $SU(2)_I$
Global Spacetime	Displacements: P Angular momentum: J Lorentz boosts: K	

I.2 Flavor Symmetries

With one exception, the entries of Table 1 provide a complete list of what has been discovered experimentally for these categories, ignoring the discrete spacetime symmetries such as parity. The exception is provided by global internal symmetries. Including color and weak degrees of freedom, 45 species of quarks and leptons have been found; experiments have therefore uncovered a $U(45)$ global internal, or flavor, symmetry, which is broken to $B \times L_i$ by the known gauge interactions and particle masses. The existence and masses of these 45 states, together with the way the known gauge forces act on them, is the flavor puzzle of particle physics. It is instructive to consider separately the breaking of $U(45)$ by gauge interactions and by masses. The known gauge interactions divide the 45 states into three identical periods, or generations, each of which contains five multiplets

transforming irreducibly under the gauge group: q, u, d, l, and e, as shown in Table 2. I have chosen to write each fermion as a left-handed spinor of the Lorentz groups, so that u, d, and e are left-handed antiquarks and antileptons. In Table 2, the number of states for each of the five representations is shown in parenthesis, the total being 15 for each of the three generations.

Table 2. The Aperiod Table

	$SU(3)$	$SU(2)$	$U(1)_Y$
$q(6)$	3	2	$\frac{1}{6}$
$u(3)$	$\overline{3}$	—	$-\frac{2}{3}$
$d(3)$	$\overline{3}$	—	$\frac{1}{3}$
$l(2)$	—	2	$-\frac{1}{2}$
$e(1)$	—	—	1

The known gauge interactions distinguish between the 15 states of a generation, but do not distinguish between the three generations; they break the flavor symmetry group from $U(45)$ to $U(3)^5$, with one $U(3)$ factor acting in generation space on each of the five multiplets q, u, d, l, and e.

This $U(3)^5$ symmetry is broken in hierarchical stages by the quark and lepton mass matrices. For example, the up quark matrix provides an explicit breaking of $U(3)_q \times U(3)_u$ transforming as a (3, 3). The largest entry in the matrix is clearly the top quark mass, which strongly breaks this group to $U(2)_q \times U(2)_u \times U(1)_{q_3 - u_3}$. The fermion mass problem, which is part of the flavor puzzle, is the question of why the quark and lepton mass matrices break $U(3)^5$ in the hierarchical fashion measured by experiment. Since we are dealing with matrices, a solution of this problem would provide an understanding of both quark and lepton masses and the Kobayashi-Maskawa mixing matrix. All questions about the quark and lepton masses and mixings can be rephrased in terms of $U(3)^5$ breaking. For example, "why is $m_t \gg m_b$?" becomes "why is the breaking $U(3)_u \to U(2)_u$ stronger than that of $U(3)_d \to U(2)_d$?" In the context of the Standard Model, this rephrasing does not seem very important; however, in the context of supersymmetry, it is of great importance.

I.3 The Major Problems of the High-Energy Frontier

All physicists should spend a great deal of time debating and deciding what are the most important issues in their subfield. At the high-energy frontier, I think the four most important puzzles are:

1. What breaks $SU(2) \times U(1)$?

   The weak interactions appear weak and are short range, because they, alone among the known forces, are generated from a symmetry group which is broken. Perturbative gauge forces do not break themselves—new interactions are required to break them. Such a fifth force must exist and be accessible to experiments designed to probe the weak scale. It is guaranteed to be exciting: it has a dynamic which is different from any of the known forces, and it should shed light on the fundamental question of what sets the mass scale of weak symmetry breaking. I will call this mass scale M_Z, even though the weak symmetry breaking mechanism of the fifth force is responsible for the dominant contribution to the mass of all of the known massive elementary particles.

2. What breaks the $U(3)^5$ flavor symmetry?

   We know that this flavor symmetry is broken at least to $B \times L_i$ because of the observed quark and lepton masses and the Kobayashi-Maskawa mixing matrix. However, such masses and mixings cannot simply be inserted into the theory because they break $SU(2)$; they must originate from some new interactions which break $U(3)^5$. In the Standard Model, these new interactions are the Yukawa couplings of the Higgs boson, but there are other possibilities. We might call these $U(3)^5$ breaking interactions the "sixth force." I think that future experiments will uncover this force also, at least the pieces of it which are strong and are responsible for the large top quark mass. Whatever the description of $U(3)^5$ breaking at the weak scale, there is still the puzzle as to why $U(3)^5$ is hierarchically broken. I think that physics at the weak scale could shed light on some aspects of this; but this is much more uncertain. It is likely that some, and perhaps all, of the understanding of flavor physics occurs at some very much higher energy scale. Nevertheless, at the very minimum, experiments must be done which uncover the weak scale description of $U(3)^5$ breaking, i.e., the sixth force. I find a sense of excitement building up in our field as experiments enter the domain where signals of the fifth and sixth forces will be discovered.

3. Why are the symmetries and fundamental constants of nature what they are?

   The most basic properties of nature can be summarized in terms of a set of gauge, flavor, and spacetime symmetries, and a set of fundamental parameters, such as the gauge couplings and the quark and lepton masses. The

next question is embarrassingly obvious: Why these symmetries and why these values of the parameters? The anthropic argument, that without them we could not exist to make the observations, is fraught with problems; it seems to me better to look boldly for a true theory. A complete answer to these questions requires going beyond four-dimensional, point particle quantum field theory, and at the moment, superstring theory provides the unique such direction. However, string theory is very ambitious, and despite exciting developments, the time scale for making definitive connections to physics is completely unknown. The central thesis of these lectures is that we may already have the basic tools required to make considerable progress in furthering our understanding of nature. The familiar tools of unified gauge symmetries, flavor symmetries, and the properties of supersymmetry and the renormalization group can carry us very far and can be tested by experiment. The gauge group $SO(10)$ explains the quantum numbers of Table 2. If the 15 known states of a generation, together with a right-handed neutrino, are placed in the 16 dimensional spinor representation of $SO(10)$, then every entry of Table 2 follows from the simple group theoretic embedding of $SU(3) \times SU(2) \times U(1)$ into $SO(10)$. This is an extraordinary achievement. The vertical unification of a generation also reduces the flavor symmetry group from $U(3)^5$ to $U(3)$, which is much more constraining. Such grand unified theories can reduce the number of free parameters on which all of low-energy physics depends. Several supersymmetric theories based on the flavor group $U(3)$, or on one of its subgroups, have been developed recently and make many predictions for the flavor-changing interactions of the superpartners. Such grand unified theories of flavor are not the ultimate theory, but they can explain a great deal very simply. For grand unified and flavor symmetries, the real question is: how can they be subjected to experimental tests? I will begin the answer to this question in these lectures.

4. How is a quantum theory of gravity to be constructed?

   Superstring theory provides the only known direction for progress.

I.4 Supersymmetry

The current interest in supersymmetry is largely because it offers interesting new directions for attacking each of the above problems. In summary, these new directions are:

1. Supersymmetry is the only symmetry which can give rise to a light, elementary Higgs boson for electroweak symmetry breaking. The puzzle of the scale of weak interactions is replaced with the puzzle of the origin of the scale of supersymmetry breaking.

2. The hierarchical breaking of $U(3)^5$ governs not only the form of the Yukawa interactions of the Higgs, but also the squark and slepton mass matrices. Since the latter are severely constrained by flavor-changing phenomenology, severe restrictions are placed on the group theoretic structure of the pattern of $U(3)^5$ breaking. In addition, supersymmetry allows for the possibility that above the weak scale, some of the $U(3)^5$ breaking which generates the quark and lepton masses arises from the scalar mass matrices rather than from the Higgs-Yukawa interactions.

3. Supersymmetric grand unification provides a successful prediction, at the percent level, of the weak mixing angle. Although less significant, m_b/m_t and m_t can also be successfully predicted in supersymmetric unified models. With further simplifying assumptions, such as the nature and breaking of the flavor group, other predictions can also be obtained.

4. A supersymmetric string theory offers the prospect of a quantum theory of gravity, unified with the other forces.

In these lectures, I will elaborate on the first three of the above: $SU(2) \times U(1)$ breaking, flavor symmetry breaking, and supersymmetric grand unification, in Chaps. II, III, and IV, respectively.

There are many excellent books and review articles on supersymmetry,[1] the supersymmetric extension of the Standard Model,[2] and supersymmetric grand unification. The aim of the present lectures is not to refine or update these works, but to explain why I think the study of supersymmetry is interesting, why the direct search for superpartners is of crucial importance, and what may be learned from a variety of other measurements. Nevertheless, it may be useful to say a few words about supersymmetry and the supersymmetric extension of the Standard Model.

Supersymmetry is an extension of the Poincaré group of spacetime transformations. Spinorial generators, Q and $\overline{Q}$, are added to the usual generators p, J, and K of translations, rotations, and boosts. The only nontrivial extension of the Poincaré algebra involving Q or $\overline{Q}$ is the anticommutation $\{Q, \overline{Q}\} = p$. Consider the evolution of our understanding of the spacetime properties of the electron. When discovered nearly a century ago[†] by J. J. Thompson, it was conceived as a negatively charged particle with just two properties: its mass and electric charge. We view the charge as a consequence of the behavior with respect to the electromagnetic $U(1)$ charge generator, and the mass as a consequence of the translation generator p. The discoveries of Stern and Gerlach dictated that it should be given another attribute, intrinsic spin, which describes its properties with respect to the angular momentum generator, J. The splitting of an atomic beam by an inhomogeneous magnetic field, which they discovered in 1922, is caused by the doubling of the number of electron states which follows from their nontrivial properties under the angular momentum generator: $e \xrightarrow{J} (e^{\uparrow}, e^{\downarrow})$. In the relativistic case, this description is inadequate. The Lorentz boost generator K requires a further doubling of the number of particle states; we call the resulting Lorentz-partners the antiparticles: $e \xrightarrow{K} (e, \overline{e})$. Their properties are dictated by Lorentz symmetry, having equal mass and opposite charge to the particles.

The extension of spacetime symmetries which results from the introduction of the supersymmetry generator, Q, causes a further doubling of the particles: $e \xrightarrow{Q} (e, \widetilde{e})$; while $\overline{e}$ is the Lorentz-partner of the electron, $\widetilde{e}$ is the supersymmetry-partner, or superpartner, of the electron. It has properties which are determined by the supersymmetry algebra: the mass and charge are identical to that of the electron, but because Q is spinorial, it has intrinsic spin which differs by 1/2 relative to the electron; it is a Lorentz scalar. Many people laugh when they hear about supersymmetry and how it leads to the introduction of a new hypothetical particle for each of the observed particles. However, it is just history repeating itself; perhaps physicists of old laughed at the prospect of antielectrons and antiprotons, but the sniggering soon stopped.

The super-electron is not degenerate with the electron; supersymmetry, if it exists, must be sufficiently broken that the s-electron mass is larger than about 65 GeV. The discovery of supersymmetry would be doubly exciting: not only would it herald an exciting new era of spectroscopy, but it would represent the

[†]I expect we will have celebrations in 1997 for the centenary of the discovery of the first particle which, as far as we know today, is elementary.

discovery of a completely new type of symmetry: a broken spacetime symmetry. The empty box of Table 1 would be filled by Q; nature would have provided examples of all six varieties of symmetries. What could be more interesting?

I.5 Summary

Three types of symmetries are shown in Table 1: local internal, global internal, and global spacetime, which I shall frequently call gauge, flavor, and spacetime symmetries, respectively. Each of these types of symmetry may be broken at scales beneath the Planck scale M_{Pl}. In these lectures, I consider the breaking of a unified group

$$G_{unified} \stackrel{M_G}{\longrightarrow} SU(3) \times SU(2) \times U(1) \stackrel{M_Z}{\longrightarrow} SU(3) \times U(1), \qquad (I.1)$$

the breaking of the flavor symmetry group $G_f \subset U(3)^5$

$$G_f \stackrel{M_F}{\longrightarrow} B \times L_i, \qquad (I.2)$$

and the breaking of supersymmetry

$$(p, J, K, Q, \overline{Q}) \stackrel{M_S}{\longrightarrow} (p, J, K). \qquad (I.3)$$

The mass scales represent the scales of the vacuum expectation values of fields which break the symmetry. There could be several stages of breaking of the unified gauge group, and there will almost certainly be several stages in the sequential breaking of the flavor group, so M_G and M_F represent a set of scales. Assuming that only one supersymmetry survives beneath M_{Pl}, M_S is unique. In the limit that $M_S \to 0$, the superparticle and particle masses become degenerate; however, in most schemes of supersymmetry breaking, the mass scale m_s of the superpartners of the known particles is not given by M_S. For example, in supergravity $m_s = M_S^2/M_{Pl}$, and in dynamical supersymmetry breaking models $m_s = \alpha M_S^2/M_X$, where M_X is some other mass scale larger than M_S. The scale M_X or M_{Pl} is known as the messenger scale, M_{mess}; it is the energy scale below which the superpartners possess local supersymmetry breaking masses and interactions.

There is no guarantee that M_F is less than M_{Pl}. The physics of flavor may be understood only at the Planck scale. Indeed, of all the mass scales introduced in this subsection, M_F is perhaps the most uncertain. If $M_F \approx M_{Pl}$, then G_f breaking interactions must occur explicitly at the boundary at M_{Pl}, with small dimensionless coefficients. An advantage to having M_F beneath M_{Pl} is that the

small dimensionless fermion mass ratios can then appear as ratios of these scales. In Chap. III, we will explore the case of $M_F < M_{Pl}$, which allows for an understanding of at least some aspects of flavor beneath M_{Pl}.

II. $SU(2) \times U(1)$ Breaking and the Weak Scale

II.1 A Symmetry Description

In the Standard Model, the $SU(2) \times U(1)$ electroweak symmetry is broken by introducing a Higgs sector to the theory, which involves an electroweak scalar doublet, h. The mass squared parameter for this field, m_h^2, determines the order parameter of the symmetry breaking: if it is negative, the electroweak symmetry breaks, while if it is positive, all the elementary particles are massless. The Higgs sector certainly provides an economical description of electroweak symmetry breaking, but it is inadequate for two reasons. There is no dynamical understanding of why symmetry breaking occurs; one simply inserts it into the theory by hand by making m_h^2 negative. Secondly, there is no symmetry understanding of the scale of the breaking, which I refer to as the Z mass, M_Z.

In physics, we have learned that mass scales should be both described and understood in terms of symmetries. Great progress has been made in providing symmetry descriptions of phenomena, but understanding the origin of the symmetry behavior at a deeper level often eludes us, as we illustrate with a few examples.

Why is the photon massless? The symmetry description is clear: electromagnetic gauge invariance is unbroken. However, the deeper question is: *why* is it unbroken? This brings us back to the breaking of $SU(2) \times U(1)$ electroweak symmetry. Why is it accomplished by a single doublet, reducing the rank by one but not by two?

Why are the neutrinos massless? A symmetry description is that nature possesses lepton number as an exact global symmetry. At a deeper level, however, many questions arise: why are there no right-handed neutrinos, why is the lepton number exact? If the neutrinos do have small masses, why are the lepton numbers such good approximate symmetries? An interesting feature of supersymmetric theories is that the standard answers to these questions are inadequate, as discussed in Secs. II.2 and III.7.

Why do the quark and charged leptons have their observed masses? Since the masses break the electroweak symmetry, they can be written as λv, where v is the dimensionful order parameter of the symmetry breaking and λ is a dimensionless

parameter, different for each quark and lepton. The overall scale of the masses is determined by v, while the mass ratios are determined by ratios of λ couplings. Many of the λ are small, which we describe in Chap. III in terms of approximate flavor symmetries. But what is the origin for these symmetries and their breaking? Why are there three generations? Why is the up quark so much lighter than the top quark: $\lambda_{up}/\lambda_{top} \approx 10^{-5}$?

What is the origin of the hadronic mass scale of the proton and neutron? This scale is the scale at which the QCD coupling constant, α_s, becomes large and non-perturbative. It arises, through renormalization, as a dimensional transmutation of this gauge coupling, and hence, is described in terms of the QCD symmetry group, $SU(3)$.

These examples illustrate how we turn to symmetries for both a description and a deeper understanding of the phenomena. This applies to all phenomena of particle physics, but here I stress the application to masses.

Now we can better appreciate the inadequacy of the Standard Model Higgs sector description of electroweak symmetry breaking. What symmetry description or understanding does it proscribe for the order parameter v which determines M_Z and the fermion masses? *None.* The crucial point is that it does not even provide a symmetry description for the scale v, let alone any deep understanding. Because the Standard Model Higgs sector is so economical, and because the Standard Model provides an accurate description of so much data, many have concluded that the Standard Model will be the final story—there will be no physics beyond the Standard Model. I strongly disagree with this viewpoint. First, there is not a shred of evidence for the Standard Model Higgs sector, but, more importantly, our experience in physics tells us that the physics responsible for electroweak symmetry breaking will, at the very least, allow a description of the mass scale in terms of a symmetry.

What will this new symmetry be? There are many possibilities, but it is useful to group them according to the fate of the hypothetical Higgs boson. There are three logical possibilities:

1. There is no Higgs boson.
2. The Higgs boson is composite (at a scale close to the weak scale).
3. The Higgs boson is elementary.

The first option is realized in technicolor theories where the weak scale arises by dimensional transmutation from a gauge coupling, just like in QCD. The second

option can also be realized by having a new strong gauge force. In this case, the new strong force first produces a composite scalar bound state, which then becomes the Higgs boson of electroweak symmetry breaking. In both of these examples, the symmetry description of the weak scale is in terms of the symmetry group of some new gauge force.

The third option is quite different. The only known symmetry description for a fundamental Higgs boson involves supersymmetry. The lightness of the Higgs may be related to a chiral symmetry acting on its fermionic superpartner, or it may be due to the Higgs being a pseudo-Goldstone boson. In either case, the weak scale is the scale at which supersymmetry is broken. To get a deeper understanding of the weak scale, one must then address the question of how supersymmetry is broken. Presumably, the reason for why the weak scale is much less than the Planck scale is the same as for the technicolor and composite Higgs options: it occurs as a dimensional transmutation due to the strong dynamics of a new interaction. Whereas in the technicolor case, one can simply appeal to the analogy with QCD; in the supersymmetry case there is no analogy—nature has not provided us with other examples of broken spacetime symmetries—hence, there is no substitute for understanding the dynamics of the field theory.

II.2 Matter vs. Higgs

In the Standard Model, it is obvious what distinguishes matter fields, the quarks and leptons, from the Higgs field: matter fields are fermions, while Higgs fields are bosons. In supersymmetry, this distinction disappears! Once superpartners are added, there is no spacetime distinction between quarks $(q, \widetilde{q})$, leptons $(\ell, \widetilde{\ell})$, and Higgs $(\widetilde{h}, h)$ supermultiplets, since each contains a fermion (q, ℓ, or $\widetilde{h}$) and a boson ($\widetilde{q}, \widetilde{\ell}$, or h). Indeed, the distinction between the lepton doublet and the Higgs doublet becomes a puzzle of fundamental importance. Since these have the same gauge quantum members, what is the theoretical distinction between the Higgs and the lepton superfield?

Supersymmetry apparently allows us to do without a Higgs supermultiplet: why not identify the Higgs boson with one of the sneutrino fields, $\widetilde{\nu}$? If there are three generations of matter, then this is not possible: a sneutrino vev $\langle\widetilde{\nu}\rangle$ leads to a Dirac mass of size M_Z coupling the corresponding ν state to the $\widetilde{Z}$. Such a theory would only have two neutrinos of mass less than M_Z. The sneutrino as Higgs idea is so attractive, that it is worth considering the Higgs to be the sneutrino of a fourth generation. In this case, it is the fourth neutrino which marries the

$\widetilde{Z}$ to acquire mass M_Z, which has the added advantage of explaining why only three neutrinos are seen in the Z width. The problem with this scheme is that supersymmetry forbids a tree-level coupling of the sneutrino to the up type quarks: the t and t' masses would have to occur via radiative corrections. Given these large masses, this would necessarily involve new nonperturbative interactions. With just four generations of chiral superfields, and the known gauge interactions, the only interactions which could break the chiral symmetry on u_R is the trilinear scalar interaction $\tilde{q}\tilde{u}\tilde{\ell}^\dagger$. Such nonholomorphic supersymmetry breaking interactions are not usually considered—however, they do not introduce quadratic divergences. This interaction is asymptotically free, so that it could become nonperturbative at low energies. However, it is very unclear whether it could give rise to sufficiently large masses for t and t' quarks.

Perhaps the above line of reasoning has not been developed further because the unification of gauge couplings in supersymmetric theories suggests that there are two light Higgs supermultiplets at the weak scale which are distinct from the matter. The conventional picture of weak scale supersymmetry has Higgs superfields, h_1 and h_2, which are distinct from the lepton superfields, although the origin of the distinction indicates that there must be yet another symmetry. The nature of this symmetry is discussed in Sec. III.7.

II.3 A Heavy Top Quark Effect

As mentioned in Sec. II.1, supersymmetry is the only known tool that allows a fundamental Higgs boson at the electroweak scale to be understood in terms of symmetries. This understanding has two aspects:

- The size of $|m_h^2|$ is controlled by the scale of supersymmetry breaking, which is presumably determined by some strong dynamics leading to a dimensional transmutation. Candidate field theories for this exist, but we are far from having a standard picture for the origin of supersymmetry breaking, and I will not discuss it further in these lectures.
- The sign of m_h^2 is controlled by the dynamics which connects the particles of the Standard Model to the supersymmetry breaking interactions, and also by radiative corrections to m_h^2. A given model makes this dynamics explicit, and, if it is perturbative, the sign of m_h^2 is calculable.

In the most popular schemes for giving mass to the superpartners, the supergravity and gauge messenger schemes mentioned in Sec. I.5, the messenger

dynamics is perturbative and leads to positive mass squareds for all scalars in the theory. This makes the issue of how $SU(2) \times U(1)$ breaks, i.e., of why m_h^2 is negative, particularly pressing. In particular, what distinguishes the Higgs boson from the other scalars in the theory, the scalar quarks and leptons, which must have positive mass squareds?

The answer to this puzzle is made plausible by its simplicity. There are two important radiative corrections to any scalar mass, m^2:

- gauge contributions, which increase m^2, and
- Yukawa contributions, which typically decrease m^2.

The only important Yukawa radiative corrections are induced by the large top Yukawa coupling λ_t.[‡] Hence, all m^2 are kept positive by the gauge radiative corrections, with the possible exceptions of m_h^2 and $m_{\tilde{t}}^2$, since only h and $\tilde{t}$ couple to λ_t. The λ_t^2 radiative correction is more powerful for m_h^2 than for $m_{\tilde{t}}^2$, meaning that it is m_h^2 which has the greater tendency to go negative. This is due to the fact that colored triplets have a larger multiplicity than weak doublets: $SU(2)$ breaks rather than $SU(3)$ because it is a smaller group. Once m_h^2 is negative, the Yukawa corrections to $m_{\tilde{t}}^2$ actually change sign, preventing $m_{\tilde{t}}^2$ from becoming negative. In addition, $m_{\tilde{t}}^2$ has QCD radiative corrections which also make it more positive than m_h^2.

Electroweak symmetry breaking is therefore understood to be a large top quark mass effect; a result which was obtained before the top quark was known to be very heavy.[3,4] Keeping other parameters of the theory fixed, λ_t is the order parameter for electroweak symmetry breaking in supersymmetric models. For low values of λ_t, $SU(2) \times U(1)$ is unbroken, whereas for high values of λ_t, it is broken. The critical value for λ_t does depend on other parameters of the theory, for example, the superpartner masses. However, now that we know that the top quark is about 175 GeV, λ_t is above the critical value for a very wide range of parameters. I am tempted to say that electroweak symmetry breaking is hard to avoid, but such a statement would require a detailed numerical study.

The size of $|m_h^2|$, and therefore M_Z, and the superpartner masses are both determined by the scale of supersymmetry breaking. Does this allow a prediction of the masses of the superpartners? Since there is more than one supersymmetry breaking parameter, the answer is no. Nevertheless, the understanding of the weak scale from symmetry principles requires that the superpartners not be much

[‡]The b and τ Yukawa couplings could also be large, in which case the conclusions of this section are strengthened.

heavier than M_Z. Denote the set of supersymmetry breaking parameters by the scale m_s and the dimensionless parameters a. For example, m_s could be defined to be the mass of the lightest chargino, and one of the a parameters would be the ratio of the top squark mass to this chargino mass. Since M_Z has its origin in supersymmetry breaking, it is necessarily given by a formula of the form $M_Z^2 = m_s^2 f(a)$. The scale of the superpartner masses, m_s, can be made much larger than M_Z only at the expense of a fine tuning amongst the a parameters to make $f(a)$ small. Hence:

- We cannot predict the mass of the superpartners. (Certain superpartner mass ratios are predicted in given messenger schemes, and in certain theories with flavor symmetries, and are important tests of these theories.)
- The superpartner mass scale, m_s, can be made much larger than M_Z only by a fine tune between dimensionless parameters which increases as m_s^2/M_Z^2.

The amount of fine tuning can be characterized by the sensitivity of M_Z^2 to small changes in the a parameters: $c_a = (a/M_Z^2)\delta M_Z^2/\delta a$ (Ref. 5). A refined definition of the sensitivity parameter, $\gamma_a = c_a/\bar{c}_a$, has been advocated, where $\bar{c}_a$ is an average of c_a (Ref. 6). Although there are no rigorous, mathematical upper bounds on the superpartner masses, it is possible to give upper bounds on the superpartner masses if the amount of fine tuning, taken to be $\tilde{\gamma}$, the largest of the γ_a, is restricted to be less than a certain value. Such naturalness bounds are shown for the Higgs scalar masses as well as the superpartner masses in the figure. The upper extent of the line corresponds to $\tilde{\gamma} = 10$, the error bar symbol to $\tilde{\gamma} = 5$, and the squares give values of the masses for which the fine tuning is minimized. This plot applies to the case of universal boundary conditions on the scalar masses at very high energies. Relaxing this condition will allow some superpartner masses, for example, the scalars of the first two generations, to increase substantially. However, there will still be several superpartners, such as the lighter charginos (χ^+), the lighter neutralinos (χ^0), and the top squarks, which will prefer to be lighter than 300 GeV. The absence of any superpartners beneath 1 TeV would mean that the understanding of the weak scale described in this chapter has very serious problems. LEP II and the Fermilab Main Injector are well-positioned to discover supersymmetry, although the absence of superpartners at these machines would not be conclusive.

III. Flavor in Supersymmetric Theories

III.1 The Fermion Mass and Flavor-Changing Problems

In nature, fermions exist in 45 different helicity states. What is the origin of these states, and why do they assemble into three generations of quarks and leptons with such diverse masses, mixings, gauge, and global quantum numbers? This is the flavor problem. Two important aspects of the flavor problem are:

1. The fermion mass problem. What is the origin of the observed hierarchy of quark and lepton masses and mixings?

   Models of particle physics can be divided into two groups. *Descriptive models* are those which describe the observed quark and lepton masses and mixings with 13 free parameters and make no attempt to understand the hierarchies. The Standard Model is a descriptive model. *Predictive models* are those which either describe the 13 observed masses and mixings with fewer than 13 parameters, or which provide some understanding of the mass and mixing angle hierarchies.

2. The flavor-changing problem. Why are processes which involve flavor-changing neutral currents (FCNC) so rare? Three such highly suppressed quantities are $\Delta m_K, \epsilon_K$, and the rate for $\mu \to e\gamma$.

Coupling constants which distinguish between generations are called flavor parameters, and include the parameters which generate the observed quark and lepton masses and mixing. In the Standard Model, there are 13 flavor parameters, precisely one for each of the 13 observed fermion masses and mixings, and they all originate from the Yukawa coupling matrices. In extensions of the Standard Model, there may be more flavor parameters, so that they cannot all be experimentally determined from the quark and lepton masses and mixings.

A model is considered natural if it suppresses FCNC processes for *generic* values of the flavor parameters, i.e., for a wide range of the parameters that is consistent with the observed fermion masses and mixing. The Standard Model is natural in this sense: all the Yukawa parameters are determined from the experimentally measured fermion masses and mixings, and the GIM mechanism[7] ensures the smallness of FCNC processes. For models with more flavor parameters, we must address the question of what values of the parameters are generic.

In this chapter, I assume that below some high scale Λ, physics is described by a softly broken, supersymmetric $SU(3) \times SU(2) \times U(1)$ gauge theory of minimal field content: three generations of quark and lepton superfields q_i, u_i, d_i, l_i, and e_i,

and two Higgs doublet superfields h_1 and h_2. Assuming invariance under R parity, the flavor parameters of this theory can be written as 11 matrices in generation space. Three of these are Yukawa coupling matrices of the superpotential

$$W = q\boldsymbol{\lambda}_U u h_2 + q\boldsymbol{\lambda}_D d h_1 + \ell\boldsymbol{\lambda}_E e h_1. \qquad (III.1)$$

The supersymmetric interactions have identical flavor structure to the Standard Model and lead to a supersymmetric GIM mechanism suppressing FCNC effects. The other eight matrices contain soft supersymmetry breaking parameters

$$\begin{aligned} V_{soft} &= \tilde{q}\boldsymbol{\xi}_U \tilde{u} h_2 + \tilde{q}\boldsymbol{\xi}_D \tilde{d} h, + \tilde{\ell}\boldsymbol{\xi}_E \tilde{e} h_1 + h.c. \\ &\quad + \tilde{q} m_q^2 \tilde{q}^\dagger + \tilde{u}^\dagger m_u^2 \tilde{u} + \tilde{d}^\dagger m_d^2 \tilde{u} + \tilde{\ell} m_\ell^2 \tilde{\ell}^\dagger + \tilde{e}^\dagger m_e^2 \tilde{e}. \end{aligned} \qquad (III.2)$$

If these eight matrices are given values which are "generic," that is, the size of any entry in a matrix is comparable to the size of any other entry, then loop diagrams involving superpartners lead to very large FCNC effects, even for superpartners as heavy as 1 TeV (Ref. 8). For example, the quantities ϵ_K and $\Gamma(\mu \to e\gamma)$ are about 10^7 larger than allowed by experiment. This is the flavor-changing problem of supersymmetry.

Over the last few years, an interesting new development has occurred. Progress has been made simultaneously on the fermion mass and flavor-changing problems of supersymmetry by introducing flavor symmetries which constrain the forms of both the Yukawa couplings of Sec. III.1 and the scalar masses and interactions of Sec. III.2. In the symmetry limit, many of the Yukawa coupling entries vanish, and the form of the scalar masses are strongly constrained. Small hierarchical breakings of the flavor symmetry introduce small parameters that govern both the small masses and mixings of the fermions, and the small violations of the superGIM mechanism which give small contributions to FCNC processes. This linking of two problems is elegant and constraining; it is so simple that it is hard to understand why it was not explored in the early '80s. Perhaps we are taking supersymmetry more seriously these days.

In Sec. III.5, I will discuss the literature on this subject, which began in 1990 and has grown into a minor industry recently. Each of the papers to date studies a particular flavor symmetry, G_f, and a particular breaking pattern. Many of the models illustrate a special point or aim for a particular fermion mass prediction. In Secs. III.2 and III.3 below, my aim is to demonstrate the generality and power of this approach. In fact, from this viewpoint, I argue that the flavor-changing problem has arisen because of an unreasonable definition of "generic." We know

from the observed masses and mixings of quarks that $\lambda_{D_{12}}$ and $\lambda_{D_{21}}$ are very small. A solution to the fermion mass problem would give us an understanding of why this is so, but no matter what the understanding, the flavor symmetries acting on the down and strange quarks are broken only very weakly. Experiment has taught us that approximate flavor symmetries (AFS) are a crucial aspect of flavor physics. It is therefore quite unreasonable to take $m^2_{q_{12}} \approx m^2_{q_{11}}$; the former breaks strange and down flavor symmetries and hence should be very suppressed compared to the latter, which does not. (A crucial difference between scalar and fermion mass matrices is that the diagonal entries of fermion mass matrices break Abelian flavor symmetries, while diagonal entries of scalar mass matrices do not.)

In this chapter, I explore the consequences of linking the flavor-changing problem to the fermion mass problem. I require that *all flavor parameters of the theory are subject to the same approximate flavor symmetries.* I take this to be an improved meaning of the word "generic" in the statement of the flavor-changing problem. With this new viewpoint, it could be that there is no flavor-changing problem in supersymmetry. Perhaps if one writes down the most generic soft parameters at scale Λ, the FCNC processes are sufficiently suppressed.

Let G_f be the approximate flavor symmetry group of the theory below scale Λ, and suppose that G_f is explicitly broken by some set of parameters $\{\epsilon(R)\}$, which transform as some representation R of G_f, and take values which lead naturally to the observed pattern of fermion masses and mixings. We will discover that for some G_f and $\{\epsilon(R)\}$, the flavor problem is solved, while for others it is not. Hence, the flavor-changing problem of supersymmetry is transformed into understanding the origin of those G_f and $\{\epsilon(R)\}$ which yield natural theories.

Below scale Λ, models are typically (but not always) *descriptive*; they do not provide an understanding of the fermion masses. However, knowing which $G_f, \{\epsilon(R)\}$ solve the flavor-changing problem serves as a guide to building *predictive* models above Λ. The theory above Λ should possess an exact flavor symmetry G_f that is broken spontaneously by fields $\{\phi\}$, which transform as R under G_f and have vacuum expectation values $\langle\phi\rangle = \epsilon\Lambda$.

In Sec. III.2, I introduce the ideas of Approximate Flavor Symmetries (AFS), and in Sec. III.3, I give a set of simple conditions which are sufficient for an AFS to solve the flavor-changing problem. In Sec. III.4, I show that the flavor-changing problem is solved when G_f is taken to be the maximal flavor symmetry. I delay a discussion of previous work on this subject until Sec. III.5. In Sec. III.6, I discuss the case $G_f = U(2)$, where the flavor-changing constraints dictate a

special and interesting texture for the fermion mass matrices. In Sec. III.7, I show that R parity finds a natural home as a subgroup of the flavor symmetry. Sections III.5 and III.7 are taken from Ref. 27. This chapter is the most technical of these lectures; a brief statement of the conclusions is given in Sec. III.8.

III.2 Approximate Flavor Symmetries

Using approximate flavor symmetries to describe the breaking of flavor is hardly new, but it is certainly powerful. QCD with three flavors has an approximate flavor symmetry $G_f = SU(3)_L \times SU(3)_R$, explicitly broken by a various parameter $\{\epsilon(R)\}$, which includes the quark mass matrix $M(3,\bar{3})$, and electric-charge matrices $Q_L(8,1)$ and $Q_R(1,8)$. Below Λ_{QCD}, the flavor symmetries are spontaneously broken to the vector subgroup and G_f is realized nonlinearly. The interactions of the Goldstone bosons can be described by constructing an invariant chiral Lagrangian ($\mathcal{L}$) for $\Sigma(3,\bar{3}) = exp(2i\pi/f)$. For our purposes, the crucial point is that the flavor symmetry breaking beneath Λ_{QCD} can be described by constructing the chiral Lagrangian to be a perturbation series in the breaking parameters $\{\epsilon\} = \{M, Q_L, Q_R ...\}$. Thus, $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2 + ...$ where $\mathcal{L}_N$ contains terms of order ϵ^N. For example,

$$\mathcal{L}_1 = a_1 \Lambda_{QCD}^3 Tr(M\Sigma^\dagger) + ... \qquad (III.3a)$$

$$\mathcal{L}_2 = a_2 \Lambda_{QCD}^2 Tr(M\Sigma^\dagger M\Sigma^\dagger) + a_3 \Lambda_{QCD}^4 Tr(Q_L \Sigma Q_R \Sigma^\dagger) + ... \qquad (III.3b)$$

where all the unknown dynamics of QCD appear in the set of dimensionless strong interaction parameters $\{a\}$, which are $O(1)$. This illustrates the basic tool which we use in this chapter.

The full-flavor symmetry of the 45 fermions of the Standard Model is $U(45)$. This is broken to the group $U(3)^5$ by the Standard Model gauge interactions. Each $U(3)$ acts in the three-dimensional generation space and is labeled by A, which runs over the five types of fermion representation (q, u, d, ℓ, e).

The $U(3)^5$ flavor symmetry of the Standard Model gauge interactions is broken explicitly by the Yukawa couplings of the Standard Model, which have the transformation properties

$$\begin{aligned} &\boldsymbol{\lambda}_U \ (\bar{3}, \bar{3}, 1, 1, 1) \\ &\boldsymbol{\lambda}_D \ (\bar{3}, 1, \bar{3}, 1, 1) \\ &\boldsymbol{\lambda}_E \ (1, 1, 1, \bar{3}, \bar{3}). \end{aligned} \qquad (III.4)$$

In this section, we speculate that these Yukawa parameters result from some new physics above scale Λ, which possesses an AFS G_f, broken explicitly by a set of parameters $\{\epsilon(R)\}$. The theory beneath Λ can be written as a perturbation series in the ϵ. The Standard Model gauge Lagrangian appears at zeroth order, while the flavor-violating fermion masses appear at higher order.

Such a picture is not new: the composite technicolor standard models were based on this picture.[9] In this case, the theory above Λ was taken to be a preonic theory with strong dynamics which leaves a $U(3)^5$ flavor symmetry unbroken. The strong dynamics produces composite quarks, leptons, and Higgs bosons. The preonic theory contains parameters $\{\epsilon(R)\}$ which explicitly break $U(3)^5$; in fact, these parameters are assumed to be preon mass matrices $M_{U,D,E}$ with the same transformation properties as $\boldsymbol{\lambda}_{U,D,E}$. At first order in perturbation theory, $\boldsymbol{\lambda}_{U,D,E}$ are generated proportional to $M_{U,D,E}$. At higher order, various phenomenologically interesting four-quark and four-lepton operators are generated. For example, the operator $1/\Lambda^6(qM_UM_U^\dagger q)(qM_UM_U^\dagger q)$ leads to an additional contribution to ϵ_K.

This picture is very close to that adopted here, except that:

(a) The theory beneath Λ is one with softly broken supersymmetry, and contains eight flavor matrices in the soft supersymmetry breaking interactions in addition to the three supersymmetric Yukawa matrices.

(b) A large variety of AFS groups G_f and explicit symmetry breaking parameters $\{\epsilon(R)\}$ are of interest. In Sec. III.4, we consider the obvious possibility that $G_f = G_{max} = U(3)^5$, and $\{\epsilon(R)\} = \boldsymbol{\epsilon}_U, \boldsymbol{\epsilon}_D, \boldsymbol{\epsilon}_E$ transforming as $\boldsymbol{\lambda}_{U,D,E}$ are the only symmetry breaking parameters.

(c) The more fundamental theory above Λ need not involve strong, nonperturbative dynamics. Each possible term in the low-energy theory will be given an arbitrary dimensionless coefficient (labelled by $\{a\}$), which we think of as being $O(1)$ if the dynamics at Λ is strong. However, if the dynamics at Λ is perturbative, then $\{a\}$ will be less than unity, and the flavor-changing effects will be milder.

As a final example of the previous use of AFS, we consider the Standard Model extended to contain several Higgs doublets. It was frequently argued that these theories had a flavor-changing problem. Those doublets orthogonal to the one with a vev could have Yukawa matrices unconstrained by fermion masses. With all such couplings of order unity, the tree-level exchange of such Higgs bosons generates large FCNC for fermion interactions, such as $(1/m_h^2)(q_1d_2)^2$ for Δm_K

and ϵ_K. For theories with several Higgs doublets at the weak scale, this flavor problem was frequently solved by imposing a discrete symmetry which allowed only a single Higgs to couple to the u_i and only a single Higgs to the d_i quarks.[10]

From the viewpoint of AFS, however, such discrete symmetries are unnecessary.[11,12] Suppose the Higgs doublet which acquires a vev is labelled h_1. The hierarchical pattern of quark masses implies that the Yukawa interactions of h_1 possess an AFS. It is unreasonable that $h_{2,3...}$ should have interactions which are all $O(1)$ and are unconstrained by these AFS. If one set of interactions possesses an AFS, it is only natural that the entire theory is constrained by the same AFS. One possibility is that the AFS of the quark sector $G_Q = U(1)^9$, a $U(1)$ factor for each of q_i, u_i, and d_i (Refs. 11 and 12), with each $U(1)$ having its own symmetry breaking parameter. Thus ϵ_{q_i} transforms under $U(1)_{q_i}$ but not under any other $U(1)$, etc. In this case, all Yukawa couplings of h_a to up quarks would have the structure $(\lambda_U^a)_{ij} \approx \epsilon_{q_i}\epsilon_{u_j}$ and to down quarks $(\lambda_D^a)_{ij} \approx \epsilon_{q_i}\epsilon_{dj}$. The nine parameters $\{\epsilon_{q_i}, \epsilon_{u_i}, \epsilon_{d_i}\}$ can be estimated from the six quark masses and the three Euler angles of the Kobayashi-Maskawa matrix. The flavor-changing problem of these multi-Higgs models is solved by such a choice of AFS, if the masses of the additional scalars are several hundred GeV. This simple Abelian symmetry is insufficient to solve the supersymmetric flavor-changing problem. It provides for no approximate degeneracy between $\tilde{d}$ and $\tilde{s}$ squarks, and allows Cabibbo-sized mixing between them, which, as shown in the next section, leads to a disastrously large contribution to ΔM_K.

III.3 The Flavor-Changing Constraints

A brief, somewhat heuristic, view of the general conditions required to solve the supersymmetric flavor-changing problem will be given in this section. The results will allow us to understand whether AFS's are likely to be of use in solving this problem. My aim is to provide a set of sufficient conditions which I find to be both simple and useful; I do not attempt to determine the necessary conditions.

Consider the case when $\boldsymbol{\xi}_{U,D,E} = O$. Unitary transformations are performed on the fermion fields to diagonalize $\boldsymbol{\lambda}_{U,D,E}$ and on the scalar fields to diagonalize $m_a^2, a = q, u, d, \ell, e$. In this *mass basis*, there will be unitary mixing matrices at the gaugino vertices, which, for the neutral gauginos, we write as W^α where $\alpha = u_L, u_R, d_L, d_R, e_L, e_R$. Flavor and CP-violating effects are induced by Feynman diagrams involving internal gauginos and scalar superpartners. These are box diagrams for $\Delta m_K, \epsilon_K, \Delta m_B$ ... and penguin-type diagrams for $\mu \to e\gamma, d_e, b \to$

$s\gamma$ .... The exchange of a scalar of generation k between external fermions (of given α) of generations i and j leads to a factor in the amplitude of

$$X_{ij}^{\alpha} = m_s^2 \sum_k W_{ki}^{\alpha} W_{kj}^{\alpha *} P_k^{\alpha} \ , \qquad (III.5)$$

where P_k^{α} is the propagator for the scalar of mass m_k^{α}. X^{α} is made dimensionless by inserting a factor m_s^2, where m_s describes the scale of supersymmetry breaking. Studies of flavor and CP-violating processes allows bounds to be placed on the magnitudes and imaginary parts of X_{ij}^{α} of the form

$$X_{ij}^{\alpha} \lesssim X_{oij}^{\alpha} \left(\frac{m_s}{m_{so}} \right)^P , \qquad (III.6)$$

where the bound is X_0 when m_s is taken to be the reference value m_{so}. The quantity p is a positive integer, so that the bounds become weaker for higher m_s. For box diagram contributions, $p = 1$, while for penguin-like diagrams, $p = 2$. Useful results for these bounds are tabulated in Refs. 13–15, as are references to earlier literature. For our purposes, we extract the following results:

If W^{α} are "KM-like," that is, if

$$|W_{ij}^{\alpha}| \lesssim |V_{ij}| (i \neq j), \qquad (III.7)$$

where V is the Kobayashi-Maskawa matrix, important limits only result for processes where the external fermions are of the first two generations (i.e., neither i nor j is three).

The most important flavor-changing limits arise when $(i, j) = (1, 2)$. For example, taking the relevant phases to be of order unity, ϵ_K implies

$$|X_{12}^{\alpha}| = m_s^2 |W_{21}^{\alpha} W_{22}^{\alpha *} (P_2^{\alpha} - P_1^{\alpha}) + W_{31}^{\alpha} W_{32}^{\alpha *} (P_3^{\alpha} - P_1^{\alpha})| \lesssim 10^{-4}. \qquad (III.8)$$

Here and below, I take $m_s = 1$ TeV. For W^{α} KM-like, $|W_{31}^{\alpha} W_{32}^{\alpha *}| \lesssim |V_{td}| \, |V_{ts}| \approx 4 \times 10^{-4}$, so there is no constraint from the last term of Eq. (8) even if there is large nondegeneracy between the scalars of the first and third generation. It is the first term which is typically the origin of the supersymmetric flavor-changing problem. This first term I call the "1–2" problem; while the second term I call the "1, 2–3" signature, because if the W^{α} are CKM-like this contribution is close to the experimental value. One way to solve the problem is to make W_{21}^{α} small

$$|W_{21}^{\alpha}| \lesssim |V_{td}| \, |V_{ts}| \, . \qquad (III.9a)$$

Another is to make the scalars $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ degenerate:

$$|D^{\alpha}_{21}| \lesssim \frac{|V_{td}|\,|V_{ts}|}{|V_{us}|}, \qquad (III.9b)$$

where $D^{\alpha}_{ij} = (m^{\alpha^2}_i - m^{\alpha^2}_j)/m^{\alpha^2}_i$, and in the limit of near degeneracy $D^{\alpha}_{12} \approx m^2_s(P^{\alpha}_2 - P^{\alpha}_1)$. In fact, the condition (8) and $(9a)$ or $(9b)$ need only be applied for $\alpha = d_L, d_R, e_L$, and e_R. The limits to flavor-changing processes in the up sector are much weaker and are not problematic. Of course, the flavor problem can also be solved by having smaller suppressions of both W^{α}_{21} and D^{α}_{21}. Nevertheless, I find it useful to keep in mind that, for $\boldsymbol{\xi}_{U,D,E} = 0$, the flavor problem is solved if:

I. All W^{α} are KM-like.

II. Either Eq. (III.9a) or Eq. (III.9b) holds in the d and e sectors.

Since the X^{α}_{12} quantities are small, it is often convenient to work in the *gaugino basis*. In this basis, superfield unitary transformations are performed to diagonalize $\boldsymbol{\lambda}_{U,D,E}$ so that the neutral gaugino vertices are flavor conserving. The scalar mass matrices now have off-diagonal entries which, assuming they are small, can be treated in perturbation theory as flavor-violating interactions. In this basis, Eq. (III.8) and Eq. (III.9a) or Eq. (III.9b) are replaced by

$$\left|\frac{m^{\alpha^2}_{12}}{m^2_s}\right| \lesssim 4 \times 10^{-4}. \qquad (III.9c)$$

Until now, we have avoided discussing the flavor matrices $\boldsymbol{\xi}_{U,D,E}$ of Eq. (III.1). Inserting the Higgs vev induces mass mixing between left and right scalars; hence 6×6 rotations are required to reach the mass basis. It is easier to use the gaugino basis and treat these masses in perturbation theory, writing them as:

$$\boldsymbol{\xi}_{U,D,E} = W'^{u_L,d_L,e_L}\overline{\boldsymbol{\xi}}_{U,D,E}W'^{u_R,d_R,e_R}, \qquad (III.10)$$

where $\overline{\boldsymbol{\xi}}_{U,D,E}$ are diagonal matrices. Experiments place many limits on the elements $\overline{\xi}_{U,D,Eii}$. For our purposes, it is useful to know that all these limits are satisfied if:

III. All W'^{α} are KM-like.

IV. $\overline{\xi}_{U,D,Eii}$ are of order $m_s\overline{\lambda}_{U,D,Eii}$.

The basic reason for this is that the only large contributions to flavor-changing processes involving the first two generations then come from terms of order $|W'^{\alpha_L}_{31}W'^{\alpha_R}_{32}|\lambda_{b,t}$ which are $\lesssim |V_{td}V_{ts}|$.

Now that we have argued that the four statements I–IV are sufficient to solve the supersymmetric flavor problem, we can ask whether it is reasonable to expect

that AFS will be of use. It should be apparent that the general expectation is that any AFS which leads to the hierarchy of fermion masses, as parameterized by $\overline{\lambda}_{U,D,E_i}$, and to the KM pattern of flavor violation, described by V_{ij}, will automatically lead to I, III, and IV being satisfied. The only remaining question is whether AFS can satisfy II, i.e., whether they can produce either Eq. (III.9a) or Eq. (III.9b) [or Eq. (III.9c) in the insertion approximation]. The Abelian G_f discussed earlier ($U(1)^9$ in the quark sector) is clearly insufficient since it gives $D^{\alpha}_{21} \approx 1$ and $W^{\alpha}_{21} \approx V_{us}$. In the next section, I show that the maximal AFS is easily sufficient.

III.4 The Maximal Approximate Flavor Symmetry

We assume that below some high scale, Λ, physics is described by a softly broken, supersymmetric $SU(3) \times SU(2) \times U(1)$ gauge theory with minimal field content. The flavor interactions are those of the superpotential and soft supersymmetry breaking interactions shown in Eqs. (III.1) and (III.2). We assume that the dynamics above Λ, which may be strong, possesses an approximate flavor symmetry G_f. Below Λ, the breaking of this AFS is characterized by a set of parameters $\{\epsilon(R)\}$ transforming as R under G_f. In this section, we take G_f to be $G_{max} = U(3)^5$, the maximal AFS which commutes with the Standard Model gauge group. Although strong dynamics could preserve a larger AFS, the breaking parameters $\{\epsilon(R)\}$ cannot violate $SU(3) \times SU(2) \times U(1)$, so that G_{max} is the largest group under which the set $\{\epsilon\}$ form complete representations. Each factor of G_{max} is labelled as $U(3)_a$ where $a = q, u, d, l$, or e. We assume that the $\{\epsilon\}$ fill out three irreducible representations: $\boldsymbol{\epsilon}_U \sim (3_q, \overline{3}_u), \boldsymbol{\epsilon}_D \sim (3_q, \overline{3}_d)$, and $\boldsymbol{\epsilon}_L \sim (3_l, \overline{3}_e)$. In the case of QCD with approximate $SU(3)_L \times SU(3)_R$ broken explicitly by the quark mass matrix M, there is no loss of generality in choosing a basis for the quark fields in which M is real and diagonal. Similarly, we may choose a basis for the lepton fields in which $\boldsymbol{\epsilon}_E$ is real and diagonal $\overline{\boldsymbol{\epsilon}}_E$. We may choose the quark basis so that $\boldsymbol{\epsilon}_U = \overline{\boldsymbol{\epsilon}}_U$ is diagonal and $\boldsymbol{\epsilon}_D = V^* \overline{\boldsymbol{\epsilon}}_D$, where $\overline{\boldsymbol{\epsilon}}_D$ is diagonal and V is a unitary matrix. All flavor-changing effects of this theory are described by a single matrix, which to high accuracy is the KM matrix. Criteria I and III of the previous section are satisfied. This theory has no violation of the lepton numbers.

To zeroth order in $\{\epsilon\}$, the only interactions of the quarks and leptons are the gauge interactions and the zeroth order supersymmetry breaking potential

$$V_0 = q m_q^2 1 q^\dagger + u^\dagger m_u^2 1 u + d^\dagger m_d^2 1 d + \ell m_\ell^2 1 \ell^\dagger + e^\dagger m_e^2 1 e. \qquad (III.11)$$

We see that the non-Abelian nature of G_f enforces squark and slepton degeneracy at zeroth order in ϵ. However, Eq. (III.11) differs from the universal boundary condition of supergravity because the five parameters m_a^2 are all independent and are not constrained to be equal. Similarly, they can differ from the Higgs mass parameters. Equation (III.9b), and therefore criterion II, is satisfied at zeroth order, but corrections appear at higher order.

At first order in ϵ, superpotential interactions are generated:

$$W_1 = a_1\ q\boldsymbol{\epsilon}_U u h_2 + a_2\ q\boldsymbol{\epsilon}_D d h_1 + a_3\ \ell\boldsymbol{\epsilon}_L e h_1, \qquad (III.12)$$

where $a_{1,2,3}$ are "strong interaction" parameters of order unity. The $U(3)$ transformations are shown explicitly in Appendix A at the end of this chapter.

The assumed transformation properties of the $\{\epsilon\}$ are sufficient to guarantee that W preserves R parity invariance to all orders in ϵ. There is no need to impose R parity as a separate exact symmetry. The Yukawa couplings can be written as expansions in ϵ, for example, $\lambda_U = a_1\epsilon_U + a_4\epsilon_U\epsilon_U^\dagger\epsilon_U + a_5\epsilon_D\epsilon_D^\dagger\epsilon_U + ...$ If we work only to second order, we can simply take $\lambda_U = a_1\epsilon_U$, etc. Even if we work to higher order, we can rearrange the perturbation series as an expansion in $\boldsymbol{\lambda}_{U,D,E}$ rather than $\boldsymbol{\epsilon}_{U,D,E}$. Either way, to second order in the expansion:

$$W_1 = q\boldsymbol{\lambda}_U u h_2 + q\boldsymbol{\lambda}_D d h_1 + \ell\boldsymbol{\lambda}_E e h_1 \qquad (III.13a)$$

$$W_2 = \frac{a_1}{\Lambda^2}(q\boldsymbol{\lambda}_U u)(q\boldsymbol{\lambda}_D d) + ... \qquad (III.13b)$$

$$V_1 = m_s(a_U q\boldsymbol{\lambda}_U u h_2 + a_D q\boldsymbol{\lambda}_D d h_1 + a_E \ell\boldsymbol{\lambda}_E e h_1) \qquad (III.13c)$$

$$\begin{aligned} V_2 \ &= m_s^2\Big(q(a_2\boldsymbol{\lambda}_U\boldsymbol{\lambda}_U^\dagger + a_3\boldsymbol{\lambda}_D\boldsymbol{\lambda}_D^\dagger)q^\dagger + a_4 d^\dagger\boldsymbol{\lambda}_D^\dagger\boldsymbol{\lambda}_D d + a_5 u^\dagger\boldsymbol{\lambda}_U^\dagger\boldsymbol{\lambda}_U u + a_6\ell\boldsymbol{\lambda}_E\boldsymbol{\lambda}_E^\dagger\ell^\dagger \\ &\quad + a_7 e^\dagger\boldsymbol{\lambda}_E^\dagger\boldsymbol{\lambda}_E e) + \frac{m_s^2}{\Lambda^2}a_8(q\boldsymbol{\lambda}_U\boldsymbol{\lambda}_U^\dagger q^\dagger)(u^\dagger u) + \frac{m_s^2}{\Lambda^2}a_9(q\boldsymbol{\lambda}_U u)(q\boldsymbol{\lambda}_D d). \qquad (III.13d) \end{aligned}$$

Given the nonrenormalization theorems, one might question whether the interactions in W really are generated. In general, the answer is yes: they are generated by integrating out heavy particles at tree level and by radiative corrections to D-terms followed by field rescalings. However, in specific simple models, one discovers that the structure of the supersymmetric theory is such that not all interactions allowed by the symmetries of the low-energy theory are generated. Hence, if the symmetry structure of the low-energy theory is insufficient to solve the flavor-changing problem, it may still be that a theory above Λ with this symmetry can be constructed which does not generate the troublesome interactions.

In QCD, the strong interaction parameters are real—the strong dynamics of QCD preserves CP. Also, the strong dynamics is well-separated from the origin of the explicit breaking parameters $\boldsymbol{\epsilon} = M, Q$. The "strong" dynamics of the supersymmetric theory above Λ may conserve CP so that $a_1 ... a_9$ are real. This would explain the smallness of the neutron electric dipole moment which has contributions from $Im(a_u)$ and $Im(a_d)$ (Ref. 16). However, it may be that the dynamics above Λ which generates these coefficients is not very separate from that which generates the $\{\epsilon\}$. Since the KM phase comes from $\{\epsilon\}$, in this case there would also be phases in $\{a\}$.

Does the boundary condition of Eq. (III.11) and Eq. (III.13) at scale Λ solve the flavor-changing problem? In the lepton sector, the answer is obviously yes: $\boldsymbol{\lambda}_E$ can be made real and diagonal so there is no lepton-flavor violation.

In the quark sector, the only mixing matrix is the KM matrix, so that criteria I and III are satisfied. In fact, the only unitary transformations needed to reach the mass basis are a rotation of V on d_L quarks, and a rotation of q squarks. This latter rotation is awkward; it is more convenient to make the V rotation on d_L to be a superfield rotation, and to treat the remaining scalar mass flavor violation as a perturbation:

$$\frac{\delta m^{d_L^2}_{21}}{m_s^2} = a_2 (V^T \overline{\boldsymbol{\lambda}}_U^2 V^*)_{21} \approx a_2 \left|V_{ts} V_{td}\right|^* \lambda_t^2 \approx 4 \times 10^{-4}. \qquad (III.14)$$

We can see that the condition of Eq. (III.9c), and therefore criterion II, is satisfied. Finally, the trilinear scalar interactions of V_1 in Eq. (III.13c) clearly satisfy the criterion IV. The matrices $W'^{\alpha} = I + O(\boldsymbol{\epsilon}^2)$ so that criterion III is also satisfied.

The flavor structure of this theory with $G_f = G_{max} = U(3)^5$ is very similar to that which results from the universal boundary conditions of supergravity discussed below. In that theory, the terms $a_2 ... a_9$ are assumed to be absent at the boundary, but are generated via renormalization groups scalings from $\Lambda = M_{Pl}$ to m_s, and end up being of order unity. What features of this flavor sector are crucial to solving the flavor-changing problem?

(i) At zeroth order in ϵ, the scalars of each A are degenerate and the soft operators have no flavor violation.

(ii) At linear order in ϵ, the superfield rotations which diagonalize the quark masses also diagonalize the soft scalar trilinear couplings. Hence, at this order, the soft operators contain no flavor-changing neutral currents.

(iii) The corrections to m_a^2, induced at second order in ϵ, induce flavor-changing effects proportional to $\boldsymbol{\lambda}_U \boldsymbol{\lambda}_U^\dagger$ and $\boldsymbol{\lambda}_D \boldsymbol{\lambda}_D^\dagger$. If we restrict $\boldsymbol{\lambda}_U$ and $\boldsymbol{\lambda}_D$ to their light 2×2 subspaces, then all contributions are less than 10^{-4}. Hence, we need only consider contributions involving the heavy generation. For external light quarks, this gives small contributions because V_{ts} and V_{td} are small.

We finish this section by briefly comparing the AFS method to several well-known solutions of the supersymmetric flavor-changing problem. The low-energy structure of these theories can be understood as examples of the AFS technique.

The most popular treatment of the supersymmetric flavor-changing problem is to assume that at some high scale, usually taken to be the reduced Planck mass, the flavor matrices possess a "universal" form[17,18]:

$$m_a^2 = m_0^2 I \qquad (III.15a)$$

$$\boldsymbol{\xi}_{U,D,E} = A \ \boldsymbol{\lambda}_{U,D,E} \quad , \qquad (III.15b)$$

which generalizes the idea of squark degeneracy.[8] This form is the most general which results from hidden sector supergravity theories, provided the Kähler potential is $U(N)$ invariant, where N is the total number of chiral superfields.[18] However, imposing this $U(N)$ invariance as an exact symmetry on one piece of the Lagrangian is ad hoc because it is broken explicitly by the gauge and superpotential interactions.

We advocate replacing this $U(N)$ idea with an approximate flavor symmetry G_f acting on the entire theory, broken explicitly by a set of parameters $\{\epsilon(R)\}$, allowing the Lagrangian to be written as a power series in ϵ: $\mathcal{L}_0 + \mathcal{L}_1 + ...$ At each order, the most general set of interactions is written which is consistent with the assumed transformation properties of $\{\epsilon(R)\}$. Taking $G = U(3)^5$, we have found that a modified universal boundary condition emerges. At zeroth order in ϵ, we found Eq. (III.15a) to be replaced by

$$m_a^2 = m_a^2 I, \qquad (III.16a)$$

and at first order in ϵ, Eq. (III.15b) is replaced by

$$\boldsymbol{\xi}_{U,D,E} = A_{U,D,E} \boldsymbol{\lambda}_{U,D,E}. \qquad (III.16b)$$

These boundary conditions are corrected at higher orders by factors of $(1 + O(\epsilon^2))$ but are sufficient to solve the supersymmetric flavor-changing problem. While Eq. (III.15) was invented as the most economical solution to the flavor-changing

problem, the symmetry structure of the theory demonstrates that it is ad hoc, and from the phenomenological viewpoint, it is overkill. The flavor structure of the low-energy theory provides a motivation for Eq. (III.16), together with the $1 + O(\epsilon^2)$ correction factors. Phenomenological results, which follow from assuming the boundary condition (15) but do not result from Eq. (III.16), should be considered suspect. For example, the flavor-changing problem provides no motivation for the belief that the squarks of the lightest generation ($\tilde{q}_L, \tilde{d}_R,$ and $\tilde{u}_R$) are degenerate (up to electroweak renormalizations and breaking). Similarly, the flavor-changing problem provides no motivation for a boundary condition where $m^2_{h_1}$ and $m^2_{h_2}$ are both set equal to squark and slepton masses.

Perhaps the most straightforward idea to solve the flavor-changing problem is to assume that supersymmetry breaking is transferred to the observable sector by the known gauge interactions.[4] Suppose this happens at scale Λ, and that below Λ the observable sector is the minimal field content supersymmetric $SU(3)\times SU(2)\times U(1)$ theory. At scale Λ, the dominant soft supersymmetry breaking operators are the three gaugino mass terms, which are generated by gauge mediation at the one-loop level. At higher loop levels, at scale Λ, the eight flavor matrices m^2_a and $\boldsymbol{\xi}_{U,D,E}$ are generated. However, since the only violation of the $U(3)^5$ flavor symmetry is provided by $\boldsymbol{\lambda}_{U,D,E}$, *the most general theory of this sort is described at scale Λ by Eqs. (III.11) and (III.13), and hence, possesses the boundary condition (III.16).* The parameters $\{a\}$ are now each given by a power series in the Standard Model gauge couplings, α_i, with coefficients which depend on the representation structure of the supersymmetry breaking sector. The gaugino masses M_i are very large, and at low energy, the parameters m^2_A of Eq. (III.11) receive contributions $\propto \sum_i C_{iA}\alpha_i M_i^2 \ln \Lambda/m_s$, where C_{iA} involve quantum numbers. This may dominate m^2_a boosting the importance of V_0, and thereby decreasing the flavor-violating effects induced by $V_{1,2}$.

The AFS technique is sufficiently general that it can be used no matter how supersymmetry is broken and transmitted to the observable sector. This almost guarantees that it will be a useful tool in studying the flavor questions of supersymmetry. It may be that nature chooses a more complicated G_f and ϵ than the above example. At scale Λ, the observable sector may involve additional fields, and there may be additional flavor-breaking matrices. Simple group theory can be used to determine the additional terms which these induce in V_1 and V_2, allowing an easy estimation of potential flavor-changing difficulties.

In the previous section, we argued that approximate flavor symmetries which lead to the observed hierarchy of quark and lepton masses and mixings are very likely to give supersymmetric theories where all mixing matrices are KM like, and the eigenvalues of $\boldsymbol{\xi}_{U,D,E}$ possess a hierarchy similar to the eigenvalues of $\boldsymbol{\lambda}_{U,D,E}$. Hence, the criteria I, III, and IV are easily satisfied, and the real flavor problem is that either Eqs. (III.9a) or (III.9b) must be imposed. This means that either the mixing in the first two generations, W_{21}^{α}, is much smaller than expected from the Cabibbo angle, or the squarks of the first two generations must be highly degenerate. This degeneracy can be understood as the consequence of a non-Abelian symmetry, continuous or discrete, which acts on the first two generations. The low-energy limit of any such theories can be analyzed using AFS. An alternative possibility is to seek Abelian symmetries, allowing squark nondegeneracies, which lead to the suppression of $W_{21}^{\alpha}A$.

It is well-known that the experimental constraints on FCNC imply that W_{21}^{α} need be suppressed only in the d and e sectors ($\alpha = d_L, d_R, e_L, e_R$): $W_{21}^{u_L} \approx W_{21}^{u_R} \approx V_{us}$ leads to interesting $D^0\overline{D}^0$ mixing but is not a problem. This opens the possibility that symmetries can be arranged so that Cabibbo mixing originates in the u sector, while mixing of the generations is highly suppressed in the d and e sectors. This idea has been used to construct models with Abelian flavor symmetries and nondegenerate squarks.[21]

III.5 A Brief Introduction to the Literature

In supersymmetric models of particle physics, there are two aspects to the flavor problem. The first is the problem of quark and lepton mass and mixing hierarchies: why are there a set of small dimensionless Yukawa couplings in the theory? The second aspect of the problem is why the superpartner gauge interactions do not violate flavor at too large a rate. This requires that the squark and slepton mass matrices not be arbitrary; rather, even though all eigenvalues are large, these matrices must also possess a set of small parameters which suppresses flavor-changing effects. What is the origin of this second set of small dimensionless parameters?

An extremely attractive hypothesis is to assume that the two sets of small parameters, those in the fermion mass matrices and those in the scalar mass matrices, have a common origin: they are the small symmetry breaking parameters of an approximate flavor symmetry group G_f. This provides a link between the fermion mass and flavor-changing problems; both are addressed by the same symmetry. Such an approach was first advocated using a flavor group $U(3)^5$, broken

only by the three Yukawa matrices $\lambda_{U,D,E}$ in the up, down, and lepton sectors,[19] as discussed in the previous section. This not only solved the flavor-changing problem, but suggested a boundary condition on the soft operators which has a more secure theoretical foundation than that of universality. However, this framework did not provide a model for the origin of the Yukawa matrices themselves and left open the possibility that G_f was more economical than the maximal flavor group allowed by the Standard Model gauge interactions.

The first explicit models in which spontaneously broken flavor groups were used to constrain both fermion and scalar mass matrices were based on $G_f = SU(2)$ (Ref. 20) and $G_f = U(1)^3$ (Ref. 21). In the first case, the approximate degeneracy of scalars of the first two generations was guaranteed by $SU(2)$. In retrospect, it seems astonishing that the flavor-changing problem of supersymmetry was not solved by such a flavor group earlier. The well-known supersymmetric contributions to the $K_L - K_S$ mass difference can be rendered harmless by making the $\tilde{d}$ and $\tilde{s}$ squarks degenerate. Why not guarantee this degeneracy by placing these squarks in a doublet of a non-Abelian flavor group $(\tilde{d}, \tilde{s})$? Perhaps one reason is that $SU(2)$ allows large degenerate masses for d and s quarks. In the case of Abelian G_f, the squarks are far from degenerate; however, it was discovered that the flavor-changing problem could be solved by arranging for the Kobayashi-Maskawa mixing matrix to have an origin in the up sector rather than the down sector.

A variety of supersymmetric theories of flavor have followed, including ones based on $G_f = 0(2)$ (Ref. 22), $G_f = U(1)^3$ (Ref. 23), $G_f = \Delta(75)$ (Ref. 24), $G_f = (S_3)^3$ (Refs. 25-27), and $G_f = U(2)$ (Refs. 28, 29). Progress has also been made on relating the small parameters of fermion and scalar mass matrices using a gauged $U(1)$ flavor symmetry in a $N = 1$ supergravity theory, taken as the low-energy limit of superstring models.[30] Development of these and other theories of flavor is of great interest because they offer the hope that an understanding of the quark and lepton masses, and the masses of their scalar superpartners, may be obtained at scales well beneath the Planck scale, using simple arguments about fundamental symmetries and how they are broken. These theories, to varying degrees, provide an understanding of the patterns of the mass matrices, and may, in certain cases, also lead to very definite mass predictions. Furthermore, flavor symmetries may be of use to understand a variety of other important aspects of the theory.

The general class of theories which address both aspects of the supersymmetric

flavor problem have two crucial ingredients: the flavor group, G_f, and the flavor fields, ϕ, which have a hierarchical set of vacuum expectation values allowing a sequential breaking of G_f. These theories can be specified in two very different forms. In the first form, the only fields in the theory beyond ϕ are the light matter and Higgs fields. An effective theory is constructed in which all gauge and G_f invariant interactions are written down, including nonrenormalizable operators scaled by some mass scale of flavor physics, M_f. An example of such a theory, with $G_f = U(3)^5$, was discussed in Sec. III.4. The power of this approach is that considerable progress is apparently possible without having to make detailed assumptions about the physics at scale M_f which generates the nonrenormalizable operators. Much, if not all, of the flavor structure of fermion and scalar masses comes from such nonrenormalizable interactions, and it is interesting to study how their form depends only on G_f, G_f breaking, and the light field content.

A second, more ambitious approach is to write a complete, renormalizable theory of flavor at the scale M_f. Such a theory possesses a set of heavy fields which, when integrated out of the theory, lead to the effective theory discussed above.[31] However, it is reasonable to question whether the effort required to construct such full theories is warranted. Clearly, these complete theories involve further assumptions beyond those of the effective theories, namely the G_f properties of the fields of mass M_f, and it would seem that the low-energy physics of flavor is independent of this, depending only on the properties of the effective theory. In nonsupersymmetric theories, such a criticism may have some validity, but in supersymmetric theories it does not. This is because in supersymmetric theories, on integrating out the states of mass M_f, the low-energy theory may not be the most general effective theory based on flavor group G_f. Several operators which are G_f invariant, and could be present in the effective theory, are typically not generated when the heavy states of mass M_f are integrated out. Which operators are missing depends on what the complete theory at G_f looks like. This phenomena is well-known and is illustrated, for example, in Refs. 24, 29, and 32, and it casts doubt on the effective theory approach to building supersymmetric theories of flavor. Finally, one might hope that a complete renormalizable theory of flavor at scale M_f might possess a simplicity which is partly hidden at the level of the effective theory.

III.6 The Minimal U(2) Theory of Flavor

The largest flavor group which acts identically on each component of a generation, and is therefore consistent with grand unification, is $U(3)$, with the three generations forming a triplet. This is clearly strongly broken to $U(2)$ by whatever generates the Yukawa coupling for the top quark. Hence, the largest such flavor group which can be used to understand the small parameters of the fermion and scalar mass matrices is $U(2)$. In this section, I briefly mention aspects of the $U(2)$ theory constructed in Ref. 29.

While the third generation is a trivial $U(2)$ singlet, ψ_3, the two light generations are doublets, ψ_a:

$$q_a = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \quad u_a = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad d_a = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \quad \ell_a = \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} \quad e_a = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}. \qquad (III.17)$$

In the symmetry limit, only the fermions of the third generation have mass, while the scalars of the first two generations are degenerate: clearly a promising zeroth order structure.

The dominant breaking of $U(2)$ is assumed to occur via the vev of a doublet: $\langle \phi^a \rangle$. If we study the most general theory beneath some flavor scale M_f, then the nonrenormalizable operators for fermion masses are:

$$\frac{1}{M_f} \, [\psi_3 \phi^a \psi_a h]_F, \qquad (III.18)$$

which generates V_{cb}, and

$$\frac{1}{M_f^2} \, [\psi_a \phi^a \phi^b \psi_b \, h]_F, \qquad (III.19)$$

which generates a 22 entry in the Yukawa matrices. An immediate difficulty is that $U(2)$ also allows the supersymmetry breaking scalar mass

$$\frac{1}{M_f^2} \, [\psi^{\dagger a} \phi^\dagger_a \phi^b \psi_b \, z^\dagger z]_D, \qquad (III.20)$$

where z is a supersymmetry breaking spurion, taken dimensionless, $z = m\theta^2$, which leads to a splitting of the degeneracy of the scalar masses of the first two generations:

$$\frac{m^2_{\tilde{e}} - m^2_{\tilde{\mu}}}{m^2_{\tilde{e}} + m^2_{\tilde{\mu}}} \approx O\left(\tfrac{m_\mu}{m_\tau}\right) \qquad (III.21)$$

in the lepton sector, and

$$\frac{m^2_{\tilde{d}} - m^2_{\tilde{s}}}{m^2_{\tilde{d}} + m^2_{\tilde{s}}} \approx O\left(\tfrac{m_s}{m_b}\right) \qquad (III.22)$$

in the down quark sector. These lead to violations of the flavor-changing constraints of Sec. III.3 (Ref. 28). However, if these operators are generated by Froggatt-Nielsen type theories,[31] one discovers that Eq. III.21 and Eq. III.22 are not generated if the exchanged heavy vector generations transform as $U(2)$ doublets.

If the final breaking of $U(2)$ occurs via a two-indexed antisymmetric tensor, $\langle A_{ab} \rangle$, then the final operator contributing to fermion masses is

$$\frac{1}{M_f} \left[\psi_a A^{ab} \psi_b h \right]_F . \qquad (III.23)$$

It is remarkable that theories of flavor can be based on the two interactions of Eq. III.18 and Eq. III.23, in addition to the third-generation coupling $[\psi_3 \psi_3 h]_F$. The Yukawa matrices take the form

$$\boldsymbol{\lambda} = \begin{pmatrix} 0 & \epsilon' & 0 \\ -\epsilon' & 0 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} , \qquad (III.24)$$

where $\epsilon = \langle \phi^2 \rangle / M_f$ and $\epsilon' = \langle A^{12} \rangle / M_f$, and the scalar mass matrices are

$$m^2 = \begin{pmatrix} m_1^2 + \epsilon'^2 m^2 & 0 & \epsilon\epsilon' m^2 \\ 0 & m_1^2 + \epsilon'^2 m^2 & 0 \\ \epsilon\epsilon' m^2 & 0 & m_3^2 + \epsilon^2 m^2 \end{pmatrix} . \qquad (III.25)$$

The splitting between the masses of the scalars of the lightest two generations is

$$\frac{m_{\tilde{e}}^2 - m_{\tilde{\mu}}^2}{m_{\tilde{e}}^2 + m_{\tilde{\mu}}^2} \approx O\left(\tfrac{m_e m_\mu^2}{m_\tau^3} \right) \qquad (III.26)$$

in the lepton sector, with similar equations in the quark sector. The "1-2" aspect of the supersymmetric flavor-changing problem is completely solved. However, because λ_{22} vanishes, the mixings to the third generation are larger than those of the CKM matrix, so that the conditions of Sec. III.3 are not immediately satisfied. The splittings between the third-generation scalar mass and the lightest two generations should not be of order unity, or the contribution to ϵ_K from the "1, 2-3" effects in this model will be too large. This splitting cannot be computed within a $U(2)$ theory but will be an important constraint on $U(3)$ theories.

This $U(2)$ theory of flavor has a significant economy of parameters. Two of the Standard Model flavor parameters are predicted:

$$\left| \frac{V_{td}}{V_{ts}} \right| = s_1 = \sqrt{\frac{m_d}{m_s}} = 0.230 \pm 0.008 \qquad (III.27a)$$

$$\left|\frac{V_{ub}}{V_{cb}}\right| = s_2 = \sqrt{\frac{m_u}{m_c}} = 0.063 \pm 0.009. \qquad (III.27b)$$

As measurements of these quantities improve, it will be interesting to see whether they remain within the uncertainties of the above predictions. There are six unitary 3×3 flavor-mixing matrices at neutralino vertices; in the $U(2)$ theory, they are real and given by six angles s_{Iij} and s^c_{Iij} where $I = U, D, E$ labels the up, down, and lepton sectors, and $ij = 12, 23, 31$ labels the generations being mixed. These angles are predicted in terms of just three free parameters r_I

$$s_{I12} = -s^c_{I12} = \left(\sqrt{\frac{m_1}{m_2}}\right)_I \qquad (III.28a)$$

$$s_{I23} = \left(\sqrt{r\frac{m_2}{m_3}}\right)_I \qquad (III.28b)$$

$$s^c_{I23} = \left(\sqrt{\frac{1}{r}\frac{m_2}{m_3}}\right)_I, \qquad (III.28c)$$

where $(m_{1,2,3})_I$ are the fermion mass eigenvalues of generations (1,2,3), renormalized at the flavor scale M_f.

Further aspects of this $U(2)$ theory of flavor can be found in Ref. 29, on which this section was based.

III.7 The Suppression of Baryon and Lepton Number Violation

The Standard Model, for all its shortcomings, does provide an understanding for the absence of baryon and lepton number violation: the field content simply does not allow any renormalizable interactions which violate these symmetries. This is no longer true when the field content is extended to become supersymmetric; squark and slepton exchange mediate baryon and lepton number violation at unacceptable rates, unless an extra symmetry, such as R parity, is imposed on the theory. It is worth stressing that some new symmetry, which in general we label by X, really is required: the known gauge and spacetime symmetries are insufficient. The need for X was first realized in the context of a supersymmetric $SU(5)$ grand unified theory.[33] As will become clear, there are a wide variety of possibilities for the X symmetry. Matter parity,[8] Z_N symmetries other than matter parity,[34,35] and baryon or lepton numbers[36] provide well-known examples; each giving a distinctive phenomenology. One of the most fundamental questions in constructing supersymmetric models is Ref. 37. *What is the origin of this extra symmetry needed to suppress baryon and lepton number-violating processes?*

The X symmetry must have its origin in one of the three categories of symmetries which occur in field theory models of particle physics: spacetime symmetries, gauge (or vertical) symmetries, and flavor (or horizontal) symmetries. The X symmetry is most frequently referred to as R parity,[§] R_p, which is a Z_2 parity acting on the anticommuting coordinate of superspace: $\theta \to -\theta$. We view this as unfortunate, since it suggests that the reason for the suppression of baryon and lepton number violation is to be found in spacetime symmetries, which certainly need not be the case. R_p can be viewed as a superspace analogue of the familiar discrete spacetime symmetries, such as P and CP. In the case of P and CP, we know that they can appear as accidental symmetries in gauge models which are sufficiently simple. For example, P is an accidental symmetry of QED and QCD, while CP is an accidental symmetry of the two-generation Standard Model. Nevertheless, in the real world P and CP are broken. This suggests to us that discrete spacetime symmetries are not fundamental and should not be imposed on a theory, so that if R_p is a good symmetry, it should be understood as being an accidental symmetry resulting from some other symmetry. These arguments can also be applied to alternative spacetime origins for X, such as a Z_4 symmetry on the coordinate θ (Ref. 34).[¶] Hence, while the symmetry X could have a spacetime origin, we find it more plausible that it arises from gauge or flavor symmetries.

In this case, what should we make of R_p? If it is a symmetry at all, it would be an accidental symmetry, either exact or approximate. If R_p is broken by operators of dimension 3, 4, or 5, then a weak-scale, lightest superpartner (LSP) would not be the astrophysical dark matter. The form of the R_p breaking interactions will determine whether the LSP will decay in particle detectors or whether it will escape leaving a missing energy signature. The realization that X may well have an origin in gauge or flavor symmetries has decoupled the two issues of the suppression of B and L violation, due to X, and the lifetime of the LSP, governed by R_p.[35,39]

At first sight, the most appealing origin for X is an extension of the Standard Model gauge group, either at the weak scale[37] or at the grand unified scale.[40] An interesting example is provided by the crucial observation that adding $U(1)_{B-L}$ (Ref. 40), or equivalently $U(1)_{T_{3R}}$, is sufficient to remove all renormalizable B and

[§] R_p was first introduced in a completely different context.[38]

[¶] Clearly, these arguments need not be correct: for example, it could be that both P and CP are fundamental symmetries, but they have both been spontaneously broken. However, in this case the analogy would suggest that R_p is also likely to be spontaneously broken.

L violation from the low-energy theory. Matter parity is a discrete subgroup of $U(1)_{B-L} \times U(1)_{T_{3R}}$. This is clearly seen in SO(10) (Ref. 41), where the requirement that all interactions have an even number of spinor representations immediately leads to matter parity, generated by the Z_2 element

$$X(SO(10)) = e^{i\pi(2T_{3_L}+2T_{3_R})} = e^{i\pi(N_{16}+N_{144}...)}, \qquad (III.29)$$

where $N_{16,144,...}$ is 1 for a 16, 144, ... representation.

However, this example has a gauge group with rank larger than that of the Standard Model, and the simplest way to spontaneously reduce the rank, for example, via the vev of a spinor 16-plet in $SO(10)$, leads to a large spontaneous breaking of the discrete matter parity subgroup of $SO(10)$ (Refs. 42, 43). Thus, theories based on $SO(10)$ need a further ingredient to ensure sufficient suppression of B and L violation of the low-energy theory. One possibility is that the spinor vev does not introduce the dangerous couplings, which typically requires a discrete symmetry beyond $SO(10)$. Alternatively, the rank may be broken by larger Higgs multiplets,[42] for example, the 126 representation of $SO(10)$. Finally, if the reduction of rank occurs at low energies, the resulting R_p-violating phenomenology may be acceptable[43]; however, the weak mixing angle prediction is then lost. The flipped $SU(5)$ gauge group allows for models with renormalizable L violation, but highly suppressed B violation[44]; however, these theories also lose the weak mixing angle prediction.

There are other possibilities for X to be a discrete subgroup of an enlarged gauge symmetry. Several Z_N examples from E_6 are possible.[35] Such a symmetry will be an anomaly-free discrete gauge symmetry, and it has been argued that if X is discrete, it should be anomaly free in order not to be violated by Planck scale physics.[45] With the minimal low-energy field content, there are only two such possibilities which commute with flavor: the familiar case of matter parity, and a Z_3 baryon parity,[46] which also prohibits baryon number violation from dimension five operators. While the gauge origin of X remains a likely possibility, we are not aware of explicit compelling models which achieve this.

Finally, we discuss the possibility that the X symmetry is a flavor symmetry: the symmetry which is ultimately responsible for the small parameters of the quark and lepton mass matrices, and also of the squark and slepton mass matrices, might provide sufficient suppression for B and L violation. Indeed, this is an extremely plausible solution for the suppression of L violation since the experimental constraints on the coefficients of the L-violating interactions are quite

weak, and would be satisfied by having amplitudes suppressed by powers of small lepton masses. However, the experimental constraints involving B violation are so strong that suppression by small quark mass factors are insufficient.[47] Hence, the real challenge for these theories is to understand the suppression of B violation.

Some of the earliest models involving matter parity violation had a discrete spacetime[34] or gauge[44] origin for B conservation, but had L violation at a rate governed by the small fermion masses. This distinction between B and L arises because left-handed leptons and Higgs doublets are not distinguished by the Standard Model gauge group, whereas quarks are clearly distinguished by their color. This provides a considerable motivation to search for supersymmetric theories with matter parity broken only by the L-violating interactions.

It is not difficult to understand how flavor symmetries could lead to exact matter parity. Consider a supersymmetric theory, with minimal field content and gauge group, which has the flavor group $U(3)^5$ broken only by parameters which transform like the usual three Yukawa coupling matrices. The Yukawa couplings and soft interactions of the most general, such effective theory can be written as a power series in these breaking parameters, leading to a theory known as weak scale effective supersymmetry.[19] The flavor group and transformation properties of the breaking parameters are sufficient to forbid matter parity-violating interactions to all orders: each breaking parameter has an even number of $U(3)$ tensor indices, guaranteeing that all interactions must have an even number of matter fields.[‖] To construct an explicit model along these lines, it is perhaps simplest to start with a $U(3)$ flavor group, with all quarks and leptons transforming as triplets, but Higgs doublets as trivial singlets. The X symmetry is generated by the Z_2 element

$$X(U(3)) = e^{i\pi N_T}, \qquad (III.30)$$

where N_T is the triality of the representation. An exact matter parity will result if the spontaneous breaking of this flavor group occurs only via fields with an even triality.

III.8 Conclusions

The use of flavor symmetries to study both the fermion and scalar masses leads to a new viewpoint. While fermion mass hierarchies remain a very fundamental puzzle, the flavor-changing constraints are definitely *not* a problem for supersymmetry;

[‖]This point was missed in Ref. 19 where R_p was imposed unnecessarily as an additional assumption. We believe that the automatic conservation of R_p makes this scheme an even more attractive framework as a model-independent low-energy effective theory of supersymmetry.

rather they are an advantage. Instead of a flavor-changing problem, we have a tool that allows us to identify which flavor symmetries are acceptable. Furthermore, many acceptable flavor symmetries lead to flavor-changing phenomena beyond the Standard Model which should be discovered in the not too distant future. Such discoveries provide the best hope for progress on the fermion mass puzzle.

In this chapter, I have pursued the idea that both fermion and scalar masses should be constrained by the same approximate flavor symmetries. However, fermion masses are supersymmetric while the soft scalar masses are not, so that some decoupling of their symmetry behavior is possible. Suppose that fermion masses are understood in terms of physics at some flavor scale M_f. If $M_f < M_{mess}$, the messenger scale of supersymmetry breaking discussed in Sec. I.5, then both fermion and scalar masses are subject to the same flavor symmetries. However, if $M_{mess} < M_f$, as in models with low-energy gauge mediation of supersymmetry breaking,[4] the soft operators can be protected from the physics of fermion mass generation, leading to flavor-changing effects which are milder than those dictated by approximate flavor symmetries.

Broken flavor symmetries are the natural way to describe flavor sectors of supersymmetric theories. For this reason, the MSSM with universal boundary conditions is badly flawed. We advocate replacing the universal boundary condition of Eq. (III.15) with the modified boundary condition of Eq. (III.16) which results from the minimal necessary breaking of $G_{max} = U(3)^5$ (Ref. 19). Any relations between $A_{U,D,E}$ or between m_a^2 should be viewed as probes of gauge unification in the vertical direction. In general, corrections to Eq. (III.16) are expected, as shown in Eq. (III.13d). Finally, in the simplest schemes, the Higgs doublets are not related by flavor symmetries to the three generations of matter, so the Higgs mass parameters should be taken to be independent of m_a^2.

III.9 Appendix A

As an example of the $U(3)$ transformation conventions used in this chapter, I consider the first interaction of Eq. (III.12). Making the transposition explicit, this is

$$W = a\, q^T \epsilon_U u\, h_2. \qquad (A1)$$

Under $U(3)_q$ I take

$$q \to L^* q. \qquad (A2)$$

Under $U(3)_u$ I take

$$u \to R u. \qquad (A3)$$

Hence, if I assign the transformation property

$$\epsilon_U \rightarrow L \epsilon_U R^\dagger \ , \tag{A4}$$

(A1) transforms to $q^T L^\dagger L \ \epsilon_U \ R^\dagger R u \ h_2$ and is therefore invariant. I say that ϵ_U transforms as $(3, \overline{3})$ under $(U(3)_q, U(3)_u)$.

I write the scalar masses as

$$V = q^T m_q^2 q^* + u^\dagger m_u^2 u \tag{A5}$$

so that $m_q^2 \rightarrow L m_q^2 L^\dagger, m_u^2 \rightarrow R m_u^2 R^\dagger$. In building invariant terms, it is useful to notice that $\epsilon_U \epsilon_U^\dagger, \epsilon_D \epsilon_D^\dagger$ transform like m_q^2, while $\epsilon_U^\dagger \epsilon_U$ transforms like m_u^2.

IV. Supersymmetric Grand Unification

IV.1 Introduction

How will we ever be convinced that grand unification, or string theory, or some other physics at very high energies, is correct? Two ways in which this could happen are:

1. The structure of the theory is itself so compelling and tightly constrained, and the links to observed particle interactions are sufficiently strong, that the theory is convincing and is accepted as the standard viewpoint. String theory is a candidate for such a theory, but connections to known physics will require much further understanding of the breaking of its many symmetries.
2. The theory predicts new physics beyond the Standard Model, which is discovered. If the structure of the theory is not very tightly constrained, several such predictions will be necessary for it to become convincing. Grand unification is a candidate for such a theory, but as yet there have been no discoveries beyond the Standard Model. Supersymmetric grand unified theories do have a constrained gauge structure, and this has led to the successful prediction of the weak mixing angle at the percent level of accuracy.[8,48–50**] While significant, this is hardly convincing. Nevertheless, supersymmetric grand unified

**While giving the lectures at SLAC, a bright spark in the audience asked why I chose to quote $\sin^2\theta = 0.231 \pm 0.003$, which suggests a significance of 1%, rather than using the well-measured weak mixing angle as input and quoting a prediction for the less well-measured strong coupling $\alpha_s = 0.126 \pm 0.013$, which looks to only have a significance of 10%. This is an excellent question. The reason I believe that the significance is 1% rather than 10% is as follows. Consider the $\sin^2\theta/\alpha_s$ plane, with $\sin^2\theta$ varying from zero to one, and α_s varying from zero to some large value α_s^c which is still perturbative. The area of this plane is α_s^c, and it could have been that the

theories offer the prospect of many further tests. In this talk, I make the case that experiments of this decade, and the next, allow for the possibility that we might become convinced that grand unification is correct.

Any grand unified theory must have at least two sectors: the gauge sector, which contains the gauge interactions, and the flavor sector containing the interactions which generate the quark and lepton masses. In supersymmetric versions, there are also the supersymmetry breaking interactions. I include the gaugino masses in the gauge sector, the supersymmetry breaking squark, slepton and Higgs masses, and interactions in the flavor sector. There are no known direct observable consequences of the interactions of the superheavy gauge bosons: they are predicted to be too heavy even to mediate proton decay at an observable rate.

I know of only one prediction in the gauge sector, other than $\sin^2\theta$: ratios of the gaugino mass parameters, $M_i, i = 1,2,3$ for $U(1), SU(2)$, and $SU(3)$. If the supersymmetry breaking is hard up to scales above the unification mass, M_G, and if the breaking of supersymmetry in the gauge kinetic function is dominantly $SU(5)$ preserving, then M_i will be independent of i at M_G. Beneath M_G, renormalizations induce splittings between the M_i; in fact, they scale exactly like the gauge couplings: $M_i = \alpha_i M$. The prediction of two gaugino mass ratios is a very important consequence of super unification. These predictions occur in the gauge sector; however, unlike the weak mixing angle, these predictions involve the supersymmetry breaking sector, and even if the supersymmetry breaking is hard at M_G, there are situations when they are broken.[51] Furthermore, these relations can occur without grand unification.[††]

---

parameters lie anywhere in this plane. The condition that the three gauge couplings unify can be represented as a band in this plane, with the width of the band representing the theoretical uncertainties, such as the various threshold corrections. By sketching the plane, you can convince yourself that the area of this band is given by $\alpha_s^c \Delta$, where Δ is the theoretical uncertainty in $\sin^2\theta$. Hence, the fraction of the area of the plane which the theory allows is Δ, which is of order 1%, and this is a measure of the significance of the prediction. This argument can be rephrased by starting in some other basis for the parameters, *e.g.*, the space of g_1, g_2, and g_3 with α held fixed, but the conclusion will be the same.

[††]Suppose supersymmetry is broken in a sector which communicates with the observable sector only via Standard Model gauge interactions. Then one expects $M_i \propto \alpha_i$ as before. The constant of proportionality is not guaranteed to be independent of i, although such an independence follows if the particles communicating the supersymmetry breaking fill out complete $SU(5)$ multiplets, as suggested by the weak mixing angle prediction.

IV.2 Flavor Signals Compared

Fortunately, the flavor sector has many signatures, listed in Table 3 in five categories. Proton decay[52,53] and neutrino masses[54,55] are the earliest and most well-known signatures of grand unification. However, the theoretical expectation for these classic signals is plagued by a power dependence on an unknown superheavy mass scale. For neutrino masses, this is the right-handed Majorana mass M_R. If we naively set $m_{\nu_i} = m_{u_i}^2/M_R$ with $M_R = M_G = 2 \times 10^{16}$ GeV, then all three neutrino masses are too small to be detected in any laboratory experiment, although they could lead to MSW oscillations in the sun.

While the many hints for detection of neutrino oscillations are extremely interesting, and theorists are full of ideas for suppressing M_R, if we fail to detect neutrino masses then we learn very little about grand unification. On the other hand, several observations hint at the presence of neutrino masses, and measurements of neutrino mass ratios and mixing angles would provide a very important probe of the flavor structure of unified models.

	Requires BSM discovery	"Present" in all models	Requires SUSY breaking hard at M_G
(I) p decay	$\surd$	No	No
(II) ν masses	$\surd$	No	No
(III) u, d, e masses and mixings	No	No	No
(IV) $\tilde{u}, \tilde{d}, \tilde{e}$ masses	$\surd$	$\surd$	$\surd$
(V) $L_{e,\mu,\tau}$ and CP violation	$\surd$	$\surd$	$\surd$

Table 3. Characteristic features of the five flavor tests of supersymmetric grand unification.

The leading supersymmetric contribution to the proton decay rate is proportional to M_H^{-2} (Refs. 37 and 40), where M_H is a model-dependent parameter, which arises from the unified symmetry breaking sector of the theory. The simple

expectation that $M_H \simeq M_G$ is excluded as it produces too short a proton lifetime.[37,40] There are many mechanisms that effectively allow M_H to be enhanced, thereby stabilizing the proton, but there is no argument, which I would defend, demonstrating that proton decay will be within reach of future experiments. If we are lucky, proton decay may be discovered, and the decay modes and branching ratios will probe flavor physics in an important way. However, as for neutrino masses, if a signal is not seen, little of use is learned about the question of grand unification, hence the "No" in the middle column of Table 3.

The third signature of the flavor sector of grand unified theories is provided by relations amongst the masses and mixings of the quarks and charged leptons, which was also first studied in the 1970s (Ref. 56). This signature has the very great advantage over all others that data exists: there is no need for discoveries beyond the Standard Model. Since the late '70s, this field has developed considerably, in step with our continually increasing knowledge of the quark and lepton masses and the Kobayashi-Maskawa matrix elements. These signatures are based on the hope that the flavor interactions which generate the fermion masses are relatively simple, involving few enough parameters that relations among the 13 observables can be derived. While there is no guarantee that this is true, it is an assumption which is reasonable and which could have an enormous payoff. A considerable fraction of high-energy physics experiments aim at extracting more precise values for the quark masses and mixings; each time an error bar is reduced, this probe of grand unification becomes more incisive. Among the interesting results obtained so far are:

- Evolution of the b and τ Yukawa couplings to high energies in the Standard Model does not lead to their unification, as expected from the simple $SU(5)$ boundary condition. Such a unification does work well if evolution is done with weak scale supersymmetry and a heavy top quark.[57–60]
- The unification of the three Yukawa couplings of the heavy generation in the MSSM,[61] expected from a simple $SO(10)$ boundary condition, can occur perturbatively only if 165 GeV $< m_t <$ 190 GeV (Ref. 62).
- It is possible to construct $SO(10)$ models where all observed fermion masses and mixings are generated from just four interactions. Seven of the 13 flavor parameters are predicted.[32]

- The observed quark masses and mixings may be consistent with several patterns of the Yukawa matrices at the unification scale in which many of the entries are zero, suggesting they have a simple origin.[63]

I have discussed the first three signatures of Table 1, stressing that only for fermion mass relations do we have any useful data, and stressing that none of these signatures is a necessary consequence of grand unification. These features are shown in the first two columns of the table. We must now discuss supersymmetry breaking, which is relevant for the third column of Table 3. The fundamental origin of the first three signatures (baryon number violation, lepton number violation, and Yukawa coupling relations) does not depend on supersymmetry breaking. However, for the last two signatures, the supersymmetry breaking interactions of the low-energy effective theory contain all the information relevant to the signals.

A crucial question for these two signatures is: at what scale do the interactions which break supersymmetry become soft? This has nothing to do with the size of the parameters which violate supersymmetry—they are of order of the weak scale. At any energy scale, μ, we can consider our theory to be a local effective field theory. What is the "messenger scale," M_{mess}, above which the supersymmetry breaking parameters, such as squark and gluino masses, do not arise from a single local interaction? Consider models where supersymmetry is broken spontaneously in a sector with a single mass scale, M, and is communicated to the observable sector by the known gauge interactions.[4,64] It is only when the particles of mass M are integrated out of the theory that local interactions are generated for squark and gluino masses. Hence, for these models, the messenger scale is given by $M_{mess} = M$, which is of order M_W/α, or 10 TeV.

The breaking of supersymmetry in a hidden sector of $N = 1$ supergravity theories[17,18] has become a popular view (although it is not satisfactory in several respects). The interactions which generate squark and slepton masses are produced when supergravity auxiliary fields are eliminated from the theory, and hence are local at all energies up to the Planck scale, giving a messenger scale $M_{mess} = M_{Pl}$. For signatures IV and V, the critical question is whether M_{mess} is larger or smaller than M_G, the unification mass. If $M_{mess} \ll M_G$, then the local interactions which break supersymmetry are produced at energies beneath M_G, and hence these interactions are not renormalized by the interactions of the unified theory. On the other hand, if $M_{mess} \geq M_G$, then the supersymmetry breaking interactions appear as local interactions in the grand unified theory itself. At energies above M_G, they take a form which is constrained by the unified

symmetry. Furthermore, they are modified by radiative corrections induced by the unified theory, giving low-energy signals which are not power suppressed by M_G (Ref. 65).

For example, in any grand unified theory in which $\tilde{u}$, $\tilde{u}^c$, and $\tilde{e}^c$ are unified in the same irreducible representation, the unified theory will possess $m^2_{\tilde{u}} = m^2_{\tilde{u}^c} = m^2_{\tilde{e}^c}$. When the unified gauge symmetry is broken, such relations can be modified both radiatively and at tree level. However, it has been shown that in all models where the weak mixing angle is a significant prediction of the theory, there will be two scalar superpartner mass relations for each of the lightest generations.[66]

It is possible that the gauge forces are unified, but the low-energy matter particles are not; for example, $\tilde{u}$, $\tilde{u}^c$, and $\tilde{e}^c$ could lie in different irreducible representations of the unified group. In this case, the unified gauge group clearly does not lead to scalar mass relations amongst the light states. While this situation is a logical possibility, I do not find it very plausible. It is not straightforward to construct such theories and maintain an understanding for the smallness of the flavor-mixing angles of the Kobayashi-Maskawa mixing matrix. Much more likely is the possibility that the light mass eigenstate fields $\tilde{u}$, $\tilde{u}^c$, and $\tilde{e}^c$ lie dominantly in one irreducible representation, but have small components in other representations.[67] This happens automatically in Froggatt-Nielsen theories of fermion masses[31] which rely heavily on mass mixing between heavy and light states. Such small mixings will lead to corresponding small deviations from the exact unified scalar mass relations of Ref. 66. In principle, these shifts in the scalar mass eigenvalues would allow s-particle spectroscopy to be used as a probe of the unified theory.[67] However, I doubt they will be big enough to be directly seen in spectroscopy. This is because the mass mixings also induce flavor-changing effects in the scalar sector, and these are powerfully constrained by experiment. Since this phenomenon occurs at tree level, it is likely to dominate over the flavor-changing effects that the unified theory will induce at the loop level,[65] and hence will become one of the most important constraints on building theories of fermion masses using the Froggatt-Nielsen method. Hence, I think that simple scalar mass relations are likely to result in unified theories, while the flavor-changing phenomenology will probe details of the flavor structure of the unified theory.

IV.3 Flavor-Changing and CP-Violating Signals

Riccardo Barbieri and I have recently shown that a new class of signatures arises in supersymmetric theories which unify the top quark and τ lepton, and which have a high messenger scale $M_{mess} > M_G$ (Ref. 68). These effects are induced

by radiative corrections involving the large top Yukawa coupling of the unified theory, λ_{tG}. The most promising discovery signatures are lepton flavor violation, such as $\mu \to e\gamma$ (Refs. 68 and 69) and electric dipole moments for the electron and neutron, d_e and d_n (Refs. 69 and 70).

These signatures are complementary to the classic tests of proton decay and neutrino masses, as shown in the last two columns of Table 1. We believe that these new signatures are much less model dependent than the classic tests: they are present in a very wide range of models with $M_{mess} > M_G$. A second crucial point, when comparing with the classic tests, is the size of these signals, which does not depend on the power of an unknown superheavy mass.

A complete calculation in the minimal $SU(5)$ and $SO(10)$ models[69] concludes that searches for the L_i and CP-violating signatures provide the most powerful known probes of supersymmetric quark-lepton unification with supersymmetry breaking generated at the Planck scale. For example, an experiment with a sensitivity of 10^{-13} to B.R. $(\mu \to e\gamma)$ would probe (apart from a small region of parameter space where cancellations in the amplitude occur) the $SU(5)$ model to $\lambda_{tG} = 1.4$ and $m_{\tilde{e}_R} = 100$ GeV, and would explore a significant portion of parameter space for $m_{\tilde{e}_R} = 300$ GeV. In the $SO(10)$ case, where the present bound on $\mu \to e\gamma$ is already more stringent than the limits from high-energy accelerator experiments, a sensitivity of 10^{-13} would probe the theory to $\lambda_{tG} = 1.25$ and $m_{\tilde{e}_R}$ close to 1 TeV.

Which search probes the theory more powerfully: rare muon processes or the electric dipole moments? In the minimal SU(5) theory, the electric dipole moments are very small so that the rare muon processes win. In the minimal SO(10) theory, the electric dipole moments are proportional to $\sin\phi$ where $\phi = \phi_d - 2\beta$, where $-\beta$ is the phase of the Kobayashi-Maskawa matrix element V_{td}, and where ϕ_d is a new phase. There is a simple relation between B.R. $(\mu \to e\gamma)$ and d_e:

$$\frac{|d_e|}{10^{-27}\text{e cm}} = 1.3\ \sin\phi \sqrt{\frac{\text{B.R.}(\mu \to e\gamma)}{10^{-12}}}. \qquad (IV.1)$$

For $\sin\phi = 0.5$, the present limits imply that the processes have equal power to probe the theory. The analysis of the data from the MEGA experiment should put the rare muon decay ahead, but eventually d_e may win because it falls only as the square of the superpartner mass, whereas the rare muon decay rate falls as the fourth power. At some point, these processes could force the s-electron masses to be

higher than is reasonable from the viewpoint of electroweak symmetry breaking, discussed in Sec. II.3.

Similar new flavor-changing tests of supersymmetric quark-lepton unification occur in the hadronic sector, where the best probes are nonstandard model contributions to $\epsilon, b \to s\gamma$ and to CP violation in neutral B meson decays.[71] These signals could provide a powerful probe of the flavor sector of unified theories. However, unlike the lepton flavor-violating and electric dipole signatures, they must be distinguished from the Standard Model contribution, and they are small when the gluino is heavy due to a gluino focusing effect on the squark masses.

Unified flavor sectors which are more complicated than the minimal ones lead to a larger range of predictions for these signals. There may be additional sources of flavor and CP violation other than those generated by the top Yukawa coupling. While cancelling contributions cannot be ruled out, they are unlikely to lead to large suppressions. Many other sources could provide effects which are larger than those generated by λ_{tG}, and hence, it is reasonable to take the top contribution as an indication of the minimum signal to be expected.

IV.4 The Top Quark Origin of New Flavor and CP Violation

At first sight, it is surprising that the top quark Yukawa coupling should lead to any violation of L_e or L_μ. What is the physical origin of this effect, and why is it not suppressed by inverse powers of M_G? The answer lies in new flavor-mixing matrices, which are analogous to the Kobayashi-Maskawa matrix.

In the Standard Model, the quark mass eigenstate basis is reached by making independent rotations on the left-handed up and down type quarks, u_L and d_L. However, these states are unified into a doublet of the weak SU(2) gauge group: $Q = (u_L, d_L)$. A relative rotation between u_L and d_L therefore leads to flavor mixing at the charged W gauge vertex. This is the well-known Cabibbo-Kobayashi-Maskawa mixing. With massless neutrinos, the Standard Model has no analogous flavor mixing amongst the leptons: the charged lepton mass eigenstate basis can be reached by a rotation of the entire lepton doublet $L = (\nu_L, e_L)$.

How are these considerations of flavor mixing altered in supersymmetric unified theories? There are two new crucial ingredients. The first is provided by weak-scale supersymmetry, which implies that the quarks and leptons have scalar partners. The mass eigenstate basis for these squarks and sleptons requires additional flavor rotations. As an example, consider softly broken supersymmetric QED with three generations of charged leptons. There are three arbitrary mass matrices, one for the charged leptons, $e_{L,R}$, and one each for the left-handed and

right-handed sleptons, $\tilde{e}_L$ and $\tilde{e}_R$. To reach the mass basis therefore requires a relative rotation between $e_{L,R}$ and $\tilde{e}_{L,R}$, resulting in a flavor-mixing matrix at the photino gauge vertex. These matrices were called W^{e_L,e_R} in Sec. III.3.

In supersymmetric extensions of the Standard Model, these additional flavor-changing effects are known to be problematic. With a mixing angle comparable to the Cabibbo angle, a branching ratio for $\mu \to e\gamma$ of order 10^{-4} results. In the majority of supersymmetric models which have been constructed, such flavor-changing effects have been suppressed by assuming that the origin of supersymmetry breaking is flavor blind. In this case, the slepton mass matrix is proportional to the unit matrix. The lepton mass matrix can then be diagonalized by identical rotations on $e_{L,R}$ and $\tilde{e}_{L,R}$, without introducing flavor-violating mixing matrices at the gaugino vertices. *Slepton degeneracy renders lepton flavor-mixing matrices nonphysical.*

The unification of quarks and leptons into larger multiplets provides the second crucial new feature in the origin of flavor mixing. The weak unification of u_L and d_L into q_L is extended in SU(5) to the unification of q_L with u_R and e_R into a ten-dimensional multiplet $T(q_L, u_R, e_R)$. Since higher unification leads to fewer multiplets, there are fewer rotations which can be made without generating flavor-mixing matrices.

In any supersymmetric unified model, there must be at least two coupling matrices, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, which describe quark masses. If there is only one such matrix, it can always be diagonalized without introducing quark mixing. One of these coupling matrices, which we take to be $\boldsymbol{\lambda}_1$, must contain the large coupling λ_t, which is responsible for the top quark mass. We choose to work in a basis in which $\boldsymbol{\lambda}_1$ is diagonal. The particles which interact via λ_t are those which lie in the same unified multiplet with t_L and t_R. In all unified models, this includes a right-handed charged lepton, which we call e_{R_3}. This cannot be identified as the mass eigenstate τ_R, because significant contributions to the charged lepton masses must come from the matrix $\boldsymbol{\lambda}_2$, which is not diagonal.

The assumption that the supersymmetry breaking mechanism is flavor blind leads to mass matrices for both $\tilde{e}_L$ and $\tilde{e}_R$, which are proportional to the unit matrix at the Planck scale, M_{Pl}. As we have seen, without unified interactions, lepton superfield rotations can diagonalize the lepton mass matrix without introducing flavor-mixing matrices. However, the unification prevents such rotations: the leptons are in the same multiplets as quarks, and the basis has already been

chosen to diagonalize $\boldsymbol{\lambda}_1$. As the theory is renormalization group scaled to lower energies, the λ_t interaction induces radiative corrections which suppress the mass of $\tilde{e}_{R_3}$ beneath that of $\tilde{e}_{R_1}$ and $\tilde{e}_{R_2}$. Beneath M_G, the superheavy particles of the theory can be decoupled, leaving only the interactions of the minimal supersymmetric Standard Model. Now that the unified symmetry which relates quarks to leptons is broken, a lepton mass basis can be chosen by rotating lepton fields relative to quark fields. However, at these lower energies, the sleptons are no longer degenerate, so that these rotations do induce lepton flavor-mixing angles. *Radiative corrections induced by λ_t lead to slepton nondegeneracies, which render the lepton mixing angles physical.*

This discussion provides the essence of the physics mechanism for $L_{e,\mu,\tau}$ violation in superunified models. It shows the effect to be generic to the idea of quark-lepton unification, requiring only that supersymmetry survive unbroken to the weak scale, and that supersymmetry breaking be present at the Planck scale. The imprint of the unified interactions is made on the soft supersymmetry breaking coefficients, including the scalar trilinears, which are taken to be universal at the Planck scale. Eventually, this imprint will be seen directly by studying the superpartner spectrum, but it can also be probed now by searching for $L_{e,\mu,\tau}$ and CP-violating effects.

The above discussion assumed a universal scalar mass at high energies. We argued in Chap. III that it is preferable to replace this ad hoc form with scalar masses that are the most general allowed by an appropriate flavor group, G_f. This group solves the "1–2" flavor problem, as discussed in Sec. III.3, but the "1, 2–3" flavor signature discussed here, which results from the large splitting between the scalars of the third generation and those of the lighter two generations, will persist.

IV.5 Summary

Supersymmetric grand unified theories are a leading candidate for physics beyond the Standard Model because:

- They provide an elegant group theoretic understanding of the gauge quantum numbers of a generation.
- $\sin^2\theta$ is the only successful prediction of any parameter of the Standard Model at the percent level of accuracy.

I have not yet mentioned the most crucial experimental hurdle which these theories must pass: superpartners must be discovered at the weak scale. Without

this, I will never be convinced that these theories are correct. As I write, I imagine the skeptics who may read this (I dare to hope!) saying "suppose by 2010 we have measured neutrino masses and mixing angles, seen proton decay and other rare processes such as $\mu \to e\gamma$, d_e and d_n, found nonstandard CP violation in B meson decays, and that we have even discovered superpartners and measured their masses. This still will not convince me that the theory behind this physics is quark-lepton unification." My reply is:

- These discoveries will not necessarily make quark-lepton unification convincing, but they will make it the standard picture.
- These discoveries might make a particular model of quark-lepton unification completely convincing.

There is certainly no guarantee of the latter point, but let me illustrate it with an optimistic viewpoint. There are millions of possible flavor sectors of unified models. Some are so complicated that, if this is the way nature is, we are unlikely to ever uncover this structure from low-energy experiments alone. Others are very simple with few interactions and parameters. Why should nature be kind to us and provide a simple flavor sector with few interactions? Quite apart from our general belief that the underlying laws of physics will be simple, I think that the answer is illustrated by the $U(2)$ model of Sec. III.7. A flavor symmetry provides a convincing solution to the flavor-changing problem. Since it must severely constrain the scalar sector, it is expected to also severely restrict the fermion mass operators. The most constrained scheme which I know has ten parameters (eight flavor and two supersymmetry breaking) to describe all the flavor physics signals. As an example, consider something in between with, say, 15 parameters (e.g., 12 flavor and three supersymmetry breaking). This has two parameters more than the flavor sector of the Standard Model. Suppose that we discover such a unified model with these two parameters correctly describing the entire superpartner spectrum, the neutrino masses and mixing angles and the magnitudes of the nonstandard model signals for $\mu \to e\gamma, d_e, d_n$ and B meson CP violation, and the masses of the two Higgs bosons, the pseudoscalar boson and the charged Higgs boson. It is certainly an optimistic scenario, but it is one which I would find convincing.

V. The High-Energy Frontier

What are the liveliest debates at the high-energy frontier today? Particle physics, like other branches of physics, is driven first and foremost by experimental discov-

eries. Many experimental discoveries laid the groundwork for the development of the gauge structure of the Standard Model, and we will need many further experiments to guide us beyond. Hence, it is not surprising that the dominant debate of the field is about which accelerators should be built and which experiments should be done.

The phenomena uncovered by experiments have led to a stunning array of theoretical developments over the last 30 years. These theoretical tools allowed the construction of the Standard Model. A dominant debate in theoretical circles is whether the tools of point particle field theories and their symmetries will take us much further, or whether further tools, such as string theory, are necessary.

There is no doubt that there are limits to point particle gauge theory, the clearest of which is that they cannot describe gravity. Nevertheless, point particle gauge theories and their symmetries are an extraordinarily rich and powerful tool. In these lectures, I have explored the possibility that they provide a deeper understanding of many of the outstanding questions of particle physics.

- A dynamical origin of electroweak symmetry breaking as a heavy top quark effect.
- A flavor symmetry origin for the pattern of fermion masses and mixing.
- A unified gauge symmetry—allowing for a highly constrained and predictive theory of flavor, in addition to the well-known picture of a unified family and unified gauge couplings.

It is extraordinary that such a comprehensive vision of particle interactions has been developed. It seems unlikely that a complete picture of particle physics can be constructed without nonperturbative dynamics entering at some point; but what is that point? It is possible that the failure to develop a comprehensive vision of particle physics beyond the Standard Model based on either technicolor or a composite Higgs is because in these cases, the issue of nonperturbative dynamics provides a barrier at the very first step. The vision developed here is largely perturbative and is based on weak-scale supersymmetry, a heavy top quark leading to perturbative dynamics for electroweak symmetry breaking, and perturbative unification. The only new nonperturbative dynamics beneath the Planck scale occurs in the supersymmetry breaking sector, which I have not discussed. Fortunately, there are many experimentally testable aspects of the theory which follow from a few minimal assumptions, and no detailed understanding, about how supersymmetry breaking occurs. Measurements of the superpartner masses

will provide a crucial guide as to how the supersymmetry breaking interactions should be generated.

The vision of weak scale supersymmetry and perturbative unification receives considerable motivation from precision electroweak measurements, but only further experiments will prove whether these ideas are correct. The discovery of supersymmetry at the weak scale would be a revolution for High-Energy Physics, as important as any the field has seen, heralding a new era. Decades of experimentation would be needed to fully elucidate the ramifications of this new symmetry; for example, measurements of the many new flavor observables would provide a new handle on the flavor problem.

Acknowledgments

I am grateful to the organizers for arranging a very stimulating school.

References

[1] J. Wess and J. Bagger, *Supersymmetry and Supergravity* (Princeton University Press, 1992).

[2] H. Haber and G. Kane, Phys. Rep. **117**, 75 (1985).

[3] L. Ibanez and G. G. Ross, Phys. Lett. B **110**, 215 (1982).

[4] L. Alvarez-Gaumé, M. Claudson, and M. B. Wise, Nucl. Phys. B **207**, 96 (1982).

[5] R. Barbieri and J. Guidice, Nucl. Phys. B **306**, 63 (1988).

[6] G. Anderson and D. Castano, Phys. Rev. D **52**, 1693 (1995); MIT-CTP-2464 (1995).

[7] S. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[8] S. Dimopoulos and H. Georgi, Nucl. Phys. B **193**, 150 (1981).

[9] R. S. Chivukula and H. Georgi, Phys. Lett. B **188**, 99 (1987) and Phys. Rev. D **36**, 2102 (1987).

[10] S. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977).

[11] A. Antaramian, L. J. Hall, and A. Rašin, Phys. Rev. Lett. **69**, 1971 (1992).

[12] L. J. Hall and S. Weinberg, Phys. Rev. D **48**, 979 (1993).

[13] J. S. Hagelin, S. Kelley, and T. Tanaka, Nucl. Phys. B **415**, 293 (1994).

[14] E. Gabrielli, A. Masiero, and L. Silvestrini, ROME1-1109/95, hep-ph/9509379 (1995).

[15] N. Arkani-Hamed, H.-C. Cheng, and L. J. Hall, LBL-37894 (1995).

[16] J. Ellis, S. Ferrara, and D. V. Nanopoulos, Phys. Lett. B **114**, 231 (1982); J. Polchinski and M. B. Wise, Phys. Lett. B **125**, 393 (1983); M. Dugan, B. Grinstein, and L. J. Hall, Nucl. Phys. B **255**, 413 (1985).

[17] R. Barbieri, S. Ferrara, and C. Savoy, Phys. Lett. B **110**, 343 (1982); P. Nath, R. Arnowitt, and A. Chamseddine, Phys. Rev. Lett. **49**, 970 (1982).

[18] L. J. Hall, J. Lykken, and S. Weinberg, Phys. Rev. D **27**, 2359 (1983).

[19] L. J. Hall and L. Randall, Phys. Rev. Lett. **65**, 2939 (1990).

[20] M. Dine, R. Leigh, and A. Kagan, Phys. Rev. D **48**, 4269 (1993).

[21] Y. Nir and N. Seiberg, Phys. Lett. B **309**, 337 (1993).

[22] P. Pouliot and N. Seiberg, Phys. Lett. B **318**, 169 (1993).

[23] M. Leurer, Y. Nir, and N. Seiberg, Nucl. Phys. B **398**, 319 (1993).

[24] D. Kaplan and M. Schmaltz, Phys. Rev. D **49**, 3741 (1994).

[25] L. J. Hall and H. Murayama, Phys. Rev. Lett. **75**, 3985 (1995).

[26] C. Carone, L. J. Hall, and H. Murayama, LBL 38047 (1995).

[27] C. Carone, L. J. Hall, and H. Murayama, LBL 38380 (1996).

[28] A. Pomarol and D. Tommasini, CERN-TH/95-207 (1995).

[29] R. Barbieri, G. Dvali, and L. J. Hall, LBL 38065, Phys. Lett. to appear (1996); R. Barbieri and L. J. Hall, LBL-38381 (1996).

[30] E. Dudas, S. Pokorski, and C. A. Savoy, Saclay T95/094, hep-ph/9509410.

[31] C. D. Froggatt and H. B. Nielsen, Nucl. Phys. B **147**, 277 (1979).

[32] G. Anderson, S. Dimopoulos, L. J. Hall, S. Raby, and G. Starkman, Phys. Rev. D **49**, 3660 (1994).

[33] E. Witten, Nucl. Phys. B **188**, 513 (1981).

[34] L. J. Hall and M. Suzuki, Nucl. Phys. B **231**, 419 (1984).

[35] M. Bento, L. Hall, and G. G. Ross, Nucl. Phys. B **292**, 400 (1987).

[36] S. Dimopoulos and L. J. Hall, Phys. Lett. B **207**, 210 (1987).

[37] S. Weinberg, Phys. Rev. D **26**, 287 (1982).

[38] G. Farrar and P. Fayet, Phys. Lett. B **76**, 575 (1978).

[39] L. J. Hall, presented at the Workshop on Quarks and Galaxies, LBL, August 1986, p. 106, edited by I. Hinchliffe (World Scientific, 1987).

[40] N. Sakai and T. Yanagida, Nucl. Phys. B **197**, 533 (1982).

[41] R. N. Cahn, I. Hinchliffe, and L. J. Hall, Phys. Lett. B **109**, 426 (1982).

[42] S. P. Martin, Phys. Rev. D **46**, 2769 (1992).

[43] R. Kuchimanchi and R. Mohapatra, Phys. Rev. D **48**, 4352 (1993).

[44] D. E. Brahm and L. J. Hall, Phys. Rev. D **40**, 2449 (1989).

[45] L. Krauss and F. Wilczek, Phys. Rev. Lett. **62**, 1221 (1989).

[46] L. E. Ibanez and G. G. Ross, Phys. Lett. B **260**, 291 (1991); Nucl. Phys. B **368**, 3 (1992).

[47] I. Hinchliffe and T. Kaeding, Phys. Rev. D **52**, 5319 (1995).

[48] H. Georgi, H. Quinn, and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974).

[49] S. Dimopoulos, S. Raby, and F. Wilczek, Phys. Rev. D **24**, 1681 (1981).

[50] L. Ibanez and G. G. Ross, Phys. Lett. B **105**, 439 (1981).

[51] J. Hisano, H. Murayama, and T. Goto, Phys. Rev. D **49**, 1446 (1994).

[52] J. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974).

[53] H. Georgi and S. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[54] M. Gell-Mann, P. Ramond, and R. Slansky, *Supergravity*, edited by P. van Nieuwenhuizen and D. Z. Freedman, 315 (North Holland, 1979).

[55] T. Yanagida, in *Proceedings of the Workshop on Unified Theory*, KEK (1979).

[56] M. Chanowitz, J. Ellis, and M. K. Gaillard, Nucl. Phys. B **128**, 566 (1977).

[57] L. Ibanez and C. Lopez, Phys. Lett. B **126**, 54 (1983).

[58] H. Arason *et al.*, Phys. Rev. Lett. **67**, 2933 (1991).

[59] A. Giveon, L. J. Hall, and U. Sarid, Phys. Lett. B **272**, 138 (1991).

[60] S. Kelly, J. L. Lopez, and D. V. Nanopoulos, Phys. Lett. B **274**, 387 (1992).

[61] B. Ananthanarayan, G. Lazarides, and Q. Shafi, Phys. Rev. D **44**, 1613 (1991).

[62] L. J. Hall, R. Rattazzi, and U. Sarid, Phys. Rev. D **50**, 7048 (1994).

[63] P. Ramond, R. Roberts, and G. G. Ross, Nucl. Phys. B **406**, 19 (1993).

[64] M. Dine and W. Fischler, Nucl. Phys. B **204**, 346 (1982).

[65] L. J. Hall, V. A. Kostelecky, and S. Raby, Nucl. Phys. B **267**, 415 (1986).

[66] H.-C. Cheng and L. J. Hall, Phys. Rev. D **51**, 5289 (1995).

[67] S. Dimopoulos and A. Pomarol, Phys. Lett. B **353**, 222 (1995).

[68] R. Barbieri and L. J. Hall, Phys. Lett. B **338**, 212 (1994).

[69] R. Barbieri, L. J. Hall, and A. Strumia, Nucl. Phys. B **445**, 219 (1995).

[70] S. Dimopoulos and L. J. Hall, Phys. Lett. B **344**, 185 (1995).

[71] R. Barbieri, L. J. Hall, and A. Strumia, Nucl. Phys. B **449**, 437 (1995).

Figure Caption

Upper bounds on superpartner and Higgs boson masses which follow from requiring a limit to the amount of fine tuning among parameters. This figure applies to the supersymmetric extension of the Standard Model with minimal field content, with all scalar masses taken equal at the unification scale, and similarly for the three gaugino masses. The upper extent of the lines for each particle correspond to $\tilde{\gamma} = 10$, the error bar symbol to $\tilde{\gamma} = 5$, and the squares to the masses which result from minimizing the amount of fine tuning. This figure was supplied to me by Greg Anderson; for further figures, see Ref. 6.

TOP PHYSICS AT CDF

Richard E. Hughes*
Department of Physics and Astronomy
University of Rochester
Rochester, New York 14627

Representing the CDF Collaboration

ABSTRACT

We report on top physics results using a 100 pb^{-1} data sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.8\ TeV$ collected with the Collider Detector at Fermilab (CDF). We have identified top signals in a variety of decay channels, and used these channels to extract a measurement of the top mass and production cross section. A subset of the data (67 pb^{-1}) is used to determine $M_{top} = 176 \pm 8(stat) \pm 10(syst)$ and $\sigma(t\bar{t}) = 7.6\ ^{+2.4}_{-2.0}\ pb$. We present studies of the kinematics of $t\bar{t}$ events and extract the first direct measurement of V_{tb}. Finally, we indicate prospects for future study of top physics at the Tevatron.

*Supported by Department of Energy Grant No. DE-FG02-91ER40685.

©1995 by Richard E. Hughes.

1 Introduction

1.1 Indirect Information on Top

Much of our indirect information on the existence of the top quark has come from studying the bottom quark. By studying the forward-backward asymmetry A_{FB} ($e^+e^- \rightarrow b\bar{b}$) at low energy,[1] and at the Z^0 resonance, combined with the partial width Γ ($Z^0 \rightarrow b\bar{b}$) (Ref. 2), one can determine that the weak-isospin of the b quark is consistent with $T_3 = -\frac{1}{2}$. In addition, the absence of decays such as $b \rightarrow sl^+l^-$ (at a rate comparable to $b \rightarrow cl^-\nu$) can be explained if the b quark is a member of an SU(2) doublet.[3] Finally, cancellation of anomalies requires that the electric charges in each generation sum to zero. By definition, the top quark is the $T_3 = +\frac{1}{2}$ weak isospin partner to the b quark.

Given fits to electroweak observables from LEP and SLC, that also include information from neutrino scattering, and the measurement of M_W from $p\bar{p}$ colliders, one obtains[2] a favored value for the top mass of $M_{top} = 178\ ^{+11}_{-11}\ ^{+18}_{-19}\ GeV/c^2$. This value is obtained assuming a value for the Higgs mass of $M_H = 300\ GeV/c^2$. The first uncertainty is experimental, while the second comes from varying the Higgs mass over the range of 60 to 1000 GeV/c^2. If instead one leaves the Higgs mass free in the fit, the resulting top mass is $M_{top} = 156\ ^{+14}_{-15}\ GeV/c^2$ (Ref. 4).

1.2 Direct Searches for Top

A number of direct searches for the top quark have been performed at $p\bar{p}$ colliders. In 1990, CDF obtained a limit[5] of $M_{top} > 91\ GeV/c^2$ at 95 % CL. In 1993, DØ obtained a limit[6] of $M_{top} > 131\ GeV/c^2$ at 95 % CL. The first evidence for top-quark production was obtained by CDF in 1994 (Ref. 7). In that analysis, we found a 2.8σ excess of signal over the expectation from background, using a data sample with an integrated luminosity of $19.3 \pm 0.7\ pb^{-1}$ (referred to as the Run 1A sample). The interpretation of the excess as top-quark production was supported by a peak in the mass distribution for fully reconstructed events. Additional evidence was found in the jet energy distributions in lepton + jet events.[8]

A second collider run took place from January 1994 until June 1995. This is referred to as Run 1B and yielded an additional $\sim 90\ pb^{-1}$ of data. On March 2nd, 1995, both the CDF[9] and DØ (Ref. 10) Collaborations reported observation of the top quark, using the Run 1A data plus a subset of the Run 1B data set. In

the remainder of this document, we will describe the CDF results, including an updated Run 1B data set.

1.3 Top Production and Decay

In $p\bar{p}$ collisions, top quarks are pair produced by gluon-gluon fusion and $q\bar{q}$ annihilation. The relative importance of the two processes is dependent on the mass of the top quark. For $M_{top} = 175\ GeV/c^2$, one expects approximately 90% of the rate from $q\bar{q}$ annihilation. The $t\bar{t}$ production cross section has been calculated at next-to-leading order, with the inclusion of diagrams due to soft gluon emission.[11] The result is shown in Fig. 1. The dashed curves in this figure represent the uncertainties obtained in varying parameters related to the perturbative part of the cross-section calculation. They do not represent the full uncertainty in the central value of the cross section, which could be as large as 30% (Ref. 12).

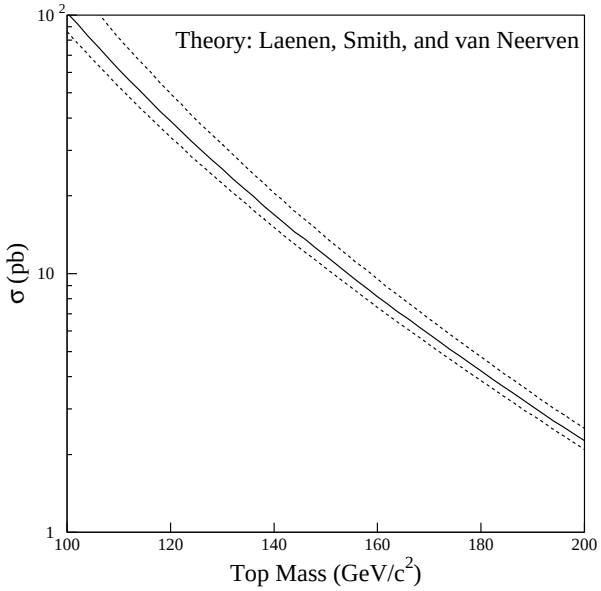


Fig. 1. The theory cross section for $t\bar{t}$ production.

For top masses greater than the combined mass of the W boson and the b quark, the top quark is expected to decay almost exclusively to a real W and a b. The $t\bar{t}$ decay signature is then determined by how the two W bosons in the event decay, as shown in Fig. 2. There are three primary signatures:

- Both W's Decay $W \to \ell\nu$
  In this case, the final state is $\ell^+ \nu \ell^- \nu b\bar{b}$, where ℓ is an electron or a muon. This is referred to as the dilepton channel. The branching ratio is approximately 5%.
- One W Decays $W \to \ell\nu$
  In this case, the final state is $\ell^+\nu \; q\bar{q}' \;\; b\bar{b}$. This is referred to as the lepton plus jets channel, and the branching ratio is $\sim 30\%$.
- Both W's decay $W \to q\bar{q}'$
  In this case, the final state is $q\bar{q}' \; q\bar{q}' \;\; b\bar{b}$. This is referred to as the all-hadronic channel, and the branching ratio is $\sim 44\%$.

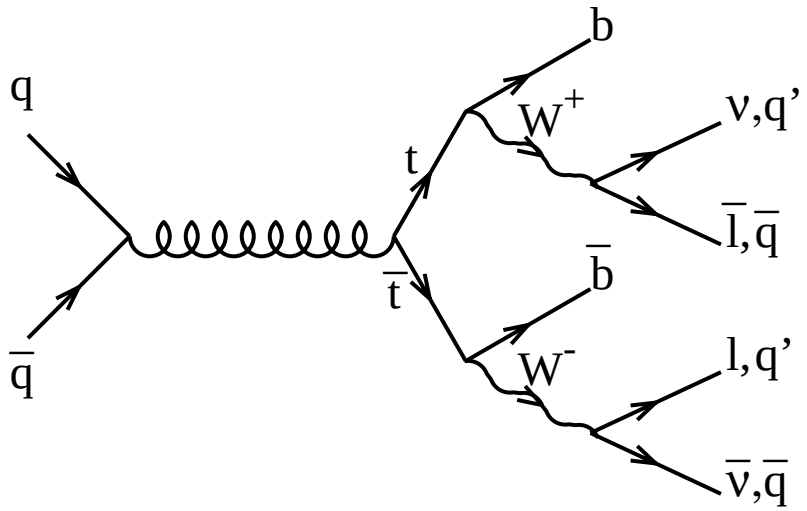


Fig. 2. Tree level $t\bar{t}$ production via $q\bar{q}$ annihilation, followed by the Standard Model decay chain.

1.4 The CDF Detector

The CDF detector is a general-purpose detector designed to study the physics of $p\bar{p}$ collisions. It has both azimuthal and forward-backward symmetry. A side-view cross section of the CDF detector is shown in Fig. 3. The CDF detector consists

of a magnetic spectrometer surrounded by calorimeters and muon chambers.[13] A new low-noise, radiation-hard, four-layer silicon vertex detector (SVX), located immediately outside the beampipe, provides precise track reconstruction in the plane transverse to the beam and is used to identify secondary vertices from b and c quark decays.[14] The momenta of charged particles are measured in the central tracking chamber (CTC), which is in a 1.4 T superconducting solenoidal magnet. Outside the CTC, electromagnetic and hadronic calorimeters cover the pseudorapidity region $|\eta| < 4.2$ (Ref. 15), and are used to identify jets and electron candidates. The calorimeters are also used to measure the missing transverse energy, $\not\!\!E_T$, which can indicate the presence of undetected energetic neutrinos. Outside the calorimeters, drift chambers in the region $|\eta| < 1.0$ provide muon identification. A three-level trigger selects the inclusive electron and muon events used in this analysis.

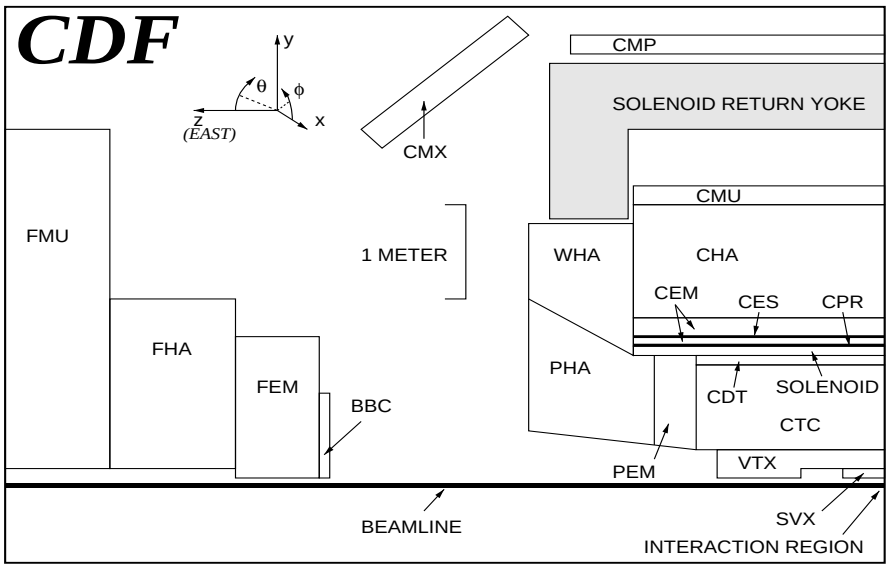


Fig. 3. A side view of the CDF detector. The detector is forward-backward symmetric about the interaction region, which is at the lower-right corner of the figure.

2 The Dilepton Channel

The dilepton selection is based on the expected decay topology $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\nu\ell^-\nu b\bar{b}$, where ℓ is an electron or muon. We require that each event contains at least one primary, isolated electron (muon) with E_T (P_T) ≥ 20 GeV (GeV/c). Electrons consistent with photon conversions are removed. In addition, $Z \rightarrow e^+e^-$ ($\mu^+\mu^-$) are removed if the invariant mass of the lepton pair is in the range of 75 to 105 GeV/c^2. For the secondary lepton, the same E_T (P_T) cut is applied, but a slightly looser set of identification cuts are imposed. It must have a charge opposite in sign to that of the primary lepton.

Since there are two remaining b quarks in the final state, we require that each event contain at least two jets, with uncorrected $E_T \geq 10$ GeV and $|\eta| < 2.0$. The presence of two ν's from the decay of the W bosons motivates a cut on the missing energy $\not\!\!E_T \geq 25$ GeV. An additional cut is applied for the backgrounds of $Z \rightarrow \tau^+\tau^-$ and Drell-Yan. In these events, $\not\!\!E_T$ can be generated along the direction of the leptons (from the τ decay ν's) or along the direction of the jets (from mismeasurement of jet energies). To reduce this background, we require $\not\!\!E_T \geq 50$ GeV, when the azimuthal angle between the $\not\!\!E_T$ and the nearest jet or lepton is $< 20^\circ$.

After all cuts, the relative acceptance among the three possible dilepton categories is 57% for $e\mu$, 28% for $\mu\mu$, and 15% for ee (for a top mass of 175 GeV/c^2). The expected number of events (using the central value of the theoretical cross section) from $t\bar{t}$ passing all cuts is shown in Table 1, for a luminosity of 100 pb^{-1}.

Top Mass (GeV/c^2)	$\sigma_{t\bar{t}}$ (pb)	Expected Events (#)
160	8.2	6.6
170	5.8	4.5
180	4.2	3.6

Table 1. The predicted central value of the $t\bar{t}$ production cross section from Laenen *et al.*, and the number of dilepton $t\bar{t}$ events expected after all cuts, as a function of top-quark mass.

The backgrounds to the dilepton channel are, in order of importance, Drell-Yan, $p\bar{p} \rightarrow Z \rightarrow \tau\tau$, fake leptons, $p\bar{p} \rightarrow WW$, and $p\bar{p} \rightarrow b\bar{b}$. The first three are calculated primarily from the data, while the last two come primarily from Monte Carlo. The individual backgrounds are given in Table 2 and yield a total background of 1.9 ± 0.3.

Category	Expected Events
Drell-Yan	0.70±0.27
$p\bar{p} \rightarrow Z \rightarrow \tau\tau$	0.56±0.11
Fake leptons	0.35±0.11
$p\bar{p} \rightarrow WW$	0.31±0.10
$p\bar{p} \rightarrow b\bar{b}$	0.03±0.02
Total	1.9±0.3

Table 2. Number of background events expected in 100 pb^{-1}, broken down by category.

We observe a total of nine events in 100 pb^{-1}. There are six $e\mu$ events, two $\mu\mu$ events, and one ee event, consistent with the relative acceptance quoted earlier. Figures 4, 5, and 6 show plots of $\Delta\phi$ ($lepton\ or\ jet, \not\!\!E_T$) vs. $\not\!\!E_T$ for the candidates. One of the $\mu\mu$ events is consistent with the radiative decay $Z \rightarrow \mu\mu\gamma$, where the γ is reconstructed as a jet (with large electromagnetic fraction). Although the background expected from this process is small ($\sim$ 0.1 event), and included in the background estimate above, we remove this event from the total sample and are left with eight events passing all selection criteria. The probability of the background estimate fluctuating to the number of observed events or more is $\mathcal{P} = 1$ x 10^{-3}.

As described in the next section, we have two methods for identifying b jets at CDF. Although we do not require b-tagging in the dilepton channel, we note that three of the eight events have one or more b-tagged jets, providing evidence for W^+W^-b, as expected from $t\bar{t}$ production.

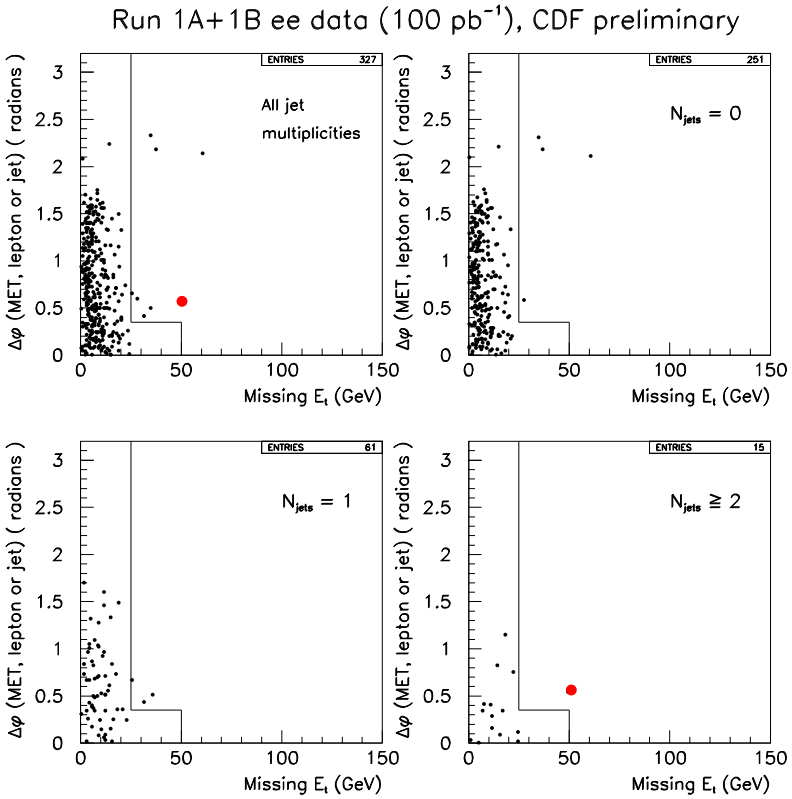


Fig. 4. The azimuthal angle between the $\not\!\!E_T$ and the closest lepton or jet vs. $\not\!\!E_T$ for the ee candidates. Upper left: all-jet multiplicities. Upper right: zero-jet events. Lower left: one-jet events. Lower right: $\geq$ 2-jet events (signal region).

3 The Lepton Plus Jets Channel

The lepton plus jets selection is based on the expected decay topology $t\bar{t} \rightarrow \ell\nu bq\bar{q}'\bar{b}$. Once again, we require that each event contain one isolated electron (muon) with E_T (P_T) $\geq$ 20 GeV (GeV/c). The requirements on this lepton are the same as in the dilepton selection. We also require that $\not\!\!E_T \geq$ 20 GeV. Since there are four partons in the final state, we require that each event have $\geq$ 3 jets, with $E_T \geq 15$ GeV, and $|\eta| \leq 2.0$. The three-jet requirement is approximately 75% efficient for $M_{top} = 175$ GeV/c^2, but strongly suppresses inclusive QCD W production, as shown in Fig. 7. Any events which also pass the dilepton selection criteria are explicitly removed at this point, to keep the two search channels statistically independent.

At this stage in the selection, 296 events remain. For $M_{top} = 175$ GeV/c^2, the theoretical cross section is $\sigma_{t\bar{t}} = 4.8$ pb, and thus, we only expect ≈ 40 $t\bar{t}$ events after all cuts. Additional background rejection is still needed.

The dominant background in a lepton plus jets search is non-top QCD W + multijet production. We reject this background using b-tagging, since every $t\bar{t}$ event

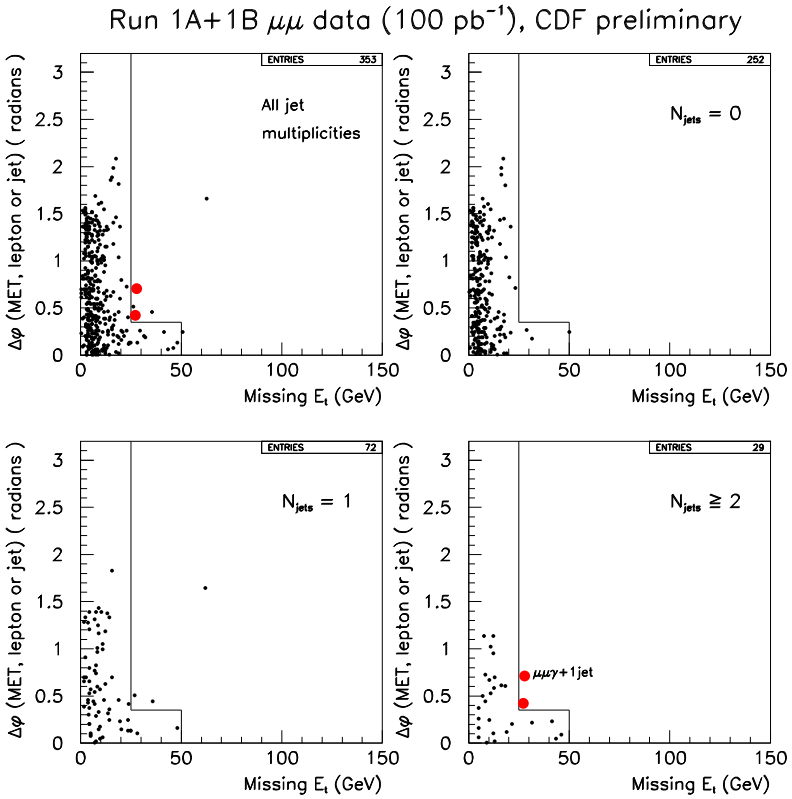


Fig. 5. The azimuthal angle between the $\not\!\!E_T$ and the closest lepton or jet vs. $\not\!\!E_T$ for the $\mu\mu$ candidates. Upper left: all-jet multiplicities. Upper right: zero-jet events. Lower left: one-jet events. Lower right: $\geq$ 2-jet events (signal region).

contains two b quarks, while only $\sim 2\%$ of QCD W + jets events are expected to contain b quarks. We identify or tag b quarks using two methods at CDF: the first requires the location of a displaced vertex using the SVX, and the second requires a soft lepton (e or μ) from b-quark decay.

3.1 Silicon Vertex Tagging

The primary method used for identifying b quarks in top events utilizes the Silicon Vertex Detector and is therefore referred to as SVX-tagging. The method relies on the excellent SVX hit resolution of ~ 8.5 μm per point, as shown in Fig. 8. This in turn yields an expected resolution on the impact parameter of charged tracks of $\sigma_d \sim 16(1 + (\frac{0.8}{P_T})^2)$ μm, where the second term is due to multiple scattering, and P_T is the transverse momentum of the track.

The b lifetime is approximately 450 μm. This, along with the large boost the b receives in the decay of the top quark, yields a displacement in the lab frame which can be quite large. In Fig. 9, we show the transverse decay length in the lab frame of b hadrons from top decay ($M_{top} = 160$ GeV/c^2). Note that this is before

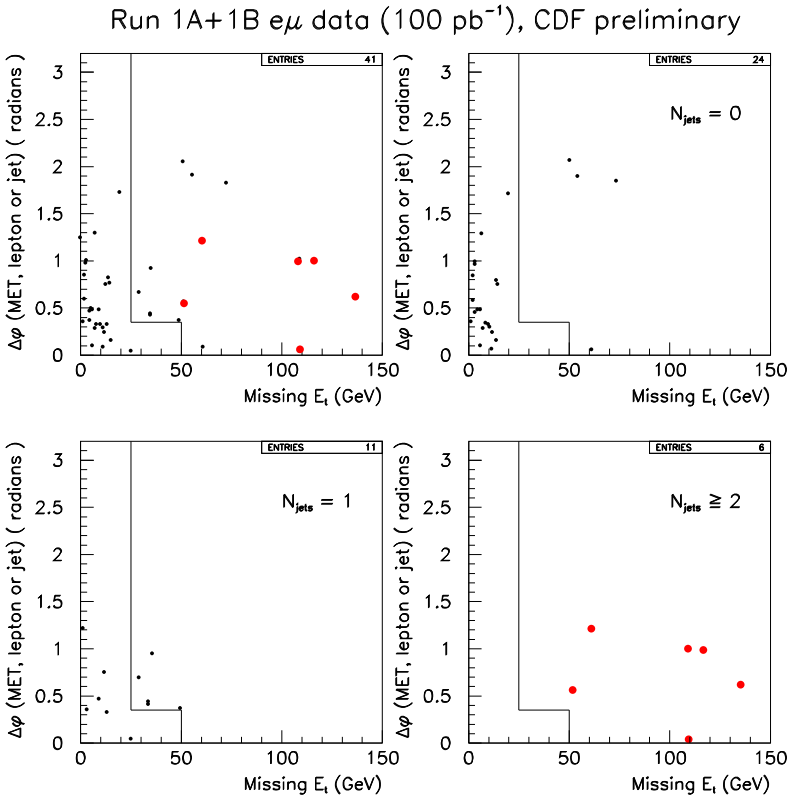


Fig. 6. The azimuthal angle between the $\not\!\!E_T$ and the closest lepton or jet vs. $\not\!\!E_T$ for the $e\mu$ candidates. Upper left: all-jet multiplicities. Upper right: zero-jet events. Lower left: one-jet events. Lower right: $\geq$ 2-jet events (signal region).

detector effects smear the resolution. The mean of this distribution is $\sim 2000\ \mu m$, while the expected resolution on this quantity is about $\sim 150\ \mu m$.

The SVX-tagging algorithm begins by selecting displaced tracks in a cone of radius $\sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ about the axis of a given jet. The tracks are required to have impact parameter significance $\frac{d}{\sigma_d} \geq 2.5$, and $P_T \geq 0.5\ GeV/c$. The algorithm then attempts to find a displaced vertex with three or more tracks. If this fails, the track requirements are tightened to $\frac{d}{\sigma_d} \geq 4.0$, and $P_T \geq 1.5\ GeV/c$, and the algorithm attempts to find a displaced vertex with only two tracks. In each case, the criteria for a tag is that the transverse decay length be greater than three times its uncertainty: $L_{xy}/\sigma_{L_{xy}} > 3.0$. From Monte Carlo studies, we find that the algorithm tags $(42 \pm 5)\%$ of all top events passing the lepton plus jets selection criteria. The expected number of events from $t\bar{t}$ passing all cuts is shown in Table 3, for a luminosity of 100 pb^{-1}.

The background to b-tagged events comes primarily from inclusive QCD W events containing real heavy flavor. The processes which contribute are $p\bar{p} \rightarrow Wg$ $(g \rightarrow b\bar{b},\ c\bar{c})$ and $p\bar{p} \rightarrow Wc$. In addition, there are contributions to the background from fake tags (i.e., tags in events which contain no true displaced

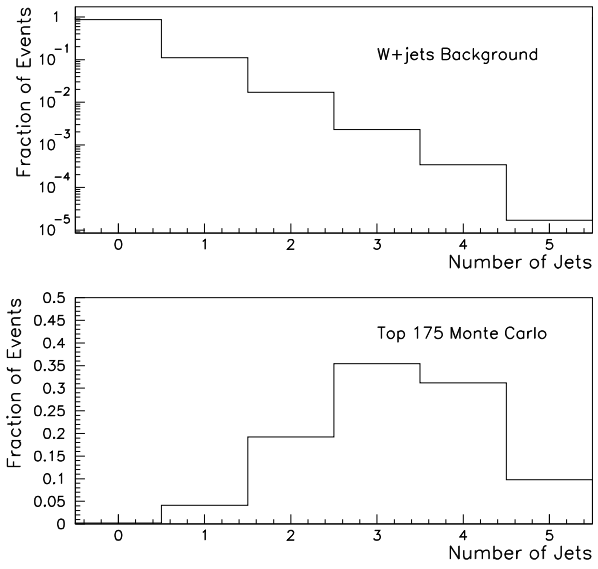


Fig. 7. Fraction of events vs. jet multiplicity observed in W plus jets data. The bottom plot shows the distribution expected from $t\bar{t}$ Monte Carlo with $M_{top} = 175\ GeV/c^2$.

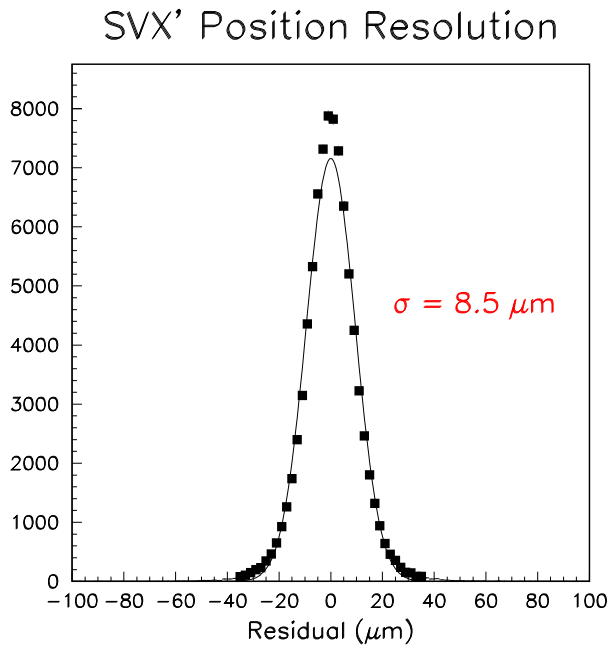


Fig. 8. Tracking residuals for all four layers of the SVX.

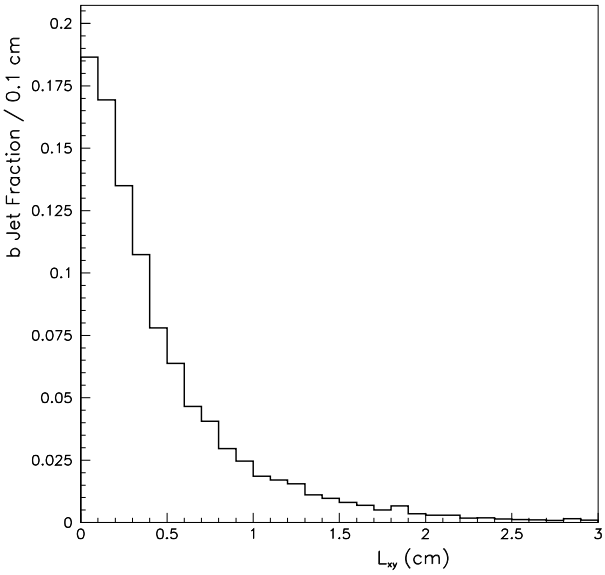


Fig. 9. The transverse decay length distribution for b hadrons, before detector resolution effects, from $t\bar{t}$ Monte Carlo events with $M_{top} = 175\ GeV/c^2$.

Top Mass (GeV/c^2)	$\sigma_{t\bar{t}}$ (pb)	Expected Events (above background)
160	8.2	30
170	5.8	24
180	4.2	15

Table 3. The predicted central value of the $t\bar{t}$ production cross section from Laenen *et al.*, and the number of SVX-tagged lepton plus jets $t\bar{t}$ events expected in 100 pb^{-1} after all cuts, as a function of top-quark mass.

vertices), and a small component from the following processes: $Z \to \tau\tau$, non-W, WW, WZ, and Drell-Yan.

Monte Carlo is used to determine the fraction of observed W + jet events containing $Wb\bar{b}$,$Wc\bar{c}$, and Wc. The tagging efficiency as measured in the data is then used to calculate the expected number of tags from these sources. The background from mis-tags is determined from a parameterization of fake tags observed in "generic" QCD jet data. The additional backgrounds are derived from a combination of data and Monte Carlo.

	Before Tag	Total Background	Observed Tags (Events)
W+0 Jet	88049		
W+1 Jet	9531	74.5±16.9	61 (61)
W+2 Jet	1469	29.7±7.9	43 (38)
W+ $\geq$ 3 Jet	296	9.9±2.8	40 (32)

Table 4. Observed SVX-tags in the W + jets sample, compared with expected background, vs. jet multiplicity bin. The signal region for $t\bar{t}$ is the W + $\geq$ 3 jet sample.

The total background is calculated separately for each W jet multiplicity bin. The results are shown in Table 4, along with the number of tags observed in the data. In the $W + 1$ jet bin, where we expect little contribution from top, there is good agreement between the observed tags and the calculated background. There is a small excess of tags in the $W + 2$ bin, consistent with the background estimate plus a small contribution from top. In the signal region of $W+ \geq 3$ jets, we observed 40 tags in 32 events over a total background of 9.9 ± 2.8 tags. A plot of observed tags and background, vs. W jet multiplicity, is shown in Fig. 10.

We have examined a number of features of the tagged events. The hypothesis is that the tagged events represent $t\bar{t}$ events where one W from a top decays to a lepton and a neutrino. In Fig. 11, we compare the transverse mass (computed using the x and y components of the lepton and $\not\!\!E_T$ in the event) observed for the tagged events vs. the distribution expected from a Monte Carlo for $M_{top} = 175\ GeV/c^2$. The agreement is good. We have the $\mathcal{H}$. In Fig. 12, we see good agreement between the $\mathcal{H}$ (described below) observed in the tagged events and the distribution expected from $t\bar{t}$ Monte Carlo plus background. In Fig. 13, we also

see good agreement between E_T of the tagged jets observed in the data and the distribution expected from $t\bar{t}$ Monte Carlo plus background. Finally, we can also examine observed tags vs. expected background in the Z + jets sample. There is very little contribution from top expected in this sample,[16] although there is a large reduction in statistics relative to the W + jets sample. The results are shown in Table 5. There is good agreement between tags and background in each jet multiplicity bin.

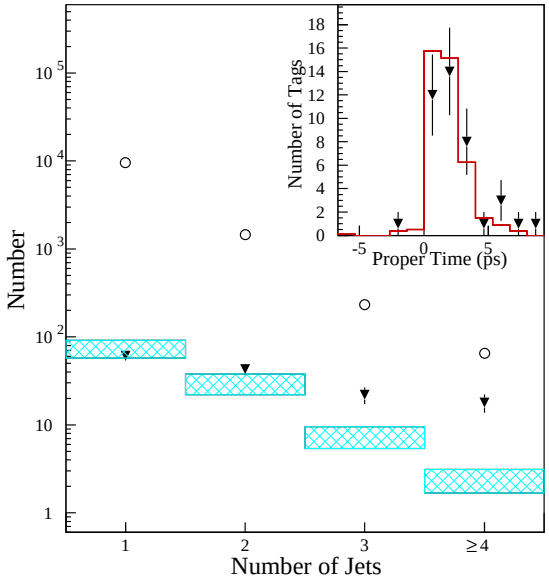


Fig. 10. The W + jets distribution observed in the data. The open circles are before SVX-tagging and the solid triangles are after SVX-tagging. The hatched boxes represent the tagging background estimate. The inset compares the $c\tau$ of the tagged jets in the W + $\geq$ 3 jet sample with the distribution observed in top Monte Carlo.

3.2 Soft Lepton Tagging

The second method used to tag b quarks in top events is to identify electrons and muons from b decay. The leptons are referred to as $soft$ due to the low momentum requirement of $P_T \geq 2\ GeV$. These additional leptons come primarily from the process $b \to \ell\nu X$ and the cascade process $b \to cX$, $c \to \ell\nu X$, but also from decays such as $W \to c\bar{s}$, $c \to \ell\nu X$ and $W \to \tau\nu$, $\tau \to \ell\nu\nu$.

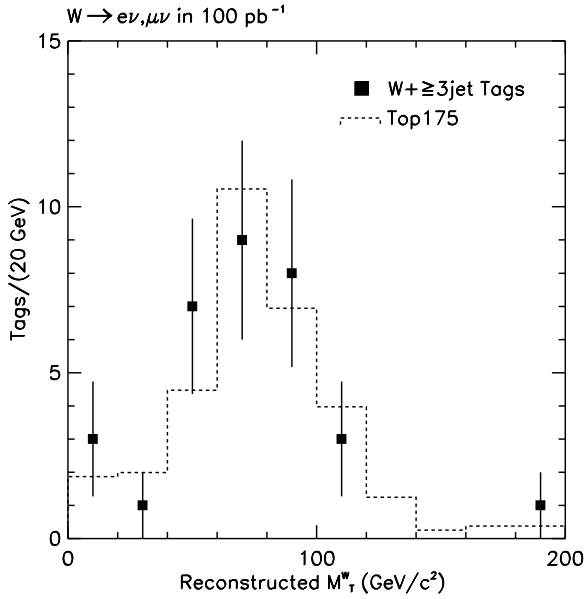


Fig. 11. Transverse mass of the SVX-tagged $W + \geq 3$ jet events (points) compared with shape expected from top Monte Carlo (solid histogram).

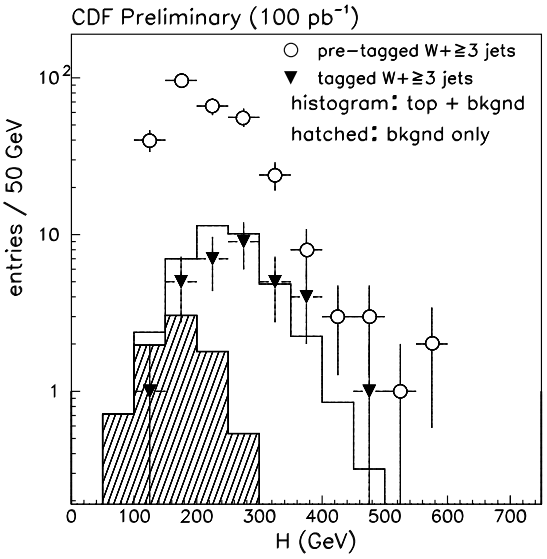


Fig. 12. $\mathcal{H}$ distributions for pretagged $W + \geq 3$ jet events (open circles), SVX-tagged $W + \geq 3$ jet events (solid triangles), compared to tags in top plus background Monte Carlo (solid histogram), and tags in background-only Monte Carlo (hatched histogram).

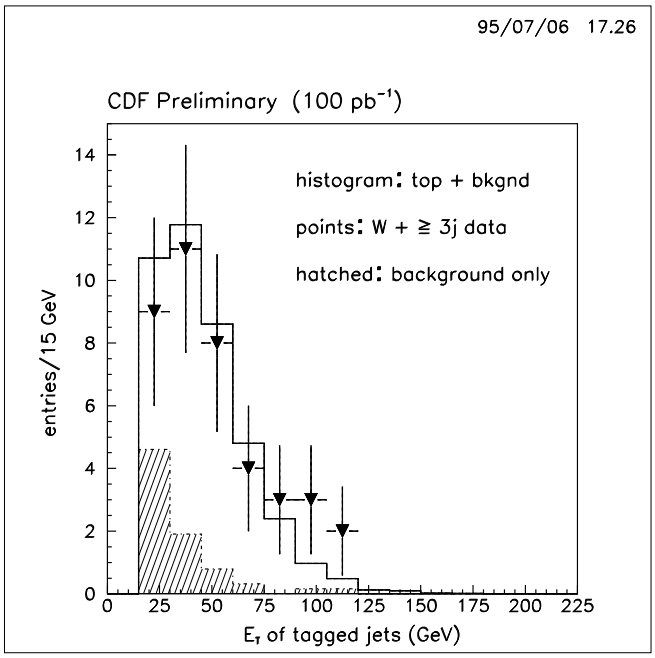


Fig. 13. SVX-tagged jet E_t of the tagged $W + \geq 3$ jet events (solid triangles), compared to tags in top plus background Monte Carlo (solid histogram), and tags in background only Monte Carlo (hatched histogram).

	Total Background	Observed Tags
$Z + 1$	8.4 ± 0.84	6
$Z + 2$	2.3 ± 0.23	3
$Z + \geq 3$	0.94 ± 0.09	1

Table 5. Observed SVX-tags in the $Z +$ jets sample, compared with expected background, vs. jet multiplicity bin.

The efficiency for identifying electrons is measured in the data using photon conversions $\gamma \rightarrow e^+e^-$. The efficiency for identifying muons is measured in the data using $J/\psi,\ Z \rightarrow \mu^+\mu^-$ events. The efficiency for finding an additional e or μ in a $t\bar{t}$ event passing the lepton plus jets selection criteria is (20±2)%. The primary background (about 75%) for soft lepton tags are hadrons misidentified as leptons, and electrons from unidentified photon conversions. A smaller background comes from processes which generate real heavy flavor, such as $Wb\bar{b}$ and $Wc\bar{c}$. The rate for these backgrounds is measured using generic jet samples and parameterized as a function of the P_T of the lepton candidates. Other much smaller backgrounds are calculated in the same manner as for the SVX analysis.

As before, the total background is calculated separately for each W jet multiplicity bin. The results are shown in Table 6, along with the number of SLT-tags observed in the data. There is good agreement between the observed tags and the calculated background in the $W+1$ and $W+2$ jet bins. In the signal region of $W+ \geq 3$ jets, we observe 40 SLT-tags in 36 events over a total background of 23.8 ± 3.6 tags. A plot of observed tags and background vs. W jet multiplicity is shown in Fig. 14.

	Before Tag	Total Background	Observed Tags
W+0 Jet	88049		
W+1 Jet	9531	250±38	232
W+2 Jet	1469	71±11	84
W+≥ 3 Jet	296	23.8±3.6	40

Table 6. Observed SLT-tags in the $W+$ jets sample, compared with expected background, vs. jet multiplicity bin. The signal region for $t\bar{t}$ is the $W + \geq 3$ jet sample.

4 Measurement of the $t\bar{t}$ Production Cross Section

We calculate the cross section implied by the excess of events observed in the three counting experiments: W + jets with an SVX B-tag, W + jets with an SLT

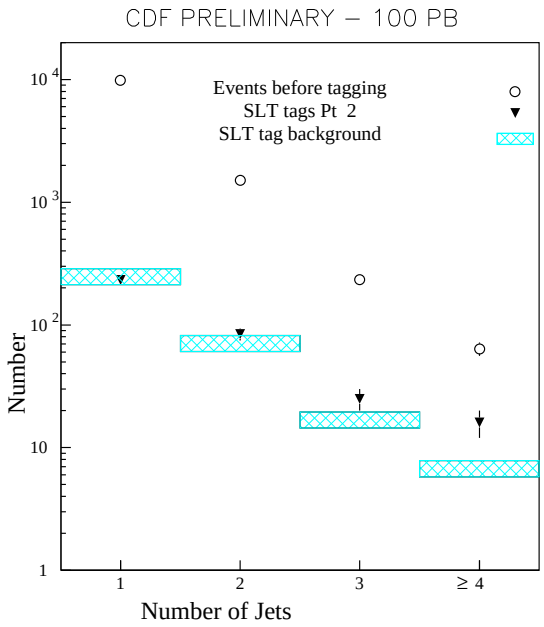


Fig. 14. The W + jets distribution observed in the data. The open circles are before SLT tagging and the solid triangles are after SLT tagging. The hatched boxes represent the tagging background estimate.

B-tag, and dileptons. The cross section is calculated using the relation:

$$\sigma = \frac{n - b}{\epsilon^{tot} \cdot \int \mathcal{L} dt}$$

where n is the observed number of events, b is the estimated background, and ϵ^{tot} is the total acceptance (including branching ratios) for a given channel. $\int \mathcal{L} dt$ is the integrated luminosity, and for the result stated below it is 67 pb^{-1}. The cross section is determined by maximizing the likelihood expression:

$$L = G(\int \mathcal{L} dt, \overline{\int \mathcal{L} dt}, \sigma_{\int \mathcal{L} dt}) \cdot L^{SVX} \cdot L^{SLT} \cdot L^{DIL}$$

and

$$L^i = G(\epsilon, \bar{\epsilon}, \sigma_\epsilon) \cdot G(b, \bar{b}, \sigma_b) \cdot P(n, \epsilon \cdot \int \mathcal{L} dt \cdot \sigma + b),$$

where $i = SVX$, SLT, and DIL, $G\ (p, \bar{p}, \sigma_p)$ is a Gaussian for parameter p, with mean $\bar{p}$ and width σ_p, and $P\ (n, m)$ is a Poisson with mean m and number of observed events n. The quantities ϵ and b represent the total efficiency and background for a given channel i. In practice, the above expressions are separated into pieces that either are or are not common between Run 1A and Run 1B.

Item	SVX	SLT	Dilepton
N_{events}^{tagged}	21	22	7
ϵ_{total} (in %)	3.4 ± 0.9	1.7 ± 0.3	0.78 ± 0.08
Background	5.5 ± 1.8	14.7 ± 2.2	1.3 ± 0.3
Luminosity	$67 \pm 5\ pb^{-1}$	$67 \pm 5\ pb^{-1}$	$67 \pm 5\ pb^{-1}$

Table 7. Parameters used in the calculation of the combined SVX/SLT/dilepton $t\bar{t}$ production cross section.

In addition, the acceptance for the SVX and SLT channels, before tagging, is taken as 100% correlated. A subset of the parameters used in the calculation are listed in Table 7. Note that the background in the SVX and SLT channels have been corrected for the top content in the sample before tagging using an iterative technique.[17] Combining the three channels results in $\sigma(t\bar{t}) = 7.6\ ^{+2.4}_{-2.0}\ pb$. In Fig. 15, we compare this result, along with the CDF measurement of the top-quark mass (described below), vs. the theoretical expectation.

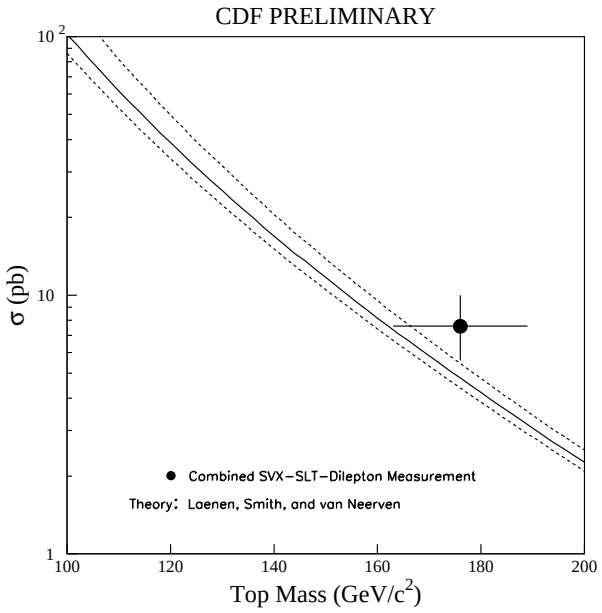


Fig. 15. Comparison of the CDF measurements of the $t\bar{t}$ production cross section and the top-quark mass, vs. the expectation of theory, from Laenen *et al.*

5 Determination of V_{tb}

In the Standard Model, a 176 GeV/c^2 top quark decays almost exclusively to Wb, because $V_{tb} \sim 1$ and there is no kinematic suppression of this decay. In fact, the statement that $V_{tb} \sim 1$ assumes that the CKM matrix is 3×3 and unitary, and this is something we would like to test. The procedure used to measure V_{tb} is relatively straightforward. We measure the ratio:

$$b = \frac{BF(t \to Wb)}{BF(t \to Wq)} = \frac{V_{tb}^2}{V_{td}^2 + V_{ts}^2 + V_{tb}^2},$$

where the relative branching fraction of top to Wb is measured by examining the distribution of dilepton events with zero, one, and two b-tagged jets, and by examining the distribution of lepton plus ≥ 3 jets events with one and two b-tagged jets. This technique takes advantage of the fact that values of BF $(t \to Wb)$ significantly smaller than 1.0 would have a noticeable effect on the relative distribution of events in each category. A maximum likelihood estimator is used to determine the best fit ratio, and in 100 pb^{-1}, we find

$$b = \frac{BF(t \to Wb)}{BF(t \to Wq)} = 0.94 \ \pm 0.27(stat) \ \pm 0.13(syst)$$

and at 95% CL, we obtain $b > 0.34$.

We then use the above relation to determine V_{tb}. Assuming unitarity, we obtain $|V_{tb}| = 0.97 \pm 0.15 \pm 0.07$. Additionally, we can relax the assumption of three-generation unitarity. Assuming that there is no coupling between the first two and a possible fourth generation, then $|V_{td}| = 0.004 - 0.015$ and $|V_{ts}| = 0.030 - 0.048$ (Ref. 18). To extract a 95% CL limit on $|V_{tb}|$, we assume that $|V_{td}|$ and $|V_{ts}|$ take on their smallest allowed values, and use our 95% CL limit measurement of b. We find that at 95% CL, $|V_{tb}| > 0.022$ (95% CL).

6 Top-Mass Reconstruction

We measure the top-quark mass using the $W+ \geq 3$ jet events which also contain at least one additional jet with $E_T > 8$ GeV, $|\eta| < 2.4$. This sample will be referred to as the $W+ \geq 4$ jet sample. This sample can then be fit to the hypothesis:

$$p\bar{p} \rightarrow t_1 + t_2 + X$$
$$t_1 \rightarrow W_1 + b_1$$
$$t_2 \rightarrow W_2 + b_2$$
$$W_1 \rightarrow l + \nu$$
$$W_2 \rightarrow j_1 + j_2,$$

assuming the four highest E_T jets correspond to the partons b_1, b_2, j_1, and j_2. When calculating masses, all jet energies are corrected for detector effects, for contributions from the underlying event, and for energy falling outside of the fixed cone size of 0.4. Finally, there are specific corrections for jets tagged as b's. When tagging information is required, both SLT and SVX-tagged events are used. All possible assignments of jets to partons are tried, with the restriction that if a jet is tagged, it is required to be one of the b jets. The $\not\!\!E_T$ is assumed to represent the transverse components of the neutrino, and the constraint that the lepton and $\not\!\!E_T$ reconstruct to the W mass yields a two-fold ambiguity for the P_z of the neutrino. The two jets selected as the decay products of the hadronic W are also constrained to the W mass. Finally, both top masses (representing the hadronic and leptonic W decays) are required to be the same. Within estimated uncertainties, the jet energies are allowed to vary in order to satisfy the constraints. Each particular assignment of jets to partons yields a top mass and a χ^2, and the solution with the lowest χ^2 is chosen for each event. This χ^2 is required to be less than ten. The resulting distribution of top masses observed in the data is then compared to the expectation for top plus background for various top masses. The background shape is from Monte Carlo W + jet events satisfying the same selection criteria as the data, and the normalization is fixed using a method described in our previous publication.[7]

In 67 pb^{-1}, there are 99 W + ≥ 4 jet events using the criteria described above. There are 19 events which contained either an SVX or SLT tag, of which $6.3^{+2.1}_{-1.7}$ are expected to be background. Figure 16 shows the data along with the best fit expectation for signal plus background. The inset shows the negative ln (likelihood) returned by the fit at each top mass. The minimum of the negative ln (likelihood) yields the best fit, and a change in the ln (likelihood) of 0.5 gives the statistical uncertainty. The systematic uncertainties are extrapolated from our earlier publication,[7] the dominant sources being the effects of gluon radiation on the determination of parton energies, and the jet energy scale. The final result for the mass is $M_{top} = 176 \pm 8(stat) \pm 10(syst)\ GeV/c^2$. In Fig. 17, we show

the reconstructed mass distribution for the sample of $W + \geq 4$ jet events before b-tagging. This distribution also shows an excess of events over the background shape in the region of reconstructed mass $\sim 175\ GeV/c^2$.

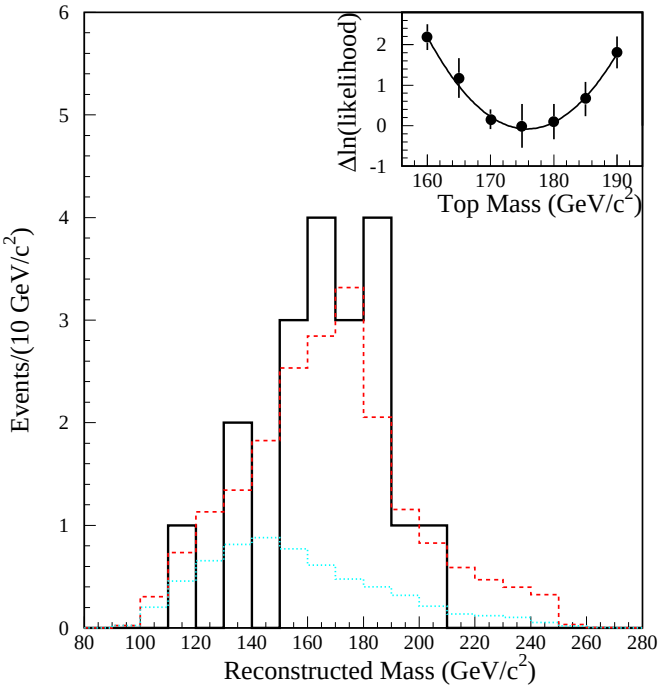


Fig. 16. Reconstructed mass distribution for events containing at least one SVX or SLT b-tagged jet. The dashed histograms represent the fitted distributions from background and the sum of background and signal. The inset plot shows the likelihood for similar fits to different top-mass hypotheses.

We have also looked for evidence of the decay of the hadronic W in this sample. In the first method, we use the same mass fitting procedure described above, except that we release the constraint on the hadronic W decay. We can then examine the mass distribution of the jets the fitter chooses (for the minimum χ^2 solution) as the decay products of the hadronic W. In Fig. 18, we can see a clear excess of events in the region of the W mass.

The second method takes advantage of the fact that when both b quarks are tagged, there is a unique jet-jet mass combination for the W decay. Both SVX and SLT tags are used, although the criteria for an SVX tag are loosened to improve acceptance. In 100 pb^{-1}, ten events satisfy these selection criteria. The mass distribution for the remaining two non-b jets is shown in Fig. 19. A clear peak is seen at the W mass.

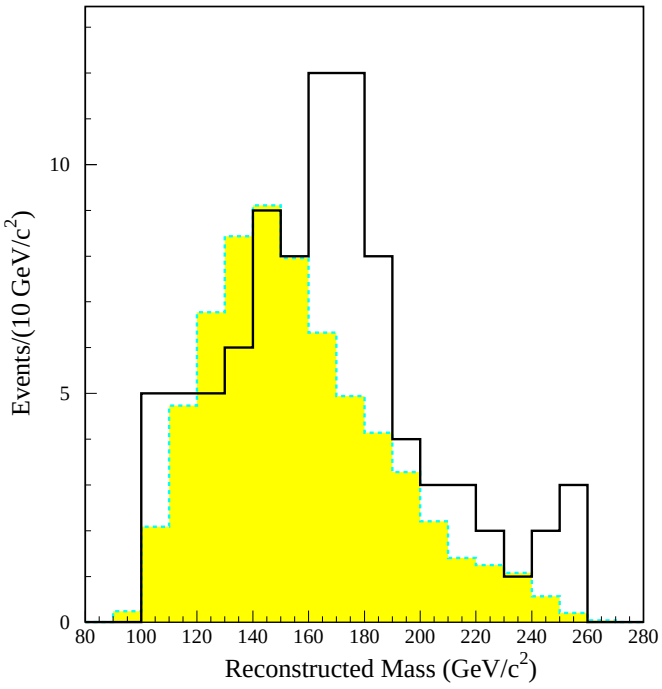


Fig. 17. Reconstructed mass distribution for the sample of $W + 4$ jet events before b-tagging. The shaded histogram is that expected for background with the normalization taken from a fit made with a background constraint.

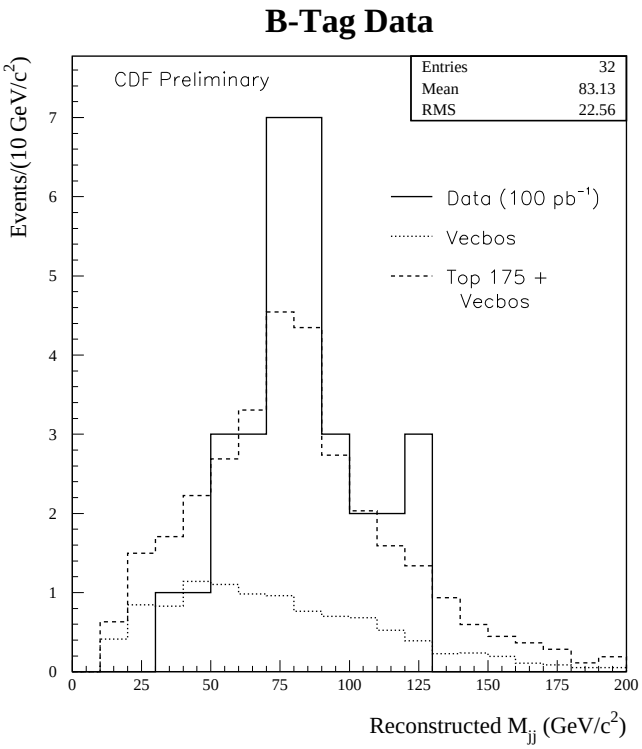


Fig. 18. Plot of the invariant mass of the two jets assigned as the decay jets of the hadronic W, when the W mass constraint is removed from the fit. Note the excess of events in the region of the W mass.

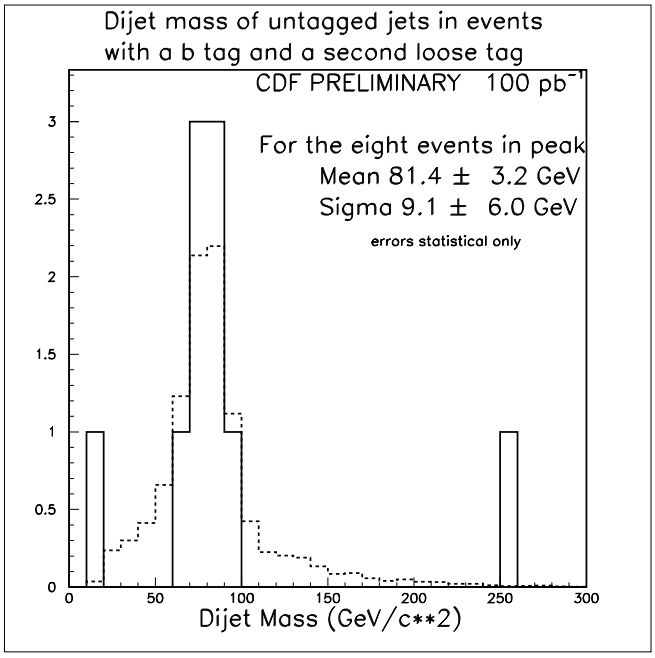


Fig. 19. Dijet mass distribution of untagged jets in events with two identified b quarks. The solid curve is for the ten events observed in the data, and the dashed curve is for a Monte Carlo calculation.

7 The $t\bar{t}$ Invariant Mass

As a result of the constrained fit in the lepton plus jets sample, one obtains the four-momenta of the t and the $\bar{t}$ quarks. Using these quantities, we can then calculate the $t\bar{t}$ invariant mass for each event. This quantity is sensitive to non-Standard Model top quark production mechanisms. A number of authors[19,20] have pointed out the possibility of heavy resonances which could have sizable decay branching ratios to $t\bar{t}$. In determining this quantity, an improvement in resolution of a factor of ≈ 2 can be obtained by constraining the top mass to the value measured in the previous section. The resulting distribution observed in 100 pb^{-1} is shown in Fig. 20. The data shows good agreement with the shape expected from Standard Model $t\bar{t}$ plus background. In order to gain acceptance, we can perform the same analysis using the pretagged data. This is shown in Fig. 21, and once again, there is good agreement between data and Standard Model $t\bar{t}$ plus background.

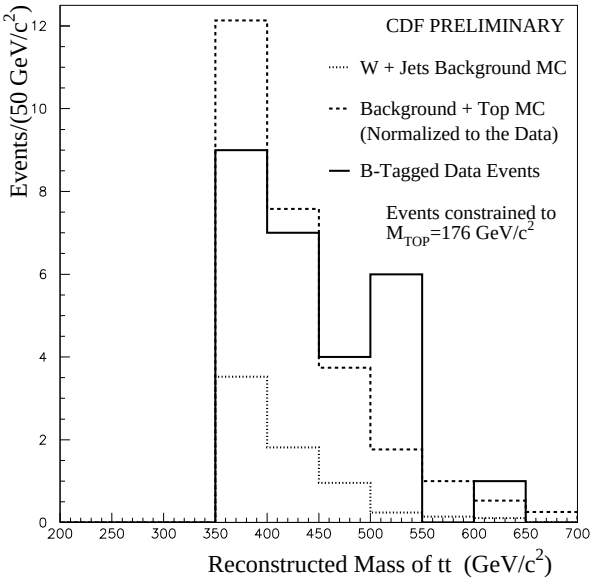


Fig. 20. $t\bar{t}$ invariant mass after b-tagging for the data (solid histogram), W + jets Monte Carlo normalized to the expected background rate (dotted histogram), and top plus W + jets Monte Carlo (dashed histogram).

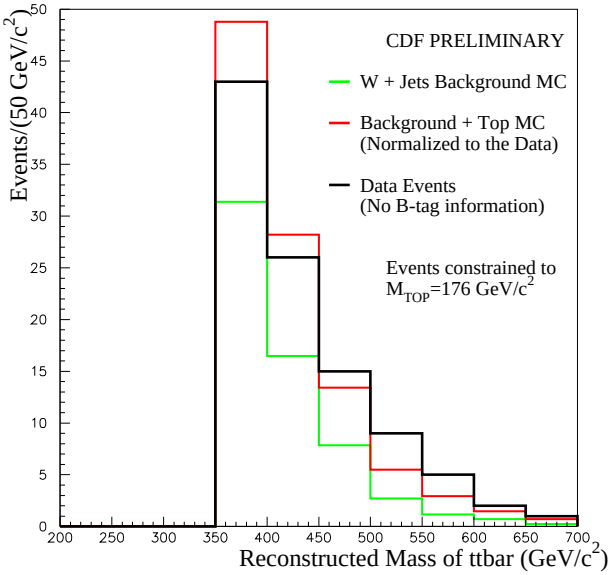


Fig. 21. $t\bar{t}$ invariant mass before b-tagging for the data (solid histogram), W + jets Monte Carlo normalized to the expected background rate (dotted histogram), and top plus W + jets Monte Carlo (dashed histogram).

8 The $\mathcal{H}$ Analysis

Due to the large top-quark mass, $t\bar{t}$ events are produced with large $\sqrt{s}$ compared to background processes. A simple variable which scales with $\sqrt{s}$ is $\mathcal{H}$, which is defined as

$$\mathcal{H} = E_T(lepton) + \not\!\!E_T + \sum_{jets} E_T(jets),$$

where the sum over jets includes all jets with $E_T \geq 8\ GeV$ and $|\eta| \leq 2.4$. This analysis is performed on the pretagged mass sample. A feature of this analysis is that it is sensitive to non-b tagged top events, as well as providing an independent check of M_{top} and $\sigma(t\bar{t})$.

In Fig. 22, we show the separation obtained using $\mathcal{H}$, for $t\bar{t}$ Monte Carlo vs. $W + 4$ jet background. Also shown in this figure is the distribution of $\mathcal{H}$ observed in the full data sample, as well as for the subset of events containing a b tag. These events cluster near large $\mathcal{H}$ as expected for the $t\bar{t}$ component of the distribution. We perform a two-component binned maximum-likelihood fit of the $\mathcal{H}$ distribution of this sample, to a sum of distributions expected for $t\bar{t}$ and background. As in the mass reconstruction analysis, this is done for several values of top mass. The resulting negative ln(likelihood) vs. top mass is plotted in Fig. 23. The top mass we obtain using 99 events in 67 pb^{-1} is $M_{top} = 180 \pm 12(stat)^{+19}_{-15}(syst)\ GeV/c^2$, which is in good agreement with the mass determination from the lepton plus jets sample.

9 The All-Hadronic Channel

The selection for this channel is based on the expected decay topology $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'\ q\bar{q}'\ \ b\bar{b}$. Each event is required to pass the following criteria:

- $N_{jet} \geq 6$, where N_{jet} is the number of jets with $E_T > 15\ GeV$ and $|\eta| < 2.0$. The jets must be separated in η, ϕ space by $\Delta R_{min} \geq 0.5$.
- $\sum E_T \geq 150\ GeV$, where $\sum E_T$ is the transverse energy sum over the N_{jet} jets.
- $\sum E_T/\sqrt{s} \geq 0.75$.
- $A \geq -0.003 \times \sum E_T + 0.45$, where A is aplanarity calculated in the center-of-mass system of the N_{jet} jets.
- Require at least one SVX-tagged jet.

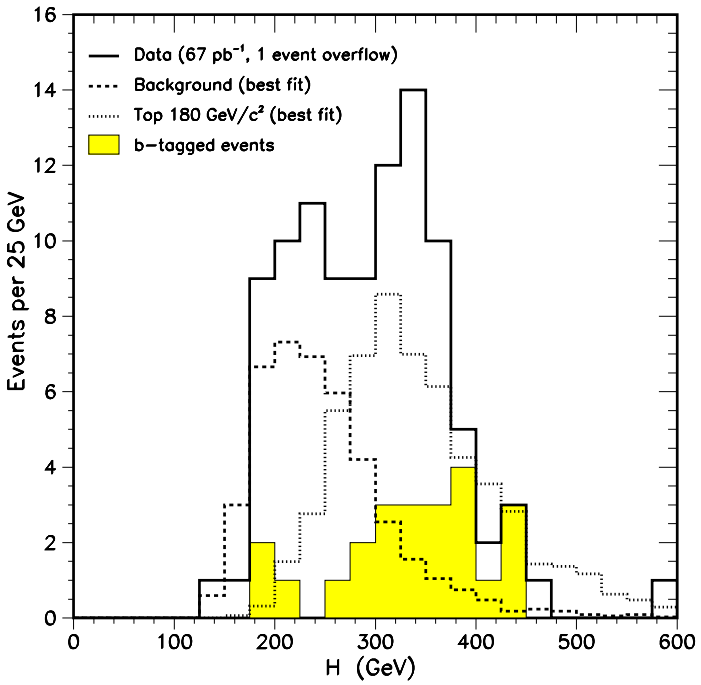


Fig. 22. $\mathcal{H}$ distribution of CDF data (solid line), W + 4 jet background (dashed line), and $t\bar{t}$ Monte Carlo with $M_{top} = 180\ GeV/c^2$ (dotted line). The shaded distribution shows the events in the data containing a b tag.

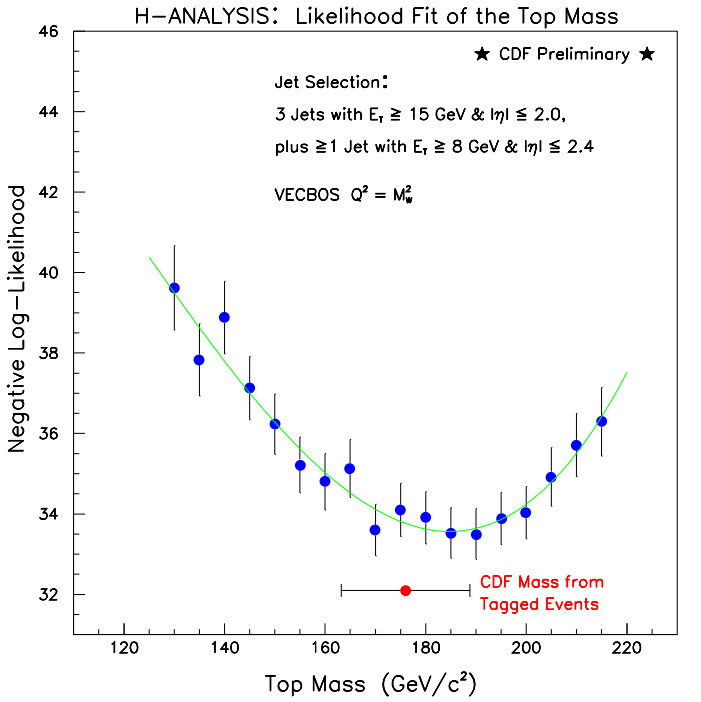


Fig. 23. Resulting negative ln(likelihood) vs. top mass for fits of $\mathcal{H}$ distribution of CDF data to a sum of distributions expected for $t\bar{t}$ and background. Also shown is the CDF mass result from the $W+ \geq 4$ jet sample described earlier.

The efficiency for this selection (including all decay modes) is $\epsilon_{HAD} = (8.6 \pm 0.4(stat))\%$ for $M_{top} = 175\ GeV/c^2$. The data comes from a multijet trigger, and the results below are based on an integrated luminosity of $\mathcal{L} = 81\ pb^{-1}$.

These events are then fit to the $t\bar{t}$ hypothesis using a method similar to that described above. Each of the six highest E_T jets is assigned as one of the six decay partons of the t and $\bar{t}$. Momentum conservation is required at the two top decay vertices, and the masses of the resulting t and $\bar{t}$ from the fit are constrained to be equal. The two W mass constraints are not imposed so as to reduce the total number of combinations. The solution with the best χ^2 is chosen.

We have used two methods to determine the shape of the background mass spectrum. The first method uses events which pass all of the selection criteria, but which contain no SVX-tagged jet. This sample is expected to contain less than 5% $t\bar{t}$ signal. The background distribution is shown in Fig. 24. In the second method, we parameterize the tag rate (as observed in an independent sample) and apply it event by event to the pretagged multijet sample. This procedure should include any biases on the background mass spectrum due to the tagging itself. The resulting shapes of the background from the two methods are in excellent agreement and are shown in Fig. 24. We have found that a Landau-like distribution adequately models the shape of the background and use this in our fit. The mass distribution observed in the data is then fit to a sum of the Landau describing the background, and a Gaussian distribution describing the $t\bar{t}$ signal. The normalization of the background and top components are left free in the fit. The result of the fit is shown in Fig. 25. The mean of the Gaussian yields a top mass in good agreement with the result of lepton plus jets analysis. From the fit result for the number of top events, the acceptance quoted above, and the integrated luminosity of the multijet sample, we calculate the $t\bar{t}$ production cross section to be:

$$\sigma_{t\bar{t}} = 9.6 \pm 3.5 \text{ (stat. only) } pb,$$

also in good agreement with the cross-section result from the SVX, SLT, and dilepton channels quoted earlier.

10 The Future for Top at CDF

The current limiting factor to higher luminosities at the Tevatron is the total number of antiprotons in the ring. Fermilab is currently building a new Main

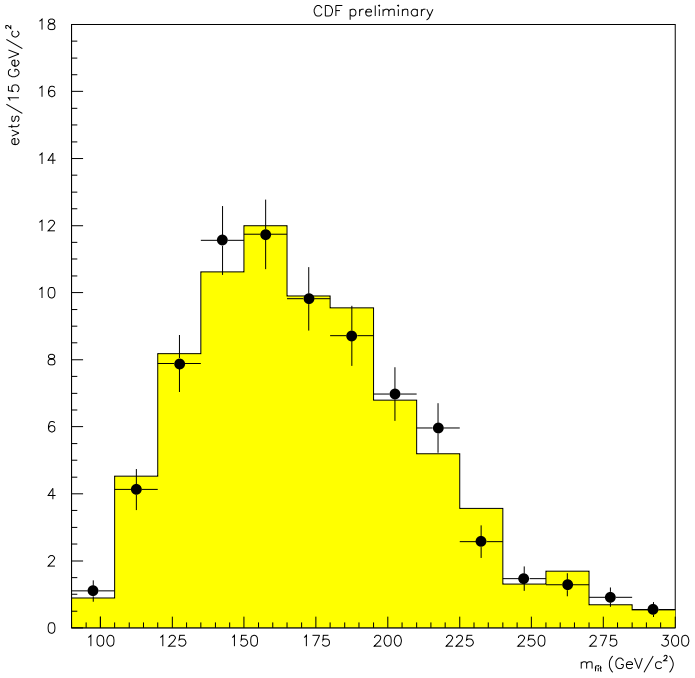


Fig. 24. Background mass distribution for the all-hadronic analysis. The points are from the no b-tag sample, and the shaded histogram is from the tag rate parameterization.

Injector accelerator that will lead to antiproton production rates about a factor of two above what is currently possible. It will also improve the transfer efficiency for antiprotons. This should lead to instantaneous luminosities of roughly a factor of four over what can currently be attained. In addition, a proposed device called the Recycler Ring may allow more efficient production and use of antiprotons, and could yield another factor of two to three in instantaneous luminosity. Finally, the beam energy will be increased from the current 900 GeV to 1000 GeV in Run II. Some parameters of the current and future Tevatron are listed in Table 8.

Run	Energy (GeV)	Instant Lum. ($cm^{-2}sec^{-1}$)	Integ. Lum. (pb^{-1}/week)
1B (Present)	900	2×10^{31}	4
II (MI)	1000	8×10^{31}	17
II (MI+Recycler)	1000	20×10^{31}	41

Table 8. Performance parameters of the Tevatron in the 1994–95 run compared with expected parameters for Run II.

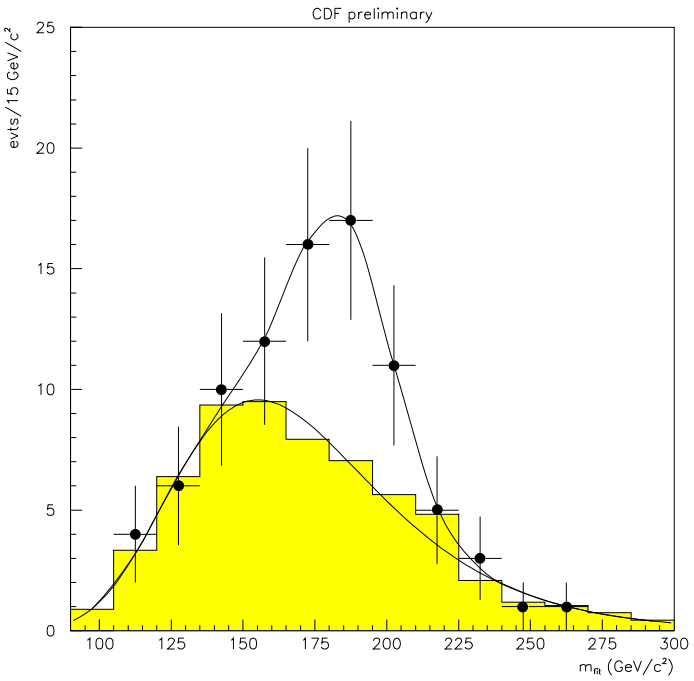


Fig. 25. Fitted top-mass distribution in the all-hadronic channel. The points are the b-tagged signal sample. The shaded histogram is the background normalized to the outcome of the fit. The lower curve is a Landau distribution representing the background shape, and the upper curve is a sum of this Landau and a Gaussian distribution representing the $t\bar{t}$ signal.

The CDF detector has been used to study proton-antiproton collisions at Fermilab since 1985. During this time, the detector has been upgraded to increase physics capability and to keep pace with changes to the Tevatron. The improvements to the Tevatron noted above necessitate replacement or modification of several detector systems. These detector systems include:

- Silicon Vertex Detector (SVX II)
  - The new detector will have five layers and will be double-sided to provide both $r-\phi$ and $r-z$ readout, allowing three-dimensional vertex reconstruction. In addition, the detector will be ~ 90 cm long, allowing much more complete coverage of the $p\bar{p}$ interaction region.
- Intermediate Fiber Tracker (IFT)
  - This will be a scintillating fiber tracker in the region $r \sim 16-27\ cm$. The IFT plus SVX II tracking combination should allow b-tagging out to $|\eta| = 2$, as well as improving electron identification in the region covered by the new plug calorimeter.
- Plug Calorimeter Upgrade
  - This is a scintillating-tile fiber calorimeter with a shower-max detector. It should allow for greatly improved electron identification at large rapidity.
- Muon Detection System
  - The muon detectors in the central region will have additional chambers installed to allow more complete ϕ and η coverage. In addition, the more compact design of the new plug calorimeter will allow the forward muon system to move closer to the interaction region, increasing our muon acceptance.

The above changes to the accelerator and the detector will impact the top analysis in a number of significant ways. First, at $\sqrt{s}$ =2 TeV, the $t\bar{t}$ production cross section increases by $\sim 40\%$ relative to $\sqrt{s}$ =1.8 TeV. The new plug calorimeter could increase the acceptance for $W \rightarrow e\nu$ from top by about 25%, assuming one can maintain a signal-to-noise ratio similar to what is currently achieved in the central calorimeter. Finally, one should note that the current Silicon Vertex Detector used in Run 1B only covers about 60% of the $p\bar{p}$ interaction region, while the SVX II detector will cover almost all of the $p\bar{p}$ interaction region. This will

almost double the b-tagging efficiency to about 80% per top event (compared to 42% at present).

Taking all of these improvements into account, the expected yield of $t\bar{t}$ events in the lepton + 4 jets + b-tag mode (i.e., the tagged mass sample) should be $\sim$ 600 per fb^{-1} (compared to $\sim$ 200 per fb^{-1} in the current run). With yields such as these, the expected precision on a number of top measurements can be estimated. For example, with 2 fb^{-1} of data, we expect to measure the top mass to better than $\sim$ 4 GeV/c^2 (including systematic and statistical uncertainties), and the production cross section to better than $\sim$ 7%.

11 Summary

The evidence for the top quark that CDF presented in April of 1994 was confirmed in all aspects by the results of the CDF and DØ Collaborations in March of 1995. We have observed top in a number of different decay channels. We have used the lepton plus jets decay channel to measure the top-quark mass:

$$M_{top} = 176 \pm 8 \pm 10 \ GeV/c^2.$$

In this same sample, we can observe a very clean peak of the process $W \rightarrow jj$, providing a sample of $WbWb$ events as expected for $t\bar{t}$. We have used a kinematic analysis of the lepton plus jets sample to derive a top-mass measurement in agreement with the above result. The $t\bar{t}$ production cross section is measured from the dilepton and lepton plus jets decay channels to be:

$$\sigma(t\bar{t}) = 7.6 \ ^{+2.4}_{-2.0} \ pb.$$

We have also observed top in the difficult all-hadronic decay channel, and derive mass and cross-section measurements consistent with the above two results. We are using the invariant mass distribution of the $t\bar{t}$ sample as a probe for new physics. Finally, we have performed the first direct measurement of the CKM matrix element V_{tb}.

We expect to record another $\sim$ 30 pb^{-1} of data in the fall (winter) of 1995(6). We plan to continue to improve our top analysis with new data and better understanding of systematic uncertainties. The Fermilab accelerator complex and the CDF detector are currently planning for Run II, which promises data samples at least a factor of ten larger than we now have. A important focus of Run II

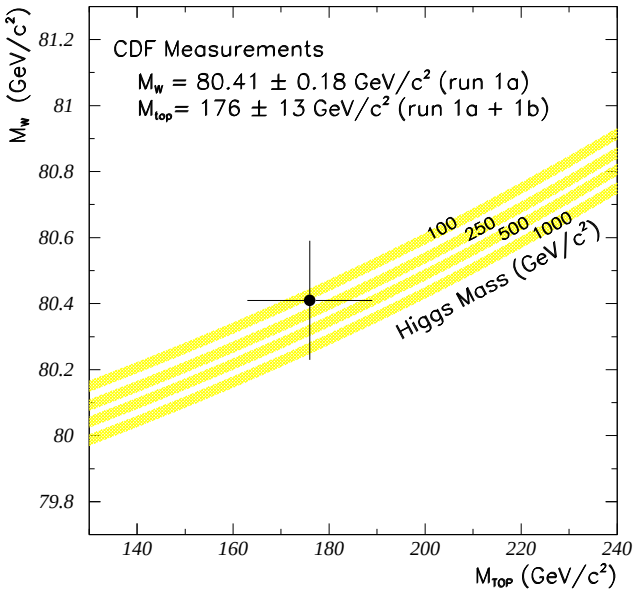


Fig. 26. Correlation between the W boson mass and the top-quark mass, for several values of the Higgs boson mass, along with current CDF measurements.

will be to probe electroweak parameter space and hopefully constrain the allowed mass range of the only unobserved Standard Model particle, the Higgs. Figure 26 shows the correlation between the W boson mass and the top-quark mass for several values of the Higgs boson mass, along with current CDF measurements.

References

[1] W. Bartel *et al.*, Phys. Lett. B **146**, 437 (1984).

[2] The LEP Collaborations ALEPH, DELPHI, L3, OPAL, and The LEP Electroweak Working Group, CERN-PPE/94-187.

[3] A. Bean *et al.*, Phys. Rev. D **35**, 3533 (1987).

[4] J. Ellis, G. L. Fogli, and E. Lisi, CERN-TH/95-202.

[5] F. Abe *et al.*, Phys. Rev. D **45**, 3921 (1991).

[6] S. Abachi *et al.*, Phys. Rev. Lett. **72**, 2138 (1994). See also S. Abachi *et al.*, Phys. Rev. Lett. **74** (1995).

[7] F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994); F. Abe *et al.*, Phys. Rev. Lett. **73**, 225 (1994).

[8] F. Abe *et al.*, Phys. Rev. D **51** (1995).

[9] F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[10] S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[11] E. Laenen, J. Smith, and W. L. van Neerven, Phys. Lett. B **321**, 254 (1994).

[12] R. K. Ellis, in *Proceedings of the 1994 International Conference on High Energy Physics*, edited by Peter J. Bussey and Ian G. Knowles (Institute of Physics, London, 1995), p. 1203.

[13] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **271**, 387 (1988).

[14] P. Azzi *et al.*, FERMILAB-CONF-94/205-E. Our previous silicon vertex detector is described in D. Amidei *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **350**, 73 (1994).

[15] In the CDF coordinate system, θ is the polar angle with respect to the proton beam direction. The pseudorapidity, η, is defined as $-\ln\tan\frac{\theta}{2}$. The transverse momentum of a particle is $P_T = P\sin\theta$. If the magnitude of this vector is obtained using the calorimeter energy rather than the spectrometer momentum, it becomes the transverse energy (E_T). The difference between the vector sum of all the transverse energies in an event and zero is the missing transverse energy ($\not\!\!E_T$).

[16] There is a small contribution of $t\bar{t}$ to the Z sample, for $t\bar{t}$ decays to dileptons e^+e^- and $\mu^+\mu^-$, where the dilepton invariant mass falls between 75 GeV/c^2 and 105 GeV/c^2. For 100 pb^{-1}, the contribution should be less than one event.

[17] The backgrounds in these channels were both calculated under the assumption that the entire sample before tagging contains no $t\bar{t}$ component. The resulting background calculation is thus an overestimate.

[18] Review of Particle Properties, Phys. Rev. D **50**, 1315 (1994).

[19] C. T. Hill and S. J. Parke, Phys. Rev. D **49**, 4454 (1994). C. T. Hill, Phys. Lett. B **345**, 483 (1995).

[20] E. Eichten and K. Lane, Phys. Lett. B **327**, 129 (1994).

OBSERVATION OF THE TOP QUARK WITH THE DØ DETECTOR

Nicholas J. Hadley
Physics Department
The University of Maryland, College Park, MD 20742

Representing the DØ Collaboration

ABSTRACT

The DØ Collaboration reports on the observation of the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV at the Fermilab Tevatron. We measure the top quark mass to be $199^{+19}_{-21}(stat.)^{+14}_{-21}(syst.)$ GeV/c^2 and its production cross section to be 6.4 ± 2.2 pb. Our result is based on approximately 50 pb^{-1} of data. We observe 17 events with an expected background of 3.8 ± 0.6 events. The probability of an upward fluctuation of the background to produce the observed signal is 2×10^{-6} (equivalent to 4.6 standard deviations). The kinematic properties of the events are consistent with top quark decay, and the distribution of events across the seven decay channels is consistent with the Standard Model top quark branching fractions. We describe the analysis that led to the observation of the top quark as well as the properties of the top quark events.

©1995 by Nicholas J. Hadley.

On Thursday, March 2, 1995, in two seminars given at Fermilab, the DØ and CDF Collaborations announced the discovery of the top quark.[1,2] This discovery was the culmination of nearly two decades of intense search by a large number of different experiments located at accelerators throughout the world.

The DØ experiment measures a mass of $199^{+19}_{-21}(stat.)^{+14}_{-21}(syst.)$ GeV/c^2 for the top quark with a production cross section of 6.4 ± 2.2 pb for that mass. (The CDF experiment measures a mass of 176 ± 13 GeV/c^2 and a cross section of $6.8^{+3.6}_{-2.4}$ pb.) The DØ result is based on approximately 50 pb^{-1} of data, about four times the previous sample.[3,4] We observe 17 events with an expected background of 3.8 ± 0.6 events. The probability of an upward fluctuation of the background to produce the observed signal is 2×10^{-6} (equivalent to 4.6 standard deviations). The kinematic properties of the events are consistent with top quark decay. The distribution of events across the seven decay channels we study is consistent with the Standard Model top quark branching fractions. We will describe the analysis that led to the observation of the top quark in detail, starting with a brief summary of past searches for the top quark with DØ. We will also show that our top data contains W's that decay hadronically.

1 Introduction

At Tevatron energies, top quarks are primarily produced in pairs. In what follows, we assume that top quarks decay into a W boson and a b quark with 100% branching fraction. The decay modes of the top are then characterized by the decays of the two *W's* in each event. Events where both *W's* decay to leptons (e or μ) are called dilepton events, denoted ee, $e\mu$, and $\mu\mu$ events. Events where one W decays to an e or a μ and the other decays to jets are called lepton + jets events. Decays to tau leptons are considered only as sources of jets, electrons, or muons. Events where both *W's* decay to jets have large backgrounds due to QCD multijet events and were not used in the discovery analysis.

In January 1994, the DØ Collaboration published[3] an upper limit on the top quark pair production cross section which can be translated into a 95% confidence level (CL) lower limit on the top quark mass of 131 GeV/c^2. This paper used 13.5 pb^{-1} of data from the 1992–1993 run and was based on the analysis of the number of events seen in the ee, $e\mu$, e + jets, and μ + jets decay channels.

In April 1994, the CDF Collaboration submitted papers claiming evidence for the top quark with a mass of 174 ± 16 GeV/c^2 and a cross section of $13.9^{+6.1}_{-4.8}$ pb. The statistical significance of the signal was about 2.8 standard deviations.[5]

In a paper[4] submitted to *Physical Review Letters* in November 1994, we re-optimized our analysis for higher masses, based on our previous result (top quark mass $>$ 131 GeV/c^2), added more decay channels, and improved our understanding of the backgrounds in the various channels. This analysis provided a background-subtracted estimate of the top quark production cross section, based on the same data set as the January 1994 paper, but using the information from all decay channels involving at least one electron or muon (ee, $e\mu$, $\mu\mu$, e + jets, and μ + jets with both event shape selection and b quark tagging). For all seven channels together, we found nine events with an expected background of 3.8 ± 0.9 events. Assuming the excess to be due to $t\bar{t}$ production, we obtained a cross section of 8.2 ± 5.1 pb for an 180 GeV/c^2 top quark mass. This cross section was consistent with both the Standard Model expectations for the top quark at this mass and with the CDF result. We concluded that this measurement did not demonstrate the existence of the top quark.

2 Optimization

DØ began its second data run (Run 1B) in December 1993, and by early 1995, we had more than tripled our data sample compared with the 1992–1993 run (Run 1A). To exploit the extended mass reach of this larger data sample, about 50 pb^{-1}, we optimized our cuts for the top quark search for top masses above 140 GeV/c^2. We optimized signal to background using Monte Carlo simulations to model the signal along with our standard background calculation methods. We achieved an improvement of a factor of four in signal to background while retaining 70% of our previous acceptance for 180 GeV/c^2 top. The improved background rejection arises primarily from requiring events to have large total transverse energy, H_T. H_T is defined as the scalar sum of the transverse energies, E_T, of the jets for the single lepton + jets and $\mu\mu$ channels and the scalar sum of the transverse energies of the leading electron and the jets for the $e\mu$ and ee channels. To be included in the calculation of H_T, jets were required to have $E_T > 15$ GeV. Electrons are identified by their longitudinal and transverse shower shapes in the calorimeter. They are required to be isolated, to have a matching

track, and to have dE/dx as measured in the tracking chambers consistent with a single electron. Electrons are required to have $|\eta| < 2.5$ for the dilepton channels and $|\eta| < 2.0$ for the e + jets channels. Muons are defined by a good quality track in the muon chambers, which points to the event vertex. Muons are also required to leave a minimum amount of energy in the calorimeter. Muons are restricted to $|\eta| < 1.7$ for data taken in the 1992–1993 run and $|\eta| < 1.0$ for data taken in the 1993–1995 run. Due to wire aging, the muon chamber efficiency at large values of η decreased with time. Isolated muons were required to be more than 0.5 in η-ϕ space from the center of the nearest jet. Jets are defined using a fixed cone algorithm of radius 0.5 in η-ϕ space.

3 Dilepton Decay Channels

We will now describe the analysis of the seven different decay modes, starting with the dilepton channels. The branching ratios for $t\bar{t}$ events to decay to dileptons are small, but the backgrounds are small as well. In each of the dilepton channels (ee, $e\mu$, and $\mu\mu$), we required two leptons, two jets, and a minimum value of H_T. In the ee and $e\mu$ channels, we required a large missing E_T, $\not\!\!E_T$, while in the $\mu\mu$ channel, the two muons were required to be inconsistent with a Z decay based on a global kinematic fit. The kinematic requirements for the three dilepton channels are given in Table 1. After all cuts, two events remain in the $e\mu$ channel. A plot of E_T^e vs. $1/p_T^\mu$ is given in Fig. 1. The stars in the plot show the two top candidates in this channel. Figure 2 shows a plot of $\not\!\!E_T$ vs. the invariant mass of the two electrons, M_{ee} for the ee events after the electron E_T and jet cuts. The data show a cluster of events consistent with Z + jets production. The one remaining event is removed by the H_T requirement, leaving no events in this channel. Figure 3 shows a plot of the probability that the pair of muons comes from the decay of a Z boson for $Z \rightarrow \mu\mu$ events and for $\mu\mu$ top events from the Monte Carlo. The location of the cut is shown by the arrow. After all cuts, one event remains in the $\mu\mu$ channel. This event has a probability of 0.008 of coming from the decay of a Z boson. The H_T distribution of the dilepton events, along with distributions for signal and background Monte Carlo events, is shown in Fig. 4. The one event with H_T below 100 GeV is the ee event that fails the H_T cut.

The backgrounds for the dilepton channels come from two sources: physics backgrounds and fake backgrounds. Physics backgrounds come from physics pro-

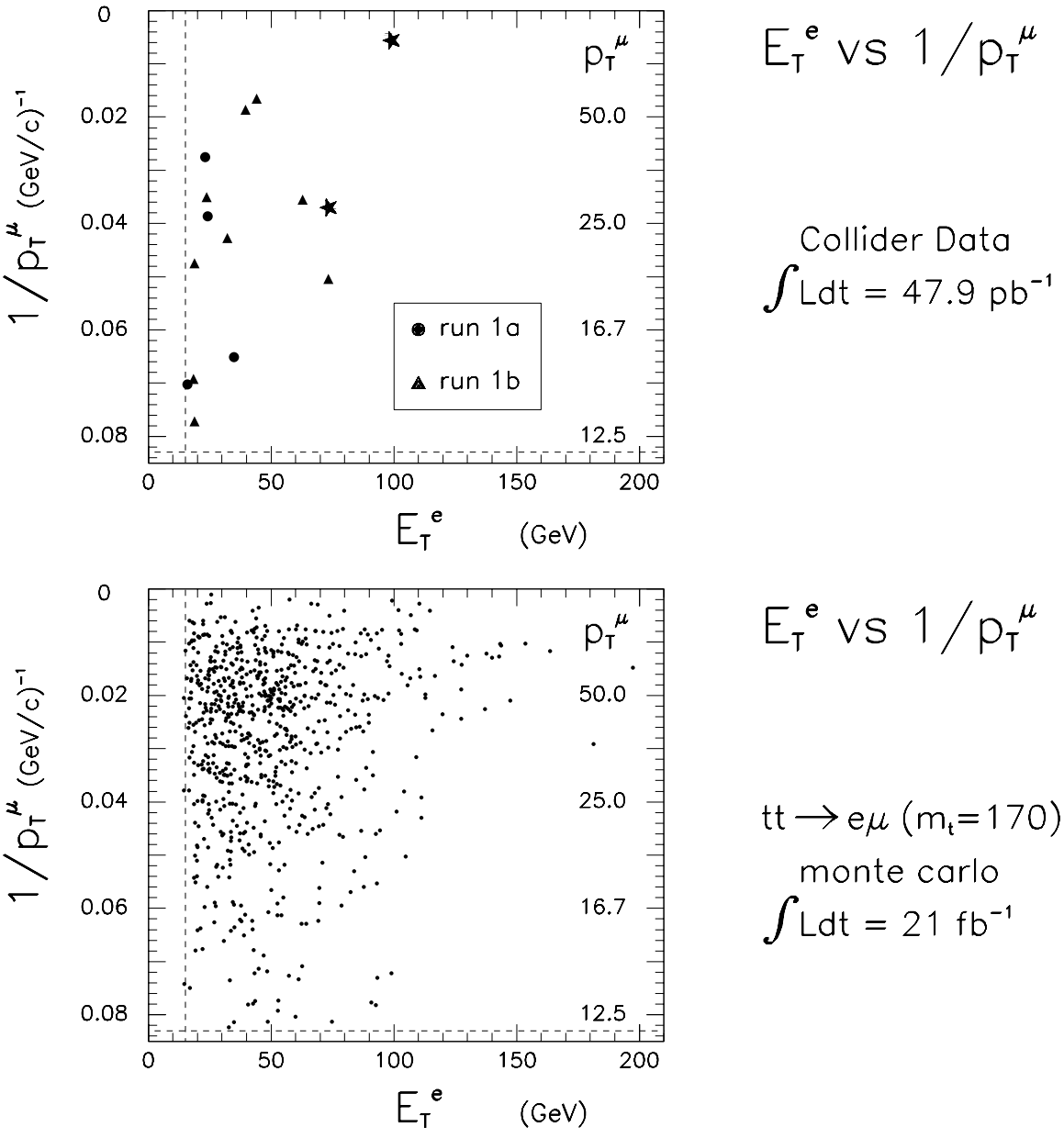


Fig. 1. Scatter plots of $1/p_T^\mu$ vs. E_T^e prior to H_T, jet, and $\not\!\!E_T$ cuts for $e\mu$ events from data and $t\bar{t}$ Monte Carlo, top quark mass of 170 GeV/c^2 and luminosity of 21 fb^{-1}. The dashed lines correspond to our cuts.

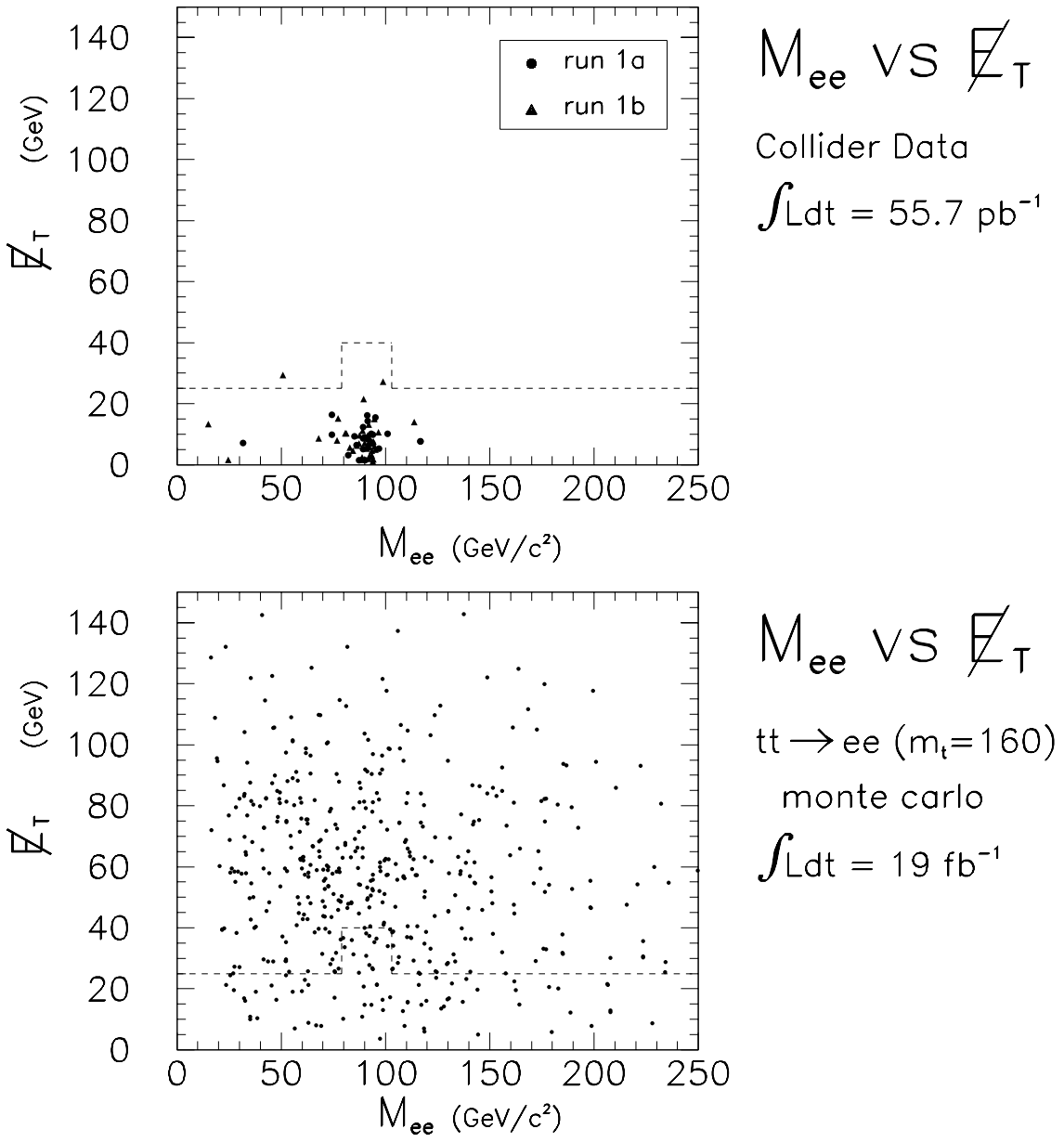


Fig. 2. Scatter plots of $\not\!\!E_T$ vs. M_{ee} after E_T and jet cuts, but before the H_T cut for ee events from the data and $t\bar{t}$ Monte Carlo, top quark mass of 160 GeV/c^2 and luminosity of 19 fb^{-1}. The dashed lines correspond to our cuts.

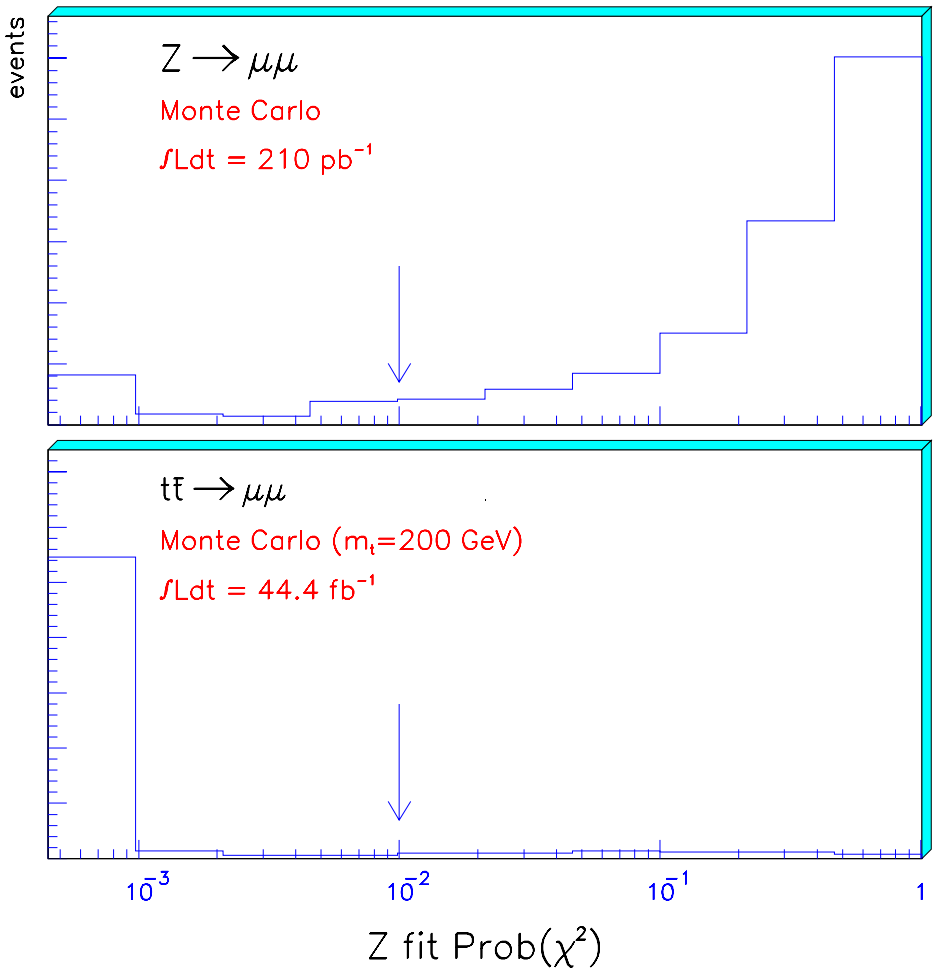


Fig. 3. Plots of probability of the dimuon pair coming from the decay of a Z boson for $Z \to \mu\mu$ and top Monte Carlo events. The location of the cut is given by the arrow.

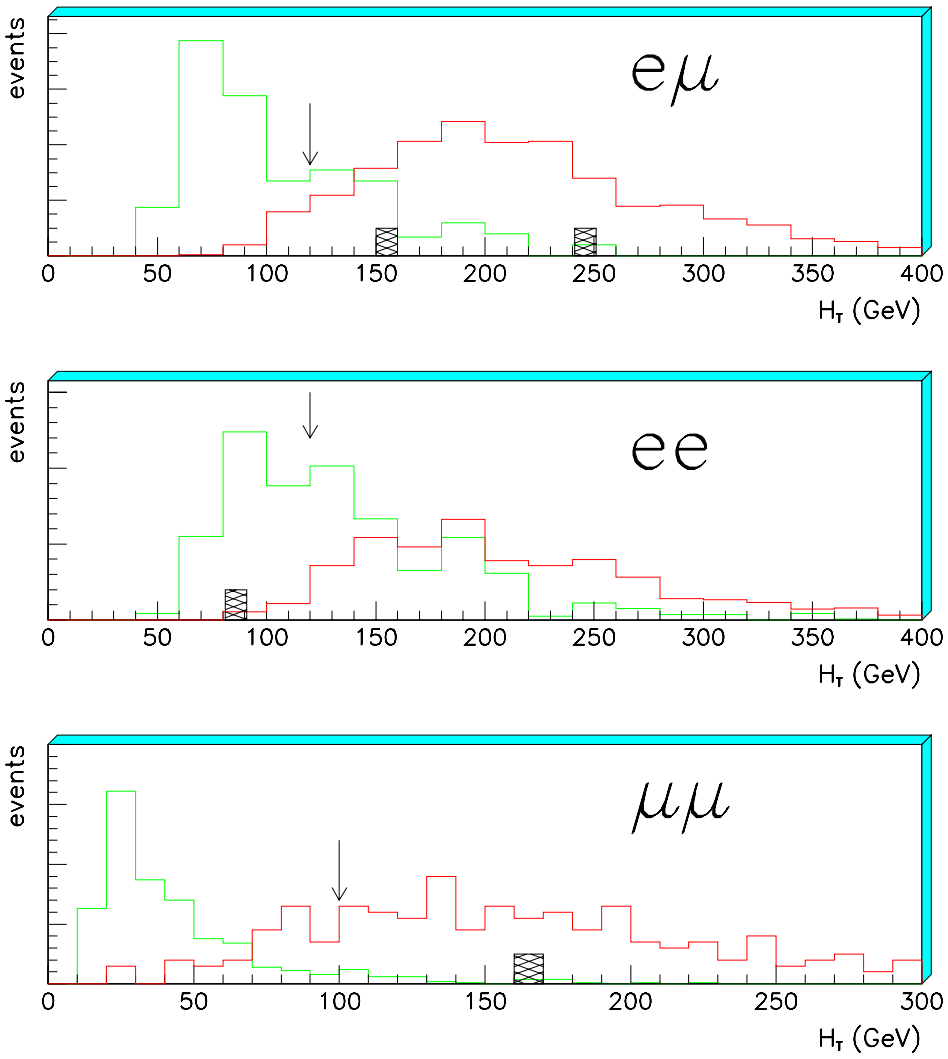


Fig. 4. H_T distribution for data events (boxes) and top (dashed lines) and background Monte Carlo events (dotted lines). The H_T distribution of the background events peaks at lower values of H_T than the top events. The arrow shows the location of the cut.

Table 1. Minimum kinematic requirements for the standard event selection (energies in GeV).

	Leptons		Jets				
Channel	$E_T(e)$	$p_T(\mu)$	N_{jet}	E_T	$\not\!\!E_T$	H_T	$\mathcal{A}$
$e\mu$ + jets	15	12	2	15	20	120	-
ee + jets	20		2	15	25	120	-
$\mu\mu$ + jets		15	2	15	-	100	-
$e + jets$	20		4	15	25	200	0.05
$\mu + jets$		15	4	15	20	200	0.05
$e + jets/\mu$	20		3	20	20	140	-
$\mu + jets/\mu$		15	3	20	20	140	-

cesses which have the same signature as top, for example $WW \rightarrow e\mu + jets$. These backgrounds are estimated using Monte Carlo. Fake backgrounds come from processes where one object is misidentified in the detector as another object. Jets, for example, are sometimes misidentified as electrons. These backgrounds are estimated directly from the data, as are the probabilities for resolution fluctuations to give large $\not\!\!E_T$. The main backgrounds are from Z and continuum Drell-Yan production, vector boson pairs (WW, WZ), heavy flavor ($b\bar{b}$, $c\bar{c}$) production, and backgrounds from jets misidentified as electrons. The total estimated background in all three dilepton channels is 0.65 ± 0.15 events. The expected top yields are calculated using the ISAJET event generator[6] coupled to a GEANT[7] simulation of the DØ detector. With the standard cuts, we observe a total of three events. The probability of the calculated background fluctuating upward to the observed signal is 0.03. From the dilepton events alone, we calculate a top cross section of 7.5 ± 5.7 pb. If we remove the H_T requirement and the cut on the probability that a pair of muons come from a Z, we have four observed events, with a calculated background of 2.66 ± 0.40 events. This set of cuts without the H_T requirement is called the loose cuts. For the loose cuts, the observed cross section is 4.4 ± 6.8 pb. We note that, although the statistical uncertainty is large, the cross sections obtained using the standard and loose cuts are consistent.

4 Lepton + Jets + Event Shape

Compared to the dilepton channels, the branching ratios are large for the lepton + jets channels, where one W decays leptonically and the other hadronically. However, the backgrounds are also large. There is a large background from W + multijet events. There is also a background from QCD multijet events where one jet fakes an electron or muon, and the missing $\not\!\!E_T$ fluctuates high. We use two different methods to distinguish $t\bar{t}$ events from background. In the first method, we exploit the different kinematics of the $t\bar{t}$ events to separate them from the background. In the second method, we use muons near jets to tag the presence of b quark jets. A $t\bar{t}$ event has two b jets, while background events have far fewer. We will discuss the kinematic, or event shape, method of separating top events from background first.

To separate top events in the lepton + jets channel from backgrounds without relying on the presence of a muon near a jet to tag the b jets, we note the following characteristics of the top events. Top events should have an isolated lepton, large missing E_T, and four jets. Since top is heavy, its decay products should tend to be central, and not at large rapidities. Top events should have a large value of the total transverse energy H_T, and the events should be nonplanar. Here, we define aplanarity, $\mathcal{A}$, where $\mathcal{A} = 3/2 \times$ the smallest eigenvalue of the normalized momentum tensor constructed from the jets in the events. $\mathcal{A} = 0.5$ for spherical events and is near zero for planar and linear events. The kinematic requirements on the lepton + jets events are given in Table 1. The principal difference between this analysis and previous analyses[4] is the tighter cut on H_T. The background events from W + four jet production should have lower values of H_T. Backgrounds from multijet events where one jet fakes an electron or muon are suppressed by the missing E_T requirement and the $\mathcal{A}$ requirement. In Fig. 5, we show the H_T distribution for Monte Carlo W + jets events and for $t\bar{t}$ events where we assume a top mass of 200 GeV/c^2. In Fig. 6, we show plots of H_T for two and three jet events where the contamination from top events is small. The agreement between our calculated background and the observed H_T distribution is good, demonstrating that we are able to calculate our backgrounds reliably. To check for possible systematic biases, we define a loose set of cuts in the lepton + jets channels as well. The loose cuts require that we make no requirement on H_T and require $\mathcal{A} > 0.03$. The standard cuts require $\mathcal{A} > 0.05$.

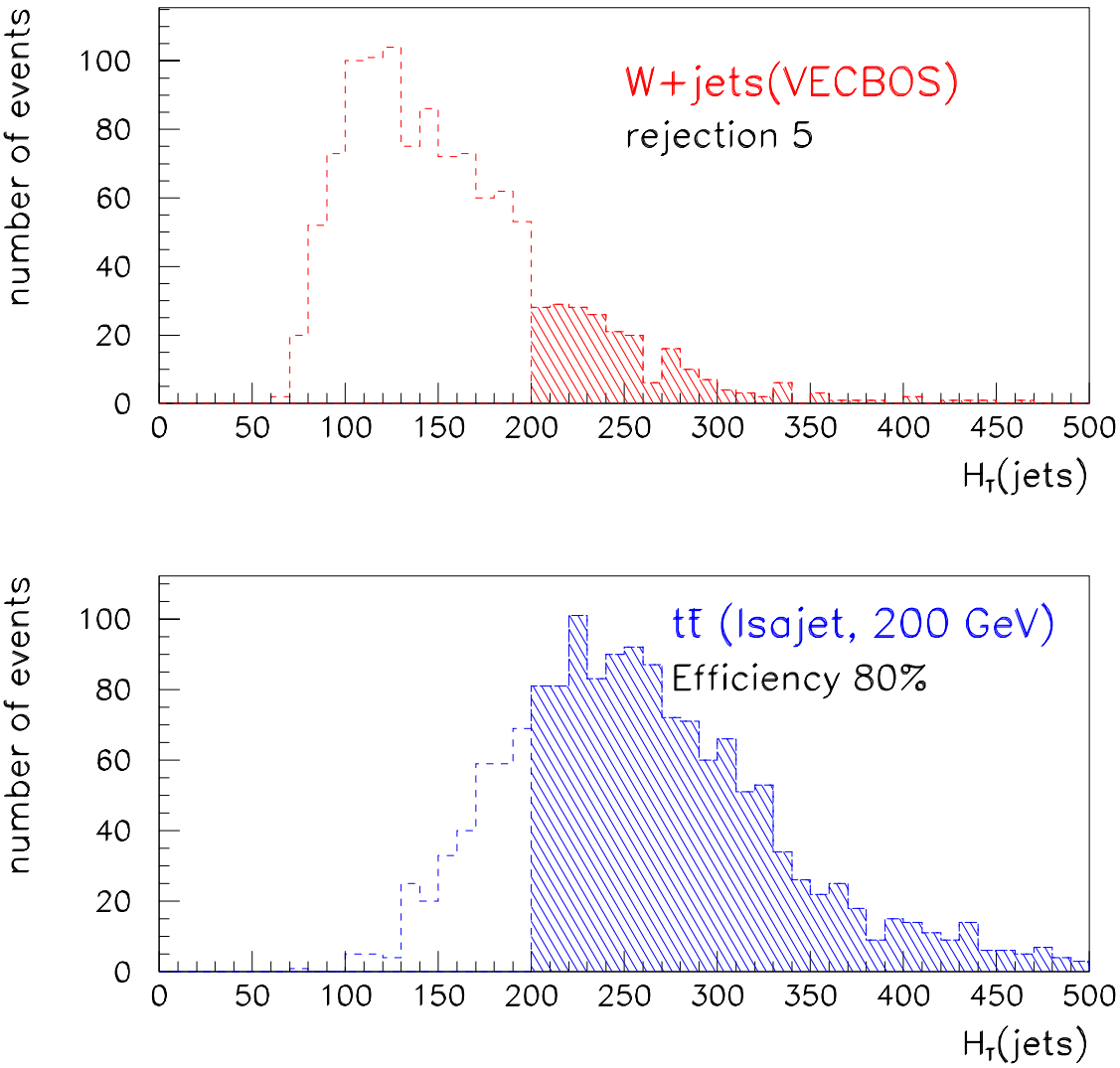


Fig. 5. Plot of H_T for e + four jet events for W + jets Monte Carlo events and $t\bar{t}$ Monte Carlo, top quark mass = 200 GeV/c^2. The shaded region is above our cut.

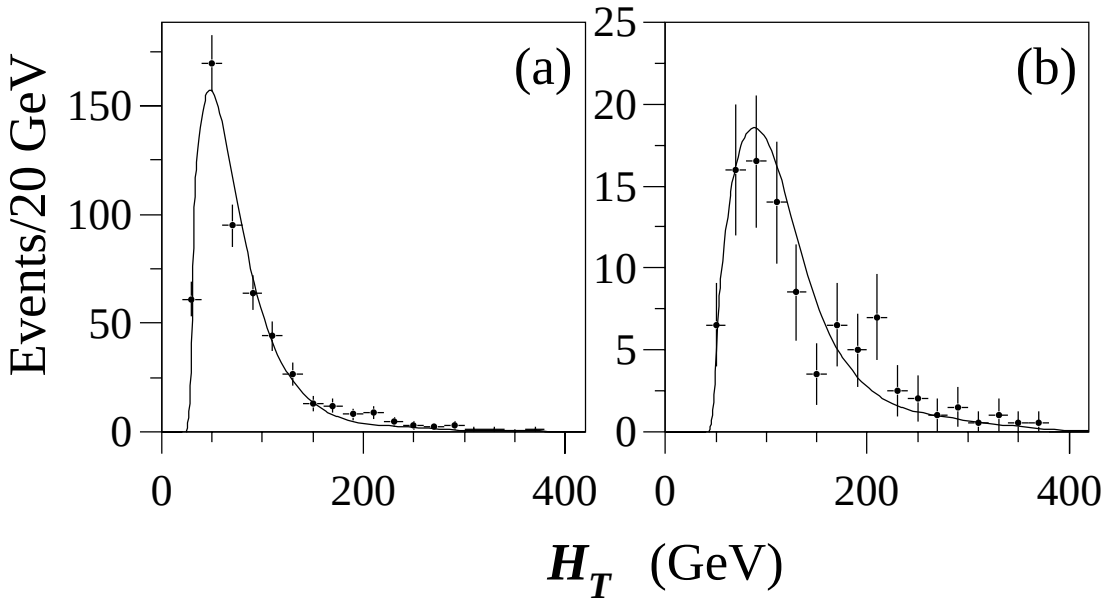


Fig. 6. Plots of H_T for e + jets events with $\not\!\!E_T > 25$ GeV. (a) e + two jet events, and (b) e + three jet events. The points are data. The curves are background calculated using VECBOS and multijet data.

The backgrounds due to multijet production where a jet fakes an electron are determined from the $\not\!\!E_T$ distribution of events containing a fake electron. Fake electrons are defined as electromagnetic clusters in the calorimeter that fail the electron identification cuts. The number of such events with $\not\!\!E_T > 25$ GeV is then scaled by the probability of a jet faking an electron, which is determined from multijet events with low $\not\!\!E_T$. The backgrounds due to multijet production where a jet fakes an isolated muon are determined by counting the number of events with muons that pass all cuts except the isolated muon requirement. This number of events is then multiplied by the probability that a jet will fake an isolated muon, which is determined from low jet multiplicity events where top and W events are negligible, but bottom and charm are present.

Backgrounds from W + jet production, which contain real isolated electrons and muons, are determined using the fact that QCD background processes follow an exponential scaling law in the number of observed jets.[8] This leads to the approximate prediction:

$$\frac{N_n}{N_{n-1}} = constant, \tag{1}$$

where N_n is the number of lepton + n jet events and N_{n-1} is the number of lepton + $(n-1)$ jet events. QCD multijet events and Z + jets events in our data are consistent with this assumption. Since W + jets production is also a QCD process, these events are predicted to follow this law as well, and, as can be seen from Fig. 7, they do to within the limit of available statistics. The slope of the line in Fig. 7 is then used to determine the number of W + four jet events in our sample before the $\mathcal{A}$ and H_T cuts. A 20% systematic uncertainty is assigned to the slope of the line, determined from difference in the slopes of the W + jets events and the multijet events. The Z and W slopes agree within statistics. The fraction of W + four jet events passing the $\mathcal{A}$ and H_T cuts is determined from Monte Carlo simulations.

We check the calculated backgrounds by fitting the observed distribution of events in the $\mathcal{A}$, H_T plane. We divide the $\mathcal{A}$, H_T plane into four quadrants whose boundaries are defined by our $\mathcal{A}$ and H_T cuts. The ratio of events of each type (top, W + jets, QCD fake + jets) in each quadrant is then taken from Monte Carlo or the fake electron data. The overall number of events of each type is

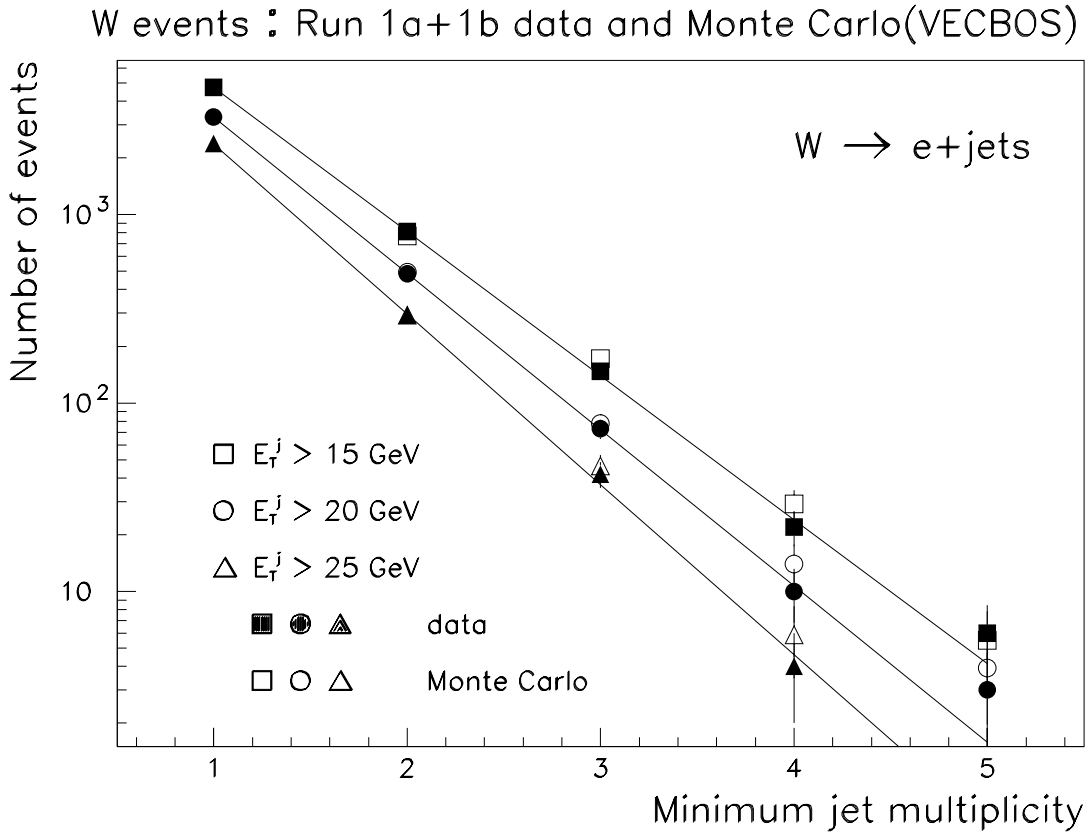


Fig. 7. Plot of number of W + jets events as a function of jet multiplicity before $\mathcal{A}$ and H_T cuts.

determined by fitting the observed distribution. See Fig. 8. The results for the backgrounds agree with those obtained using the scaling law.

The results from the lepton + jets channels without muon tagging are listed in Table 2. With the standard cuts, the total number of untagged e + jets and μ + jets events is eight with a background of 1.9 ± 0.5 events. The probability of an upward fluctuation of background having resulted in the observed signal is 0.002 (2.9 σ). The cross section from the lepton + jets channels is 4.9 ± 2.5 pb. With the loose cuts ($\mathcal{A} > 0.03$ and no H_T cut), the total number of e + jets and μ + jets events is 23 with a background of 15.7 ± 3.1 events. This corresponds to a cross section of 4.0 ± 3.2 pb. The agreement in the cross section calculated with the standard and loose cuts indicates that our backgrounds are correctly accounted for within the limits of statistics.

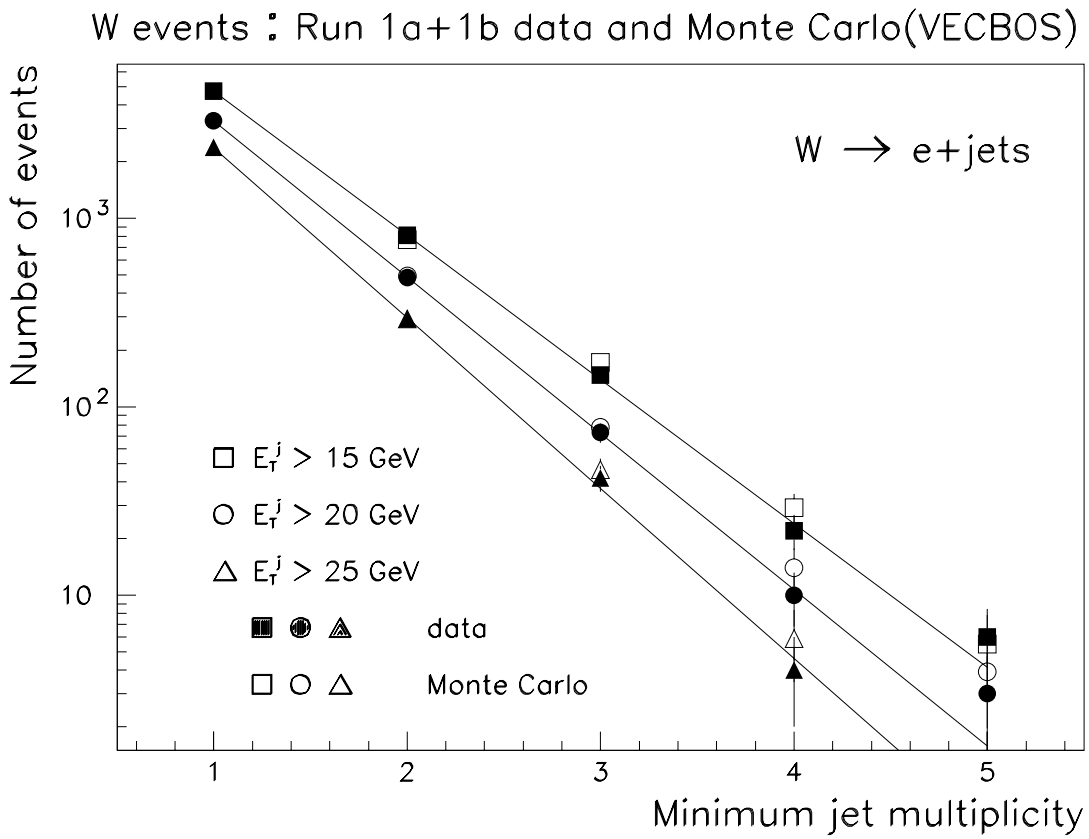


Fig. 8. $\mathcal{A}$ vs. H_T distribution for data, $t\bar{t}$ events, W + jet events, and QCD multijet events.

Table 2. Summary of number of events observed, the predicted background, and the probability for the background to account for the data for both standard and loose cuts. A $t\bar{t}$ production cross section ($\sigma_{t\bar{t}}$) is also given for a top quark mass (M_t) of 200 GeV/c^2.

	Standard Selection	Loose Selection
Dileptons	3	4
Lepton + Jets (Shape)	8	23
Lepton + Jets (Muon tag)	6	6
All channels	17	33
Background	3.8 ± 0.6	20.6 ± 3.2
Probability	2×10^{-6} (4.6σ)	0.023 (2.0σ)
$\sigma_{t\bar{t}}$ $(M_t = 200$ GeV/c$^2)$	6.3 ± 2.2 pb	4.5 ± 2.5 pb

5 Lepton + Jets with b Quark Tag

By requiring the presence of a b quark jet in our events, we can substantially reduce the major backgrounds. We tag b events by requiring a muon to be located within 0.5 in η-ϕ space of a jet in the event and to have a minimum P_T of 4 GeV. For the lepton + jets channel without tags, we require that no such muons be present. The two sets of channels are then independent.

Standard Model $t\bar{t}$ events that decay according to the lepton + jets signature contain, after the decays of the top quarks and the W's, two b quarks and approximately 2.5 c quarks. Each b or c has a branching ratio into a muon of about 10%. Thus, 44% of the $t\bar{t}$ lepton + jets events contain a muon from a b or c decay. The DØ muon system acceptance and detection efficiency (about 45%) is such that about 20% of the $t\bar{t}$ events have a detectable muon tag. The kinematic requirements on the lepton + jets with muon tag events are given in Table 1. The loose kinematic cuts are the same as the standard cuts, except the cut on H_T is not used.

Backgrounds are calculated by multiplying the observed number of W + jets events and multijet events with a fake lepton by the fraction of background events containing muon tags. As can be seen in Fig. 9, the tagging rate is consistent with being proportional to the number of jets, as would be expected if heavy quarks from gluon splitting and fakes dominate the background. Corrections are made for the change in the tagging probability with jet E_T and event $\not\!\!E_T$.

The results from the lepton + jets with muon tag channels are listed in Table 2. With the standard cuts, the total number of e + jets and μ + jets with muon tag events is six with a background of 1.2 ± 0.2 events. The probability of an upward fluctuation of background having resulted in the observed signal is 0.002 (2.9 σ). The cross section from the lepton + jets channels with tag is 8.9 ± 4.8 pb. With the loose cuts (no H_T cut), the total number of e + jets and μ + jets events is also six with a background of 2.2 ± 0.3 events. The cross section for the loose cuts for the lepton + jets channels with muon tag is 6.3 ± 4.2 pb. Here again, we assume a top mass of 200 GeV/c^2. The agreement in the cross section calculated with the standard and loose cuts indicates that our backgrounds are correctly accounted for within the limits of statistics. In Fig. 10, we show the distribution of loose

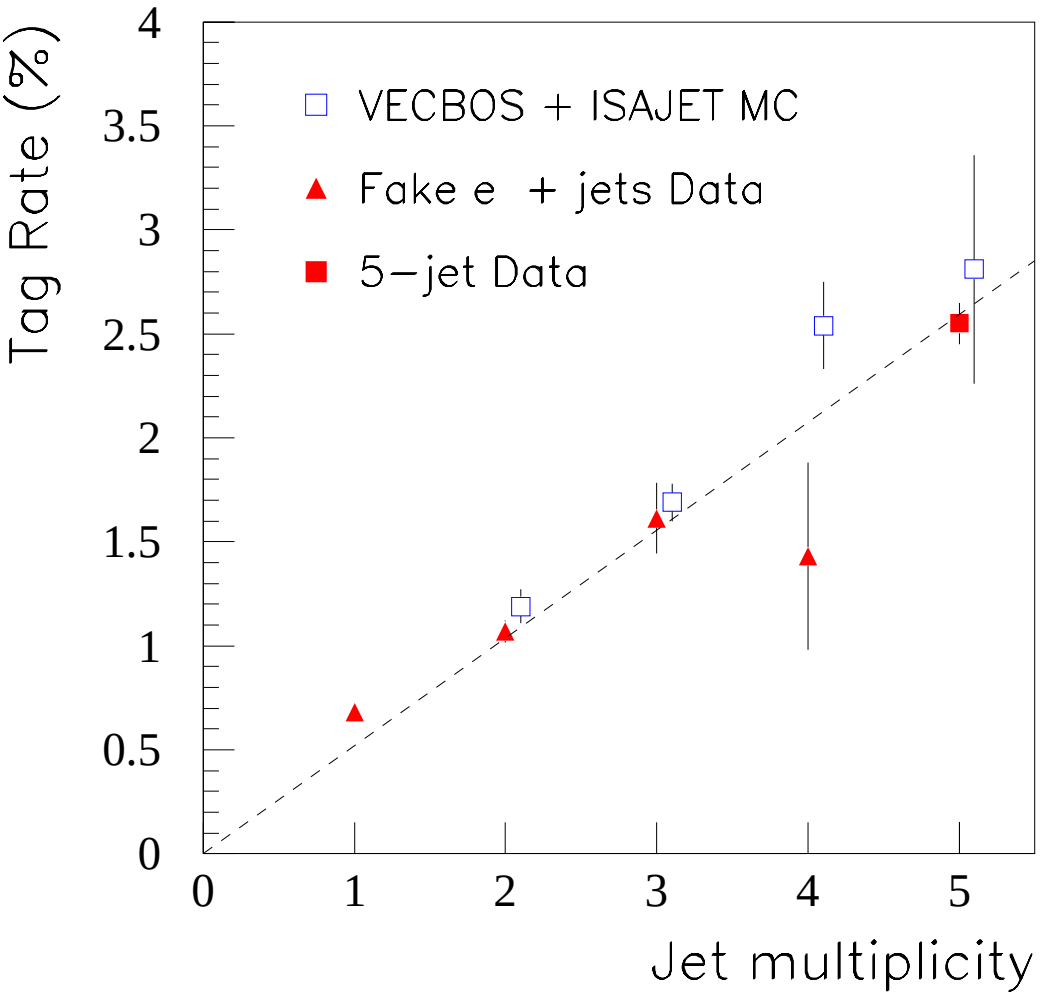


Fig. 9. Muon tag rate for background events as a function of jet multiplicity.

cut events after the QCD background contribution has been subtracted compared with the W + jets background. The excess of events for jet multiplicities greater than two is clear.

6 Cross Section and Significance

Combining the above seven channels, we observe a total of 17 events with an expected background of 3.8 ± 0.6 events. The probability of an upward fluctuation of the background giving 17 or more events is 2×10^{-6}. This corresponds to a 4.6 standard deviation effect for a Gaussian probability distribution. Our measured cross section plotted as a function of assumed top quark mass is shown in Fig. 11. Also shown is a theoretical cross section curve.[9] Assuming a top quark mass of 200 GeV/c^2, our measured cross section is 6.3 ± 2.2 pb. The error in the cross section includes a 12% uncertainty in the integrated luminosity. We have included the difference in top detection efficiencies when using the HERWIG[10] Monte Carlo instead of ISAJET in the systematic error. For the loose cuts, we observe a total of 33 events with an expected background of 20.6 ± 3.2 events. This leads to a cross section of 4.5 ± 2.5 pb for a 200 GeV/c^2 top mass for the loose cuts, in good agreement with the value obtained from the standard cuts. Figure 12 shows the cross sections for the various decay channels calculated individually. We have calculated the probability of seeing our distribution of events across the seven channels and find that our result is consistent with the expected top branching fractions at the 53% confidence level. (Note that the branching fractions are determined by the assumption of top decay to $W + b$, and the Standard Model W branching fractions.) We observe a statistically significant excess of events, and the distribution of these events among the decay channels studied is consistent with top quark production. We conclude that we observe the top quark.

We also have results from searches using multivariate techniques and from searches in the channel where both W's decay to jets. Details of these analyses can be found in the references.[11,12]

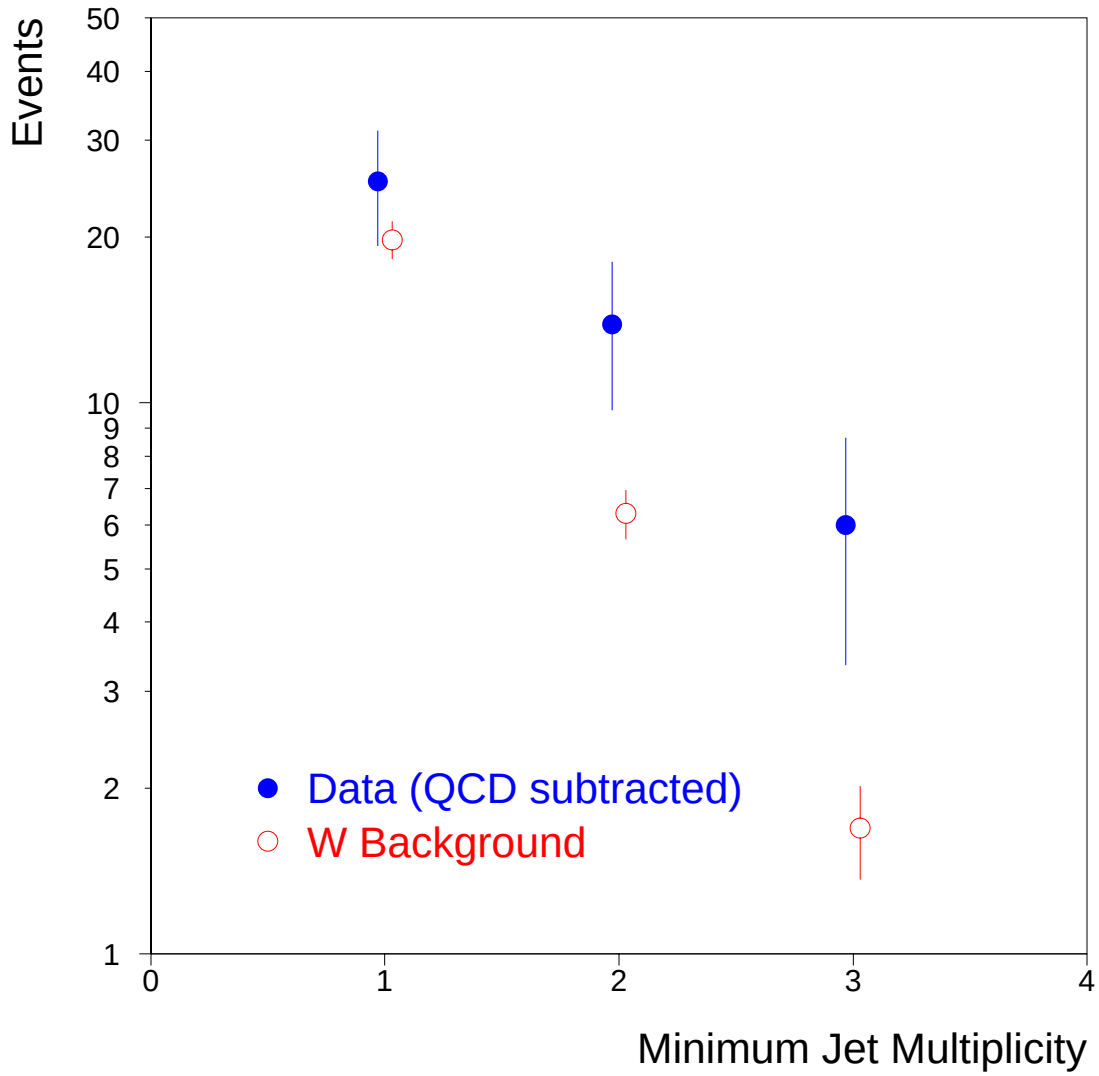


Fig. 10. Plot of lepton + jets + muon tag events as a function of inclusive jet multiplicity.

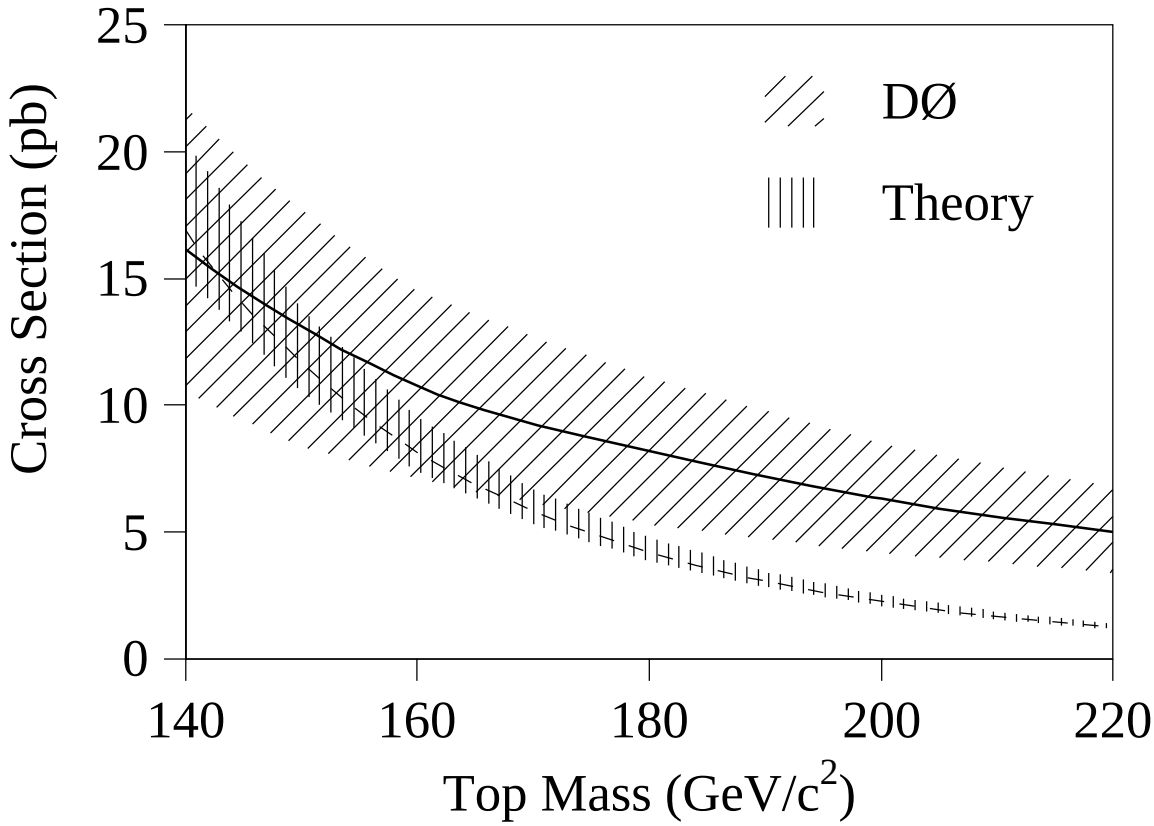


Fig. 11. Measured $t\bar{t}$ production cross section as a function of assumed top quark mass. Also shown is the theoretical cross section curve.

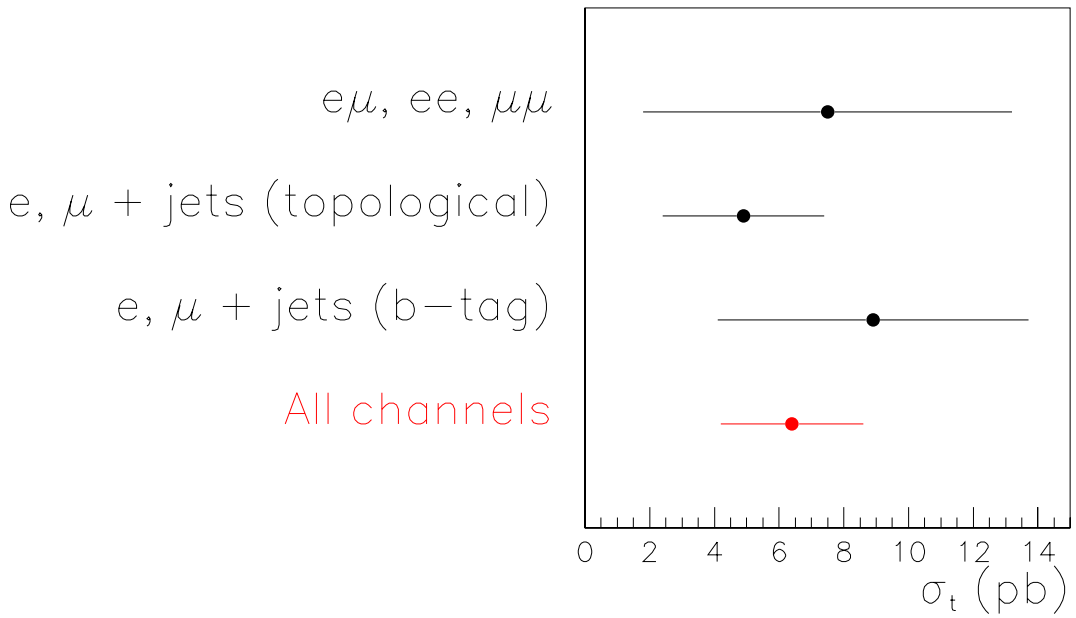


Fig. 12. Measured $t\bar{t}$ production cross section as a function of decay mode for 200 GeV/c^2 assumed top quark mass.

7 Mass Analysis

Having determined that there is an excess of events in our data, and that the observed distribution of events is consistent with that expected from the Standard Model top quark, we now study the kinematic properties of our lepton + jets events in order to determine the top quark mass.

We assume that our excess events are due to the process

$$t\bar{t} \rightarrow (W^+b)(W^-\bar{b}) \rightarrow (l\nu b)(q\bar{q}\bar{b}).$$

Using both W mass constraints, and requiring that the masses of the t and $\bar{t}$ quarks be equal, we perform a two constraint (2C) kinematic fit for the top quark mass. We select lepton + jet events requiring at least four jets with $E_T > 15$ GeV and $|\eta| < 2.5$. We use jets of cone radius 0.3. We use only the four highest E_T jets in the fit. For each event, there are 12 distinct ways of assigning jets to the original partons. We use up to three combinations with $\chi^2 < 7$ and calculate a χ^2 weighted average mass for each event.

We have performed extensive Monte Carlo studies of the method and tested many possible variations. For example, we tested using just the best χ^2 combination instead of the weighted average of the three best for each event and found that the weighted average method gave slightly better results. We tested our jet energy corrections by studying Z + jet events where the Z decays to e^+e^- and comparing the E_T of the Z with the E_T predicted from the jets. As shown in Fig. 13, after jet corrections, E_T is well-balanced in our events. Jets with a tag muon have twice the muon p_T added to the jet energy to compensate for the muon and the missing neutrino.

We have studied in detail the effects of initial state radiation (ISR), final state radiation (FSR), and the combinatorical background due to the wrong combinations. Note that the solution with the lowest χ^2 corresponds to the correct jet assignment less than 20% of the time. Figure 14 shows the effects of wrong jet assignment, and QCD radiation on our mass resolution for 180 GeV/c^2 top events generated with the ISAJET and HERWIG Monte Carlos.

We apply our kinematical fitting procedure to ISAJET Monte Carlo top events with a full GEANT detector simulation to obtain resolution functions for different assumed top masses. These distributions are shown in Fig. 15 for a range of top mass values from 140 to 240 GeV/c^2. Note that the average value of the calculated mass is shifted from the input mass due to the effects of ISR, FSR, and jet assignment combinatorics. The Monte Carlo top mass distributions are then fit; the fits are smoothed and parametrized as a continuous function of top mass. The mass distributions from background events are obtained by applying the same kinematic fit to W + jet events from VECBOS[13] and QCD multijet events obtained from the bad electron sample.

Eleven of the 14 lepton + jets candidate events selected using the standard cuts were successfully fit. Figure 16 shows the mass distribution, along with the likelihood distribution from the fit. A maximum likelihood fit is then used to extract the top mass from our data. The likelihood fit gives a top mass of 199^{+31}_{-25} GeV/c^2 and describes the data well. To increase the statistics available for the fit and to test for any possible bias from the H_T cut, we also performed the mass analysis on events selected using the loose requirement. Out of the 27 loose lepton + jets events that have at least four jets, 24 were successfully fit. The likelihood fit to the loose sample gave a value of 199^{+19}_{-21} GeV/c^2 for the top mass. The statistical uncertainty is smaller for the loose cuts, since the H_T cut used in

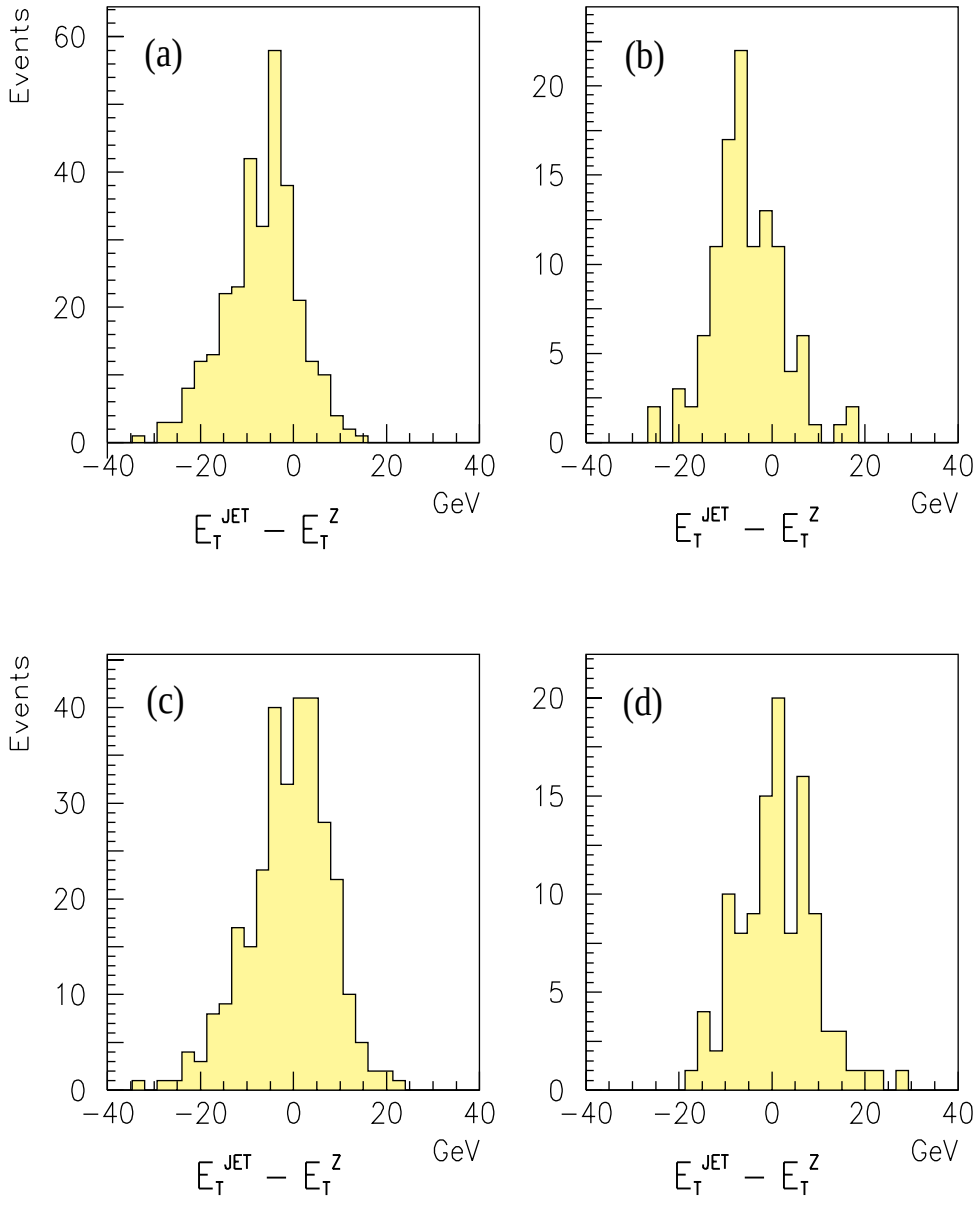


Fig. 13. $E_T^{jet} - E_T^Z$ for (a) Monte Carlo before final jet energy corrections, (b) data before final jet energy corrections, (c) Monte Carlo after all jet energy corrections, and (d) data after all jet energy corrections.

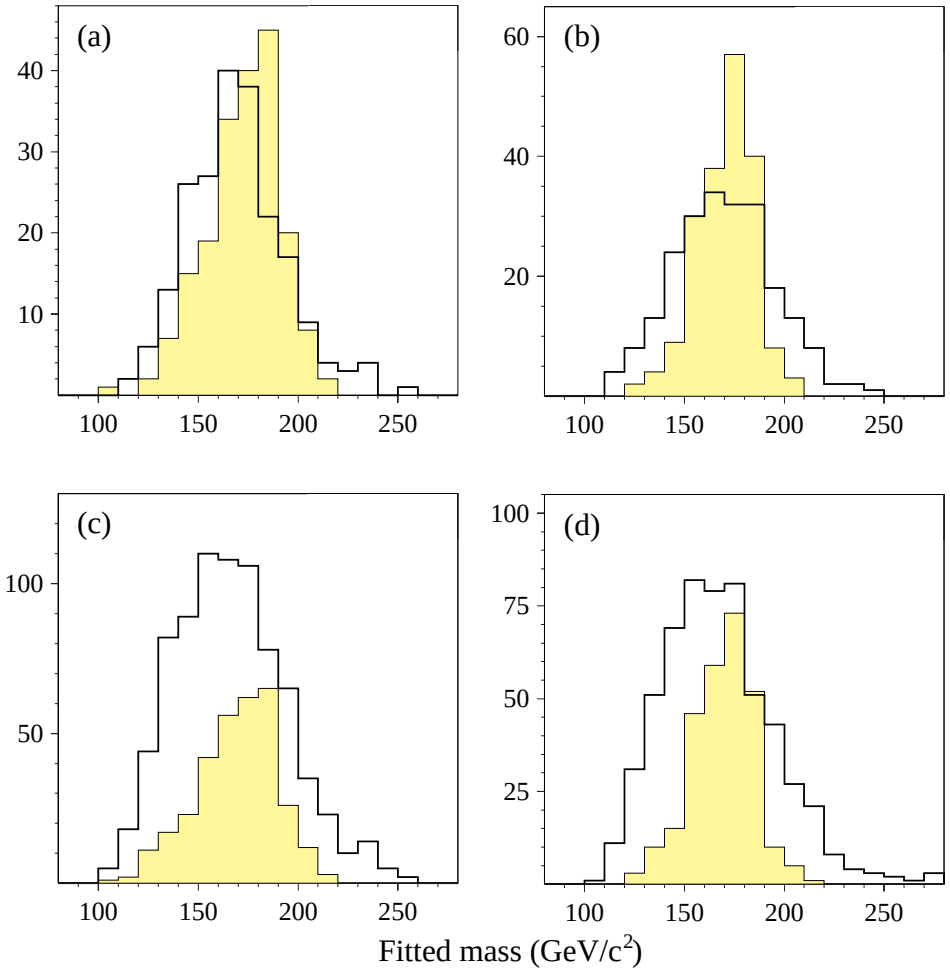


Fig. 14. Distributions of fitted top quark mass for (a),(c) ISAJET and (b), (d) HERWIG 180 GeV/c^2 top Monte Carlo. In (a) and (b), events have exactly four jets, uniquely matched to the four primary jets (two b jets + two W jets) required for mass analysis. In (b) and (d), four or more jets are allowed as in actual analysis, without any matching requirement. Open histograms show the fitted top quark mass as in actual analysis. Shaded histograms show the fitted mass for the correct jet assignment for those events in which the four highest E_T jets are uniquely matched to the four primary jets.

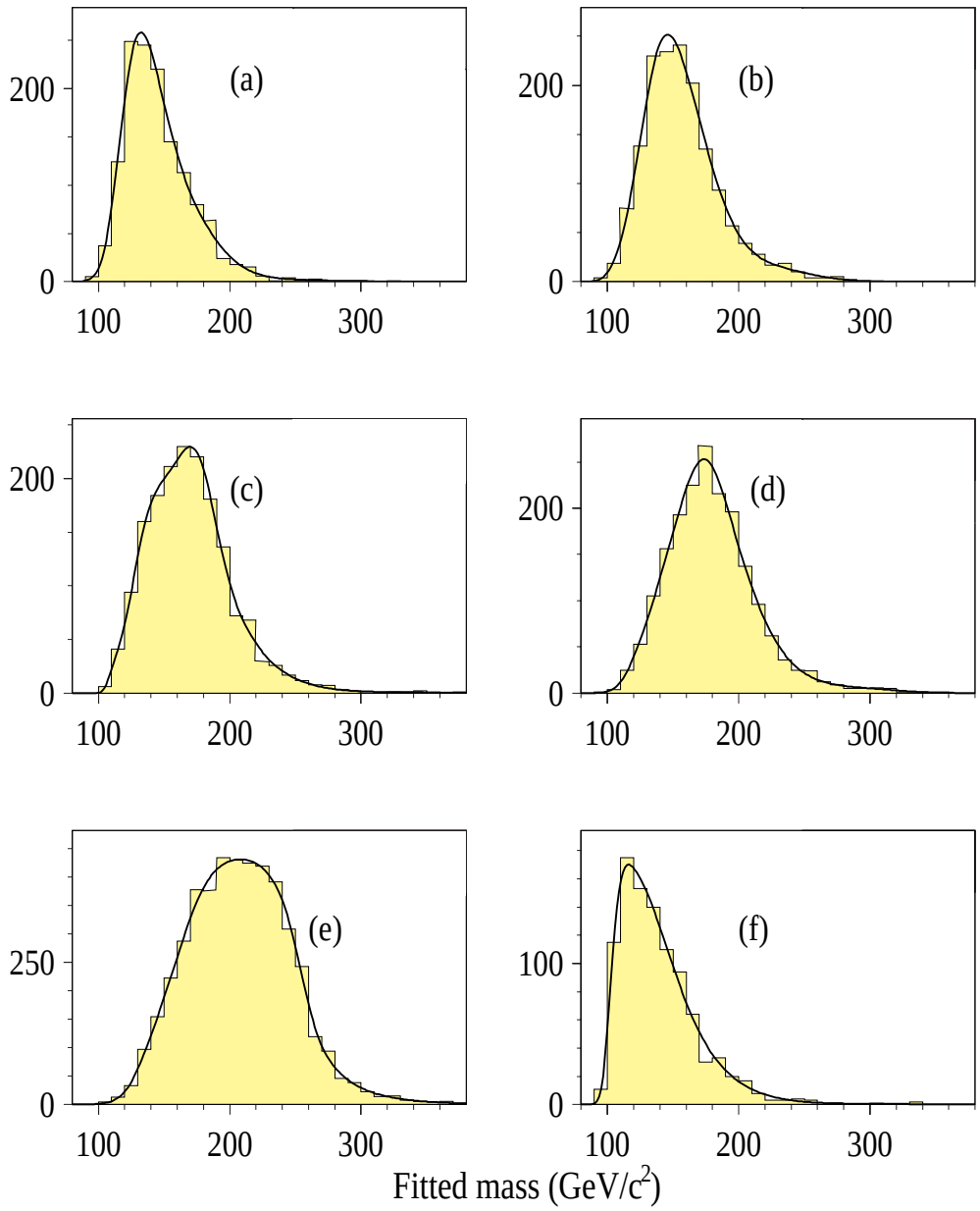


Fig. 15. Results of kinematic fit for top mass for ISAJET Monte Carlo events (a) 140 GeV/c^2 top, (b) 160 GeV/c^2 top, (c) 180 GeV/c^2 top, (d) 200 GeV/c^2 top, (e) 240 GeV/c^2 top, and (f) for W + four jet Monte Carlo.

the standard analysis biases the fitted mass distributions. The events are shown in Fig. 16, along with the likelihood distribution from the fit. The results of the fit did not depend significantly on whether or not the background normalization was constrained to the calculated value. As can be seen in Fig. 16, the masses of the tagged events are consistent with those of the untagged events. Using HERWIG instead of ISAJET resulted in a 195 GeV/c^2 value for the top mass. The systematic uncertainty on the top mass is $^{+14}_{-21}$ GeV/c^2 and is dominated by the uncertainty in the jet energy scale.[14]

Figures 17 and 18 give preliminary results for the transverse momentum distribution and $t\bar{t}$ mass distribution for top candidate events that pass the standard cuts. Within the limits of the low statistics, the observed distributions are in agreement with the expected mixture of Standard Model top plus background events.

8 Hadronic W Decays

Top events have two W's in them, while the main backgrounds do not. In this section, we present preliminary results of a search in our single lepton data for evidence of W to two jet decays. We select events requiring the loose cuts and at least four jets. The jets are required to have $|\eta| < 2.5$. All jet assignments consistent with the muon tag (if one is present) are used. The solutions are weighted according to $e^{-\chi^2/2}$ with $\chi^2 \propto ln^2(M(bl\nu)/M(bjj))$. Each event's weights are normalized to unity. For the top mass, we plot the weighted average of the $bl\nu$ mass and the bjj mass. When the b jet in $t \to bjj$ is untagged, often the highest energy jet (jet 1) as measured in the top CM frame is assigned to the b jet, and we plot M_{23} for the W mass. However, if $(E_1 - E_2) < (E_2 - E_3)$ in the top CM frame, we plot M_{23} and M_{13} with equal weight. The dijet mass vs. top mass distribution is shown in Fig. 19 for 200 GeV/c^2 HERWIG top + background events and for background events only. The same plots for data events and for background events are shown in Fig. 20. Note that the data are inconsistent with the prediction for background alone. If we now plot the top mass for dijet masses greater than 58 GeV/c^2 and the dijet mass for top masses greater than 150 GeV/c^2, we see both a peak in the top mass distribution consistent with 200 GeV/c^2 top and a peak in the dijet mass consistent with W decays. See Fig. 21. We conclude that

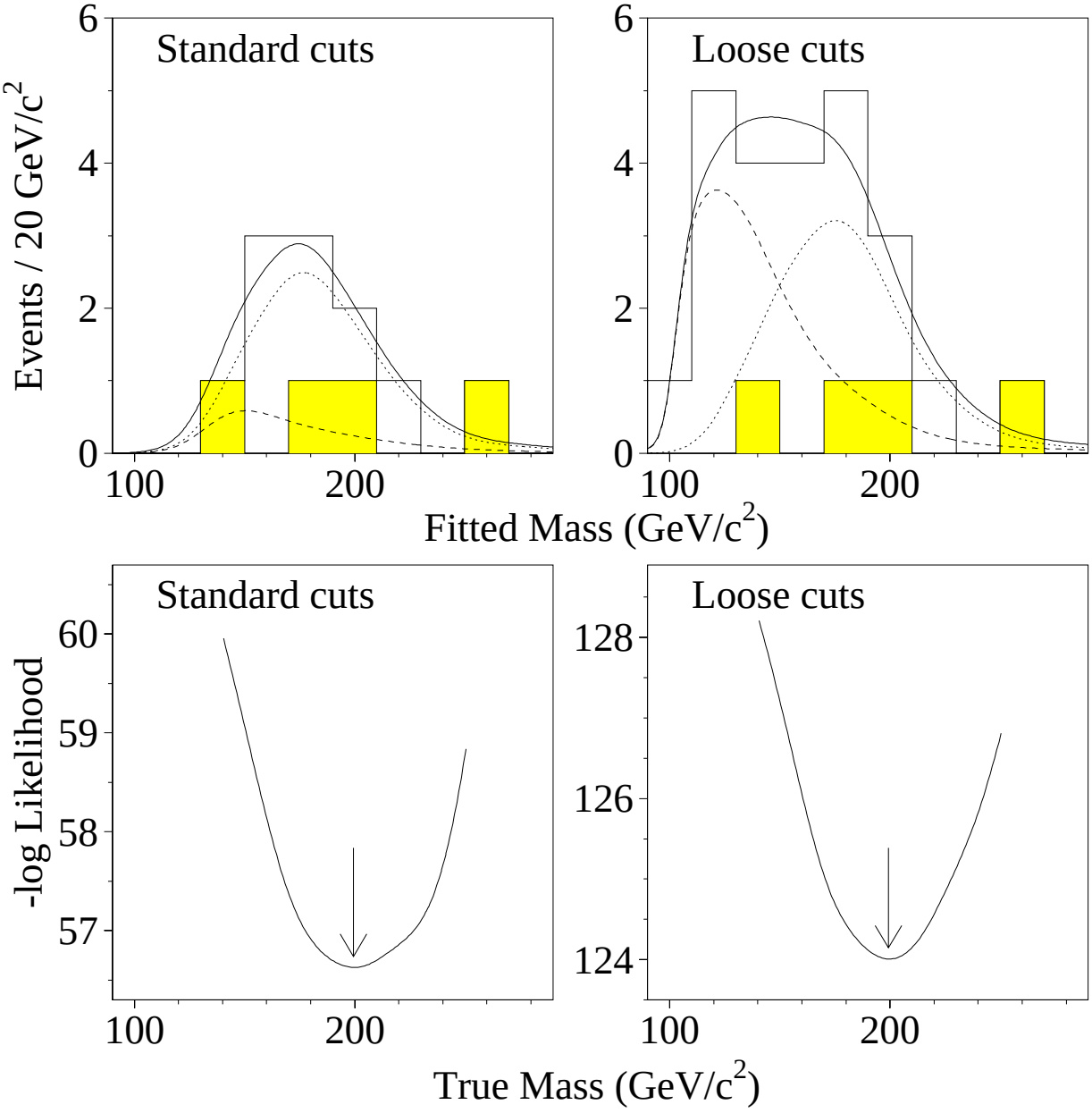


Fig. 16. Top mass and likelihood distribution for the standard and loose cuts. The dashed line is the expected background distribution, the dotted line the expected distribution from 199 GeV/c^2 Monte Carlo top events, and the solid line the sum of top + background. The shaded events have tag muons.

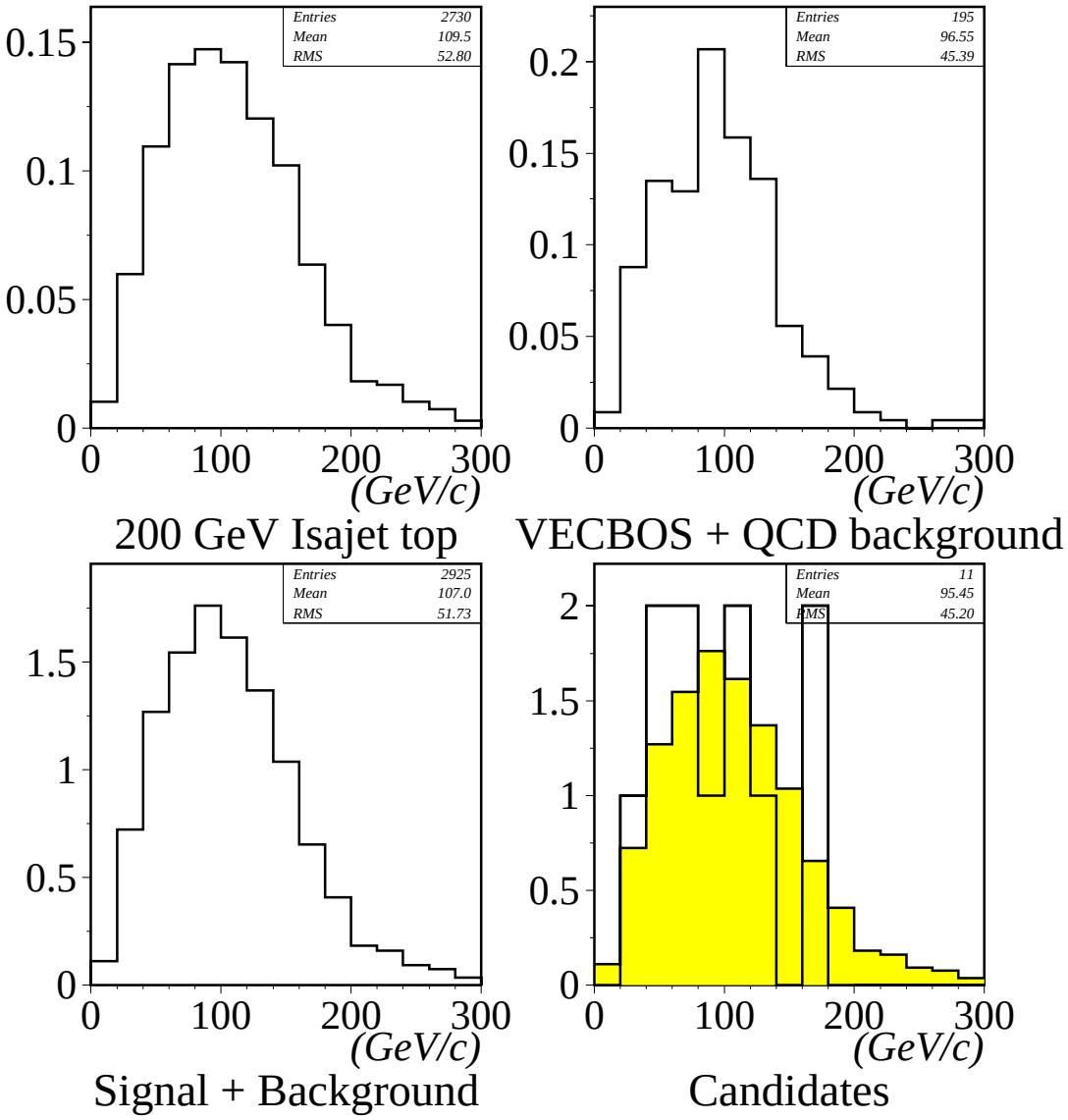


Fig. 17. Top transverse momentum distribution for the standard cuts. The shaded region is the expected distribution of top plus background events normalized to the data.

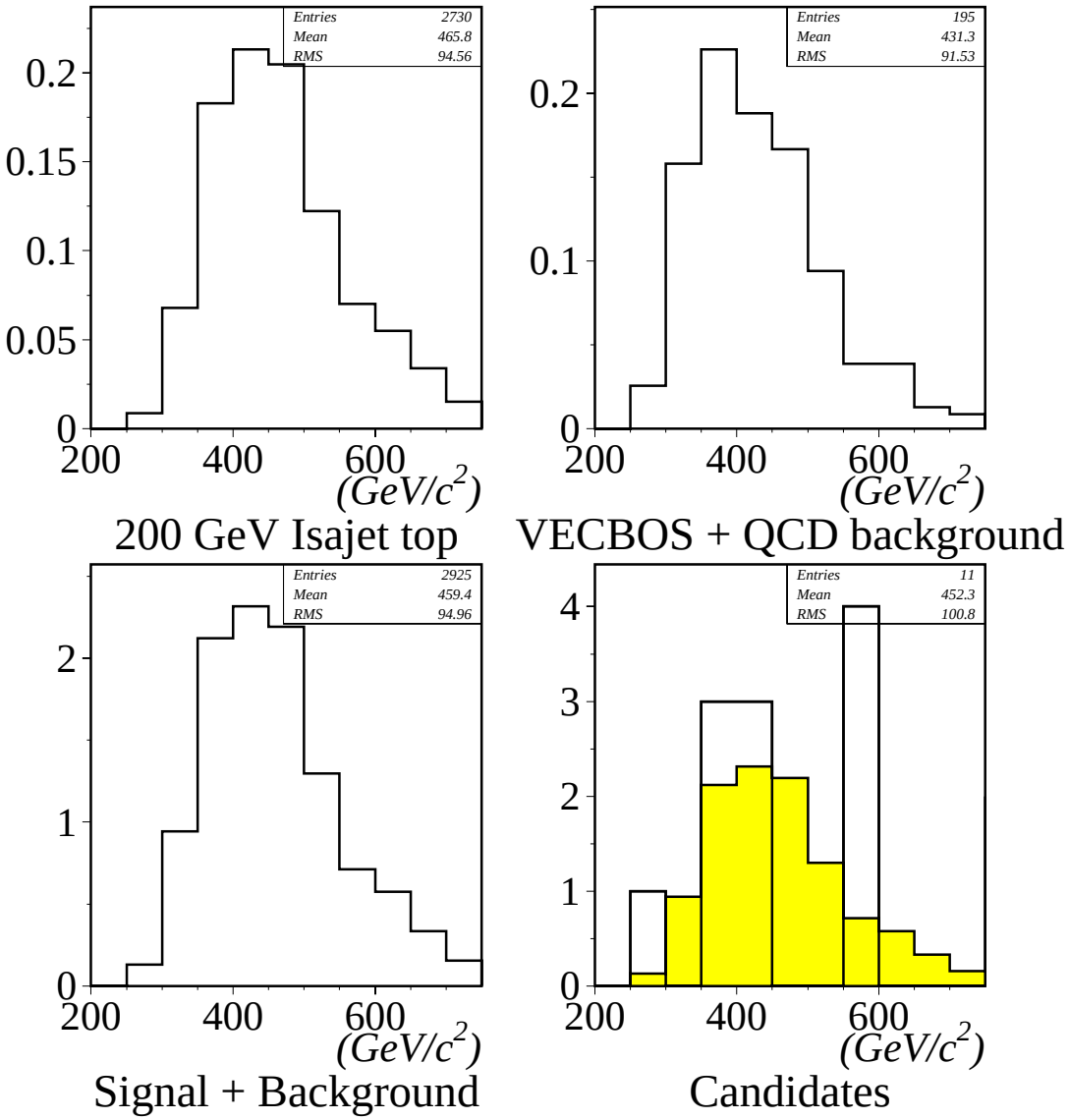


Fig. 18. $t\bar{t}$ mass distribution for the standard cuts. The shaded region is the expected distribution of top plus background events normalized to the data.

the events that are from the top signal region also show a W to dijet mass peak, as expected for top events.

9 Conclusion

We report the observation of the top quark with the DØ detector. We measure the top mass to be $199^{+19}_{-21}(stat.)\ ^{+14}_{-21}$(syst.) GeV/c^2 and measure the production cross section to be 6.4 ± 2.2 pb at our central mass. We show the existence of a peak in the dijet mass distribution consistent with hadronic decays of the W in our top data. Note that by the time these proceedings appear in print, we will have updated results from a data sample of approximately 100 pb^{-1}, which is twice as large as that discussed here. An upgraded version of the DØ detector will run in 1999, and we expect to begin making detailed measurements of the properties of the top quark with a data sample of 2 fb^{-1} at that time.

I would like to thank all of my colleagues on the DØ experiment whose efforts made these results possible.

DØ thanks the Fermilab Accelerator, Computing, and Research Divisions, and the support staffs at the collaborating institutions for their contributions to the success of this work. We also acknowledge the support of the U.S. Department of Energy; the U.S. National Science Foundation; the Commissariat à L'Energie Atomique in France; the Ministry for Atomic Energy and the Ministry of Science and Technology Policy in Russia; CNPq in Brazil; the Departments of Atomic Energy and Science and Education in India; Colciencias in Colombia; CONACyT in Mexico; the Ministry of Education, Research Foundation, and KOSEF in Korea; and the A. P. Sloan Foundation.

References

[1] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[2] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995).

[3] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **72**, 2138 (1994).

[4] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **74**, 2422 (1995).

[5] F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **50**, 2966 (1994); Phys. Rev. Lett. **73**, 225 (1994).

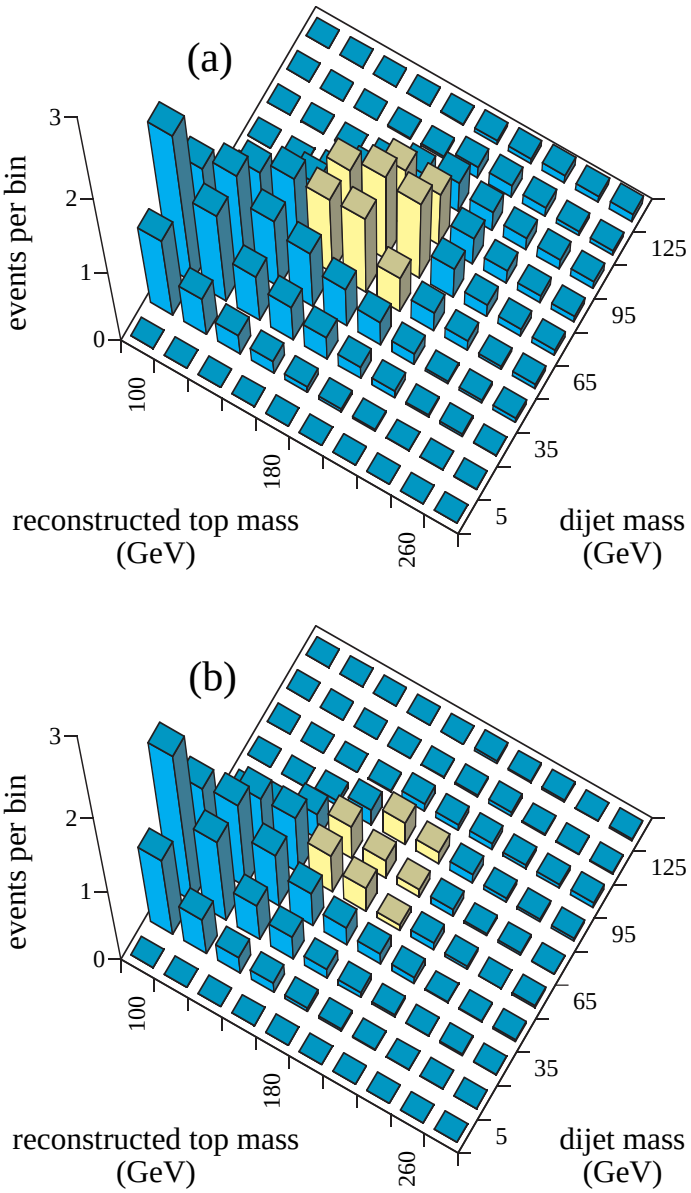


Fig. 19. Lego distributions in reconstructed top quark mass and dijet mass of (a) sum of HERWIG 200 GeV/c^2 top Monte Carlo and background events; (b) background alone. The background includes both W + four jet VECBOS Monte Carlo and QCD multijet data, each normalized using control samples. The events usually increment more than one bin because of multiple solutions; the increments for each event are normalized so that they sum to unity.

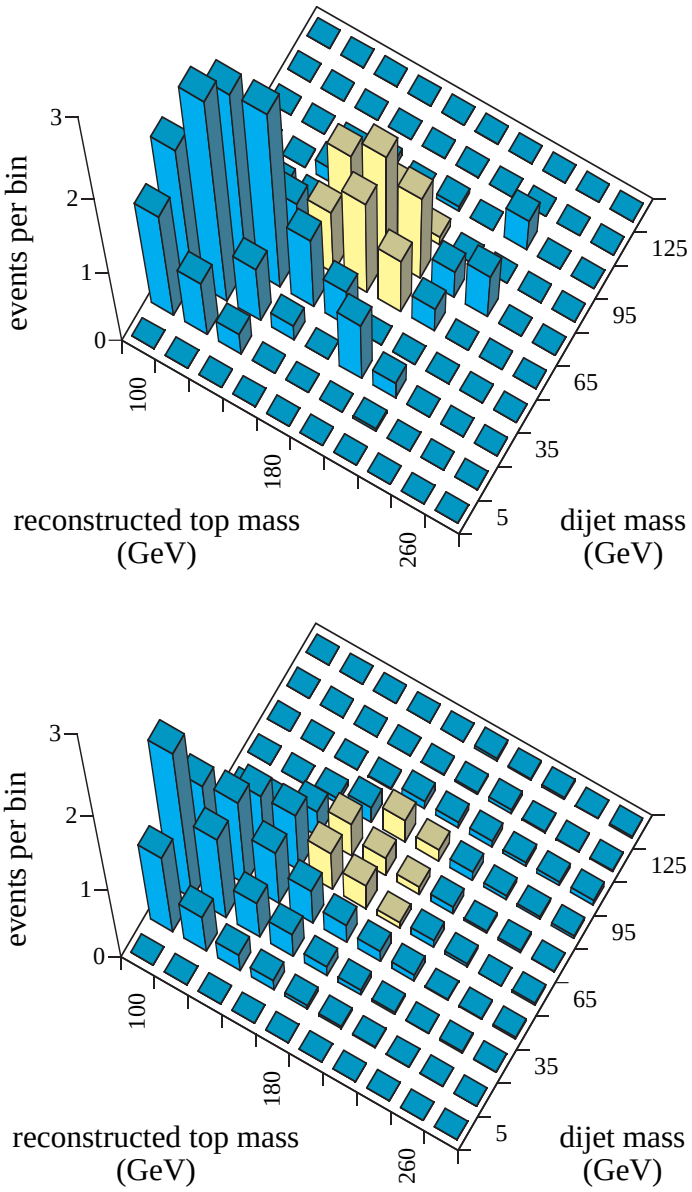


Fig. 20. Lego distribution in reconstructed top quark mass and dijet mass of the data (26 events). The events usually increment more than one bin because of multiple solutions; the increments for each event are normalized so that they sum to unity. The background-only distribution from Fig. 19(b) is reproduced for comparison.

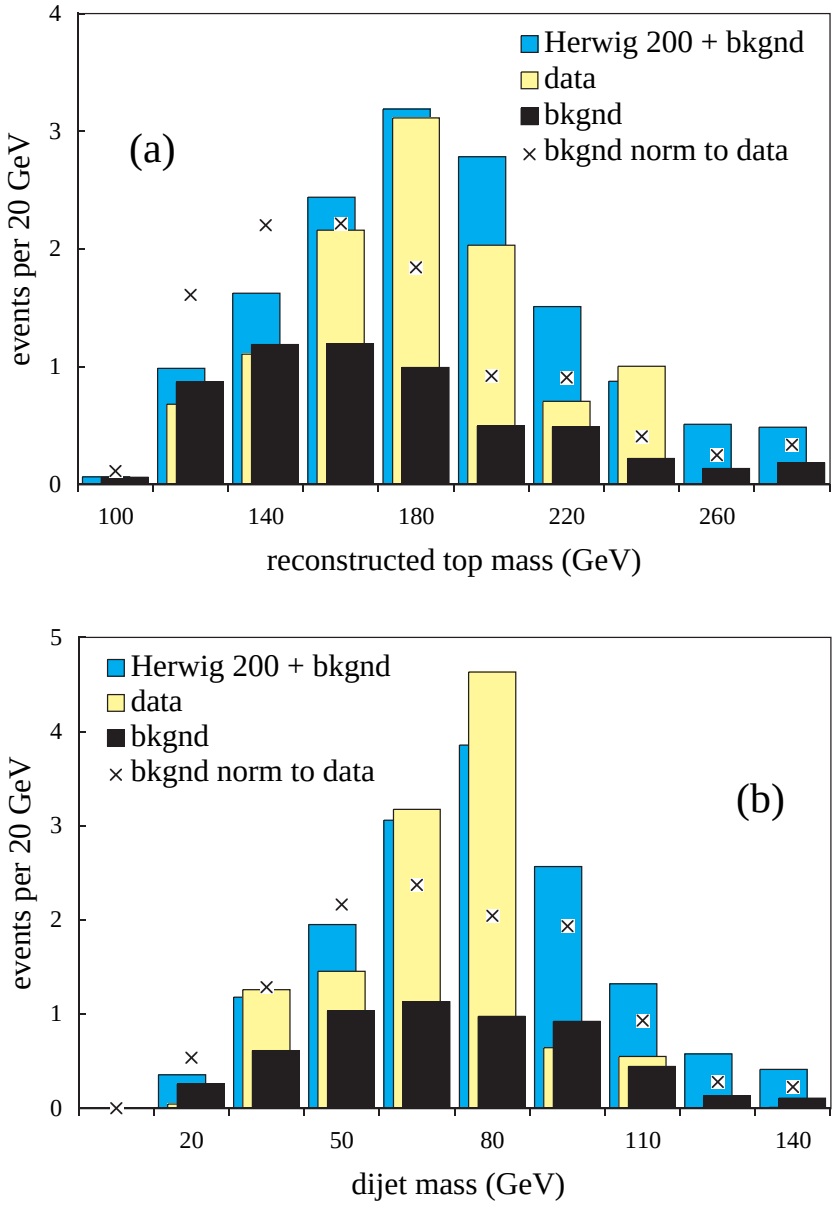


Fig. 21. Distributions of (a) reconstructed top quark mass M_t and (b) dijet mass M_{jj} with (a) $M_{jj} > 58$ GeV/c^2 and (b) $M_t > 150$ GeV/c^2, for (light shaded) data, (medium) sum of background and HERWIG 200 GeV/c^2 top Monte Carlo, (black) background alone, and (x's) background normalized to match the area of the data.

[6] F. Paige and S. Protopopescu, BNL Report No. BNL38034, 1986 (unpublished), release v. 6.89.

[7] F. Carminati *et al.*, "GEANT guide," CERN Library, 1991 (unpublished).

[8] F. A. Berends *et al.*, Phys. Lett. B **357**, 32 (1991).

[9] E. Laenen, J. Smith, and W. van Neerven, Phys. Lett, B **321**, 254 (1994).

[10] G. Marchesini *et al.*, Comput. Phys. Commun **67**, 465 (1992).

[11] P. C. Bhat (DØ Collaboration), *Proceedings of the 8th Meeting of the Division of Particles and Fields of the APS*, Vol. 1 (1994), p. 705; FERMILAB-CONF-95-211-E. To be published in the *Proceedings of the 10th Workshop on $p\bar{p}$ Collider Physics*, Fermilab, Batavia, IL, May 9–13, 1995.

[12] B. Klima (DØ Collaboration), FERMILAB-CONF-95/303-E. To be published in the *Proceedings of the 17th International Symposium on Lepton-Photon Interactions*, Beijing, China, Aug. 10–16, 1995.

[13] W. Giele, E. Glover, and D. Kosower, Nucl. Phys. B **403**, 633 (1993).

[14] S. Snyder (DØ Collaboration), to be published in the *Proceedings of the International Europhysics Conference on High Energy Physics*, Brussels, Belgium, July 1995, and M. Strovink (DØ Collaboration), to be published in the *Proceedings of the 10th Workshop on $p\bar{p}$ Collider Physics*, Fermilab, Batavia, IL, May 9–13, 1995.

Monopoles from Heaven, or S-Duality, and What It Means to You

M. Peskin

Stanford Linear Accelerator Center
Stanford, CA 94309

No paper was received from the author in time for printing.

Precision Electroweak Experiments at SLD

T. Schalk

Institute for Particle Physics
University of California
Santa Cruz, CA 95064

No paper was received from the author in time for printing.

LEP PRECISION ELECTROWEAK MEASUREMENTS FROM THE Z^0 RESONANCE

David Strom*
University of Oregon, Eugene, Oregon 97403

Representing the ALEPH, DELPHI, L3, and OPAL Collaborations, and the LEP Energy Working Group

ABSTRACT

Preliminary electroweak measurements from the LEP Collaborations from data taken at the Z^0 resonance are presented. Most of the results presented are based on a total data sample of 12×10^6 recorded Z^0 events which included data from the 1993 and 1994 LEP runs. The Z^0 resonance parameters, including hadronic and leptonic cross sections and asymmetries, τ polarization and its asymmetry, and heavy-quark asymmetries and partial widths, are evaluated and confronted with the predictions of the Standard Model. This comparison incorporates the constraints provided by the recent determination of the top-quark mass at the Tevatron. The Z^0 resonance parameters are found to be in good agreement with the Standard Model prediction using the Tevatron top-quark mass, with the exception of the partial widths for Z^0 decays to pairs of b and c quarks.

*Supported by DOE Contract DE-FG06-85ER40224.

©1995 by David Strom.

1 Introduction

The LEP measurements of the Z^0 resonance parameters allow precision tests of the Standard Model to be made in a number of ways. The prediction of the top-quark mass from radiative corrections to processes at the Z^0 resonance is perhaps the most notable. The prediction of the top-quark mass from the Z^0 resonance parameters has recently become a test with the discovery and subsequent determination of top-quark mass from the Tevatron.[1,2] The addition of a constraint from the top-mass measurement allows measurements of the Z^0 resonance parameters to provide a first glimpse at what the mass of the Higgs boson might be. Other interesting tests of the Standard Model are made by forming ratios of quantities where radiative corrections which depend on the unknown Higgs mass largely cancel. These ratios, which include $R_{inv} = \Gamma_{inv}/\Gamma_{l^+l^-}$, $R_b = \Gamma_{b\bar{b}}/\Gamma_{had}$, and $R_c = \Gamma_{c\bar{c}}/\Gamma_{had}$, are ideal for searching for physics beyond the Standard Model. In the context of the Standard Model, the ratio $R_\ell = \Gamma_{had}/\Gamma_{l^+l^-}$ provides a precise measurement of the strong coupling constant α_s.

2 Theory Review

This section reviews the essentials of electroweak theory needed for understanding measurements of the Z^0 resonance parameters in unpolarized electron–positron annihilation. A more complete review of electroweak theory is given elsewhere in these proceedings.[3]

At tree level, only three inputs are needed to calculate electroweak quantities. These three inputs are typically taken to be the electromagnetic coupling constant, α_{em}, the Fermi Constant, G_F, and the Z^0 mass, m_Z. Calculations of higher-order corrections require that the masses of the fermions, the mass of the Higgs boson, and the strong coupling constant, α_s, also be known. Almost all radiative corrections involving the light quarks can be absorbed into the value of the electromagnetic coupling constant, α_{em}, by using a "running" value of this constant. (For more details, see Ref. 3.) For this report, we take[4] $\alpha_{em}(m_Z^2) = 1/128.896 \pm 0.090$ where the error (largely due to uncertainties in the measured total cross section for electron–positron annihilation at low energies) is propagated through all of the fit results reported in Sec. 8. The uncertainty on G_F is too small to influence the final results.

After correcting for the purely electromagnetic effects of initial state radiation, and ignoring the effects of final state photon and gluon radiation, the cross section for electron–positron annihilation to fermions at the Z^0 can be written as:

$$\begin{aligned}\frac{2s}{\pi\alpha_{em}(m_Z^2)^2 N_c}\frac{d\sigma}{d\cos\theta} = & \\ & q_f^2\,(1+\cos^2\theta) \\ & -8Re\{\chi(s)q_f(g_{ve}g_{vf}\,(1+\cos^2\theta) \quad + 2g_{ae}g_{af}\cos\theta)\} \\ & +16|\chi(s)|^2[({g_{ve}}^2+{g_{ae}}^2)({g_{vf}}^2+{g_{af}}^2)\,(1+\cos^2\theta) \quad + 8g_{ve}g_{ae}g_{vf}g_{af}\cos\theta],\end{aligned} \tag{1}$$

where q_f is the fermion charge, g_{vf} and g_{af} are the vector and axial vector fermion couplings, N_c is the color factor, $\alpha_{em}(m_Z^2)$ is the value of the electromagnetic coupling constant at the Z^0 resonance, and

$$\chi(s) = \frac{G_F m_Z^2}{8\pi\alpha_{em}(m_Z^2)\sqrt{2}}\frac{s}{s-m_Z^2+is\Gamma_Z/m_Z}, \tag{2}$$

where m_Z and Γ_Z are the mass and width of the Z^0. The first term in Eq. (1) is from photon exchange, the next term from interference between the Z^0 and photon exchange, and the third term from Z^0 exchange. Near the Z^0 peak, the third term dominates; in most of the following, the photon term and the interference term have been set to their Standard Model values.

The effects of radiative corrections can be seen by expressing the partial widths for the Z^0 to fermion pairs as

$$\Gamma_{f\overline{f}} = \frac{G_F m_Z^3}{6\pi\sqrt{2}}({g_{af}}^2+{g_{vf}}^2)(1+\delta_{QCD})N_c(1+\delta_{QED}). \tag{3}$$

Here, δ_{QED} reflects the effects of final state photon radiation, δ_{QCD} the effects of final state QCD corrections, and N_c a color factor. The δ_{QED} is almost negligible, but for quark pairs $\delta_{QCD} = \frac{\alpha_s}{\pi} + ...$ is substantial. The additional radiative corrections can be absorbed into the definition of g_{af} and g_{vf}. At tree level, these quantities are given by

$$\begin{aligned} {g_{vf}}_{tree} &= I_3 - 2q_f\sin^2\theta_W \\ {g_{af}}_{tree} &= I_3, \end{aligned} \tag{4}$$

where q_f and I_3 are the charge and weak isospin of the fermion. If the radiative corrections are absorbed into the definition of g_{af} and g_{vf}, we then *define* the effective value of $\sin^2\theta_W$ for leptons as

$$\sin^2\theta_{eff}^{lept} \equiv \frac{1}{4}(1 - \frac{g_{v\ell}}{g_{a\ell}}). \tag{5}$$

In the context of the Improved Born Approximation,[5] the radiative corrections which affect both g_{vf} and g_{af} can be described by the ρ parameter as follows:

$$\begin{aligned} g_{vf} &= \sqrt{\rho}(I_3 - 2q_f \sin^2\theta_{eff}^{lept}) \\ g_{af} &= \sqrt{\rho}(I_3). \end{aligned} \tag{6}$$

Note that the effect of radiative corrections to the asymmetries are determined by $\sin^2\theta_{eff}^{lept}$, while the total and partial widths of the Z^0 depend primarily on ρ. The effect of the top and Higgs mass on ρ are substantial, making the total width Z^0 sensitive to these quantities. This dependence is illustrated taking

$$\rho = 1 + \Delta\rho_{top} + \Delta\rho_{Higgs} \tag{7}$$

where $\Delta\rho_{top} \simeq 3\frac{G_F}{8\pi^2\sqrt{2}}m_t^2$ and $\Delta\rho_{Higgs} \simeq -11\frac{G_F m_Z^2 \sin^2\theta_w}{24\pi^2\sqrt{2}} \ln\frac{m_H^2}{\cos^2\theta_w m_Z^2}$. New physics may affect the values of $\sin^2\theta_{eff}^{lept}$ and ρ in different ways, making it important to make accurate measurements of both the asymmetries and the partial and total widths of the Z^0.

It is possible to construct quantities which have reduced or altered dependence on the top and Higgs mass by taking the ratio of widths. In the absence of new physics, we can determine α_s from

$$R_l \equiv \frac{\Gamma_{had}}{\Gamma_{l^+l^-}} = \frac{N_{had}}{N_{l^+l^-}}, \tag{8}$$

where l refers to any charged lepton. N_{had} and $N_{l^+l^-}$ are the acceptance-corrected number of multihadrons and lepton pairs. The radiative corrections to this quantity are dominated by the $1 + \delta_{QCD}$ factor in Γ_{had}. The ratio R_l does have a slight top-mass dependence from $\sin^2\theta_{eff}^{lept}$, and from nonuniversal corrections to $\Gamma_{b\bar{b}}$, which are discussed below. The remaining radiative corrections are the same for Γ_{had} and $\Gamma_{l^+l^-}$. Another similar quantity is the total hadronic cross section at the peak of the Z^0 resonance defined by

$$\sigma_{had}^{pole} \equiv \frac{12\pi}{m_Z^2}\frac{\Gamma_{ee}\Gamma_{had}}{\Gamma_Z^2}. \tag{9}$$

This quantity is determined from the acceptance-corrected number of hadronic events, and the measured luminosity determined from small-angle Bhabha scattering. Again, most of the radiative correction to the partial widths cancel in the ratio. Since both Γ_{had} and Γ_Z depend on α_s, the dependence of σ_{had}^{pole} on α_s is 40% of that of R_l.

The quantity with the least dependence on the top and Higgs mass is

$$R_{inv} \equiv \frac{\Gamma_{inv}}{\Gamma_{l^+l^-}} = \frac{\Gamma_Z \; - \; \Gamma_{had} \; - \; 3\Gamma_{l^+l^-}}{\Gamma_{l^+l^-}}, \tag{10}$$

which can be used to search for new "invisible" particles which couple to the Z^0 and to check that the couplings for neutrinos correspond to the Standard Model predictions.

The ratios $R_b = \Gamma_{b\bar{b}}/\Gamma_{had}$ and $R_c = \Gamma_{c\bar{c}}/\Gamma_{had}$ are also good places to look for new physics. R_b has some sensitivity to the top mass; vertex corrections involving top quarks give R_b a quadratic correction of approximately 2%, for a top mass of 180 GeV. Before the measurement of the top-quark mass at the Tevatron, this dependence was useful for determining the top-quark mass without an assumption about the Higgs mass. Now that the top-quark mass has been determined directly, the measurement of R_b constitutes a direct test of the Standard Model.

To probe the ratio of vector to axial vector couplings of the Z^0 to quarks and leptons, the forward–backward asymmetries are measured. The asymmetry due to the Z^0 exchange is given by

$$A_{fb} \equiv \frac{3}{4}\mathcal{A}_e\mathcal{A}_f = \frac{N_{forward} - N_{backward}}{N_{forward} + N_{backward}} \tag{11}$$

where

$$\mathcal{A}_f \equiv \frac{2g_{vf}g_{af}}{(g_{vf})^2 + (g_{af})^2} = \frac{2g_{vf}/g_{af}}{1 + (g_{vf}/g_{af})^2}. \tag{12}$$

The measured asymmetries must be corrected for the residual effects of the $Z^0 - \gamma$ interference, and in the case of electrons, for t-channel γ exchange. The polarization of the outgoing τ leptons may also be used as a probe of the ratio of vector and axial vector couplings of the leptons. For the unpolarized beams available at LEP, the τ polarization is given by

$$P_\tau(\cos\theta) \equiv \frac{\sigma_{right} - \sigma_{left}}{\sigma_{right} + \sigma_{left}} = -\frac{\mathcal{A}_\tau + \mathcal{A}_e\frac{2\cos\theta}{1+\cos^2\theta}}{1 + \mathcal{A}_\tau\mathcal{A}_e\frac{2\cos\theta}{1+\cos^2\theta}}. \tag{13}$$

3 LEP Luminosity and Energy Calibration

The results presented in the following sections are based on data collected at LEP by the ALEPH, DELPHI, L3, and OPAL Collaborations. Various improvements to the LEP machine have increased both the peak luminosity and the general reliability so that in 1994, a typical experiment was able to record 55 pb^{-1} of data. The amount of data taken on and off peak is tabulated below in Table 1. In most cases, the data from 1992 and before has been fully analyzed. However, almost all of the results based on the 1993 and 1994 data are *preliminary*.

Year	On Peak	Off Peak
'90	4 pb^{-1}	3 pb^{-1}
'91	8 pb^{-1}	5 pb^{-1}
'92	24 pb^{-1}	
'93	13 pb^{-1}	18 pb^{-1}
'94	55 pb^{-1}	
Total	104 pb^{-1}	26 pb^{-1}

Table 1: Recorded luminosity for a typical LEP experiment.

The electroweak analysis here benefits from the large amount of data taken at energy points approximately 1.8 GeV above and below the Z^0 resonance. Measurements of the Z^0 mass, m_Z, and the total width of the Z^0, Γ_Z, depend primarily on the amount of data taken off peak, and on the energy calibration of the LEP machine.

The energy calibration of the LEP machine for 1993 has now been finalized,[7] and the error in the LEP energy scale contributes approximately 1.4 MeV to the systematic error on m_Z and 1.5 MeV to the systematic error on Γ_Z. The beam energy is determined by allowing transverse polarization to build up in the electron or positron beam, and then using resonant depolarization to determine the beam energy. The systematic error on individual resonant depolarization measurements is approximately 1.1 MeV.

The measurements of the beam energy must be then transported from the time of the resonant depolarization to the time at which the data were taken. Significant changes in the energy occur due to tidal distortions of the LEP ring. Because the

length of the beam orbit is fixed by the RF system, distortions in LEP's shape will cause the beams to travel slightly off center in quadrupole magnets and shift the beam energy. Because of the strong focusing employed in the LEP machine, a 1 MeV shift in the beam energy corresponds to an average shift in the quadrupoles of only 13 μm. Calculations of the effect of the tidal variation of the beam energy agree well with repeated measurements made in dedicated tide experiments. The results of these experiments are shown in Fig. 1.

Additional variations in the machine energy are seen on much longer time scales. Measurements of the beam orbit indicate that the distortions in the LEP ring are possibly due to changes in the ground water level and the level of Lake Geneva. The extrapolation from the time of the resonant depolarization measurements to the time of data-taking contributes the dominant systematic error in Γ_Z and m_Z. Other effects which are important in the energy calibration can be found in Ref. 7.

The results presented here do not include data taken in the 1995 LEP scan which is expected to include about 18 pb^{-1} per experiment of off-peak data. This scan has more frequent energy calibrations, including some at the start of fills. This will allow a reduction of the systematic error associated with slow distortions in the LEP shape.

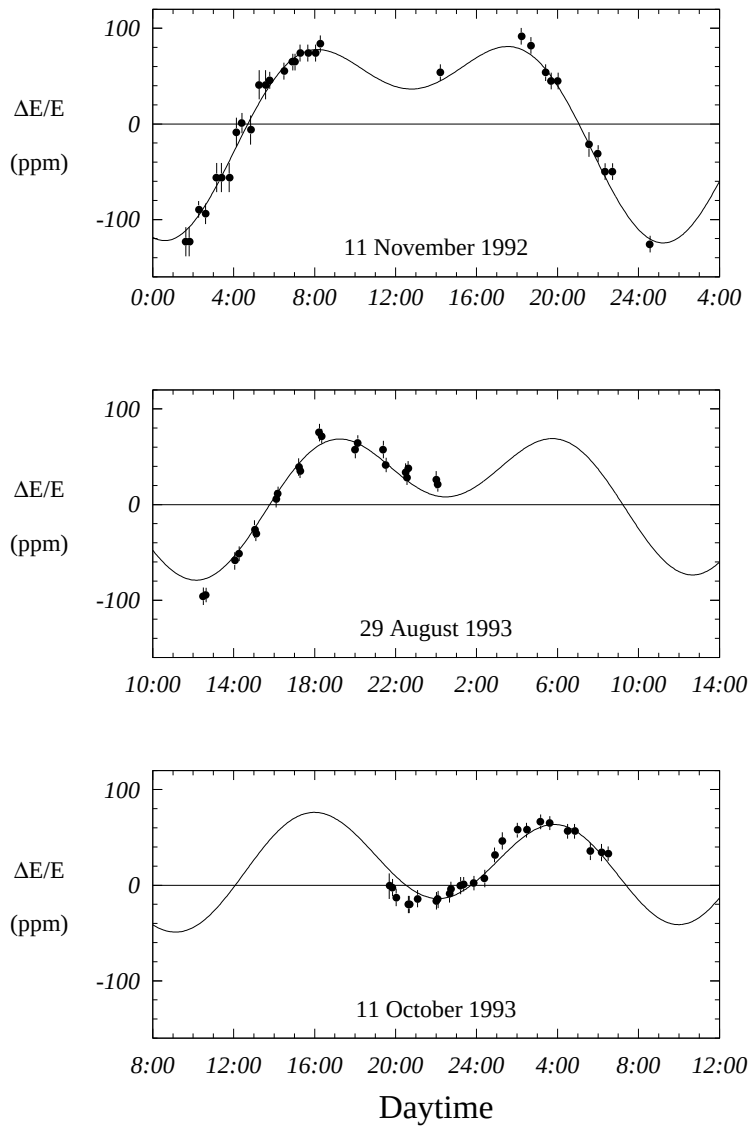


Figure 1: Results of dedicated tide experiments. The beam energy measured with resonant depolarization has been corrected for changes of the integrated dipole field. The solid line shows the prediction of the tidal model. Adapted from Ref. 7.

4 Line-Shape Measurements

The line shape of the Z^0 is measured using decays of the Z^0 to hadrons and leptons. The hadronic decays comprise the largest sample and give us the most information about Γ_Z and m_Z. These events are separated from the beam-related background, the background from the process $e^+e^- \rightarrow e^+e^-X$ and leptonic decays of the Z^0, using a selection based on charged and neutral multiplicity, energy balance, and total energy. The decays of the Z^0 to electron and muon pairs are identified by requiring high total energy and the presence of identified electrons and muons. Decays of the Z^0 to τ pairs are separated from the hadronic decays on the basis of multiplicity, and from the other lepton pairs on the basis of missing energy. Details of the selection procedures used by the LEP Collaborations can be found in Refs. 8–11.

For the determination of m_Z and Γ_Z, it is only necessary to know the relative efficiency and backgrounds of the off-peak energy points to on-peak points. These relative acceptances are generally known with a greater precision than the corresponding statistical errors. For measurement of the absolute cross sections, it is necessary to have an absolute luminosity measurement and a calculation of the absolute acceptances. In general, it has been possible to reduce the systematic error on the luminosity and the hadronic and leptonic acceptances considerably beyond what had been anticipated at the start of the LEP operation.

The experimental and theoretical luminosity measurement has been dramatically improved. Previous to the operation of LEP, luminosity measurements at the 1% level were rare. Now that a second generation of luminosity monitors are in use at LEP, the typical experimental systematic error has been reduced to below 0.1%. The theoretical error is currently 0.16% (Ref. 12), and further improvement is perhaps possible. The detailed breakdown of the experimental systematic errors is given in Table 2. When combining measurements, it has been assumed that the experimental systematic errors in these acceptances are uncorrelated. The leptonic acceptances are 0.15% to 0.8%, and are similar to the corresponding statistical errors. The error in the hadronic acceptance is at the 0.1% level, which is slightly larger than the corresponding statistical error.

The individual LEP Collaborations extract values for the parameter set m_Z, Γ_Z, R_ℓ, σ^{pole}_{had}, and the lepton forward-backward asymmeteries, A_{FB}, by fitting the measured cross section and asymmetries as a function of energy using the program

	ALEPH		DELPHI		L3		OPAL	
	'93 prel.	'94 prel.	'93 prel.	'94 prel.	'93 prel.	'94 prel.	'93 prel.	'94 prel.
$\mathcal{L}^{exp.}$	0.087%	0.116%	0.21%	0.09%	0.12%	0.15%	0.076%	0.079%
σ_{had}	0.073%	0.073%	0.13%	0.15%	0.08%	0.2%	0.15%	0.16%
σ_e	0.50 %	0.48%	0.44%	[a]	0.3%	0.4%	0.23%	0.24%
σ_μ	0.25 %	0.26%	0.28%	0.40%	0.3%	0.6%	0.16%	0.15%
σ_τ	0.34 %	0.32%	0.8%	[a]	0.8%	1.5%	0.43%	0.46%

Table 2: The experimental systematic errors for the analysis of the Z^0 line shape at the Z^0 peak. The errors quoted do not include the common uncertainty due to the LEP energy calibration. The treatment of correlations between the errors for different years is described in Refs. 8-11. Adapted from Ref. 6.
[a]No preliminary result quoted yet.

ZFITTER.[13] For the analysis of the process $e^+e^- \to e^+e^-$, it is also necessary to correct for the photon t-channel. In the fitting procedure, the correlated uncertainties in the LEP energy scale at the various scan points are taken into account. The result of a nine-parameter fit, which does not assume lepton universality, is given in Table 3. The average values for the parameters shown in Table 4 are obtained by taking into account the correlated uncertainties between the experiments (primarily due to uncertainty in the LEP energy scale) and the theoretical uncertainty in the luminosity calculation. Also shown in Table 4 are the average parameters assuming lepton universality. The results for the individual leptonic channels are in good agreement with the assumption of lepton universality. (Because of the nonzero mass of the τ lepton, a 0.2% difference is expected between R_τ and R_ℓ.)

	ALEPH	DELPHI	L3	OPAL
$m_Z\,(\mathrm{GeV})$	91.1924±0.0037	91.1849±0.0034	91.1936±0.0036	91.1852±0.0036
$\Gamma_Z\,(\mathrm{GeV})$	2.4954±0.0057	2.4913±0.0054	2.5022±0.0054	2.4960±0.0053
$\sigma^{pole}_{had}\,(nb)$	41.56±0.09	41.39±0.10	41.48±0.11	41.47±0.10
R_e	20.54±0.11	20.88±0.16	20.89±0.12	20.90±0.10
R_μ	20.88±0.09	20.70±0.09	20.80±0.11	20.796±0.073
R_τ	20.77±0.10	20.61±0.16	20.73±0.17	21.00±0.11
$A^{0,\,e}_{FB}$	0.0196±0.0044	0.0233±0.0070	0.0125±0.0070	0.0081±0.0051
$A^{0,\,\mu}_{FB}$	0.0189±0.0029	0.0166±0.0030	0.0168±0.0038	0.0137±0.0027
$A^{0,\,\tau}_{FB}$	0.0206±0.0039	0.0210±0.0057	0.0287±0.0064	0.0183±0.0035
χ^2/d.o.f.	181/185	151/135	118/138	$10/6^{(a)}$

Table 3: Line-shape and asymmetry parameters from nine-parameter fits to the data of the four LEP experiments. Adapted from Ref. 6.

$^{(a)}$This parameter set has been obtained from a parameter transformation applied to the 15 parameters of the OPAL fit,[11] which treats the γZ^0 interference terms for leptons as additional free parameters. The extra parameters for the γZ^0 interference terms have been fixed to their Standard Model values in the transformation. The χ^2/d.o.f. for the 15-parameter fit to the data is 87/132.

Parameter	Average Value	Average Value with Lepton Universality
m_Z(GeV)	91.1885±0.0022	91.1884±0.0022
Γ_Z(GeV)	2.4963±0.0032	2.4963±0.0032
$\sigma_{had}^{pole}(nb)$	41.488±0.078	41.488±0.078
R_e	20.797±0.058	
R_μ	20.796±0.043	
R_τ	20.813±0.061	
R_ℓ		20.788±0.032
$A_{FB}^{0,e}$	0.0157±0.0028	
$A_{FB}^{0,\mu}$	0.0163±0.0016	
$A_{FB}^{0,\tau}$	0.0206±0.0023	
$A_{FB}^{0,\ell}$		0.0172±0.0012
χ^2/d.o.f.	36/27	39/31

Table 4: Average line-shape and asymmetry parameters from the data of the four LEP experiments given in Table 3. Also shown is the average of the measurements assuming lepton universality. Adapted from Ref. 6.

5 τ Polarization Measurements

The study of the polarization of τ decays in unpolarized collisions provides additional information about the lepton couplings to the Z^0. The momentum and angles of the visible τ decay products are used to obtain the average polarization of the τ.

For two-body decays ($\tau \to h\nu_\tau$), the τ energy spectrum is given by

$$\frac{1}{N}\frac{dN}{dx} \simeq 1 + \xi_s \mathcal{P}_\tau (2x - 1), \tag{14}$$

where $x = E_h/E_{beam}$ and

$$\begin{array}{rcll} \xi_s & = & 1 & \text{spinless hadrons } (\pi,\ \mathrm{K}) \\ \xi_s & = & \frac{m_\tau^2 - 2m_h^2}{m_\tau^2 + 2m_h^2} & \text{spin one hadrons } (\rho,\ a_1). \end{array} \tag{15}$$

When the τ decays to either a ρ or a_1, additional information from the subsequent decay of the hadron is used. For three-body final states from leptonic τ decays, the momentum spectrum is given by

$$\frac{1}{N}\frac{dN}{dx} = \frac{1}{3}\left[(5 - 9x^2 + 4x^3) + \mathcal{P}_\tau(1 - 9x^2 + 8x^3)\right]. \tag{16}$$

The main challenge to the experiments is to devise selection criteria for the various τ decay channels that have a well-understood dependence on momentum, and minimize the contamination from other τ decay modes.

From the correlation between angle and polarization, it is possible to extract both $\mathcal{A}_e$ and $\mathcal{A}_\tau$. From examination of Eq. (13), it can be seen that the value of $\mathcal{A}_e$ is determined from the asymmetry of the τ polarization. The uncertainty on $\mathcal{A}_e$ is limited by available statistics. The individual values of $\mathcal{A}_\tau$ have a sizable contribution from systematic errors which are roughly equivalent to the statistical errors. These systematic errors are not correlated between experiments. The individual measurements from the LEP experiments are described in Refs. 14–17 and are summarized in Table 5, which is taken from Ref. 6.

6 Heavy-Quark Partial Widths and Asymmetries

Measurements of heavy-quark partial widths and asymmetries are based on hadronic decays of the Z^0, where it has been possible to tag one or more jets as

		$\mathcal{A}_\tau$	$\mathcal{A}_e$
ALEPH	('90 - '92) final	$0.136 \pm 0.012 \pm 0.009$	$0.129 \pm 0.016 \pm 0.005$
DELPHI	('90 - '92) final	$0.148 \pm 0.017 \pm 0.014$	$0.136 \pm 0.027 \pm 0.003$
L3	('90 - '94) prel.	$0.152 \pm 0.010 \pm 0.009$	$0.156 \pm 0.016 \pm 0.005$
OPAL	('90 - '94) prel.	$0.134 \pm 0.010 \pm 0.009$	$0.134 \pm 0.015 \pm 0.004$
LEP Average		0.1418 ± 0.0075	0.1390 ± 0.0089

Table 5: LEP results for $\mathcal{A}_\tau$ and $\mathcal{A}_e$.

containing a heavy quark. To measure the asymmetry, it is also necessary to reconstruct the b- or c-quark direction. The original quark-antiquark axis can be estimated from the event thrust axis, the quark direction from the reconstructed quark charge.

In this section, the tagging techniques are briefly described followed by a discussion of the measurement of R_b and R_c. Next, we discuss the asymmetry measurements. Finally, all of the results are combined using a common set of assumptions about the errors introduced from measurements at lower energies, and other needed input such as degree of b-mixing present at LEP. (Several of the topics related to heavy-quark physics are treated in more detail elsewhere in these proceedings.[18])

6.1 Measurements of R_b and R_c

Hadronic decays of the Z^0 to pairs of b and c quarks can be tagged with a variety of techniques. The relatively long b-hadron lifetime of $c\tau \sim 0.45$ mm, and the large mean charged multiplicity of the b hadron, makes tagging techniques based on the identification of detached vertices attractive.

Another useful technique is based on the weak decays of b and c hadrons to final states including leptons. Leptons from b-hadron decays are separated from those from c decays on the basis of the momentum and transverse momentum of the leptons with respect to the jet axis. Since the transverse momentum of the lepton is a measure of the parent's mass, the large b-hadron mass ensures that the leptons at high transverse momentum are likely to be from b hadrons. The

large b-hadron mass is also exploited by event shape, techniques which are based on the difference between the jet structure of b hadrons and the lighter quarks.

Heavy quarks are also tagged by fully or partially reconstructing c hadrons. The measured momentum of the c hadron, as well as decay length and event-shape information, are used to separate the $c \to$ c hadron from cascade process $b \to c \to$ c hadron.

Recent measurements of $R_b = \frac{\Gamma_{b\bar{b}}}{\Gamma_{had}}$ are based on double-tag techniques that reduce the dependence of the analysis on the tagging efficiency. In a tagging method without background, the total number of tagged jets is given by

$$n_t = 2R_b N_{had}\varepsilon, \tag{17}$$

where ε is the tagging efficiency. The number of double-tagged events is

$$n_{tt} = R_b N_{had}\varepsilon^2. \tag{18}$$

Ignoring any correlations between the tagging efficiencies for the jets in the same event, Eqs. (17) and (18) can be solved, giving

$$R_b = \frac{n_t^2}{4n_{tt}N_{had}}. \tag{19}$$

This expression is independent of the ε, which may have large experimental and theoretical uncertainties. In practice, it is necessary to apply a small correction for correlations between the efficiencies of two jets in the same event. In addition, it is necessary to correct n_t and n_{tt} for contamination from light quarks and charm. The correction for charm contamination is the largest contribution to the error on the individual measurements of R_b and gives the measured value of R_b a dependence on the assumed value of R_c. This dependence is parameterized as

$$R_b = R_b^{measured} + a(R_c)\frac{R_c - R_c^{assumed}}{R_c}, \tag{20}$$

with the value of $a(R_c) \simeq -0.15$ for the three measurements which dominate the average.

The values of the three measurements which dominate the LEP average are shown in Fig. 2, and a summary of the correlated systematic errors is given in Table 6. The correlated uncertainty is dominated by the uncertainty in the relative fraction of charm hadrons produced and the decay multiplicity of the charm hadrons.[6,28]

	ALEPH shape (Ref. 19)	ALEPH lifetime (Ref. 20)	DELPHI multiple (Ref. 21)	L3 shape (Ref. 22)	OPAL multiple (Ref. 23)
Charm production	0.0	−0.85	−1.0	0.0	−0.94
D^0 lifetime	0.0	−0.28	−0.2	0.0	−0.23
D^+ lifetime	0.0	−0.36	−0.2	0.0	−0.29
D_s lifetime	0.0	−0.22	−0.2	0.0	−0.17
D decay multiplicity	0.0	−0.57	−0.4	0.0	−0.76
BR(D→K^0)	0.0	0.0	+0.6	0.0	+0.59
$g \to b\overline{b}, c\overline{c}$	0.0	−0.33	−0.2	0.0	−0.46
Long-lived light hadrons	0.0	−0.24	−0.4	0.0	−0.33
BR($c \to \ell$)	+0.6	0.0	−0.2	0.0	−0.28
Semileptonic model $c \to \ell$	−2.1	0.0	−0.2	0.0	−0.25
$\langle x_E(c) \rangle$	+0.8	−0.12	−0.4	+1.8	−0.75
Semileptonic model $b \to \ell$	−1.3	0.0	+0.2	0.0	0.0
$\langle x_E(b) \rangle$	0.0	0.0	0.0	−3.1	0.0
Total corr. error	2.7	1.2	1.5	3.6	1.7

Table 6: Example of breakdown of the correlated systematic error for R_b from lifetime, multiple, and shape double-tag measurements (in units of 10^{-3}). The sign is the sign of the correlation among the experiments. Adapted from Ref. 6.

The value of R_c has been determined using two different techniques. Both of the techniques take advantage of precise data available on the branching ratios for the various decay modes of charm hadrons. One of the methods is based on the measurement of leptons (μ and e) in hadronic events. The other method is based on reconstructed charm hadrons (primarily $D^\star$).

Except for leptons produced in decays and conversions, leptons in hadronic events originate predominately from heavy-quark decays. Leptons from charmed hadrons can be separated from b hadrons on the basis momentum (p), and transverse momentum with respect to the jet axis (p_t). Using models of the rest frame momentum spectra of the b- and c-hadron decays, which are based on low-energy measurements, it is possible to predict the p and p_t spectra of the leptons at LEP energies. The value of R_c is extracted from a "grand" fit to the lepton p and p_t

spectrum which includes the branching ratio $Br(b \to \ell)$ and $Br(b \to c \to \ell)$ and the mean energy carried by primary b and c hadrons. In addition, these fits include the polar angle of the event thrust axis and the charge of the lepton, so that the charm and bottom forward-backward asymmetries and the average b-mixing parameter, $\overline{\chi}$, can be determined.

It is also possible to determine R_c from the measurement of reconstructed charm mesons. Most of the LEP measurements are based on the decay $D^{*\pm} \to D^0\pi^{\pm}$, where the D^0 is fully or partially reconstructed. The transition pion emitted in the $D^{*\pm}$ decay very closely follows the direction of the $D^{*\pm}$ because of the small Q value of the decay. This has two important consequences. First, the mass difference between the reconstructed $D^\star$ and reconstructed D^0 is small, even if some of the decay products of the D^0 are missing. Second, the transition pion very closely follows the jet-axis direction, allowing a charm signal to be isolated using only the transition pions p_t.

Since $D^\star$ mesons are also produced in the b-meson decay, information about the event shape, $D^\star$ decay length, and $D^\star$ momentum are all used to extract the fraction of $D^\star$ production due to $Z^0 \to c\bar{c}$. The largest external systematic error in this procedure comes from the uncertainty in the expected production of $D^\star$ in $Z^0 \to c\bar{c}$. This production rate has been taken from measurements in lower energy e^+e^- and has also been determined from double-tag techniques at LEP.

The OPAL double-tag technique uses an identified charm decay in one jet to produce a $Z^0 \to c\bar{c}$ sample, and then exploits the characteristic transverse momentum spectrum of the transition pions in the opposite jet to determine the inclusive branching ratio of $c \to D^{\star+}$. There is good agreement between the parameters measured at LEP and those at lower energy. For example, the OPAL value of

$$Br(c \to D^\star)Br(D^\star \to D^0\pi)Br(D^0 \to K\pi) = (6.47 \pm 0.75) \times 10^{-3}$$

agrees with the value obtained at the low energy of

$$Br(c \to D^\star)Br(D^\star \to D^0\pi)Br(D^0 \to K\pi) = (7.1 \pm 0.5) \times 10^{-3}.$$

The DELPHI double-tag technique measurement uses a cut on the transverse momentum spectrum of the single particles to produce single and double tags in the same way as a decay length cut is used to produce single and double b-tagged samples.

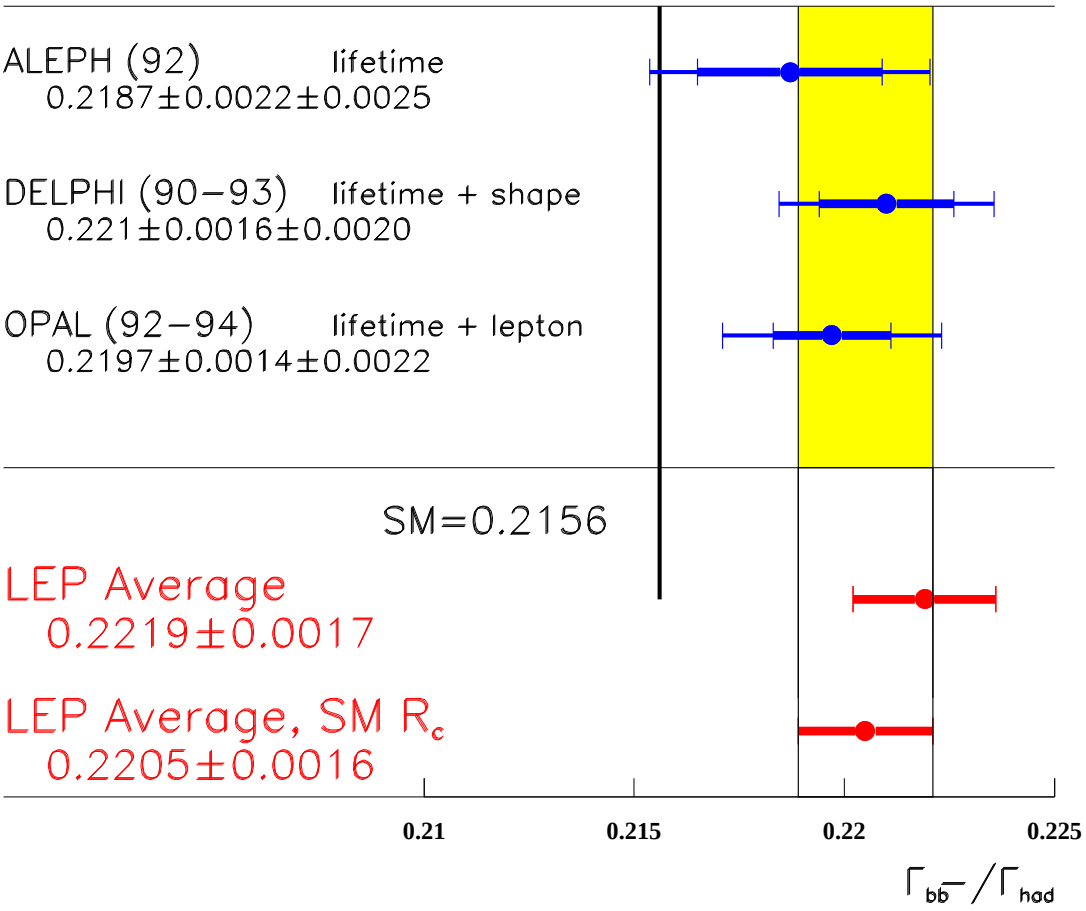


Figure 2: The individual measurements of R_b which dominate the average. The LEP average is shown with and without the Standard Model constraint of $R_c = 0.172$. The unconstrained average value incorporates the results of all of the heavy-flavor measurements and the LEP R_c analysis as described in the last portion of this section.

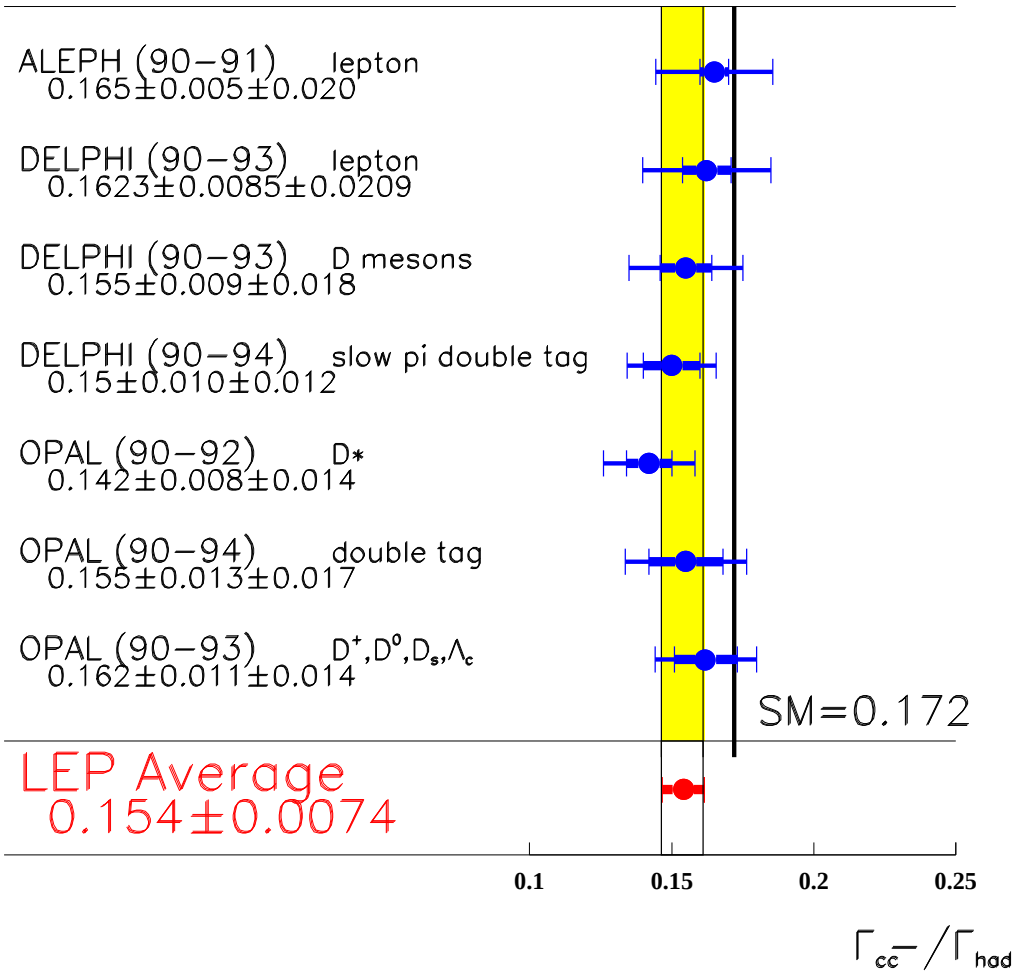


Figure 3: Individual measurements of R_c. The average value incorporates the results of the other LEP heavy flavor results as described in the last portion of this section.

The results of the measurements are summarized in Fig. 3. The resulting average value is somewhat smaller than the Standard Model prediction, and none of the individual measurements are larger than the Standard Model.

6.2 Heavy Quark Forward-Backward Asymmetries

The forward–backward asymmetry can be determined for $Z^0 \to b\bar{b}$ and $Z^0 \to c\bar{c}$ by tagging the events using methods similar to those employed for the measurement of R_b and R_c. To determine the direction of the event axis, three techniques are used: lepton charge, jet charge, and reconstruction of a charm hadron.

In measurements based on identified leptons in hadronic events, the sign of the lepton is used to infer the sign of the original quark. For example, a b quark will produce a b hadron, which will then decay to a negatively charged lepton, i.e., $b \to W^- c \to \ell^- \nu c$. Similarly, a c quark, which produces a c hadron, will then decay to a positively charged lepton. The measurement of the forward-backward asymmetry is complicated by both the mixing of neutral b mesons, which can cause the final state lepton to have the wrong sign, and the process $b \to c \to \ell^+$ which also produces a lepton of the wrong sign. Fortunately, both of these effects can be measured in the data. The average value of b mixing at LEP, $\overline{\chi}$, has been extracted from the relative number of same-sign and opposite-sign leptons in opposite hemispheres. (More details are given in Ref. 27.) The value of $Br(b \to c \to \ell^+)$ has similarly been extracted from the rate of opposite-signed leptons in the same hemisphere.

Leptons from $b \to \ell$ and $c \to \ell$ can be statistically separated from each other on the basis of p and p_t. The asymmetry is usually determined via a combined maximum likelihood fit to the p, p_t, and angular distribution of the data. These measurements depend on both the measured values of R_b and R_c as well as $\overline{\chi}$ and the assumed semileptonic branching ratios. The resulting correlations are all taken into account in determining the LEP average.

One of the disadvantages of the lepton-based technique is the relatively small branching ratio of b hadrons to leptons. In order to use tagging techniques based on lifetime, it is necessary to use jet charge to determine which jet is associated with the quark and which is associated with the antiquark. This method relies on a calculation of the average jet charge which is usually determined from a momentum-weighted average of the signs of charged tracks observed in a jet. The

momentum weighting exploits the fact that stiff tracks have a larger correlation with the sign of the quark charge than soft tracks.

At first sight, this method appears to depend completely on the Monte Carlo prediction of the reconstructed jet charge. However, the correlation between the charges of two jets (or hemispheres) can be measured to check the validity of the Monte Carlo. Similar methods that have been used for b-mixing studies are described elsewhere in these proceedings.[27]

The forward-backward asymmetries of $Z^0 \to b\bar{b}$ and $Z^0 \to c\bar{c}$ can also be measured using events tagged with charm hadrons. Most measurements rely on the reconstructed charge of a $D^\star$ meson to deduce the charge of the initial heavy quark. Event shape, lifetime, and the momentum of the charm hadron are used to separate b- and c-quark final states. The reconstruction of the b asymmetry from the charm mesons utilizes the decay chain $b \to c \to D^{\star +}$ to determine the direction of the jet axis.

6.3 Combined Fit

It is apparent from the discussion in this section that the results of the heavy-quark measurements are dependent on a common set of uncertainties which are both internal and external to the LEP electroweak measurements. To determine the LEP values, a χ^2 minimization technique is used where the internal parameters are determined directly and the external parameters (such as the charged multiplicities of D^+ decays and c- and b-hadron lifetimes) are allowed to vary within their measured constraints. Additional details about the LEP heavy-flavor fit can be found elsewhere.[28] The results of a four-parameter fit is

$$\begin{aligned} R_b &= 0.2219 \pm 0.0017 \\ R_c &= 0.1543 \pm 0.0074 \\ A_{FB}^{0,b} &= 0.0999 \pm 0.0031 \\ A_{FB}^{0,c} &= 0.0725 \pm 0.0058, \end{aligned} \tag{21}$$

where the values of A_{FB} have been translated to the Z^0 pole and corrected for radiative QCD and QED effects. The Standard Model values of the γ exchange and $\gamma - Z$ have also been assumed. The measured values of R_b and R_c have a -0.34 correlation because of the dependence of the R_b measurement on R_c. While the values of forward-backward asymmetries are in agreement with the Standard Model prediction (see Fig. 4 and Table 10) which incorporates the Fermilab top

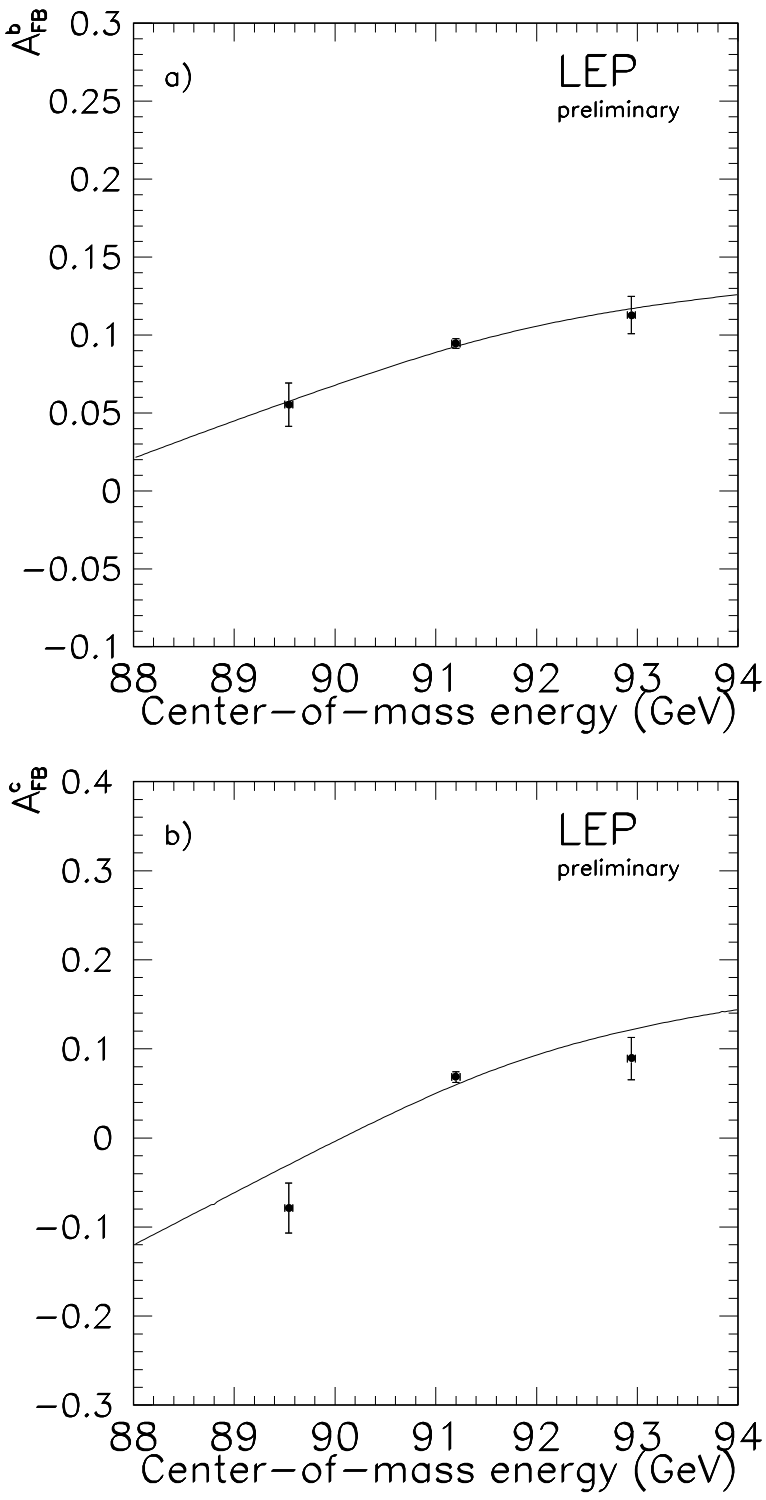


Figure 4: Forward-backward asymmetry for (a) $Z^0 \to b\bar{b}$ and (b) $Z^0 \to c\bar{c}$ as a function of center-of-mass energy. The curve shows the Standard Model prediction.

mass value of 180 ± 12, the measured values of R_b and R_c do not agree with the Standard Model prediction, as can be seen from Fig. 5. If Gaussian errors are assumed, the LEP measurements only agree with the Standard Model at the 0.1% confidence level. Since the LEP measurements of R_b are dominated by systematic error, the assumption of Gaussian errors may not be justified. On the other hand, it is worth noting that it difficult to obtain agreement with the Standard Model by adjusting only one of the quantities in Table 6. The largest common systematic error is due to uncertainties in the charm production and depends primarily on the production rate of the long-lived D^+ mesons in $Z^0 \to c\bar{c}$ events. This leads to a common uncertainty in the more precise measurements of R_b of approximately 0.001. To obtain agreement between the LEP measurements and the Standard Model, it would be necessary to change the production of D^+ by more than six times its uncertainty. Note that the D^+ rate depends primarily on the $D^\star$ rate production. This is because the decay $D^{*0} \to D^+\pi^-$ is not kinematically allowed, while the corresponding decay $D^{*+} \to D^0\pi^+$ occurs with a large branching ratio. We have already seen that the LEP measurements of the $D^{\star+}$ production are in good agreement with the lower energy measurements giving us confidence that the inputs to the R_b analysis have sensible values.

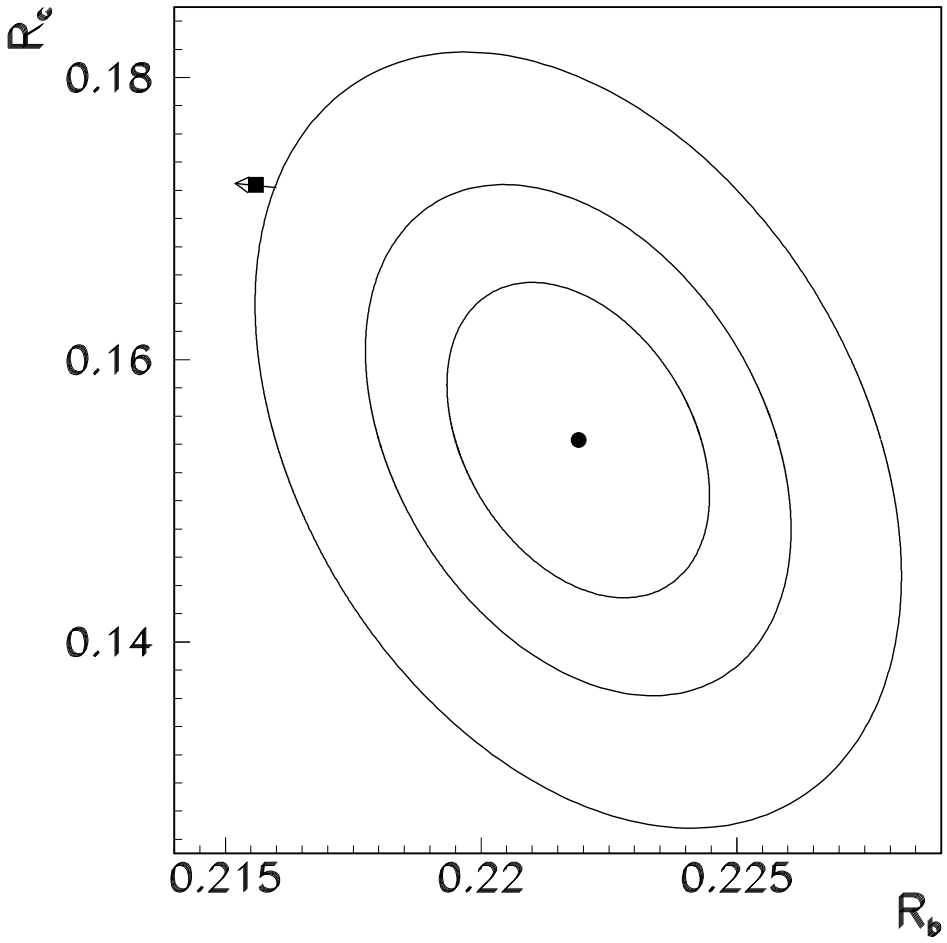


Figure 5: Contours in the R_b-R_c plane corresponding to 68%, 95%, and 99.9% confidence levels. The Standard Model prediction for $m_t = 180 \pm 12$ GeV is also shown. The arrow point shows the direction of increasing top mass.

7 Hadronic Forward-Backward Asymmetry

The jet-charge technique used in conjunction with the forward-backward asymmetry measurements discussed above can also be used to measure the hadronic forward-backward asymmetry averaged over quark flavors. These measurements are primarily sensitive to radiative corrections which affect the ratio of the vector to axial vector coupling of the electron. These measurements are reported in terms of the quantity $\sin^2\theta^{lept}_{eff}$ as defined in Eq. (5). Recall that the forward-backward asymmetry depends on the product $\mathcal{A}_e\mathcal{A}_f$ [see Eq. (11)]. For quark final states, most of the sensitivity radiative corrections come through $\mathcal{A}_e$ which has a greater dependence on $\sin^2\theta^{lept}_{eff}$ than $\mathcal{A}_f$.

Experiment		$\sin^2\theta^{lept}_{eff}$
ALEPH[33]	'90-'93 prel.	$0.2323 \pm 0.0010 \pm 0.0010$
DELPHI[34]	'90-'91	$0.2345 \pm 0.0030 \pm 0.0027$
OPAL[35]	'91-'94 prel.	$0.2326 \pm 0.0012 \pm 0.0013$
Average		0.2325 ± 0.0013

Table 7: Summary of $\sin^2\theta^{lept}_{eff}$ measurements from the inclusive hadronic charge asymmetries at LEP. The first error is statistical, the second systematic. Adapted from Ref. 6.

The values of $\sin^2\theta^{lept}_{eff}$ determined using this technique are shown in Table 7. The systematic errors are dominated by the uncertainty in the hadronization process. A more complete discussion of the systematic errors which affect these measurements and the correlation of these errors with the jet-charge method used to determine $A^{0,b}_{FB}$ is given in Ref. 6.

8 Discussion and Combined Results

In this section, we combine the LEP measurements and compare them to the predictions of the Standard Model and the electroweak measurements made at the Tevatron and the SLC. We begin with a discussion of the lepton and quark couplings. Then all of the measurements are used to derive constraints on the top and Higgs masses.

8.1 Lepton Couplings

The axial and vector couplings of the leptons can be extracted from measurements of the lepton cross sections, the lepton asymmetries, and the tau polarization. The lepton cross sections are used to obtain the leptonic width,* $\Gamma_{l^+l^-}$, which is related to the lepton couplings [Eq. (3)]. The other quantities are used to determine $\mathcal{A}_f$ which is related to the lepton couplings [Eq. (12)].

The values derived from the LEP measurements[6] are given in Table 8 and displayed in Fig. 6. There is good agreement between the LEP measurements and the Standard Model predictions which use the Tevatron top mass. The LEP values are also consistent with the constraint from $\mathcal{A}_{LR}$ as measured by SLD,[39] which is also displayed in Fig. 6.

8.2 Neutrino Couplings

Decays of the Z^0 to neutrinos are not detected directly in the LEP detectors, making asymmetry measurements of these decays impossible. It is, however, possible to derive the quantity R_{inv} from the lepton cross section and the ratio R_ℓ. Taking

$$R_{inv} = \frac{\Gamma_{inv}}{\Gamma_{l^+l^-}} = \frac{\Gamma_Z - \Gamma_{had} - 3\Gamma_{l^+l^-}}{\Gamma_{l^+l^-}} \tag{22}$$

and using the relationship $\sigma_{ll}^{pole} = \frac{12\pi}{{m_Z}^2}(\frac{\Gamma_{l^+l^-}}{\Gamma_Z})^2$, we have

$$R_{inv} = \left(\frac{12\pi}{m_Z^2}\frac{1}{\sigma_{ll}^{pole}}\right)^{\frac{1}{2}} - R_\ell - 3. \tag{23}$$

*Note that the quantity $\sigma_{ll}^{pole} \equiv \frac{12\pi}{m_Z^2}\frac{\Gamma_{l^+l^-}^2}{\Gamma_Z^2}$ can be obtained from the LEP parameter set via the relationship $\sigma_{ll}^{pole} = \frac{\sigma_{had}^{pole}}{R_\ell}$.

	Without Lepton Universality
g_{ve}	-0.0368 ± 0.0017
$g_{v\mu}$	-0.0370 ± 0.0041
$g_{v\tau}$	-0.0371 ± 0.0018
g_{ae}	-0.50115 ± 0.00052
$g_{a\mu}$	-0.50113 ± 0.00076
$g_{a\tau}$	-0.50151 ± 0.00089
	Ratios of Couplings
$g_{v\mu}/g_{ve}$	1.01 ± 0.14
$g_{v\tau}/g_{ve}$	1.008 ± 0.071
$g_{a\mu}/g_{ae}$	1.0000 ± 0.0018
$g_{a\tau}/g_{ae}$	1.0007 ± 0.0020
	With Lepton Universality
$g_{v\ell}$	-0.0369 ± 0.0010
$g_{a\ell}$	-0.50119 ± 0.00041

Table 8: Results for the effective vector and axial vector couplings derived from the combined LEP data with and without the assumption of lepton universality. Adapted from Ref. 6.

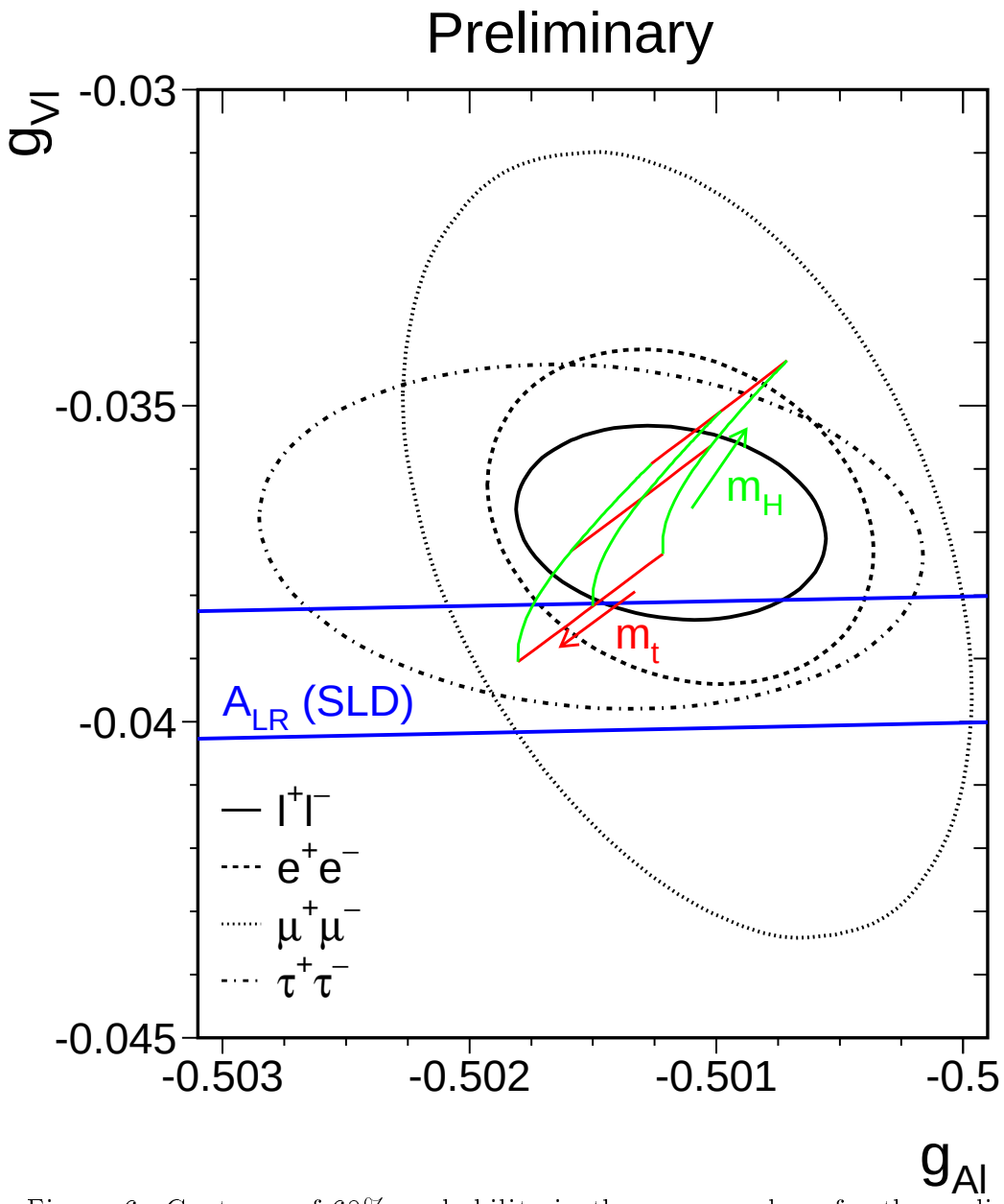


Figure 6: Contours of 68% probability in the $g_v - g_a$ plane for the preliminary LEP measurements. The constraint from the SLD $\mathcal{A}_{LR}$ measurement appears as a band in the figure. The Standard Model prediction for $m_t = 180 \pm 12\ GeV$ and for $60 < m_H < 1000\ GeV$. The arrows point in the direction of increasing top and Higgs mass. Adapted from Ref. 6.

	LEP	SLD	Standard Model
$\mathcal{A}_b$	0.910 ± 0.037	0.841 ± 0.053	0.935
$\mathcal{A}_c$	0.660 ± 0.056	0.606 ± 0.090	0.667

Table 9: Comparison of LEP and SLD measurements with the Standard Model value of $\mathcal{A}_b$ and $\mathcal{A}_c$. The Standard Model prediction is for a top mass of 178 GeV and a Higgs mass of 300 GeV.

Using the LEP line-shape data presented in Sec. 4,

$$R_{inv} = 5.956 \pm 0.031$$

is obtained. This is in good agreement with the Standard Model prediction of 5.973 ± 0.001, where the small error on Standard Model prediction corresponds to $m_t = 180 \pm 12$ GeV and $60 < m_H < 1000$ GeV. Expressed as the number of neutrino generations which couple to the Z^0, LEP obtains

$$N_\nu = 2.991 \pm 0.16.$$

To derive a constraint on the neutrino couplings, we assume that $g_{v\nu} = g_{a\nu} = g_\nu$ and obtain $g_\nu = +0.5011 \pm 0.0013$. The sign of g_ν is not determined from the LEP data and is taken from neutrino scattering measurements.[38]

8.3 Quark Couplings

The couplings of b and c quarks to the Z^0 can be determined from the measurements of R_b and R_c as well as the heavy-quark forward-backward asymmetries. The measured values of the asymmetries depend on the product $\mathcal{A}_e\mathcal{A}_f$ [see Eq. (11)]. To find $\mathcal{A}_f$, we assume lepton universality and use the LEP lepton asymmetries and τ polarization. The resulting value is $\mathcal{A}_e = 0.1464 \pm 0.0039$. The values of the LEP measurement are compared to the Standard Model prediction in Table 9. Also shown in the table are the measurements of $\mathcal{A}_b$ and $\mathcal{A}_c$ made by SLD using the forward–backward polarized asymmetries.[36]

It is possible to derive the quark couplings by adding the constraint provided by the measurement of R_b and R_c. We convert R_b and R_c to partial widths using the hadronic width of the Z^0 derived from the line-shape fit of 1744.8 ± 3.0 GeV.

In addition, the world average value[37] of $\alpha_s = 0.117 \pm 0.006$ is used to compute the QCD corrections. The resulting contours are shown in Fig. 7.

8.4 Top and Higgs Mass

All of the LEP electroweak observables presented above can be used in a fit to determine, in the context of the Standard Model, the values of α_s and m_t. The values of the quantities input to the fit are shown in Table 10 as well as the Standard Model values for the quantities. The result of the fit is

$$\begin{aligned} m_t &= 170 \quad \pm 10 \quad {}^{+17}_{-10} \quad GeV \\ \alpha_s &= 0.125 \quad \pm 0.004 \quad \pm 0.002, \end{aligned}$$

where the χ^2 is 18 for nine degrees of freedom. The central value is for $m_H = 300$ GeV, and the second error reflects the variation in the result as the Higgs mass is varied in the interval $60 < m_H < 1000$ GeV. The LEP value is in excellent agreement with the direct determination of $m_t = 180 \pm 12$ GeV made at the Tevatron. The value of α_s is slightly higher than the world average, adjusted to remove input from the Z^0 line shape, of 0.117 ± 0.006 (Ref. 37). The poor quality of the χ^2 is primarily due to the large discrepancy between the measured value of R_b, R_c, and the Standard Model prediction. Recall that R_b has only a moderate dependence on the top mass and R_c is almost independent of the top mass. As a result, refitting without R_b and R_c increases the value obtained for the top mass by only 4 GeV.

The dependence of the m_t prediction on the assumed Higgs mass shows that the LEP measurements will have some sensitivity to the Higgs mass if an external measurement of m_t is added to the fit. Combining the LEP measurements with the Tevatron top-mass measurement[1,2] of $m_t = 180 \pm 12$ GeV, the change in χ^2, $\Delta\chi^2$, as a function of Higgs mass, reveals the logarithmic dependence of the electroweak observables to the Higgs mass as shown in Fig. 8.

We can improve our estimate of the Higgs mass by including the electroweak measurements made at SLD with polarized beams,[39] the value of $1 - m_W^2/m_Z^2$ from neutrino scattering,[40–42] and the Tevatron W mass determination[43] summarized in Table 11. Adding these constraints to the Standard Model fit, $m_t = 178 \pm 8^{+17}_{-20}$ GeV and $\alpha_s = 0.123 \pm 0.004 \pm 0.002$ for $m_H = 300$ GeV are obtained. Figure 8 also shows the $\Delta\chi^2$ curves for the full electroweak set. The combined data favor relatively low values of the Higgs mass, but the data is not yet precise

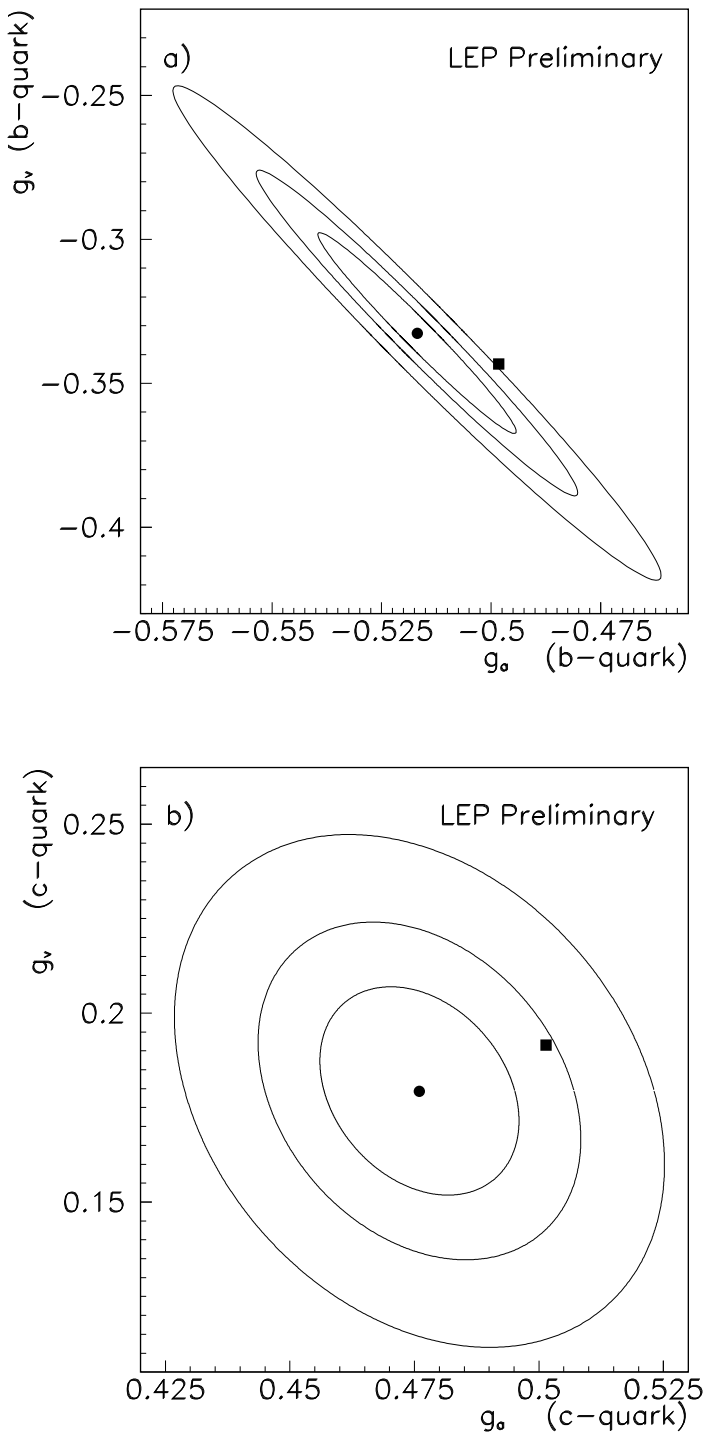


Figure 7: Contours of 68%, 95%, and 99.9% probability in the $g_v - g_a$ plane for the preliminary LEP measurements of (a) b-quark couplings and (b) c-quark couplings. The solid square shows the Standard Model prediction for $m_t = 180$ GeV and $m_H = 300$ GeV.

enough as to exclude a Standard Model Higgs at the upper range of its allowed value. At LEP, direct searches for the Standard Model Higgs have excluded it in the mass region below 60 GeV (Ref. 44).

The measurement of the Γ_Z, R_ℓ, and σ_{had}^{pole} provide additional constraints on possible deviation in R_b and R_c from physics beyond the Standard Model. One way to use these constraints is to allow for an additional contribution to $\Gamma_{b\bar{b}}$ by taking $\Gamma_{b\bar{b}} = \Gamma_{b\bar{b}}^{SM} + \Delta_{b\bar{b}}$. Fitting to all of the data, $\Delta_{b\bar{b}} = 11.7 \pm 3.8 \pm 1.4$ MeV with $m_t = 181 \pm 8^{+17}_{-19}$ GeV and $\alpha_s = 0.102 \pm 0.008$ is obtained. The resulting value of α_s is in agreement with the world average, but much lower than the value obtained from the line-shape measurements reported above. Allowing for deviation in $\Delta_{c\bar{c}}$ in $\Gamma_{c\bar{c}}$, a value of $\alpha_s = 0.18 \pm 0.04$ is obtained. This value of α_s is only consistent with the world average at the two standard-deviation level.

	Measurement	Standard Model Fit	Pull
line-shape and lepton asymmetries:			
m_Z [GeV]	91.1884 ± 0.0022	91.1882	0.1
Γ_Z [GeV]	2.4963 ± 0.0032	2.4973	-0.3
σ_{had}^{pole} [nb]	41.488 ± 0.078	41.450	0.5
R_ℓ	20.788 ± 0.032	20.773	0.5
$A_{FB}^{0,\ell}$	0.0172 ± 0.0012	0.0159	1.1
τ polarization:			
$\mathcal{A}_\tau$	0.1418 ± 0.0075	0.1455	-0.5
$\mathcal{A}_e$	0.1390 ± 0.0089	0.1455	-0.7
b and c quark results:			
$R_b{}^{(a)}$	0.2219 ± 0.0017	0.2156	3.7
$R_c{}^{(a)}$	0.1543 ± 0.0074	0.1724	-2.5
$A_{FB}^{0,b\,(a)}$	0.0999 ± 0.0031	0.1020	-0.7
$A_{FB}^{0,c\,(a)}$	0.0725 ± 0.0058	0.0728	0.0
$q\overline{q}$ charge asymmetry:			
$\sin^2\theta_{eff}^{lept}$ ($\langle Q_{FB} \rangle$)	0.2325 ± 0.0013	0.23172	0.6

Table 10: The LEP measurements used in the combined Standard Model fit for m_t and α_s. The Standard Model fit results and the pulls (difference between measurement and fit in units of the measurement error) are derived from the fit including data from the Tevatron and SLD for a fixed value of $m_H = 300$ GeV. The full fit includes the correlation matrices given in Ref. 6.

	Measurement	Standard Model Fit	Pull
(a) SLD			
$\sin^2\theta^{lept}_{eff}$ ($\mathcal{A}_{LR}$) [Ref. 39]	0.23049 ± 0.00050	0.23172	-2.5
R_b [Ref. 36]	0.2171 ± 0.0054	0.2156	0.3
$\mathcal{A}_b$ [Ref. 36]	0.841 ± 0.053	0.935	-1.8
$\mathcal{A}_c$ [Ref. 36]	0.606 ± 0.090	0.667	-0.7
(b) $p\overline{p}$ and νN			
m_W [GeV] ($p\overline{p}$) [Ref. 43]	80.26 ± 0.16	80.35	-0.5
$\sin^2\theta_W$ (νN) [Ref. 40–42]	0.2257 ± 0.0047	0.2237	0.4

Table 11: (a) SLD results for $\sin^2\theta^{lept}_{eff}$ from the measurement of the left-right polarization asymmetry, for R_b and for $\mathcal{A}_b$ and $\mathcal{A}_c$ from polarized forward-backward asymmetries. (b) Electroweak precision tests from $p\overline{p}$ colliders and νN-scattering. Correlations between the systematic errors of the SLD heavy-quark measurements and of the LEP measurements have been included in the fit. Adapted from Ref. 6.

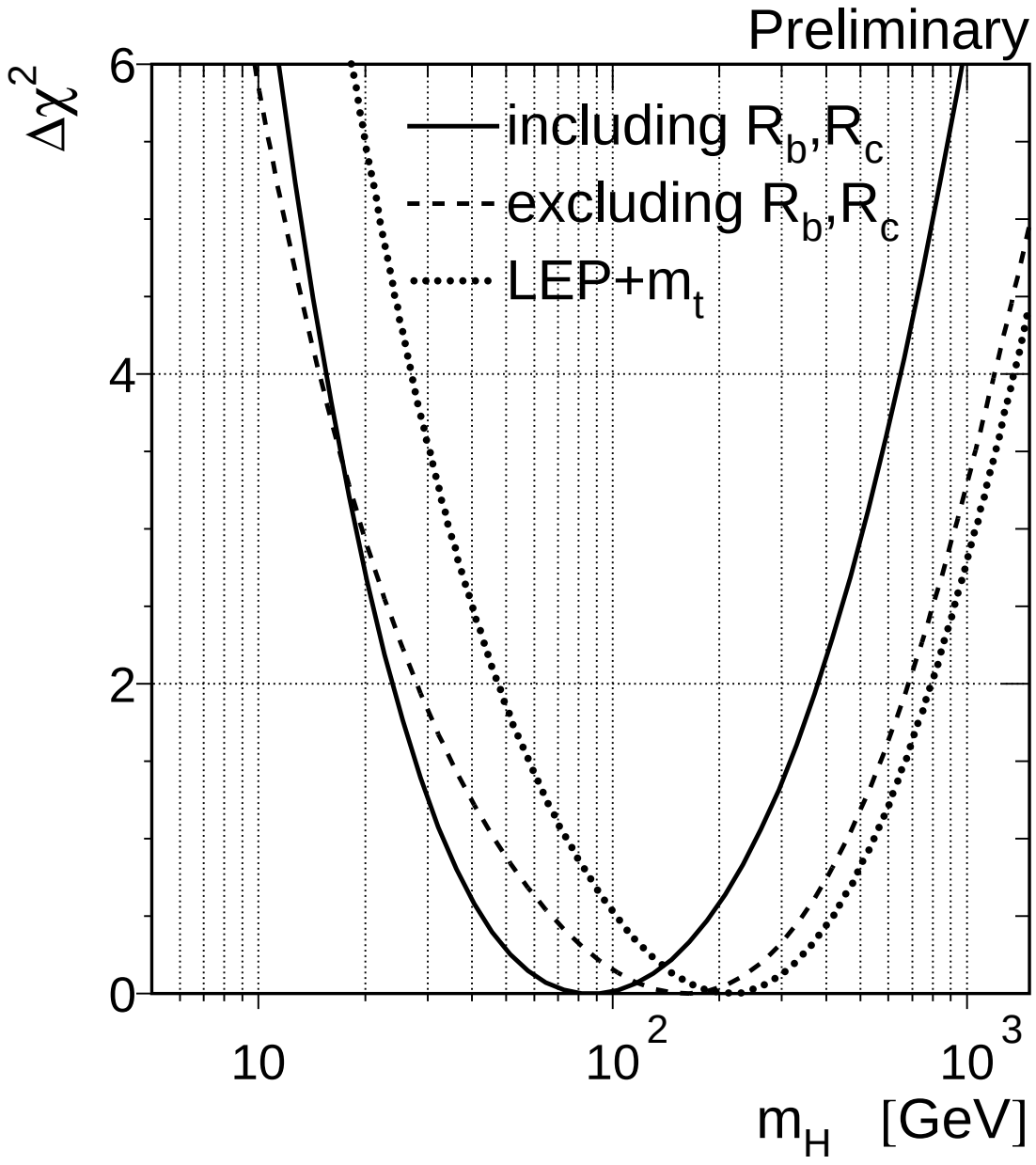


Figure 8: Sensitivity of the electroweak observables to the Higgs mass for LEP and m_t and for all electroweak data. The latter result is also shown when R_b and R_c are excluded from the fit. Note that $\Delta\chi^2$ is plotted.

9 Conclusion and Outlook

Almost all of the precision electroweak measurements which have been made at LEP I are in excellent agreement with the predictions of the Standard Model and the recent measurement of the top-quark mass by DØ (Ref. 2) and CDF (Ref. 1). The preliminary measurements of R_b and R_c, however, show a marked departure from the prediction of the Standard Model.

Improvement in the statistical and systematic errors on many of the measurements presented here can be expected as the LEP experiments complete the final analysis of the 1994 data. Improvements in the statistical errors on R_b and R_c are expected as two of the LEP experiments have not yet analyzed the large amount of data delivered in 1994. Improvements in the systematic errors on these quantities can also be expected as the experiments endeavor to reduce the dependence of their measurements on external measurements, and as the external measurements become more precise.

Data from the 1995 run, the last LEP run at the Z^0, are presently being analyzed. Because of the additional data taken off the Z^0 peak, the addition of the 1995 data will improve our knowledge of Γ_Z.

Starting in the Fall of 1995, LEP will begin running well above the Z^0 at $\sqrt{s} \sim 140$ GeV. In 1996, it is hoped that the W threshold will be reached and a first measurement of the W mass can be made. One goal of the LEP II physics program is to acquire sufficient data to make a measurement of the W mass at the 40 MeV level.

Another contribution of LEP II to electroweak measurements will be to extend the reach of the Higgs search from the present limit near 60 GeV to m_Z or higher, depending on the final center-of-mass energy of the LEP II machine.[44]

Acknowledgments

I would like to thank the LEP Collaborations ALEPH, DELPHI, L3, and OPAL for making their latest physics results available in time for the SLAC Summer Institute. I received invaluable assistance from D. Schaile and P. Wells in making the combined fits which were presented at the Institute. In this report, I have included a few results which were not available in time for the Institute. I have relied heavily on the work of the LEP Electroweak Working Group which is sum-

marized in Ref. 6. I would also like to thank the organizers of the Institute for providing a well-organized and stimulating program.

References

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[2] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[3] M. Swartz, these proceedings.

[4] S. Eidelmann and F. Jegerlehner, *Hadronic Contributions to (g-2) of the Leptons and to the Effective Fine Structure Constant* $\alpha_{em}(m_Z^2)$, PSI-PR-95-1, BUDKERINP 95-5, JANUARY 1995.

[5] M. Consoli and W. Hollik, *Z Physics at LEP 1*, CERN 89-08, edited by G. Altarelli *et al.*, Vol. 1 (1989), p. 7.

[6] The LEP Collaborations, *A Combination of Preliminary LEP Electroweak Measurements and Constraints on the Standard Model*, CERN-PPE/95-172.

[7] LEP energy working group, R. Asmann *et al.*, Z. Phys C **66**, 567 (1995).

[8] ALEPH Collaboration, D. Decamp *et al.*, Z. Phys. C **48**, 365 (1990);
ALEPH Collaboration, D. Decamp *et al.*, Z. Phys. C **53**, 1 (1992);
ALEPH Collaboration, D. Decamp *et al.*, Z. Phys. C **60**, 71 (1993);
ALEPH Collaboration, D. Decamp *et al.*, Z. Phys. C **62**, 539 (1994);
ALEPH Collaboration, D. Decamp *et al.*, *Preliminary Results on Z Production Cross-Sections and Lepton Forward–Backward Asymmetries Using the 1994 Data*, **eps0398**.

[9] DELPHI Collaboration, P. Aarnio *et al.*, Nucl. Phys. B **367**, 511 (1991);
DELPHI Collaboration, P. Abreu *et al.*, Nucl. Phys. B **417**, 3 (1994);
DELPHI Collaboration, P. Abreu *et al.*, Nucl. Phys. B **418**, 403 (1994);
DELPHI Collaboration, DELPHI, Note 95-62 PHYS 497, July 1995.

[10] L3 Collaboration, B. Adeva *et al.*, Z. Phys. C **51**, 179 (1991);
L3 Collaboration, O. Adriani *et al.*, Phys. Rep. **236**, 1 (1993);
L3 Collaboration, M. Acciarri *et al.*, Z. Phys. C **62**, 551 (1994);
L3 Collaboration, L3 Note 1809, July 1995.

[11] OPAL Collaboration, G. Alexander *et al.*, Z. Phys. C **52**, 175 (1991);
OPAL Collaboration, P. D. Acton *et al.*, Z. Phys. C **58**, 219 (1993);
OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **61**, 19 (1994);
OPAL Collaboration, OPAL Internal Physics Note PN142, July 1994;
OPAL Collaboration, OPAL Internal Physics Note PN166, February 1995.

[12] S. Jadach *et al.*, Phys. Lett. B **353**, 362 (1995).

[13] D. Bardin *et al.*, Z. Phys. C **44**, 493 (1989);
D. Bardin *et al.*, Comp. Phys. Comm. **59**, 303 (1990);
D. Bardin *et al.*, Nucl. Phys. B **351**, 1 (1991);
D. Bardin *et al.*, Phys. Lett. B **255**, 290 (1991);
D. Bardin *et al.*, CERN-TH 6443/92 (May 1992).

[14] ALEPH Collaboration, "Improved tau polarisation measurements," submitted to Z. Phys., CERN-PPE/95-023 (1995).

[15] DELPHI Collaboration, P. Abreu *et al.*, Z. Phys. C **67**, 183 (1995).

[16] L3 Collaboration, O. Accaiari *et al.*, Phys. Lett. B **341**, 24 (1994);
L3 Collaboration, *A Preliminary Update of $\mathcal{A}_\tau$ and $\mathcal{A}_e$ Using the 1994 Data*, L3 Note 1793, July 5, 1995.

[17] OPAL Collaboration, *Updated Measurement of the Tau Polarisation Asymmetries*, OPAL Internal Physics Note PN172, March 1995.

[18] P. Dornan, these proceedings.

[19] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **313**, 549 (1993).

[20] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **313**, 535 (1993).

[21] DELPHI Collaboration, "Measurement of the partial decay width $R_b = \Gamma_{b\bar{b}}/\Gamma_{had}$ with the DELPHI detector at LEP," contributed paper to EPS-HEP-95 Brussels, **eps0570**.

[22] L3 Collaboration, O. Adriani *et al.*, Phys. Lett. B **307**, 327 (1993).

[23] OPAL Collaboration, Z. Phys. C **67**, 17 (1995);
OPAL Collaboration, "An update of the measurement of $\Gamma_{b\bar{b}}/\Gamma_{had}$ using a double tagging method," contributed paper to EPS-HEP-95, **eps0278**.

[24] ALEPH Collaboration, D. Buskulic *et al.*, Z. Phys. C **62**, 179 (1994).

[25] DELPHI Collaboration, "Study of charm mesons production in Z^0 decays and measurement of Γ_c/Γ_h," contributed paper to EPS-HEP-95, **eps0557**.

[26] OPAL Collaboration, R. Akers *et al.*, Z. Phy. C **62**, 27;
OPAL Collaboration, "A measurement of $b(c \to D^\star)$ and $\Gamma_{c\bar{c}}/\Gamma_{had}$ using a double tagging technique," contributed paper to EPS-HEP-95 Brussels, **eps0289**;
OPAL Collaboration, *A Summary of OPAL Measurements of Charm Production in Z Decays*, OPAL Physics Note PN190.

[27] P. Dornan, these proceedings.

[28] LEP Electroweak Working Group, *Combination of Heavy Flavour Electroweak Measurements at LEP*, in preparation, to be submitted to Nucl. Instrum. Methods.

[29] ALEPH Collaboration, D. Buskulic *et al.*, Z. Phys. C **26**, 179 (1994);
ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **335**, 99 (1994);
ALEPH Collaboration, D. Buskulic *et al.*, Z. Phys. C **62**, 1 (1994);
ALEPH Collaboration, "$B^0\bar{B}^0$ mixing and $b\bar{b}$ asymmetry from high-$p-t$ leptons," ALEPH 94-036, contributed paper to the LaThuile and Moriond Winter conferences, 1994;
ALEPH Collaboration, *Heavy Flavour Lepton Contribution for Summer 1994 Conferences*, ALEPH 94-123 PHYSICS 94-107;
ALEPH Collaboration, "Measurement of the semileptonic b branching ratios from inclusive leptons in Z decays," contributed paper to EPS-HEP-95 Brussels, **eps0404**;
ALEPH Collaboration, "The forward-backward asymmetry for charm quarks at the Z pole: An update," contributed paper to EPS-HEP-95 Brussels, **eps0634**.

[30] DELPHI Collaboration, P. Abreu *et al.*, Z. Phys C **65**, 569 (1995);
DELPHI Collaboration, P. Abreu *et al.*, Z. Phys C **66**, 341 (1995);
DELPHI Collaboration, "Measurement of the forward-backward asymmetries

of $e^+e^- \to Z \to b\bar{b}$ and $e^+e^- \to Z \to c\bar{c}$," contributed paper to EPS-HEP-95 Brussels, **eps0571**.

[31] L3 Collaboration, O. Adriani *et al.*, Phys. Lett. B **292**, 454 (1992);
L3 Collaboration, M. Acciarri *et al.*, Phys.Lett. B **335**, 542 (1994);
L3 Collaboration, *Measurement of R_b and $BR(b \to lX)$ from b-Quark Semileptonic Decays*, L3 Note 1449, July 1993;
L3 Collaboration, *L3 Results on $A_{FB}^{0,b}$, $A_{FB}^{0,c}$, and χ for the Glasgow Conference*, L3 Note 1624;
L3 Collaboration, *L3 Results on R_b and $BR(b \to l)$ for the Glasgow Conference*, L3 Note 1625.

[32] OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **60**, 199 (1993);
OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **67**, 365 (1995);
OPAL Collaboration, "Measurement of the forward-backward asymmetries of and $e^+e^- \to Z \to b\bar{b}$ and $e^+e^- \to Z \to c\bar{c}$ from events tagged by a lepton, including 1994 data," contributed paper to EPS-HEP-95 Brussels, **eps0279**;
OPAL Collaboration, "A measurement of the charm and bottom forward-backward asymmetry using D mesons with the OPAL detector at LEP," contributed paper to EPS-HEP-95 Brussels, **eps0290**.

[33] ALEPH Collaboration, D. Decamp *et al.*, Phys. Lett. B **259**, 377 (1991);
ALEPH Collaboration, ALEPH-NOTE 93-041 PHYSIC 93-032 (1993);
ALEPH Collaboration, ALEPH-NOTE 93-042 PHYSIC 93-044 (1993);
ALEPH Collaboration, ALEPH-NOTE 93-044 PHYSIC 93-036 (1993);
ALEPH Collaboration, "*Jet charge measurements in $Z \to q\overline{q}$ decays*," contributed paper to EPS-HEP-95 Brussels, **eps0449**.

[34] DELPHI Collaboration, P. Abreu *et al.*, Phys. Lett. B **227**, 371 (1992).

[35] OPAL Collaboration, P. D. Acton *et al.*, Phys. Lett. B **294**, 436 (1992);
OPAL Collaboration, OPAL Physics Note PN195 (1995).

[36] R. Messner, these proceedings.

[37] S. Bethke, "Status of α_s measurements," XXXth Rencontres de Moriond, QCD and Hadronic Interactions, Les Arcs, March 19-26, 1995.

[38] CHARM II Collaboration, P. Viliain *et al.*, Phys. Lett. B **335**, 246 (1994).

[39] T. Schalk, these proceedings.

[40] CDHS Collaboration, H. Abramowicz *et al.*, Phys Rev. Lett. **57**, 446 (1986);
CDHS Collaboration, A. Blondel *et al.*, Z. Phys. C **45**, 351 (1990).

[41] CHARM Collaboration, J. V. Allaby *et al.*, Phys. Lett. B **177**, 446 (1986);
CHARM Collaboration, J. V. Allaby *et al.*, Z. Phys. C **36**, 611 (1987).

[42] CCFR Collaboration, C. G. Arroyo *et al.*, Phys. Rev. Lett. **72**, 3452 (1994);
D. Harris, *CCFR Measurement of R_ν and a New Extraction of $\sin^2\theta_W$*, EPS-HEP-95 Conference, Brussels.

[43] R. Hughes and N. Hadley, these proceedings.

[44] J.-F. Grivaz, "New particle searches," presented at the EPS-HEP-95 Conference, LAL 95-83.

RECENT RESULTS FROM TRISTAN*

Ryoji Enomoto†
National Laboratory for High Energy Physics, KEK
1-1 Oho, Tsukuba-city, Ibaraki 305, Japan

Representing the TRISTAN Experiments

ABSTRACT

The TRISTAN results from 1994 to 1995 are reviewed in this report. The physics results dominated the $\gamma\gamma$ physics. Therefore, only these are selected in this article. We have systematically investigated jet productions, the γ-structure function, and charm pair productions in $\gamma\gamma$ processes. The results, discussions, and future prospects are presented.

*TALK PRESENTED AT THE SLAC SUMMER INSTITUTE, TOPICAL CONF., 1995.
†Internet address: enomoto@bsun03.kek.jp.

1 TRISTAN

Initially, the TRISTAN project was aimed at finding the "top" quark.[1] Although only a three-km circumstance was available, we achieved a maximum beam energy of 33 GeV. Unfortunately, the top mass was far beyond this energy.[2] We, thus, converted our target to a high-luminosity operation of this collider. Figure 1 shows the relationship between the beam energies and luminosities for various accelerators. Assuming that CLEO and LEP are standard, we can see why TRISTAN is a

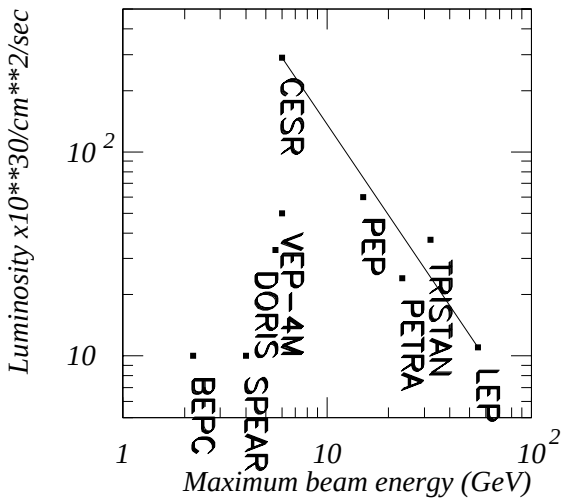


Figure 1: Luminosities versus the maximum beam energies of various e^+e^- colliders.

high-luminosity machine. We hope that the same thing occurs in the near-future B-factory.[3] As a matter of fact, we (TOPAZ, VENUS, and AMY) obtained $\int Ldt$ of 300pb^{-1} per each experiment at $\sqrt{s}$=58 GeV.

If there is a process having a cross section that is an increasing function of $\sqrt{s}$, that may be a big target until the B-factory starts. $\gamma\gamma$ physics is one of them. The luminosity function ($L_{\gamma\gamma}$) is roughly proportional to $\log(s)$. As a result, TRISTAN becomes the highest luminosity $\gamma\gamma$-factory, except for the low $W_{\gamma\gamma}$ region, where CESR still gives the highest $\gamma\gamma$ yield. Therefore, CESR has been fitted for resonance physics, and TRISTAN is for parton physics. For a higher $W_{\gamma\gamma}$ region ($>$ 6 GeV), we have obtained the largest statistics; this situation will remain forever. TRISTAN can play an important role in particle physics, especially regarding strong interactions.

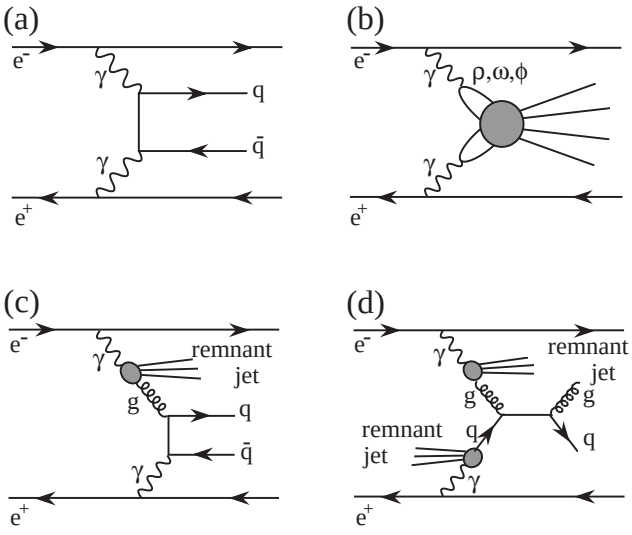


Figure 2: Feynman diagrams which contributed to $\gamma\gamma$ processes at TRISTAN.

2 $\gamma\gamma$ Physics

Here, we briefly mention the $\gamma\gamma$ processes. Figure 2 shows four typical diagrams which contribute to these. (a) is called a "direct process," where photons interact with quarks via point-like interactions.[4] The vector-meson dominance process (VDM) is shown in (b).[5] (c) and (d) are called the "resolved-photon process," where partons inside photons interact point-like.[6] (a) contributes to high-P_T production of quarks, (b) to a low P_T, and (c) and (d) to a medium P_T. Considering our sensitivity over the $W_{\gamma\gamma}$ range, in addition to our trigger system ability,[7] we can study (a), (c), and (d) at most accurate levels.

To conclude, we are sensitive for γ-structure studies, especially concerning the partonic structure of the photon, in addition to higher orders of QCD (or strong interaction). The most important topic is to determine the gluonic densities inside photons. This is the cleanest way to determine the γ-structure in contrast with ep collisions at HERA experiments.

3 Detector

Three groups (TOPAZ, VENUS, and AMY) were operating at TRISTAN. Among them, we pay special attention to the TOPAZ experiments, because of having low-angle calorimeters.

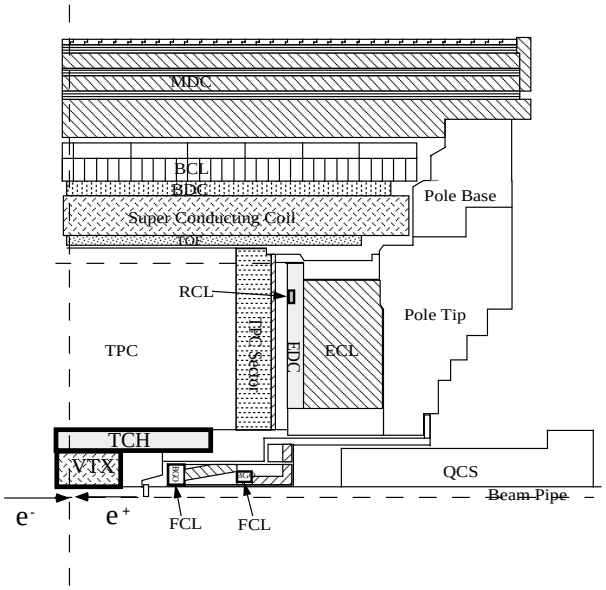


Figure 3: TOPAZ detector.

The apparatus of the TOPAZ detector is shown in Figure 3.[8] The central tracker is a TPC, which enables us to study heavy flavor productions. TOPAZ is the only detector having low-angle calorimeters (FCL).[9] This covers a polar angle region from 3.2 to 12 degrees with respect to the beam axis. The mean beam energy(E_b) of TRISTAN was 29 GeV. When we select events with an energy deposit of 0.4 E_b (beam-electron tag), the Q^2 for the photon is greater than 1.05 GeV^2.

In addition to the beam-electron tag, we have introduced a "remnant-jet-tag.". As shown in Figs. 2 (c) and (d), hadron jets which are resolved form photons flow into beam directions. Typically, hadrons from these jets have P_T's of about 0.4 GeV. Assuming that these hadrons have energies of several GeV, they hit the FCL fiducial region. The energy flow in typical $\gamma\gamma \to 2jet$ events are shown in Fig. 4. It has enhancements at low-angle regions which cannot be explained by the processes shown in Figs. 2 (a) nor (b).

The energy deposits in the FCL are also shown in Fig. 5. The soft component corresponds to these resolved-photon events. We can, therefore, tag the resolved-photon process by selecting a soft energy deposit in the FCL. The efficiency of this tagging was estimated to be $\sim$80% with a background of 10%, mostly due to the beam background. We call this "remnant-jet-tag," or "rem-tag" in short.

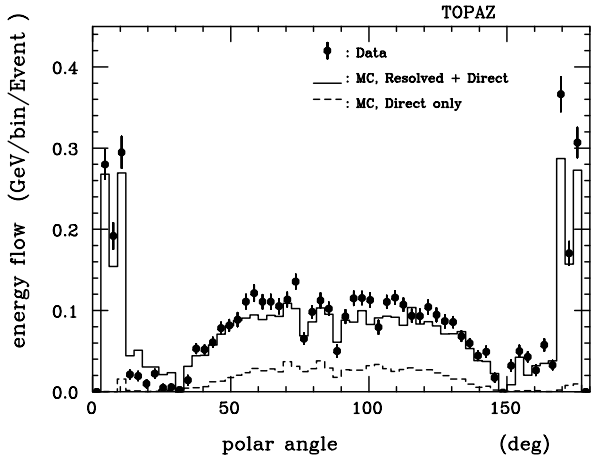


Figure 4: Energy flow of $\gamma\gamma \to 2jet$ events. The histograms are the Monte Carlo prediction; the dashed one is the direct process, and the solid one is the resolved and direct process.

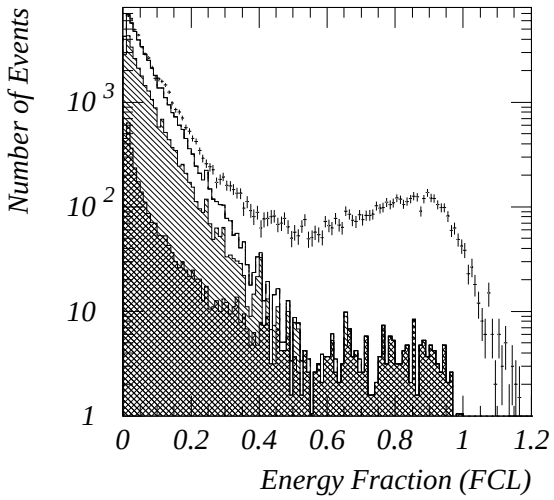


Figure 5: Distribution of the energy fractions (normalized at the beam energy) of the maximum-energy clusters in FCL. The points with error bars are experimental data. The histograms are predictions by a Monte Carlo simulation; the cross-hatched area is a single-photon-exchange process, the singly-hatched one is VDM, and the open one is a resolved-photon process.

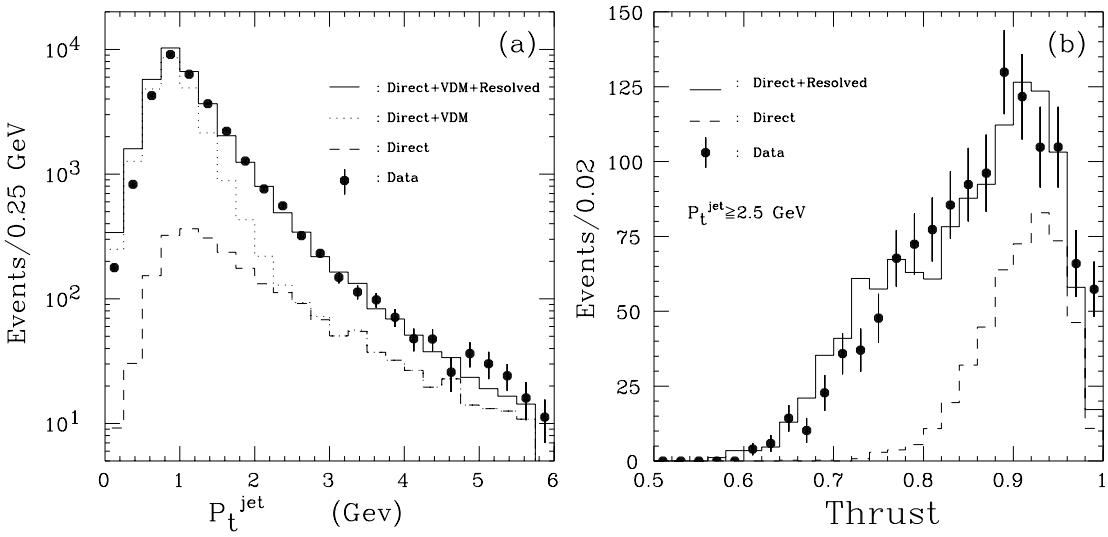


Figure 6: (a) P_T^{jet} distribution. The histograms are theoretical predictions; the dashed one is a direct process, the dotted one is the direct and VDM, and the solid one is the sum of these two and the resolved-photon process. (b) Thrust distribution of high P_T^{jet} (> 2.5 GeV) events.

4 Event Structure

4.1 Event Shapes

As has been described, various processes contribute to $\gamma\gamma$ collisions; the analysis ways are not unique. According to a historical method, hadron systems at the CMS frame were divided into two hemispheres (definition of jets). This method has an advantage for analyzing events in all P_T regions. AMY showed evidence for a resolved-photon process by this method.[10] A similar analysis was carried out by TOPAZ, and the P_T^{jet} distribution is shown in Fig. 6 (a). For example, at P_T^{jet}=2.5 GeV, the data excess is by a factor of 2.5 compared with the incoherent sum of direct and VDM processes. This excess has been explained by the resolved-photon process. Next, the thrust distribution of high P_T^{jet} (>2.5 GeV) events are plotted in Fig. 6 (b). The events are spherical, consistent with the prediction by the resolved-photon processes. Similar results have been obtained by the LEP experiments.[11,12]

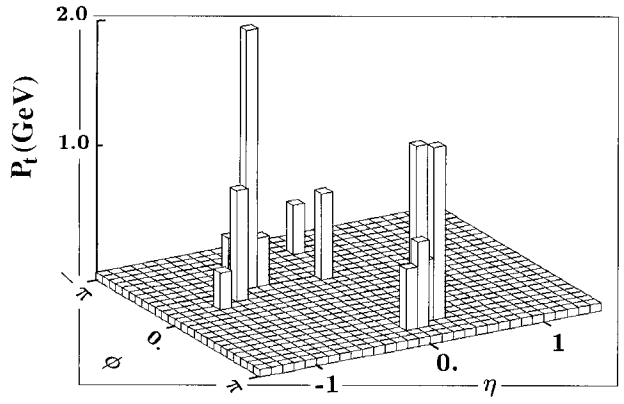


Figure 7: Typical jets observed by the TOPAZ detector.

4.2 Jet Cross Section

The processes shown in Figs. 2 (a), (c), and (d) include hard scattering of partons which are observed as jets (Fig. 7). These jets are reconstructed in ϕ and η plane. The particles within the circle $R = \sqrt{(\phi - \phi_0)^2 + (\eta - \eta_0)^2}$ are used. Figure 8 is the cross section of jet production versus P_T^{jet}. The cross section is consistent with the incoherent sum of the direct and resolved-photon processes at the $P_T^{jet} > 2$GeV region (the same result as the previous one). The theoretical models, called LAC1, LAC3, and DG shown in the figure, have significant differences in the gluon distribution inside the photon.[13,14] The hard-gluon model (LAC3) is clearly rejected. LAC1 and DG show difference at low-x gluonic-density, and it is difficult to distinguish them by this experimental method.[15] A similar result was obtained by AMY.[16]

5 Structure Function

The photon-structure function (F_2^γ) was measured by the TOPAZ collaboration.[17] We obtained a high value compared with the theoretical values at x $\sim$ 0.04 at $3 < Q^2 < 30$ GeV^2. These regions are important for determining the QCD-based models. Although the experimental ambiguities in determining x value from the mass of the measured hadronic system were found to be large, there

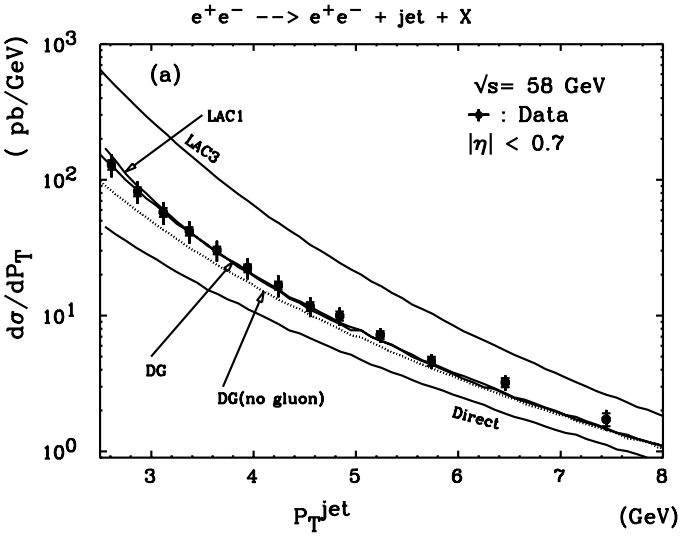


Figure 8: Jet-production cross section in $\gamma\gamma$ collisions.

will be a systematic shift. We are, therefore, going to reanalyze the data using a new algorithm to determine x while assuming missing-energy flow directions (i.e., beam-pipe direction).

6 Charm-Pair Production

According to a QCD calculation of parton-parton scattering, the cut-off parameter (P_T^{min}) was introduced for light-quark scattering. This parameter must be determined experimentally; the optimum value was obtained to be around $1.7 \sim 2$ GeV. Fortunately, for the charm-quark case, this parameter is not necessary, and we can experimentally select charm-pair events with high purity. In addition, the VDM effects are considered to be small for charmed-particle production. In the resolved-photon processes, we only have to consider gluon-gluon scattering; therefore, this is sensitive to the gluonic density in which a large model dependence exists. The NLO calculations are available at the parton level.[18]

6.1 Full and Partial Reconstruction of $D^{*\pm}$

Initial charm quark fragment to D mesons. D^* is the most probable state. This fragmentation function is experimentally well-known. We can, therefore, estimate the initial charm quarks' P_Ts.

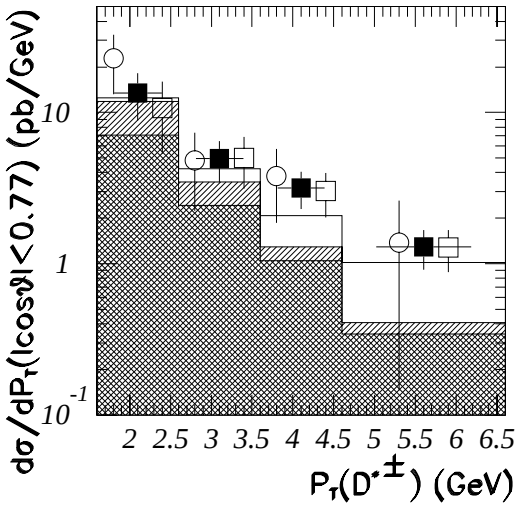


Figure 9: Differential cross sections of $D^{*\pm}$ versus P_T. The open circles were obtained by the full reconstruction method, the open squares by the softpion method, and the closed squares are the average by these two methods. The histograms are the theoretical predictions: the cross hatched is the direct process, the singly hatched is the resolved, and the open area is obtained assuming the hypothetical $\tilde{t}$ pair production.

We first tried to reconstruct $D^{*+} \to \pi_s^+ D^0 (D^0 \to K^- \pi^+ X)$.[19] 20 ± 5 D^*s were reconstructed with a good S/N ratio. The obtained cross section is plotted in Fig. 9 by open circles. Although this cross section is higher than the sum of the direct and resolved-photon predictions, the statistics are low. The NLO effect was taken into account in the prediction. We used LAC1 for the gluonic density in photons.

In order to improve the experimental accuracy, we carried out a "softpion analysis" in reconstructing the D^*s. The results are also shown in Fig. 9 by the open circles. They are consistent with that of the full reconstruction. The high P_T anomaly still existed, and the hypothetical $\tilde{t}$ pair assumption was tested by looking at the event shapes.[20] These shapes differ from $\tilde{t}$-pair prediction and rather resemble the typical $\gamma\gamma$ events. The similar high P_T anomaly was also reported by the AMY collaboration.[21]

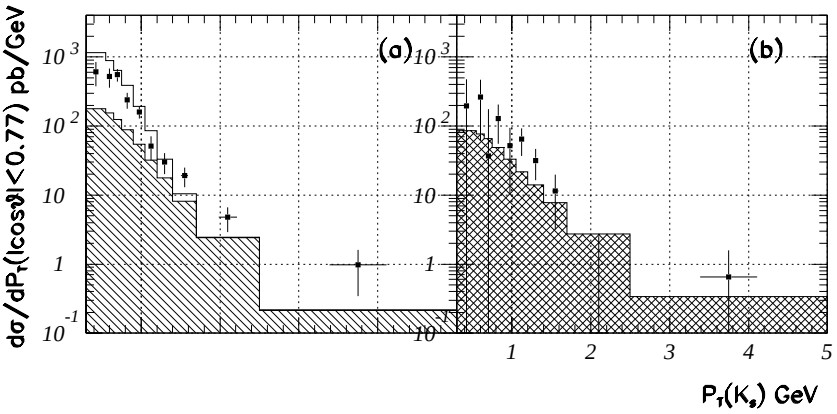


Figure 10: Differential cross sections of K_0 versus P_T: (a) for the VDM and resolved-photon process, and (b) for the direct process. The histograms are the theoretical predictions; the definition of the hatches are the same as those for the previous figure.

6.2 K_s Inclusive

The maximum integrated luminosity of the TRISTAN experiment is 300 pb^{-1}, and now most of them were analyzed. We must, therefore, seek other ways of analysis than waiting for an increase in data. An inclusive analysis of the strange particle is one of them. The P_T spectrum of these reflect that of charm quarks. Also, strange-quark pair production is strongly suppressed in $\gamma\gamma$ collisions.

In the K_s inclusive analysis, we introduced "remnant-jet-tag."[22] The details were described in the previous section. We can, therefore, derive the cross sections process by process. These are shown in Figs. 10 (a) and (b). Here, we could not separate the VDM and the resolved-photon events, because of the low-P_T particle production by the VDM.

By this study, the existing theory plus the LAC1 parametrization with the NLO correction well describe the experimental data. We further tested the parametrization difference in the gluon density by using the WHIT parametrization.[23] This gave six systematic parametrizations. Some combinations of these with various P_T^{min} cut-offs fitted the experimental data perfectly.

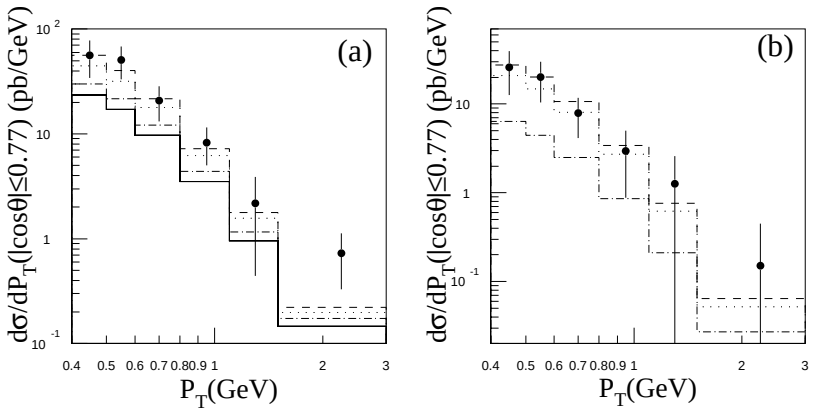


Figure 11: (a) Electron inclusive cross sections; the solid line is the direct process, the dot-dashed is the DG case, the dotted is the LAC1 with m_c=1.5 GeV, and the dashed is the LAC1 with m_c=1.3 GeV. (b) The resolved-photon cross section; the definition of lines are as same as (a).

6.3 Electron Inclusive

The electron inclusive method is a cleaner one than the inclusive K_0. Here, we do not have to consider the VDM. The TOPAZ detector can identify very low P_T electrons, such as 0.4 GeV.[24] We can, therefore, measure the gluon density at very low x ($\sim$ 0.02), where the model differences appears. Figure 11 (a) is the differential cross section versus the electron P_Ts. The experimental data clearly favor the LAC1 parametrization, also suggesting the necessity of the NLO correction and a low charm-quark mass of $\sim$ 1.3 GeV.[25] Although VENUS produced a similar result, the statistics were about half that of ours.[26]

We carried out "remnant-jet-tag," and obtained a purely "resolved-photon" cross section [Figure 11 (b)]. Again, it confirmed our parametrization of the theory. We also observed a large difference between the DG and LAC1. This is because this method is sensitive to very low-x regions where the jet analysis could not resolve. Note that this method is more powerful than the single-tag experiment (F_2^γ) in determining the gluon density inside photons.

In Fig. 11 (a) at highest P_T region, there are some excesses compared with the existing theory. A similar high P_T excess was observed by AMY.[21]

tag cond.	Experiment	Theory (LO)	Exp./Theory	subprocess
antitag	43.3±8.3	19.1	2.26±0.43	VDM+resolved+direct
rem-tag (-VDM)	15.6±3.5	6.0	2.60±0.58	resolved
rem-tag	34.8±7.8	17.3	2.01±0.45	VDM+resolved
anti-rem	27.7±7.9	13.1	2.11±0.60	VDM+direct

Table 1: Total cross section (pb) of $\Lambda(\overline{\Lambda})$ in the $|\cos\theta| < 0.77$ and $0.75 < P_T < 2.75$ GeV range. The notation (-VDM) means VDM subtraction using theory. Here, we use the LO theories in order to show the discrepancy with the experimental data.

6.4 Λ Inclusive

So far, what we have learned is that there are some high P_T excess in charm production, and that the experimental results at low P_T agree with the existing theory with the NLO correction and high gluonic density at low x. We investigated the Λ-inclusive cross section in order to qualitatively study the NLO effects. Λ's can also tag charm-pair events the same as in the K_0 case.

In addition, there is an experimental fact that a gluon-jet produces more Λ's than does a quark jet.[27] Our experimental results are shown in Table 1.[28] There are process-independent excesses compared with the prediction of the LO theories. The values are a factor of two. We can, therefore, conclude that there exists significant gluon jet production in $\gamma\gamma$ collisions, i.e., the NLO effect.

7 Double Tag

We carried out a double-tag analysis and obtained the total hadronic cross sections.[31] The Q^2 ranges for γ^* was $2 \sim 25$ GeV^2 and the W range was $2 \sim 25$ GeV. Figure 12 is the ratio of the cross sections ($e^+e^- \to e^+e^-h$) between the experiment and the LO $e^+e^- \to e^+e^-q\bar{q}$ theory. The experimental value agrees with the LO prediction in the high-Q^2 region. There are enhancements of $\sim 30\%$ in the low-Q^2 region, suggesting NLO effects.

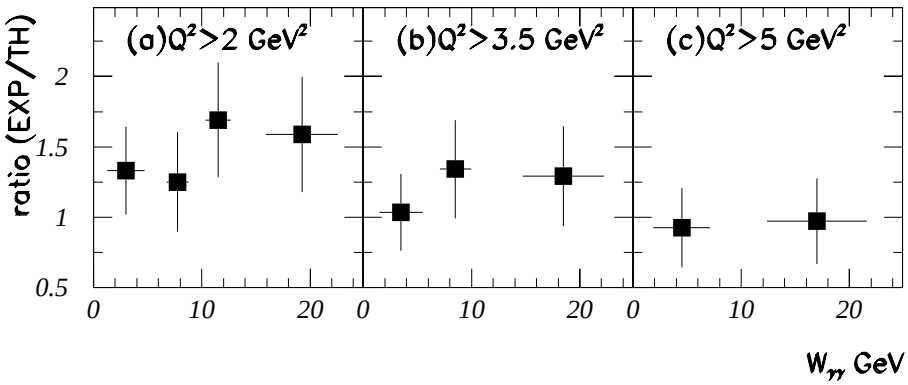


Figure 12: Ratios of the experimental and theoretical cross sections in various kinematic regions; (a) $Q^2_{\gamma,min} > 2$ GeV^2, (b) $Q^2_{\gamma,min} > 3.5$ GeV^2, and (c) $Q^2_{\gamma,min} > 5$ GeV^2. The $Q^2_{\gamma,min}$ is specified in the text.

8 Discussion

Our experimental data strongly favor a large gluon density at low x, as has been suggested by LAC1. However, the HERA experiment (ep collision) showed a lower gluon density than that which LAC1 predicted.[29] Also, the LEP experiment is inconsistent with LAC1 at low x.[30] The problem is whether we can explain all of the experimental data by simply changing the parton density functions. In addition, the high P_T excess in charm production cannot be solved by any existing theories.

The cross section of the $\gamma\gamma$ collision increases with energy in a future e^+e^- linear collider experiment. It would be a large background and may be related to such physics as $H \to \gamma\gamma$ searches. In order to reliably estimate the background, our measurement greatly helps. Systematic measurements, such as $\gamma\gamma \to h^{\pm}X$ and γX, are necessary.

9 Conclusion

At the TRISTAN e^+e^- collider, a systematic study of hadronic $\gamma\gamma$ collisions was carried out. TRISTAN is a high-luminosity $\gamma\gamma$ factory, and our data of these

processes have the largest statistics. For parton production, our data greatly contributed to our experimental and theoretical understanding of photon structures. Further systematic measurements on various processes are awaited.

Acknowledgments

I thank Drs. H. Hayashii, M. Iwasaki, and T. Nozaki for summarizing these results. I also thank the TOPAZ collaboration and the KEK accelerator division.

References

[1] TRISTAN Project Group, KEK Report 86-14 (1987).

[2] F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995); A. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[3] The BELLE collaboration, Letter of Intent, KEK Report 94-2 (1994).

[4] B. J. Brodsky, T. Kinosita, and H. Terazawa, Phys. Rev. D**4**, 1532 (1971).

[5] J. J. Sakurai, "Currents and mesons" (University of Chicago Press, 1969).

[6] S. J. Brodsky, T. A. DeGrand, J. F. Gunion, and J. H. Weis, Phys. Rev. Lett. **41**, 672 (1978); H. Terazawa, J. Phys. Soc. Jpn. **47**, 355 (1979).

[7] R. Enomoto *et al.*, Nucl. Instrum. Methods A**269**, 507 (1988); R. Enomoto, K. Tsukada, N. Ujiie, and A. Shirahashi, IEEE Trans. on NS. **35**, 419 (1988); T. Tsukamoto, M. Yamauchi, and R. Enomoto, Nucl. Instrum. Methods A**297**, 148 (1990).

[8] T. Kamae *et al.*, Nucl. Instrum. Methods A**252**, 423 (1986); S. Kawabata *et al.*, Nucl. Instrum. Methods A**270**, 11 (1988); J. Fujimoto *et al.*, Nucl. Instrum. Methods A**256**, 449 (1987); A. Yamamoto *et al.*, Jpn. J. Appl. Phys. Lett. **55**, L440 (1986).

[9] H. Hayashii *et al.*, Nucl. Instrum. Methods A**316**, 202 (1992).

[10] R. Tanaka *et al.*, Phys. Lett. B**277**, 215 (1992).

[11] D. Buskulic *et al.*, Phys. Lett. B**313**, 5099 (1993).

[12] P. Abero *et al.*, Z. Phys. C**62**, 357 (1994).

[13] H. Abramowicz, K. Charchula, and A. Levy, Phys. Lett. B**269**, 458 (1991).

[14] M. Drees and K. Grassie, Z. Phys. C**28**, 451 (1985).

[15] H. Hayashii *et al.*, Phys. Lett. B**314**, 149 (1993).

[16] B. J. Kim *et al.*, Phys. Lett. B**325**, 248 (1994).

[17] K. Muramatsu *et al.*, Phys. Lett. B**332**, 477 (1994).

[18] M. Drees, M. Krämer, J. Zunft, and P. M. Zerwas, Phys. Lett. B**306**, 371 (1993).

[19] R. Enomoto *et al.*, Phys. Rev. D**38**, 1879 (1994).

[20] R. Enomoto *et al.*, Phys. Lett. B**328**, 535 (1994).

[21] T. Aso *et al.*, KEK Preprint 95-19, talk presented at Photon '95, April 1995, Sheffield, U.K.

[22] R. Enomoto *et al.*, Phys. Lett. B**341**, 238 (1994).

[23] K. Hagiwara, M. Tanaka, I. Watanabe, and T. Izubuchi, Phys. Rev. D**51**, 3197 (1995).

[24] M. Iwasaki, E. Nakano, and R. Enomoto, KEK Preprint 94-180, to be published in Nucl. Instrum. Methods A.

[25] M. Iwasaki *et al.*, Phys. Lett. B**341**, 99 (1994).

[26] S. Uehara *et al.*, Z. Phys. C**63**, 213 (1994).

[27] B. Andersson, G. Gustafson, and T. Sjöstrand, Physica. Scripta. **Vol. 32**, 574 (1985).

[28] R. Enomoto *et al.*, Phys. Lett. B**347**, 179 (1995).

[29] J. M. Butterworth, DESY 95-043.

[30] I. Kronkvist and B. W. Kennedy, talk presented at Photon '95, April 1995, Sheffield, U.K.

[31] R. Enomoto *et al.*, submitted for publication.

ELECTROWEAK RESULTS FROM THE TEVATRON

Darien Wood*
Fermi National Accelerator Laboratory
Batavia, Illinois 60510

Representing the CDF and DØ Collaborations

ABSTRACT

Electroweak results are presented from the CDF and DØ experiments based on data collected in recent runs of the Fermilab Tevatron Collider. The measurements include the mass and width of the W boson, the production cross sections of the W and Z bosons, and the W charge asymmetry. Additional results come from studies of events with pairs of electroweak gauge bosons and include limits on anomalous couplings.

*Present address: Department of Physics, Northeastern University, Boston, MA 02115.

©1995 by Darien Wood.

1 Introduction

In high-energy $\overline{p}p$ collisions, it is possible to study electroweak physics by direct observation of the carriers of the weak force, W and Z bosons. W bosons, in particular, have been produced and detected only at the CERN S$\overline{p}p$S (closed since 1991) and at the Fermilab Tevatron Collider. The large samples of W bosons produced in hadron colliders complement the detailed studies of the Z boson at the e^+e^- colliders, LEP and SLC.

It is interesting to track the number of W bosons detected by experiments over the years, as shown in Fig. 1. From the handful of events that established the existence of W and Z bosons at CERN in 1982 (Refs. 1 and 2), the samples available now to the Tevatron experiments number in the tens of thousands. This steady increase in statistics has yielded corresponding increases in the precision of electroweak measurements and in the variety of electroweak properties that are studied at the hadron collider experiments.

Almost all results presented here come from the recent runs of the Tevatron Collider and from its two collider experiments: CDF (Ref. 3) and DØ (Ref. 4). The run which took place in 1992–93 is referred to as "Run 1A," and it resulted in integrated luminosities of about 13–20 pb^{-1} per experiment. The run which began in early 1994 and which is still in progress is called "Run 1B" and is expected to yield ≈ 100 pb^{-1}. In both runs, the $\overline{p}p$ collisions have a center-of-mass energy of $\sqrt{s} = 1.8$ TeV. Final results are available for most of the Run 1A analyses, and some preliminary results based on part of the Run 1B data are included as well.

Since their hadronic decay modes are difficult to distinguish from the large background from QCD multijet production, these gauge bosons are usually studied through their leptonic decay modes: $W \to \ell\nu$ and $Z \to \ell^+\ell^-$.

2 W and Z Boson Production Studies

At lowest order, W and Z bosons are produced via quark-antiquark annihilation. Higher order contributions, which can include gluons in the initial and final states, increase the total cross section and create a nonzero transverse momentum spectrum for the W and Z bosons. Thus, the total production rate of W and Z bosons depends on many factors outside the scope of pure electroweak theory, especially parton distribution functions and QCD corrections. Some electroweak properties

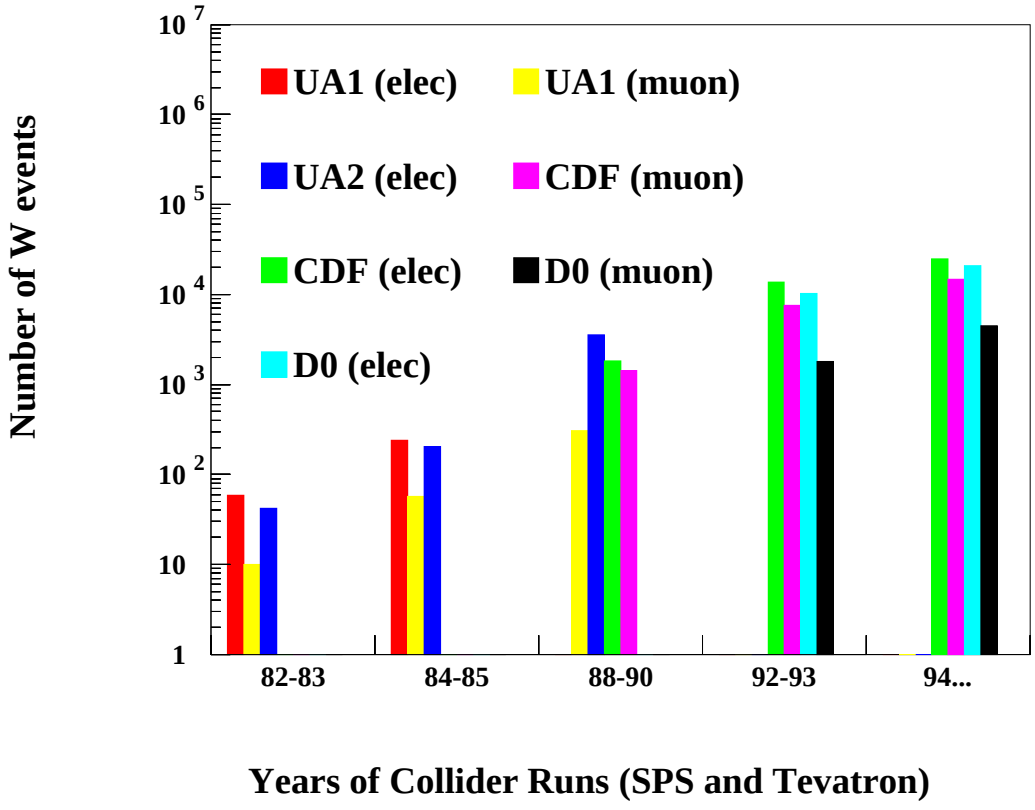


Figure 1: Number of W boson events observed by experiments as a function of the years of hadron collider runs.

can be revealed, however, by examining production rates as a function of boson type and lepton charge and rapidity.

2.1 Production Cross Sections and Indirect Γ_W Measurement

The details of selection of events vary slightly among the different analyses, but, in general, are quite similar. For W bosons, a high-p_T isolated lepton ($p_T^\ell >$ 20–25 GeV) is required along with missing transverse energy ($\not\!\!E_T >$ 20–25 GeV) which is identified as the neutrino transverse momentum (p_T^ν). The Z event selection generally requires two high-p_T ($p_T^\ell >$ 15–25 GeV) leptons and an invariant mass for the pair near the Z boson mass. The principal backgrounds are QCD multijet events with fake leptons and/or $\not\!\!E_T$, decays $W \to \tau\nu, \tau \to \ell\nu\bar{\nu}$, and (in the muon channel only) cosmic rays. Figure 2 shows an example[5] of the W transverse mass ($M_T = \sqrt{(p_T^\ell + p_T^\nu)^2 - (\vec{p}_T^{\,\ell} + \vec{p}_T^{\,\nu})^2}$) and Z invariant mass distributions after selection cuts.

The rate of W and Z bosons observed by the experiments is proportional to the product of production cross section and leptonic branching fraction. The measurements of this product in the electron and muon channels are given in Table 1 (Refs. 5–7). Also shown is the ratio of $\sigma \cdot B$ for W and Z production:

$$R_\ell = \frac{\sigma(\overline{p}p \to W \to \ell\nu)}{\sigma(\overline{p}p \to Z \to \ell\ell)} = \frac{\sigma(\overline{p}p \to W)}{\sigma(\overline{p}p \to Z)} \cdot \frac{\Gamma(W \to \ell\nu)}{B(Z \to \ell\ell)} \cdot \frac{1}{\Gamma_W}. \qquad (1)$$

R_ℓ is predicted more precisely than the individual cross sections because many of the QCD and parton-distribution effects partially cancel. Experimentally, it has the advantage that the luminosity errors cancel completely and the efficiency errors cancel partially. The world R_ℓ measurements[5–9] (excluding preliminary results) and their averages are shown in Fig. 3. The ratio of production cross sections predicted at $O(\alpha_s^2)$ (Ref. 10) is 3.33 ± 0.03 (3.26 ± 0.09) at $\sqrt{s} = 1.8$ TeV (0.63 TeV). The branching ratio $B(Z \to \ell\ell) = (3.367 \pm 0.006)\%$ can be taken from the LEP experiments.[11] With these inputs, Eq. (1) can be used to transform the R_ℓ measurement into a determination of the W leptonic branching ratio: $B(W \to \ell\nu) = (10.9 \pm 0.3)\%$. If, in addition, we assume the Standard Model prediction for the partial decay width $\Gamma(W \to \ell\nu) = 225.2 \pm 1.5$ MeV (Ref. 12), then we obtain an indirect measurement of the total width of the W boson: $\Gamma_W = 2.062 \pm 0.059$ GeV. This can be compared with the Standard Model prediction

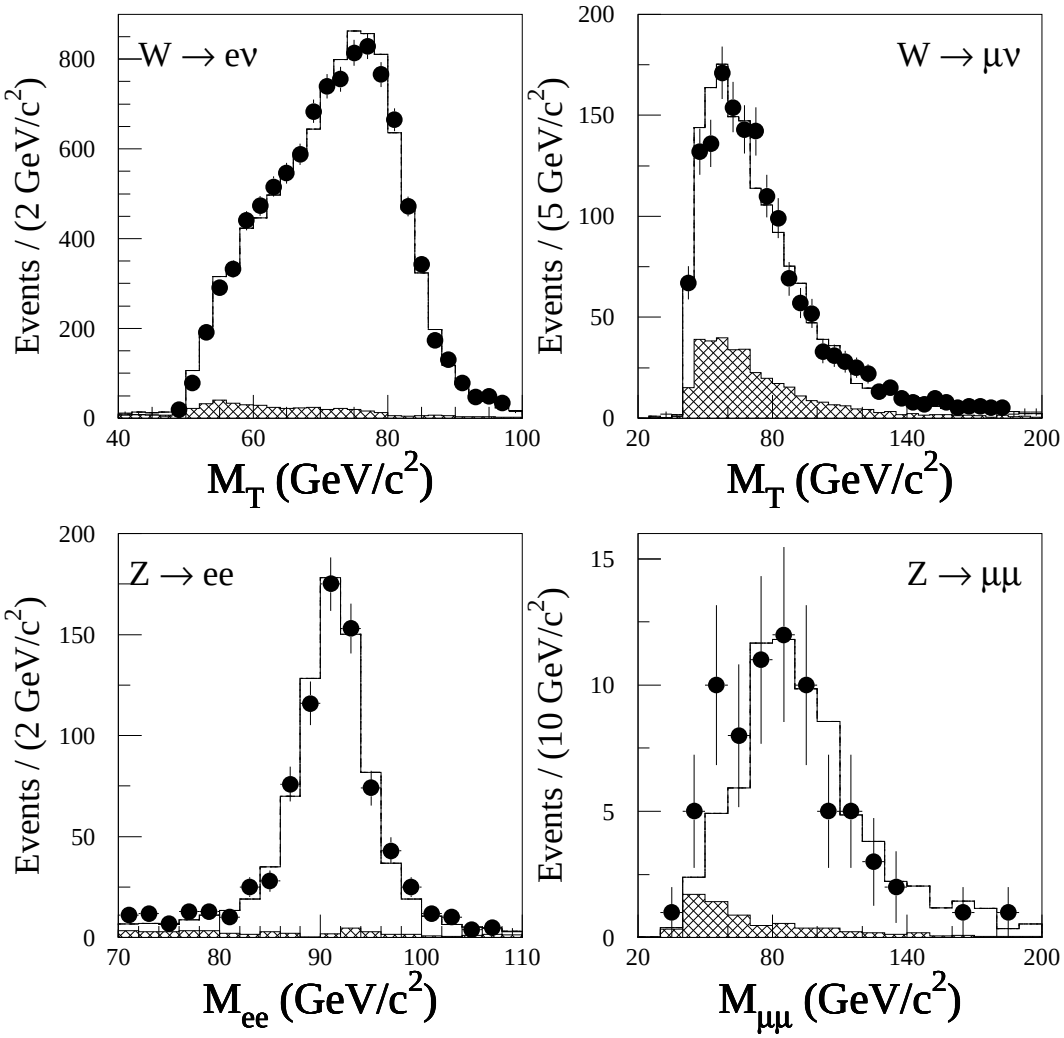


Figure 2: W transverse mass distributions and Z invariant mass distributions from the DØ Run 1A cross-section analysis.

	$\sigma \cdot B(W \to \ell\nu)$ (nb)	$\sigma \cdot B(Z \to \ell\ell)$ (nb)	R_ℓ
1988/'89 data			
CDF(e)	$2.19 \pm 0.04 \pm 0.21$	$0.209 \pm 0.013 \pm 0.017$	$10.2 \pm 0.8 \pm 0.4$
CDF(μ)	2.21 ± 0.22	0.226 ± 0.032	9.8 ± 1.2
Run 1A			
CDF(e)	2.49 ± 0.12	0.231 ± 0.012	$10.90 \pm 0.32 \pm 0.29$
DØ(e)	$2.36 \pm 0.02 \pm 0.15$	$0.218 \pm 0.008 \pm 0.014$	$10.82 \pm 0.41 \pm 0.30$
DØ(μ)	$2.09 \pm 0.06 \pm 0.25$	$0.178 \pm 0.022 \pm 0.023$	$11.8 \pm 1.6 \pm 1.1$
Run 1B			
(prelim.)			
DØ(e)	$2.24 \pm 0.02 \pm 0.20$	$0.226 \pm 0.006 \pm 0.021$	$9.9 \pm 0.3 \pm 0.8$
DØ(μ)	$1.93 \pm 0.04 \pm 0.20$	$0.159 \pm 0.014 \pm 0.022$	$12.3 \pm 1.1 \pm 1.2$

Table 1:

of $\Gamma_W = 2.077 \pm 0.014$ GeV (Ref. 12). This comparison results in an upper limit (95% CL) of 109 MeV for the excess decay width of the W boson which can be used to put limits on any new final states into which the W might decay.

2.2 Direct Γ_W Measurement

The CDF experiment also estimates the W total width with a direct fit[13] of the transverse mass spectrum of $W \to e\nu$ events, shown in Fig. 4. The high transverse mass region of the distribution is sensitive to the width of the Breit-Wigner line shape. A fit to the transverse mass above 110 GeV results in a determination $\Gamma_W = 2.11 \pm 0.28 \pm 0.16$ GeV. Although the uncertainties are larger than those from the indirect ratio method, this direct width determination requires fewer Standard Model assumptions.

2.3 W Charge Asymmetry

The measurement of the lepton charge asymmetry in W boson events gives additional information about the production properties. The different momentum distributions of *up* and *down* quarks in the proton give rise to an asymmetry in the production of W bosons: a W^+ is more likely to follow the direction of the

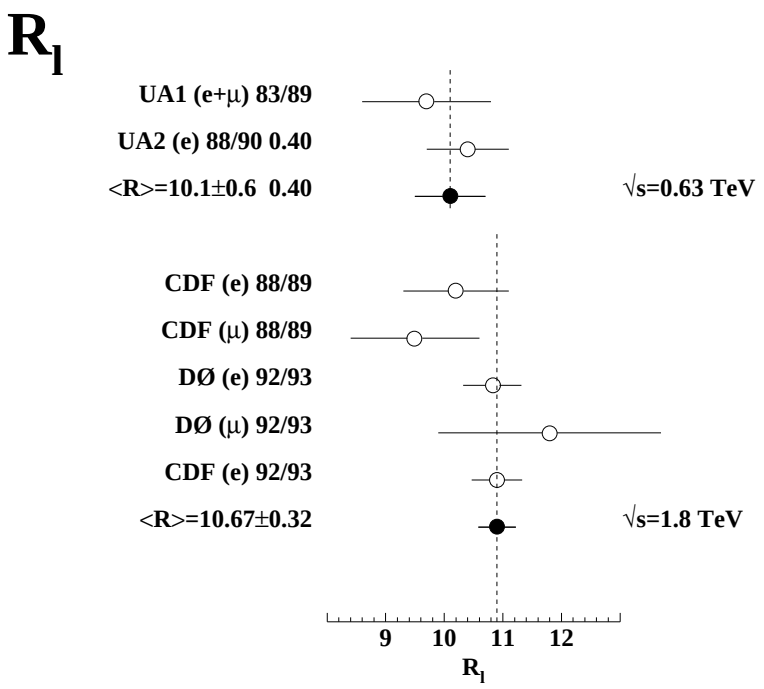


Figure 3: World measurements of R_ℓ, the ratio of W to Z cross sections times leptonic branching ratio.

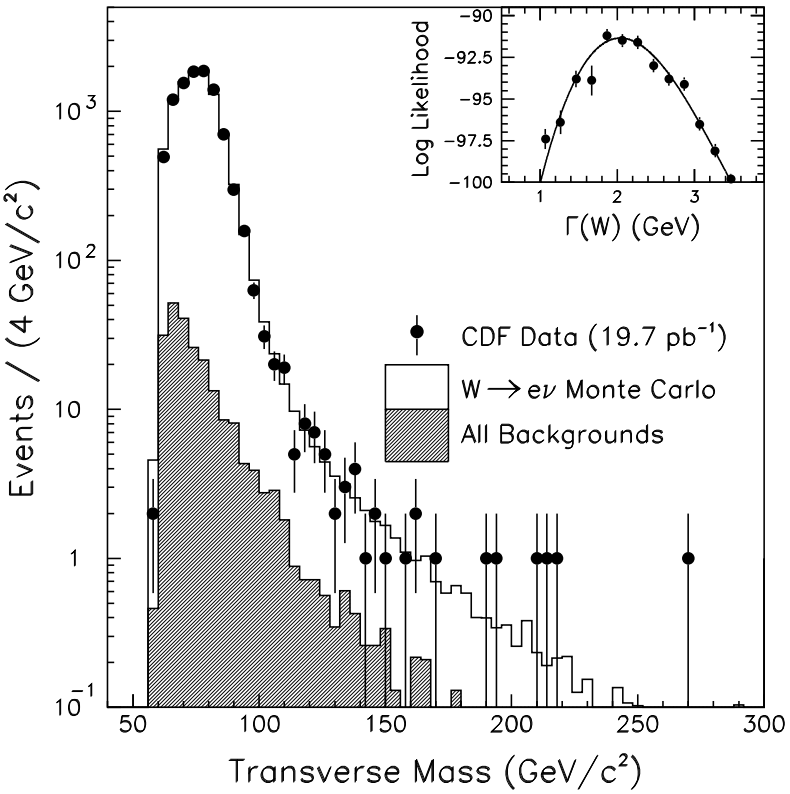


Figure 4: The transverse mass distribution used in the CDF direct fit of the W boson width. The fit is made in the region $M_T > 110$ GeV.

proton beam. When the boson decays, this is also asymmetric due to the $V-A$ couplings, and the tendency is opposite, sending the ℓ^+ back toward the antiproton direction in the W boson rest frame. In the lab frame, one sees the combined effect in the charge asymmetry of leptons from W boson decays as a function of pseudorapidity (η), with the production asymmetry dominating in most cases. For W boson production at $\sqrt{s} = 1.8$ TeV, the asymmetry is especially sensitive to the slope of $u(x)/d(x)$ for $0.007 < x < 0.27$, where x is the fraction of proton momentum carried by the interacting quark or antiquark.

The lepton charge asymmetry is defined as $A(\eta) = (N_{\ell^+}(\eta) - N_{\ell^-}(\eta))/(N_{\ell^+}(\eta) + N_{\ell^-}(\eta))$. The CDF measurement[14] of $A(\eta)$ from Run 1A has been published and has been used to constrain the parton distribution functions (pdf's). This is important, for example, in the W mass measurement, where pdf uncertainties can contribute significantly to the uncertainty in the final result. Figure 5 shows preliminary asymmetry distributions from Run 1A and part of Run 1B combined. The CDF points are from about 67 pb^{-1} from both electrons and muons, and the DØ points are from about 36 pb^{-1} from muons only. The curves show the NLO Monte Carlo predictions[15] using several pdf sets.[16–18] The older sets which were disfavored by the CDF Run 1A asymmetry[14] are not shown in the figure. The pdf sets shown have included the CDF Run 1A asymmetry as part of their input data, and all three sets are in good agreement with the new data.[19]

3 W Boson Mass

The W mass measurement is the most precise electroweak measurement from the hadron colliders. The favored technique involves fitting the M_T spectra of the W bosons to simulated spectra generated with different W masses. The p_T spectra of the charged lepton and of the neutrino also carry information about the mass, but they are more sensitive to the transverse momentum distribution of the W boson itself than is M_T.

The CDF measurement from Run 1A (Ref. 20) described here is now finalized. Figure 6 shows the M_T distributions which are fit. The electron and muon channels are fit separately, with the results $M_W(\mu) = 80.310 \pm 0.205 \pm 0.130$ GeV, $M_W(e) = 80.490 \pm 0.145 \pm 0.175$ GeV, where the first error is statistical and the second is systematic. The combined result is $M_W = 80.410 \pm 0.180$. The contributions to the uncertainties are summarized in Table 2.

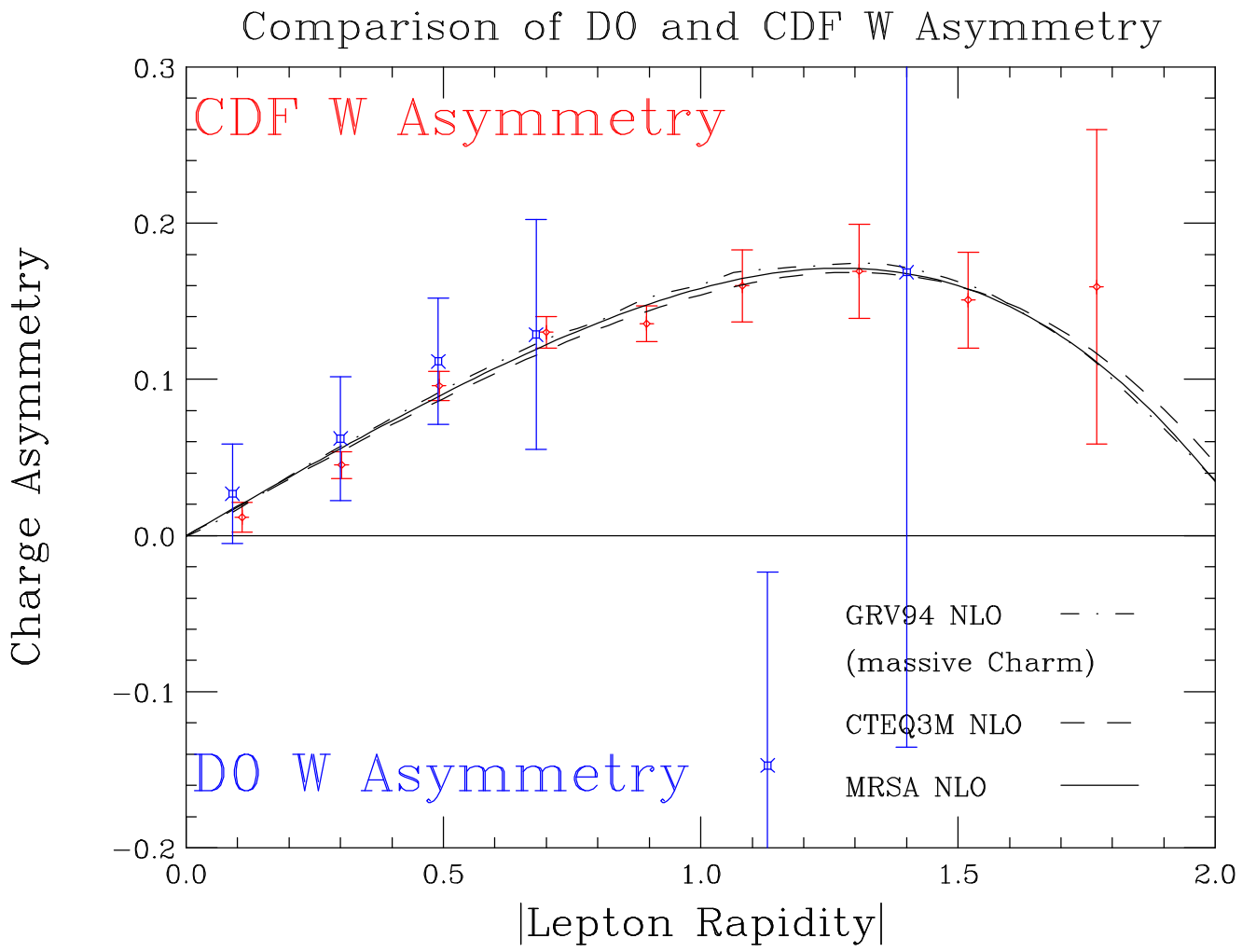


Figure 5: Preliminary lepton charge asymmetries from the W boson samples from Run 1A plus part of Run 1B. The crosses are from CDF, electrons and muons (67 pb^{-1}), and the x's from DØ, muons only (36 pb^{-1}).

Source	e	μ	Common
Statistical	145	205	–
Energy scale	120	50	50
E or p resolution	80	60	–
p_T^W and recoil model	75	75	65
pdf's	50	50	50
QCD/QED corrections	30	30	30
W width	20	20	20
Backgrounds/bias	30	40	5
Fitting procedure	10	10	–
Total	230	240	100
Combined	180		

Table 2: Uncertainties on M_W (MeV) for CDF Run 1A.

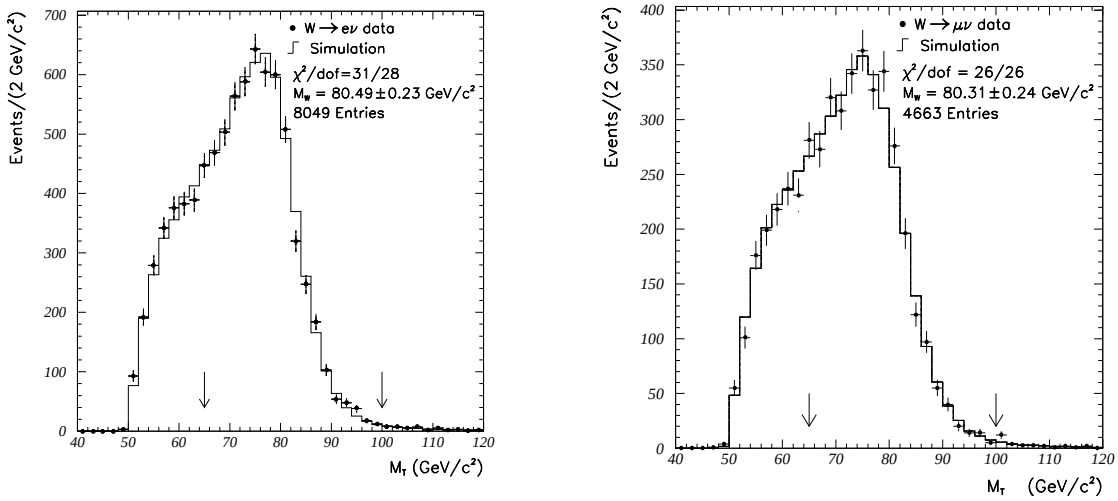


Figure 6: The CDF mass fits to the W boson transverse mass spectra from Run 1A. The electron channel is shown on the left and the muon channel on the right.

The energy scale for the leptons is calibrated first for muons. A sample of about 60,000 $J/\psi \to \mu\mu$ events is used to set the momentum scale in the CDF spectrometer. The ratio between the fit to $\mu\mu$ invariant mass spectrum is shown in Fig. 7 and the world average for the J/ψ mass yields a momentum correction factor 0.999984 ± 0.00058. The error includes the contribution from the uncertainty in the extrapolation from the transverse momenta for the muons from J/ψ decay (typically $\sim$ 3 GeV) to those from W boson decays (typically $\sim$ 40 GeV). This calibration uncertainty results in a contribution of 50 MeV to the W mass uncertainty in the muon channel. After this correction, the scale is checked with the peaks for $\Upsilon \to \mu\mu$ (shown in Fig. 8) and $Z \to \mu\mu$.

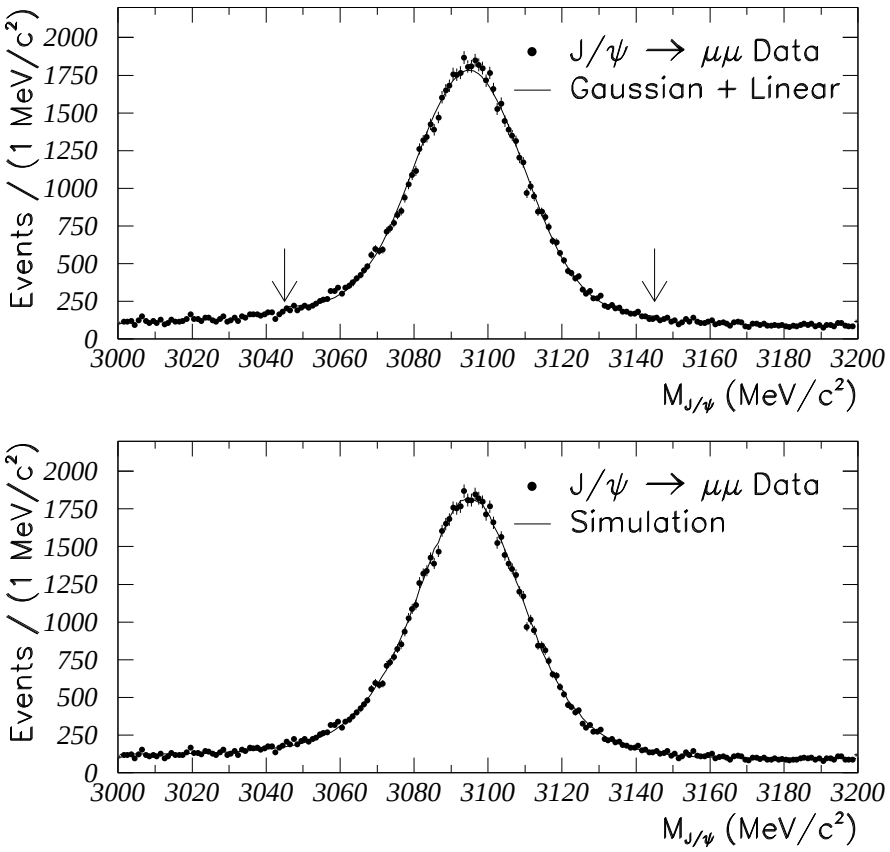


Figure 7: The $J/\psi \to \mu\mu$ invariant mass spectra used in determining the momentum scale used in the CDF W boson mass measurement. A simple fit (top) and the result of the Monte Carlo simulation (below) are compared to the data.

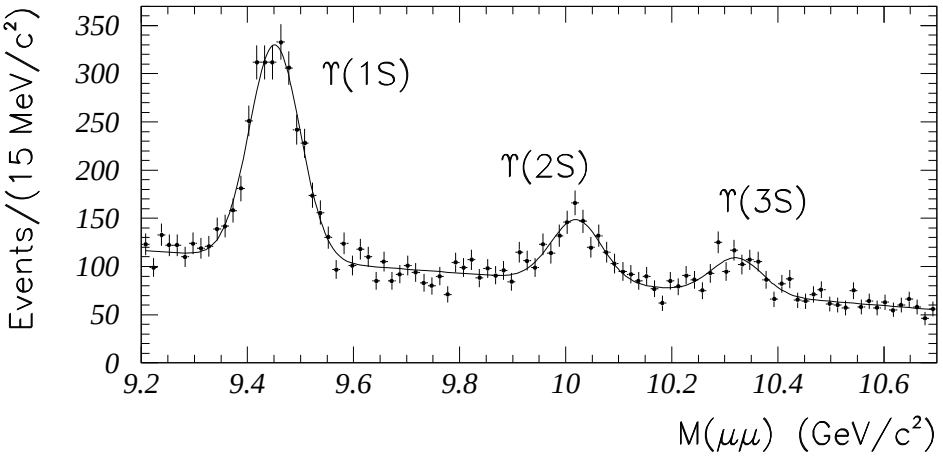


Figure 8: Comparison of measured and predicted upsilon resonance peaks are used to check the momentum calibration of the CDF spectrometer.

The electron transverse energies are determined from their measurement in the calorimeter. In order to set the calibration of the electromagnetic calorimeter, the momentum from the spectrometer is compared to the energy measured in the calorimeter for a sample of electrons from W decays. The scale of the calorimeter is adjusted until the measured E/p ratio (shown in Fig. 9) agrees with that expected. This transfer procedure contributes 110 MeV to the calibration uncertainty on M_W, yielding a total calibration uncertainty of 120 MeV in the electron channel.

Recall that M_T depends on $\vec{p}_T^{\,\ell}$ and $\vec{p}_T^{\,\nu}$. Since $\vec{p}_T^{\,\nu}$ is determined from the missing $\vec{p}_T$, it depends on $\vec{u}$, the measured transverse momentum of the hadrons recoiling against the W direction: $\vec{p}_T^{\,\nu} = \not{\vec{p}}_T = \vec{p}_T^{\,W} - \vec{p}_T^{\,\ell} = -\vec{u} - \vec{p}_T^{\,\ell}$. Many systematic studies, therefore, concentrate on properly modeling the measurement of $\vec{p}_T^{\,\ell}$ and $\vec{u}$. Once the lepton scale is established, the uncertainties from the p_T^W distribution and from the measurement of u must be established. In most cases, this is done using the $Z \to \ell\ell$ events, where each event has independent measurements of p_T^W from the hadrons and from the leptons. Underlying events from real Z events are used directly in the W simulation to model the recoil response of the detector. The distribution that is used to control this process is $u_\perp$, the component of $\vec{u}$ perpendicular to the lepton direction in $W \to \ell\nu$ events. The p_T distribution of the Z events used in the W simulation is scaled until the $u_\perp$ distribution of the simulation matches that of the W sample events, as shown in Fig. 10. The

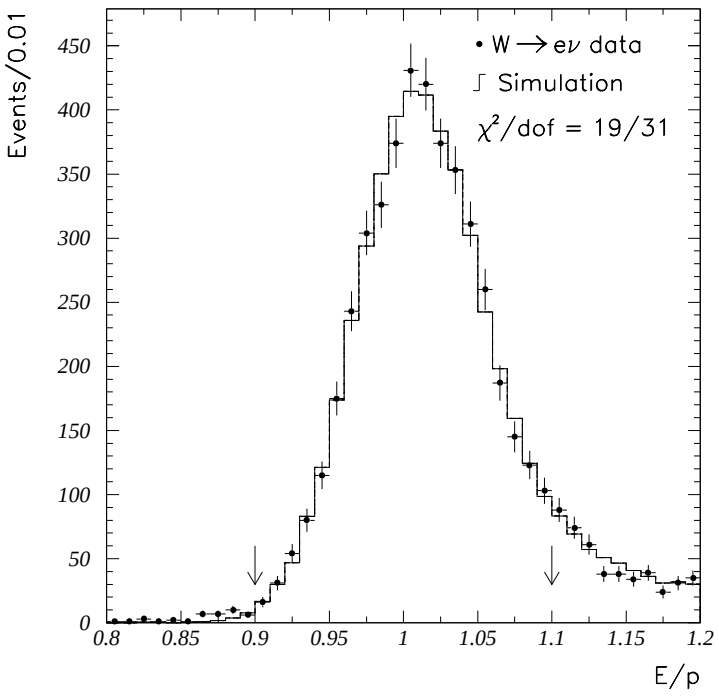


Figure 9: The CDF ratio of energy measured in the calorimeter to momentum measured in the spectrometer for the $W \to e\nu$ sample. This distribution is used in the transfer of the energy calibration scale to the calorimeter.

uncertainty in M_W resulting from the remaining uncertainty on p_T^W is 45 MeV. In addition, there is a 60 MeV uncertainty due to the modeling of the recoil measurement.

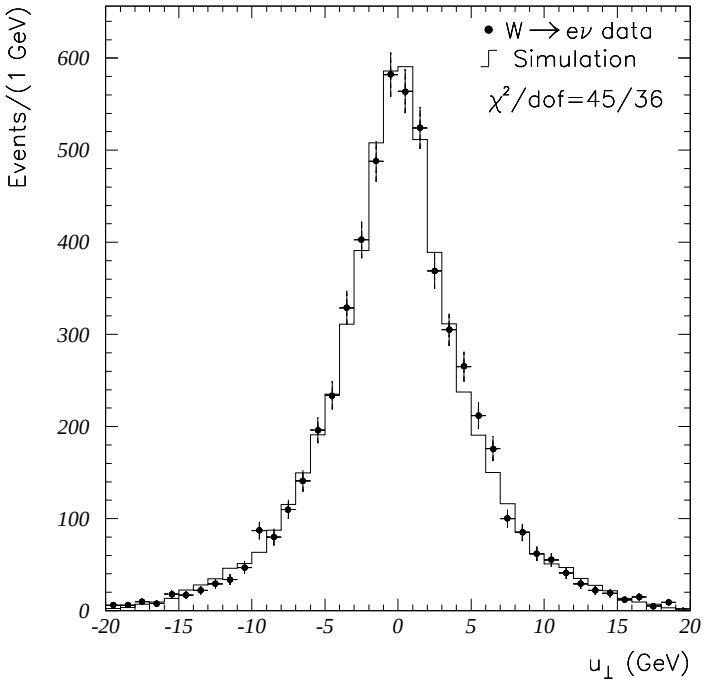


Figure 10: The CDF distribution of $u_\perp$ (see text) in $W \to e\nu$ events (points), compared with the simulation model (histogram).

The pdf's largely determine the longitudinal production distribution of the W bosons and consequently influence the observed transverse mass distribution after acceptance effects are taken into account. The best constraints on the pdf uncertainties come from the measurement of the W charge asymmetry described earlier. Figure 11 shows the change in the measured mass for different pdf sets, with respect to that obtained with the MRS D$'_-$ pdf set. The abscissa in this figure is a measure of the deviation of the CDF Run 1A measured charge asymmetry from that predicted by each pdf set, and the uncertainty on M_W is obtained by considering only those sets which are within $\pm 2\sigma$ from the best agreement with the asymmetry. The uncertainty from pdf's is thus determined to be 50 MeV.

The M_W results just described are shown in Fig. 12 along with other currently available measurements of the W (Refs. 21–23). The DØ measurement[23]

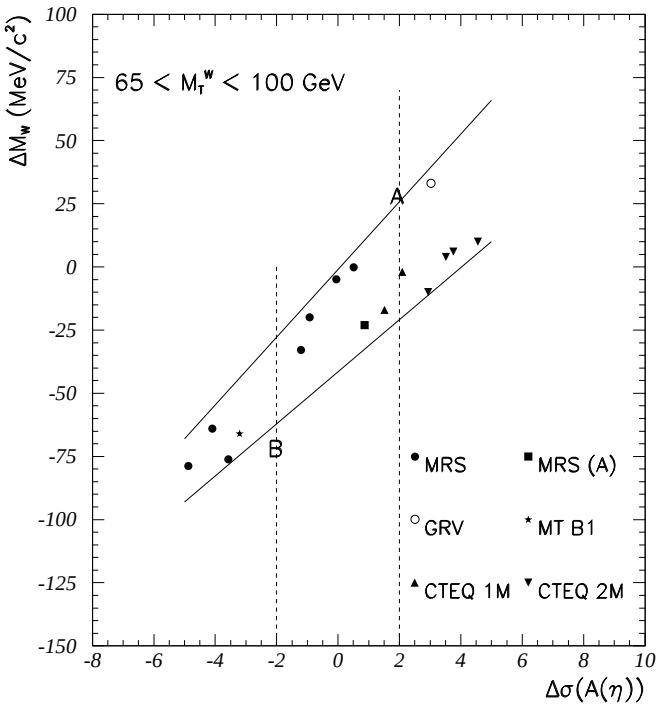


Figure 11: The change in the fitted W boson mass is plotted against the deviation from the CDF Run 1A data of the predicted asymmetry for various pdf's. The lines show the limits used in establishing the uncertainty on M_W.

is preliminary and is expected to be superseded by a final result soon. A preliminary world average of 80.26 ± 0.16 GeV is also shown, which was obtained assuming a common systematic error of 85 MeV among the different measurements. The precision electroweak measurements of the Z boson from LEP[24] and SLC[25] are also sensitive to the W mass through the relationship between M_W and $sin^2\theta_W$, and the corresponding predictions for M_W are shown as well. The hadron collider measurement is in very good agreement with the LEP prediction but disagrees somewhat with the SLC prediction. Another indirect measurement of $M_W = 80.24 \pm 0.25$ GeV is obtained from the $sin^2\theta_W$ measurement in neutrino scattering,[26] which is also in good agreement with the direct measurement.

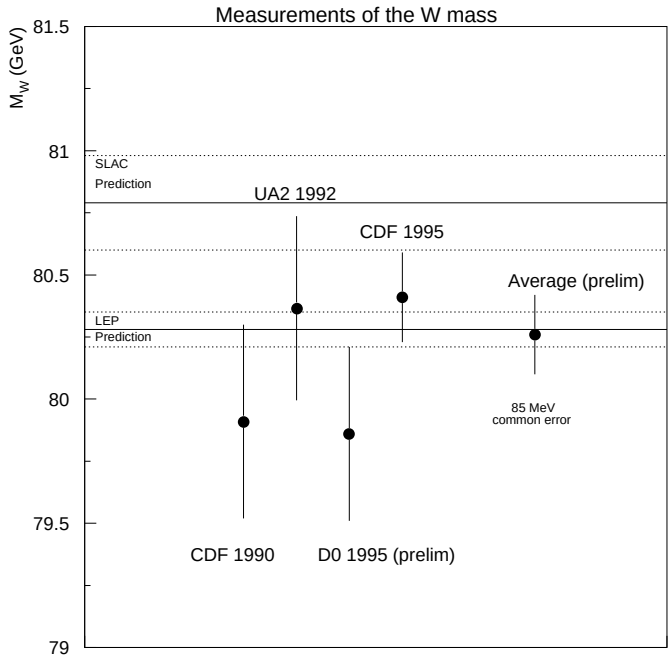


Figure 12: The direct measurements of M_W from hadron collider experiments and their average (points) is compared to predictions based on Z pole measurements at LEP and SLC (horizontal bands).

In the Standard Model, the value of M_W is sensitive to the mass of the top quark (quadratically) and to the mass of the Higgs boson (logarithmically) through radiative corrections. This relationship is shown in Fig. 13 for three different values of the Higgs mass.[27] The uncertainties on the predictions are shown as the dotted

lines and are dominated by uncertainties on the value of α_{EM} at the vector boson masses. The world average for M_W is plotted along with the average value of the top quark mass from CDF (Ref. 28) and DØ (Ref. 29). With the present uncertainties, the data are consistent with all the values of the Higgs mass shown, but more precise future measurements of M_W and M_{top} might be able to constrain M_{Higgs}.

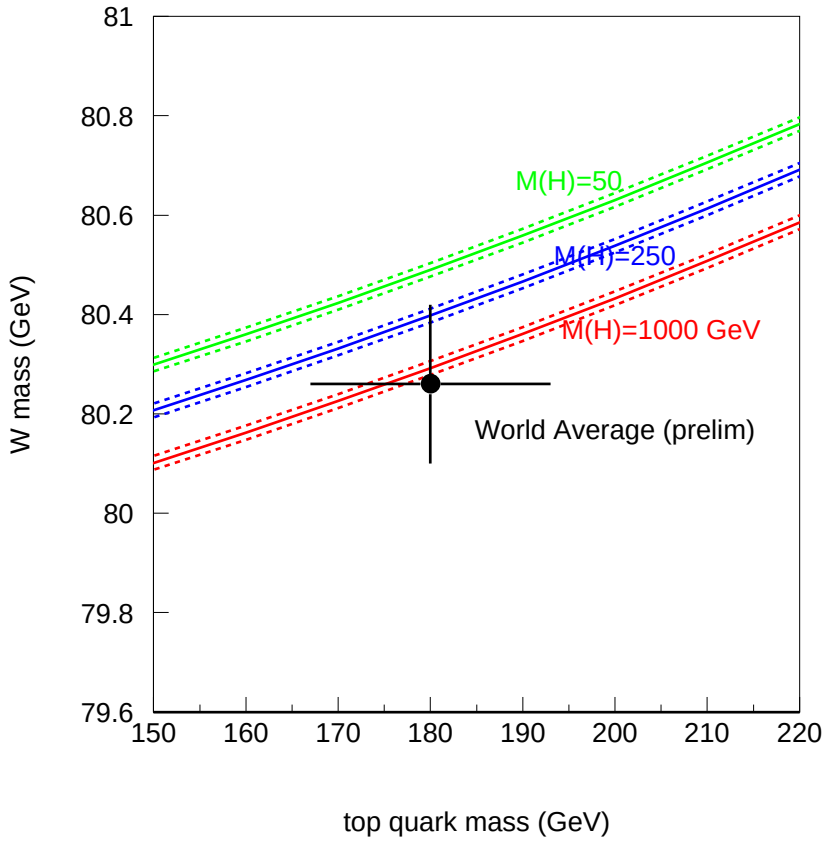


Figure 13: The relationship between M_W and M_{top} in the minimal Standard Model for different values of the Higgs mass. The dotted lines show the variation with the uncertainty on $\alpha_{EM}(M_W)$, which is the dominant uncertainty. The point shows the world averages of the measurements of M_W and M_{top}.

4 Studies of Diboson Final States

An interesting consequence of the non-Abelian gauge symmetry $SU(2)_L \times U(1)_Y$ is that the electroweak gauge bosons should be self-coupling. In particular, the SM predicts nonzero trilinear couplings for $WW\gamma$ and WWZ. It is possible to test these couplings by studying final states involving two bosons: $W\gamma$, $Z\gamma$, WW, WZ, etc. The amplitudes from the s-channel trilinear diagrams usually interfere destructively with amplitudes from other u- and t-channel diagrams, and the diboson production rate is near its minimum for the trilinear coupling strengths dictated by the Standard Model. For models with non-SM coupling values, this cancellation is spoiled, and the coupling constants must be regulated by form factors characterized by a scale Λ_{FF} in order to preserve unitarity.

A formalism has been developed to describe the $WW\gamma$ and WWZ interactions beyond the SM.[30] If Lorentz invariance, C, P, CP invariance, and $U(1)_{EM}$ gauge invariance are assumed, the most general Lagrangian describing the three-boson vertex can be written

$$\mathcal{L}_{WWV}/g_{WWV} = ig_1^V(W^\dagger_{\mu\nu}W^\mu V^\nu - W^\dagger_\mu V_\nu W^{\mu\nu}) + i\kappa_V W^\dagger_\mu W_\nu V^{\mu\nu} + \frac{i\lambda_V}{M_W^2}W^\dagger_{\lambda\mu}W^\mu_\nu V^{\nu\lambda} \tag{2}$$

where $V = \gamma$ or Z. In the SM, $\kappa_\gamma = \kappa_Z = 1$ and $\lambda_\gamma = \lambda_Z = 0$.

4.1 $W\gamma$

The most abundant diboson final state is $W\gamma$. It is studied in both the $e\nu\gamma$ and $\mu\nu\gamma$ channels. The most important selection criteria used by CDF and DØ are shown in Table 3. Note that a minimum photon E_T and a minimum separation between the photon and the lepton, $\Delta R(\ell\gamma) = \sqrt{\Delta\phi(\ell\gamma)^2 + \Delta\eta(\ell\gamma)^2}$, are required. These requirements are necessary even in the theoretical predictions in order to avoid infrared and collinear divergences from photon radiation from the final state leptons.

CDF reports 109 $W\gamma$ events with $E_T^\gamma > 7$ GeV from a preliminary analysis of 67 pb^{-1} of data from Run 1A (Ref. 31) and part of Run 1B. DØ has a final sample 23 events with $E_T^\gamma > 10$ GeV from Run 1A (14 pb^{-1}) (Ref. 32). The photon E_T spectra from these samples is shown in Fig. 14. The main background in these samples is $W + jet$ events in which the jet fakes a photon. Both the normalization and the shape of these spectra are in good agreement with the

Requirement	CDF	DØ
electron acceptance	$\|\eta_e\| < 1.1$	$\|\eta_e\| < 1.1$ or $1.5 < \|\eta_e\| < 2.5$
muon acceptance	$\|\eta_\mu\| < 0.6$	$\|\eta_\mu\| < 1.7$
photon acceptance	$\|\eta_\gamma\| < 1.1$	$\|\eta_\gamma\| < 1.1$ or $1.5 < \|\eta_\gamma\| < 2.5$
photon E_T	$E_T^\gamma > 7$ GeV	$E_T^\gamma > 10$ GeV
γ-ℓ separation	$\Delta R(\ell\gamma) > 0.7$	$\Delta R(\ell\gamma) > 0.7$

Table 3: Selection requirements for $W\gamma$ and $Z\gamma$ events.

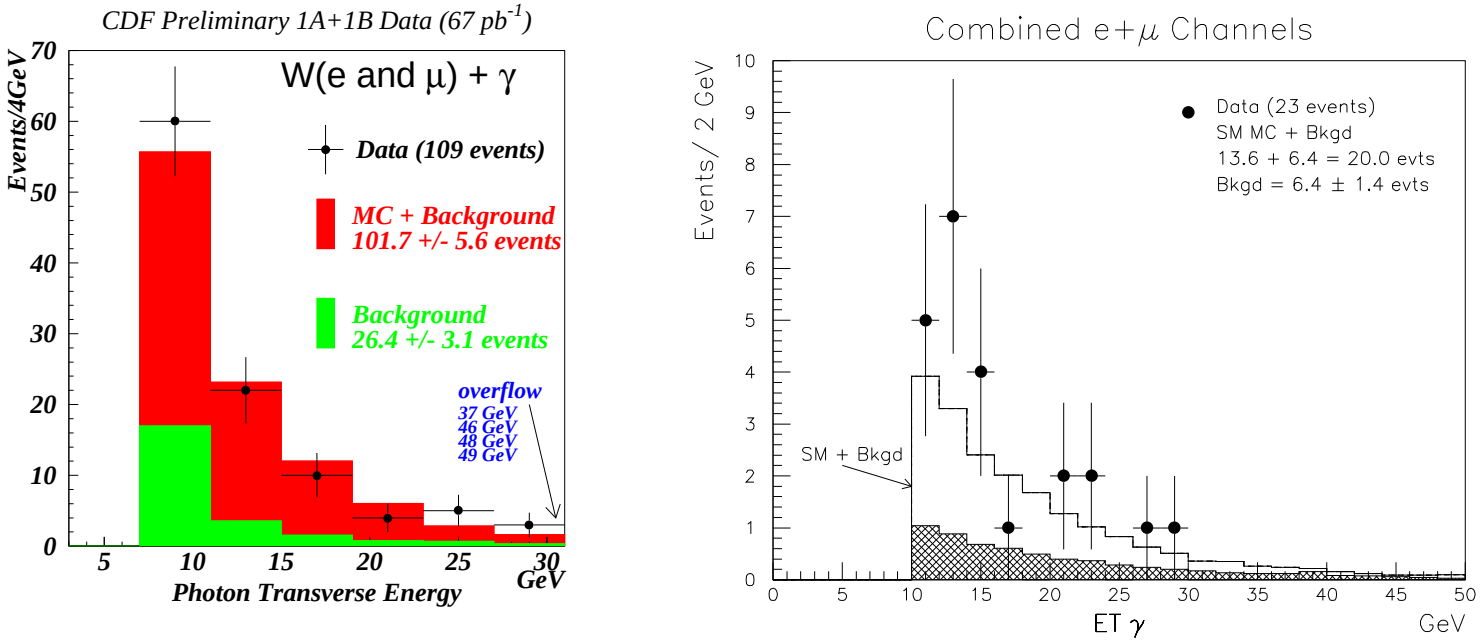


Figure 14: The transverse energy spectra of the photons in the $W\gamma$ samples from (left) preliminary CDF Run 1A plus partial Run 1B (67 pb^{-1}) and (right) DØ Run 1A (14 pb^{-1}).

Standard Model predictions. Both of these facts are significant, since a deviation from SM couplings should result in a higher overall cross section and a harder spectrum for E_T^γ. Limits on $\Delta\kappa$ ($\Delta\kappa = \kappa - 1$) and λ are obtained from fits to these E_T^γ spectra. The 95% CL contours are shown in Fig. 15. The limits on the axes are

$$\text{CDF (prelim.):} \begin{cases} -1.8 < \Delta\kappa < 2.0 & (\lambda = 0) \\ -0.7 < \lambda < 0.6 & (\Delta\kappa = 0) \end{cases}$$

$$\text{DØ:} \begin{cases} -1.6 < \Delta\kappa < 1.8 & (\lambda = 0) \\ -0.6 < \lambda < 0.6 & (\Delta\kappa = 0) \end{cases}$$

where a form factor with $\Lambda_{FF} = 1.5$ TeV has been assumed in both analyses. The magnetic dipole moment ($\mu_W = (\kappa + \lambda + 1)e/2M_W$) and the electric quadrupole moment ($Q_W = -(\kappa - \lambda)e/M_W^2$) of the W boson can be expressed as linear combinations of κ and λ. The lines for $\mu_W = 0$ and $Q_W = 0$ are shown in Fig. 15, and it can be seen that the point where both moments vanish can now be excluded.

A particular SM prediction for $W\gamma$ production is that the destructive interference between the s-channel diagram and the t- and u-channel diagrams should produce a sharp minimum in the angular distribution. This so called "gauge zero"[33] should occur at $\cos(\theta^*) = \pm 0.3$, where θ^* is the angle between the photon and the incoming quark direction in the $W^\mp\gamma$ rest frame. The calculation of θ^* requires one to solve for the unknown longitudinal component of the neutrino momentum, which generally involves a two-fold ambiguity. The cancellation of the amplitudes is destroyed as the couplings deviate from their SM values, so the gauge zero can provide another test for anomalous couplings. The prominence of the zero in the distribution is degraded, however, by the presence of background, by resolution effects, and by contributions from radiative decays. The preliminary distribution of $\cos(\theta^*)$ from CDF is shown in Fig. 16. This sample has had additional requirements placed on it to suppress the contributions from radiative decays, and the $W^+\gamma$ events have been added to the $W^-\gamma$ events after inverting the sign of $\cos(\theta^*)$.

CDF has also investigated some charge asymmetries in $W\gamma$ production using an independent sample of events in which the photons are detected in the region $1.1 < |\eta_\gamma| < 2.4$. In Fig. 17, the first plot shows the rapidity distribution of the photons, signed by the charge of the lepton from the W boson decay. This shows a strong asymmetry, which originates in part from the difference in magnitude of electric charge between *up* and *down* quarks. The forward/backward asymmetry

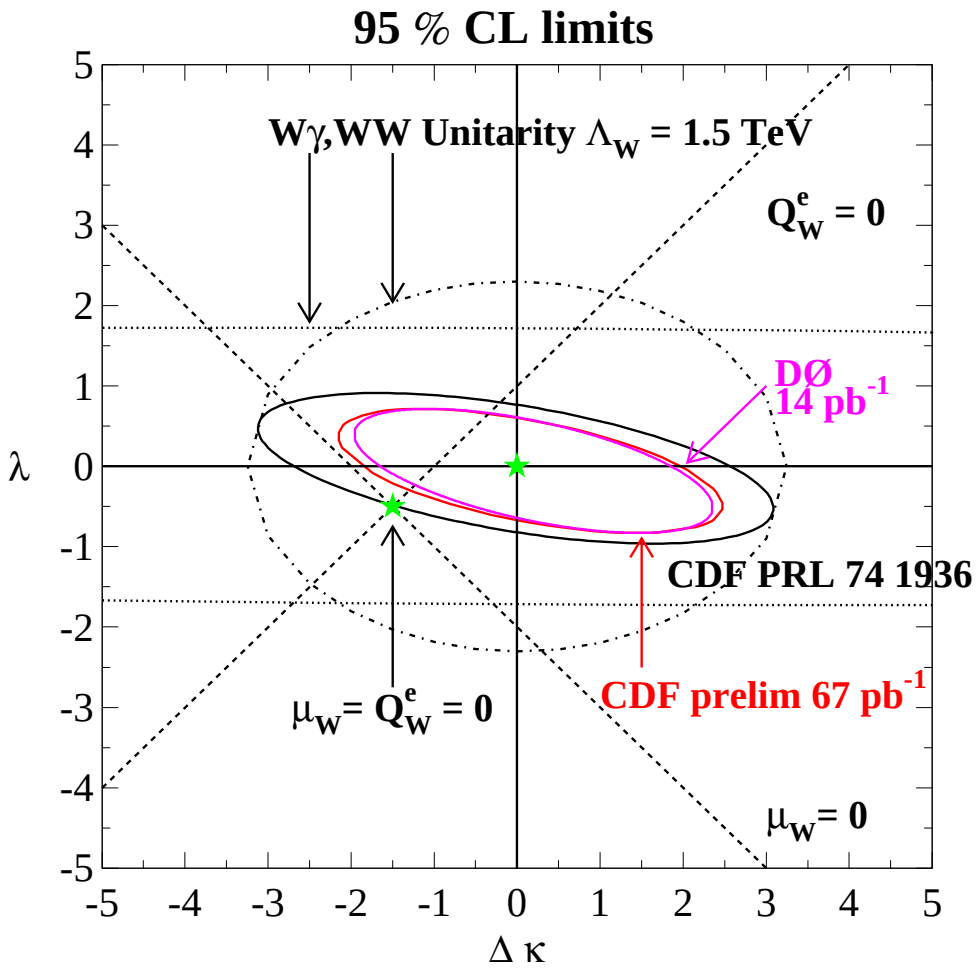


Figure 15: The limits (95% CL) on anomalous coupling parameters λ and $\Delta\kappa$ derived from the $W\gamma$ samples. The dashed lines show the places where the magnetic dipole moment and electric quadrapole moment vanish, and the star indicates the Standard Model prediction. The dot-dashed ellipses show the limits implied by unitarity when a form factor scale of 1.5 TeV is assumed.

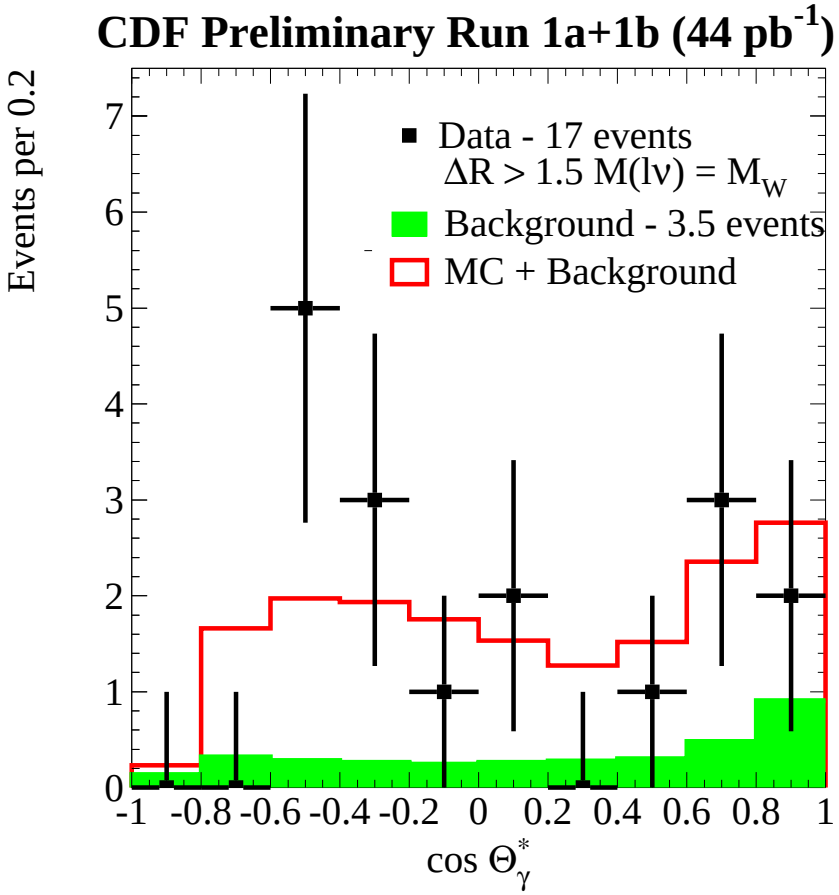


Figure 16: The preliminary CDF data distribution (points) for $\cos\theta^*$ is compared to the expected signal plus background distribution (histogram). The shaded portion of the histogram shows the expected background contribution.

measured on this sample is 0.77 ± 0.07, and it is in good agreement with the prediction of 0.76 ± 0.04. The second figure shows the rapidity difference between the lepton and the photon ($\eta_\gamma - \eta_\ell$), where again the quantity is signed by the lepton charge. The dip in the middle results largely from the requirement that the photons are in the end regions while the leptons are central. The asymmetry measured for the rapidity difference is 0.70 ± 0.04.

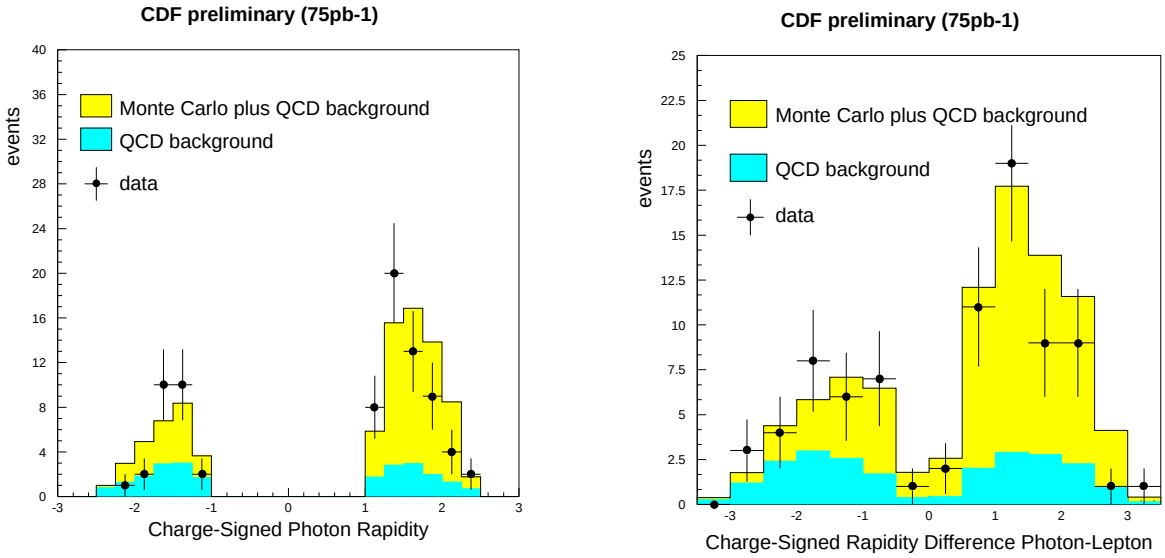


Figure 17: The CDF preliminary distributions of photons in the "plug" region ($1.1 < |\eta| < 2.4$) in $W\gamma$ events. The rapidity is shown on the left, and the rapidity difference $\eta(\gamma) - \eta(\ell)$ of the photon and lepton is shown on the right, and each is signed by the charge of the lepton. The points are the data and the histograms are the predictions for Standard Model plus background.

4.2 $Z\gamma$

The Z boson is a neutral particle, so the SM predicts no direct $Z\gamma$ couplings, although $Z\gamma$ production is still allowed through the t- and u-channels. A more general non-SM formalism,[34] similar to that used for the $W\gamma$, allows for nonzero anomalous couplings given by the parameters h^Z_{10}, h^Z_{20} (CP violating) and h^Z_{30}, h^Z_{40} (CP conserving). The same general features apply: anomalous couplings tend to increase the production cross section and make the photon spectra harder.

Both DØ and CDF have completed preliminary $Z\gamma$ analyses which include Run 1A (Refs. 35, 36) and part of Run 1B. The E_T^γ spectra from these samples is

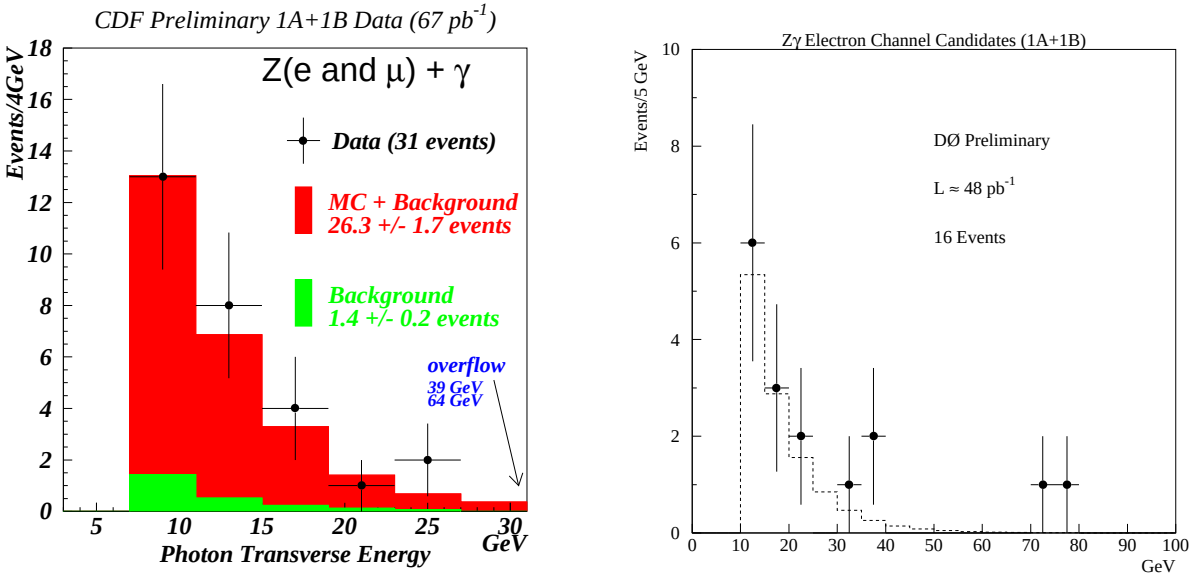


Figure 18: Preliminary distribution of the transverse energy of photons in $Z\gamma$ events from Run 1A plus partial Run 1B from (left) CDF (67 pb^{-1}) and (right) DØ (48 pb^{-1}). The points are the data and the histograms are the expectations of Standard Model plus background.

shown in Fig. 18. The limits on anomalous couplings from CDF are derived from fits to the E_T^γ distribution of the sample shown, while the DØ result is presently only from the Run 1A spectrum (14 pb^{-1}). The limit contours are shown in Fig. 19. The limits on the axes are:

$$\text{CDF (prelim.):} \begin{cases} -1.6 < h_{30}^Z(h_{10}^Z) < 1.6 & (h_{40}^Z(h_{20}^Z) = 0) \\ -0.4 < h_{40}^Z(h_{20}^Z) < 0.4 & (h_{30}^Z(h_{10}^Z) = 0) \end{cases}$$

$$\text{DØ:} \begin{cases} -1.9 < h_{30}^Z(h_{10}^Z) < 1.8 & (h_{40}^Z(h_{20}^Z) = 0) \\ -0.5 < h_{40}^Z(h_{20}^Z) < 0.5 & (h_{30}^Z(h_{10}^Z) = 0) \end{cases}$$

where a form factor with $\Lambda_{FF} = 0.5$ TeV has been assumed.

4.3 WW and WZ

The cleanest channels for detecting pairs of W bosons are those where both bosons decay leptonically. The signatures are then $ee+\not\!\!p_T$, $e\mu+\not\!\!p_T$, and $\mu\mu+\not\!\!p_T$, where the $\not\!\!p_T$ comes from the vector sum of the two neutrino momenta. DØ has searched for WW production in these modes in Run 1A (Ref. 37) (14 pb^{-1}), and CDF has a preliminary result based on Run 1A and part of Run 1B (67 pb^{-1}). DØ

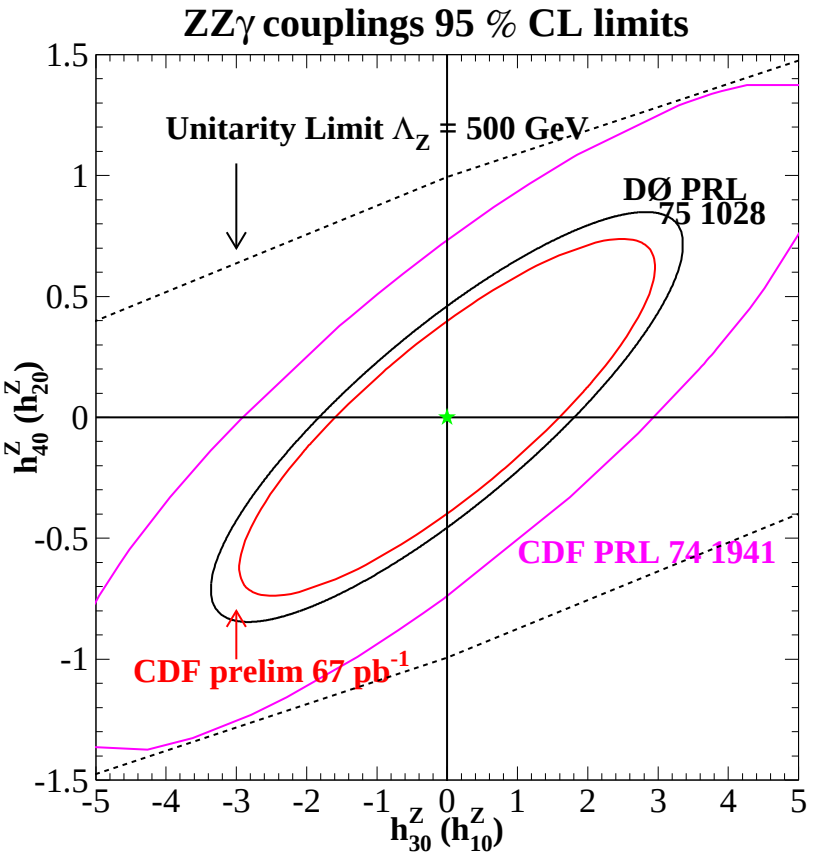


Figure 19: The limits (95% CL) on anomalous coupling parameters $h_{10}^Z, h_{20}^Z, h_{30}^Z, h_{40}^Z$ derived from the $Z\gamma$ samples. The Standard Model prediction is at the center (0,0). The dotted ellipses show the limits implied by unitarity when a form factor scale of 0.5 TeV is assumed.

observes one event, with an expected background of 0.56 ± 0.13 events and expected SM signal of 0.47 ± 0.07 events. CDF observes five events, with expectations of 1.23 ± 0.43 events background and 2.6 ± 0.9 events signal. From these samples, DØ calculates an upper limit (95% CL) of 87 pb on the cross section for WW production. CDF calculates a cross section of $13.8 \pm^{9.2}_{7.4} \pm 2.9$ pb. DØ also quotes limits on anomalous coupling from this analysis, with the assumption that $\kappa^\gamma = \kappa^Z$ and $\lambda^\gamma = \lambda^Z$:

$$\text{DØ:} \begin{cases} -2.6 < \Delta\kappa < 2.8 & (\lambda = 0) \\ -2.1 < \lambda < 2.1 & (\Delta\kappa = 0). \end{cases}$$

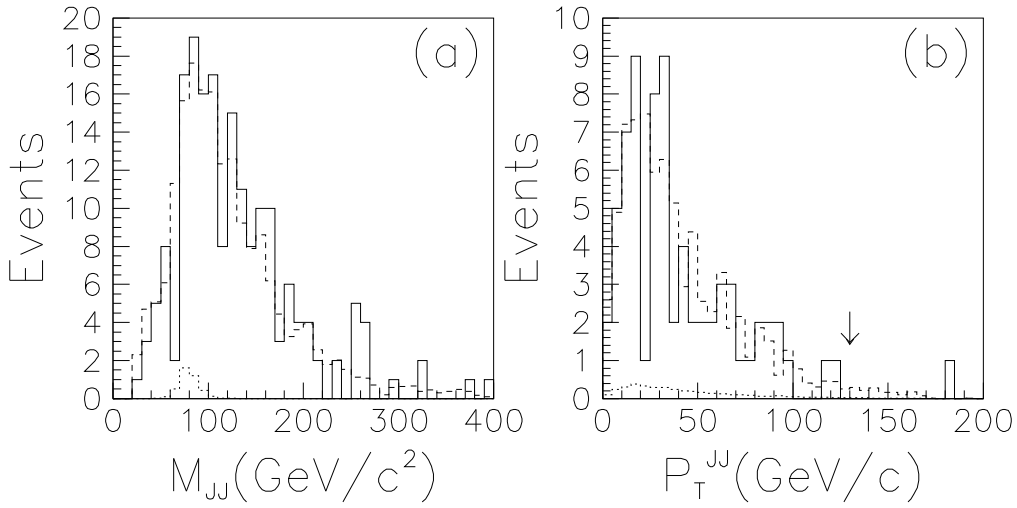


Figure 20: The sample selection for the CDF $WW \to \ell\nu jj$ analysis. The invariant mass spectrum of the two jets is shown in (a) prior to any cut on $M(jj)$. The $p_T(jj)$ spectrum after the cut of $60 < M(jj) < 110$ GeV is shown in (b). The arrow indicates the final cut of $p_T(jj) > 130$. The solid histograms are the data, the dashed histograms are the expected background, and the dotted histograms are the expectations from SM diboson production.

The small rates in the pure leptonic channels of WW decay make it attractive to consider the case where one W boson decays hadronically to two jets. The signature is then $\ell jj + \not{p}_T$. In this case, the experiments do not distinguish $W \to jj$ from $Z \to jj$, so it is the sum of WW and WZ which contributes to signal. The largest background is from production of single W bosons accompanied by two jets. This background is reduced by both CDF and DØ by requiring that the jj

invariant mass be consistent with that of a W or Z boson. Figure 20 shows the jj mass spectrum from CDF[38] before the requirement $60 < M_{jj} < 110$ GeV was imposed, and the p_T^{jj} spectrum afterwards. Even after the M_{jj} requirement, the sample is dominated by single W plus two-jet events. Since the high p_T portion of the spectrum is greatly enhanced by anomalous couplings, CDF requires $p_T^{jj} > 130$ GeV, after which one event remains. The preliminary DØ analysis is similar, except that no cut is made on p_T^{jj}, and instead, a fit is made to the $p_T^{e\nu}$ spectrum (shown in Fig. 21) from which the anomalous coupling limits are derived. With the same assumption of the equality of λ and κ for photons and Z bosons, the limits obtained from the $\ell\nu jj$ analyses are:

$$\text{CDF:} \begin{cases} -1.11 < \Delta\kappa < 1.27 & (\lambda = 0) \\ -0.81 < \lambda < 0.84 & (\Delta\kappa = 0) \end{cases}$$

$$\text{DØ (prelim.):} \begin{cases} -0.89 < \Delta\kappa < 1.07 & (\lambda = 0) \\ -0.66 < \lambda < 0.67 & (\Delta\kappa = 0) \end{cases}$$

where a form factor with $\Lambda_{FF} = 1$ TeV is used by CDF and $\Lambda_{FF} = 1.5$ TeV is used by DØ.

5 Conclusion

Recent analyses of W and Z boson events from the Fermilab Tevatron Collider have resulted in a considerable improvement in the measurements of the properties of the W boson. The W mass is now measured to about 0.2%, while the W width is measured to about 15% (3%) directly (indirectly). The best limit on the anomalous coupling parameters $\Delta\kappa$ and λ are around 1.1 and 0.6, respectively. The anomalous couplings of the Z bosons have also been tested.

In most cases, these results were obtained from only a fraction of the data that will be available from the complete Run 1 of the Tevatron Collider. A total sample of about 100 pb^{-1} is expected for each experiment. When these data are analyzed (within the next year, probably), the precision of the electroweak measurements should be considerably improved. The uncertainty on the W mass, for example, should be reduced to around 80 MeV. The next major improvement is then expected in Run 2 of the Tevatron Collider, which is scheduled to begin in 1999 and to provide samples of about 2000 pb^{-1} for the upgraded versions of CDF and DØ.

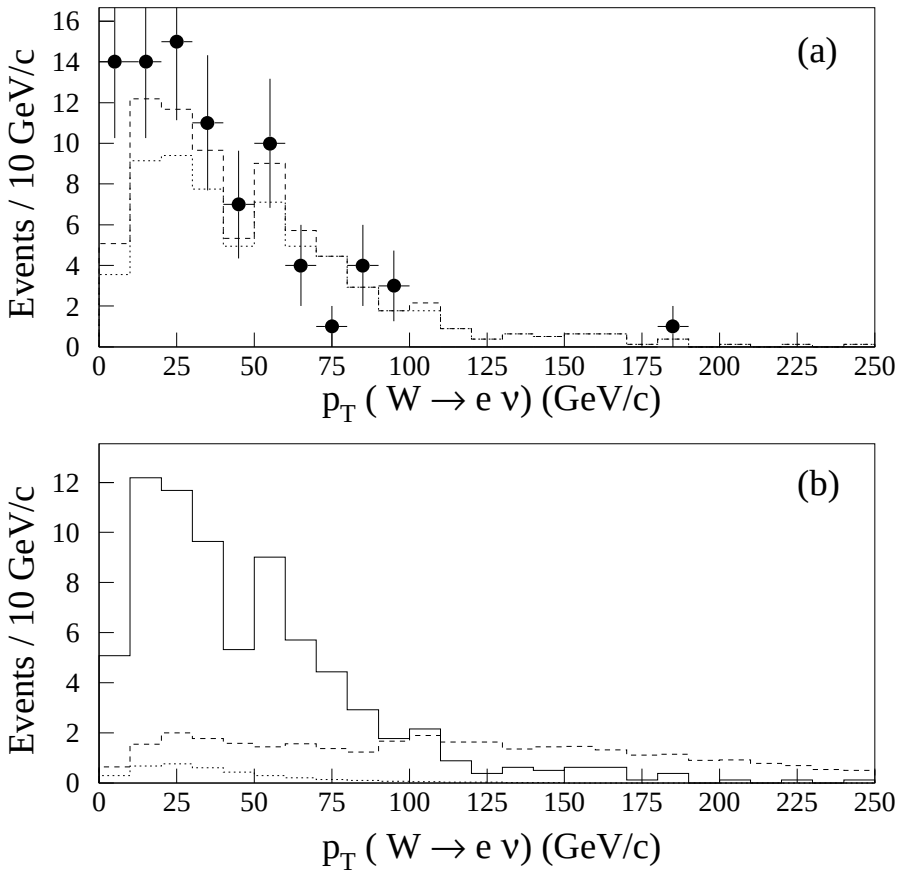


Figure 21: The preliminary DØ $p_T(e\nu)$ distributions for $WW \rightarrow e\nu jj$. The upper figure shows the data (points) compared to the expected background (dashed histogram). The lower figure shows the expected background (solid histogram) compared to predictions for SM WW production (dotted) and as an example ($\Delta\kappa = 2, \lambda = 1.5$) of anomalous couplings (dashed).

References

[1] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **122**, 103 (1983); G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **129**, 273 (1983).

[2] M. Banner *et al.* (UA2 Collaboration), Phys. Lett. B **122**, 476 (1983); P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B **129**, 130 (1983).

[3] F. Abe *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **271**, 387 (1988).

[4] S. Abachi *et al.* (DØ Collaboration), Nucl. Instrum. Methods A **338**, 185 (1994).

[5] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **75**, 1456 (1995).

[6] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **64**, 152 (1990); F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **44**, 29 (1991); F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **69**, 28 (1992).

[7] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 220 (1994); F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **52**, 2624 (1995); W. F. Badgett in *Proceedings of the 8th DPF Meeting*, August 2-6 (1994), 431 (Albuquerque, New Mexico).

[8] C. Albajar *et al.* (UA1 Collaboration), Phys. Lett. B **253**, 503 (1991).

[9] J. Alitti *et al.* (UA2 Collaboration), Phys. Lett. B **276**, 365 (1992).

[10] R. Hamberg, W. L. van Neerven, and T. Matsuura, Nucl. Phys. B **359**, 343 (1991); W. L. van Neerven and E. B. Zijlstra, Nucl. Phys. **382**, 11 (1992).

[11] Particle Data Group, L. Montanet *et al.*, Phys. Rev. D **50**, 1173 (1994).

[12] J. L. Rosner, M. P. Worah, and T. Takeuchi, Phys. Rev. D **49**, 1363 (1994).

[13] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 341 (1995).

[14] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 850 (1995).

[15] W. Giele, E. Glover, and D. A. Kosower, Nucl. Phys. B **403**, 663 (1993).

[16] M. Glück, E. Reya, and A. Vogt, Z. Phys. C **67**, 433 (1995); S. Kretzer, E. Reya, and M. Stratmann, DO-TH 94/26, December 1994.

[17] H. L. Lai *et al.*, Phys. Rev. D **51**, 4763 (1995).

[18] A. D. Martin, R. G. Roberts, and W. J. Stirling, Phys. Rev. D **50**, 6734 (1994).

[19] P. de Barbaro in *Proceedings of the 10th Topical Workshop on Proton-Antiproton Collider Physics*, Batavia, Illinois, 9-13 May, 1995.

[20] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **75**, 11 (1995); F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **52**, 4784 (1995).

[21] J. Alitti *et al.* (UA2 Collaboration), Phys. Lett. B **276**, 354 (1992).

[22] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **65**, 2243 (1990); F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **43**, 2070 (1991); the 88/89 cross sections should be increased by 1/0.907 for consistent comparison with the Run 1 cross sections.

[23] S. Rajagopalan in *Proceedings of Les Rencontres de Physique de la Vallee d'Aoste*, La Thuile, Italy, 5-11 March, 1995.

[24] The LEP Electroweak Working Group, CERN Report No. CERN/PPE/94-187 (1994), unpublished.

[25] K. Abe *et al.* (SLD Collaboration), Phys. Rev. Lett. **73**, 25 (1994).

[26] C. Arroyo *et al.* (CCFR Collaboration), Phys. Rev. Lett. **72**, 3452 (1994).

[27] F. Halzen and B. A. Kniehl, Nucl. Phys. B **353**, 567 (1991).

[28] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995).

[29] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[30] K. Hagiwara, R. D. Peccei, D. Zeppenfeld, and K. Hikasa, Nucl. Phys. B **274**, 253 (1987); U. Baur and D. Zeppenfeld, Phys. Lett. B **201**, 383 (1988).

[31] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 1936 (1995).

[32] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **75**, 1034 (1995).

[33] K. O. Mikaelian, Phys. Rev. D **17**, 750 (1978); K. O. Mikaelian, M. A. Samuel, and D. Sahdev, Phys. Rev. Lett. **43**, 746 (1979); R. W. Brown, K. O. Mikaelian, and D. Sahdev, Phys. Rev. D **20**, 1164 (1979); T. R. Grose and K. O. Mikaelian, Phys. Rev. D **23**, 123 (1981); S. J. Brodsky and R. W. Brown, Phys. Rev. Lett. **49**, 966 (1982); M. A. Samuel, Phys. Rev. D **27**, 2724 (1983); R. W. Brown, K. L. Kowalski, and S. J. Brodsky, Phys. Rev. D **28**, 624 (1983); R. W. Brown and K. L. Kowalski, Phys. Rev. D **29**, 2100 (1984).

[34] K. Hagiwara *et al.*, Nucl. Phys. B **282**, 253 (1987).

[35] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 1941 (1995).

[36] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **75**, 1028 (1995).

[37] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **75**, 1023 (1995).

[38] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **75**, 1017 (1995).

DYNAMICAL ELECTROWEAK SYMMETRY BREAKING AND THE TOP QUARK*

R. Sekhar Chivukula
Department of Physics, Boston University
590 Commonwealth Ave., Boston MA 02215
E-mail: sekhar@bu.edu

Talk presented at the SLAC Topical Workshop
Stanford, July 19-21, 1995

BUHEP-95-23 & hep-ph/9509384

ABSTRACT

In this talk, I discuss theories of dynamical electroweak symmetry breaking, with emphasis on the implications of a heavy top-quark on the weak-interaction ρ parameter.

*AN ABBREVIATED VERSION[1] OF THIS TALK WAS PRESENTED AT THE *WORKSHOP ON TOP QUARK PHYSICS*, IOWA STATE UNIVERSITY, AMES, IA, MAY 25-26, 1995 AND THE *YUKAWA INTERNATIONAL SEMINAR '95*, YUKAWA INSTITUTE, KYOTO, AUG. 21-25, 1995.

1 What's Wrong with the Standard Model?

In the standard one-doublet Higgs model, one introduces a fundamental scalar doublet of $SU(2)_W$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad , \tag{1.1}$$

which has a potential of the form

$$V(\phi) = \lambda \left(\phi^\dagger \phi - \frac{v^2}{2} \right)^2 \quad . \tag{1.2}$$

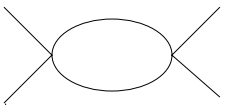
In the potential, v^2 is assumed to be positive in order to favor the generation of a nonzero vacuum expectation value for ϕ. This vacuum expectation value breaks the electroweak symmetry, giving mass to the W and Z.

This explanation of electroweak symmetry breaking is unsatisfactory for a number of reasons. For one thing, this model does not give a dynamical explanation of electroweak symmetry breaking. For another, when embedded in theories with additional dynamics at higher energy scales, these theories are technically unnatural.[2]

Perhaps most unsatisfactory, however, is that theories of fundamental scalars are probably "trivial,"[3] *i.e.*, it is not possible to construct an interacting theory of scalars in four dimensions that is valid to arbitrarily short-distance scales. In quantum field theories, fluctuations in the vacuum screen charge—the vacuum acts as a dielectric medium. Therefore, there is an effective coupling constant which depends on the energy scale (μ) at which it is measured. The variation of the coupling with scale is summarized by the β–function of the theory

$$\beta(\lambda) = \mu \frac{d\lambda}{d\mu} \quad . \tag{1.3}$$

The only coupling in the Higgs sector of the Standard Model is the Higgs self-coupling λ. In perturbation theory, the β-function is calculated to be



$$\rightarrow \beta = \frac{3\lambda^2}{2\pi^2} \quad . \tag{1.4}$$

Using this β–function, one can compute the behavior of the coupling constant as a function of the scale.* One finds that the coupling at a scale μ is related to the coupling at some higher

*Since these expressions were computed in perturbation theory, they are only valid when $\lambda(\mu)$ is sufficiently small. For large couplings, we must rely on nonperturbative lattice Monte Carlo studies,[4,5] which show behavior similar to that implied by the perturbative expressions derived here.

scale Λ by

$$\frac{1}{\lambda(\mu)} = \frac{1}{\lambda(\Lambda)} + \frac{3}{2\pi^2}\log\frac{\Lambda}{\mu} \quad . \tag{1.5}$$

In order for the Higgs potential to be stable, $\lambda(\Lambda)$ has to be positive. This implies

$$\frac{1}{\lambda(\mu)} \geq \frac{3}{2\pi^2}\log\frac{\Lambda}{\mu} \quad . \tag{1.6}$$

Thus, we have the bound

$$\lambda(\mu) \leq \frac{2\pi^2}{3\log\left(\frac{\Lambda}{\mu}\right)} \quad . \tag{1.7}$$

If this theory is to make sense to arbitrarily short distances, and hence arbitrarily high energies, we should take Λ to ∞ while holding μ fixed at about 1 TeV. In this limit, we see that the bound on λ goes to zero. In the continuum limit, this theory is trivial; it is free field theory.

The theory of a relatively light weakly coupled Higgs boson can be self-consistent to a very high energy. For example, if the theory is to make sense up to a typical GUT scale energy, 10^{16} GeV, then the Higgs boson mass has to be less than about 170 GeV.[6] In this sense, although a theory with a light Higgs boson does not really answer any of the interesting questions (*e.g.*, it does not explain *why* $SU(2)_W \times U(1)_Y$ breaking occurs), the theory does manage to postpone the issue up to higher energies.

2 Dynamical Electroweak Symmetry Breaking

2.1 Technicolor

Technicolor[7] theories strive to explain electroweak symmetry breaking in terms of physics operating at an energy scale of order a TeV. In technicolor theories, electroweak symmetry breaking is the result of chiral symmetry breaking in an asymptotically-free, strongly-interacting gauge theory with massless fermions. Unlike theories with fundamental scalars, these theories are technically natural: just as the scale Λ_{QCD} arises in QCD by dimensional transmutation, so too does the weak scale v in technicolor theories. Accordingly, it can be exponentially smaller than the GUT or Planck scales. Furthermore, asymptotically-free non-Abelian gauge theories may be fully consistent quantum field theories.

In the simplest theory,[7] one introduces the doublet of new massless fermions

$$\Psi_L = \begin{pmatrix} U \\ D \end{pmatrix}_L \qquad U_R, D_R \tag{2.8}$$

which are N's of a technicolor gauge group $SU(N)_{TC}$. In the absence of electroweak interactions, the Lagrangian for this theory may be written

$$\begin{aligned} \mathcal{L} &= \bar{U}_L i\not{D} U_L + \bar{U}_R i\not{D} U_R + \qquad (2.9) \\ &\quad \bar{D}_L i\not{D} D_L + \bar{D}_R i\not{D} D_R \qquad (2.10) \end{aligned}$$

and thus has an $SU(2)_L \times SU(2)_R$ chiral symmetry. In analogy with QCD, we expect that when technicolor becomes strong,

$$\langle \bar{U}_L U_R \rangle = \langle \bar{D}_L D_R \rangle \neq 0, \qquad (2.11)$$

which breaks the global chiral symmetry group down to $SU(2)_{L+R}$, the vector subgroup (analogous to isospin in QCD).

If we weakly gauge $SU(2) \times U(1)$, with the left-handed technifermions forming a weak doublet, and identify hypercharge with a symmetry generated by a linear combination of the T_3 in $SU(2)_R$ and technifermion number, then chiral symmetry breaking will result in the electroweak gauge group's breaking down to electromagnetism. The Higgs mechanism then produces the appropriate masses for the W and Z bosons if the F-constant of the technicolor theory (the analog of f_π in QCD) is approximately 246 GeV. (The residual $SU(2)_{L+R}$ symmetry ensures that, to lowest-order, $M_W = M_Z \cos\theta_W$ and the weak interaction ρ-parameter equals one at tree-level.[8])

2.2 Top-Mode and Strong-ETC Models

There is also a class of theories in which the scale (M) of the dynamics responsible for (all or part of) electroweak symmetry breaking can, in principle, take any value of order a TeV or greater. These models, inspired by the Nambu-Jona-Lasinio (NJL) model[9] of chiral symmetry breaking in QCD, involve a strong, but *spontaneously broken*, nonconfining gauge interaction. Examples include top quark condensate (and related) models,[10–14] as well as models with strong extended technicolor interactions.[15] When the strength of the effective four-fermion interaction describing the broken gauge interactions—*i.e.* the strength of the extended technicolor interactions in strong ETC models or the strength of other gauge interactions in top-condensate models—is adjusted close to the critical value for chiral symmetry breaking, the high-energy dynamics may play a role in electroweak symmetry breaking without driving the electroweak scale to a value of order M.

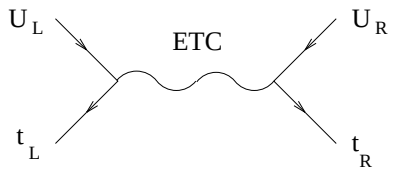
The high-energy dynamics must have the appropriate properties in order for it to play a role in electroweak symmetry breaking:[16] If the coupling constants of the high-energy theory are small, only low-energy dynamics (such as technicolor) can contribute to electroweak symmetry breaking. If the coupling constants of the high-energy theory are large and the interactions are attractive in the appropriate channels, chiral symmetry will be broken by the high-energy interactions and the scale of electroweak symmetry breaking will be of order M. If the transition between these two extremes is continuous, *i.e.*, if the chiral symmetry breaking phase transition is *second order* in the high-energy couplings, then it is possible to adjust the high-energy parameters so that the dynamics at scale M can contribute to electroweak symmetry breaking. The adjustment of the high-energy couplings is a reflection of the fine-tuning required to create a hierarchy of scales.

What is crucial is that the transition be (at least approximately) second order in the high-energy couplings. If the transition is first order, then as one adjusts the high-energy couplings, the scale of chiral symmetry breaking will jump discontinuously from approximately zero at weak coupling to approximately M at strong coupling. Therefore, if the transition is first order, it will generally not be possible to maintain any hierarchy between the scale of electroweak symmetry breaking and the scale of the high-energy dynamics.

If the transition is second order and if there is a *large* hierarchy of scales ($M \gg 1$ TeV), then close to the transition the theory may be described in terms of a low-energy effective Lagrangian with composite "Higgs" scalars—the Ginsburg-Landau theory of the chiral phase transition. However, if there is a large hierarchy, the arguments of triviality given in the first section apply to the effective low-energy Ginsburg-Landau theory describing the composite scalars: the effective low-energy theory would be one which describes a *weakly* coupled theory of (almost) fundamental scalars, despite the fact that the "fundamental" interactions are strongly self-coupled!

3 m_t in Models of Dynamical EWSB

In technicolor models, the masses of the ordinary fermions are due to their coupling to the technifermions, whose chiral-symmetry breaking is responsible for electroweak symmetry breaking. This is conventionally[17] assumed to be due to additional, broken, extended-technicolor (ETC) gauge-interactions:



(3.1)

which lead to a mass for the top-quark

$$m_t \approx \frac{g^2}{M_{ETC}^2}\langle \bar{U}U\rangle_{M_{ETC}} \ , \tag{3.2}$$

where we have been careful to note that it is the value of the technifermion condensate renormalized at the scale M_{ETC} which is relevant.

For a QCD-like technicolor, there is no substantial difference between $\langle \bar{U}U\rangle_{M_{ETC}}$ and $\langle \bar{U}U\rangle_{\Lambda_{TC}}$, and we can use naive dimensional analysis[18] to estimate the technifermion condensate, arriving at a top-quark mass

$$m_t \approx \frac{g^2}{M_{ETC}^2} 4\pi F^3 \ . \tag{3.3}$$

We can invert this relation to find the characteristic mass-scale of top-quark mass-generation

$$\frac{M_{ETC}}{g} \approx 1 \text{ TeV} \left(\frac{F}{246 \text{ GeV}}\right)^{\frac{3}{2}} \left(\frac{175 \text{ GeV}}{m_t}\right)^{\frac{1}{2}} . \tag{3.4}$$

We immediately see that the scale of top-quark mass generation is likely to be *quite* low, unless the value of the technifermion condensate ($\langle \bar{U}U\rangle_{M_{ETC}}$) can be raised significantly above the value predicted by naive dimensional analysis. The prospect of such a low ETC scale is both tantalizing and problematic. As we will see in the next section, constraints from the deviation of the weak interaction ρ parameter from one suggest that the scale may have to be larger than one TeV.

There have been two approaches to enhance the technifermion condensates which have been discussed in the literature: "walking"[19] and "strong-ETC".[15] In a walking theory, one arranges for the technicolor coupling constant to be approximately constant and large over some range of momenta. The maximum enhancement that one might expect in this scenario is

$$\langle \bar{U}U\rangle_{M_{ETC}} \approx \langle \bar{U}U\rangle_{\Lambda_{TC}} \left(\frac{M_{ETC}}{\Lambda_{TC}}\right)^{\gamma(\alpha_{TC})} \ , \tag{3.5}$$

where $\gamma(\alpha_{TC})$ is the anomalous dimension of the technifermion mass operator (which is possibly as large as one). As described above, however, we expect that M_{ETC} cannot be too much higher than Λ_{TC}, and therefore that the enhancement due to walking is not sufficient to reconcile the top-quark mass and an ETC scale higher than a TeV.

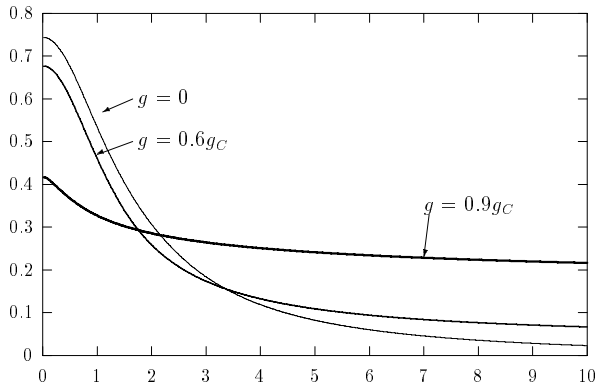


Figure 1: Plot[20] of technifermion self energy vs. momentum (both measured in TeV), as predicted by the gap-equation in the rainbow approximation, for various strengths of the ETC coupling relative to their critical value g_C.

The strong-ETC alternative is potentially more promising. As the size of the ETC coupling at the ETC scale approaches the critical value for chiral symmetry breaking, it is possible to enhance the running technifermion self-energy $\Sigma(k)$ at large momenta (see Fig. 1). Since the technifermion condensate is related to the trace of the fermion propagator.

$$\langle \bar{U}\, U \rangle_{M_{ETC}} \propto \int^{M_{ETC}^2} dk^2 \Sigma(k) \ , \tag{3.6}$$

a slowly-falling running-mass translates to an enhanced condensate.[†]

Unfortunately, there is no such thing as a free lunch. As we see from Fig. 2, the enhancement of the technifermion self-energy in strong-ETC theories comes at the cost of a "fine-tuning" of the strength of the ETC coupling relative to the critical value where the ETC interactions would, in and of themselves, generate chiral symmetry breaking. In the context of the NJL approximation, we find that enhancement of the top quark mass is directly related to the

[†]More physically, in terms of the relevant low-energy theory, it can be shown that the enhancement of the top-quark mass is due to the dynamical generation of a light scalar state.[16,21]

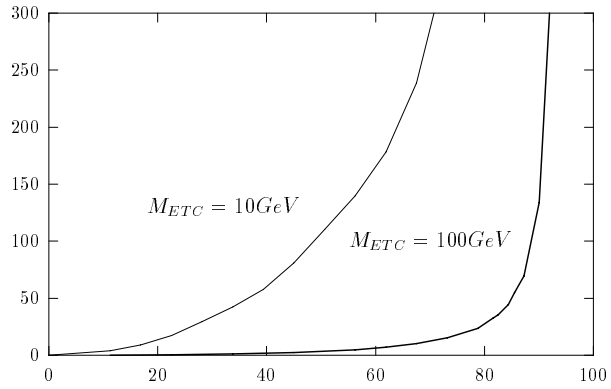


Figure 2: Plot[20] of top mass (in GeV) vs. ETC coupling (g/g_C in %), as predicted by gap-equation in the rainbow approximation, for ETC scales of 10 and 100 TeV.

severity of this adjustment. In particular, if we denote the critical value of the ETC coupling by g_C, in the NJL approximation, we find[9]

$$\frac{\langle \bar{U}U \rangle_{\Lambda_{TC}}}{\langle \bar{U}U \rangle_{M_{ETC}}} \approx \frac{\Delta g^2}{g_c^2} \tag{3.7}$$

where $\Delta g^2 \equiv g^2 - g_C^2$.

4 $\Delta\rho_*$

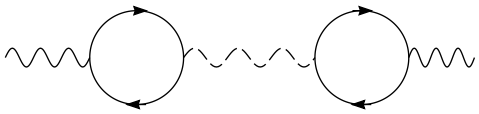
The physics which are responsible for top-quark mass generation must violate custodial $SU(2)$ since, after all, these physics must give rise to the disparate top- and bottom-quark masses. The danger is that this isospin violation will "leak" into the W and Z gauge-boson masses and give rise to a deviation of the weak interaction ρ-parameter from one.

4.1 Direct Contributions

As emphasized by Appelquist, Bowick, Cohler, and Hauser,[22] ETC operators which violate custodial isospin by two units ($\Delta I = 2$) are particularly dangerous. Denoting the right-handed technifermion doublet by Ψ_R, consider the operator

$$\frac{g^2}{M^2} \left(\bar{\Psi}_R \gamma_\mu \sigma_3 \Psi_R\right)^2 \quad , \tag{4.1}$$

which can result in the (mass-)mixing of the Z with an isosinglet ETC gauge-boson



(4.2)

and hence a contribution to $\Delta\rho$. Contributions of this sort arise naturally in ETC-models which give rise to the top-quark mass.[24]

If there are N_D doublets of the technifermions Ψ, and they give rise to a contribution to M_W^2 proportional to $N_D F^2$, the contribution of the operator in Eq. (4.1) to the ρ parameter can be estimated to be

$$\Delta\rho_* \approx \frac{2g^2}{M^2}\frac{N_D^2 F^4}{v^2} \tag{4.3}$$

$$\approx 12\%\ g^2 \left(\frac{N_D F^2}{(246\ \mathrm{GeV})^2}\right)^2 \left(\frac{1\ \mathrm{TeV}}{M}\right)^2 . \tag{4.4}$$

Current limits (see Fig. 3) on the parameter T ($\Delta\rho_* = \alpha T$) imply that $\Delta\rho_* \lesssim 0.4\%$.

There are two ways[‡‡] in which one may try to satisfy this constraint. The equation above implies

$$\frac{M}{g} \gtrsim 5.5\ \mathrm{TeV} \left(\frac{N_D F^2}{(246\ \mathrm{GeV})^2}\right) . \tag{4.5}$$

If $N_D F^2 \approx (246\ \mathrm{GeV})^2$, that is, if the sector giving rise to the top-quark mass is responsible for the bulk of EWSB, then the scale M must be much larger than the naive 1 TeV expectation in QCD-like technicolor. Comparing this with Eqs. (3.4) and (3.7) above, we see that the enhancement of the condensate needed requires a fine-tuning of order 3% ($\approx (1/5.5)^2$) in order to produce a top-quark mass of order 175 GeV.

Alternatively, we may re-write the bound as

$$F \lesssim \frac{105\ \mathrm{GeV}}{\sqrt{N_D}} \left(\frac{M/g}{1\ \mathrm{TeV}}\right)^{\frac{1}{2}} . \tag{4.6}$$

[‡‡]It is also conceivable[28] that there are additional isospin-asymmetric contributions—say, from relatively light pseudo-Goldstone bosons—which give rise to *negative* contributions to T and cancel some or all of the positive contributions discussed here.

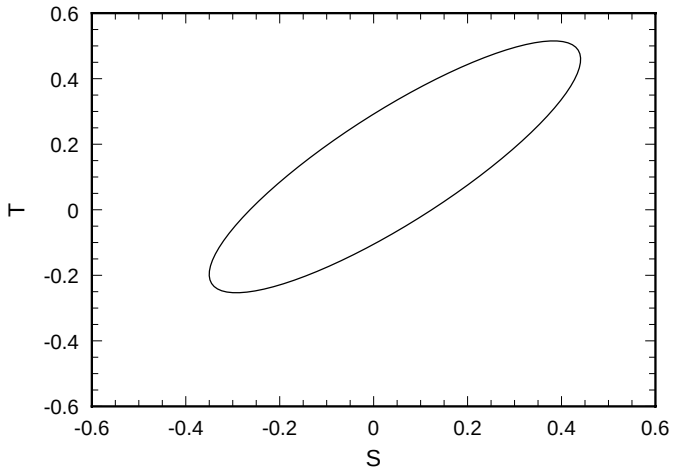


Figure 3: The ellipse[23] in the $S-T$ plane which projects onto the 95% confidence range for T. Note that $\Delta\rho_* = \alpha T$.

If M/g is of order 1 TeV, it is necessary that the sector responsible for top quark mass generation *not* give rise to the bulk of EWSB. While this case is counterintuitive (after all, the third generation is the heaviest!), it may in fact provide a resolution to the issue of how large isospin breaking can exist in the fermion mass spectrum without leaking into the W and Z masses. This is essentially what happens in multiscale models[25,26] and in top-color assisted technicolor.[27] Such hierarchies of technifermion masses are also useful for reducing the predicted value of S in technicolor models.[§][28]

4.2 Indirect Contributions

A second class of potentially dangerous contributions come from isospin violation in the technifermion mass spectra. In a manner analogous to the contribution[29] of the $t-b$ mass splitting to $\Delta\rho$, any difference in the dynamical masses of two technifermions in the *same* doublet will give rise to deviations in the ρ parameter from one. The size of this effect can be estimated a la Pagels-Stokar.[30] Using this approximation, we find that the contributions to the loop diagram from low-momenta dominate and

[§]Recently, the experimental upper bound on S has been relaxed, so that positive values of S are allowed ($S < 0.4$ at the 95% confidence level).[23]

$$\Delta\rho_* \propto \frac{N_D\, d}{16\pi^2}\left(\frac{\Sigma_U(0)-\Sigma_D(0)}{v}\right)^2 , \tag{4.7}$$

where N_D and d are the number of doublets and dimension of the technicolor representation respectively. Since we require $\Delta\rho_* \lesssim 0.4\%$, the equation above implies

$$N_D\, d\left(\frac{\Delta\Sigma(0)}{m_t}\right)^2 \lesssim 1.3 . \tag{4.8}$$

From this, we see that $\Delta\Sigma(0)$ must be less than of order m_t (perhaps, given the crude approximations involved, one may be able to live with $d = 2$ in the fundamental of and $SU(2)$ technicolor group with one doublet).

However, if the t and b get their mass from the same technidoublet, then at the ETC-scale, we expect that there is no difference between the t, b and the corresponding technifermions[15]

$$\begin{aligned} \Delta\Sigma(M_{ETC}) &\equiv \Sigma_U(M_{ETC}) - \Sigma_D(M_{ETC}) \approx \\ \Delta m(M_{ETC}) &\equiv m_t(M_{ETC}) - m_b(M_{ETC}) . \end{aligned} \tag{4.9}$$

Furthermore, if QCD is the only interaction which contributes to the scaling of the t and b masses, we expect $\Delta m(M_{ETC}) \approx m_t^{pole}$, and from scaling properties of the technifermion self-energies, we expect $\Delta\Sigma(0) \gtrsim \Delta\Sigma(M_{ETC})$.

There are two ways to avoid these constraints. One is that perhaps there are *additional* interactions which contribute to the scaling of the top- and bottom-masses below the ETC scale, and hence that $\Delta m(M_{ETC}) \ll m_t^{pole}$. This would be the case if the t and/or b get only a *portion* of their mass from the technicolor interactions, and would imply that the third generation must have (strong) interactions different from the technifermions (and possibly from the first and second generations). Another possibility is that the t and b get mass from *different* technidoublets, each of which have isospin-symmetric masses. The first alternative is the solution chosen in top-color assisted technicolor models (see below), while the latter has only recently begun to be explored.[31]

5 Case Study: Top-Color Assisted Technicolor

Recently, Hill has combined aspects of two different approaches to dynamical electroweak symmetry breaking into a model which he refers to as top-color assisted technicolor.[27] In this model, a top-condensate is driven by the combination of a strong, but spontaneously broken

and nonconfining, isospin-symmetric top-color interaction and an additional (either weak or strong) isospin-breaking $U(1)$ interaction which couple only to the third-generation quarks.

At low energies, the top-color and hypercharge interactions of the third-generation quarks may be approximated by four-fermion operators,[27]

$$\mathcal{L}_{4f} = -\frac{4\pi\kappa_{tc}}{M^2}\left[\overline{\psi}\gamma_\mu\frac{\lambda^a}{2}\psi\right]^2 - \frac{4\pi\kappa_1}{M^2}\left[\frac{1}{3}\overline{\psi_L}\gamma_\mu\psi_L + \frac{4}{3}\overline{t_R}\gamma_\mu t_R - \frac{2}{3}\overline{b_R}\gamma_\mu b_R\right]^2 \ , \tag{5.1}$$

where ψ represents the top-bottom doublet, κ_{tc} and κ_1 are related respectively to the top-color and $U(1)$ gauge-couplings squared, and where (for convenience) we have assumed that the top-color and $U(1)$ gauge-boson masses are comparable and of order M. The first term in Eq. (5.1) arises from the exchange of top-color gauge bosons, while the second term arises from the exchange of the new $U(1)$ hypercharge gauge boson which has couplings proportional to the ordinary hypercharge couplings. In order to produce a large top quark mass without giving rise to a correspondingly large bottom quark mass, the combination of the top-color and extra hypercharge interactions are assumed to be critical in the case of the top quark but not the bottom quark. The criticality condition for top quark condensation in this model is then:

$$\kappa^t_{eff} = \kappa_{tc} + \frac{1}{3}\kappa_1 > \kappa_c = \frac{3\pi}{8} > \kappa^b_{eff} = \kappa_{tc} - \frac{1}{6}\kappa_1 \ . \tag{5.2}$$

The contribution of the top-color sector to electroweak symmetry breaking can be quantified by the F-constant of this sector. In the NJL approximation,[9] for M of order 1 TeV and $m_t \approx 175$ GeV, we find

$$f_t^2 \equiv \frac{N_c}{8\pi^2}\, m_t^2 \log\left(\frac{M^2}{m_t^2}\right) \approx (64 \text{ GeV})^2 \ . \tag{5.3}$$

As f_t is small compared to 246 GeV, there must be additional dynamics which are largely responsible for giving rise to the W and Z masses. In top-color assisted technicolor, technicolor interactions play that role.

5.1 Direct Isospin Violation

Technifermions are necessary to produce the bulk of EWSB and to give mass to the light fermions. However, the heavy and light fermions must mix—hence, we would naturally expect that at least some of the *technifermions* carry the extra $U(1)$ interaction. If the additional $U(1)$ interactions violate custodial symmetry,¶ the $U(1)$ coupling will have to be quite small to keep

¶It has been noted[31] that if the top and bottom quarks receive their masses from *different* technidoublets, it is possible to assign the extra $U(1)$ quantum numbers in a custodially invariant fashion.

this contribution to $\Delta\rho_*$ small.[23] We will illustrate this in the one-family technicolor[32] model, assuming that techniquarks and technileptons carry $U(1)$-charges proportional to the hypercharge of the corresponding ordinary fermion.[‖] We can rewrite the effective $U(1)$ interaction of the technifermions as

$$\mathcal{L}_{4T1} = -\frac{4\pi\kappa_1}{M^2}\left[\frac{1}{3}\overline{\Psi}\gamma_\mu\Psi + \overline{\Psi}_R\gamma_\mu\sigma^3\Psi_R - \overline{L}\gamma_\mu L + \overline{L}_R\gamma_\mu\sigma^3 L_R\right]^2 \ , \tag{5.4}$$

where Ψ and L are the techniquark and technilepton doublets respectively.

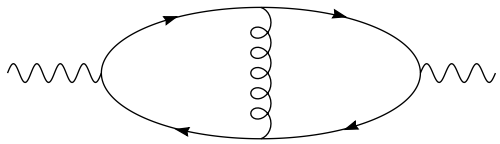
From the analysis given above (Eq. (4.4)), we see that the contribution to $\Delta\rho_*$ from degenerate technifermions is:[23]

$$\Delta\rho_*^{\rm T} \approx 152\% \ \kappa_1 \left(\frac{1 \ {\rm TeV}}{M}\right)^2 \ . \tag{5.5}$$

Therefore, if M is of order 1 TeV and the extra $U(1)$ has isospin-violating couplings to technifermions, κ_1 must be extremely small.

5.2 Indirect Isospin Violation

In principle, since the isospin-splitting of the top and bottom are driven by the combination of top-color and the extra $U(1)$, the technifermions can be degenerate. In this case, the only indirect contribution to the ρ parameter at one loop is the usual contribution coming from loops of top and bottom quarks.[29] However, since there are additional interactions felt by the third-generation of quarks, there are "two-loop" contributions of the form



. (5.6)

This contribution yields[23]

$$\Delta\rho_*^{\rm tc} \approx 0.53\% \left(\frac{\kappa_{tc}}{\kappa_c}\right)\left(\frac{1 \ {\rm TeV}}{M}\right)^2\left(\frac{f_t}{64 \ {\rm GeV}}\right)^4 \ . \tag{5.7}$$

Combining this with Eq. (5.3), we find that

$$M \ \gtrsim \ 1.4 \ {\rm TeV} \tag{5.8}$$

[‖]Note that this choice is anomaly-free.

for $\kappa_{tc} \approx \kappa_c$. This immediately puts a constraint on the mass of the top-color gluon which is comparable to the direct limits currently obtained by CDF.[33]

5.3 Fine Tuning

Finally, we must require that the sum of the effects of Eqs. (5.5) and (5.7) do not give rise to an experimentally disallowed contribution to the ρ parameter. Equation (5.5) implies that κ_1 must either be very small, or M very large. However, we must also simultaneously satisfy the constraint of Eq. (5.2), which implies that

$$\frac{\Delta\kappa_{tc}}{\kappa_c} = \left|\frac{\kappa_{tc} - \kappa_c}{\kappa_c}\right| \leq \frac{1}{3}\frac{\kappa_1}{\kappa_c} \ . \tag{5.9}$$

Therefore, if M is low and κ_1 is small, the top-color coupling must be tuned close to the critical value for chiral-symmetry breaking. On the other hand, if κ_1 is not small and M is relatively large, the *total* coupling of the top quark must be tuned close to the critical NJL value for chiral-symmetry breaking in order to keep the top-quark mass low. The gap-equation for the NJL model implies that

$$\frac{\Delta\kappa_{eff}}{\kappa_c} = \frac{\kappa^t_{eff} - \kappa_c}{\kappa_c} = \frac{\frac{m_t^2}{M^2}\log\frac{M^2}{m_t^2}}{1 - \frac{m_t^2}{M^2}\log\frac{M^2}{m_t^2}} \ . \tag{5.10}$$

These two constraints are shown in Fig. 4. For $M > 1.4$ TeV, we find that either $\Delta\kappa_{tc}/\kappa_c$ or $\Delta\kappa_{eff}/\kappa_c$ must be tuned to less than 1%. This trade-off in fine tunings is displayed in Fig. 4. For the "best" case where both tunings are of order 1%, $M = 4.5$ TeV.

6 Conclusions

We have seen that a large top quark mass has a number of important implications for dynamical electroweak symmetry breaking:

- A large top-quark mass naturally implies, in models of dynamical electroweak symmetry breaking, the possibility of a correspondingly low scale for the scale of top-flavor physics. While I have emphasized the constraints on such physics arising from potential contributions to the weak interaction ρ parameter, there are also significant constraints arising from the size of the $Z \to b\bar{b}$ branching ratio,[34,24] as well as from contributions to $b \to s\gamma$ and $B - \bar{B}$ mixing.[35–37]

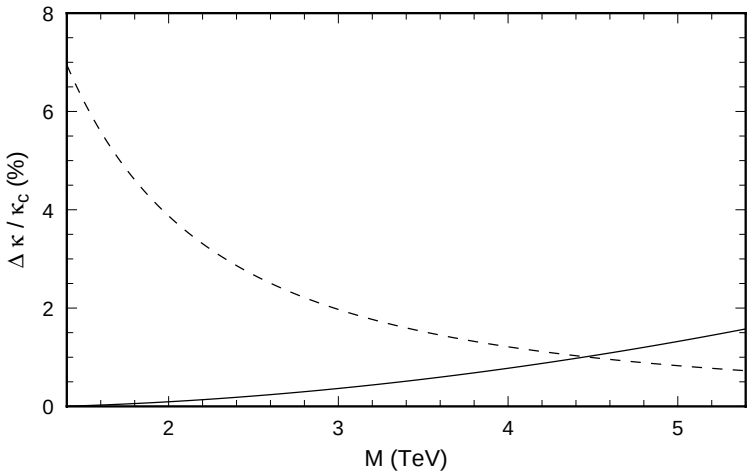


Figure 4: The amount of fine tuning required[23] in the TC^2 model. The dashed line is the amount of fine tuning in $\Delta\kappa_{eff}$ required to keep m_t much lighter than M, see Eq. (5.10). The solid curve shows the amount of fine tuning (see Eq. (5.9)) in $\Delta\kappa_{tc}$ required to satisfy the bound $\Delta\rho_* < 0.4\%$. The region excluded by the experimental constraint on $\Delta\rho_*$ is above the solid curve.

- The physics responsible for the large isospin breaking in the $t-b$ mass splitting can lead to potentially dangerous "direct" and "indirect" effects in the W and Z masses.
- The direct and indirect effects can be mitigated if the sector which is responsible for the top and bottom masses does *not* provide the bulk of electroweak symmetry breaking and, conversely, if the sector responsible for the W and Z masses gives rise to only a *small portion* of the top and bottom masses. This can happen only if the top and bottom feel *strong* interactions which are not shared by the technifermions, and possibly, the first two generations.
- In top-color assisted technicolor, the extra top-color interactions give rise to additional indirect contributions to $\Delta\rho$, and we must require that $M_g \gtrsim 1.4$ TeV. Furthermore, if the extra $U(1)$ has isospin-violating couplings to technifermions, we require fine tuning of order 1%.

Acknowledgments

I thank Tom Appelquist, Nick Evans, and Ken Lane for helpful conversations; Mike Dugan for help in preparing the manuscript; and John Terning and Bogdan Dobrescu for collaboration[23] on some of the work reported in this talk. I also acknowledge the support of an NSF Presidential Young Investigator Award and a DOE Outstanding Junior Investigator Award. *This work was supported in part by the National Science Foundation under grant PHY-9057173, and by the Department of Energy under grant DE-FG02-91ER40676.*

References

[1] R. S. Chivukula, B. A. Dobrescu, and J. Terning, "Isospin breaking and the top quark mass in models of dynamical electroweak symmetry breaking," Boston University preprint BUHEP-95-22 and hep-ph/9506450. Talk presented by R. S. Chivukula at the *Workshop on Top Quark Physics*, Iowa State University, Ames, May 25-26, 1995, and the *Yukawa International Seminar*, Kyoto, Aug. 21-25, 1995.

[2] G. 't Hooft, in *Recent Developments in Gauge Theories*, edited by G. 't Hooft *et. al.*, eds. (Plenum Press, New York, NY, 1980).

[3] K. G. Wilson, Phys. Rev. B **4**, 3184 (1971); K. G. Wilson and J. Kogut, Phys. Rep. **12**, 76 (1974).

[4] M. Lüscher and P. Weisz, Nucl. Phys. B**318**, 705 (1989); J. Kuti, L. Lin, and Y. Shen, Phys. Rev. Lett. **61**, 678 (1988); A. Hasenfratz *et. al.*, Phys. Lett. B**199**, 531 (1987); A. Hasenfratz *et. al.*, Nucl. Phys. B**317**, 81 (1989); G. Bhanot *et. al.*, Nucl. Phys. B**353**, 551 (1991) and B**375**, 503 (1992) E.

[5] U. M. Heller, H. Neuberger, and P. Vranas, Nucl. Phys. B**399**, 271 (1993); K. Jansen, J. Kuti, and C. Liu, Phys. Lett. B**309**, 119 (1993).

[6] L. Maiani, G. Parisi, and R. Petronzio, Nucl. Phys. B**136**, 115 (1978).

[7] S. Weinberg, Phys. Rev. D**19**, 1277 (1979); L. Susskind, Phys. Rev D**20**, 2619 (1979); E. Farhi and L. Susskind, Phys. Rep. **74**, 277 (1981).

[8] M. Weinstein, Phys. Rev **8**, 2511 (1973).

[9] Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1961).

[10] Y. Nambu, Enrico Fermi Institute Preprint EFI 88-39; V. A. Miransky, M. Tanabashi, and K. Yamawaki, *Phys. Lett.* **B221**, 177 (1989) and Mod. Phys. Lett. A**4**, 1043 (1989).

[11] W. A. Bardeen, C. T. Hill, and M. Lindner, *Phys. Rev.* **D41**, 1647(1990).

[12] C. T. Hill, M. Luty, and E. A. Paschos, *Phys. Rev.* **D43**, 3011 (1991); T. Elliot and S. F. King, *Phys. Lett.* **B283**, 371 (1992).

[13] C. T. Hill *et. al.*, *Phys. Rev.* **D47**, 2940 (1993).

[14] C. T. Hill, *Phys. Lett.* **B266**, 419 (1991); S. Martin, *Phys. Rev.* **D45**, 4283 (1992) and **D46**, 2197 (1992); N. Evans, S. King, and D. Ross, Z. Phys. C**60**, 509 (1993).

[15] T. Appelquist, M. Einhorn, T. Takeuchi, and L. C. R. Wijewardhana, Phys. Lett. B**220**, 223 (1989); V. A. Miransky and K. Yamawaki, Mod. Phys. Lett. A**4**, 129 (1989); K. Matumoto, Prog. Theor. Phys. Lett. **81**, 277 (1989) .

[16] R. S. Chivukula, A. Cohen, and K. Lane, *Nucl. Phys.* **B343**, 554 (1990).

[17] S. Dimopoulos and L. Susskind, *Nucl. Phys.* **B155** (1979) 237; E. Eichten and K. Lane, *Phys. Lett.* **B90** 125 (1980).

[18] A. Manohar and H. Georgi, *Nucl. Phys.* **B234** 189 (1984).

[19] B. Holdom, Phys. Rev. D**24**, 1441 (1981); B. Holdom, Phys. Lett. B**150**, 301 (1985); K. Yamawaki, M. Bando, and K. Matumoto, Phys. Rev. Lett. **56**, 1335 (1986); T. Appelquist, D. Karabali, and L. C. R. Wijewardhana, Phys. Rev. Lett. **57**, 957 (1986); T. Appelquist and L. C. R. Wijewardhana, Phys. Rev. D**35**, 774 (1987); T. Appelquist and L. C. R. Wijewardhana, Phys. Rev. D**36**, 568 (1987).

[20] N. Evans, *Phys. Lett.* **B331** 378 (1994).

[21] T. Appelquist, J. Terning, and L. Wijewardhana, Phys. Rev. **44**, 871 (1991).

[22] T. Appelquist *et al.,* Phys. Rev. D**31** 1676 (1985).

[23] R. S. Chivukula, B. A. Dobrescu, and J. Terning, hep-ph/9503203, Phys. Lett. B**353**, 289 (1995).

[24] See, for example, G.-H. Wu, hep-ph/9412206 Phys. Rev. Lett. **74**, 4137 (1995).

[25] K. Lane and E. Eichten, Phys. Lett. B**222**, 274 (1989).

[26] J. Terning, Phys. Lett. B**344**, 279 (1995).

[27] C. T. Hill, *Phys. Lett.* **B345**, 483 (1995).

[28] T. Appelquist and J. Terning, Phys. Lett. B**315**, 139 (1993).

[29] M. Einhorn, D. Jones, and M. Veltman, Nucl. Phys. B**191**, 146 (1981).

[30] H. Pagels and S. Stokar, Phys. Rev. D**20**, 2947 (1979); B. Holdom, *Phys. Lett.* **B226** (1989) 137.

[31] K. Lane and E. Eichten, hep-ph/9503433, Phys. Lett. B**352**, 382 (1995).

[32] E. Farhi and L. Susskind, Phys. Rev. D**20**, 3404 (1979).

[33] F. Abe *et. al.*, CDF collaboration, FERMILAB-PUB-94/405-E.

[34] R. S. Chivukula, S. B. Selipsky, and E. H. Simmons, Phys. Rev. Lett. **69**, 575 (1992); R. S. Chivukula, E. Gates, J. Terning, and E. H. Simmons *Phys. Lett.* **B311**, 157 (1993); R. S. Chivukula, J. Terning, and E. H. Simmons, *Phys. Lett.* **B331**, 383 (1984) and hep-ph/9506427.

[35] L. Randall and R. Sundrum, *Phys. Lett.* **B312**, 148 (1993).

[36] B. Balaji, "Technipion contribution to $b \to s\gamma$,"Boston University preprint BUHEP-95-18, hep-ph/9505313.

[37] D. Kominis, "Flavor changing neutral current constraints on topcolor assisted technicolor," Boston University preprint BUHEP-95-20, hep-ph/9506305.

PHOTONS AND POMERONS IN PHOTOPRODUCTION AT HERA

Usha Mallik
The University of Iowa
Representing the H1 and ZEUS collaborations

ABSTRACT

The physics of transition between nonperturbative and perturbative QCD at low x has been explored at the ep collider HERA. Some recent results from the H1 and ZEUS experiments are discussed.

1 Introduction

Deep Inelastic Scattering (DIS) has been one of the most resourceful ways to examine the internal structure of hadrons. Leptons, structureless compared to hadrons, are used as probes. The experiments are difficult and time consuming, as seen from the experiences of SLAC, EMC, BCDMS, E665, CDHS, CCFR,[1] etc. All these are fixed target experiments with electron,[2] muon,[3] and neutrino beams,[4] performed in the kinematic ranges shown in Fig. 1(a). HERA, the first *ep* collider, covers a very large unexplored kinematic domain. It can extend to the fixed target region by varying either the lepton or the proton beam energy in order to establish continuity with the fixed target experimental results. Since electrons interact with charged particles predominantly through photon exchange, HERA also provides an opportunity to study the characteristics of photons from quasireal to virtual.

The low x region ($10^{-4} \sim 10^{-5}$) at moderate Q^2 is a newly explored kinematic regime where pertubative calculations of Quantum ChromoDynamics (QCD) are expected to work. This is also a region where scaling violation is large and the sea density is high where the transition from nonperturbative to perturbative QCD (pQCD) is expected to take place. New phenomena observed in this transition region might be calculated with the help of pQCD. In other words, theoretical calculations can be tested under new and novel experimental situations in this region. At HERA, we operationally describe processes up to a virtuality of $Q^2 \leq 4\text{GeV}^2$ as photoproduction; the reason is explained later. Since HERA covers the kinematic region where photon propagators range from quasireal ($Q^2 \simeq 0$) to virtual (DIS), the changing nature of the photon can be critically studied. Photoproduction results have historically been explained by a combination of the Vector Dominance Model (VDM)[7] and Regge phenomenology,[9] i.e., nonpertubative theory. The DIS region, on the other hand, has been described by pQCD. A challenge for HERA is to study and understand the change or the transition region between the nonperturbative and the perturbative domain. In many cases, the same interaction is studied with variation of parameters and the results then compared with pQCD predictions. The hope is to understand which of these, if any, parameters sets the appropriate scale for QCD. We denote the nonperturbative domain as the "old" domain and the transition region as the "new" domain. This "new" or the transition domain is critically studied both in photoproduction, i.e., interactions with $Q^2 \leq 4$ GeV^2, and in DIS interactions as well. In this talk, I describe the

HERA photoproduction results; the next speaker will discuss the DIS results. We

Figure 1: (a) A comparison of the kinematic reach of HERA with the fixed target DIS experiments. (b) The luminosity delivered by HERA to date.

start with the photoproduction reactions which were understood with VDM and Regge phenomenology[8,9]; these are measured at HERA and compared with the nonperturbative and perturbative models. These measurements are repeated as each parameter (e.g., Q^2, p_t) is varied and the results tested with various predictions. Both light and heavy quark vector mesons productions are discussed. Comparisons between photon-induced reactions at HERA and hadron-induced reactions from hadron colliders are made, when photoproduction events of different virtuality are compared with the hadron collider results; the difference and the similarities between the two are noted. Evidence of partonic structure and the structure function of the photon are described. Finally, the structure and the contents of Pomerons are discussed.

1.1 HERA, H1, and ZEUS

The HERA collider is located in Hamburg, Germany, at the DESY laboratory. The center-of-momentum (c.m.) energy of HERA is 300 GeV, with the ability

to run either polarized electrons or positrons to collide with the proton beam. Figure 1(b) shows the luminosity delivered by HERA to date. While the increase is impressive, the goal of 170 pb^{-1} a year, necessary to critically test the Standard Model, remains for the future with the HERA upgrade. The parameters are noted in Table 1. The two colliding beam experiments, H1 and ZEUS, occupy the North and the South intersection regions (IR), respectively. The East IR is being used by HERMES, a polarization experiment which uses the polarized lepton beam impinging on a polarized gas jet target. HERMES has started data taking in 1995. Longitudinal polarizations of higher than 60% have routinely been achieved with the help of spin rotators (which convert the naturally acquired transverse polarization of the beam to longitudinal polarization) at the East IR. The West IR is scheduled for HERA B, a high luminosity CP violation experiment using an internal target in the proton beam. They expect to collect data starting from 1997.

Figure 2 shows the H1 and the ZEUS detectors; the lepton beam approaches the IR from the left and exits through the beam pipe to the right. The outgoing proton/lepton direction is noted as the Forward/Rear direction and the respective detector components are noted with the prefix F/R (ZEUS). Detailed descriptions of H1[5] and ZEUS[6] are given elsewhere.

Table 1: HERA parameters.

Parameters	electron Design/Achieved	proton Design/Achieved
nominal energy(GeV)	30/27.5	820
energy range (GeV)	10 - 33	300 - 820
beam current (mA)	58/30	163/50
c.m. energy (GeV)	314/300	
circumference (m)	6336	
no. of coll. bunches	200	
bunch crossing time(ns)	96	

The two detectors are similar, both are multipurpose and magnetic, but complementary to each other in specific details. Both detectors display a forward-backward asymmetry to accommodate the forward motion of the c.m. in the lab frame. Both detectors use a right-handed coordinate system with the origin (0,0,0) as the Interaction Point (IP). The Z-axis is positive in the forward direction where the polar angle θ is zero. The negative Z direction corresponds to a θ value of 180°.

Figure 2: Axial view of the (a) H1 and (b) ZEUS detector.

1.2 Kinematics

The kinematics are illustrated in Fig. 3 and are described below. Here:

k/k' = four-momentum of the incoming/outgoing e;
p/p' = four-momentum of the incoming/outgoing p;
$s = (p+k)^2$ = square of c.m. energy;
$q = k - k'$ = four-momentum of the virtual photon;
$Q^2 = -q^2$;
$y = \frac{2p\cdot q}{2p\cdot k}$;
$x = \frac{Q^2}{2p\cdot q}$;
$W^2_{\gamma p} = (q+p)^2$; $W_{\gamma p}$ = photon-proton cm energy;
$t = (p-p')^2$;
$\eta = -ln(tan\theta/2)$ = pseudorapidity.

Figure 3: Diagram of ep collision where a neutral vector meson is produced.

The ZEUS detector covers the pseudorapidity region from $\eta \simeq -3.8, \theta \simeq 176^\circ$ in the rear direction to $\eta = 4.3, \theta \simeq 2.5^\circ$ in the forward direction. H1 coverage is similar. (The photon fragmentation region, in the rear direction, generally is characterized by $\sim -3.8 \leq \eta \leq -0.75$; similarly, the proton fragmentation region in the forward direction and the central region are characterized, typically, by $\sim 1.1 \leq \eta \leq 4.3$, and $-1.0 \leq \eta \leq 1.0$, respectively.) The scaling variables x and y are related by $Q^2 = xys \simeq 4xyE_eE_p$, where E_e and E_p are the energies of the incoming e and p, respectively.

2 Photoproduction

As mentioned before, we operationally discuss processes as photoproduction up to a virtuality of $Q^2 \leq 4\text{GeV}^2$, beyond which the scattered electron gets detected in the main calorimeter. We define the photon wavefunction to consist of two parts,

i.e., $\gamma = \gamma_{direct} + \gamma_{resolved}$, where the direct part represents the pointlike bare photon; photon structure containing partons or virtual mesons as in the VDM is included in the resolved component. Figure 4 shows diagrams in which the photon interacts as a point particle (direct) and in which the photon interacts as a composite, i.e., a parton from the photon interacts with a parton from the proton and the remnants of both the photon and the proton (resolved) are observed. In the direct photon picture, the entire photon partakes in the interaction; so the momentum fraction of the photon measured from the detected particles from the interaction, $x_\gamma \simeq 1$. In the resolved interaction, part of the photon momentum is carried away by its remnant; hence, the momentum carried by the partons in the central collision provide $x_\gamma < 1$. The photon remnant typically is observed in the rear part of the detector.

Figure 4: Leading-order diagrams of direct and resolved photon interactions.

One hopes that a systematic comparison of experimental results under varying parameters with theoretical calculations will lead to the understanding of the scale to allow pQCD calculations. Various kinematic signatures in an interaction can set such a scale. An interaction where high mass vector mesons are produced using quasireal photons (e.g., photoproduction of J/ψ), or light quark vector meson production at high virtuality or at high p_t^2, where p_t is the transverse momentum of the detected system (e.g., jets in hard photoproduction) can be used to test the validity of the pQCD models. We explore each of these interactions starting with the "old" topics, i.e., the validity of the well-known VDM models in the framework of Regge theory in photoproduction.

2.1 Total γp Cross Section

Early experiments showed that quasireal photons behave much like hadrons: the total photoproduction cross section, $\sigma(\gamma p)$, was shown to be related to the total $\pi^- p$ and $\pi^+ p$ cross section through VDM-inspired relations.[8] The equivalent photon flux, obtained from the Weizacker-Williams formula,[10] decreases logarithmically as the virtuality increases; hence the largest contribution to the photoproduction cross section is made by almost real photons. Prior to HERA, the only $\sigma(\gamma p)$ measurements were available for c.m. energy $W_{\gamma p} \leq 20$ GeV, with laboratory photon beam energies up to 100 GeV.[11] The ZEUS and H1 measurements, shown in Fig. 4(a), achieve $150 \leq W_{\gamma p} \leq 200$ GeV, corresponding to a laboratory photon beam energy above 5 TeV in a fixed target experiment. Earlier data were parameterized using the Regge model prescription of the total cross section[12] given by $\sigma(\gamma p) = Xs^{\epsilon} + Ys^{-\eta}$, where X, Y, ϵ, and η are constants determined from the data with $\epsilon = 0.0808$ and $\eta = 0.4525$. The first term accounted for the slow rise of the cross section at high energies by incorporating the Regge idea of Pomeron exchange, and the second term denoted the contribution from all other meson exchanges. A Pomeron, in the Regge model, is a hypothetical particle with the quantum numbers of vacuum. The so-called Regge trajectories are represented by $\alpha(t) = \alpha(0) + \alpha' t$; for a Pomeron, $\alpha(0) = 1 + \epsilon = 1.0808$ denotes the intercept on the t-axis, and $\alpha' = 0.25\text{GeV}^{-2}$ denotes the slope of the trajectory. A fit to recent high energy total cross section measurements in $p\bar{p}$[13] prefers $\alpha(0) \simeq 1.11$. This value is also incorporated into the prediction shown in Fig. 5 along with the data at low and high energies*, and some of the parameterizations.[14,15] In the low Q^2 domain, where the nonperturbative approach seems adequate, the Pomeron is thought of as "soft." Hence the "soft Pomeron" approach seems to still suffice in explaining the total photoproduction cross section at even the very high HERA energies.

2.2 Vector Meson Production: Light Quark Sector

The VDM models originated before any of the photoproduction measurements of ρ, ω, and ϕ. According to VDM, the photon fluctuates into a vector meson (ρ, ω,

*In the results, the first uncertainty quoted is statistical, and the second is systematic. In the figure, the error bars indicate the uncertainties. The inner error bars indicate statistical errors; the outer error bars show the systematic uncertainty (excluding the correlated uncertainty) added in quadrature.

and ϕ); in elastic photoproduction, the vector meson materializes by exchange of the Pomeron between the γ and the proton. The soft Pomeron is characterized by a weak dependence of the cross section on $\gamma^* p$[†] c.m. energy, $\sigma_{(\gamma p \to Vp)} \simeq W_{\gamma p}^{0.22}$, similar to that of the total photoproduction cross section described above. In analogy to the soft Pomeron in "old" photoproduction, exclusive (or elastic) vector meson production calculated in pQCD requires exchange of an object with the quantum numbers of the Pomeron. This is alluded to as a "hard" Pomeron.

At the H1 and ZEUS experiments, the ρ, ω, and ϕ signals are observed in $\pi^+\pi^-$, $\pi^+\pi^-\pi^0$, and K^+K^- decay modes, respectively.[19,18,20] Figure 6(a) shows a compilation of the elastic ρ photoproduction cross sections with measurements from both H1 and ZEUS. The ρ^0 polarization measured from its decay angular distribution was observed to be transverse. This is in accordance with the expected polarization of the quasireal photons producing the ρ's, assuming s channel helicity conservation (SCHC).

The $W_{\gamma p}$ dependence of the cross sections is shown for all available exclusive vector meson final states in Fig. 7(a). The data for ρ, ω, and ϕ agree well with the soft Pomeron picture. The measured cross section ratios are in good agreement with the VDM predictions within the quoted uncertainties.[‡]

The t dependence of the ρ^0 cross sections was fit to the form $d\sigma/dt = A \cdot exp(-b|t'|)$, with $t' = t - t_{min}$. A compilation of the measurements of the slope parameter b is shown in Fig. 6(b). For the 1994 ZEUS data, the scattered p was detected in the Leading Proton Spectrometer (LPS),[21] and thus t' was directly measured in the range $0.07 \leq |t'| \leq 0.4\text{GeV}^2$. For ω and ϕ, the differential cross sections were measured for the $|t'|$ ranges of $0 \leq |t'| \leq 0.6$ and $0.1 \leq |t'| \leq 0.5$, respectively. The results are summarized in Table 2.

[†] γ^* refers to an off-shell photon.

[‡] The ϕ total cross section was determined from extrapolating the differential cross section $d\sigma/dt$ to the forward direction $t = t_{min}$ and then integrating the spectrum.

Figure 5: Total photoproduction cross sections. The solid line is the prediction of the DL[14] combined fit of the hadron-hadron and low-energy photoproduction data, and the dashed line is the ALLM[15] parameterization. The dotted line presents the DL parameterization obtained after the recent measurement[13] from CDF.

Table 2: Elastic photoproduction of light quark vector mesons.

Reaction	H1	ZEUS
$\gamma p \to \rho^0 p$		
$< W_{\gamma p} >$ (GeV)	55	70
$\sigma(\gamma p \to \rho^o p)$ (μb)	$9.1 \pm 0.9 \pm 2.5$	$14.7 \pm 0.4 \pm 2.4$
b (GeV^{-2})	$10.9 \pm 2.4 \pm 1.1$	$9.9 \pm 1.2 \pm 1.4$
r_{00}^{04}	$-0.11 \pm 0.11 \pm 0.04$	$0.055 \pm 0.028(stat.)$
$\gamma p \to \phi p$		
$< W_{\gamma p} >$ (GeV)		70
$\sigma(\gamma p \to \phi p)$ (μb)		$0.96 \pm 0.19^{+0.21}_{-0.18}$
b (GeV^{-2})		$7.3 \pm 1.0 \pm 0.8$
r_{00}^{04}		$0.03 \pm 0.05(stat.)$
$\gamma p \to \omega p$		
$< W_{\gamma p} >$ (GeV)		80
$\sigma(\gamma p \to \omega p)$ (μb)		$1.2 \pm 0.1 \pm 0.3$
b (GeV^{-2})		$9.3 \pm 1.7 \pm 1.3$
r_{00}^{04}		$0.07 \pm 0.07(stat.)$

Figure 6: (a) Compilation of elastic ρ photoproduction cross sections. The solid line is the soft Pomeron parameterization by Schuler and Sjöstrand[16]. (b) Compilation of b from $d\sigma/dt = A \cdot exp(-b|t'|)$ measurements.

2.3 Heavy Quark Sector: Charm Production

Elastic J/ψ Production: Elastic photoproduction of J/ψ has been observed[22] with fixed target experiments, but the dependence of the cross section on $W_{\gamma p}$ was not established because of the limited phase space available at the fixed target energies relative to the J/ψ mass ($M^2_{J/\psi}$). HERA experiments measured the elastic photoproduction cross section[25,26] in the range 40 GeV$\leq W_{\gamma p}$ 180 GeV, clearly establishing the energy dependence shown in Fig. 7(a). The J/ψ was detected in both leptonic decay modes e^+e^- and $\mu^+\mu^-$.

Perturbative calculations[23,24] of this diffractive process were thought to be valid in J/ψ production because of the high charm quark mass. Accordingly, the cross section should depend on the square of the gluon momentum density, and the contributing gluon x range in the proton ($x \sim 1/W^2_{\gamma p}$,[23] i.e., typically $10^{-3} \sim 10^{-4}$). The lower half of Fig. 7(a) shows the predictions using three popular parameterizations of the gluon momentum density as input to Ryskin's QCD model.[23] The observed fast rise of this cross section with $W_{\gamma p}$ is thought to arise from the increase in gluon momentum density as x decreases, and is not in agreement with the soft Pomeron prediction.[17]

The helicity of the J/ψ's is determined from the J/ψ decay angular distribution.The $cos\theta$ distribution[§] from e^+e^- and $\mu^+\mu^-$ decay modes of the J/ψ in the

[§]The $cos\theta$ distribution is fit to $[1 + r^{04}_{00} + (1 - 3r^{04}_{00})cos^2\theta)]$ according to a prescription,[38] where $r^{04}_{00} = 0$ means transverse and 1 means longitudinal polarizations, respectively.

helicity frame, displayed in Fig. 7(b), shows the J/ψ's polarizations as primarily transverse.

Inelastic J/ψ and Charmed Meson Production: The photon-gluon fusion diagram, an inclusive process, (Fig. 4) is thought to be the primary source of inelastic charm production, as opposed to the elastic or exclusive production of J/ψ which is diffractive. Background processes are expected from the resolved photon process as well as the double diffractive process (with a Pomeron exchange) where the proton diffractively dissociates.

(a) (b)

Figure 7: (a) Cross sections of elastic photoproduction of vector mesons. The dotted lines are for the vector meson photoproduction are the DL predictions[17] for a soft Pomeron exchange. The J/ψ cross section is described by Ryskin's pQCD model[23] using the structure functions shown. (b) Decay angular distribution of the J/ψ's in the helicity frame.

(a) (b)

Figure 8: (a) Inelastic photoproduction of J/ψ; the curves represent NLO calculations with proton structure functions as marked. (b) Inelastic photoproduction of D^*'s; the curves represent NLO calculations with various proton and photon structure functions.

Several calculations of J/ψ production[27] and charmed meson production[28] through the photon-gluon fusion process in next-to-leading-order (NLO) of pQCD are available. However, theoretical uncertainties abound. These arise from the parameterizations of the photon and the proton structure functions, the scale of α_s, the value of the charm mass, etc.

The ZEUS and H1 experiments both observed J/ψ signals in the leptonic decay modes.[26] The D^*, D_s, and Λ_c signals were also observed by ZEUS and D^* by H1; the D^* signals were observed from $D^{*+} \to D^0 + \pi^+$ and the charged conjugate decays.[29] Qualitative agreement between experimental data and theoretical predictions were observed as shown in Figs. 8(a) and (b). The typical x of the proton probed in these reactions is 10^{-3} and the proton structure functions¶ use a gluon momentum density that increases at low x. The rise in the J/ψ cross section is in agreement with the rise in the gluon density[51,52] seen in measurements by ZEUS and H1, and will be described by the next speaker.

2.4 Exclusive Vector Meson Production at High Virtuality

Exclusive production of ρ^0, ϕ and J/ψ in ep DIS interactions have been observed at HERA, with the scattered electron detected in the calorimeter ($Q^2 \geq 4$ GeV^2). Predictions using pQCD[23,24,30,31] are available for these reactions. The 1993 ZEUS results[32] and the 1994 H1 results[33] for ρ production are displayed in Fig. 9 (a) and (b) (with the elastic photoproduction cross section also shown for comparison). Here the electroproduction cross section has been converted into a virtual photon proton ($\gamma^* p$) cross section. The NMC values at lower energies compared to the HERA values indicate a rapid rise of cross section with $W_{\gamma p}$ compared to that in the photoproduction cross section. Presumably, this illustrates a difference between the soft and the hard Pomeron exchange. The results have been compared with a theoretical model[30] which applies pQCD calculation based on a similar gluon ladder exchange as that of Ryskin.[23] The model uses longitudinally polarized photons only.[30] The study of the angular distribution[38] of the ρ in the ZEUS data yields $\sim 60\%$ longitudinal polarization at a Q^2 of ~ 9 and 17 GeV^2.

¶Most recent parameterizations of the proton structure function used here are MRSG (MRS: Martin-Robert-Sterling), MRSA,[48] CTEQ,[49] and GRV[50] (GRV: Glück-Reya-Vogt). They all obey the Gribov-Dokshitzer-Lipatov-Altarelli-Parisi (GDLAP) evolution equations.[46] The GRV prescription starts with initial hard parton distribution at 0.3 GeV^{-2} and propagate it according to GDLAP evolution in NLO. In others, the gluon momentum density rises as $x^{-\lambda}$ at low x.

Assuming SCHC, 40% of the photon polarization is still transverse. For the H1 data, the longitudinal fraction is observed to be 70%. The data points in the Q^2 and the $W_{\gamma p}$ ranges for the H1 and the ZEUS experiments are noted in Table 3. The Q^2 dependence of the cross section, described by Q^{2n} is shown in Fig. 9(b); the n value determined from the data is also listed in Table 3. (For soft Pomeron exchange the cross section falls as Q^4, the pQCD model[30] predicts a faster fall off.)

The DIS ϕ and J/ψ productions, observed by both H1[33] and ZEUS,[34] are also listed in Table 3. The longitudinal polarization of the ϕ's observed by ZEUS and H1 (prelim) are, respectively, $0.6 \pm 0.1^{+0.2}_{-0.1}$ and $0.77 \pm 0.11 \pm 0.02$, at the Q^2 and $W_{\gamma p}$ shown in Table 3.

Table 3: Exclusive vector meson production in DIS.

Reaction	H1		ZEUS	
$\gamma^* p \to \rho^0 p$				
$\langle Q^2 \rangle (\mathrm{GeV}^2)$	$\simeq 10$	$\simeq 20$	$\simeq 9$	$\simeq 17$
$W_{\gamma p}$ (GeV)	$\simeq 65 \quad \simeq 115$	$\simeq 65 \quad \simeq 115$	$\simeq 62 \quad \simeq 100$	$\simeq 62 \quad \simeq 100$
(Q^{2n}) : n value	$2.5 \pm 0.5 \pm 0.2$		$2.1 \pm 0.4^{+0.7}_{-0.3}$	
$\gamma^* p \to \phi p$	preliminary			
$\langle Q^2 \rangle (\mathrm{GeV}^2)$	8.3	14.8	8.2	14.7
$\langle W_{\gamma p} \rangle (\mathrm{GeV})$	95	100	94	99
$\sigma(\gamma^* p \to \phi p)(\mathrm{nb})$	~ 10	~ 3	10.2 ± 2.2	3.1 ± 0.7
(Q^{2n}) : n value	2.0 ± 0.6		2.1 ± 0.6	
$\gamma^* p \to J/\psi p$			prelim	
$\langle Q^2 \rangle (\mathrm{GeV}^2)$	10	20	7	
$\langle W_{\gamma p} \rangle (\mathrm{GeV})$	88	88	25	
$\sigma(\gamma^* p \to J/\psi p)(nb)$	17 ± 6.3	6.6 ± 1.6	$8.0^{+5.2}_{-2.0}$	
(Q^{2n}) : n value	2.0 ± 0.6		2.1 ± 0.6	

(a) (b)

Figure 9: (a) Cross section of exclusive ρ production in $\gamma^* p \to \rho p$ vs. $W_{\gamma p}$ (at $Q^2 \sim 10$ and 20 GeV^2). The rise in cross section with $W_{\gamma p}$ observed in DIS is faster than that at $Q^2 \simeq 0$. (b) Cross section for exclusive ρ production in $\gamma^* p \to \rho p$ as a function of Q^2.

2.5 Photoproduction at High Transverse Momentum

Vector mesons, e.g., ρ^0 and ϕ productions with $p_t^2 > 1.0$ GeV^2, have been observed; detailed studies have begun. Preliminary indications are that the diffractive dissociation process dominates.

Figure 10: The solid dots are corrected ZEUS data. The shaded band shows the correlated uncertainty from measurement of energy in the calorimeter and the integrated luminosity. ($\bar{\eta}$ is the average η of the dijets). In (a), the data are compared to LO QCD calculations using several parton distribution sets for the proton and the GS2[37] set for the photon. In (b), the data are compared to the LO QCD calculation of (a) using the GRV LO parton distribution set for the proton, the same calculation but omitting the resolved contribution with $x_\gamma \geq 0.75$, including only gluon-induced direct photon processes and the "k_T factorized" curve. In (c), the data are compared to HERWIG Monte Carlo estimates of the cross section using partons and final state jets. For these HERWIG histograms, we have used the GRV LO (LAC1) proton (photon) parton distribution set. In (d), the data are compared to HERWIG jet cross sections using the GRV (LAC1) and MRSD_0 (LAC1) proton (photon) parton distribution sets.

Dijet Production in Direct and Resolved Photon Interactions: In hard photoproduction, i.e., in events with high p_t, jet measurements are compared with pQCD predictions in an attempt to understand whether the measurements agree with the predictions, and if so, in which kinematic region is the agreement observed. Events with two jets with a minimum p_t of 7 GeV and with reasonable p_t balance are selected. The details of the jet algorithm, event selection, and Monte Carlo are described elsewhere.[35,36] The direct and the resolved photoproduction events are distinguished by the values of x_γ obtained from the observed jets. The pQCD calculations are made in leading order (LO)[35]; the measurements and the theoretical predictions are shown in Fig. 10. Direct photon events are selected by $x_\gamma \geq 0.75$; the agreement between the data and the theoretical models is at best within $\sim 20\%$ for direct photon events (with the choice of the "best" photon structure function). The resolved photon events agree with the predictions only in shape. Studies are continuing in order to better understand the various parameters, scales, cut-offs, jet algorithms, and the use of Monte Carlos.

2.6 Photons as Hadrons

Comparison of Photon and Hadron Induced Reactions: Quasireal photons, known to behave like hadrons[8] in interactions, have been described in VDM models. The H1 experiment extended the comparison of various kinematic quantities and event shapes between photons of different virtualities in the photon induced reactions with similar quantities in hadronic reactions.[39] In the photon fragmentation region, here $-3 \leq \eta \leq -2$, the photons clearly show a rise in p_t^2 per unit rapidity as a function of Q^2. This is shown in Fig. 11(a). However, the virtuality seems to have little effect in the central region[‖], here $-0.5 \leq \eta \leq 0.5$ [Fig. 11(b)]. Figure 12 shows a comparison of the H1 results in the central region with those from the hadron colliders ordered according to the c.m. energies. It is clear that the hadron and the photon induced reactions look alike in the central region independent of Q^2. Not only does the transverse energy distribution not seem to depend on Q^2, but also the event distributions according to the maximum event rapidity, η_{max}, (i.e., the event shapes) between the hadron and the photon induced reactions look alike, independent of the c.m. energy, as observed from the distributions shown in Figs. 13(a) and (b). Hence photon induced reactions and hadron induced reactions behave very alike in some aspects, independent of

[‖]The MC, however, does not describe the data very well.

the photon virtuality.

Resolved Component of the Photon: Partonic structure of the photon is expected to be suppressed as the virtuality increases.[40] The ZEUS experiment observed the resolved behavior of the photon[41] at a Q^2 known by tagging the scattered electrons,** in the Q^2 regions of $10^{-6} < Q^2 < 0.02$ GeV^2, and $0.3 < Q^2 < 0.6$ GeV^2. The x_γ distributions of these events are shown in Figs.14(a) and (b). The low x_γ, indicative of resolved photons, is clearly observed at Q^2 as high as $0.3 - 0.5$ GeV^2.

Photon Structure, Gluon Content: The quark structure of the photon has been measured in $e\gamma$ interactions by the e^+e^- experiments[42] in the fractional momentum range $0.007 \leq x \leq 1$. Photoproduction of jets at high energies offers a new tool for determination of the gluon density in the photon in γp interactions at HERA. (Recent high-energy e^+e^- experiments are also sensitive to the jets from $\gamma\gamma^*$ interactions[43]). The H1 experiment used hard photoproduction dijet events, with each jet having a minimum transverse momentum, $p_t > 7$ GeV, to define the event sample to extract the photon structure function.[44] Figure 15(a) shows the x_γ distribution of these events as measured from the jets. The direct photon events are seen at $x_\gamma \sim 1$, where the photon behaves as a point-like object. Using the quark-density photon structure function measured in e^+e^- experiments as input, a fraction of events shown by the solid curve in Fig. 15(a) can be explained. The rest of the events in the spectrum, at lower x_γ likely indicate the presence of gluons in the photon. The GRV structure function†† of the photon explains the events well, as shown in Fig. 15(b). Input use of several other structure functions, LAC1 and LAC3,[47] are also shown in the same figure; they do not describe the data so well. The probability of the observed event distribution to be compatible with no gluonic component in the photon is 0.1%. The gluonic component at low x_γ appears to rise faster than the quark content of the photon.

2.7 The Transition Region: The Soft and the Hard Pomeron

The F_2 measurements from DIS can be turned into a $\gamma^* p$ total cross section by the relation, $\sigma(\gamma^* p) = \frac{4\pi^2\alpha}{Q^2} F_2(W^2_{\gamma p}, Q^2)$, similar to the way the electroproduction cross sections in Fig. 9 were turned into the $\gamma^* p$ cross section. Figure 16 is a compilation

**The electron calorimeter of the luminosity monitor and the Beam Pipe Calorimeter (BPC) detects scattered electrons with $10^{-6} < Q^2 < 0.02$ GeV^2 and $0.3 < Q^2 < 0.6$ GeV^2, respectively.

††The Glu̇ck-Reya-Vogt (GRV) structure function of the photon in LO[45] starts with an input parton density at $Q^2 = 0.03$ GeV^2 and then is evolved according to the GDLAP prescription.[46]

of the $W_{\gamma p}$ dependence for all the $\gamma^* p$ cross sections obtained from photoproduction and DIS. At low $Q^2 (\simeq 0)$, the cross section decreases with an increase in $W_{\gamma p}$ to a minimum after which there is a slow rise with $W_{\gamma p}$. This was explained earlier as being consistent with soft Pomeron (and meson) exchange. As Q^2 increases, the $\sigma(\gamma^* p)$ shows a slow rise followed by a fast rise with increasing $W_{\gamma p}$. The faster rise, thought to be a general consequence of QCD, was described earlier as a signature of hard Pomeron exchange. This trend is observed throughout Fig. 16, with the transition from soft to hard Pomeron exchange taking place at increasingly lower $W_{\gamma p}$ as Q^2 increses. It is from this region that we hope to gain an understanding of the transition phenomena from nonperturbative to perturbative QCD. In photoproduction, this region is approached starting with quasireal photons and then going towards the transition region. In DIS, the approach is to go as low as possible in Q^2 and x. This will be discussed by the next speaker in detail. At the lowest Q^2 value of 2 GeV^2, the $\sigma(\gamma^* p)$ in Fig. 16 still shows a fast rise.

3 The Rapidity Gap Events

The Neutral Current (NC) DIS events are characterized by the detected scattered electron, the current jet caused by the hadronization of the struck quark, and the debris of the remains of the proton (called the proton remnant jet) around the forward beam pipe. As the struck quark leaves the proton, color strings are stretched between the struck quark and the remaining part of the proton: these eventually break up and form hadrons, which in turn deposit energies in the calorimeter between the current jet and the proton remnant. Hence, in a typical NC DIS event, the observed particle with the highest pseudorapidity (η_{max}) and a minimum of (e.g.) 400 MeV energy deposit in the calorimeter has η_{max} above 3 or 4 as shown in Fig.17 (a). In the 1992 NC events sample, a category of events ($\sim 10\%$) were observed with $\eta_{max} \leq 2.5$, i.e., as if the proton after the reaction went down the forward beam pipe unchanged. A special class of these events would be exclusive vector meson production in DIS. In QCD terminology, a color singlet exchange (e.g., gluon ladder) between the lepton and the proton could explain these events; in Regge terminology, these events would represent a hard Pomeron exchange.

First observed by ZEUS, both experiments have now seen these events with Large Rapidity Gaps[53] (LRG), characteristic of a Pomeron exchange. Figure 17(b)

shows a diagram of this mechanism for such LRG events. Kinematics of these events are very like the NC DIS events; this will be discussed in detail by the next speaker.

3.1 Structure Function of the Pomeron

The kinematics of a Pomeron exchange are shown in Fig. 18 and are explained below.

$t = P - P' \simeq 0$
$x_{\mathcal{P}} = \frac{(P-P')\cdot q}{P\cdot q} \simeq \frac{M_X^2+Q^2}{W_{\gamma p}^2+Q^2}$ = momentum fraction of the $\mathcal{P}$ in prton
$\beta = \frac{x}{x_{\mathcal{P}}} \simeq \frac{Q^2}{M_X^2+Q^2} \simeq x_{q/\mathcal{P}}$ = momentum fraction of the struck quark in the $\mathcal{P}$.

A structure function for the Pomeron $\mathcal{P}$ from the LRG events in the DIS data sample can be defined from:

$$\frac{d^3\sigma}{dx_{\mathcal{P}}dxdQ^2} = \frac{4\pi\alpha^2}{Q^4x}(1-y-y^2/2)F_2^{D(3)}(Q^2,\beta,x_{\mathcal{P}}).$$

Since the diffracted proton is not observed, $F_2^{D(3)}$ has been integrated over t. If factorization at the Pomeron-proton vertex can be assumed, then,

$$F_2^{D(3)}(Q^2,\beta,x_{\mathcal{P}}) = f_{\mathcal{P}/\mathbf{p}}(x_{\mathcal{P}})\cdot F_2^{D(2)}(Q^2,\beta),$$

where $f_{\mathcal{P}/\mathbf{p}}(x_{\mathcal{P}})$ denotes the Pomeron flux. The rapidity gap DIS event samples were divided into bins of Q^2 and β by both H1 and ZEUS experiments. For each data sample, the event distribution as a function of $x_{\mathcal{P}}$ could be fit to a straight line with a slope independent of the Q^2 and the β value. Figure 19 shows the ZEUS[55] and the H1 distributions[54]; a fit to the form $b(1/x_{\mathcal{P}})^a$ (with b varying from bin to bin) yielded $a = 1.19 \pm 0.06 \pm 0.07$ for H1 data and $a = 1.30 \pm 0.08^{+0.08}_{0.14}$ from the ZEUS data. Hence, the data do not violate the factorization assumption. A second more recent analysis from ZEUS is also shown in Fig. 19. The event selection and analysis criteria were entirely different and will not be described here.[56] The results were consistent with the first analysis; subtracting the nondiffractive background in a way similar to the first analysis[55] gave $a = 1.28 \pm 0.04$, while using a new method gave $a = 1.42 \pm 0.08$. The uncertainties quoted in the second method are statistical only.

Since the data were consistent with factorization in the available kinematic region, we obtain $F_2^{D(2)}$ by integrating over the observed kinematic $x_{\mathcal{P}}$ range. The

$F_2^{D(3)}$ distributions vs. β and Q^2 for H1 and ZEUS data and the $F_2^{D(3)}$ vs. β for ZEUS data are shown in Fig. 20. A soft momentum distribution of the Pomeron constituents like $(1-\beta)^n$, where n is a positive integer, does not fit the data shown; a harder distribution like $\beta(1-\beta)$ is preferred. Both are superimposed on the ZEUS data. The data seem to prefer a combination of hard with a slight mixture of soft distribution.

Contents of the Pomeron: In DIS, the highly virtual photon couples to quarks in lowest order. But, in hard photoproduction when jets are observed, both quarks and gluons in the proton can couple to the virtual photon in the lowest nonvanishing order. The LRG events have also been observed in photoproduction events.[57] So, by comparing the rapidity gap events of DIS and hard photoproduction, ZEUS has estimated the relative gluon content of the Pomeron inside the proton. Events with jets having $E_t \geq 8$ GeV in the central region $(-1 \leq \eta \leq 1)$ were selected. Similar to the DIS LRG events, the distribution of the photoproduction LRG events preferred a hard partonic component in the $\mathcal{P}$. Assuming all of the Pomeron momenta are carried by its partonic contents, i.e., using the momentum sum rule $\int d\beta \Sigma_i \beta f_{\mathcal{P}/\mathbf{p}}(\beta) = 1$, the event distribution in terms of η_{max} of the events favors hard gluon content. The nondiffractive contribution was modeled by the MC and a double-diffractive contribution of $\sim 15\%$[55] where the proton diffractively dissociated, was subtracted.

As the momentum sum rule need not apply to the constituents of the $\mathcal{P}$, this was not imposed to the data for further analysis. The total momentum carried by the Pomeron was then measured as a function of C_g, the relative ratio of gluonic to quark content of the Pomeron without the momentum sum rule restriction. This was performed for both the photoproduction and the DIS jet sample by the ZEUS collaboration.[58] Figure 21 shows the result with 1σ uncertainty for both. The two curves intersect around $C_g = 0.6$, i.e., both sets of data are satisfied with a relative gluonic component of 0.6 and a quark content of 0.4 in the Pomeron without the need for the momentum sum rule. The assumption made is that the same $\mathcal{P}$ is observed in both photoproduction and DIS; the momentum sum rule has not been applied. It should be noted that the total momentum in the DIS sample rises very steeply as C_g approaches 1, reflective of the fact that the lowest order coupling of the virtual photon in DIS to the proton is to its valence quarks.

4 Conclusion and Future Plans

The ep colliding beam experiments, H1 and ZEUS, have made remarkable progress towards understanding the transition betweeen the regions where nonperturbative and perturbative QCD apply. For quasireal photons, low p_t, and light quark vector mesons, the elastic photoproduction follows the VDM-Regge nonperturbative prediction. However, as we get to heavy quarks, e.g., photoproduction of J/ψ, or to high virtuality, e.g., exclusive production of light quark vector mesons in DIS, or photoproduction at high transverse momenta, some pQCD calculations begin to be applicable. In any case, the nonperturbative predictions do not apply as observed from the steeper rise of cross section of photoproduction of J/ψ with increasing $W_{\gamma p}$ than predicted by DL[17] calculations. The DIS exclusive production of J/ψ does not hold any surprise since the pQCD scale has already been set. The diffractive scattering or Pomeron exchange at an appropriate scale can then be described in terms of pQCD. The changing nature of Pomeron is being studied extensively. The study of the structure function of Pomeron in DIS has started; many new results are forthcoming in this area. A large amount of literature is available on the diffractive interactions.[9,12,59–63] We have also studied the Pomeron contents and observed substantial presence of gluonic partons. The structure of photons have revealed new information that at low x, the gluonic content of the photon increases. And, the presence of partonic structure in the photon can be observed at a Q^2 as high as $0.3 - -0.6$ GeV^2.

Upgrades to cover more phase space with better resolution have been added in 1995. The H1 experiment has recently added a silicon vertex detector and a silicon rear detector. The ZEUS group completed all six stations of LPS which helps detect the intact outgoing proton. The Beam Pipe Calorimeter has been replaced with a W-scintillator sampling calorimeter with a larger coverage of $0.1 \leq Q^2 \leq 0.8$ GeV^2. In addition, a scattered electron tagger 44 m away from the IP in the rear direction at $80 < W_{\gamma p} < 100$ GeV, a proton remnant tagger around the forward direction, and a presampler to the F and R calorimeter (to understand the energy scale) have also been installed. The early look at the 1995 data indicates the technical as well as physics success of these upgrades. The 1994 data at HERA provided a unique kinematic domain to understand QCD, and to study photon, Pomeron, and proton structures. With the 1995 upgrades, both H1 and ZEUS detectors are well on their way to provide the best understanding of these topics. (Many interesting topics, like strange quark studies and different approaches for

understanding color singlet exchange in both H1 and ZEUS experiments, are not covered due to lack of time and space.)

Acknowledgments

It is a pleasure to acknowledge the effort of my ZEUS colleagues and the organizers of the SLAC Summer Institute.

References

[1] M. Riordan, *The Hunting of Quark* (Simon & Schuster, New York); R. Taylor, in *Proceedings of the 19th SLAC Summer Institute on Particle Physics*, edited by J. Hawthorne, SLAC-REP-398, 1991 and references therein.

[2] J. T. Friedman and H. W. Kendall, Annu. Rev. Nucl. Sci. **22**, 203 (1972).

[3] For a review, see T. Sloan, G. Smadja, and V. Ross, Phys. Rep. **130**, 45 (1988); E665 Collab., M. R. Adams *et al.*, Phys. Rev. D **48**, 5057 (1993).

[4] For a review, see S. R. Mishra and F. Sciulli, Annu. Rev. Nucl. Part. Sci. **39**, 259 (1989).

[5] H1 Collab., T. Ahmed *et al.*, The H1 Detector at HERA, DESY 93-103.

[6] ZEUS Collab., The ZEUS Detector, Status Report 1993, DESY 1993; ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **293**, 465 (1992).

[7] J. J. Sakurai, Phys. Rev. Lett. **22**, 981 (1969).

[8] For a review, see D. W. G. S. Leith, High Energy Photoproduction: Diffractive Processes, SLAC-Pub-1878, *Electromagnetic Interaction of Hadrons*, edited by A. Donnachie and G. Shaw (Plenum Publ. Co.).

[9] K. Goulianos, Phys. Rep. **101**, 169 (1983).

[10] C. F .von Weizsa̋cker, Z. Phys. **88**, 612 (1934); E. J. Williams, Phys. Rev. **45**, 729 (1934).

[11] D. O. Caldwell *et al.*, Phys. Rev. Lett. **40**, 1222 (1978).

[12] A. Donnachie and P. V. Landshoff, Nucl. Phys. B **244**, 322 (1984); A. Donnachie and P. V. Landshoff, Phys. Lett. B **185**, 403 (1987); A. Donnachie and P. V. Landshoff, Nucl. Phys. B **311**, 509 (1989); P. V. Landshoff, Nucl. Phys. B (Proc. Suppl.) **18C**, 211 (1990).

[13] CDF Collab., Phys. Rev. D **50**, 5535 (1994).

[14] A. Donnachie and P. V. Landshoff, Phys. Lett. B **296**, 227 (1992).

[15] H. Abramovich, E. M. Levin, A. Levy, and U. Maor, Phys. Lett. B **269**, 465 (1991).

[16] G. A. Schuler and T. Sjöstrand, Phys. Lett. B **300**, 169 (1993); G. A. Schuler and T. Sjöstrand, Nucl. Phys. B **407**, 539 (1993).

[17] A. Donnachie and P. V. Landshoff, Phys. Lett. B **348**, 213 (1995).

[18] ZEUS Collab., M. Derrick *et al.*, Z. Phys. C **69**, 39 (1995).

[19] H1 Collab., S. Aid *et al.*, DESY-95-251, to appear in Nucl. Phys. B.

[20] ZEUS Collab., M. Derrick *et al.*, DESY-96-067, to appear in Phys. Lett. B.

[21] ZEUS Collab., The ZEUS Detector, Status Report 1993.

[22] E401 Collab., M. Binkley *et al.*, Phys. Rev. Lett. **48**, 73 (1982); E516 Collab., B. H. Denby *et al.*, Phys. Rev. Lett. **52**, 795 (1984).

[23] M. Ryskin, Z. Phys. C **57**, 89 (1993).

[24] J. Nemchik, N. N. Nikolaev, and B. G. Zakharov, Phys. Lett. B **341**, 228 (1994).

[25] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **350**, 120 (1995).

[26] H1 Collab., S. Aid *et al.*, DESY-96-037, to appear in Nucl. Phys. B.

[27] H. Jung *et al.*, Z. Phys C **60**, 721 (1993); M. Kraṁer *et al.*, Phys. Lett. B **348**, 677 (1995); M. Kraṁer, Nucl. Phys. B **459**, 3 (1996).

[28] R. K. Ellis and P. Nason, Nucl. Phys. B **312**, 551 (1989); P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. B **303**, 607 (1988); J. Smith and W. L. van Neerven, Nucl. Phys. B **374**, 36 (1992).

[29] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **349**, 225 (1995).

[30] S. J. Brodsky *et al.*, Phys. Rev. D **50**, 3134 (1994).

[31] I. F. Ginzburg *et al.*, TPI-MINN-94-14-T, submitted to Nucl. Phys. B.

[32] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **356**, 601 (1995).

[33] H1 Collab., S. Aid *et al.*, DESY-96-023, to appear in Nucl. Phys. B.

[34] ZEUS Collab., M. Derrick *et al.*, DESY-96-067, to appear in Phys. Lett. B.

[35] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **348**, 665 (1995).

[36] ZEUS Collab., M. Derrick *et al.*, DESY-96-094, to appear in Phys. Lett. B.

[37] L. E. Gordon and J. K. Storrow, Z. Phys. C **56**, 307 (1992).

[38] K. Schilling *et al.*, Nucl. Phys. B **15**, 397 (1970); K. Schilling and G. Wolf, Nucl. Phys. B **61**, 381 (1973); P. Joos *et al.*, Nucl. Phys. B **113**, 53 (1976).

[39] H1 Collab., S. Aid *et al.*, Phys. Lett. B **358**, 412 (1995).

[40] M. Drees and R. Godbole, hep-ph/9508221.

[41] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **354**, 163 (1995).

[42] C. Berger and W. Wagner, Phys. Rep. **146**, 1 (1987) and references therein.

[43] AMY Collab., B. J. Kim *et al.*, Phys. Lett. B **325**, 248 (1994); TOPAZ Collab., H. Hayashi *et al.*, Phys. Lett. B **314**, 149 (1993); ALEPH Collab., D. Buskulic *et al.*, Phys. Lett. B **313**, 509 (1993); DELPHI Collab., P. Abreu *et al.*, Phys. Lett. B**342**, 320 (1995).

[44] H1 Collab., T. Ahmed *et al.*, Nucl. Phys. B **445**, 195 (1995).

[45] M. Glück, E. Reya, and A. Vogt, Phys. Rev. D **46**, 1973 (1992).

[46] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972) and 675; G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977); Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1971).

[47] H. Abramowicz, K. Charchula, and A. Levy, Phys. Lett. B **269**, 458 (1991).

[48] A. D. Martin and W. J. Stirling, Phys. Lett. B **354**, 155 (1995).

[49] Various parameterizations by CTEQ Collab., see e.g., J. Botts *et al.*, Phys. Lett. B **304**, 159 (1993), etc.

[50] M. Glück, E. Reya, and A. Vogt, Z. Phys. C **67**, 433 (1995).

[51] H1 Collab., I. Abt *et al.*, Phys. Lett. B **321**, 161 (1994); M. Roco, Ph.D. Thesis, Univ. of Iowa, DESY Int. Rep. F35D-94-04.

[52] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **345**, 576 (1995).

[53] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **315**, 481 (1993); ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **332**, 228 (1994); H1 Collab., T. Ahmed *et al.*, Nucl. Phys. B **429**, 477 (1994).

[54] H1 Collab., T. Ahmed *et al.*, Phys. Lett. B **348**, 681 (1995).

[55] ZEUS Collab., M. Derrick *et al.*, Z. Phys. C **68**, 569 (1995).

[56] ZEUS Collab., M. Derrick *et al.*, DESY-96-018, to appear in Z. Phys. C.

[57] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **346**, 399 (1995); H1 Collab., T. Ahmed *et al.*, Nucl. Phys. B **435**, 3 (1995).

[58] ZEUS Collab., M. Derrick *et al.*, Phys. Lett. B **356**, 129 (1995).

[59] P. D. B. Collins, *An Introduction to Regge Theory and High Energy Physics* (Cambridge University Press, Cambridge, 1977).

[60] G. Ingelman and P. Schlein, Phys. Lett. B **152**, 256 (1985).

[61] L. N. Lipatov, Sov. J. Nucl. Phys. **23**, 338 (1976); E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977); Y. Y. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[62] N. N. Nikolaev and B. G. Zakharov, Z. Phys. C **53**, 331 (1992).

[63] A. Capella *et al.*, Phys. Lett. B **343**, 403 (1995).

Figure 11: (a) Normalized transverse energy E_T flow in the hadronic c.m. vs. η^* at various values of Q^2. (b) E_T distribution per unit of η^* in the central region ($-0.5 < \eta^* < 0.5$) and in the photon fragmentation region ($-3 < \eta^* < -2$). Two different MC models are shown; the lower (upper) curves correspond to the central (photon fragmentation) region.

Figure 12: Normalized transverse energy E_T in the central region as a function of the hadronic c.m. energy. The HERA data at various Q^2 are compared with data from hadron-hadron collisions ($p\bar{p}$: UA1, pp: NA22 and AFS). Systematic erros are not shown.

Figure 13: (a) Uncorrected E_T distribution in the central region ($-0.5 < \eta^* < 0.5$) per unit of η^*. Open circles: photoproduction data. Full circles: DIS data with $10 < Q^2 < 100$ GeV^2 for $0.3 < y < 0.5$. (b) η^*_{max} distribution per unit of η^* normalized to the total number of events.

Figure 14: x_γ distribution from events with the scattered electron tagged in the (a) beam pipe calorimeter and in the (b) electron calorimeter of the luminosity monitor.

(a) (b)

Figure 15: (a) x_γ distribution of the two-jet events. Only statistical errors are displayed. The full line shows the contribution from the quark-resolved photon process; the dashed line shows the direct photon contribution as obtained from PYTHIA. (b) The gluon density of the photon divided by the fine structure constant ($\alpha = 1/137$) at $< p_t >^2 = 75$ GeV^2. For comparison, the GRV-LO (full), the LAC1 (dashed), and the LAC3 (dotted) parameterizations of the photon's gluon density are shown.

Figure 16: Compilation of total cross sections as a function of $W_{\gamma p}$ at various Q^2 from photoproduction and DIS.

(a) **(b)**

Figure 17: (a) Schematics of ordinary DIS process, without any rapidity gap. (b) Schematics of events with a large rapidity gap.

Figure 18: Kinematics of an event with a Pomeron exchange.

Figure 19: $F_2^{D(3)}$ vs. $x_{\mathcal{P}}$ for various β values are shown from the H1 data and two different analyses with two selection criteria from ZEUS data.

Figure 20: $F_2^{D(3)}$ vs. β obtained from the ZEUS analyses.

Figure 21: The total momentum of the Pomeron as a function of the ratio of its gluon partonic content and quark partonic content from the LRG events from the DIS event sample and the hard photoproduction event sample. The intersection of the two samples yield a ratio of $\sim 2:1$ gluon to quark partonic content of the Pomeron without the need for any momentum sum rule for the partons from the Pomeron.

NEWS FROM THE PROTON—RECENT DIS RESULTS FROM HERA

Karlheinz Meier
Heidelberg University, Heidelberg, Germany

Representing the H1 and ZEUS Collaborations at HERA

ABSTRACT

Recent results from the two large, general-purpose detectors H1 and ZEUS at HERA (DESY, Hamburg, Germany) are presented. Emphasis is given to the analysis of deep inelastic scattering defined by the observation of the scattered electron or positron in the main calorimeters. Results on purely inclusive cross sections lead to a determination of the charged (quarks) parton distribution $F_2(x, Q^2)$. Access to the electrically neutral parton content (gluons) is obtained indirectly by an analysis of the expected scaling violation behavior of F_2 or directly from multijet rates originating from well-defined initial parton configurations. Finally, the recently uncovered subclass of large rapidity gap (LRG) events has been analyzed in terms of F_2. The result supports the concept of a color neutral object (Pomeron $I\!P$) being probed by a hard scattering electron. Evidence for factorization of the Pomeron radiation process as well as for scaling in the inclusive $I\!P$ structure function has been found.

©1995 by Karlheinz Meier.

1 HERA and HERA Experiments

The electron-proton collider HERA at DESY provides a choice of electron or positron beams with momenta of 26.7 GeV colliding with high momentum protons of 820 GeV. The machine features multibunch operation with a bunch crossing repetition frequency of 100 ns. Since its initial operation in 1992, integrated luminosities of 55 nb^{-1} in 1992, 880 nb^{-1} in 1993, and 5500 nb^{-1} in 1994 have been delivered by the HERA machine to the experiments. The 1994 luminosity period saw the switchover from electrons to positrons. Expectations for the 1995 run period are around 10 pb^{-1}. Figure 1 shows the layout of the HERA accelerator complex.

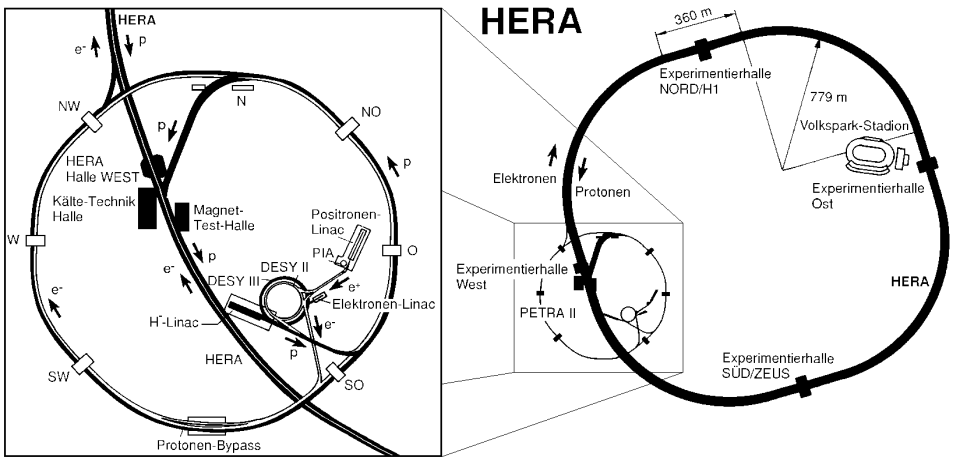


Figure 1: View of the HERA electron-proton collider complex. The left figure displays the accelerator system on the DESY site. The right figure provides an overview of the entire HERA ring with the four interaction regions available for experiments. The two large general-purpose detectors H1 and ZEUS are marked.

The HERA collider serves two large general-purpose detectors, H1 and ZEUS. Both experiments are classical solenoidal spectrometers with electromagnetic and hadronic calorimetry for electron and jet measurements. Reflecting the asymmetry of beam energies, both detectors have far better track recognition and thicker calorimeters in the proton (forward) direction. A major difference between the two detectors is the concept of calorimetry. H1 has choosen a LAr calorimeter inside

a huge solenoidal coil in order to optimize the measurement of electromagnetic showers without degradation by dead material. ZEUS has built a high-resolution uranium scintillator calorimeter with photomultiplier readout providing an excellent measurement of final state hadrons even at low energies. Detailed descriptions of both detectors can be found in their respective technical papers.[1,2]

HERA's physics goals in deep inelastic scattering are twofold. The main frontier is given by the availability of very large momentum transfers Q^2. Theoretically, values up to 90,000 GeV^2 can be reached. In practice, the current range in this area is statistically limited to approximately 10,000 GeV^2 because of the characteristic $1/Q^4$ suppression in the Rutherford scattering formula. The other area unexplored in the pre-HERA time is deep inelastic scattering with very low values of the scaling variable x representing the longitudinal momentum fraction of the struck parton. The specific kinematics of HERA allows measurements at x values as low as 10^{-4} with still sizable momentum transfers of $\mathrm{O}(GeV)$.

2 Inclusive Cross Sections in DIS

Scattering experiments have been tools for the exploration of the structure of nuclear and subnuclear matter for almost one century. The size of accessible substructures is essentially given by the inverse momentum of the exchanged vector boson representing a characteristic wavelength. HERA kinematics provides a resolution power of approximately $10^{-3}\ fm$. Using the four-momenta k of the incident electron q of the virtual photon and P of the incident proton, the following kinematical variables can be defined:

$$Q^2 = -q^2 \qquad x = \frac{Q^2}{2P \cdot q} \qquad y = \frac{P \cdot q}{P \cdot k} \qquad W^2 = (P+q)^2. \quad (1)$$

Q^2 is the squared virtuality of the photon or the squared momentum transfer. x is the Bjorken scaling variable which can be interpreted in the naïve quark parton model (QPM) as the fraction of the struck parton momentum relative to the proton. In the proton rest frame, y is the normalized energy transfer from the electron to the proton. W is the invariant mass of the virtual photon-proton system. It corresponds to the invariant mass of the total final state excluding the scattered electron.

The term "deep inelastic scattering" (DIS) is an experimental definition based on the observability of the scattered lepton (electron or positron in the case of HERA neutral current events) in the main detectors. Current calorimeters are able to catch electrons corresponding to momentum transfers Q as low as 1 GeV. Without detected leptons, the events are usually called "Photoproduction Events" referring to the fact that at very low momentum transfers, the mass of the virtual photon is close to zero. Such photons are therefore almost real. For an overview of photoproduction results from HERA, see Ref. 3.

Both experiments, H1 and ZEUS, have analyzed the new regime of very large momentum transfers and compared the results to expectations from the Standard Model (see Refs. 4–6). Both experiments have analyzed neutral current (NC) and charged current (CC) data. NC events are balanced in the observed vector sum of all transverse momenta (p_T) in the detector. The high p_T lepton is mainly balanced by the current jet emerging from the struck parton. CC events with an undetected high p_T neutrino in the final state exhibit a huge transverse momentum imbalance. These are clear experimental signatures which make the detection and separation of NC and CC events a relatively easy task. Neutral currents are mediated by t-channel exchange of virtual photons and Z bosons. Those two diagrams give rise to identical final states so that there is an additional interference term contributing to the cross section. Weak and electroweak contributions become important for Q^2 values of approximately $10,000$ GeV^2. Below 1000 GeV^2, those contributions are negligible. Charged currents have a distinct sensitivity to the composition of the proton target. Left-handed electrons are mostly sensitive to up and charm quarks. Scattering on antidown and antistrange quarks is suppressed by a factor $(1-y)^2$. Right-handed positrons see predominantly anti-up and anticharm quarks, and the contribution from down and strange quarks is suppressed. There is no CC scattering of right-handed electrons and left-handed positrons in the Standard Model. Current HERA beams are not yet longitudinally polarized in the H1 and ZEUS interaction regions. From the above, it can be understood that the electron-proton CC cross section is considerably larger than the positron-proton cross section.

Figure 2 shows inclusive NC and CC cross section measurements carried out with the ZEUS detector. The CC propagator suppression relative to the NC data is clearly visible at low values of Q^2. A comparison to the Standard Model predictions shows good agreement.

The H1 experiment has carried out a detailed investigation of NC cross sections for the high statistics 1994 data sample. Figure 3 shows the measurements for electrons and positrons. The sample is restricted to events with $Q^2 > 160\ GeV^2$ and $0.0.5 < y < 0.80$. The agreement with the Standard Model is very good. In particular, the large Q^2 tail is potentially sensitive to new propagators or a finite charge radius of quarks. The data has been used to place various limits on new effects. Typical values of such limits are:

- mass of leptoquark exchanged in the t-channel $> 1\ TeV$ (normalized to the effective electromagnetic coupling),
- scale for electron/positron-quark compositeness $\Lambda_c < 1\ TeV - 2.5\ TeV$, and
- charge radius of quarks $< 2.6 \times 10^{-18}\ m$.

The limits depend on assumptions on the coupling constants for leptoquarks, the particular type of leptoquark, the chiral structure of the new contact interaction, and possible interference effects with the Standard Model processes.

The classical way to present cross sections in deep inelastic lepton-proton scattering is the calculation of structure functions. The charged lepton probes at HERA are directly sensitive to the charged parton content in the proton (i.e., quarks and antiquarks). In the naïve quark parton model, the structure function $F_2(x)$ describes the momentum distributions of all flavors of quarks weighted with their respective electrical charge squared:

$$F_2(x) = \sum_{flavors} e_i^2\ [xq_i(x) + x\bar{q}_i(x)]. \tag{2}$$

This simple picture is modified by the existence of strong interactions mediated by gluons and described in the framework of Quantum Chromo Dynamics (QCD). This fact leads to the well-known scaling violations demonstrated by the Q^2 dependence of the measured structure functions. The corresponding QCD evolution of structure functions is in principle calculable in perturbative QCD which should be able to tell the experiments how F_2 evolves with Q^2. In practice, however, assumptions have to be made in order to obtain testable predictions. In the picture of parton splitting processes only (i.e., no recombination), two approaches for the calculation of purely inclusive cross sections are currently available. The DGLAP approach[7] orders parton emission processes according to Q^2 and sums over logarithmic terms $(\alpha_S\ lnQ^2)^n$. The BFKL approach[8] orders emissions in the

Bjorken scaling variable x of the emitted partons and sum terms in $(\alpha_S\ ln(1/x))^n$. The GLR picture[9] includes parton recombination effects and does therefore depend on the square of parton distributions. This "non-linear" evolution leads to a saturation of parton densities in the proton at very low x.

Experimental challenges in the measurement of structure functions at HERA lie mainly in the detection of the scattered electron/positron, in particular, at very low angles corresponding to very low Q^2 which in turn corresponds also to very low x. Beyond simple detection, the measurement of kinematical variables requires in addition a precise determination of the energy and position of electrons and positrons. The HERA detectors have achieved an excellent understanding in particular of the calibration of their electromagnetic calorimeters. Figure 4 shows the raw electron spectrum recorded with the H1 backward electromagnetic calorimeter. The comparison to a simple parametrized simulation demonstrates the good understanding of the electromagnetic energy scale and resolution.

The data-taking strategy for DIS physics follows the constraints given by the Rutherford cross-section formula. Whereas for larger values of Q^2 the available statistics is the major concern, the situation is different for the low Q^2 (and low x) part. Here, experimental limitations are at least partly given by the angular acceptance of the backward (electron direction) electromagnetic calorimeters. Cross sections are however so large that it is worthwhile to perform dedicated runs with substantially smaller integrated luminosity, but the HERA interactions vertex shifted away from the calorimeters. This configuration allows the experiments to record data at Q^2 values as low as 2 GeV^2 (H1) and 1 GeV^2 (ZEUS). Other strategies to extend the lower end of the Q^2 range are the use of trailing bunches (satellite bunches) at HERA and the analysis of events with a photon radiated from the incoming lepton, thus reducing the effective beam energy (ZEUS "ISR" data).

The HERA F_2 data are presented in two ways. The traditional way is to take all data points at a given value of x and to plot them as a function of the momentum transfer Q^2. In this way, scaling violations (i.e., the evolution of the parton densities with Q^2) can be seen directly and compared to perturbative QCD predictions. Figure 5 presents the H1 data from the 1994 data-taking period.

The two HERA frontiers are clearly visible. At relatively large x values, the H1 data extend to Q^2 values up to almost $10,000$ GeV^2. At the low Q^2 end, measurements at x values as low as 0.5×10^{-4} are presented. For comparison,

measurements from fixed-target muon scattering experiments are also included. The full line represents a fit to the experimental data based on a next-to-leading (NLO) DGLAP calculation. The fit is based on parametrizations for the parton densities (nonsinglet quarks at large x, singlet quarks at all values of x, and gluons at small x). Based on the perturbative DGLAP NLO calculation, the evolution is started at $Q_0^2 = 4\ GeV^2$. The fit describes the data very well.

QCD does not provide a direct prediction for the x dependence of parton densities. The theoretical approach is to parametrize the parton densities according to:

$$xq_i(x, Q_0^2) \ \propto \ x^{\alpha_i}(1 - x\ )^{\beta_i}. \tag{3}$$

An assumption has to be made on the low-x behavior of the structure function. The case $\alpha = -0.5$ is usually referred to as "Lipatov behavior" and corresponds to a steeply rising parton density at low x values. Such approaches together with a perturbative QCD evolution starting at values of Q_0^2 of approximately 4 GeV^2 (Ref. 10) can describe the measured x dependence of F_2 at fixed Q^2. Another approach is to start the QCD evolution from very low Q_0^2 values (i.e., $\propto 0.23\ GeV^2$) where a valence type x distribution is used.[11] This strategy (GRV) does not require an assumption on the shape of the very low-x behavior of parton densities. The x dependence of the HERA data at fixed values of Q^2 is shown in Figs. 6 and 7 for H1 and ZEUS.

Both experiments do observe a very strong rise of F_2 with falling x. No indications of saturation effects are visible. The rise is well-described by either the NLO QCD fit (shown in comparison to the H1 data) or the GRV approach (shown in the case of the ZEUS data).

3 Gluons and the Hadronic Final State

The theory of perturbative QCD uses the elementary couplings between quarks and gluons and between the gluons themselves to describe the observed effects of scaling violations in the structure function F_2. As mentioned above, this structure function is directly sensitive only to the (anti)quark content of the proton but does not provide a straightforward measurement of the gluon distribution $xg(x)$. Two strategies have been used by the HERA experiments to obtain a determination of the gluon content in the proton.[12–14]

The first strategy is based on the assumption of perturbative QCD describing the observed scaling violations. The gluon distribution can either be extracted from the QCD fit described in the previous chapter or by measuring the local logarithmic derivatives of F_2 (Prytz method).[15] The latter method is based on the assumption that scaling violations arise mainly through the splitting process of gluons into quark-antiquark pairs. In this picture, the logarithmic derivative is directly proportional to the gluon structure function $xg(x)$.

$$\frac{\partial F_2}{\partial log Q^2} \;\propto\; \alpha_S\,(Q^2)\;xg(x). \tag{4}$$

The second strategy provides direct access to the gluon content of the proton without assumptions on the mechanism of scale breaking in inclusive cross sections. The method is based on leading order α_S QCD corrections to the naïve QPM process which is of purely electromagnetic nature. To order α_s, two basically different subprocesses come into play. Gluons can be radiated either from the incoming or the outgoing quark taking part in the hard scattering process (QCD Compton process). Secondly, an initial gluon from the proton can interact with the virtual photon via the exchange of a virtual quark line giving rise to an observable quark-antiquark pair in the final state (boson-gluon fusion process). Typically, the final state arising from such processes contains two hadronic jets in addition to the proton fragment which, to a large extent, disappears undetected down the beampipe. Such events are denoted as 2 + 1 jet events. The QCD Compton process is initiated from an (anti)quark in the proton in contrast to the BGF process originating from a gluon. Consequently, the 2 + 1 cross section observable in the experiment receives contributions from the quark and the gluon content in the proton:

$$\sigma_{2+1} \;\propto\; \alpha_S\,(Ag \;+\; Bq). \tag{5}$$

Using the known (and measured) quark distribution F_2, the gluon structure function $xg(x)$ can be obtained. Figure 8 summarizes the measurements of gluon structure functions at HERA. The strong rise towards low values of x is also seen for the gluons. Both experiments and the two methods give consistent results. The two theoretical approaches described in the previous chapter are in agreement with the experimental data.

4 Diffractive Scattering

The parton model together with QCD has proven to be able to provide a good description of a variety of different processes involving hadrons in the final and/or initial state. On the other hand, especially in hadron-hadron collisions, a large amount of data is well-described by Regge theory which describes interactions between hadrons through the exchange of Regge trajectories (associated with mesons or a Pomeron $I\!P$), the latter carrying quantum numbers of the vacuum.

The structure of the Pomeron $I\!P$ has been suggested to be of partonic nature, evidence for which was found in proton-antiproton collisions by the UA8 experiment.[16] At HERA, diffractive electron-proton scattering would result in events having a region without hadrons around the proton beam direction [called a "large rapidity gap" (LRG)]. In normal DIS, this region of phase space in rapidity between the struck parton and the proton remnant (both colored objects) is filled with particles from the hadronization of the color field between the parton and the remnant. This class of DIS events has been observed at HERA.[17,18] A pictorial representation of this process in comparison to the standard DIS process is shown in Fig. 9. For diffractive DIS, additional kinematical variables can be defined using the four-momentum P' of the colorless remnant (either a nucleon or a higher mass baryon excitation) in the final state:

$$x_{I\!P} = \frac{q \cdot (P - P')}{q \cdot P} \qquad \beta = \frac{Q^2}{2q \cdot (P - P')} \qquad t = (P - P')^2. \quad (6)$$

The variables x, $x_{I\!P}$,and β are related via $x = x_{I\!P} \cdot \beta$. With the setup of the HERA detectors, the remnant system is not detected; thus the squared momentum transfer t from the incident proton to the remnant system is not measured. Defining M_X to be the invariant mass of the virtual boson-Pomeron system, $x_{I\!P}$ and β can be written as

$$x_{I\!P} = \frac{Q^2 + M_X^2 - t}{Q^2 + W^2 - M_p^2} \qquad \beta = \frac{Q^2}{Q^2 + M_X^2 - t} \quad , \quad (7)$$

where M_p is the mass of the proton. When M_p^2 and $|t|$ are small ($M_p^2 \ll Q^2(W^2)$ and $|t| \ll Q^2(M_X^2)$), $x_{I\!P}$ can be interpreted as the fraction of the proton four-momentum transferred to the Pomeron, and β can be viewed as the four-momentum fraction of the quark entering the hard scattering relative to the

Pomeron in analogy to the definition of the Bjorken scaling variable x for a parton relative to the proton. In the kinematic region under investigation, both M_p^2 and $|t|$ can be neglected, and therefore, $x_{I\!P}$ and β can be calculated from M_X^2, Q^2, and W^2 as

$$x_{I\!P} \approx \frac{M_X^2 + Q^2}{W^2 + Q^2} \qquad\qquad \beta \approx \frac{Q^2}{M_X^2 + Q^2}. \tag{8}$$

The measured differential cross-section $\frac{d^3\sigma}{d\beta dQ^2 dx_{I\!P}}$ for the DIS events with a LRG has been shown by both HERA experiments to be compatible with a universal dependence on $x_{I\!P}$ (Refs. 17, 20). The differential cross section is expressed in terms of a structure function depending on three variables $F_2^{D(3)}$. This universal dependence can be interpreted as an intercept of a leading Regge trajectory. The value obtained for this intercept is compatible with the intercept of the Pomeron describing soft hadronic interactions, and thus, gives evidence for the diffractive nature of the process. The remaining term of the differential cross section then depends only on β and Q^2. It can be converted to a structure function $\tilde{F}_2^D(\beta, Q^2)$. This structure function exhibits scaling behavior (i.e., no substantial Q^2 dependence) and thus, leads to the evidence for a partonic substructure in the process. Figure 10 shows the measurement of $F_2^{D(3)}$ from the two HERA experiments H1 and ZEUS.

Acknowledgments

It is a pleasure to thank the SLAC Summer Institute for the kind invitation to present HERA results at the Topical Conference. The Institute is a special event in the annual calendar of high-energy physics with a very fruitful and lively atmosphere. Special thanks to Lilian DePorcel for her tireless efforts which made this school such a success. I also wish to thank my colleagues from the H1 and ZEUS Collaborations for providing me with their latest results.

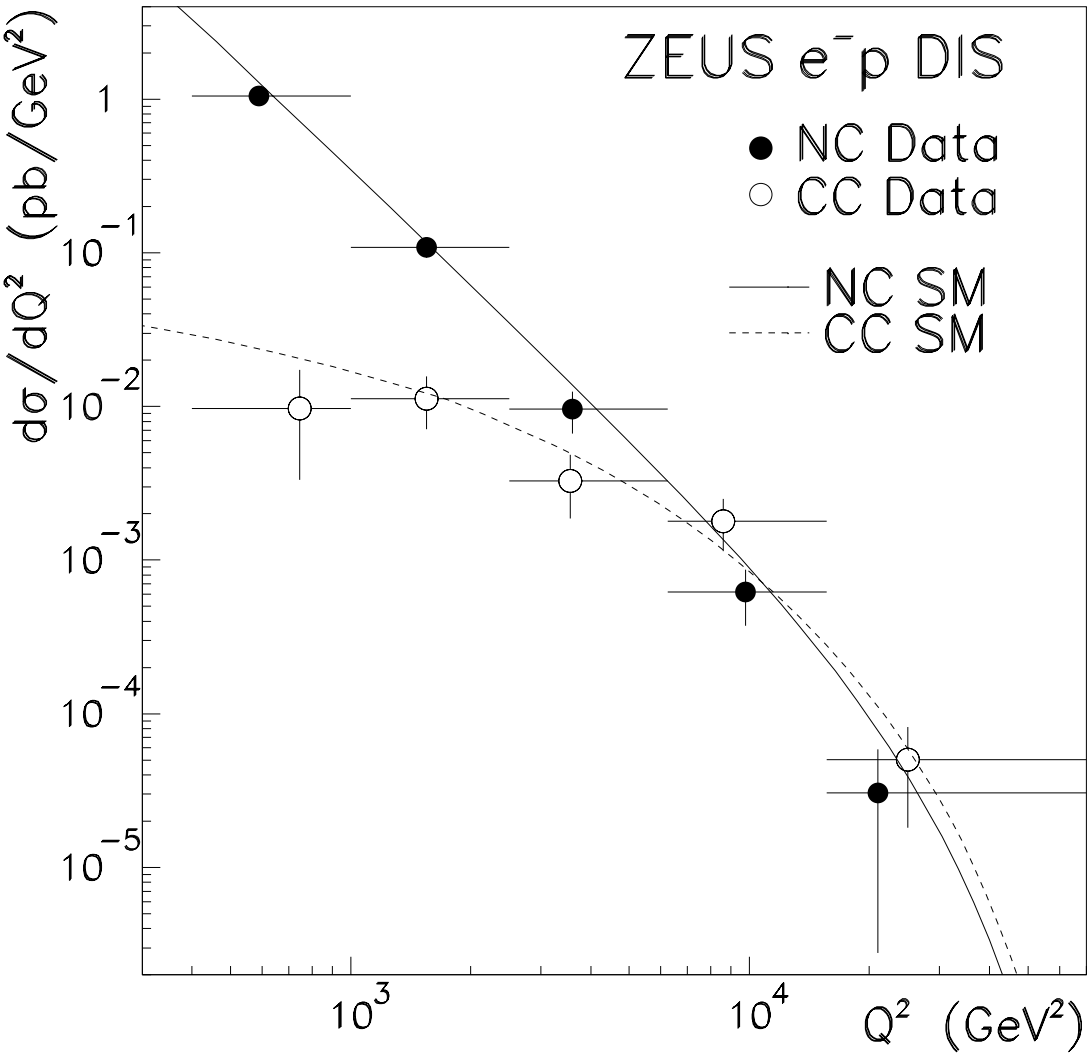


Figure 2: Inclusive neutral current and charged current cross sections in electron-proton scattering from ZEUS at HERA (1993 data).

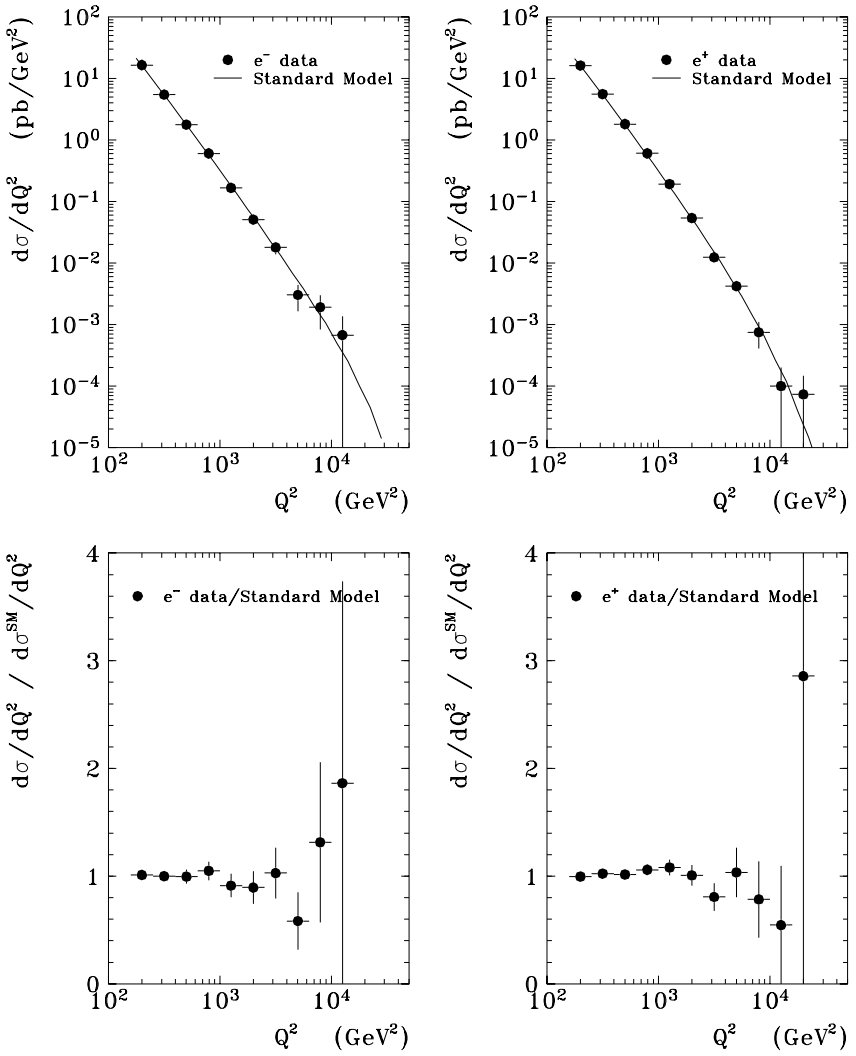


Figure 3: Inclusive neutral-current cross section in positron-proton and electron-proton scattering from H1 at HERA (1994 data) and ratios to Standard Model prediction.

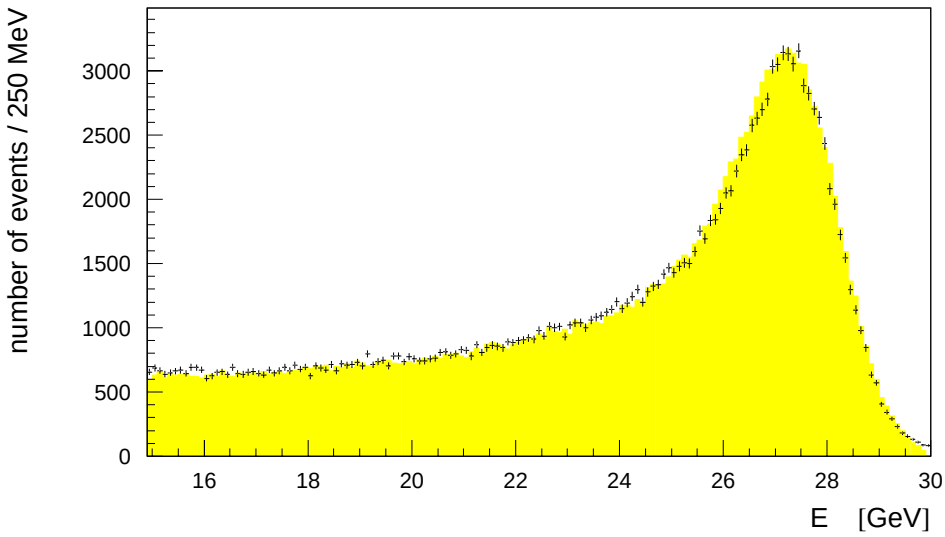


Figure 4: Raw electron spectrum observed in the H1 backward calorimeter compared to a simulation (shaded histogram) with parametrized detector response (1994 data).

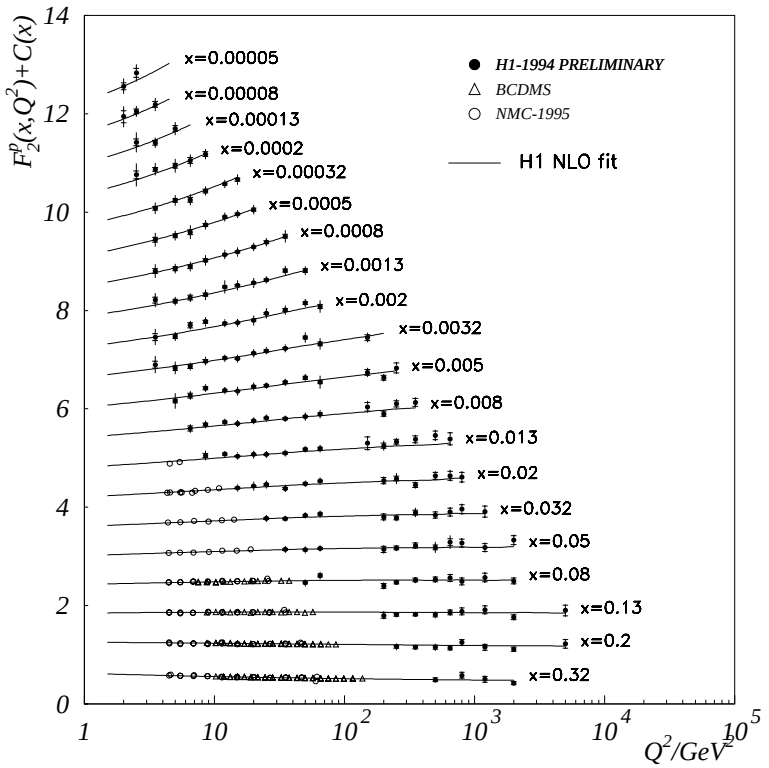


Figure 5: Structure function $F_2(x, Q^2)$ for fixed values of x as measured by the H1 experiment at HERA. Low-x data from the fixed-target muon scattering experiments NMC and BCDMS are also plotted. The full lines represent the NLO fit by H1 as described in the text.

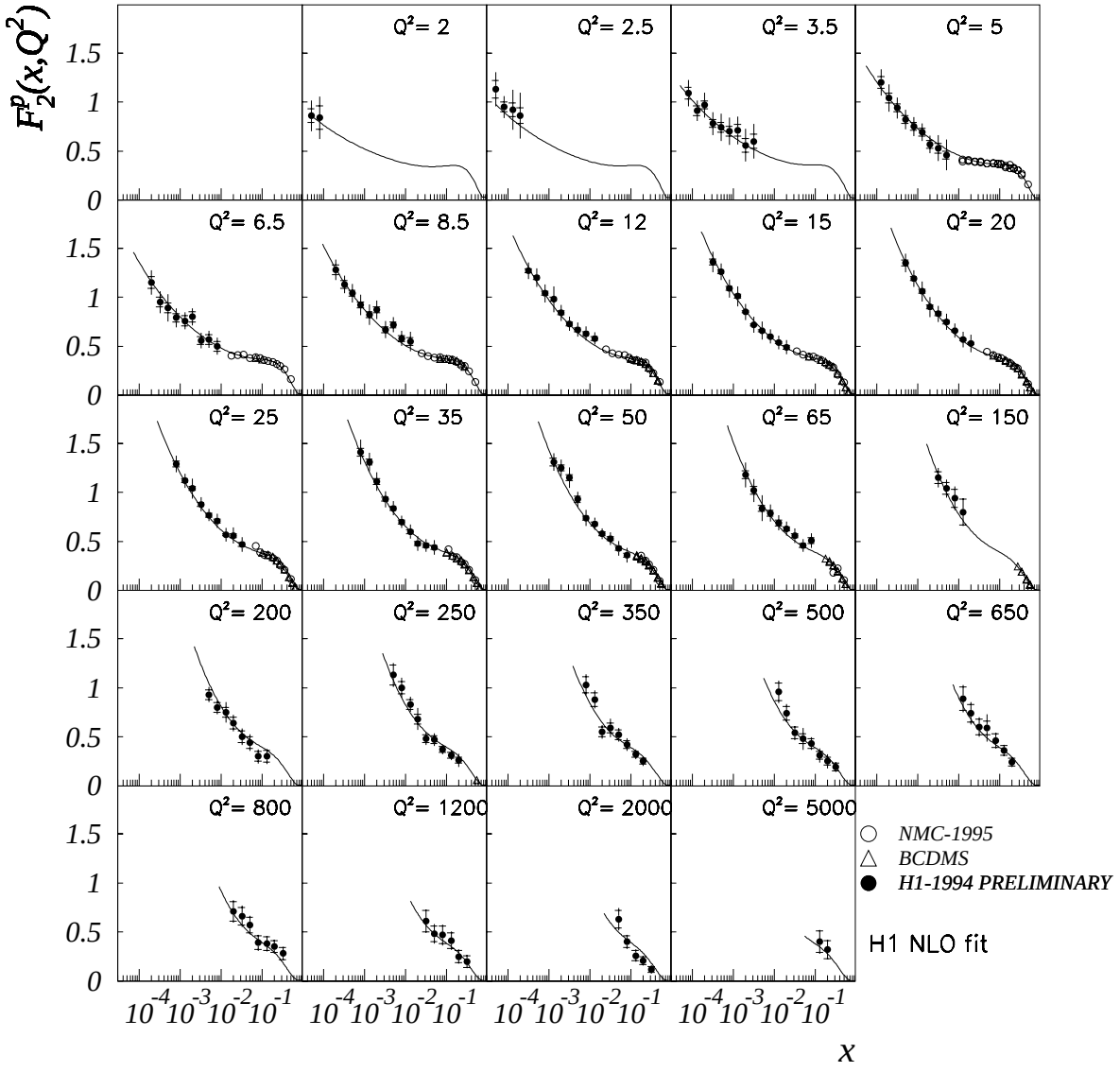


Figure 6: Structure function $F_2(x, Q^2)$ for fixed values of Q^2 as measured by the H1 experiment at HERA. Low-x data from the fixed target muon scattering experiments NMC and BCDMS are also plotted. The full lines represent the NLO fit by H1 as described in the text.

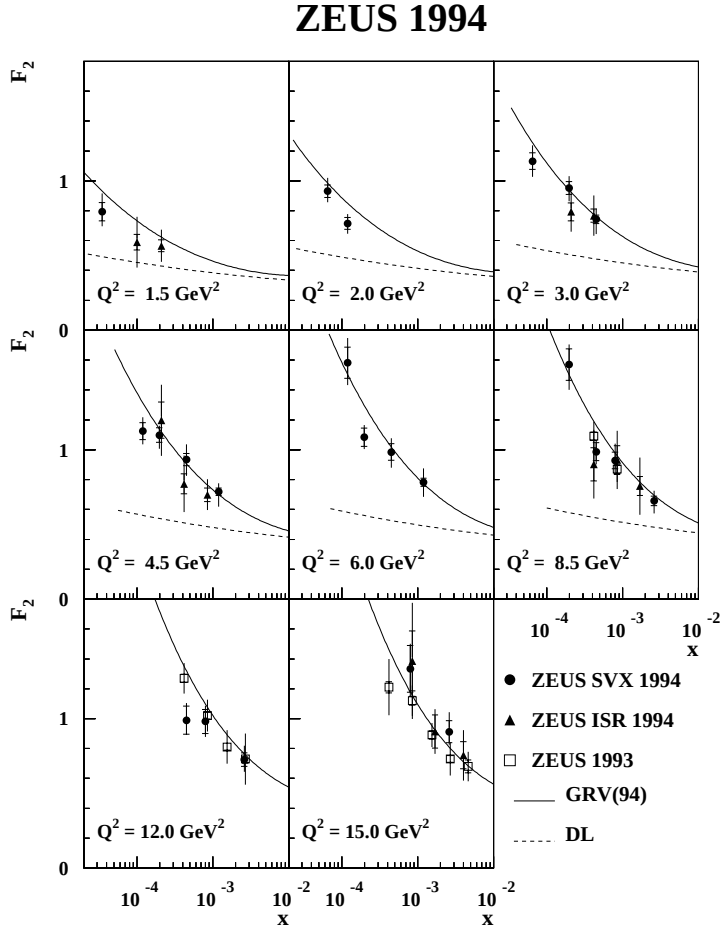


Figure 7: Very low Q^2 data of $F_2(x, Q^2)$ at fixed Q^2 from the ZEUS experiment at HERA. The special data-taking techniques are explained in the text.

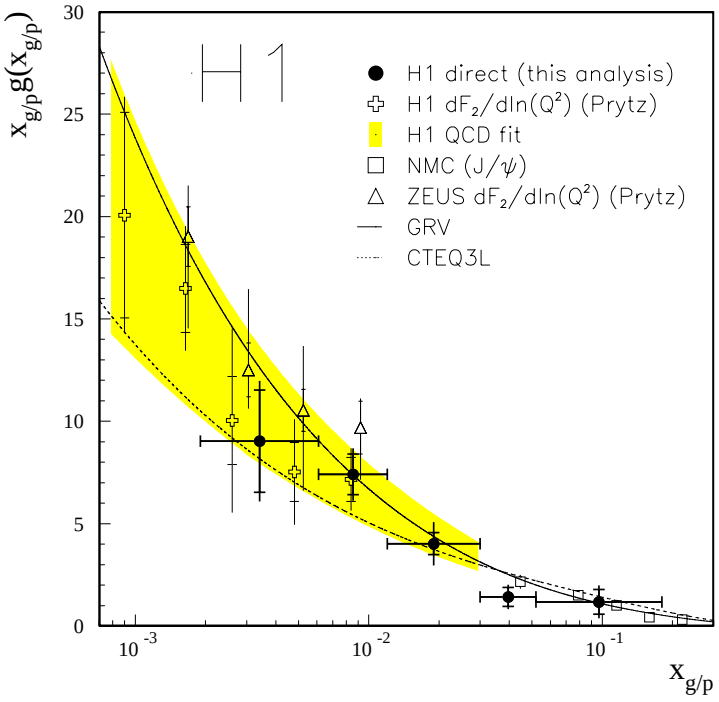


Figure 8: Summary plot of gluon structure function determinations at HERA.

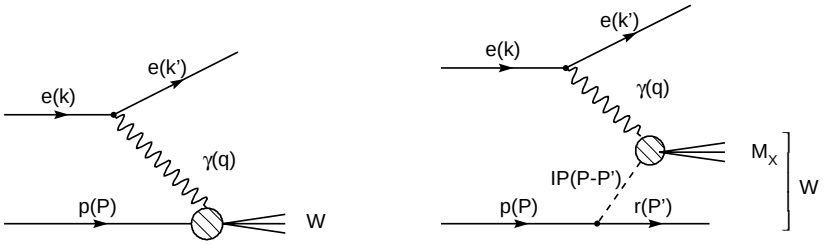


Figure 9: Factorization of Pomeron radiation in diffractive events (right figure) in contrast to the standard DIS process (left figure).

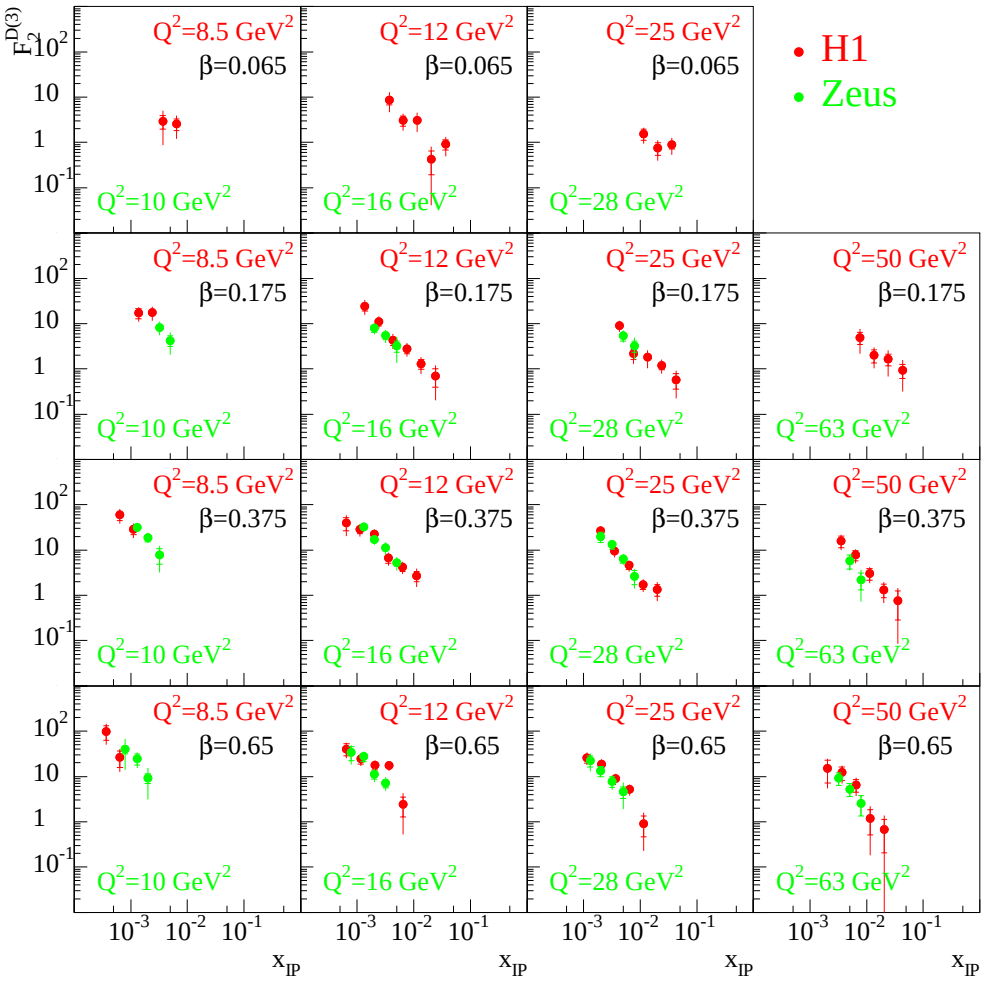


Figure 10: Summary of $F_2^{D(3)}$ data from H1 and ZEUS. Q^2 and β are fixed.

References

[1] H1 Collaboration, T. Ahmed *et al.*, DESY preprint 93-103, to be published in Nucl. Instrum. Methods.

[2] ZEUS Detector Status Report 1993, available on the WWW under http://zow00.desy.de.8000/bluebook/bluebook.html.

[3] U. Mallik, talk given at this conference.

[4] H1 Collaboration, T. Ahmed *et al.*, Phys. Lett. B **353**, 578 (1995).

[5] H1 Collaboration, T. Ahmed *et al.*, DESY-95-102, subm. to Zeit. Phys. C.

[6] ZEUS Collaboration, M. Derrick *et al.*, Phys. Rev. Lett. **75**, 1006 (1995).

[7] Y. L. Dokshitzer, JTEP **46**, 641 (1977); V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **45**, 438 (1972); and G. Altatrelli and G. Parisi, Nucl. Phys. B **126**, 297 (1977).

[8] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Phys. Lett. B **60**, 50 (1975); Sov. Phys. JTEP **44**, 443 (1976); Sov. Phys. JTEP **45**, 199 (1977); and I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[9] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Nucl. Phys. B **188**, 555 (1981); and Phys. Rep. **100** (1983).

[10] A. D. Martin, W. J. Stirling, and R. G. Roberts, Phys. Lett. B **306**, 145 (1993); Phys. Lett. B **309**, 492 (1993); and J. Botts *et al.*, Phys. Lett. B **304**, 159 (1993); preprint MSU-HEP 93/18 (1993).

[11] M. Glueck, E. Reya, and A. Vogt, Z. Phys. C **53**, 127 (1992); Phys. Lett. B **306**, 391 (1993).

[12] H1 Collaboration, T. Ahmed *et al.*, Phys. Lett. B **353**, 578 (1995).

[13] ZEUS Collaboration, Phys. Lett. B **345**, 576 (1995).

[14] H1 Collaboration, S. Aid *et al.*, Nucl. Phys. B **449**, 3 (1995).

[15] K. Prytz, Phys. Lett. B **311**, 286 (1993).

[16] UA8 Collaboration, R. Bonino *et al.*, Phys. Lett. B **211**, 239 (1988).

[17] ZEUS Collaboration, M. Derrick *et al.*, Phys. Lett. B **315**, 481 (1993).

[18] H1 Collaboration, T. Ahmed *et al.*, Nucl. Phys. B **429**, 477 (1994).

[19] H1 Collaboration, T. Ahmed *et al.*, Phys. Lett. B **348**, 681 (1995).

[20] ZEUS Collaboration, M. Derrick *et al.*, DESY-95-093, subm. to Zeit. Phys. C.

SPIN STRUCTURE MEASUREMENTS FROM E143 AT SLAC

Linda M. Stuart*
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309

Representing the E143 Collaborations

ABSTRACT

Measurements have been made of the proton and deuteron spin structure functions, g_1^p and g_1^d at beam energies of 29.1, 16.2, and 9.7 GeV, and g_2^p and g_2^d at a beam energy of 29.1 GeV. The integrals $\Gamma_p = \int_0^1 g_1^p(x,Q^2)dx$ and $\Gamma_d = \int_0^1 g_1^d(x,Q^2)dx$ have been evaluated at fixed $Q^2 = 3$ (GeV/c)^2 using the 29.1 GeV data to yield $\Gamma_p = 0.127 \pm 0.004(\text{stat.}) \pm 0.010(\text{syst.})$ and $\Gamma_d = 0.041 \pm 0.003 \pm 0.004$. The Q^2 dependence of the ratio g_1/F_1 has been studied and is found to be small for $Q^2 > 1$ (GeV/c)^2. Within experimental precision the g_2 data are well-described by the twist-2 contribution, g_2^{WW}. Twist-3 matrix elements have been extracted and are compared to theoretical predictions. The asymmetry A_2 has also been measured and is found to be significantly smaller than the positivity limit $\sqrt{R}$ for both targets. A_2^p is found to be positive and inconsistent with zero.

*Work supported in part by Department of Energy contract DE-AC03-76SF00515.

1 Introduction

Measurements of nucleon spin-dependent structure functions are valuable tools used to understand the complex nature of nucleon structure. These structure functions are probes of the longitudinal and transverse quark and gluon polarization distributions inside the nucleons. Measurements of these structure functions allow us to test sum rules, quark model predictions, and QCD predictions.

The spin dependent structure functions $g_1(x,Q^2)$ and $g_2(x,Q^2)$ are measured by scattering longitudinally polarized leptons from a target which is polarized either longitudinally or transversely. The longitudinal ($A_\parallel$) and transverse ($A_\perp$) asymmetries are formed from combining data taken with opposite beam helicity, and the structure functions are determined from these asymmetries:

$$\begin{aligned} g_1(x,Q^2) &= \frac{F_1(x,Q^2)}{d}\ [A_\parallel + \tan(\theta/2)A_\perp], \\ g_2(x,Q^2) &= \frac{yF_1(x,Q^2)}{2d} \quad \left[\frac{E+E'\cos\theta}{E'\sin\theta}A_\perp - A_\parallel\right], \end{aligned} \tag{1}$$

where E is the incident electron energy, E' is the scattered electron energy, θ is the scattering angle, x is the Bjorken scaling variable, Q^2 is the four-momentum transfer squared, $y=(E-E')/E$, $d=[(1-\epsilon)(2-y)]/[y(1+\epsilon R(x,Q^2))]$, $\epsilon^{-1} = 1+2[1+\gamma^{-2}]\tan^2(\theta/2)$, $\gamma = 2Mx/\sqrt{Q^2}$, M is the nucleon mass, $F_1(x,Q^2)$ is one of the spin-averaged structure functions, and $R(x,Q^2)=\sigma_L/\sigma_T$ is the ratio of the longitudinal to transverse virtual photon absorption cross sections. Also of interest are the virtual photon absorption asymmetries

$$A_1 = \frac{\sigma_{1/2}-\sigma_{3/2}}{\sigma_{1/2}+\sigma_{3/2}} \qquad \text{and} \qquad A_2 = \frac{2\sigma_{TL}}{\sigma_{1/2}+\sigma_{3/2}}, \tag{2}$$

where $\sigma_{1/2}$ and $\sigma_{3/2}$ are the virtual photon-nucleon absorption cross sections for total helicity between photon and nucleon of 1/2 and 3/2 respectively, and σ_{TL} is an interference term between the transverse and longitudinal photon-nucleon amplitudes. These asymmetries are also determined from the measured asymmetries:

$$\begin{aligned} A_1 &= \frac{1}{d}\left[A_\parallel(1+xM/E) - A_\perp\frac{xM}{E\tan(\theta/2)}\right], \\ A_2 &= \frac{\gamma(2-y)}{2d}\left[A_\perp\frac{y(1+xM/E)}{(1-y)\sin\theta} + A_\parallel\right]. \end{aligned} \tag{3}$$

1.1 Physical Interpretation of g_1

The structure function $g_1(x)$ is interpreted in the naive parton model as the charge weighted difference between momentum distributions for quarks and nucleon helicities aligned parallel ($\uparrow$) and anti-parallel ($\downarrow$):

$$g_1(x) = \frac{1}{2}\sum_i e_i^2[q_i^{\uparrow}(x) - q_i^{\downarrow}(x)] \equiv \sum_i e_i^2 \Delta q_i(x), \tag{4}$$

where e_i is the charge of quark flavors i, and $q_i^{\uparrow(\downarrow)}(x)$ are the quark plus antiquark momentum distributions. The quantity $\int_0^1 \Delta q_i(x)dx = \Delta i$ refers to the helicity of quark species $i = u, d, s$ in the proton, and $\Delta q = \Delta u + \Delta d + \Delta s$ is the net helicity of quarks. Using measurements of $\int_0^1 g_1(x)dx$, g_A/g_V and F/D as well as the QCD corrections to the sum rules, one can separately extract the quantities Δ_i.[1]

1.2 Physical Interpretation of g_2

Unlike g_1, the interpretation of g_2 in the naive parton model is ambiguous.[2] A more advanced light-cone parton model[3,4] as well as an operator product expansion (OPE) analysis[5] indicate that there are three components contributing to g_2. These components include the leading twist-2 part, $g_2^{WW}(x, Q^2)$, coming from the same set of operators that contribute to g_1, another twist-2 part coming from the quark transverse polarization distribution $h_T(x, Q^2)$, and a twist-3 part coming from quark-gluon interactions $\xi(x, Q^2)$:

$$g_2(x, Q^2) = g_2^{WW}(x, Q^2) - \int_x^1 \frac{\partial}{\partial y}\left(\frac{m}{M}h_T(y, Q^2) + \xi(y, Q^2)\right)\frac{dy}{y}. \tag{5}$$

The quark mass is denoted by m, and the g_2^{WW} expression of Wandzura-Wilczek[6] is given by

$$g_2^{WW}(x, Q^2) = -g_1(x, Q^2) + \int_x^1 \frac{g_1(y, Q^2)}{y}dy. \tag{6}$$

2 Sum Rules

2.1 Bjorken Sum Rule

A sum rule developed by Bjorken[7] relates the integral over the proton minus neutron spin structure functions to the nucleon beta decay weak coupling constants.

It is believed to be strictly valid at infinite Q^2:

$$\int \left(g_1^p(x) - g_1^n(x)\right) dx = \frac{1}{6}\frac{g_A}{g_V} \qquad Q^2 = \infty, \tag{7}$$

where g_A and g_V are the nucleon axial-vector and vector coupling constants and $g_A/g_V = 1.2573 \pm 0.0038$.[8] The advent of QCD corrections has brought this sum rule into the regime where it and thus the QCD corrections can be experimentally tested. These non-singlet corrections[9] to order three for three quark flavors are $C_{NS} = \left[1 - \alpha_s/\pi - 3.58\left(\alpha_s/\pi\right)^2 - 20.22\left(\alpha_s/\pi\right)^3\right]$ where $\alpha_s(Q^2)$ is the strong coupling constant.

2.2 Ellis-Jaffe Sum Rule

Other sum rules of interest for g_1, although less rigorous than the Bjorken sum rule, are the Ellis-Jaffe sum rules[10] which were derived using SU(3) symmetry and assuming the strange sea in the nucleons is unpolarized.

$$\begin{aligned}
\Gamma_1^p(Q^2) &= \int_0^1 g_1^p(x,Q^2)dx = \frac{1}{18}\left[C_{NS}(3F+D) + 2C_S(3F-D)\right] \\
\Gamma_1^n(Q^2) &= \int_0^1 g_1^n(x,Q^2)dx = \frac{1}{9}\left[-DC_{NS} + C_S(3F-D)\right],
\end{aligned} \tag{8}$$

where F and D are weak hyperon decay constants extracted from data[11] $F/D = 0.575 \pm 0.016$, $F + D = g_A/g_V$, and the second order singlet QCD corrections[12] are given by $C_S = \left[1 - 0.3333\alpha_s/\pi - 0.5495\left(\alpha_s/\pi\right)^2\right]$.

2.3 OPE Sum Rules

The OPE[2,5,13] is a useful technique within QCD because it separates the physics into a perturbative part which is easily treatable and a nonperturbative part which is parameterized in terms of unknown matrix elements of Lorentz-covariant operators. The OPE analysis of g_1 and g_2 yields an infinite number of sum rules

$$\begin{aligned}
\int_0^1 x^n g_1(x,Q^2)dx &= \frac{a_n}{2}, & n = 0,2,4,... \\
\int_0^1 x^n g_2(x,Q^2)dx &= \frac{1}{2}\frac{n}{n+1}(d_n - a_n), & n = 2,4,...
\end{aligned} \tag{9}$$

where a_n are the twist-2 and d_n are the twist-3 matrix elements of the renormalized operators. The OPE only gives information on the odd moments of the spin structure functions. The Wandzura-Wilczek relation in Eq. (6) can be derived from these sum rules by setting $d_n = 0$.

2.4 Burkhardt-Cottingham Sum Rule

The Burkhardt-Cottingham sum rule[14] for g_2 at large Q^2,

$$\int_0^1 g_2(x)dx = 0, \tag{10}$$

has been derived from from virtual Compton scattering dispersion relations. This sum rule does not follow from the OPE since the $n = 0$ sum rule is not defined for g_2 in Eq. (9). The Burkhardt-Cottingham sum rule relies on g_2 obeying Regge theory which may not be a good assumption. A non-Regge divergence of g_2 at low x would invalidate this sum rule[2,5] and such a divergence could be very difficult to detect experimentally.

2.5 Efremov-Teryaev Sum Rule

The Efremov-Teryaev sum rule[15] is derived in leading order QCD in which quark-gluon correlators have been included. This sum rule relates the g_1 and g_2 structure functions:

$$\int_0^1 x[2g_2(x) + g_1(x)]dx = 0.$$

3 Other Experiments

The earliest spin structure experiments, E80[16] E130[17] and EMC[18] measured $A_{||}$ for the proton only. Using the assumption that $g_1 \simeq F_1 A_1$, the EMC extracted $g_1^p(x, Q^2)$ with sufficient precision to test the Ellis-Jaffe sum rule which was violated and the so-called "spin crisis" was born. In the naive quark model this was interpreted to mean that the total quark helicity was small and consistent with zero while the strange quark helicity was negative and inconsistent with zero. This unexpected result has generated a lot of interest in the physics community. Many theoretical papers have surfaced to explain the data, better QCD corrections have been calculated bringing predictions closer to experimental results, and extensive experimental programs at SLAC, CERN and HERA were begun to learn more about nucleon spin structure. Results are now available from the SMC[19–22] experiment at CERN, and the E142[23] experiment at SLAC. These data include significantly more precise proton data, measurements on deuterons and ^{3}He (neutrons), and the first measurement of the transverse asymmetry A_2 for the proton.

These experiments have confirmed the Bjorken sum rule, and have shown that the Ellis-Jaffe sum rules for both the proton and neutron are violated.

4 This Experiment

For this experiment, E143,[24–27] longitudinally polarized electrons were scattered from polarized protons and deuterons into two independent spectrometers at angles of 4.5° and 7°. The beam polarization, typically $P_b = 0.85 \pm 0.02$, was measured with a Møller polarimeter. Measurements were made at three beam energies of 29.1, 16.2, and 9.7 GeV. The target cells were filled with granules of either $^{15}NH_3$ or $^{15}ND_3$, and were polarized using the technique of dynamic nuclear polarization. The targets could be polarized longitudinally or transversely relative to the beam by physically rotating the polarizing magnet. Target polarization P_t, measured by a calibrated NMR, averaged around 0.65 ± 0.017 for protons and 0.25 ± 0.011 for deuterons.

The experimental asymmetries $A_{\|}$ and $A_{\perp}$ were determined from

$$A_{\|} \text{ (or } A_{\perp}) = C_1\left(\frac{N_L - N_R}{N_L + N_R}\frac{1}{fP_bP_t} - C_2\right) + A_{RC}, \tag{11}$$

where N_L and N_R are the number of scattered electrons per incident electron for negative and positive beam helicity, where corrections have been made for charge-symmetric backgrounds and deadtime; f is the dilution factor representing the fraction of measured events originating from polarizable protons or deuterons within the target; C_1 and C_2 correct for the polarized nitrogen nuclei and for residual polarized protons in the ND_3 target; and A_{RC} are the radiative corrections, which include internal[28] and external[29] contributions.

4.1 Longitudinal results at $E = 29$ GeV

From the measured values of $A_{\|}$ and $A_{\perp}$ we calculated the ratios g_1^p/F_1^p and g_1^d/F_1^d using the definition given in Eq. (1). For $F_1(x,Q^2) = F_2(x,Q^2)(1+\gamma^2)/[2x(1 + R(x,Q^2))]$ we used the NMC[30] fits to $F_2(x,Q^2)$ data and the SLAC fit[31] to $R(x,Q^2)$, which was extrapolated to unmeasured regions for $x < 0.08$. These results[24,25] are shown in Fig. 1. Also included in the plots are the data from other experiments;[16–18,20,22] which are all in good agreement with the E143 results.

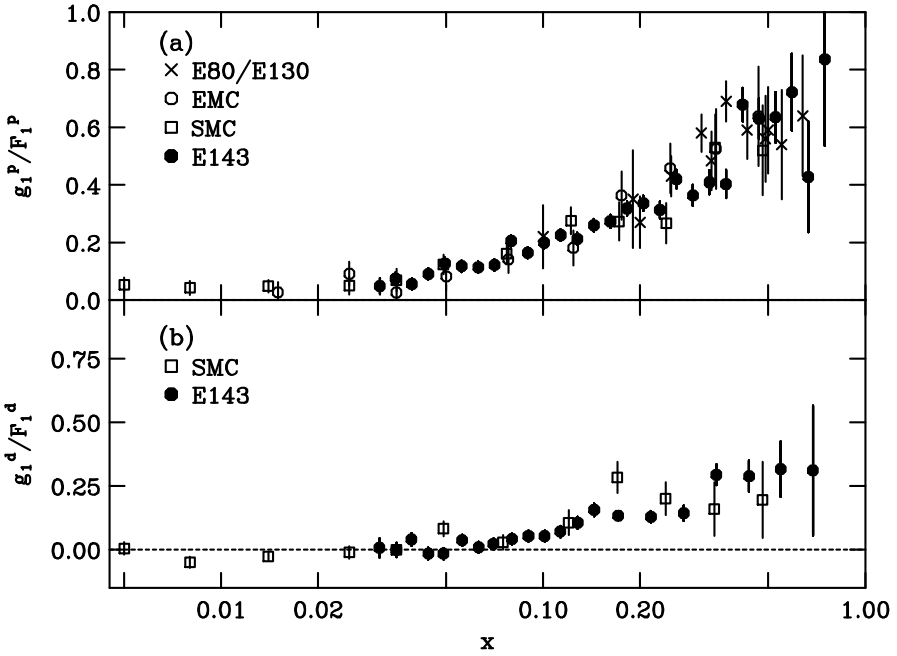


Figure 1: Measurements of g_1/F_1 for (a) proton and (b) deuteron for all experiments. The E143 data are in good agreement with all other data. Uncertainties for the E143 data include statistical contributions only.

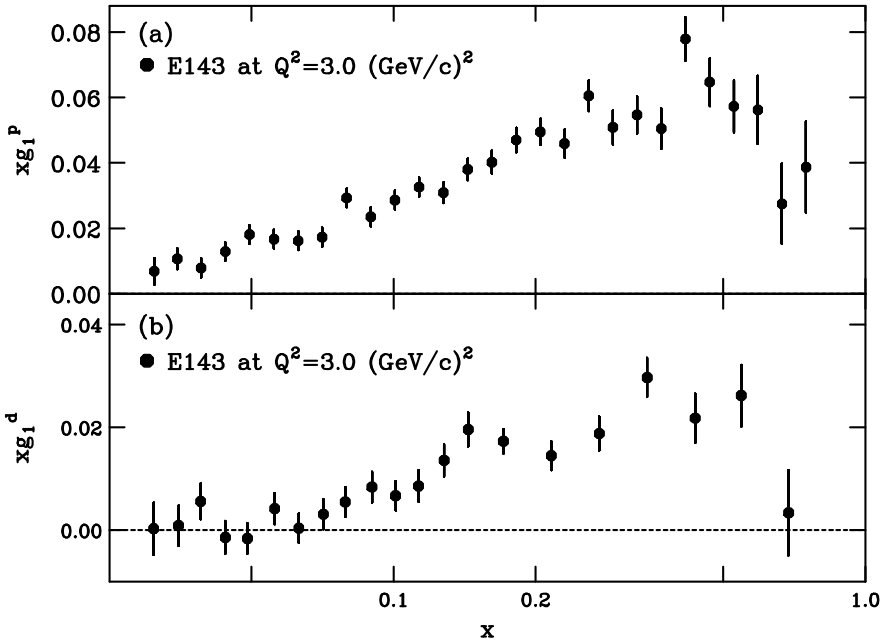


Figure 2: Measurements of xg_1 for (a) proton and (b) deuteron from experiment E143 at a constant $Q^2 = 3$ (GeV/c)^2. The uncertainties include statistical contributions only.

Values of xg_1^p and xg_1^d at the average $Q^2 = 3$ (GeV/c)^2 of this experiment are shown in Fig. 2. The evaluation at constant Q^2 is model dependent, and we have made the assumption that g_1/F_1 is independent of Q^2 which is believed to be reasonable for the kinematics of this experiment (see discussion on Q^2 dependence below).

Values of xg_1^p and xg_1^n from several experiments at an average $Q^2 = 5$ (GeV/c)^2 are shown in Fig. 3. The data were evolved to constant Q^2 assuming g_1/F_1 is independent of Q^2. The neutron results from this experiment[25] and from SMC[22] were extracted from proton and deuteron data using $g_1^d = \frac{1}{2}(g_1^p + g_1^n)(1 - \frac{3}{2}\omega_D)$, where ω_D is the probability that the deuteron is in a D state. Both experiments used $\omega_D = 0.05 \pm 0.01$.[8] We see from Fig. 3 that the data sets are in good agreement when evolved to the same Q^2.

The integrals over x of g_1 for the proton (Γ_1^p), deuteron (Γ_1^d), and neutron (Γ_1^n) were evaluated at a constant $Q^2 = 3$ (GeV/c)^2. The measured x region was $0.029 < x < 0.08$. The extrapolation from $x = 0.8$ to $x = 1$ was done assuming

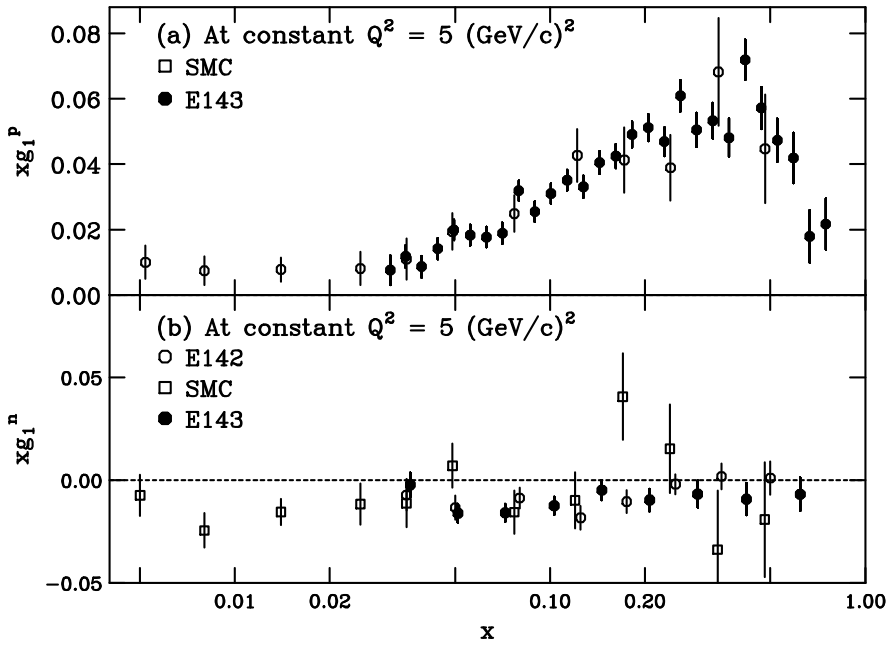


Figure 3: Measurements of xg_1 for (a) proton and (b) neutron for E143,[24,25] E142,[23] and SMC[20,22] at a constant $Q^2 = 5$ (GeV/c)^2. The data sets are in agreement. Uncertainties for the E143 data include statistical only.

that g_1 varies as $(1-x)^3$ at high x. The extrapolation from $x = 0$ to $x = 0.029$ was determined by fitting the low x data to a Regge[32] motivated form $g_1 = Cx^{-\alpha}$. An alternate form[33] $g_1 = C\ln(1/x)$, which provides a good fit to the low-x F_2 data from NMC and HERA, gives consistent results within the uncertainties. Table 1 gives a summary of the measured and extrapolated contributions to Γ_1^p and Γ_1^d. Table 2 shows the E143 measurements for Γ_1^p, Γ_1^d, Γ_1^n and $\Gamma_1^p - \Gamma_1^n$, as well as the corresponding Ellis-Jaffe and Bjorken sum rule predictions for $Q^2 = 3$ (GeV/c)^2. The data consistently demonstrate that the Ellis-Jaffe sum rule is violated. The most precise determination is given by the deuteron measurement which is more than 3 σ away from the prediction. Note that the E143 results agree with the E142 results[23] for $\Gamma_1^n = -0.022 \pm 0.011$ at $Q^2 = 2$ (GeV/c)^2, and the SMC[20,22] results for $\Gamma_1^p = 0.136 \pm 0.016$ and $\Gamma_1^d = 0.034 \pm 0.011$ at $Q^2 = 10$ (GeV/c)^2. The estimated Q^2 dependence of these quantities for $2 < Q^2 < 10$ (GeV/c)^2 is within the errors on all the experiments. Table 3 is a summary of the dominant systematic error contributions to the E143 measured integrals shown in Table 2.

Table 1: Results for Γ_1^p and Γ_1^d from experiment E143, broken up into the measured and extrapolated contributions. The measured contribution has a statistical and systematic uncertainty. The uncertainty on the extrapolated contributions is assumed systematic.

x Region	Γ_1^p	Γ_1^d
$0 < x < 0.029$	0.006 ± 0.006	0.001 ± 0.001
$0.029 < x < 0.8$	$0.120 \pm 0.004 \pm 0.008$	$0.040 \pm 0.003 \pm 0.004$
$0.8 < x < 1$	0.001 ± 0.001	0.000 ± 0.001
Total	$0.127 \pm 0.004 \pm 0.010$	$0.042 \pm 0.003 \pm 0.004$

Table 2: Summary of E143 g_1 sum rule tests

	Measured	Prediction	Sum Rule
Γ_p	$0.127 \pm 0.004 \pm 0.010$	0.160 ± 0.006	Ellis-Jaffe
Γ_d	$0.042 \pm 0.003 \pm 0.004$	0.069 ± 0.004	Ellis-Jaffe
Γ_n	$-0.037 \pm 0.008 \pm 0.011$	-0.011 ± 0.006	Ellis-Jaffe
$\Gamma_p - \Gamma_n$	$0.163 \pm 0.010 \pm 0.016$	0.171 ± 0.008	Bjorken

Table 3: Summary of the systematic error contributions to the E143 g_1 integrals				
source	Γ_p	Γ_d	Γ_n	Γ_p-Γ_n
beam polarization	0.003	0.001	0.001	0.004
target polarization	0.003	0.002	0.005	0.007
dilution factor	0.004	0.002	0.006	0.008
radiative corrections	0.002	0.002	0.006	0.007
F_2, R	0.004	0.001	0.002	0.005
Extrapolation	0.006	0.001	0.004	0.006
TOTAL	0.010	0.004	0.011	0.016

The violation of the Ellis-Jaffe sum rule implies that the assumption that the strange quark is unpolarized within the nucleon may be false. This can be seen by extracting the net quark helicity within the proton using the naive quark model[1] (See Eq. (4) and related discussion). Table 4 gives the extracted quark helicities as determined from the measurements of Γ_1^p and Γ_1^d and the SU(3) coupling constants F and D. The data include third order non-singlet and second order singlet QCD corrections. We see that the net quark helicity Δq is significantly less than a prediction[10] that $\Delta q = 0.58$ assuming zero strange quark helicity and SU(3) flavor symmetry in the baryon octet. Also, Δs is negative and significantly different from zero. Figure 4 shows a plot of Δq versus Δs as extracted from various experimental measurements at the appropriate Q^2. We see that all experiments are consistent with a small Δq and a Δs which is negative and inconsistent with zero.

Table 4: Extracted quark helicities from experiment E143.		
	Γ_p	Γ_d
Δu	0.81 ± 0.04	0.83 ± 0.02
Δd	-0.44 ± 0.04	-0.43 ± 0.02
Δs	-0.10 ± 0.04	-0.09 ± 0.02
Δq	0.27 ± 0.11	0.30 ± 0.06

4.2 Transverse results at $E = 29$ GeV

From the measured values of $A_{||}$ and $A_\perp$ at $E = 29$ GeV we have calculated g_2^p, g_2^d, A_2^p and A_2^d using Eqs. (1) and (3). The results for A_2 for the proton and

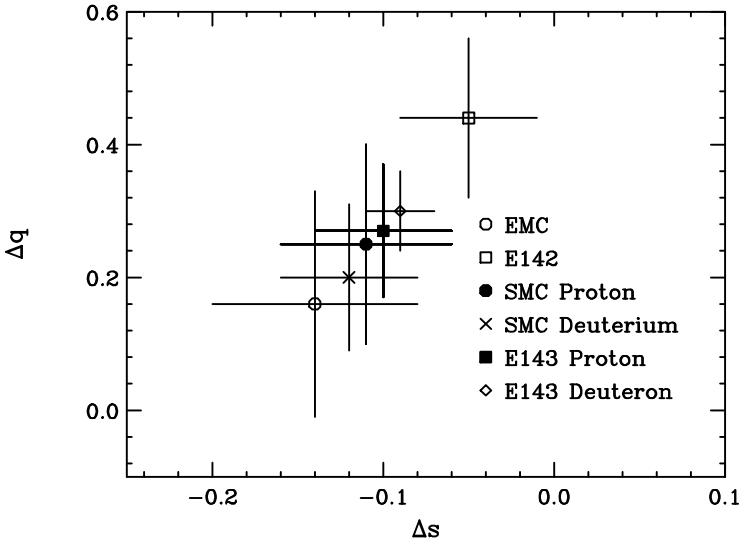


Figure 4: The quark helicity content of the proton as extracted from various measurements is shown for Δq versus Δs. The data include third order non-singlet and second order singlet QCD corrections.

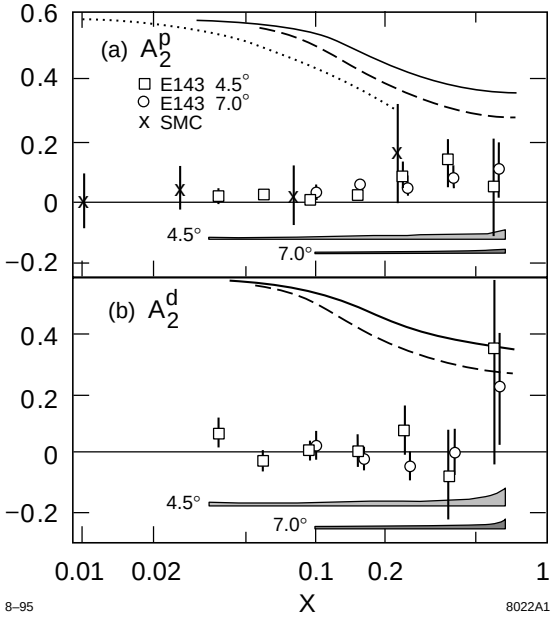


Figure 5: Measurements of (a) A_2^p, and (b) A_2^d from E143[26] and SMC.[21] Systematic errors are indicated by bands. The curves show the $\sqrt{R}$ [31] positivity constraints for the three data sets. The solid, dashed and dotted curves correspond to the 4.5° E143, 7.0° E143, and SMC kinematics, respectively. Overlapping data have been shifted slightly in x to make errors clearly visible.

deuteron are shown in Fig. 5. The systematic errors, dominated by radiative correction uncertainties, are indicated by bands for the two spectrometers used in the experiment. The data agree within errors despite the differences in Q^2 of the measurements (nearly a factor of two). Also in Fig. 5 are the proton results from SMC,[21] and the $\sqrt{R}$ [31] positivity limits for each data set. The data are much closer to zero than the positivity limit. Results for A_2^p are consistently > 0, and since A_2 is expected to be zero at high Q^2 (because $R \to 0$), these data indicate that A_2 must have Q^2 dependence. A comparison of the data with the hypothesis $A_2 = 0$ yields $\chi^2 = 73$ for the proton and $\chi^2 = 44$ for the deuteron for 48 degrees of freedom.

The results for xg_2 for the proton and deuteron are shown in Fig. 6. The g_2^d results are per nucleon. The systematic errors are indicated by bands. Also shown is the g_2^{WW} curve evaluated using Eq. (6) at $E = 29$ GeV and $\theta = 4.5°$. The same

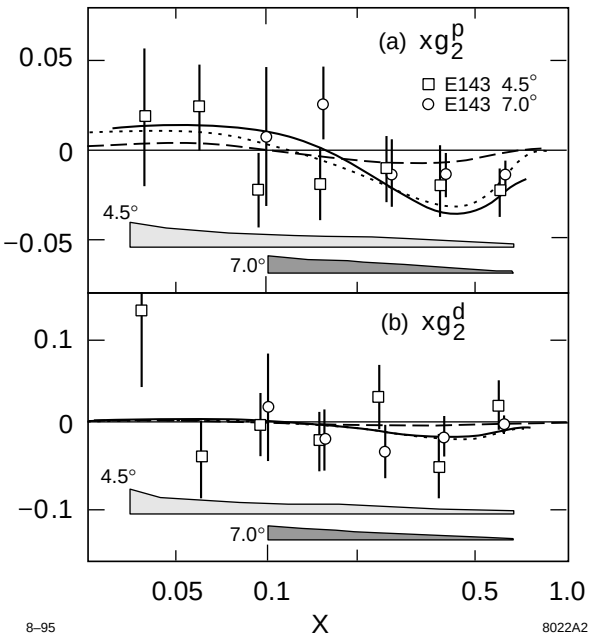


Figure 6: Spin structure function measurements for (a) xg_2^p, and (b) xg_2^d from E143. Systematic errors are indicated by bands. Overlapping data have been shifted slightly in x to make errors clearly visible. The solid curve shows the twist-2 g_2^{WW} calculation for the kinematics of the 4.5° spectrometer. The same curve for 7° is nearly indistinguishable. The bag model calculations at $Q^2 = 5.0$ (GeV/c)^2 by Stratmann[34] (dotted) and Song and McCarthy[35] (dashed) are indicated.

curve for $\theta = 7°$ is nearly indistinguishable. The values for g_2^{WW} were determined from $g_1(x, Q^2)$ evaluated from a fit to world data of A_1[27] and assuming negligible higher-twist contributions. Also shown are the bag model predictions of Stratmann[34] and Song and McCarthy,[35] which include both twist-2 and twist-3 contributions for $Q^2 = 5$ (GeV/c)^2. At high x the results for g_2^p indicate a negative trend consistent with the expectations for g_2^{WW}. The deuteron results are less conclusive because of the larger errors.

We can look for possible quark mass and higher twist effects by extracting the quantity $\overline{g_2}(x, Q^2) = g_2(x, Q^2) - g_2^{WW}(x, Q^2)$, If the term in Eq. (5) which depends on quark masses can be neglected then $\overline{g_2}(x, Q^2)$ is entirely twist-3. Our results can be seen in Fig. 7. Within the experimental uncertainty the data are consistent with $\overline{g_2}$ being zero but also with $\overline{g_2}$ being of the same order of magnitude as g_2^{WW}. Also shown in Fig. 7 are the bag model predictions of Stratmann[34] and Song and McCarthy[35] for $Q^2 = 5$ (GeV/c)^2 which compare favorably with the data given the large experimental uncertainties.

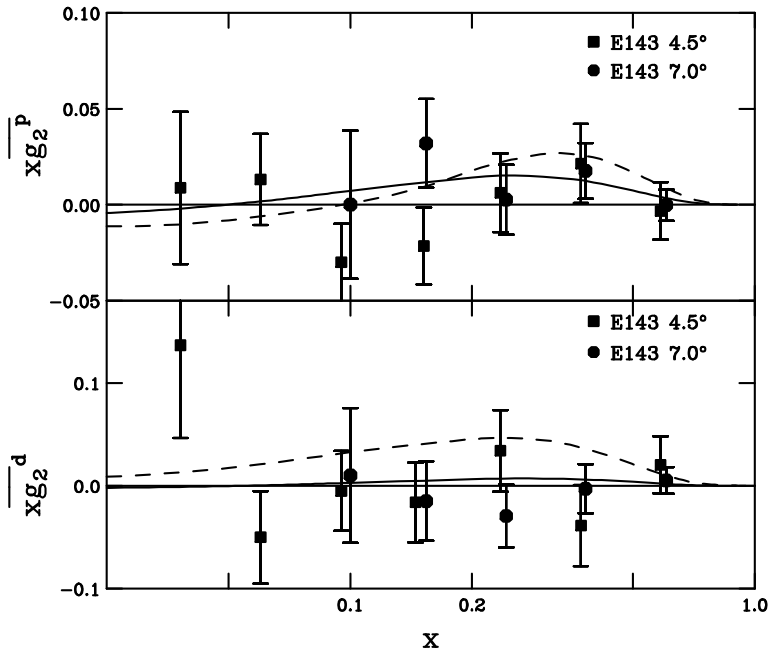


Figure 7: E143 results[26] for (a) $x\overline{g_2}^p$, and (b) $x\overline{g_2}^d$ Overlapping data have been shifted slightly in x to make errors clearly visible. The bag model calculations at $Q^2 = 5.0$ (GeV/c)^2 by Stratmann[34] (solid) and Song and McCarthy[35] (dashed) are indicated.

Using our results for the longitudinal spin structure functions g_1^p and g_1^d, we have computed the first few moments of the OPE sum rules, and solved for the twist-3 matrix elements d_n. These moments are defined to be $\Gamma_1^{(n)} = \int_0^1 x^n g_1(x)dx$ and $\Gamma_2^{(n)} = \int_0^1 x^n g_2(x)dx$. For the measured region $0.03 < x < 0.8$, we evaluated g_1 and corrected the twist-2 part of g_2 to fixed $Q^2 = 5$ (GeV/c)^2 assuming g_1/F_1 is independent of Q^2, and have averaged the two spectrometer results to evaluate the moments. Possible Q^2 dependence of $\overline{g_2}$ has been neglected. We neglect the contribution from the region $0 \leq x < 0.03$ because of the x^n suppression factor. For $0.8 < x \leq 1$, we assume that both g_1 and g_2 behave as $(1-x)^3$, and we fit the data for $x > 0.56$. The uncertainty in the extrapolated contribution is taken to be the same as the contribution itself. The results are shown in Table 5a. For comparison, in Table 5b we quote theoretical predictions[34–37] for d_2^p and d_2^d. For d_2^d the proton and neutron results were averaged and a deuteron D-state correction was applied. Our extracted values for d_n are consistent with zero, but the errors

Table 5a: Results for the moments $\Gamma_1^{(n)}$ and $\Gamma_2^{(n)}$ evaluated at $Q^2 = 5$ (GeV/c)^2, and the extracted twist-3 matrix elements d_n for proton (p) and deuteron (d) targets. The errors include statistical (which dominate) and systematic contributions.

	n	$\Gamma_1^{(n)}$	$\Gamma_2^{(n)}$	d_n
p	2	0.0121 ± 0.0010	-0.0063 ± 0.0018	0.0054 ± 0.0050
	4	0.0032 ± 0.0004	-0.0023 ± 0.0006	0.0007 ± 0.0017
	6	0.0012 ± 0.0002	-0.0010 ± 0.0003	0.0001 ± 0.0008
d	2	0.0040 ± 0.0008	-0.0014 ± 0.0030	0.0039 ± 0.0092
	4	0.0008 ± 0.0003	0.0000 ± 0.0010	0.0017 ± 0.0026
	6	0.0002 ± 0.0001	0.0001 ± 0.0005	0.0006 ± 0.0011

Table 5b: Theoretical predictions for the twist-3 matrix element d_2^p for proton and d_2^d for deuteron. The values for Q^2 are in (GeV/c)^2.

	Bag models		QCD sum rules	
	Ref.[35]	Ref.[34]	Ref.[36]	Ref.[37]
Q^2	5	5	1	1
d_2^p	0.0176	0.0060	-0.006 ± 0.003	-0.003 ± 0.006
d_2^d	0.0066	0.0029	-0.017 ± 0.005	-0.014 ± 0.006

are large. The results do not have sufficient precision to distinguish between the model predictions.

We have also evaluated the integrals $\int_{0.03}^{1} g_2(x)dx$ and $\int_{0.03}^{1} x[2g_2(x) + g_1(x)]dx$ for both the proton and deuteron structure functions. We do not attempt a low x extrapolation due to the theoretical uncertainty on the low x behavior of g_2. For the latter integral, the low x region is suppressed by x, so it is not unreasonable to assume that the low x extrapolation is negligible. The high-x extrapolation is done as discussed above. The results are given in Table 6, and are all consistent with zero within their large errors as expected from the Burkhardt-Cottingham and Efremov-Teryaev sum rules. Of course, we cannot really test the Burkhardt-Cottingham sum rule due to the uncertainty in the unmeasured low x behavior.

Table 6: Summary of E143 g_2 sum rule tests. The predictions for both sum rules are zero.

	$\int_{0.03}^{1} g_2(x)dx$	$\int_{0.03}^{1} x[2g_2(x) + g_1(x)]dx$
proton	-0.013 ± 0.028	0.008 ± 0.008
deuteron	-0.033 ± 0.082	-0.001 ± 0.014

4.3 Q^2 Dependence of g_1

Data for g_1 measured at a fixed energy of 29 GeV were discussed above. These data cover the range $1 < Q^2 < 10$ (GeV/c)^2 where the lower values of Q^2 are at the lower values of x. In order to evaluate sum rules at some fixed Q^2 it is necessary to extrapolate the data from the measured kinematics. Since this is a model-dependent procedure (e.g., assuming g_1/F_1 is independent of Q^2), it is useful to measure the Q^2 dependence by taking data at multiple beam energies. In E143 we made measurements at beam energies of 29.1, 16.2, and 9.7 GeV. The kinematic coverage of all these data sets where a Q^2-dependent measurement has been made is $0.03 < x < 0.6$ and $0.3 < Q^2 < 10$ (GeV/c)^2.

According to the GLAP equations[38] which give the predicted Q^2 dependence of the nucleon polarized and unpolarized quark and gluon distribution functions, g_1 is expected to evolve logarithmically in a similar way as the unpolarized structure functions $F_1(x, Q^2)$ and $F_2(x, Q^2)$. The Q^2 dependence of the ratio g_1/F_1 may be independent of Q^2 to a first approximation, but the precise behavior is sensitive to the underlying spin-dependent quark and gluon distribution functions. Measurements will help pin down this behavior. Fits have been made[39,40] of $g_1(x, Q^2)$

data using next-to-leading-order (NLO) GLAP equations.[41] The results indicate that NLO fits are more sensitive to the strength of the polarized gluon distribution function $\Delta G(x,Q^2)$ than previous leading-order (LO) fits.[41–45] In addition our understanding of the Q^2 dependence of g_1 is complicated by possible higher twist contributions which are not part of the GLAP equations. These terms are expected to behave as $C(x)/Q^2$, $D(x)/Q^4$, etc., where $C(x)$ and $D(x)$ are unknown functions.

The ratio g_1/F_1 has been extracted from the data taken in this experiment[27] as well as from other available data for the proton[16–18,20] and the deuteron[22] using the relations given in Eq. (1). The twist-2 model of Wandzura and Wilczek[6] given in Eq. (6) was used to describe g_2 for all data since the E143 g_2 data discussed above are in agreement with this model. The results for g_1^p/F_1^p and g_1^d/F_1^d are shown in Figs. 8 and 9, respectively, at eight values of x. Improved radiative corrections have been applied to the E80[16] and E130[17] results. Only statistical uncertainties have been plotted. For the present experiment, most systematic uncertainties (beam polarization, target polarization, fraction of polarizable nucleons in the target) for a given target are common to all data and correspond to an overall normalization error of about 5% for the proton data and 6% for the deuteron data. The remaining point-to-point systematic uncertainties (radiative corrections, model uncertainties for $R(x,Q^2)$, resolution corrections) vary over x from a few percent to 15%, and are consistently less than the statistical uncertainties for all data. We see in Figs. 8 and 9 that g_1/F_1 is approximately independent of Q^2 at fixed x, although there is a noticeable trend for the ratio to decrease for $Q^2 < 1$ (GeV/c)^2.

We have performed several simple global fits[27] to the data, in order to have a practical parameterization (needed, for example, in making radiative corrections to the data), and to study the possible Q^2 dependence of the first moments of g_1. The fits are of the general form $g_1/F_1 = ax^\alpha(1 + bx + cx^2)[1 + Cf(Q^2)]$ where a, α, b, c, and C are fit parameters and $f(Q^2)$ is defined to be either $1/Q^2$, or $\ln(1/Q^2)$. Cuts were applied to some of the fits to include only data with $Q^2 > 1$ (GeV/c)^2, and C was forced to be zero (no Q^2 dependence) for some fits. The results indicate that when all the data are included, the fits where $C \neq 0$ have significantly better χ^2 per degree of freedom than those where $C = 0$. However, good fits to the data are obtained when $C = 0$ and the $Q^2 > 1$ (GeV/c)^2 cut is applied to the data (fit II). Two of these global fits[27] are shown in Figs. 8 and 9:

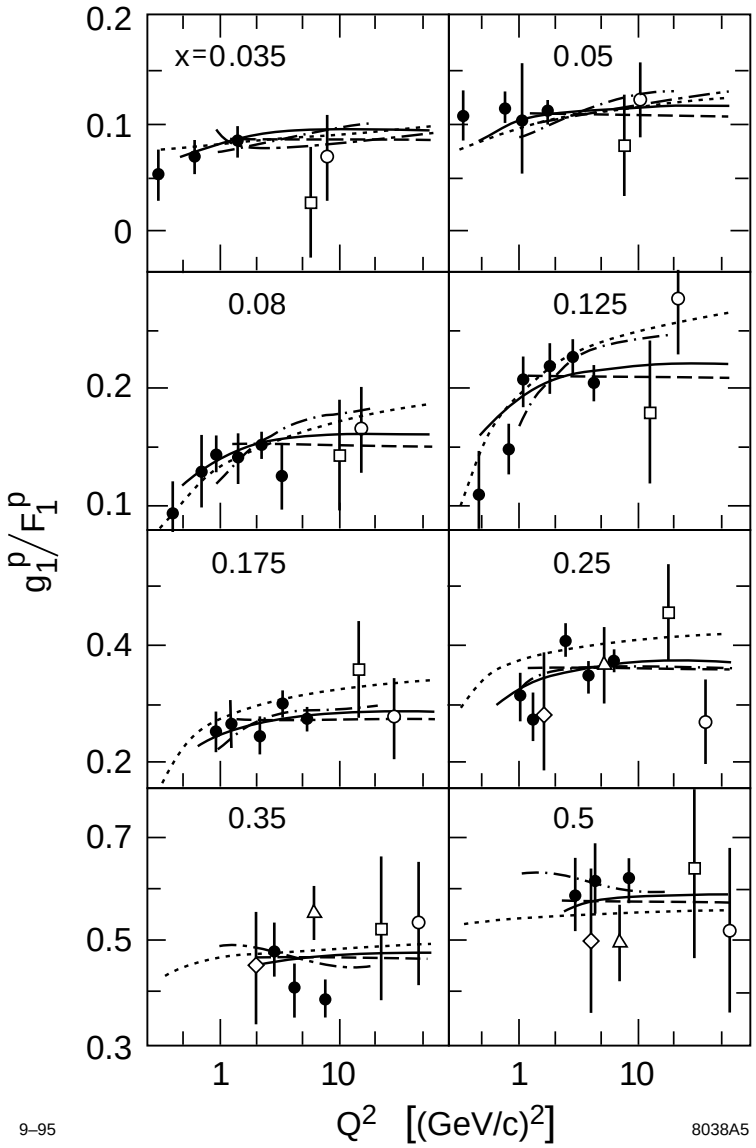


Figure 8: Ratios g_1^p/F_1^p extracted from experiments assuming $g_2 = g_2^{WW}$. The uncertainties are statistical only. Data are from E143[27] (solid circles), E80[16] (diamonds), E130[17] (triangles), EMC[18] (squares), and SMC[20] (open circles). The dashed and solid curves correspond to global fits[27] II (g_1^p/F_1^p Q^2-independent) and III (g_1^p/F_1^p Q^2-dependent), respectively. Representative NLO pQCD fits from Ref.[39] and Ref.[40] are shown as the dot-dashed and dotted curves respectively.

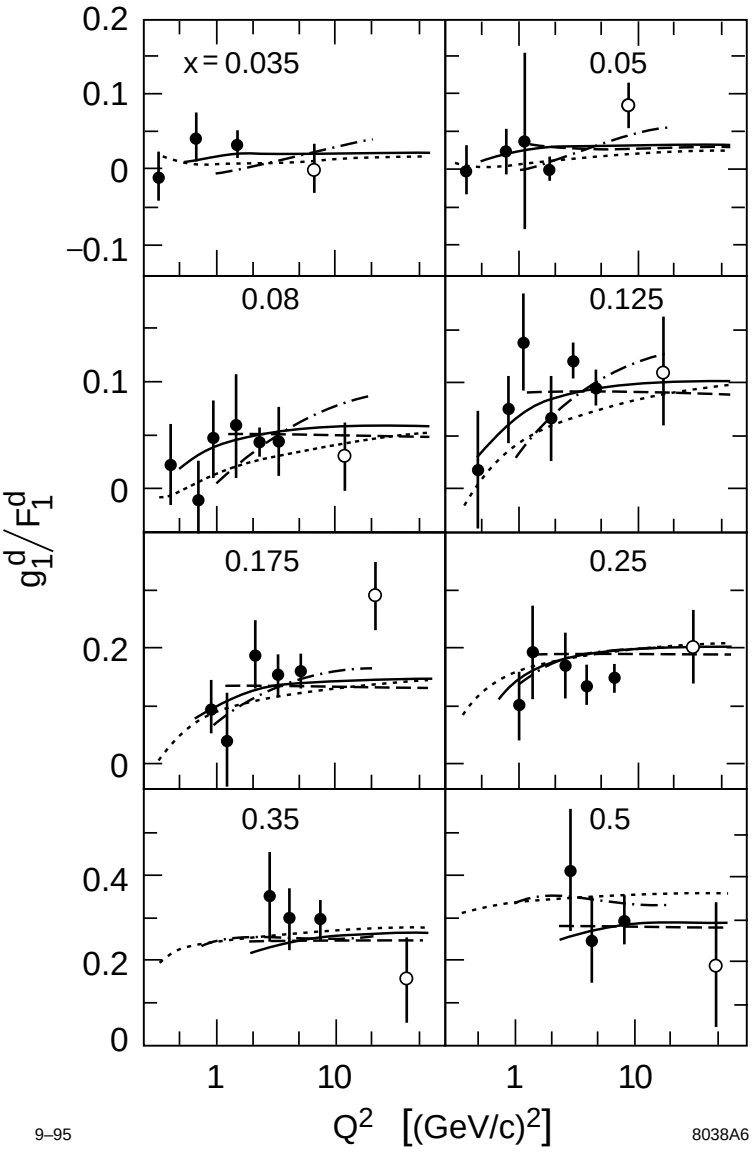


Figure 9: Ratios g_1^d/F_1^d from E143[27] (solid circles) and SMC[22] (open circles). The curves are as in Fig. 8

fit II and fit III which assumes $f(Q^2) = 1/Q^2$ and the data at all Q^2 are fit.

Also shown in Figs. 8 and 9 are representative global NLO pQCD fits[39,40] to available structure function data excluding those measured at the 9.7 GeV and 16.2 GeV beam energies of this experiment. These fits are indicated as the dot-dashed curves[39] and the dotted curves.[40] Both sets of predictions[39,40] indicate that g_1^p/F_1^p decreases with Q^2 at lower x, in agreement with the trend of our $E = 9.7$ and $E = 16.2$ results.

Another type of fit was made to the data which was motivated by possible differences in the twist-4 contributions to g_1 and F_1. We fit the data in each x bin (see Figs. 8 and 9) with the form $g_1/F_1 = a(1 + C/Q^2)$. The results for the C coefficients are shown in Fig. 10 for fits to all data (circles) and for fits to data with $Q^2 > 1$ (GeV/c)^2 (squares). The coefficients indicate significantly negative values for C at intermediate values of x when all the data are fit. The errors are much larger when data with $Q^2 < 1$ (GeV/c)^2 are excluded, and the results are consistent with no Q^2-dependence to g_1/F_1 ($C = 0$). The present data do not

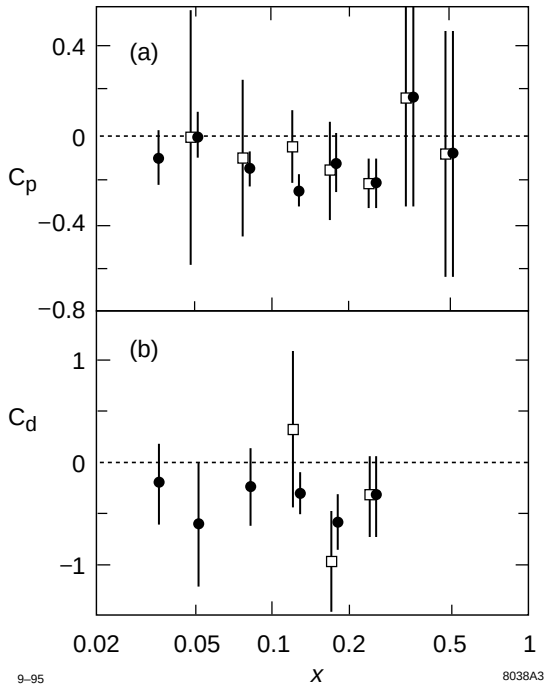


Figure 10: Coefficients C for fits to g_1/F_1 at fixed x of the form $a(1 + C/Q^2)$ for (a) proton and (b) deuteron. Solid circles are from fits to all data, and open squares are from fits to data with $Q^2 > 1$ (GeV/c)^2 only.

have sufficient precision to distinguish between a logarithmic and power law Q^2 dependence, but can rule out large differences between the Q^2-dependence of g_1 and F_1, especially for $Q^2 > 1$ (GeV/c)^2.

Using fits[27] II and III described above and a global fit[30,31] to F_1, we have evaluated the first moments Γ_1^p and Γ_1^d, and the corresponding results for $\Gamma_1^p - \Gamma_1^n$ and the net quark helicity Δq. The results for $\Gamma_1^p - \Gamma_1^n$ are shown as a function of Q^2 as the lower (fit II) and upper (fit III) bands in Fig. 11, where the width of the band reflects the combined statistical and systematic error estimate. Both fits are in reasonable agreement with the Bjorken sum rule (solid curve) evaluated using $\alpha_s(Q^2)$ evolved in Q^2 from $\alpha_s(M_Z) = 0.117 \pm 0.005$[8] for the QCD corrections.

Our results for Δq evaluated at $Q^2 = 3$ (GeV/c)^2 are shown in Table 7. Note that these results for Δq and for $\Gamma_1^p - \Gamma_1^n$ have shifted slightly from the

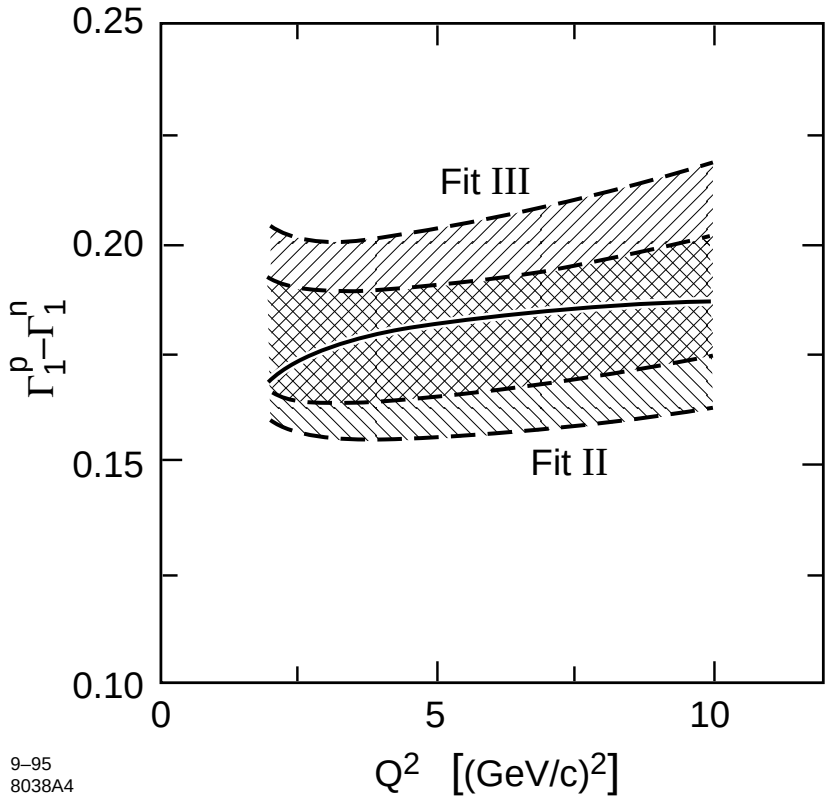


Figure 11: Evaluations of $\Gamma_1^p - \Gamma_1^n$ from the Q^2-independent fits II (lower band) and Q^2-dependent fits III (upper band). The errors include both statistical and systematic contributions and are indicated by the widths of the bands. The solid curve is the prediction of the Bjorken sum rule with third-order QCD corrections.

original results[24,25] at 29 GeV discussed above (See Tables 2 and 4) because of improved radiative corrections, the inclusion of additional data runs, and improved measurements of the beam and target polarizations. Using fits II or III makes little difference at $Q^2 = 3$ (GeV/c)^2, but we find Δq (which should be independent of Q^2) to vary less with Q^2 for fit III than for fit II, especially for the deuteron fits.

Table 7: Summary of extracted Δq results at $Q^2 = 3$ (GeV/c)^2 using fits II and III[27] for Γ_p and Γ_d.

Fit	Δq from Γ_p	Δq from Γ_d
II (Q^2-independent)	0.34 ± 0.09	0.35 ± 0.05
III (Q^2-dependent)	0.36 ± 0.10	0.34 ± 0.05

5 Summary

Measurements of $A_{||}$ have been made at beam energies of 29.1, 16.2, and 9.7 GeV and $A_\perp$ at a beam energy of 29.1 GeV for protons and deuterons. The spin structure functions, g_1 and g_2 have been extracted for the 29.1 GeV data. The integrals $\Gamma_p = \int_0^1 g_1^p(x,Q^2)dx$ and $\Gamma_d = \int_0^1 g_1^d(x,Q^2)dx$ have been evaluated at fixed $Q^2 = 3$ (GeV/c)^2. These results support the Bjorken sum rule predictions, and thus an important test of QCD is passed. The Ellis-Jaffe sum rule predictions for the proton and deuteron, however, are violated. In the context of the quark model, this implies that a non-negligible fraction of the proton helicity is carried by either strange quarks, gluons or both, and that the net quark helicity is smaller than expected. The Q^2 dependence of the ratio g_1/F_1 has been studied and is found to be small for $Q^2 > 1$ (GeV/c)^2.

Within experimental precision we find that the g_2 data are well-described by the twist-2 contribution, g_2^{WW}. Results for $\overline{g_2}$ are consistent with zero, although $\overline{g_2}$ about the same order of magnitude as g_2^{WW} are allowed within the statistical uncertainties. More precise data is needed in the future to provide a more stringent measurement of $\overline{g_2}$. Twist-3 OPE matrix elements have been extracted from the moments of g_1 and g_2. These results have a different sign than the QCD sum rule predictions, although within errors these predictions cannot be ruled out. The asymmetry A_2 has also been measured and is found to be significantly smaller than the positivity limit $\sqrt{R}$ for both targets. A_2^p is found to be positive and inconsistent with zero.

A number of experimental programs will produce new spin structure function measurements in the future. SMC is continuing to take data. Additional results are expected from SLAC using a 50 GeV incident electron beam. Measurements of the neutron spin structure functions are in progress (E154), and proton and deuterium spin structure functions measurements (E155) will be made in 1996. Also, the HERMES collaboration at HERA is currently measuring spin-dependent structure functions of the proton and neutron. The data from these experiments will improve our understanding of the nucleon spin structure, and should answer many questions that have arisen due to current experimental results.

References

[1] F. E. Close and R. G. Roberts, Phys. Lett. B316 (1993), 165.

[2] M. Anselmino, A. Efremov, and E. Leader, Phys. Rep. **261**,1 (1995).

[3] L. Mankiewicz and Z. Rysak, Phys. Rev. **D43**, 733 (1991).

[4] J. L. Cortes, B. Pire and J. P. Ralston, Z. Phys. **C55**, 409 (1992).

[5] R. L. Jaffe and X. Ji, Phys. Rev. **D43**, 724 (1991).

[6] S. Wandzura and F. Wilczek, Phys. Lett. **B72**, 195 (1977).

[7] J. D. Bjorken, Phys. Rev. **148** 1467 (1966); Phys. Rev. D **1** 1376 (1970).

[8] Particle Data Group, L. Montanet, *et al.,* Phys. Rev. **D50** 1173 (1994).

[9] S. A. Larin and J . A. M. Vermaseren, Phys. Lett. **B259** 345 (1991) and references therein.

[10] J. Ellis and. R. Jaffe, Phys. Rev. D **9**, 1444 (1974); Phys. Rev. D **10** 1669 (1974).

[11] F. E. Close and R. G. Roberts, Phys. Lett. **B316**, 165 (1993).

[12] S. A. Larin, Phys. Lett. **B334**, 192 (1994).

[13] E. V. Shuryak and A. I. Vainshtein, Nucl. Phys. **B201**, 141 (1982).

[14] H. Burkhardt and W. N. Cottingham, Ann. Phys. **56**, 453 (1970).

[15] A. V. Efremov, and O. V. Teryaev, Phys. Lett. **B200**, 363 (1988).

[16] E80, M. Alguard, *et al.*, Phys. Rev. Lett. **37**, 1261 (1976); Phys. Rev. Lett. **41**, 70 (1978).

[17] E130, G. Baum, *et al.*, Phys. Rev. Lett. **51**, 1135 (1983).

[18] EMC, J. Ashman *et al.*, Nucl. Phys. **B328**, 1 (1989).

[19] SMC, B. Adeva, *et al.*, Phys. Lett. **B302**, 533 (1993).

[20] SMC, D. Adams, *et al.*, Phys. Lett. **B329**, 399 (1994).

[21] SMC, D. Adams, *et al.*, Phys. Lett. **B336**, 125 (1994).

[22] SMC, D. Adams, *et al.*, Phys. Rev. Lett. **B357**, 248 (1995).

[23] E142, P.L. Anthony *et al,* Phys. Rev. Lett. **71** 959 (1993).

[24] E143, K. Abe *et al.*, Phys. Rev. Lett. **74**, 346 (1995).

[25] E143, K. Abe *et al.*, Phys. Rev. Lett. **75**, 25 (1995).

[26] E143 K. Abe *et al.,* SLAC-PUB-95-6982 (1995), submitted to Phys. Rev. Lett.

[27] E143 K. Abe *et al.,* SLAC-PUB-95-6997 (1995), accepted by Phys. Lett. B.

[28] T. V. Kukhto and N. M. Shumeiko, Nucl. Phys. **B219**, 412 (1983); I. V. Akusevich and N. M. Shumeiko, J. Phys. **G20**, 513 (1994).

[29] Y. S. Tsai, Report No. SLAC-PUB-848 (1971); Y. S. Tsai, Rev. Mod. Phys. **46**, 815 (1974).

[30] NMC, P. Amaudruz *et al.,* Phys. Lett. **B295**, 159 (1992).

[31] L. W. Whitlow *et al.,* Phys. Lett. **B250**, 193 (1990).

[32] R. L. Heimann, Nucl. Phys. **B64**, 429 (1973).

[33] S. D. Bass and P. V. Landshoff, Report No. DAMTP 94/50 (1994); F. E. Close and R. G. Roberts, Report No. RAL-94-071 (1994).

[34] M. Stratmann, Z. Phys. **C60**, 763 (1993), and private communication for values at $Q^2 = 5$ (GeV/c)^2.

[35] X. Song and J. S. McCarthy, Phys. Rev. **D49**, 3169 (1994), **D50**, 4718, (1994) (Erratum), X. Song, INPP-UVA-95/04 (1995).

[36] E. Stein *et al.*, Phys. Lett. **B334**, 369 (1995).

[37] I. I. Balitsky, V. M. Braun and A. V. Kolesnichenko, Phys. Lett. **B242**, 245 (1990); **B318**, 648 (1993) (Erratum).

[38] G. Altarelli and G. Parisi, Nucl. Phys. B126 (1977), 298; V. N. Gribov and L. N. Lipatov, Yad. Fiz. 15, 781 (1972), [Sov. J. Nucl. Phys. 15, 438 (1972)].

[39] R. D. Ball, S. Forte, and G. Ridolfi, Nucl. Phys. B444 (1995), 287; and CERN-TH/95-266 (1995).

[40] M. Glück, E. Reya, M. Stratmann, and W. Vogelsang, Report No. DO-TH-95/13 and RAL-TR-95-042, August 1995.

[41] E. B. Zijlstra and W. L. van Neerven, Nucl. Phys. B417 (1994), 61; R. Mertig and W. L. van Neervan, Univ. Leiden INLO-PUB-6/95 and NIKHEF-H/95-031, June, 1995.

[42] S. J. Brodsky, M. Burkardt, and I. Schmidt, Nucl. Phys. B441 (1995), 197.

[43] G. Altarelli, P. Nason, and G. Ridolfi, Phys. Lett. B320 (1994), 152.

[44] T. Gehrmann and W. J. Stirling, Z. Phys. C65 (1995), 461.

[45] B. Ehrnsperger and A. Schäfer, Report No. UFTP/370/1994 (1994).

QCD AT SLD

Richard Plano*
Rutgers University
New Brunswick, NJ 08903
Representing the SLD Collaboration

ABSTRACT

This talk reviews seven SLD papers, which provide useful, unique, and precise contributions to our understanding of hadron production in the decay of Z^0's produced in e^+e^- collisions. The data were gathered by the SLAC Large Detector (SLD) at the SLAC Linear Collider (SLC). This accelerator/detector is able to compete with LEP in many studies, in spite of smaller statistics, due to systematic advantages including a tiny and stable interaction region combined with a precise high-resolution vertex detector, excellent particle identification, and a highly polarized electron beam.

The papers include studies of:

- factorial and cumulant moments,[1]
- charged multiplicities produced by b, c, and uds quarks,[2]
- rapidity gaps,[3]
- orientations and energy partitions of three-jet events,[4]
- jet handedness,[5]
- triple-product correlation in polarized Z^0decays to three jets,[6] and
- $\pi^{\pm}$, $K^{\pm}$, p, K^0, and Λ production in Z^0 decays.[7]

Comparisons are made to LEP results where appropriate.

*Supported by NSF Grant PHY-9510439.

©1995 by Richard Plano.

1 Introduction

Quantum Chromodynamics (QCD), the local non-Abelian gauge field theory of quarks and gluons, is now reasonably well-understood and capable of making many testable predictions. The reaction which is discussed here,

$$e^+e^- \to Z^0 \to q\bar{q} \;\; (g)\cdots \to hadrons, \;\; e.g., \;\; \pi K p \rho \phi \Delta \Lambda b \cdots \quad , \tag{1}$$

is illustrated schematically in Fig. 1. The production of $q\bar{q}$ by e^+e^- annihilation via Z^0 exchange is well-understood; γ exchange is almost negligible—about 1100 times less probable than Z^0 exchange. The resulting "parton showers" [Fig. 1 (i)] as gluons and further $q\bar{q}$ pairs are produced: can be calculated using perturbative QCD (PQCD), for example, in the Modified Leading Logarithm Approximation. Next [Fig. 1 (ii)], the partons hadronize in a manner which is not fully understood. The assumption of Local Parton-Hadron Duality (LPHD), which states that the distributions of final state particles are related in some detail to the original partons, is the source of numerous predictions, some of which will be tested in this paper. The physical hadrons [Fig. 1 (iii)] produced by the hadronization process are often unstable and decay into the final state particles [Fig. 1 (iv)] which are observed by the detector. While these decays are well-understood, they add a considerable complication to the analysis.

Numerous properties of the final state particles can be measured and used to provide information about the reaction that produced them. The SLD at the SLC is well-equipped to contribute to these investigations.

2 Features of SLD

A cross section of one quarter of the SLD detector is shown in Fig. 2 and an isometric view in Fig. 3. The characteristics of this detector which are most important for the results presented here will be briefly described. See the original design report[8] for details.

The position of the Interaction Point (IP) at which the e^+e^- collision takes place is constrained by the beam size to lie within a region roughly 0.8 μm vertically, 2.6 μm horizontally ("flat" beams), and 700 μm along the beam. The transverse position of the collision region is quite stable, with variations typically less than 10 μm over periods of a few hours. To determine the transverse position

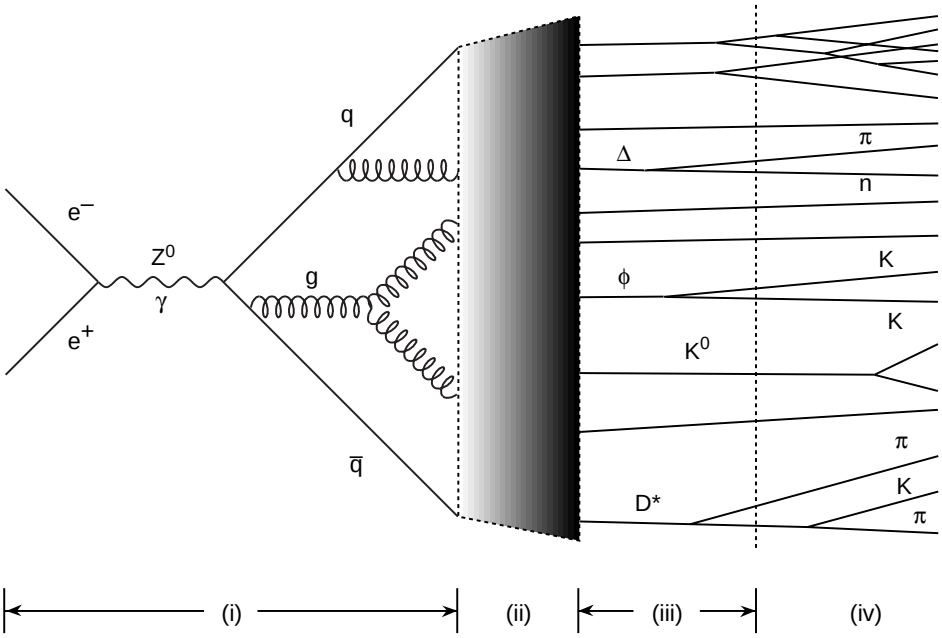


Fig. 1. QCD description of $e^+e^- \to Z^0 \to q\bar{q}g \cdots \to$ hadrons.

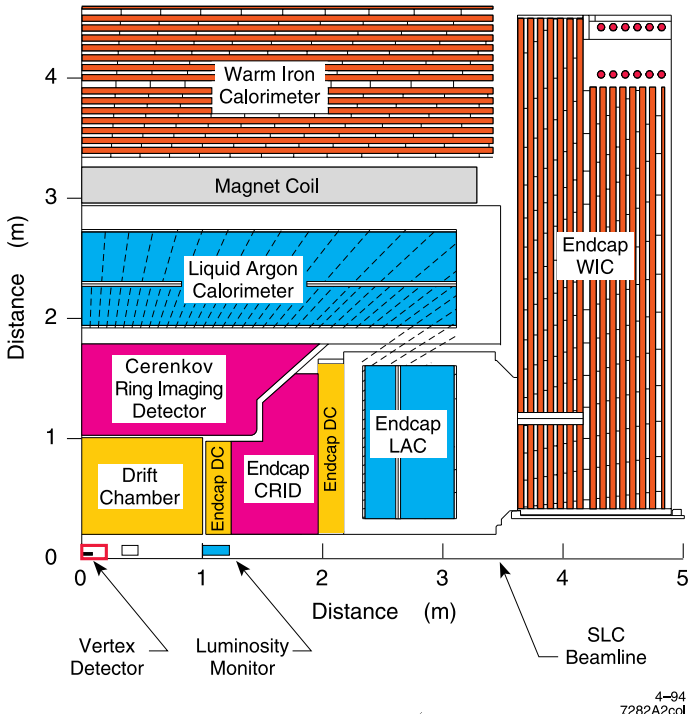


Fig. 2. The SLD detector (quadrant view).

accurately, the intersection point of tracks from hadronic events is averaged over about 30 successive hadronic events. Poor quality tracks are discarded, leaving about 300 tracks in each fit, constraining the IP to less than 10 μm.

The vertex detector (VXD) consists of 120 million pixels, each 22 μm $\times$ 22 μm, placed in three layers between 2.9 cm and 4.2 cm radius around the beamline. The VXD has a point resolution of about 5 μm in the plane perpendicular to the beamline.

The Central Drift Chamber (CDC) has 80 layers of sense wires roughly parallel to the beamline with 48 layers which are at angles of $\pm$ 50 mrad to provide stereo information. The intrinsic accuracy of a hit in the CDC, which tells the distance of a track from a wire, is 85 μm near the center of a drift cell.

Combining the VXD and the CDC provides an impact parameter resolution (distance of closest approach of a track to the IP) of (p_t in GeV/c):

$$\sigma_{r\phi} = \frac{76\mu m}{p_t\sqrt{sin\theta}} \bigoplus 11\mu m \tag{2}$$

and, using the magnetic field of 6 kG parallel to the beam, a momentum resolution of:

$$\frac{\sigma_{p_t}}{p_t} \sim 0.5\% p_t \bigoplus 1.0\%. \tag{3}$$

The Cherenkov Ring Imaging Detector (CRID) uses Cherenkov radiation to measure the speed of charged particles which, when combined with a momentum determination, identifies the particle. The CRID[9] will be described in detail later in this talk.

The Liquid Argon Calorimeter (LAC) and Warm Iron Calorimeter (WIC) together serve both as an electromagnetic calorimeter with an energy resolution of 15% at 1 GeV and an angular resolution of a few milliradians, and a hadronic calorimeter with an energy resolution of 55% at 1 GeV.

Rather than describe in detail the criteria used for event triggers, track selection, and event selection for each analysis, typical values are listed here; in no case do the results depend sensitively on the details.

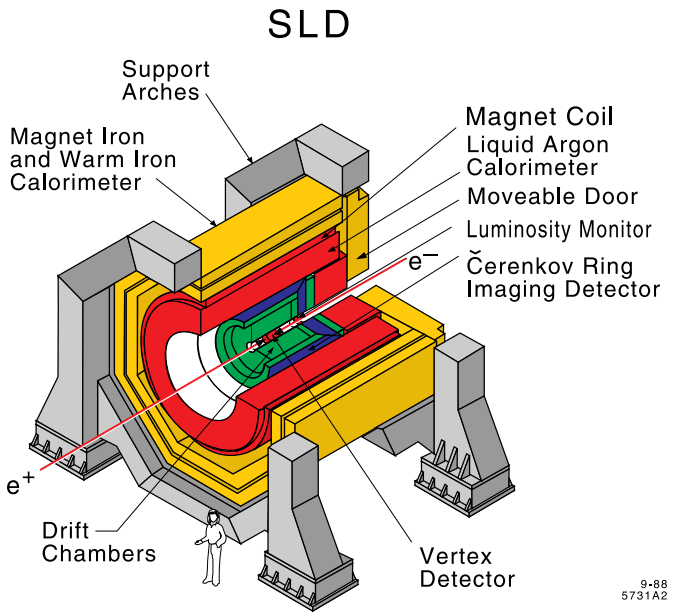


Fig. 3. The SLD detector (isometric view). The end caps have been removed for clarity.

Typical Event Trigger: an event is recorded if it passes any of the following criteria:

1. $E_{LAC} > 12$ GeV,
2. ≥ 2 tracks reconstructed in the CDC, using only the cell locations, and
3. $E_{LAC} > 4$ GeV and ≥ 1 CDC track.

Once a trigger is selected, tracks are reconstructed in the CDC and VXD. A track is selected if it passes all of:

1. $DOCA_{\perp} < 5$ cm, $DOCA_{\parallel} < 10$ cm, $\perp \equiv$ perpendicular to the beam,
2. $\mid cos\theta_{trk} \mid < 0.80$ $(37^{\circ} < \theta_{trk} < 143^{\circ})$, and
3. $p_{\perp} = p \sin\theta_{trk} > 0.15$ GeV/c.

$DOCA$ is the Distance Of Closest Approach of the extrapolated track to the IP. θ_{trk} is the track angle relative to the beam. An event is selected for further analysis if the following criteria are all met:

1. ≥ 5 selected tracks,
2. $\mid cos\theta_T \mid < 0.71 \quad (44.8^{\circ} < \theta_T < 135.2^{\circ})$, and
3. $E_{vis} > 20$ GeV, using the tracks selected above.

θ_T is the angle between the thrust axis of the event and the electron beam direction; E_{vis} is the energy calculated from the momentum measured in the CDC assuming the pion mass for all tracks.

3 Charged Multiplicity

The first two papers[1,2] concern different aspects of charged multiplicity, the detailed shape of the overall charged multiplicity distribution, and the dependence of average multiplicities on the flavor of the primary $q\overline{q}$ pair.

Multiplicity distributions contain fundamental information about interactions and have often been studied.[10] The first guess—a Poisson distribution based on stochastic processes—works very badly, as it predicts that the fractional width should decrease as the multiplicity increases. "KNO scaling,"[11] which postulates that the multiplicity distribution is a universal function of $n/ < n >$ and so predicts a constant fractional width, does much better, but also fails at the highest energies.

"Factorial and cumulant moments in $e^+e^- \to$ hadrons at the Z^0 resonance"[1] follows a proposal of I. M. Dremin[12] to make a sensitive test of the shape of the overall multiplicity distribution. This is the first study of H_q, the ratio of the cumulant moment K_q to the factorial moment F_q:

$$H_q = K_q/F_q; \qquad F_q = \frac{< n(n-1)\cdots(n-q+1) >}{< n >^q} = \sum_{m=0}^{q-1} C_{q-1}^m K_{q-m} F_m. \qquad (4)$$

C_{q-1}^m are the binomial coefficients, and $F_0 = F_1 = K_1 = 1$. H_q is predicted to have the following behavior for the indicated models of particle production.

Poisson	NBD	DLA	NLA	NNLA
0	q^{-k} $k \approx 24$	q^{-2}	Minimum at $q \approx 5$	Mininum at $q \approx 5$, then oscillates

In this table, NBD is the Negative Binomial Distribution, DLA refers to a QCD calculation done in leading order (Double Logarithmic Approximation), and NLA, NNLA refer to Next-to-Leading and Next-to-Next-to Leading Approximation QCD calculations. These calculations[12,14] illustrate the striking sensitivity of H_q to models whose predictions look quite similar when plotted as simple multiplicity distributions.

Monte Carlo (MC) calculations, including a detailed description of the SLD detector, were used to correct for effects introduced by geometrical acceptance and resolution, γ conversions, and particle interactions. Charged decay products of particles with lifetimes $< 3 \times 10^{-10}$ s were included in the multiplicity count.

Figure 4 shows the data with a fit to a NBD distribution. Although the fit is quite good ($k = 24.9 \pm 0.9, < n > = 20.6 \pm 0.1, \chi^2/NDF = 22.9/24$), note that all the measured points for $n \geq 36$ are **above** the fitted curve.

Figure 5 shows the H_q obtained from these data (points with error bars). Also shown are lines connecting H_q values calculated from the Poisson (dots) and NBD (dashes) distributions fitted to the multiplicity distribution. The Poisson distribution, which is identically zero, is unacceptable at $q \leq 3$ and, as is especially clear from the magnified view in the inset, NBD is also very poor. Interestingly, an analysis[14] of the multiplicity distribution observed by UA5 in $\overline{p}p$ interactions at 546 GeV shows very similar behavior of H_q, including the oscillations. MC

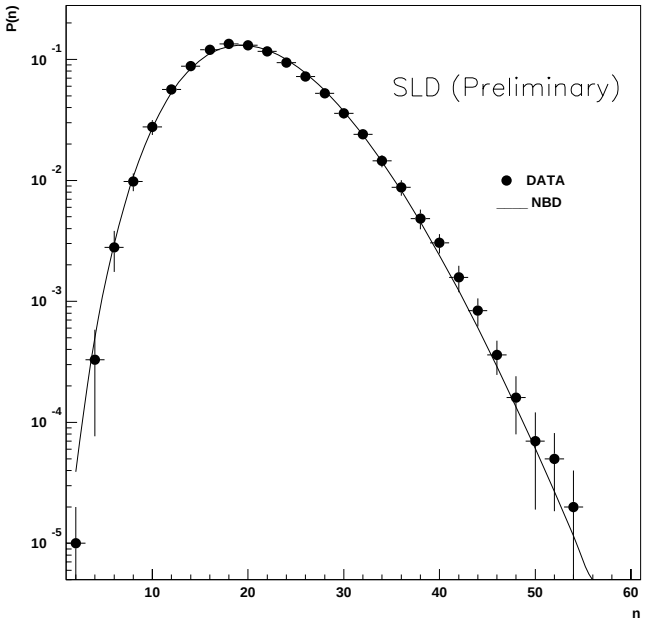


Fig. 4. Negative Binomial Distribution fit to the multiplicity distribution.

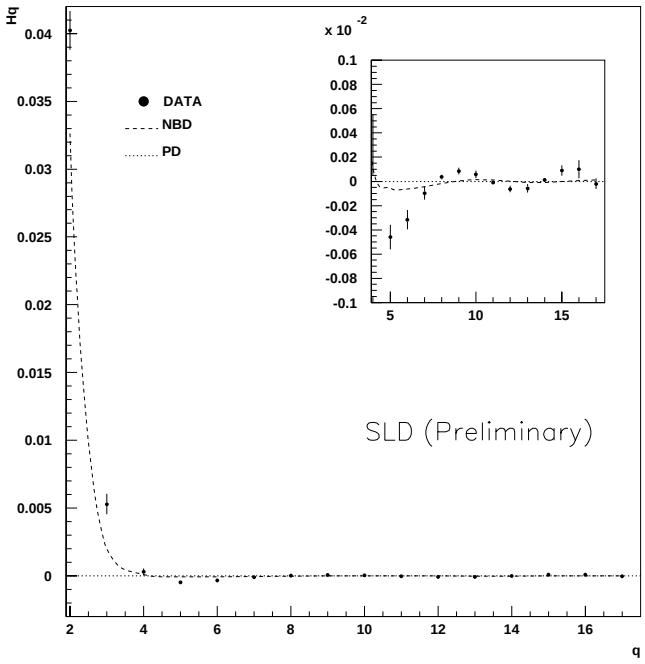


Fig. 5. H_q fit to data compared to NBD and Poisson predictions. Note enlarged view in inset.

calculations of the SLD data have shown that the oscillations in H_q are not related to the effective truncation at high multiplicity due to the finite size of the data sample.

Higher order calculations,[13] including the NNLA, describe the shape of the H_q distribution rather well. These results tend to support both the validity of QCD at the parton level and the concept of LPHD—the idea that the observable final state distributions reflect in some detail the distributions of the underlying partons.

"Measurement of the charged multiplicities of b, c, and light quark events from Z^0 events"[2] tests how the **average** multiplicities depend on flavor.

To obtain uds, b, and c quark enriched samples without bias, divide each event into two hemispheres by a plane perpendicular to the thrust axis. Then:

- Tag each hemisphere as described below.

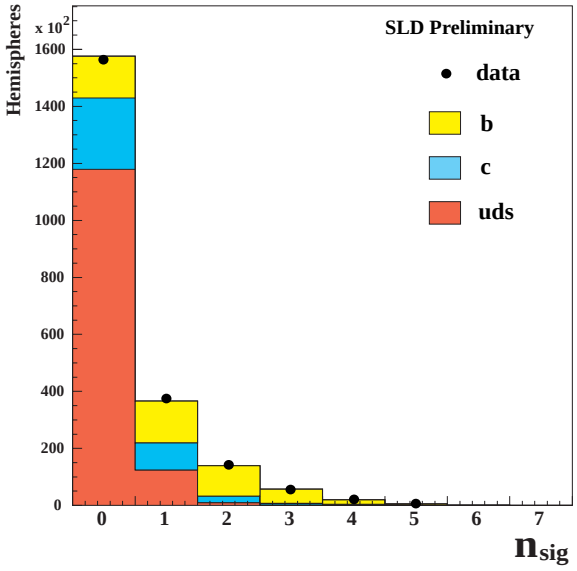


Fig. 6. MC estimates of the uds, b, and c components of hemispheres *vs.* n_{sig} and data *vs.* n_{sig}.

- Use the tag to label the **opposite** hemisphere; this is necessary to avoid bias.
- Double the count in the labeled hemisphere.

uds and b tags are based on n_{sig}, the number of tracks in a hemisphere with significant impact parameters, b, where $b/\sigma_b > 3$. The c quark tag is based on detection of a D^+ or D^{*+}. The tags, the resulting number of hemispheres tagged, and the purity of the tags are summarized in the table.

Quark	Tag	Purity	# Hemispheres
uds	$n_{sig} = 0$	75%	156K
b	$n_{sig} \geq 3$	94%	8.5K
c	$D^+ or D^{*+}$	58%	1.2K

Figure 6 shows the number of hemispheres containing n_{sig} significant tracks in the data. The shaded bars indicate MC estimates of the quark content of each n_{sig} bin. The MC is in good agreement with the data and gives estimates of the purity provided by the n_{sig} cuts.

The counts in the labeled hemispheres are then used to form $\overline{n}_{uds}$, $\overline{n}_b$, and $\overline{n}_c$, as well as $\delta\overline{n}_b = \overline{n}_b - \overline{n}_{uds}$ and $\delta\overline{n}_c = \overline{n}_c - \overline{n}_{uds}$.

Perturbative QCD (PQCD) in the modified leading logarithm approximation (MLLA) predicts[15] that $\delta\overline{n}_b = \overline{n}_b - \overline{n}_{uds}$ and $\delta\overline{n}_c = \overline{n}_c - \overline{n}_{uds}$ are independent

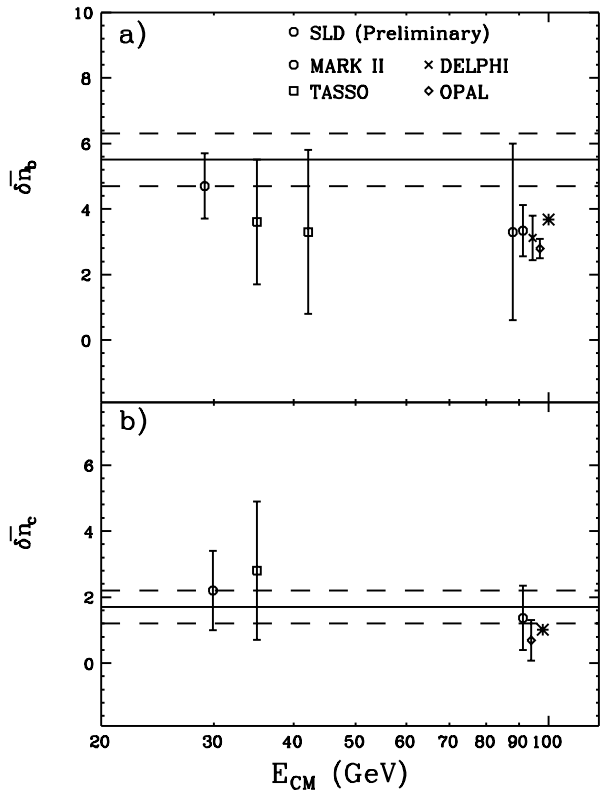


Fig. 7. $\delta\overline{n}_b = \overline{n}_b - \overline{n}_{uds}$ and $\delta\overline{n}_c = \overline{n}_c - \overline{n}_{uds}$ as a function of the center-of-mass energy. The solid lines are MLLA QCD + LPHD predictions with uncertainties indicated by the dashed lines. The stars indicate another prediction (see text).

of energy within about 0.1 track. Furthermore, MLLA QCD + LPHD[15] and an alternative approach[16] have made definite predictions which are summarized in the table and in Fig. 7.

Ref.	These data	Ref. 15	Ref. 16
$\delta\overline{n}_b$	3.34 ± 0.13	5.5 ± 0.8	3.68
$\delta\overline{n}_c$	1.37 ± 0.45	1.7 ± 0.5	1.01

The data are in reasonable agreement with the predictions within the rather large errors. As shown in Fig. 7, DELPHI and OPAL are in good agreement with the SLD numbers. Furthermore, data at lower energies from MARK II and TASSO are consistent with negligible energy dependence as predicted.

4 A Study of Rapidity Gaps

Events containing large rapidity gaps have been reported at HERA and at Fermilab, where they have been interpreted as evidence for the Pomeron.[17,18] Hadronic events from e^+e^- annihilation containing large rapidity gaps produced by a color-screening mechanism are expected to occur at a very small rate.[19] This is the **first** study of rapidity gaps in e^+e^- annihilations, and it will be useful for the analysis of the Fermilab/HERA data, since the Pomeron is not expected in e^+e^- annihilations.

Figure 8 summarizes the results obtained in this measurement.[3] Figure 8(a) shows the inclusive rapidity distribution of charged tracks in hadronic events, Fig. 8(b) the average gap (average difference in rapidity between neighboring charged tracks), and Fig. 8(c) the distribution of the largest gap in each event. As can be seen most clearly in Fig. 8(c), there are many more large gaps than are predicted by the hadronic MC. However, when contamination from $\tau^+\tau^-$ events in the hadronic sample are included in the MC, the disagreement disappears.

This data then provides natural spectra for gap distributions which are needed for the analysis of gap distributions found in hadronic events.

5 Orientation and Energy Partition of Three-Jet Events

Three-jet events have been studied since 1979 (Ref. 20) when they were used to support the existence of a spin 1 gluon produced in the reaction $e^+e^- \to Z^0 \to q\overline{q}g$. Similar events measured in the SLD detector have been analyzed in terms of the energies and orientation angles of the three jets.[4]

The Liquid Argon Calorimeter (LAC), which covers 98% of the solid angle, was used for this analysis as it measures the jet energies and angles with good precision. The events are analyzed by ordering the three jets by their energy in the overall center of mass: $E_1 > E_2 > E_3$. Then the Ellis-Karliner angle is formed:

$$\cos\theta_{EK} = \frac{x_2 - x_3}{x_1}; \qquad x_i = 2E_i/\sqrt{s}. \tag{5}$$

Distributions of these quantities are shown in Fig. 9. These plots show the data, fully corrected to the parton level, as dots and compare the data with

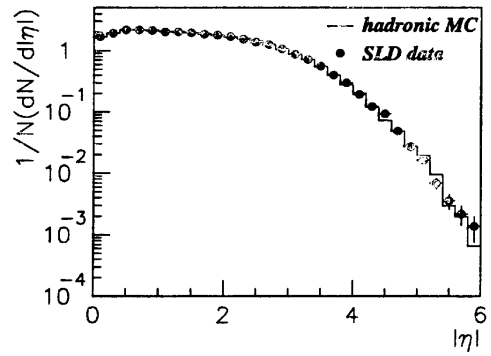


Fig 1a

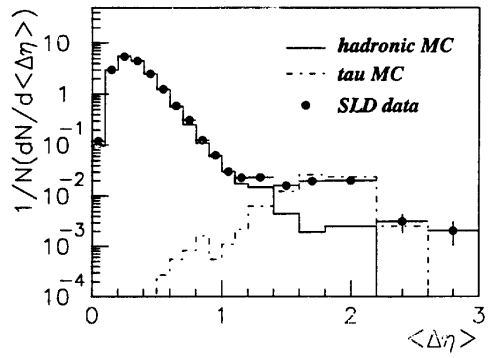


Fig 1b

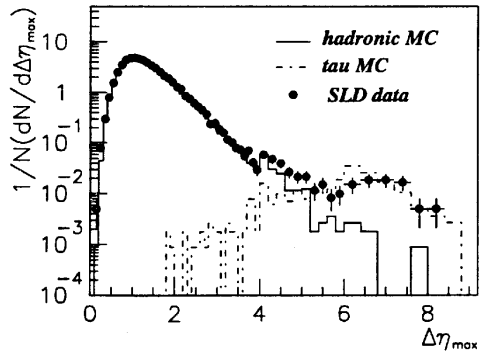


Fig 1c

Fig. 8. The (a) rapidity distribution, (b) average gap distribution, and (c) maximum gap distribution. A hadronic MC with a 0.3% $\tau^+\tau^-$ contamination provides an excellent description.

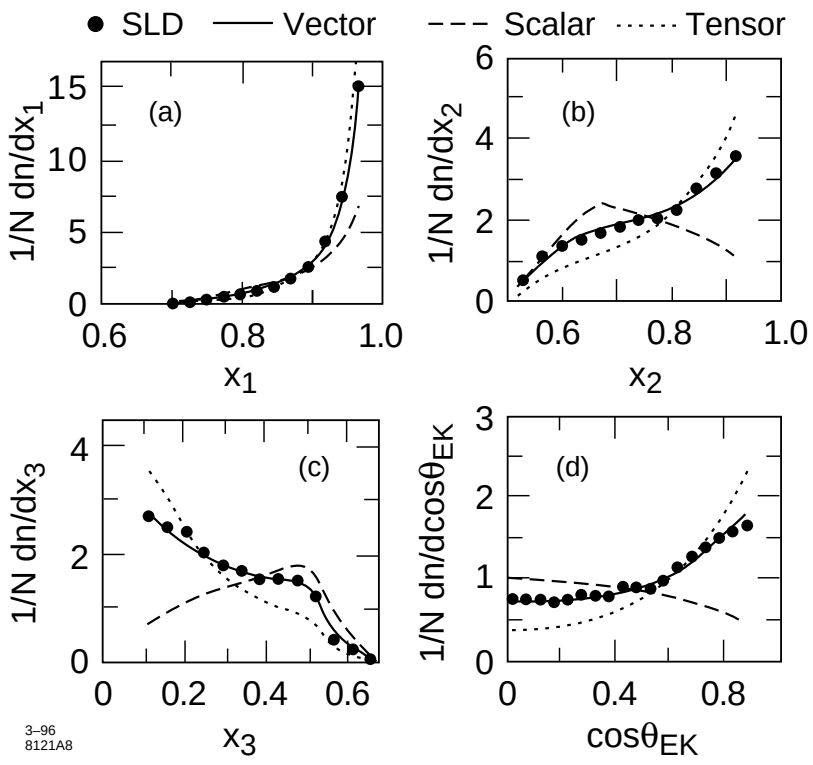


Fig. 9. Fully corrected data from this experiment are shown as dots; leading-order predictions are shown as solid lines for the vector case, dashes for scalar, and dots for tensor.

leading-order predictions assuming vector, scalar, and tensor gluons. The agreement of the data with the predictions for a vector gluon is in all cases satisfactory. Scalar and tensor fits are very poor and appear to be excluded, but precise limits are not feasible without $O(\alpha_s^3)$ calculations.

6 Jet Handedness

The polarized beam carries with it an obvious handedness. It is not unreasonable to expect that the high polarization of the electron beam might carry over into measurable final state effects. Since there is no complete theory of hadronization and very few measurements of any spin-dependent quantities, this is a fertile area for experimental studies.

The final three papers reported here[5–7] address three specific spin-dependent questions:

- Jet handedness. $\Omega = \vec{t} \cdot (\vec{k}_1 \times \vec{k}_2)$ to measure spin along the jet axis, $\vec{t}$, using the momenta of two particles in the jet, chosen to be sensitive to handedness.
- The triple product $\vec{S}_Z \cdot (\vec{k}_1 \times \vec{k}_2)$, where $\vec{k}_1$, $\vec{k}_2$ are the momenta of the highest and second-highest energy jets.
- A search for a leading particle effect to observe $\vec{S}_Z \cdot \vec{p}$, where $\vec{p}$ is the momentum of a particle composed of quarks.

SLC/SLD is an ideal place to study asymmetries in Z^0 decay as the beam polarization assures that the partons from Z^0 decay are highly longitudinally polarized. If techniques to observe this polarization could be developed, the spin dynamics of a variety of hard processes could be studied.

The polarization of the Z^0, A_Z, is related to the polarization of the beam electron, P_{e^-}, by:

$$A_Z = \frac{P_{e^-} - A_e}{1 - P_{e^-} A_e}; \qquad A_f = \frac{2 v_f a_f}{v_f^2 + a_f^2}, \tag{6}$$

where v_f, a_f are the vector and axial vector couplings of fermion f to the Z^0; $A_e \approx 0.16$. The following table shows how the large beam polarization at the SLC produces large forward-backward asymmetries A_{FB} for both u and d quarks, as well as highly polarized forward/backward quarks.

P_{e^-}	A_Z	A_{FB}(u-type)	A_{FB}(d-type)	$P_q(\cos\theta = \pm 1)$
+0.77	+0.70	-0.35	-0.49	$\pm$ 0.70
-0.77	-0.83	+0.42	+0.58	$\mp$ 0.83

By contrast, the Z^0 polarization at LEP, where the beams are unpolarized, is only 0.16. So the SLC errors, both statistical and systematic, are reduced by a factor close to five. This reduction in the statistical error could be compensated at LEP by an increase in the number of events by a factor of about 25; it is not possible to reduce the systematic errors by any simple technique.

Further, the ability to reverse the sign of the polarization is an extremely important aid in the further reduction of systematic errors at SLAC. It can be used to prove that an observed asymmetry is not due to an asymmetry in the detector or in the analysis.

"An improved limit on jet handedness in Z^0 decays"[5] doubles the sensitivity obtained with the 1993 run of the SLD detector.[21] The increased accuracy is due

both to the increased statistics and the higher polarization, which increased, on average, from 0.63 to 0.78, equivalent to a 50% increase in statistics.

Two observables were chosen to study polarization along a jet axis, one based on the magnitude of the momenta and the other on the charge of the particles. $\Omega_{hel} = \hat{t} \cdot (\vec{k_1} \times \vec{k_2})$, where $k_1 > k_2$ and $\hat{t}$ is pointing along the jet axis, has the same transformation properties as the "helicity-based" polarization:

$$P^f_{hel}(\cos\theta) \equiv \frac{\sigma^f_R + \sigma^{\overline{f}}_R - \sigma^f_L - \sigma^{\overline{f}}_L}{\sigma^f_R + \sigma^{\overline{f}}_R + \sigma^f_L + \sigma^{\overline{f}}_L} = 2\frac{A_Z \cos\theta}{1+\cos^2\theta}. \tag{7}$$

$\Omega_{chi} = \hat{t} \cdot (\vec{k_+} \times \vec{k_-})$ has the transformation properties of the "chirality-based" polarization:

$$P^f_{chi} \equiv \frac{\sigma^f_R - \sigma^{\overline{f}}_R - \sigma^f_L + \sigma^{\overline{f}}_L}{\sigma^f_R + \sigma^{\overline{f}}_R + \sigma^f_L + \sigma^{\overline{f}}_L} = -A_f \tag{8}$$

which is independent of $\cos\theta$ and beam polarization. Note that the beam polarization at SLC/SLD allows examination of helicity as well as chirality-based asymmetries; LEP can study only chirality-based asymmetries.

The measure of interest is then:

$$H = \alpha P = \frac{N_{\Omega<0} - N_{\Omega>0}}{N_{\Omega<0} + N_{\Omega>0}} \quad ; \tag{9}$$

H is the jet handedness, shown in Fig. 10. P is the expected polarization of the underlying partons, and α is the analyzing power of the method. The data was broken up into light-flavor jets and heavy-flavor jets using the number of nonzero impact parameters in the opposite hemisphere as described earlier with results given in the table. The numbers in parentheses are the 95% confidence upper limits on the magnitude of the analyzing power.

	Analyzing Power		
Analysis	Light-Flavor Jets	Heavy-Flavor Jets	All Jets
Helicity	$+0.005 \pm 0.017$ (.037)	$+0.025 \pm 0.019$ (.056)	-0.006 ± 0.022 (.047)
Chirality	$+0.017 \pm 0.026$ (.062)	$+0.014 \pm 0.013$ (.035)	$+0.005 \pm 0.017$ (.036)

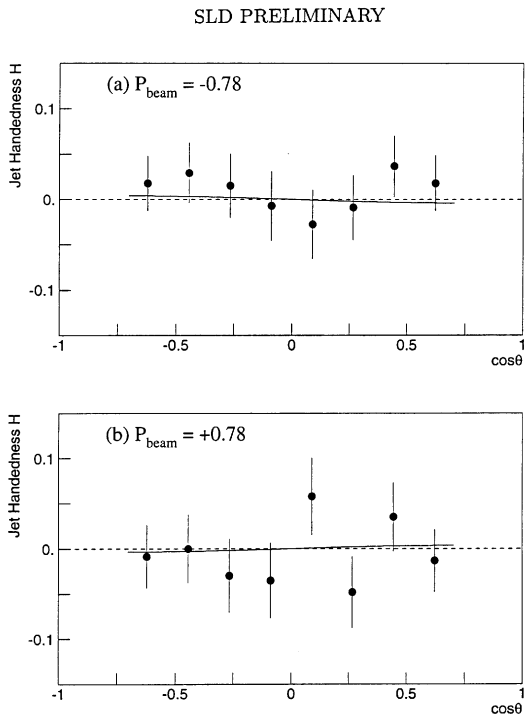


Fig. 10. Simultaneous fits to light-flavor jets for both signs of the electron beam polarization.

7 Triple-Product Correlations in Polarized Z^0 Decays

This is the first measurement of the triple-product correlation in polarized Z^0 decays into three hadronic jets.[6] The CP-even and T-odd triple product $\cos\omega \propto \vec{S_Z} \cdot (\vec{k_1} \times \vec{k_2})$ is sensitive to physics beyond the Standard Model, which predicts unobservably small values. Here, $\vec{S_Z}$ is the Z^0 polarization vector and $\vec{k_1}$, $\vec{k_2}$ are the largest and second-largest jet momenta. $\vec{S_Z}$ is in the direction of the electron beam polarization.

The variable $\cos\omega$ is even under C and P reversals, and odd under T_N, where T_N reverses momenta and spin vectors without exchanging initial and final states. Since that is not a true time-reversal operation, a nonzero value is possible without CPT violation. Though this variable was proposed[22] in 1980 as a means of observing the non-Abelian character of QCD at lower energies, no measurements have been made since a longitudinally polarized electron beam is required, but has not been available until now.

The angular distribution expected after integrating over other variables is[23]:

$$\frac{d\sigma}{d\cos\omega} \propto (1 - \frac{1}{3}\cos^2\omega) + \beta \cdot A_Z \cdot \cos\omega, \tag{10}$$

where ω is the polar angle of the vector along $\vec{k_1} \times \vec{k_2}$, and A_Z is the spin polarization of the Z^0 along the Z axis. Since the sign and magnitude of this term is different for the two beam helicities, the $\cos\omega$ distribution is examined separately for the two cases. It has been shown[23] that in the Standard Model, β vanishes identically at tree level, but could have a nonzero value due to contributions from the interference between tree level and higher-order terms. However, the effect is expected to be very small ($|\beta| < 10^{-5}$). Thus, any experimentally observable nonzero value would indicate physics beyond the Standard Model.

Figure 11 shows the angular distributions for left-handed ($P_{e^-} < 0$) and right-handed ($P_{e^-} > 0$) electron beams for the 1994-1995 data. A fit, also shown in Fig. 11 and including the 1993 data, yields $\beta = 0.008 \pm 0.015$ with 95% confidence level limits of $-0.022 < \beta < 0.039$. These limits take into account the acceptance and efficiency of the detector as well as the probability (76.4%) that the jet-energy ordering mirrors that of the partons.

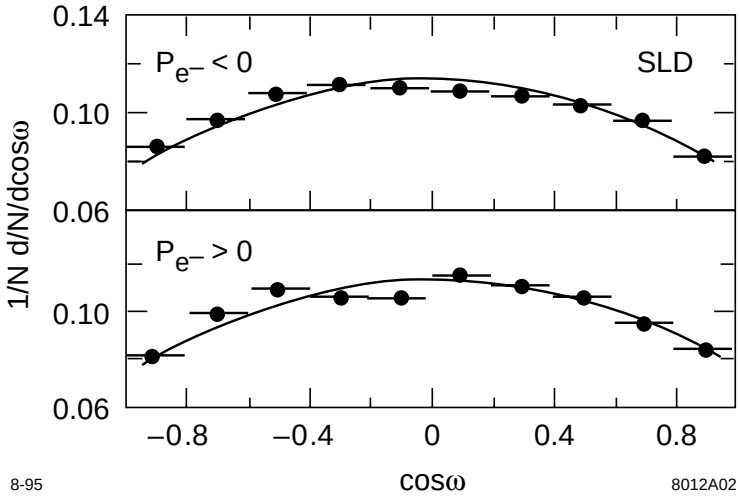


Fig. 11. $\cos\omega$ distribution of the 1994-1995 data for (top) left-handed and (bottom) right-handed electron beams.

Thus, the first measurement of the triple-product correlation in polarized Z^0 decays to three jets finds no statistically convincing asymmetry and sets a 95% confidence level limit on the rate of T_N-odd Z^0 decays to three jets of $|\beta| < 0.039$.

8 Production of $\pi^{\pm}$, $K^{\pm}$, p, K^0, Λ in Hadronic Z^0 Decays

The SLD detector combined with the SLC has a variety of excellent capabilities that are useful for identifying different types of particles; most of them were utilized in the analysis described in this study[7]:

- Charged hadrons can be identified from the cone angle of their Cherenkov radiation in a liquid and/or gas radiator as measured in the CRID (Cherenkov Ring Imaging Detector).
- Light- and heavy-flavor events can be isolated using the impact parameter analysis described earlier using combined drift chamber and vertex detector reconstruction.

- Samples enriched in quark jets or in antiquark jets can be isolated using the electron beam polarization and the fact that quarks (antiquarks) prefer to be left- (right-) handed.

Since it is so vital to this analysis, a short description of the CRID detector is appropriate. Figure 12 illustrates the principal design features of the barrel CRID. A charged particle, whose momentum is determined by the central drift chamber, passes through a liquid radiator producing an average of 11 detected photons. After passing through a drift box, the particle emits an average of nine detected photons in a gas radiator. The gas photons are focused back onto a circular image in the drift box by a set of spherical mirrors; the liquid photons also produce a roughly circular image due to the fact that the liquid radiator is thin compared to its distance from the drift box "proximity focusing." The two radiators with different indices allow particle identification over a wide range of momenta using the fact that the cone angle of the emitted radiation is given by $\cos\theta = 1/n\beta$, where n is the index of refraction and β is the speed of the particle. The number of photons expected is also used in identifying particles as the number increases from zero below threshold to an average of about ten for each of the two radiators for momenta far above threshold. The main properties of the two radiators are shown in the table.

				Max. Ring	Threshold (GeV/c)		
Radiator	Material	Length	Index	Radius	π	K	p
Liquid	C_6F_{14}	1 cm	1.277	16 cm	0.18	0.62	1.2
Gas	C_5F_{12}	$\sim$ 40 cm	1.0017	3 cm	2.4	8.4	16.0

There are altogether 40 drift boxes and 400 mirrors in the barrel CRID. The drift boxes have a small amount of Tetrakis diMethylAmine Ethylene (TMAE) which acts as a photocathode, converting the photons to electrons with a mean free path of around 1.6 cm.

A maximum likelihood is constructed for each mass hypothesis for each track, where the product, i, is over all measured photons. We expect n_j photons at angles θ_j and background b. Particle identification is then based on differences in L_j. For example, $L_K - L_\pi > 4.5$ identifies the particle as a K and rejects the π interpretation with 3σ confidence.

$$\pounds_j \propto -ln\left(\frac{\overline{n_j}}{n!}e^{-\overline{n_j}}\prod_i[exp\left(\frac{-(\theta_i-\theta_j)^2}{2\sigma^2}\right)+b]\right). \qquad (11)$$

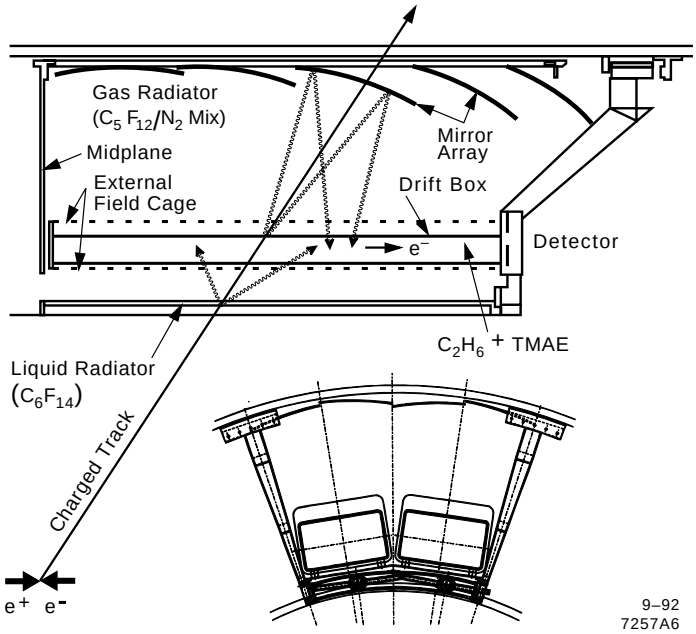


Fig. 12. Side and end views of one section of the barrel CRID.

In order to correct for the various inefficiencies, the following formalism was employed. A MC calculation, including a detailed description of the SLD detector, including the CRID, gave the number of observed and identified π's, K's, and protons for an assumed number of produced π's, K's, and protons. It is then only necessary to invert the matrix to obtain the true fractions of π's, K's, and protons produced. This formalism can be written as:

$$\begin{pmatrix} N_\pi \\ N_K \\ N_p \end{pmatrix} = \begin{bmatrix} \epsilon_{\pi\pi} & \epsilon_{\pi K} & \epsilon_{\pi p} \\ \epsilon_{K\pi} & \epsilon_{KK} & \epsilon_{Kp} \\ \epsilon_{p\pi} & \epsilon_{pK} & \epsilon_{pp} \end{bmatrix} n_{ch} \begin{pmatrix} f_\pi \\ f_K \\ f_p \end{pmatrix}.$$

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \text{ID'd Particles} & \text{Efficiencies} & \text{True Fractions} \end{array}$$

N_π is the number of particles identified by SLD as pions, f_π is the fraction of particles produced which are truly pions, and the matrix $[\epsilon]$ is the efficiency matrix. $\epsilon_{\pi K}$ is the probability that a produced K is identified as a π. The efficiency matrix is **not** constrained so $\Sigma f_i = 1$. This matrix can then be inverted to obtain an estimate of the true fractions:

$$\begin{pmatrix} f_\pi \\ f_K \\ f_p \end{pmatrix} = \begin{bmatrix} \epsilon_{\pi\pi} & \epsilon_{\pi K} & \epsilon_{\pi p} \\ \epsilon_{K\pi} & \epsilon_{KK} & \epsilon_{Kp} \\ \epsilon_{p\pi} & \epsilon_{pK} & \epsilon_{pp} \end{bmatrix}^{-1} \frac{1}{n_{ch}} \begin{pmatrix} N_\pi \\ N_K \\ N_p \end{pmatrix}.$$

Corrections for electrons and muons, which are sometimes confused with pions, were made using the MC. Only negative tracks are used below 2 GeV/c as an appreciable number of slow protons are produced in secondary interactions.

The efficiencies for identifying particles of type i as type j, where $i, j = \pi, K, p$ are shown in Fig. 13 as a function of momentum. The values at a given momentum from this array of graphs are essentially the efficiency matrix described above. Correct identification efficiencies ($i = j$) peak near 90% and misidentification rates are typically less than 5% with a peak value of 9%. These efficiencies were checked using pions from K^0_s decays with results that agreed with the MC simulation within the statistical errors.

Figure 14 shows the measured hadron fractions as a function of momentum. The solid line and dashed line show predictions of the JETSET 7.4 simulation program,[24] both in the standard version, called "Vanilla" and in the "SLD tuned" version, which includes a better β-decay model and is tuned to better reproduce a wide range of data from LEP. The lower figure shows the sum of the efficiencies. Since the efficiencies were not constrained to add up to 1.0, the consistency of the sum with 1.0 is a check on the validity of the method.

In some momentum regions, it is impossible to distinguish between two of the three species, so the procedure was reduced to a 2 x 2 analysis. In that case, only the fraction of the identified species are shown, i.e., protons in the liquid (gas) analysis above 3 (27) GeV/c and pions in the gas analysis below 10.5 GeV/c. Systematic errors dominate for the liquid points and are comparable to the statistical errors for the gas points.

Pions dominate at low momentum, but decline gradually to two-thirds by 10 GeV/c. Most of the decline is taken up by the kaon fraction, which rises to about one-third, whereas the proton fraction remains below one-tenth. These results are in agreement with results from LEP, where particle identification has been carried out both using ionization[25,26] as shown in Fig. 15 and using ring imaging[27] similar to CRID as shown in Fig. 16. Combining the techniques permits continuous coverage from 0.2 GeV/c to 35 GeV/c.

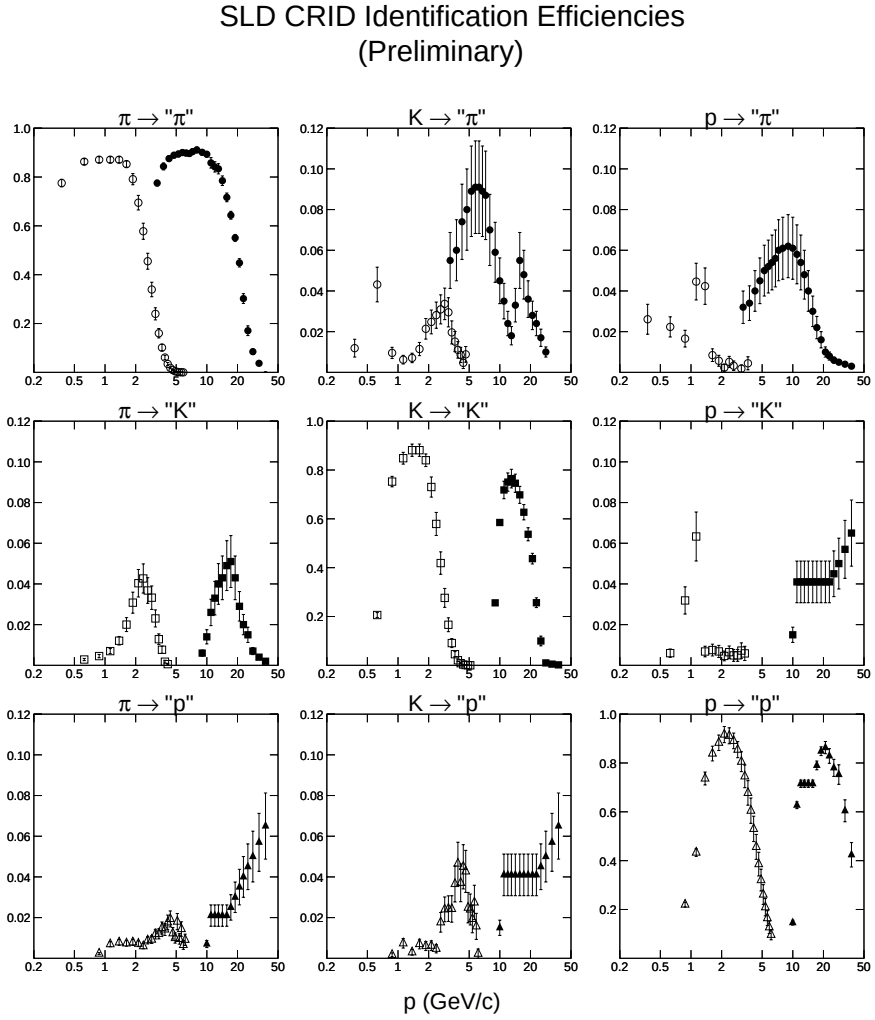


Fig. 13. Identification efficiencies for charged π, K, and $p/\overline{p}$ as a function of momentum for particles identified using the liquid radiator (open circles at lower momenta) and the gaseous radiator (solid circles). The notation $K \rightarrow$ "π" signifies true K's which are misidentified as π's.

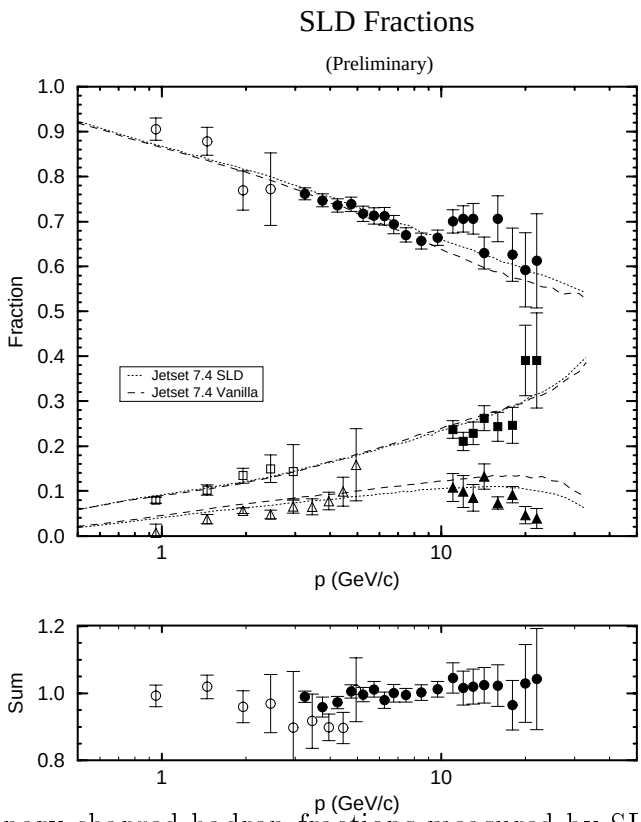


Fig. 14. Preliminary charged hadron fractions measured by SLD. Circles denote particles identified as $\pi^{\pm}$ (top points), squares are $K^{\pm}$ (middle points), and triangles are $p/\overline{p}$ (lower points). Open symbols are for the liquid radiator and solid symbols for the gas radiator. The lower figure shows the sum of the efficiencies, which should equal 1.0.

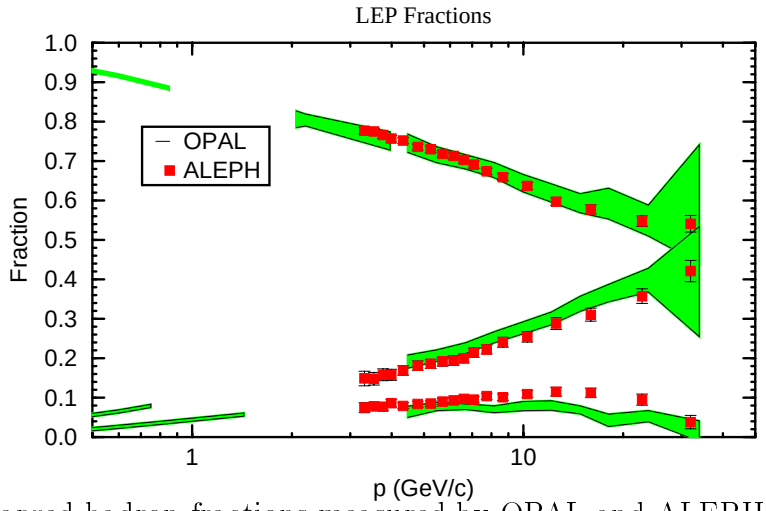


Fig. 15. Charged hadron fractions measured by OPAL and ALEPH using ionization loss techniques.

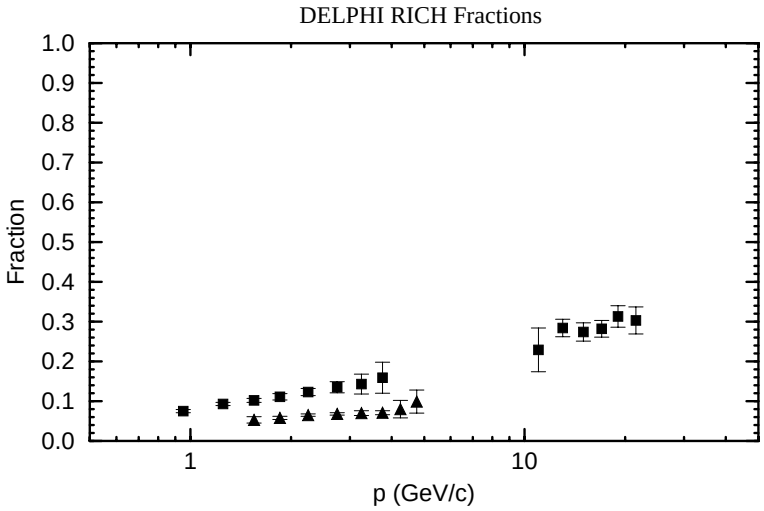


Fig. 16. Charged hadron fractions measured using the ring imaging technique in the DELPHI RICH.

Neutral strange particle production (Λ/K^0) was also measured starting with all well-measured pairs of oppositely charged particles in the barrel region ($|\cos\theta| < 0.80$). Other requirements included an acceptable vertex with fit probability greater than 2%, the V^0 momentum pointing back to the IP, and an e^+e^- invariant mass greater than 70 MeV. The kinematic overlap between Λ's and K^0 was eliminated using standard techniques.

Figure 17 shows the measured hadron fractions as a function of scaled momentum $x_p = 2p/W$ for the five measured species. Cross sections for the charged hadrons were obtained by multiplying the measured fractions by the total hadronic cross section generated by JETSET 7.4 (Ref. 24). Note that the cross sections for charged and neutral kaon production are consistent.

For comparison with QCD+LPHD predictions,[28] the analysis was repeated as a function of $\xi = \ln(1/x_p)$. The resulting spectra are shown in Fig. 18 and in general are adequately fit by Gaussians within the point-to-point errors. The peak positions are listed in the table.

Particle	ξ-Peak Position
$\pi^{\pm}$	$3.74 \pm 0.01 \pm 0.04$
K^0	$2.60 \pm 0.02 \pm 0.02$
$\Lambda\overline{\Lambda}$	$2.63 \pm 0.04 \pm 0.01$

This analysis was repeated on high-purity light (uds) and heavy (b) quark samples which were isolated as described in the section on multiplicity measurements.

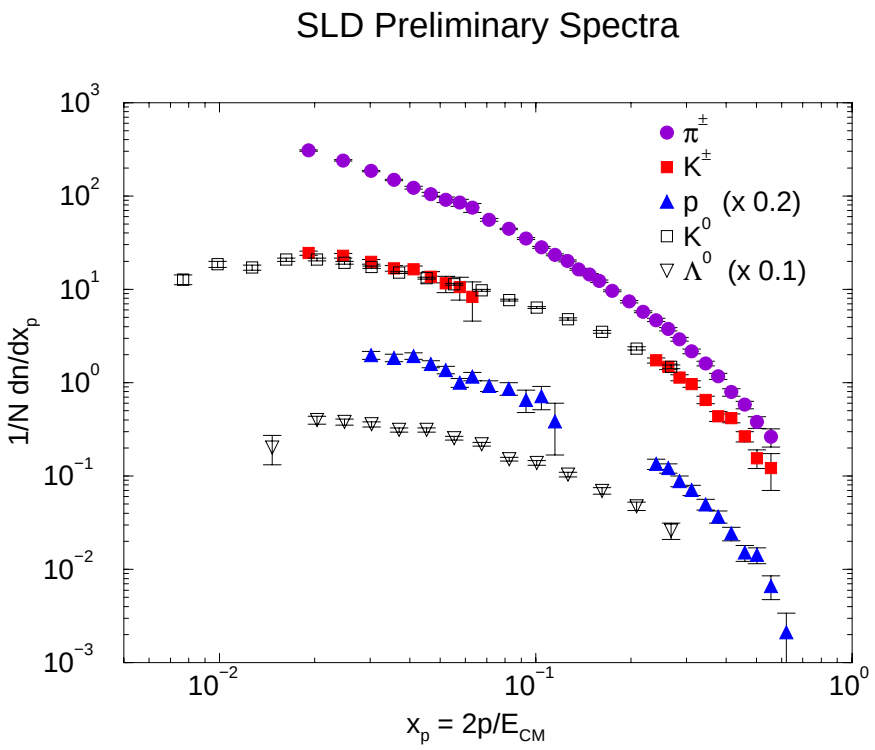


Fig. 17. Preliminary SLD spectra for five particle species as a function of $x_p = p/p_{beam}$. The $p/\overline{p}$ points are multiplied by 0.2 and the $\Lambda/\overline{\Lambda}$ points by 0.1 to make the figure clearer.

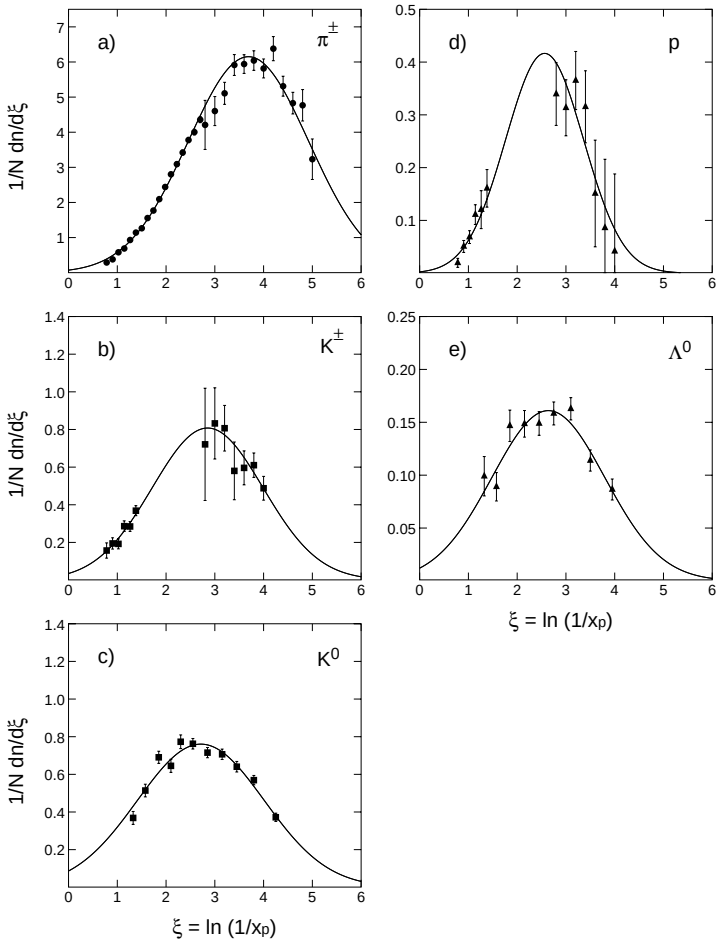


Fig. 18. Preliminary production spectra as a function of $\xi = -\ln x_p$. The curves are Gaussian fits. As in Fig. 17, the errors include systematic and statistical contributions added in quadrature. The systematic errors on the charged species are dominant and correlated point-to-point. An overall normalization uncertainty of 4% is not included.

The resulting spectra are shown in Fig. 19. Note the higher production of kaons and pions in the b-tagged sample and lower production of baryons. Except for the pion peak being lower in the b sample by 0.15 ± 0.05, there are no significant differences.

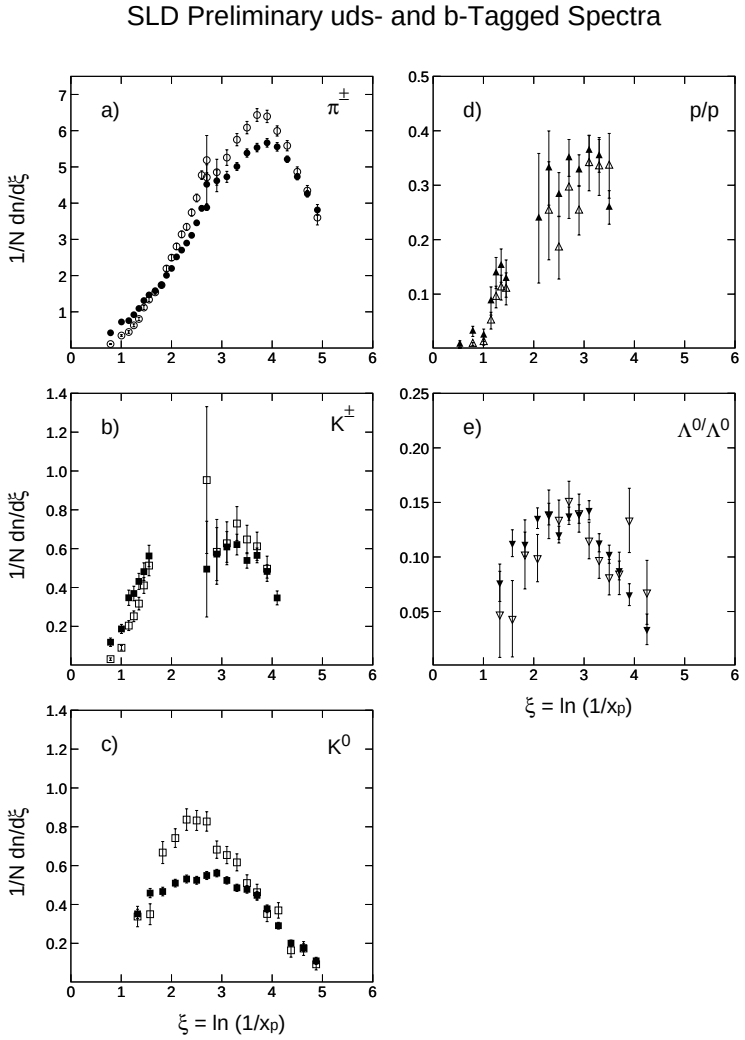


Fig. 19. Preliminary production spectra for samples enriched in $Z^0 \to u\overline{u}, d\overline{d}, s\overline{s}$ (solid symbols), and $Z^0 \to b\overline{b}$ (open symbols).

As a final exploitation of the beam polarization, an investigation was carried out to see if the quark forward-backward asymmetry relative to the beam polarization produces an observable leading-particle effect.

Figure 20 shows that fast Λ's and protons do indeed tend to follow the quark direction while fast $\overline{p}$'s and $\overline{\Lambda}$'s tend to follow the anti quark direction. The

average asymmetry for p > 10 GeV/c is 0.44 ± 0.07, thereby giving six standard deviations evidence for a leading-particle effect.

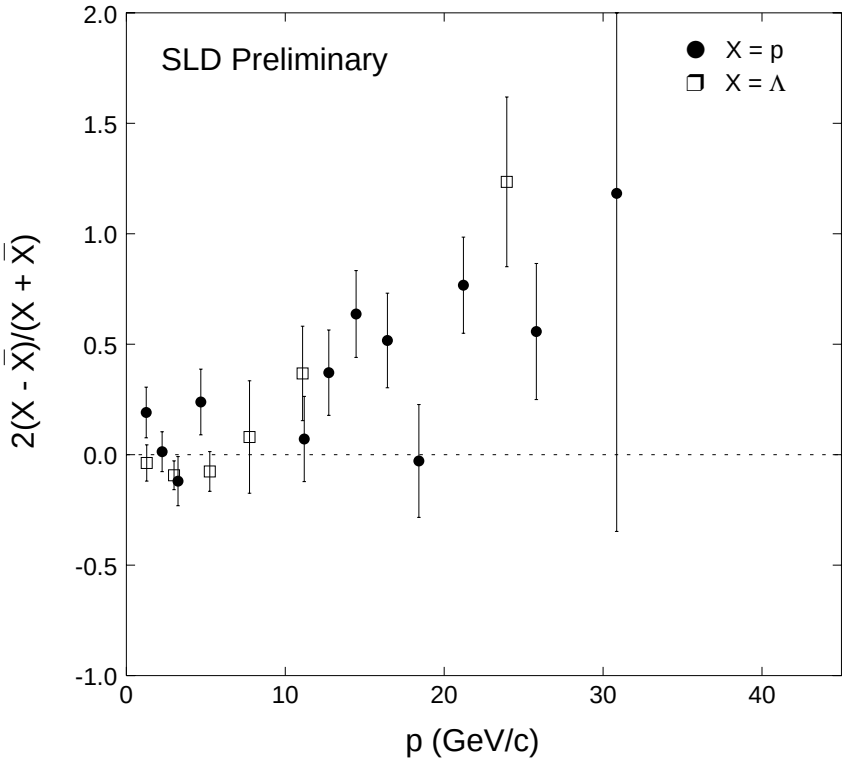


Fig. 20. The asymmetry between p, Λ, and their antiparticles in quark-tagged jets (those in the electron direction for left-handed electrons and in the opposite direction for right-handed beam polarization). Statistical errors only.

9 Summary

As should be clear from the results presented here, SLD is doing excellent QCD physics. In many cases, SLD is competitive or even superior to LEP in spite of the much higher statistics available to the LEP experiments.

Many of the results are limited by systematic errors, even at the statistics available to SLD. Since the systematics at SLC/SLD are quite different from those

at LEP, independent measurements are well worth doing. This is especially true because of the quality of the SLD detector and the unique features of SLC/SLD:

- Flavor selection using the number of significant impact parameters has been highly effective due to the tiny beam size, stable beam position, and precise vertex detector.
- Particle ID has shown high efficiency.
- The beam polarization has not only produced the best determination of $\sin^2\theta_W$, but has made possible the demonstration of a leading-particle effect associated with quark jets, and helped to show that other possible spin effects are not easily detectable.

The future looks promising, given success in meeting the following goals:

- A vertex detector with increased coverage, efficiency, and accuracy.
- Increased statistics, with continuing high beam polarization.
- Reliable operation of both SLC and SLD.

References

[1] K. Abe *et al.*, "Factorial and cumulant moments in $e^+e^- \to$ hadrons at the Z^0 resonance," SLAC-PUB-7027 (1996); to appear in Phys. Lett. B.

[2] K. Abe *et al.*, "Measurement of the charged multiplicities of b, c, and light quark events from Z^0 decays," SLAC-PUB-6924 (1995).

[3] K. Abe *et al.*, "First study of rapidity gaps in e^+e^- annihilation," SLAC-PUB-7076 (1996); to appear in Phys. Rev. Lett.

[4] K. Abe *et al.*, "A study of the orientation and energy partition of three-jet events in hadronic Z^0 decays," SLAC-PUB-6922 (1995).

[5] K. Abe *et al.*, "An improved limit on jet handedness in Z^0 decays," SLAC-PUB-6921 (1995).

[6] K. Abe *et al.*, "First measurement of the T-odd correlation between the Z^0 spin and the three-jet plane orientation in polarized Z^0 decays to three jets," Phys. Rev. Lett. **75**, 4173 (1995).

[7] K. Abe *et al.*, "Production of $\pi^{\pm}$, $K^{\pm}$, p, K^0, Λ in hadronic Z^0 decays," SLAC-PUB-6920 (1995).

[8] SLD Design Report, SLAC Report 273 (1984).

[9] K. Abe *et al.*, Nucl. Instrum. Methods A **288**, 288 (1990).

[10] G. Giacomelli, Int. J. Mod. Phys. A **5**, 223 (1990).

[11] Z. Koba, M. B. Nielsen, and P. Oleson, Nucl. Phys. B **240**, 317 (1972).

[12] I. M. Dremin, Phys. Lett. B **313**, 209 (1993).

[13] I. M. Dremin *et al.*, Phys. Lett. B **336**, 119 (1994).

[14] I. M. Dremin, Mod. Phys. Lett. A **8**, 2747 (1993).

[15] Y. L. Dokshitzer *et al.*, J. Phys. G **17**, 1481 (1991); *ibid* G **17**, 1602 (1991).

[16] V. A. Petrov and A. V. Kisselev, CERN-TH-7318-94 (1994).

[17] F. Abe *et al.*, Phys. Rev. Lett. **74**, 855 (1994).

[18] T. Ahmed *et al.*, Nucl. Phys. B **249**, 477 (1994).

[19] J. Bjorken *et al.*, Phys. Lett. B **286**, 153 (1992).

[20] TASSO Collaboration, R. Brandelik *et al.*, Phys. Lett. B **86**, 243 (1979); Mark II Collaboration, D. P. Barber *et al.*, Phys. Rev. Lett. **43**, 830 (1979);
PLUTO Collaboration, C. Berger *et al.*, Phys. Lett. B **86**, 418 (1979); JADE Collaboration, W. Bartel *et al.*, Phys. Lett. B **91**, 142 (1980).

[21] SLD Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **74**, 1512 (1994).

[22] K. Fabricius *et al.*, Phys. Rev. Lett. **45**, 867 (1980), and J. G. Körner *et al.*, Phys. Lett. B **94**, 207 (1980).

[23] A. Brandenburg, L. Dixon, and Y. Shadmi, SLAC-PUB-95-6725, April 1995 (to be published in Phys. Rev. D, 1996).

[24] ALEPH Collaboration, D. Decamp *et al.*, Z. Phys. C **55**, 209 (1992).

[25] OPAL Collaboration, P. D. Acton *et al.*, Z. Phys. C **63**, 181 (1994).

[26] ALEPH Collaboration, D. Decamp *et al.*, CERN-PPE-94/201 (submitted to Z. Physics C).

[27] DELPHI Collaboration, P. Abreu *et al.*, CERN-PPE-95/28 (submitted to Nucl. Phys. B).

[28] T. I. Azimov, Y. L. Dokshitzer, V. A. Khoze, and S. I. Troyan, Z. Phys. C **27**, 65 (1985).

SEARCHES FOR VERY RARE DECAYS OF KAONS

Karol Lang *
Department of Physics
University of Texas at Austin
Austin, Texas 78712

ABSTRACT

The physics motivation for searches for very rare kaon decays, either forbidden or suppressed within the Standard Model, is briefly discussed. Simple arguments conclude that such searches probe possible new forces at a 200 TeV mass scale or constitute a precision test of the electroweak model. The examples of such processes are decays of $K_L^0 \rightarrow \mu^{\pm} e^{\mp}$, $K^+ \rightarrow \pi^+ \mu^+ e^-$, $K_L^0 \rightarrow \mu^+ \mu^-$, and $K^+ \rightarrow \pi^+ \nu \bar{\nu}$. We present the current experimental status and describe the new efforts to reach sensitivities down to one part in 10^{12}. The discussion is focused on the experimental program at the Alternating Gradient Synchrotron at Brookhaven National Laboratory, where intense beams make such studies possible.

*Supported by DOE Contract DE-FG03-93ER-40757 (Task OJI).

1 Introduction

Experimental confirmations of a theoretical picture of elementary interactions, the Standard Model, constitute an undeniable triumph of particle physics. Despite that, the Standard Model is generally perceived as incomplete or merely a *low* energy realization of a more general theory with a full symmetry at some high energy. There are indeed many basic questions regarding the Standard Model to which answers will have to come from the outside of the model. With 21 free parameters in the minimal version of the Model, an overall picture is complicated and many *fundamental* questions are obvious: Why are there three quark and lepton families? Why is the number of families three in both cases? Why are the masses of constituent fermions and intermediate bosons what they are? Why and how is CP violated? Why is separate lepton-flavor conserved? Etc.

As the consequence of this situation, many new theoretical models (or entire classes of models) have been proposed over the years. The most successful ones [i.e., the ones which have survived the challenge of existing experimental data and make (usually hard to test) predictions of new phenomena] include supersymmetry,[1] technicolor,[?] left-right symmetry,[2] horizontal symmetry,[3] compositeness,[4] and the most hopeful candidate for the ultimate theory of elementary particles, superstring theory.[5] At the same time, testing of the Standard Model continues by examining more and more subtle effects predicted by it. One such venue is studying higher-order electroweak (i.e., suppressed) processes. The level at which they occur subjects the theory to a stringent scrutiny and may lead to observed inconsistencies pointing towards "new physics." Processes like $K^+ \to \pi^+ \nu\bar{\nu}$, $K_L^0 \to \mu^+\mu^-$, and $K_L^0 \to e^+e^-$ could be used both ways: to confirm or to look beyond the Standard Model. They can shed light on the CP-violation mechanism—one of the crucial tests of the present theory.

Looking back at the history of elementary particles, one cannot escape a conclusion that kaons played a central role in many discoveries. Starting with associated (strangeness) production and "$\tau - \theta$ puzzle," through CP-violation and GIM-mechanism, the kaon system was essential in establishing the foundations of the Standard Model. Studies of rare kaon decays continue this physics-rich tradition. Rare processes provide sensitive tests of new theories and an important testing ground for the Standard Model. The field of rare decays is diverse and active as it offers potential for major discoveries. At the Alternating Gradient

Synchrotron (AGS) at Brookhaven National Laboratory (BNL), the program is centered around tests of the separate lepton-flavor conservation ($K_L^0 \rightarrow \mu^{\pm}e^{\mp}$, $K^+ \rightarrow \pi^+\mu^+e^-$, $\pi^0 \rightarrow \mu^{\mp}e^{\pm}$), and measurements of the suppressed electroweak decays ($K^+ \rightarrow \pi^+\nu\bar{\nu}$, $K_L^0 \rightarrow \mu^+\mu^-$). The newest generation of detectors optimized for these decays is completed, and now the AGS program is entering the extended data-taking period.

This review focuses on searches and studies of processes pursued by experiments at BNL. Transitions relevant for the CP-violation (like $K_L^0 \rightarrow \pi^0e^+e^-$) are primarily studied at other laboratories and will not be covered here. Some other recent articles offer more complete overviews.[6,7,8,9]

2 Motivation

2.1 Forbidden Processes

Many new theories, or "extensions" of the Standard Model, predict observation of separate lepton-flavor violation. In a variety of theories, as illustrated in Fig. 1, processes like $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ or $K^+ \rightarrow \pi^+\mu^+e^-$ occur at tree-level. Such decays are forbidden within the Standard Model although the origin of separate lepton-flavor conservation is not understood. It is not associated with any known symmetry usually expected to exist behind a conservation law. Decays $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ or $K^+ \rightarrow \pi^+\mu^+e^-$ have to be mediated by a *new* intermediate boson, thus explicitly manifesting *a new fundamental force* in nature. As such, these processes comprise some of the most sensitive tests of new theories.

Figure 1 illustrates how some models lead naturally to $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$. To construct "a dimensional argument," one can assume a $V-A$ form of a new interaction and compare it to a copious electroweak decay $K^+ \rightarrow \mu^+\nu_\mu$ as shown in Fig. 2:

$$\frac{\Gamma(K_L^0 \rightarrow \mu^{\pm}e^{\mp})}{\Gamma(K^+ \rightarrow \mu^+\nu_\nu)} \simeq \left[\frac{ff'/M_X^2}{g^2 sin\theta_C/M_W^2}\right]^2 \qquad (1)$$

where θ_C is the Cabibbo angle, M_W the mass of the W boson, and g is the electroweak coupling. Thus the branching ratio is

$$B(K_L^\circ \rightarrow \mu^{\pm}e^{\mp}) = \frac{\Gamma(K_L^\circ \rightarrow \mu^{\pm}e^{\mp})}{\Gamma(K_L^0 \rightarrow all)} \simeq \frac{\Gamma(K^+ \rightarrow \mu^+\nu_\mu)}{\Gamma(K_L^0 \rightarrow all)}\left[\frac{ff'/M_X^2}{g^2 sin\theta_c/M_W^2}\right]^2 \quad . \qquad (2)$$

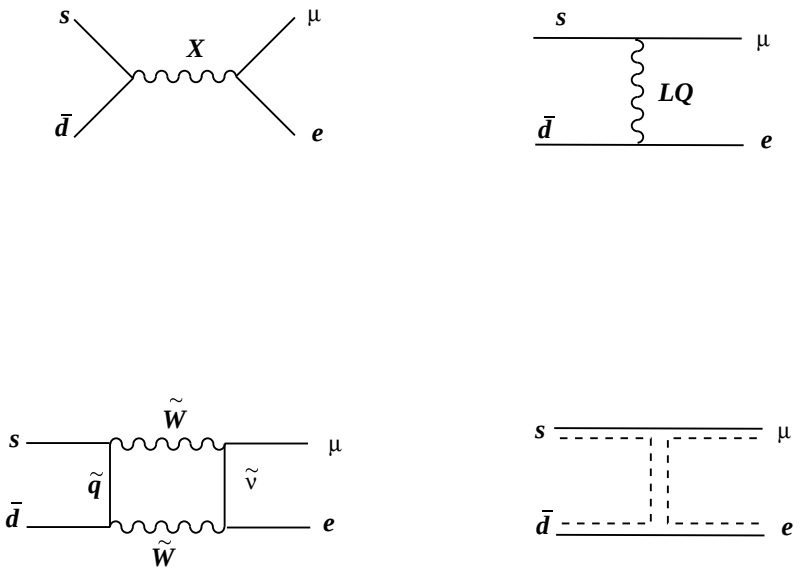


Fig. 1. Examples of tree-level diagrams leading to the separate lepton-flavor violation in four classes of theories "beyond" the Standard Model: (a) horizontal symmetry, (b) leptoquarks, (c) supersymmetry, and (d) compositeness.

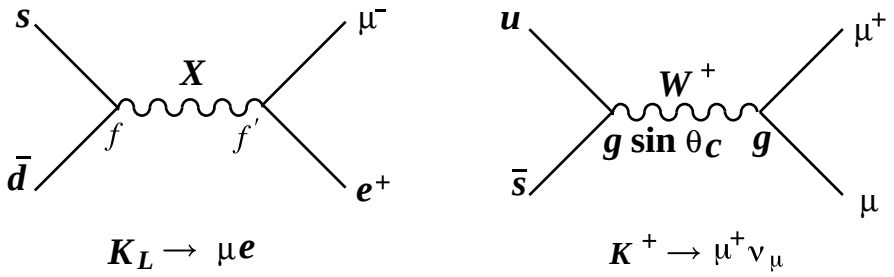


Fig. 2. Examples of a tree-level diagram with "horizontal symmetry" leading to the separate lepton-flavor violation $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$, and a copious electroweak decay $K^+ \rightarrow \mu^+\nu_\mu$.

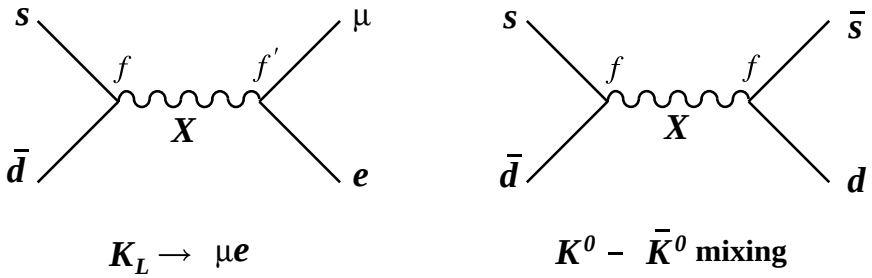


Fig. 3. A diagram with "horizontal symmetry" leading to the separate lepton-flavor violation $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ could also contribute to $K^0 \leftrightarrow \bar{K}^0$.

Inserting the known numbers gives

$$B(K_L^0 \rightarrow \mu e) \simeq (1.2 \times 10^{-2} TeV^4) \left[\frac{f^2}{M_X^2}\right]^2 \left[\frac{f'}{f}\right]^2 \quad . \tag{3}$$

Assuming for simplicity the same couplings $f = f' = g$ turns this comparison to a specific relationship of branching fractions versus the mass scale of the new interaction

$$M_X \simeq 220\ TeV \left[\frac{10^{-12}}{B(K_L^0 \rightarrow \mu e)}\right]^{1/4} \quad . \tag{4}$$

A similar comparison between $K^+ \rightarrow \pi^+\mu^+e^-$ and $K_L^0 \rightarrow \pi^{\pm}\mu^{\mp}\nu_{\mu}$ yields

$$M_X \simeq 86\ TeV \left[\frac{10^{-12}}{B(K^+ \rightarrow \pi^+\mu^{\pm}e^{\mp})}\right]^{1/4} \quad . \tag{5}$$

Thus sensitivity to *rare* processes opens a window at interactions at *very high mass scales* through virtual effects of *new particles*. It should be emphasized that $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ and $K^+ \rightarrow \pi^+\mu^+e^-$ provide complementary information on potential new interactions. The first is sensitive to axial-vector or pseudoscalar couplings; the second is sensitive to vector or scalar couplings.

For completeness, it is worth mentioning that a new force in $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ may also contribute to $K^0 \leftrightarrow \bar{K}^0$ transitions as illustrated in Fig. 3. Thus stringent limits on M_X from Δm_K exist. However, some theories circumvent this restriction.[10] If one assigns "a generation number," G, to leptons and quarks with G = 1 for (e, ν_e, u, d), G = 2 for (μ, ν_μ, c, s), and G = 3 for (τ, ν_τ, t, b), then $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ is a $\Delta G = 0$, and $K^0 \leftrightarrow \bar{K}^0$ is $|\Delta G| = 2$ transition. If G is conserved by the new force (representing some unknown symmetry), $K_L^0 \rightarrow \mu^{\pm}e^{\mp}$ would be allowed without affecting $K^0 \leftrightarrow \bar{K}^0$.

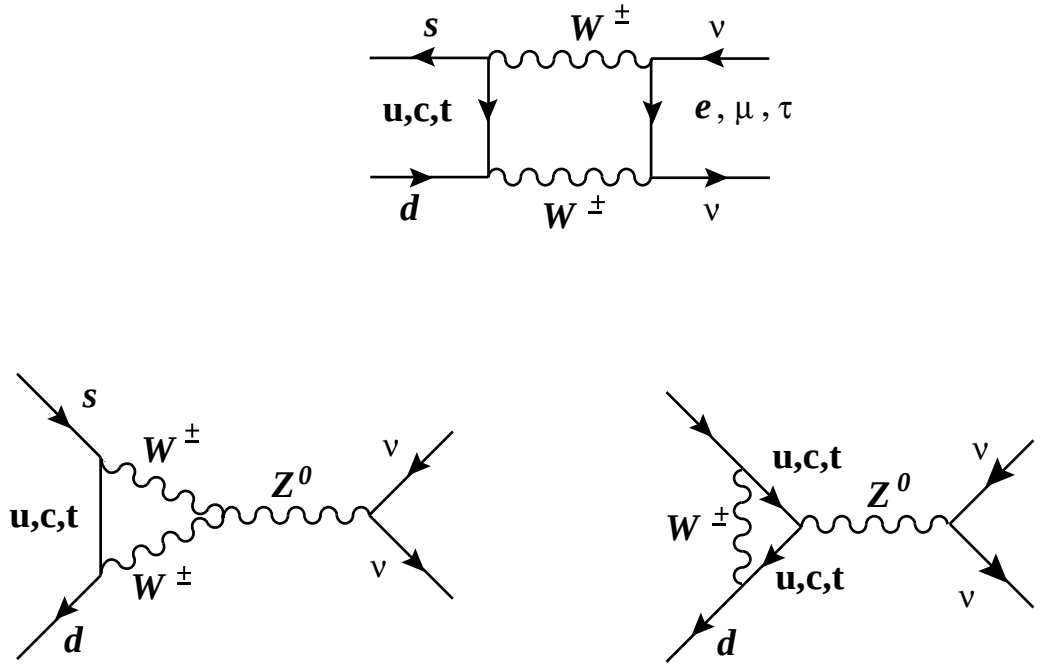


Fig. 4. Short-distance diagrams which dominate the transition $K^+ \rightarrow \pi^+\nu\bar{\nu}$.

2.2 Suppressed Processes

Decays proceeding through a higher-order (loop-level) electroweak transition constitute an essential test of the Standard Model. Rates of decays such as $K^+ \rightarrow \pi^+\nu\bar{\nu}$ or $K_L^0 \rightarrow \mu^+\mu^-$ [referred to as Flavor-Changing Neutral Currents (FCNC)] could be used as a check of the theory or to extract some of the parameters of the theory which are hard to reach through other processes.

The $K^+ \rightarrow \pi^+\nu\bar{\nu}$ decay mode is short-distance *dominated*[11,12,13,14] and proceeds mainly via diagrams shown in Fig. 4. With QCD corrections known and small,[15] the diagrams depend primarily on the mass of the charm and top quarks, m_t and m_c, and the $V_{ts}^* V_{td}$ product of the elements of the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix. If all uncertainties are included, the $B(K^+ \rightarrow \pi^+\nu\bar{\nu}\ )$ is expected in the range of $(0.5 - 5.0) \times 10^{-10}$. Thus, if observed, the process can be used to shed more light on m_t and the CKM matrix.

$K_L^0 \rightarrow \mu^+\mu^-$ mode has played a crucial role in establishing the multifamily quark content of the theory through the famous Glashow-Iliopoulos-Maiani cancellation ("GIM-mechanism").[16] The reaction is dominated by an electromagnetic two-(on-shell)photon transition, as shown in Fig. 5(a), which determines the

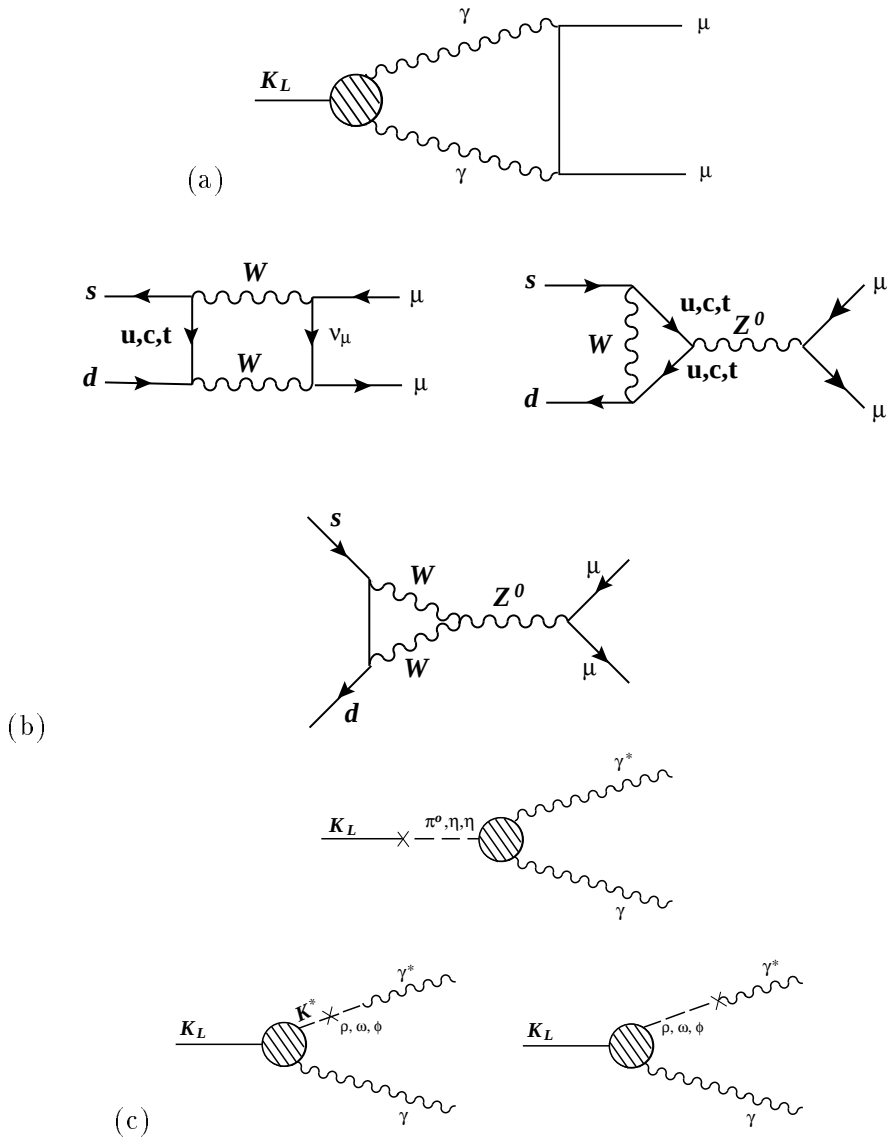


Fig. 5. Diagrams contributing to $K_L^0 \rightarrow \mu^+\mu^-$: (a) a dominant diagram which determines the unitarity boundary for this process at $(6.8 \pm 0.3) \times 10^{-9}$, (b) short-distance "box" and "penguin" diagrams (similar to $K^+ \rightarrow \pi^+\nu\bar{\nu}$), and (c) representation of long-distance contributions.

unitarity bound for the branching fraction[17]

$$\frac{\Gamma(K_L^0 \to \mu\mu)_{2\gamma}}{\Gamma(K_L^0 \to \gamma\gamma)} = \alpha^2 \left(\frac{m_\mu}{m_K}\right)^2 \frac{1}{2\beta} \left(ln\frac{1+\beta}{1-\beta}\right)^2 \quad , \tag{6}$$

where $\beta = \sqrt{1 - 4m_\mu^2/m_K^2}$. The numerical uncertainty of this process is primarily due to the experimental uncertainty of $B(K_L^0 \to \gamma\gamma)$, measured to be[18] $B(K_L^0 \to \gamma\gamma) = (5.70 \pm 0.27) \times 10^{-4}$. So

$$B(K_L^0 \to \mu\mu)_{2\gamma} = (1.2 \times 10^{-5}) B(K_L^0 \to \gamma\gamma) = (6.8 \pm 0.3) \times 10^{-9} \quad . \tag{7}$$

There are also other contributions to this transition: the short-distance diagrams[19,20] shown in Fig. 5(b) (similar to the $K^+ \to \pi^+\nu\bar{\nu}$ diagrams in Fig. 4 where neutrino lines are replaced with muon lines), and the long-distance diagrams exemplified by graphs in Fig. 5(c).

At present, it is not clear how to calculate the long-distance contributions with uncertainties small enough to extract the short-distance graphs[21,22] and thus obtain information on m_t and V_{td}. However, theoretical and experimental efforts are under way to improve this situation. Determination of $K_L^0 - \gamma^* - \gamma$ and $K_L^0 - \gamma^* - \gamma^*$ vertices from such decays as $K_L^0 \to e^+e^-\gamma$, $K_L^0 \to \mu^+\mu^-\gamma$, and $K_L^0 \to e^+e^-e^+e^-$ should provide additional guidance for the chiral perturbation theory so that the extraction of the short-distance part of $K_L^0 \to \mu^+\mu^-$ may be feasible.

In summary, the two decay modes $K^+ \to \pi^+\nu\bar{\nu}$ and $K_L^0 \to \mu^+\mu^-$ are complementary: $K^+ \to \pi^+\nu\bar{\nu}$ is harder experimentally, but cleaner theoretically; $K_L^0 \to \mu^+\mu^-$ is presently less well understood theoretically, but the near future should bring thousands of collected events.

2.3 Searches for New Particles and Other Rare Decays

As a byproduct of searches for rare decays of kaons, other channels are studied. In particular, modes like $K^+ \to \pi^+ X^0$, where X^0 is a new particle, could be examined. Particular experimental approaches impose certain sensitivity windows for masses and lifetimes of such new objects. In addition, hermeticity and acceptances of spectrometers also allow searches for other rare decays like $\pi^0 \to \mu^\mp e^\pm$ or measurement of $\pi^0 \to e^+e^-$.

Other less rare decay channels like $K^+ \to \pi^+e^+e^-$, $K^+ \to \pi^+\mu^+\mu^-$, $K^+ \to \pi^+\gamma\gamma$, $K^+ \to \pi^+\pi^0e^+e^-$, $K^+ \to e^+e^+e^-\nu$, $K^+ \to \mu^+e^+e^-\nu$,

$K^+ \rightarrow \pi^+\pi^0\gamma$, $K^+ \rightarrow \mu^+\nu_\mu\gamma$, and $K^+ \rightarrow \pi^0\mu^+\nu_\mu\gamma$ are also pursued, and could either present more stringent tests of the chiral perturbation theory or provide additional input to advance it for processes like $K_L^0 \rightarrow \mu^+\mu^-$.

3 Current Best Results

Searches for rare kaon decays have had a very long tradition. As illustrated in Figs. 6 and 7, the progression of new results is grouped in two time periods. In the late 1960s to early 1970s, experiments impressively probed branching ratios down to the 10^{-8} level. Exhausting detector technology and available beam intensities, the searches stagnated for about a decade. They were renewed in the early 1980s when more modern experiments were able to provide faster and better detectors and data acquisition, and more intense beams became available. Most of the final results from these recent searches have been published. Tables 1, 2, and 3 list these results. Figures 8, 9, 10, and 11 illustrate some of the results in a full form as published.

The lowest sensitivities have been reached at the AGS at BNL. Results from KEK played an important role in the early stages of the new generation experiments, but have been ultimately superseded by the BNL experiments. As this round of experiments exploited the detector and beam capabilities, the three groups at BNL, equipped with multiyear experience of conducting experiments with intense beams, embarked on the next round of studies.

The newest upgrades and follow-up experiments have been designed with goals to reach sensitivities which would "close the discovery window" for allowed pro-

Decay mode	90% C.L. limit	Reference
$K_L^0 \rightarrow \mu^\pm e^\mp$	$< 3.3 \times 10^{-11}$	BNL E791[23]
	$< 9.4 \times 10^{-11}$	KEK E137[24]
	$< 1.9 \times 10^{-9}$	BNL E780[25]
$K^+ \rightarrow \pi^+\mu^+e^-$	$< 2.1 \times 10^{-10}$	BNL 777[26]
$K^+ \rightarrow \pi^+X^0$	$< 5.2 \times 10^{-10}$	BNL 787[27]
	X^0 weakly interacting	

Table 1. Summary of best limits of searches beyond the Standard Model.

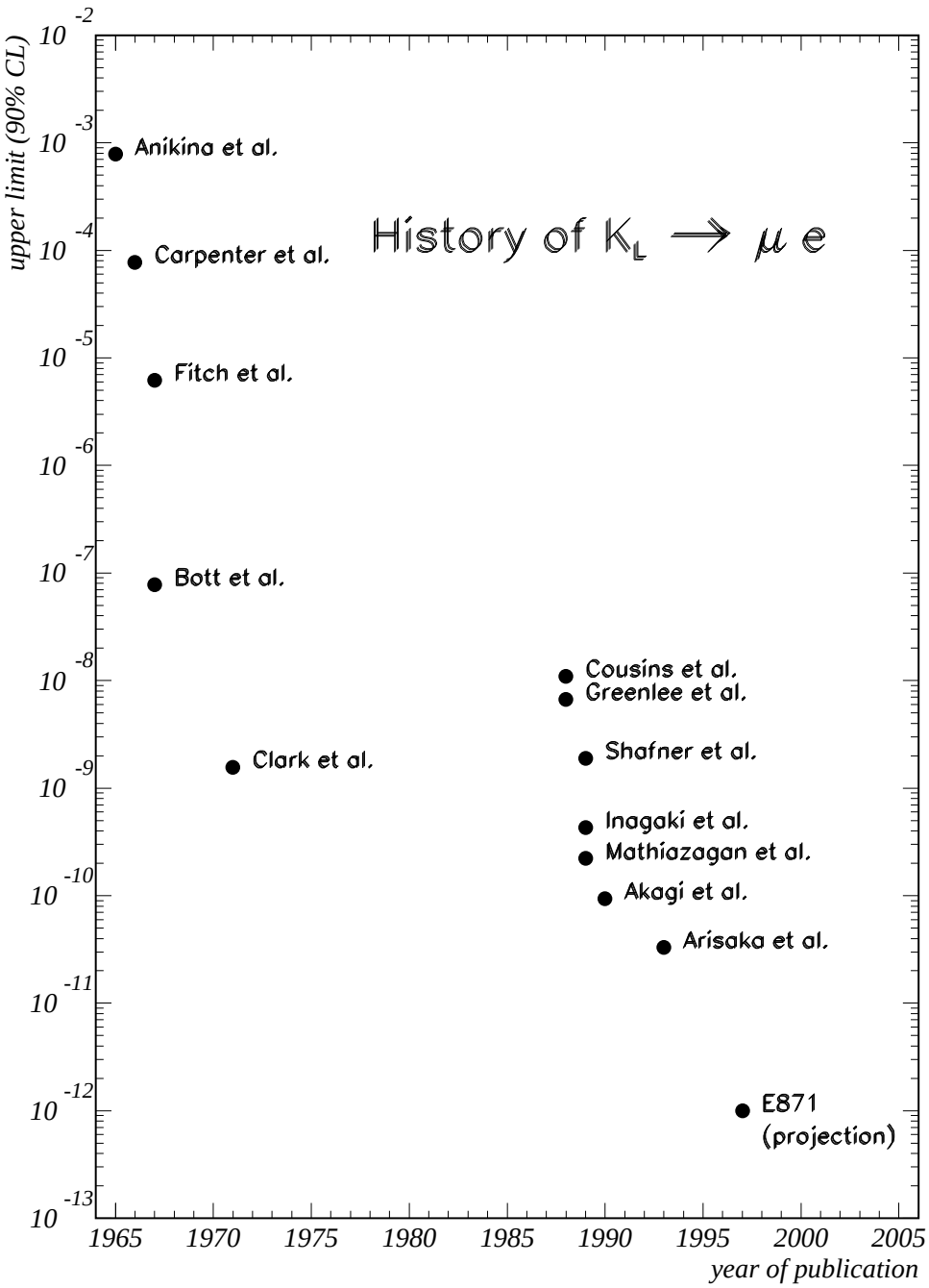


Fig. 6. History of $K_L^0 \rightarrow \mu^{\pm} e^{\mp}$ searches.[19] The points represent 90% C.L. upper limits.

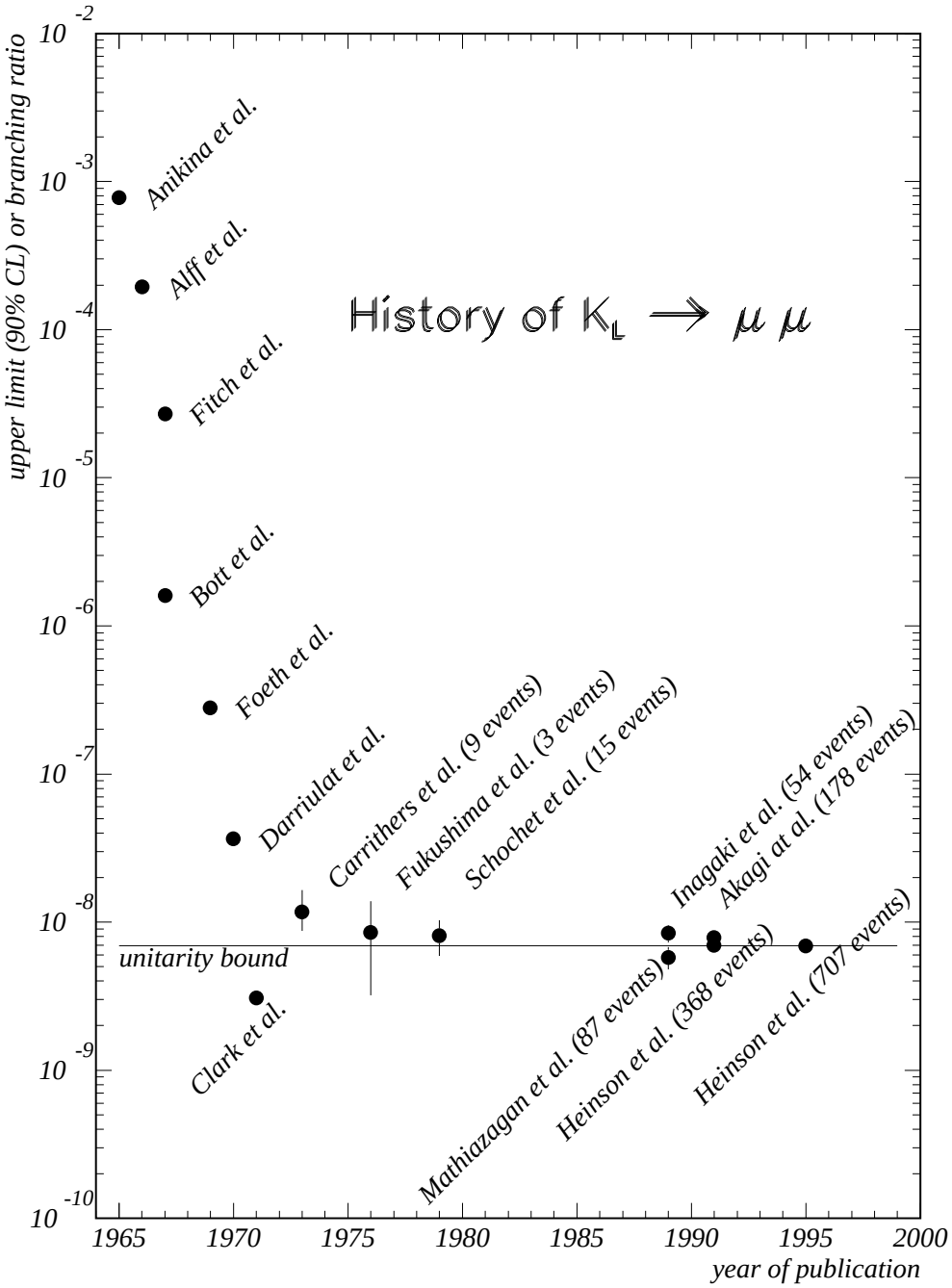


Fig. 7. History of $K_L^0 \to \mu^+\mu^-$ results.[19] The first observation of this decay was reported by Carrithers *et al.* in 1973; thus the earlier points represent 90% C.L. upper limits.

Decay mode	BR or 90% C.L. limit	Reference
$K^+ \rightarrow \pi^+\nu\bar{\nu}$	$< 2.4 \times 10^{-9}$	BNL E787[27]
$K_L^0 \rightarrow \mu^+\mu^-$	$(6.86 \pm 0.37) \times 10^{-9}$	BNL E791[28]
	$(7.9 \pm 0.6(stat.) \pm 0.3(sys.)) \times 10^{-9}$	KEK E137[29]
$K_L^0 \rightarrow e^+e^-$	$< 4.1 \times 10^{-11}$	BNL E791[30]
	$< 1.6 \times 10^{-10}$	KEK E137[24]
	$< 1.2 \times 10^{-9}$	BNL E780[31]

Table 2. Summary of best results on processes suppresed within the Standard Model.

Decay mode	BR or or 90% C.L. limit	Reference
$K^+ \rightarrow \pi^+X^0$	$< 1.5 \times 10^{-8*}$	BNL E851[32]
$X^0 \rightarrow e^+e^-$	$150\ MeV < m_{X^0} < 340\ MeV$	
$K^+ \rightarrow \pi^+X^0$	$< 1.6 \times 10^{-6}$	BNL E787[33]
$X^0 \rightarrow \gamma\gamma$	$0\ MeV < m_{X^0} < 150\ MeV$	
$K^+ \rightarrow \pi^+\gamma\gamma$	$< 1.0 \times 10^{-6}$	BNL E787[33]
$K^+ \rightarrow \pi^+H^0$	$< 1.5 \times 10^{-7}$	BNL E787[34]
$H \rightarrow \mu^+\mu^-$	$220\ MeV < m_{X^0} < 320\ MeV$	
$\pi^0 \rightarrow \mu^\mp e^\pm$	$< 1.6 \times 10^{-8}$	BNL E777[26]
$K^+ \rightarrow \pi^+e^+e^-$	$(2.75 \pm 0.23(stat.) \pm 0.13(syst.)) \times 10^{-7}$	BNL E851[32]
$\pi^0 \rightarrow e^+e^-$	$(6.9 \pm 2.3(stat.) \pm 0.6(syst.)) \times 10^{-8}$	BNL E851[35]
$\pi^0 \rightarrow e^+e^-$	$(7.6^{+3.9}_{-2.8}(stat.) \pm 0.5(syst.)) \times 10^{-8}$	FNAL E799[36]
$(m_{ee}/m_{\pi^0})^2 > 0.95$		

Table 3. Summary of other searches and limits from the rare K decay experiments.(* 99% C.L.)

Fig. 8. Final results from BNL E791. The top figure is a scatter plot of p_T^2 vs. $M_{\mu e}$ for $K_L^0 \rightarrow \mu^{\pm} e^{\mp}$ candidates[24] (p_T is a missing transverse momentum of the reconstructed final-state pair). The bottom figure shows the $K_L^0 \rightarrow \mu^+ \mu^-$ mass peak of > 700 events[29] (an angle θ_K points in the direction of the missing transverse momentum and is equivalent to p_T). The rectangular signal boxes surround $M_K = 0.497$ MeV/c^2 and $p_T = 0$ (or $\theta_K = 0$).

Fig. 9. Final results from BNL E777. The scatter plots of the variable S vs. invariant mass (S is the rms distance of closest approach of three track combinations to a common vertex). The top figure shows $K^+ \rightarrow \pi^+\pi^-\pi^0$ decays; the bottom $K^+ \rightarrow \pi^+\mu^+e^-$ candidates.[27]

Fig. 10. Final results from BNL E787. Charged-track range vs. kinetic energy for (a) data and (b) $K^+ \rightarrow \pi^+\nu\bar{\nu}$ Monte Carlo for events satifying the selection criteria and having measured momentum $211 \leq P_\pi \leq 243$ MeV/c. The rectangular box indicates the search region for $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and $K^+ \rightarrow \pi^+X^0$ ($M_{X^0} \approx 0$). The horizontal and vertical dashed lines in (b) are the theoretical end-points of $K^+ \rightarrow \pi^+\nu\bar{\nu}$ in range and energy, respectively.[28]

Fig. 11. Limits from the other searches from BNL E787: (a) the 90% C.L. upper limits on $B(K^+ \rightarrow \pi^+ X^0)$ as a function of M_{X^0} and various X^0 lifetimes;[28] (b) the 90% C.L. upper limit on $K^+ \rightarrow \pi^+ H^0$ with $H \rightarrow \mu\mu$ as a function of m_H;[35] (c) the 90% C.L. upper limit on $K^+ \rightarrow \pi^+ X^0$ with $X^0 \rightarrow \gamma\gamma$ for different X^0 lifetimes as a function of mass m_{X^0}. The dashed curve shows the upper limit for the combined branching ratio $K^+ \rightarrow \pi^+ H^0$ and $H^0 \rightarrow \gamma\gamma$.[34]

cesses ($K^+ \rightarrow \pi^+\nu\bar{\nu}$ and $K_L^0 \rightarrow e^+e^-$) and probe branching fractions at the 10^{-12} level ($K_L^0 \rightarrow \mu^\pm e^\mp$ and $K^+ \rightarrow \pi^+\mu^+e^-$). The new experiments have been commissioned and started the data-taking period which will extend into the near future. In the following part of this article, these experimental efforts will be briefly discussed.

4 Experiments in Progress

4.1 AGS

Alternating Gradient Synchrotron at BNL provides the beam for the rare kaon experiments. The accelerator was commissioned in 1960 and originally provided 10^{10} protons per pulse. The AGS has been continuously upgraded, conforming to new technology and new experimental requirements. The ~250 m diameter synchrotron ring can now accelerate protons (up to 33 GeV/c, but more reliably at 24 GeV/c), polarized protons (up to 22 GeV/c) and heavy ions (gold ions up to 14.5 GeV/c per nucleon). In the future, the AGS complex will provide an injector system to the Relativistic Heavy Ion Collider (RHIC) currently under construction.

The AGS has recently completed another phase of upgrades—construction of a new rf system and an injector ring, known as the Booster. The Booster increases the proton energy for injection to the AGS from 0.2 GeV to 1.5 GeV, and quadruples the AGS intensity by using four bunches in one AGS acceleration cycle. The new intensity record of 6.3×10^{13} protons per pulse was reached during the 1995 proton run and exceeds the design goal. The 24 GeV/c primary proton beam is delivered about every 3.2–3.6 seconds over 1.2–1.6 second-long "slow" extraction, assuring about 40% duty cycle. Essentially, the entire *new* AGS intensity is necessary to support adequately the kaon program.

4.2 The $K_L^0 \rightarrow \mu^\pm e^\mp$ Experiment (E871)

The primary goal of E871[37,38] is to search for separate lepton-flavor violation in the decay $K_L^0 \rightarrow \mu^\pm e^\mp$. The two-arm spectrometer is optimized for this search and should reach a single-event sensitivity of 10^{-12}, probing new forces in nature in the 200 TeV mass range. Other decays, such as $K_L^0 \rightarrow \mu^+\mu^-$ and

$K_L^0 \rightarrow e^+e^-$, can also be studied. For the former, a sample of several thousand decays is expected to be recorded. For the latter, predicted by the Standard Model to occur at $\sim 3 \times 10^{-12}$ level, the experiment should detect the first few events ever.

The main source of background for $K_L^0 \rightarrow \mu^\pm e^\mp$ is the copious K_{e3} ($K_L^0 \rightarrow \pi^\pm e^\mp \nu_e$) decay which has a branching ratio $B(K_L^0 \rightarrow \pi^\pm e^\mp \nu_e) = 38.7\%$.[18] If this decay proceeds in a way that the neutrino has very little energy in the laboratory frame, and the pion either decays $\pi \rightarrow \mu\nu_\mu$ or is misidentified as a muon, the *observed* final state contains a muon-electron pair. Theoretically, if experimental resolutions are ignored, the reconstructed invariant mass of the muon-electron pair, $M_{\mu e}$, could approach the mass of the parent kaon to within 8.4 MeV/c^2. Equally dangerous is the case in $K_L^0 \rightarrow \pi^\pm e^\mp \nu_e$ decay when *both* charged particles are misidentified. If the pion is mistaken by an electron, and the electron is mistaken for a pion, the invariant mass of this doubly-misidentified pair is not bounded by the kaon mass, M_K, and could exceed M_K if momentum of a pion is much larger than momentum of an electron. Analogously, the background for $K_L^0 \rightarrow \mu^+\mu^-$ originates in (a) $K_L^0 \rightarrow \pi^\pm e^\mp \nu_e$ if $\pi \rightarrow \mu\nu_\mu$ and the electron is mistaken for a muon or (b) in $K_{\mu 3}$ decays $K_L^0 \rightarrow \pi^\pm \mu^\mp \nu_\mu$ ($B(K_L^0 \rightarrow \pi^\pm \mu^\mp \nu_\mu) = 27.0\%$[18]) if the pion decays or is mistaken for a muon. The background for $K_L^0 \rightarrow e^+e^-$ is due to $K_L^0 \rightarrow \pi^\pm e^\mp \nu_e$ with one of the pions mistaken for an electron, and ironically, from other rare processes: $K_L^0 \rightarrow e^+e^-e^+e^-$ and $K_L^0 \rightarrow e^+e^-\gamma$ which occur at[18] $(9.1 \pm 0.5) \times 10^{-6}$ and $(3.9 \pm 0.7) \times 10^{-8}$, respectively.

To reach 10^{-12}, an experiment has to assure efficient background rejection. The essential requirements on an apparatus are good kinematics reconstruction and reliable particle identification, both to be accomplished in a high rate environment. The spectrometer for E871, shown in Fig. 12, has been designed to satisfy such demands.[37] E871 is the successor to E791, which set a 90% C.L. upper limit for $B(K_L^0 \rightarrow \mu^\pm e^\mp)$ at 3.3×10^{-11} (see Table 1). The experiment has been fully commissioned for the 1995 AGS running cycle. The main features of the new apparatus, shown in Fig. 12, are:

- two magnets for momentum-finding and providing "parallelism" (see below) of two-body decays,
- a beam-stop placed in the first magnet to absorb the entire neutral beam,

- redundant finely-segmented fast straw drift chambers and conventional drift chambers in regions of low rate,
- redundant particle identification of muons and electrons,
- multilevel trigger with fast on-line reconstruction, and
- fast custom-designed massively parallel data acquisition system.

The neutral kaon beam is produced by an intense primary ~24 GeV/c beam of about 1.7×10^{13} protons delivered onto a (1.5 interaction-length) platinum target. A system of sweeping magnets and collimators placed at 3.75° with respect to the proton beam direction forms the neutral beam with mostly neutrons and kaons. Particles emerging from the 11 m long evacuated "decay volume" are tracked and identified in a two-magnet spectrometer. The neutral beam is absorbed in the beam stop specially designed and tested for this configuration.[39] The strengths of the magnetic fields (+440 MeV/c and −230 MeV/c) imposes that trajectories of two-body kaon decays emerge nearly parallel downstream of the second magnet. Such an arrangement simplifies triggering and provides the first rejection stage of the three-body decays. The intense primary proton beam produces $\sim 2 \times 10^7$ K_L decays per the AGS pulse[37] resulting in high hit rates in the upstream straw drift chambers. In addition, rates in chambers result from leakage of low-energy particles (charged, neutrons, and photons) from the beam-stop. The beam-stop shields the downstream part of the spectrometer, where the rates are substantially reduced. This minimizes the probability of pattern-recognition or particle-identification errors.

Particle identification of electrons and muons uses redundancy to minimize errors. Particles to be identified as electrons are required to have hits in a threshold Cherenkov counter filled with hydrogen (muons with momenta larger than 6.3 GeV/c and pions with momenta above 8.3 GeV/c can also produce Cherenkov radiation). In addition, electron candidates are identified with a segmented two-layer lead-glass calorimeter. Total energy deposited in lead-glass has to be consistent with the momentum measured from the reconstructed trajectory in the magnetic field. The $K_L^0 \rightarrow \pi^{\pm} e^{\mp} \nu_e$ and $K_L^0 \rightarrow \pi^{\pm} \mu^{\mp} \nu_\mu$ modes provide a constant calibration source for this system. Muons are identified as particles penetrating (the iron slabs in the upstream part and marble slabs in the remaining part of) the "muon range finder." Scintillation hodoscopes are located at depths of material corresponding to muon momenta of 0.75, 0.94, 1.4, 2.8, and 5.8 GeV/c.

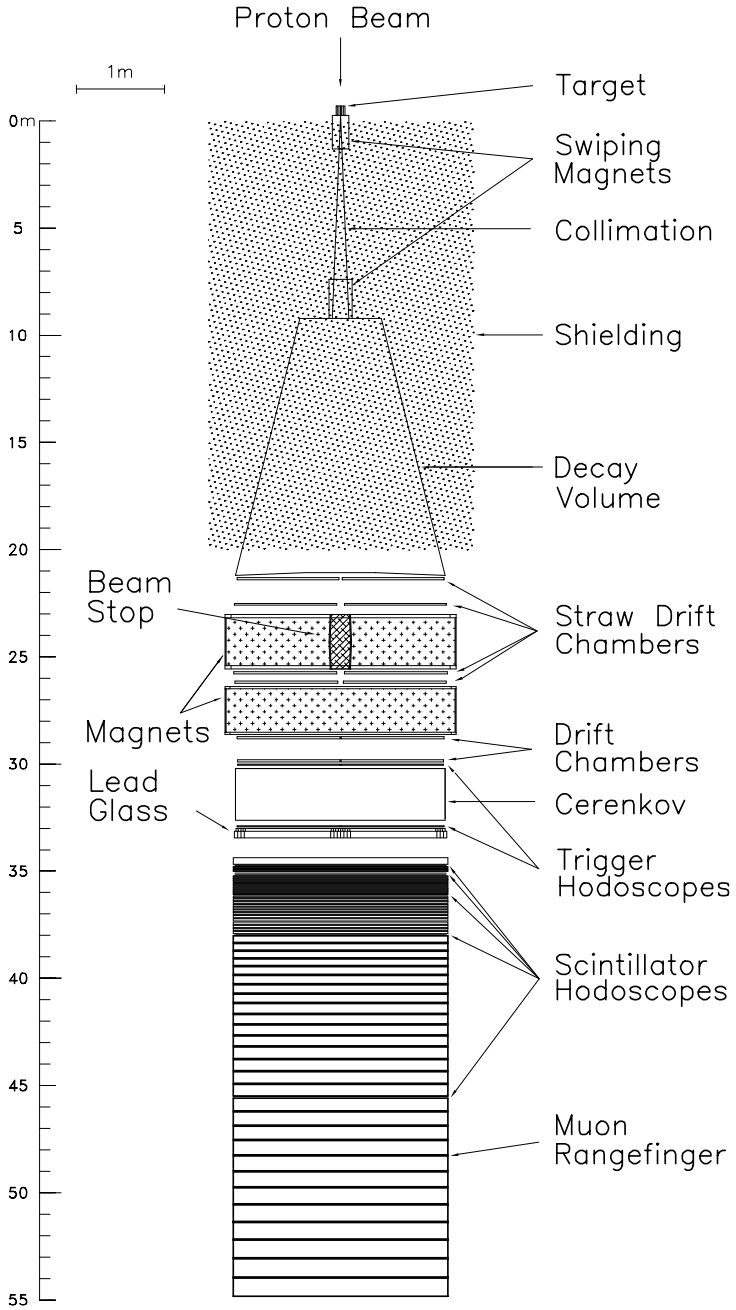


Fig. 12. BNL E871 apparatus.

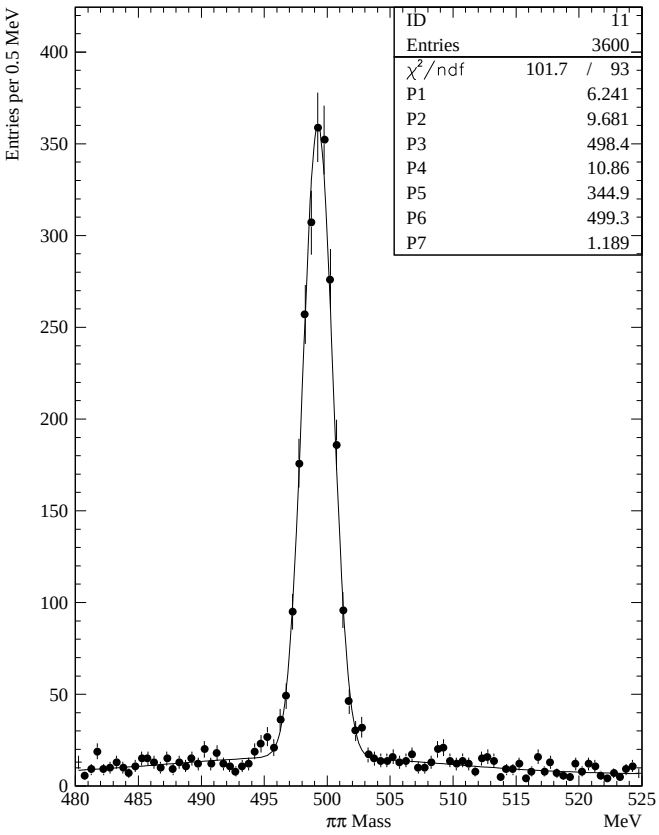


Fig. 13. Preliminary results from BNL E871 show the $K_L^0 \rightarrow \pi^+\pi^-$ invariant mass distribution with resolution of about 1.2 MeV/c^2.

Proportional drift tubes are placed throughout the range stack providing 5% measurement resolution of the penetration range. The range of the particle, its timing, and trajectory have to be consistent with the kinematics determined from the magnetic spectrometer. The power of the background suppression comes also from kinematics: quality of reconstructed trajectories, consistency of momenta measured in two magnets, quality of the vertex, missing transverse momentum, and reconstructed invariant mass of the final-state particles.

The custom-designed data acquisition system used by E871 is the same as for E791.[40] Briefly, fast conversion-time front-end ADC's, TDC's, and latches are sparsely read out into dual-port memories residing in VME crates and serviced by Silicon Graphics V-35 processors. Once the memory buffer is filled with trigger events, a software algorithm reconstructing kinematics is applied. Events passing loose invariant-mass and angular requirements are retained as candidate events. A certain fraction of minimum bias events is unconditionally retained and ulti-

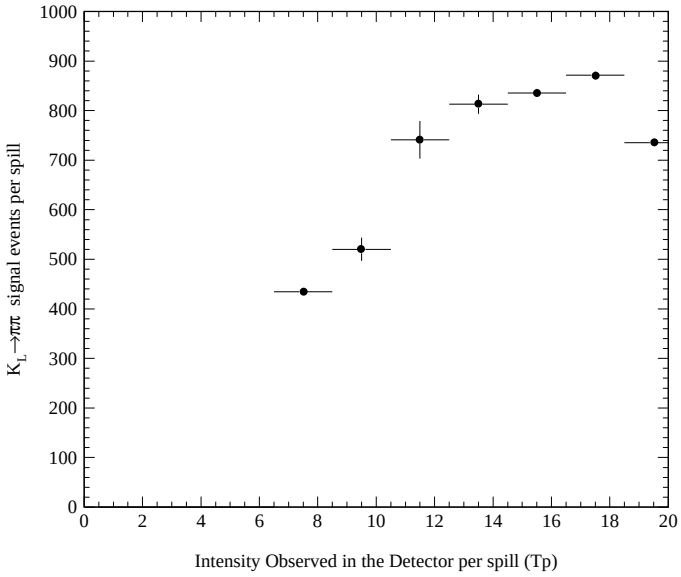


Fig. 14. The number of reconstructed $K_L^0 \rightarrow \pi^+\pi^-$ events in E871 as a function of the beam intensity[42] measured in units of Tp (1 Tp = 10^{12} protons) on target per spill. The experiment typically operates ~100% live at ~17 Tp.

mately used for the sensitivity measure of the experiment. With 17×10^{12} protons on target, there are about 750 k nonparallel triggers formed by the coincidences in trigger hodoscopes. The requirement of parallelism (i.e., position correlations between scintillator slats hit in the upstream and downstream trigger banks) reduces this number to about 200 k. Further position-correlated coincidences with the Cherenkov counter and/or muon hodoscopes form about 12 k triggers which are read out during the AGS spill of 1.2–1.6 seconds. About 400 events per spill are retained and written to tape as either candidate events, minimum bias events, or calibration events. E871 collected data between January and mid-June of the 1995 AGS and the offline analysis is currently under way. Preliminary results, as shown in Figs. 13 and 14, indicate that the apparatus operates at the expected level.[41]

4.3 The $K^+ \rightarrow \pi^+\mu^+e^-$ Experiment (E865)

Experiment 865[42,43] is an upgrade of the previous experiments E777 and E851. The new spectrometer has been designed with the goal of reaching 10^{-12} single-event sensitivity in the channel $K^+ \rightarrow \pi^+\mu^+e^-$. Observation of such a decay would explicitly demonstrate lepton-flavor violation. Due to the helicity structure of the hadronic currents, this mode is sensitive to scalar or vector interactions, while $K_L^0 \rightarrow \mu^\pm e^\mp$ process tests pseudoscalar or axial-vector forces making the two searches complementary. As argued earlier, $K^+ \rightarrow \pi^+\mu^+e^-$ probes the mass scale of 86 TeV (Eq. 5) at the 10^{-12} branching fraction.

The background to $K^+ \rightarrow \pi^+\mu^+e^-$ process comes primarily from $K^+ \rightarrow \pi^+\pi^-\pi^0$ and $K^+ \rightarrow \pi^+\pi^0$. If in the τ-decay ($K^+ \rightarrow \pi^+\pi^-\pi^0$) one of the π^+ mesons decays in flight ($\pi \rightarrow \mu\nu_\mu$) or is misidentified as a muon, and π^- is misidentified as an electron, the observed final-state particles are π^+, μ^+, and e^-. Similarly, in the $K^+ \rightarrow \pi^+\pi^0$ case, if π^+ decays or is misidentified as a muon, and π° decays via a Dalitz mode $\pi^0 \rightarrow e^+e^-\gamma$, and e^+ is mistaken for π^+, this leads again to the same effective final state which mimics the real $K^+ \rightarrow \pi^+\mu^+e^-$. The latter source hints at the necessity to conduct the search of the μ^+e^- and not the μ^-e^+ pair. This stems from the fact that the K_{e3}^+ decay ($K^+ \rightarrow \pi^0e^+\nu_e$) mode has a branching fraction of $B(K^+ \rightarrow \pi^0e^+\nu_e$) $= 4.82 \pm 0.06)\%$ and thus is the source of positrons at a factor of $\sim$20 higher level than the Dalitz source of electrons. It is worth mentioning that a search for $K^+ \rightarrow \pi^+\mu^-e^+$ conducted in the mid-1970s set an impressive 90% C.L. upper limit for this process at[44] 6.9×10^{-9}.

This charge asymmetry propagates into the detector design. Figure 15 shows the spectrometer designed and commissioned by the E865 collaboration. The key features of the apparatus are:[42,45]

- the left arm of the spectrometer, to which negative particles are deflected by the first magnet, is optimized for an electron identification (or π^-/e^- separation);
- the right arm (with positive particles) is optimized for a muon and pion identification (or e^+/π^+ separation);
- two dipole magnets with p_T kick of $+250\ MeV$ and $-250\ MeV$ provide the charge separation and a momentum measurement; and
- redundant particle identification employs two sequential threshold Cherenkov

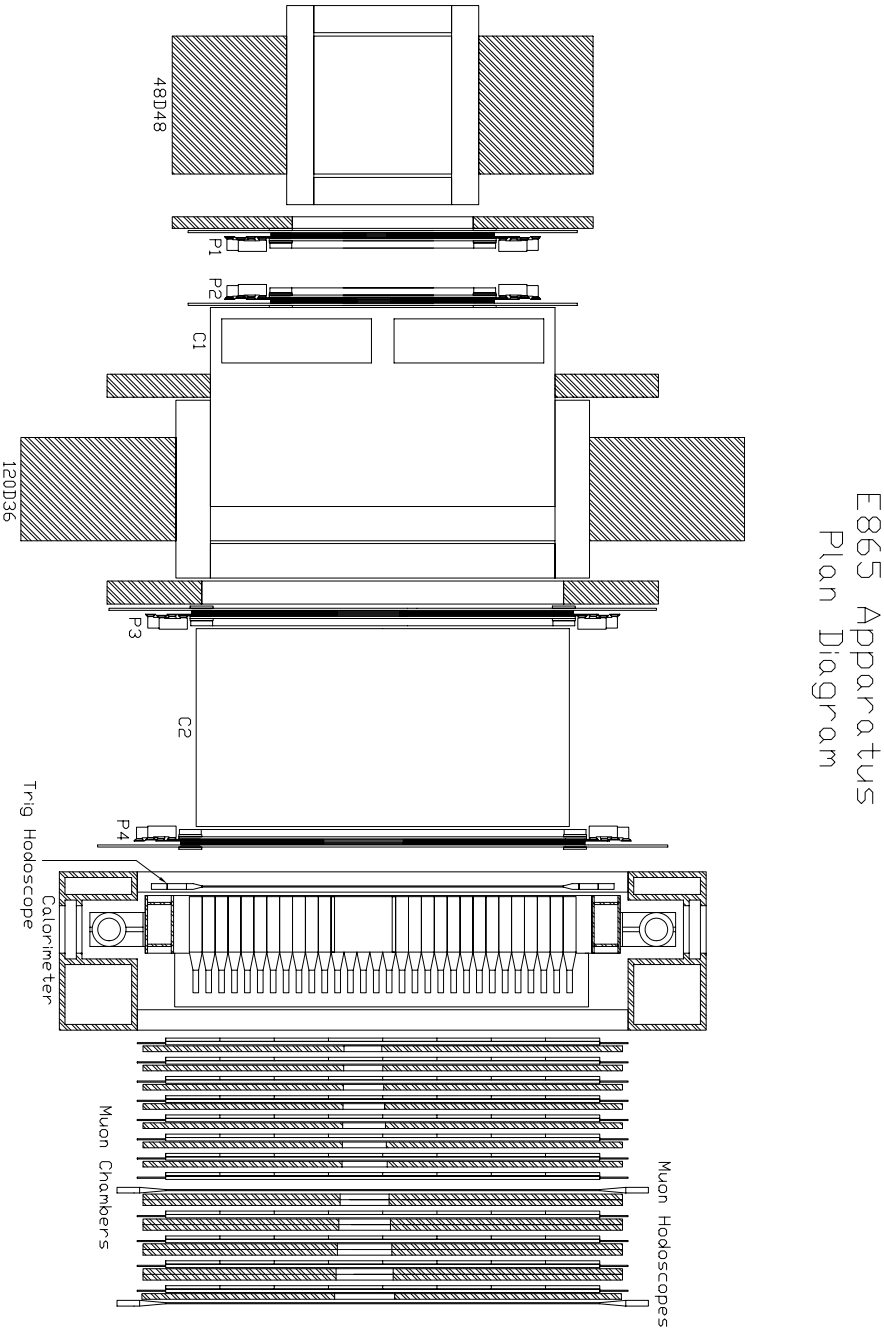


Fig. 15. BNL E865 apparatus.

counters, electromagnetic calorimeter, and muon range stack.

The 6 GeV/c K^+ beam was substantially upgraded for E865. It delivers kaons with about 5% momentum spread. For 1×10^{13} protons on target, the flux of K^+ is about 6×10^7, with about 2×10^9 protons and pions. K^+ decay products pass through the first dipole magnet which deflects negatively charged particles to the left, and positively charged particles to the right. A particle to be identified as an electron is required to induce signals in two left-arm hydrogen-filled Cherenkov counters. In addition, the electromagnetic calorimeter ("shishkebab" configuration of the lead-scintillator tiles with a fiber readout) should reconstruct deposited energy consistent with the momentum measured in the upstream part of the spectrometer, where particles are tracked using proportional wire chambers with resistive Mylar cathode. The right-arm particle identification is designed to discriminate against positrons. The two right-side Cherenkov counters are filled with CH_4 providing the momentum threshold for muons at 3.7 GeV, which is near the upper end of the muon momentum spectrum for $K^+ \rightarrow \pi^+\mu^+e^-$. Further, the electromagnetic calorimeter and muon stack with scintillation hodoscopes and proportional wire chambers provide additional discrimination. The experiment ran with about 2×10^6 three-particle trigger-counter coincidences per spill. Additional coincidences with Cherenkov counters, muon hodoscopes, and PWC's suppressed the trigger rate to about 1000 per spill. About 80% live-time events were read out and processed by a farm of processors which discarded events useless for further offline analysis, leaving about 50% triggers written to tape.

E865 collected data March through mid-June 1995. In addition to lepton-flavor violation, the E865 collaboration plans to study a number of less-rare decay modes which could provide an important input (or tests) of chiral perturbation theory. These modes include $K^+ \rightarrow \pi^+e^+e^-$, $K^+ \rightarrow \pi^+e^+e^-\gamma$, $K^+ \rightarrow \pi^+\mu^+\mu^-$, $K^+ \rightarrow \pi^+\gamma\gamma$, $K^+ \rightarrow \pi^+\pi^0e^+e^-$, $K^+ \rightarrow e^+e^+e^-\nu$, and $K^+ \rightarrow \mu^+e^+e^-\nu$. Figure 16 shows a preliminary result of $K^+ \rightarrow \pi^+e^+e^-$ reconstruction from a fraction of collected data.[45]

4.4 The $K^+ \rightarrow \pi^+\nu\bar{\nu}$ Experiment (E787)

Experiment 787[46] is a continuing effort started in the mid-1980s. The primary goal of the experiment is to make the first observation of the $K^+ \rightarrow \pi^+\nu\bar{\nu}$ decay mode. This GIM-suppressed decay is expected at $\sim 10^{-10}$ level, but the final state

Fig. 16. Preliminary results from E865—a sample of reconstructed events $K^+ \rightarrow \pi^+ e^+ e^-$ from a fraction of the collected data in 1995[46].

Fig. 17. Range and momentum spectra of π^+ from $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and from the major sources of backgrounds.

presents formidable experimental challenges. With well-understood underlying theory and a measured mass of the top quark, the observation of even a few events could be turned into a measurement of the V_{td} of the CKM matrix.

Because of the weak topological constraints, the search for $K^+ \rightarrow \pi^+\nu\bar{\nu}$ is carried with a stopping K^+ beam, and with the limited range of pion momenta above the $K^+ \rightarrow \pi^+\pi^0$ peak. As illustrated in Fig. 17, most of the $K^+ \rightarrow \pi^+\nu\bar{\nu}$ phase-space lies below the $K^+ \rightarrow \pi^+\pi^0$ momentum peak of 205 MeV/c, but the severe background due to π^+ from $K^+ \rightarrow \pi^+\pi^0$ (π^+ can interact with the detector material and shift down its energy) makes this region experimentally difficult at low sensitivities. The main background above the $K^+ \rightarrow \pi^+\pi^0$ peak is due to misidentified muons from $K^+ \rightarrow \mu^+\nu_\mu$ and $K^+ \rightarrow \mu^+\nu_\mu\gamma$, as well as from mismeasured π^+'s from $K^+ \rightarrow \pi^+\pi^0$ if the two photons from π° are missed. In addition, pions in the beam can be mistaken for kaons. Also, misidentified muons and/or protons originating from $K^+n \rightarrow K^\circ p$ and followed by $K_L^0 \rightarrow \mu\nu$ pose a background threat.

The experimental approach undertaken by the E787 collaboration emphasizes:[47]

- redundant determination of the pion kinematics by independent measurements of the pion's momentum, kinetic energy, range, and dE/dx,
- observation and measurement of timing and range relationships of the $\pi \to \mu \to e$ decay sequence, and
- nearly 4π solid-angle photon hermeticity of the apparatus.

The present E787 apparatus is shown in Fig. 18. This is an upgraded version of the spectrometer used in 1989–1991.[48] Two electrostatic separators are used to deliver 800 MeV/c K^+'s. The beam provides three times more kaons than pions. Kaons are identified by Cherenkov and dE/dx counters placed in the beam. Passing the BeO degrader, kaons are stopped in an active target made of scintillation fibers located in the center of the apparatus. The target is surrounded by the central drift chamber covering about 2π solid-angle and used to determine momenta of particles through trajectory curvature in the 1 T solenoidal magnetic field in the apparatus. Further, in the central (barrel) part of the apparatus, a scintillator range stack is used to provide kinetic energy and range measurement of pions. In addition, pions are distinguished from muons by identifying the $\pi \to \mu \to e$ decay sequence using 500 MHz transient digitizers. Finally, photon detectors (lead-scintillator stack in the barrel part, and CsI blocks in the endcap) assure hermeticity to photons.

The thickness of the stack is optimized to catch pions with momentum $214 \leq P_\pi \leq 231$ MeV/c and range $34 \leq R_\pi \leq 40$ cm in the scintillator, and to achieve $\sim$1 cm range resolution. Additional constraints on tagging of transitions $\pi^+ \to \mu^+$ and $\mu^+ \to e^+$ in transient digitizers, beam and photon vetoing bring an overall acceptance for $K^+ \to \pi^+\nu\bar{\nu}$ to only few % level. E787 took data with an upgraded apparatus between January and mid-June 1995. The beam provided about 1.3 million stopping K^+'s per spill. A multilevel trigger reduced this number to about a couple of hundred events which were written to tape. Figure 19 shows an example of improvement in timing and energy resolutions achieved in the 1995 data. Besides the $K^+ \to \pi^+\nu\bar{\nu}$ mode, the collaboration will also study several other topologically and kinematically related kaon decay modes. For example, $K^+ \to \pi^+ X^0$ would give similar signature to $K^+ \to \pi^+\nu\bar{\nu}$ if X^0 is a weakly interacting (new) particle. Motivated by their significance in chiral perturbation theory, and the fact that they are experimentally accessible at the same time as $K^+ \to \pi^+\nu\bar{\nu}$ decays, channels like $K^+ \to \pi^+\mu^+\mu^-$, $K^+ \to \pi^+\pi^0\gamma$,

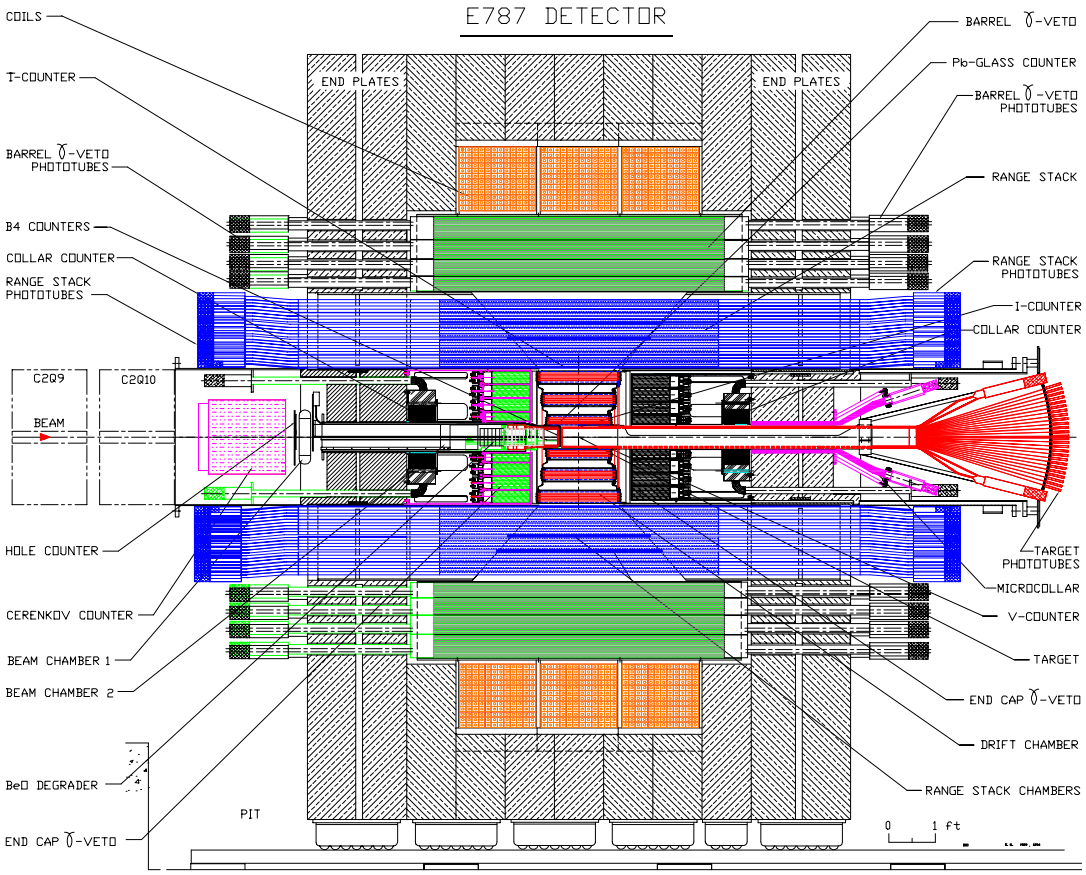
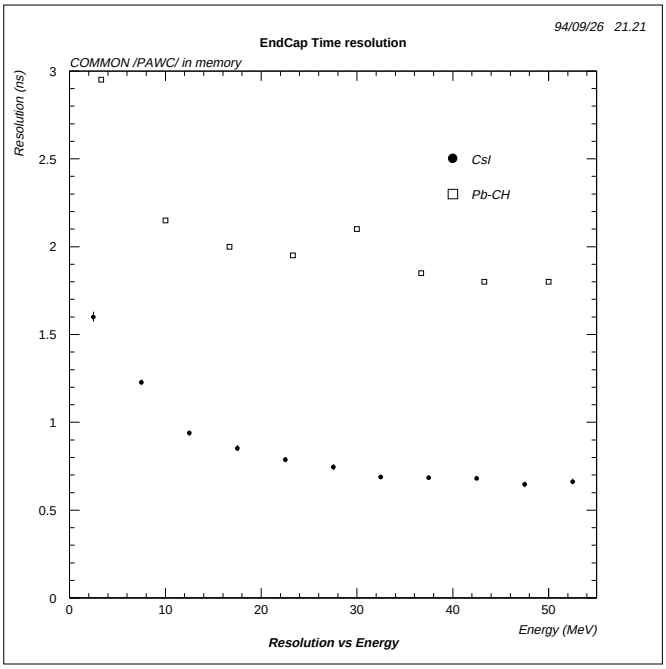
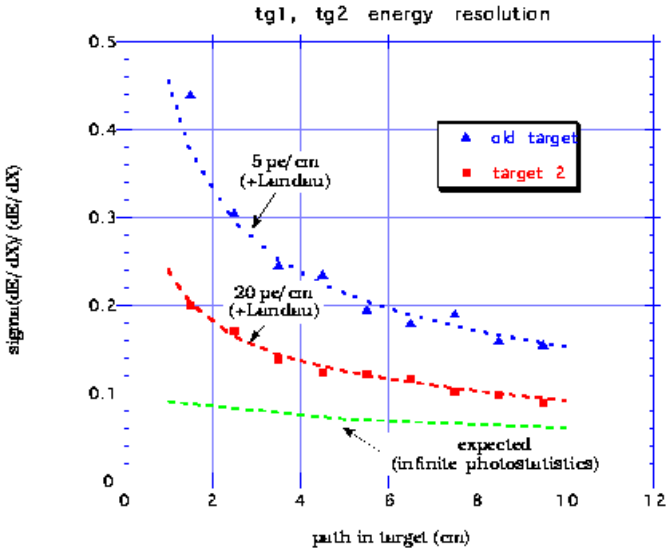


Fig. 18. BNL E787 apparatus.



(a)



(b)

Fig. 19. Improvements of the upgraded E787 apparatus as determined from the 1995 data:[48] (a) improved photon timing of the newly installed CsI blocks (circles) vs. old lead-scintillator stack (squares); (b) improved energy resolutions.

Decay mode	Projected Sensitivity	Experiment
$K_L^0 \rightarrow \mu^{\pm} e^{\mp}$	$\sim 1 \times 10^{-12}$	E871
$K_L^0 \rightarrow e^+ e^-$	$\sim 1 \times 10^{-12}$.	
$K_L^0 \rightarrow \mu^+ \mu^-$	several thousand events	
$K^+ \rightarrow \pi^+ \mu^+ e^-$	$\sim 1 \times 10^{-12}$	E865
$K^+ \rightarrow \pi^+ \nu \bar{\nu}$	$\sim 1 \times 10^{-10}$	E787

Table 4. Projected sensitivities of the three experiments at the AGS.

$K^+ \rightarrow \pi^+ \gamma\gamma$, $K^+ \rightarrow \mu^+ \nu_\mu \gamma$, and $K^+ \rightarrow \pi^0 \mu^+ \nu_\mu \gamma$ will be studied.

5 Summary and Conclusions

We have presented a brief summary of best results and described current experiments at the AGS at BNL in the area of very rare kaon decays. This active program, which was started in the early 1980s, has brought a major advancement in testing "rare physics" within the Standard Model and in searching for processes outside of it. The current round of experiments will probe mass scales unattainable in direct searches and will possibly observe the lowest particle decay branching fraction of any kind. Table 4 summarizes sensitivities expected for the main rare decay modes studied by the three collaborations at BNL.

For the currently scheduled running, Expriment 871 should be able to reach 10^{-12} single event sensitivity in $K_L^0 \rightarrow \mu^{\pm} e^{\mp}$, probing ~200 TeV mass scale for the new interactions. Reaching low branching ratios will allow us to collect several thousands of $K_L^0 \rightarrow \mu^+ \mu^-$ events and substantially improve the branching ratio measurement. In addition, the first observation of $K_L^0 \rightarrow e^+ e^-$ expected at $\sim 3 \times 10^{-12}$ should be possible. Complementarily, the near future of Experiment 865 should bring the ability to probe ~80 TeV mass range in $K^+ \rightarrow \pi^+ \mu^+ e^-$ and study other decays to test and expand the reliability of chiral perturbation theory. With improved capabilities, the collaboration is considering also studying CP asymmetries in $K^+ \rightarrow \pi^+ \pi^- \pi^0$. Experiment 787 plans a continuing program of upgrades and expects ultimately to measure the $B(K^+ \rightarrow \pi^+ \nu \bar{\nu})$ with 20% uncertainty. This will be an important test of the Standard Model and will provide an important input to determining the $ts^* V_{td}$ quantity of the CKM

matrix, shedding more light onto the current picture of the CP violation.

I would like to thank L. Littenberg and M. Zeller for providing useful information, and J. Ritchie for comments on the manuscript.

References

[1] H. Haber and G. L. Kane, Phys. Rep. er and G. L. Kanersus` œßëMΔÌÌ¯e oÖisamissingtransvrersemomentumofthe¡œßëreconstructedΛê
nal-statepair).Π9 ThebSottom
gureshorwstheKû2Ü0ϒRAL

[2] J. C. Pati and A. Salam, Phys. Rev. D**10**, 275 (1974); A. Barroso, G. C. Branco, and M. C. Bento, Phys. Lett. B**134**, 123 (1984).

[3] T. Maehara and T. Yanagida, Prog. Theo. Phys. **60**, 1434 (1979); A. Davidson and K. C. Wali, Phys. Rev. Lett. **26**, 691 (1981); W.-S. Hou and A. Soni, Phys. Rev. Lett. **54**, 2083 (1985).

[4] Z. Y. Zhu, Z. Phys. C**13**, 321 (1982); I. Bars, M. J. Bowick, and K. Freese, Phys. Lett. B **138**, 159 (1984); O. W. Greenberg, R. N. Mohapatra, and S. Nussinov, Phys. Lett. B**148**, 465 (1984); J. C. Pati, Phys. Rev. D**30**, 1144 (1984).

[5] B. A. Campbell *et al.*, Int. J. Mod. Phys. A **2**, 831 (1987).

[6] "Rare K decays," J. L. Ritchie and S. G. Wojcicki, Rev. Mod. Phys. **65**, 1149 (1993); "Rare and radiative kaon decays," L. Littenberg and G. Valencia, Ann. Rev. Nucl. Part. Sci. **43**, 729 (1993); "The search for direct CP violation," B. Winstein and L. Wolfenstein, Rev. Mod. Phys. **65**, 1113 (1993).

[7] A. J. Buras, preprint MPI-PhT/95-30, April 1995.

[8] L. M. Sehgal, preprint PITHA 94/52, November 1994; to appear in *Procedings of The Third Workshop on High Energy Particle Physics*, Madras, Jan. 10-22, 1994.

[9] N. Bilic and B. Guberina, Fortsch. Phys. **42**, 209 (1994).

[10] R. N. Cahn and H. Harari, Nucl. Phys. B **176**, 135 (1980).

[11] D. Rein and L. M. Sehgal, Phys. Rev. D**39**,3325 (1989); M. Lu and M. B. Wise, Phys. Lett. B**324**, 461 (1994).

[12] J. Hagelin and L. Littenberg, Prog. Part. Nucl. Phys. **23**, 1 (1989).

[13] I. Bigi and F. Gabbiani, Nucl. Phys. B**367**, 3 (1991).

[14] G. Buchalla, A. Buras, and M. Harlander, Nucl. Phys. B**349**, 1 (1991).

[15] G. Buchalla, A. Buras, and M. Harlander, Nucl. Phys. B**412**, 106 (1994).

[16] S. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[17] L. M. Sehgal, Phys. Rev. **183**, 1511 (1969); Nuovo Cimento **45**, 785 (1966).

[18] Review of Particle Properties, Phys. Rev. D**50**, 1173-1826 (1994).

[19] A. Buras and M. Harlander, in *Heavy Flavours*, edited by A. Buras and M. Lindner (World Scientific, 1992), p. 58.

[20] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981), Erratum 1772.

[21] G. Belanger and C. Q. Geng, Phys. Rev. D**43**, 140 (1991).

[22] P. Ko, Phys. Rev. D**45**, 174 (1992).

[23] K. Arisaka *et al.*, Phys. Rev. Let. **70**, 1049 (1993).

[24] T. Akagi *et al.*, Phys. Rev. Let. **67**, 2614 (1991).

[25] S. F. Schaffner *et al.*, Phys. Rev. D**39**, 990 (1989).

[26] A. M. Lee *et al.*, Phys. Rev. Let. **64**, 165 (1990).

[27] S. Adler *et al.*, BNL-62327, PRINCETON/HEP/95-8, TRI-PP-95-83.

[28] A. P. Heinson *et al.*, Phys. Rev. D**51**, 985 (1995).

[29] T. Akagi *et al.*, Phys. Rev. Let. **67**, 2618 (1991).

[30] K. Arisaka *et al.*, Phys. Rev. Let. **71**, 3910 (1993).

[31] E. Jastrzembski *et al.*, Phys. Rev. Lett. **61**, 2300 (1988).

[32] C. Alliegro *et al.*, Phys. Rev. Let. **68**, 278 (1992).

[33] M. S. Atiya *et al.*, Phys. Rev. Let. **70**, 2521 (1993).

[34] M. S. Atiya *et al.*, Phys. Rev. Let. **63**, 2177 (1989).

[35] A. Deshpande *et al.*, Phys. Rev. Let. **71**, 27 (1993).

[36] K. S. McFarland *et al.*, Phys. Rev. Let. **71**, 31 (1993).

[37] A. Heinson *et al.*, AGS Proposal 871, September 1990.

[38] E871 is a collaboration of *The University of California at Irvine*— M. Bachman, D. Connor, P. de Cecco, N. Kanematsu, R. Lee, W. Molzon; *Stanford Univeristy*— K. Ecklund, K. Hartman, M. Hebert, G. Irwin, D. Ouimette, M. Pommot-Maia, S. Wojcicki; *The University of Texas at Austin*— D. Ambrose, S. Graessle, M. Hamela, G. Hoffmann, K. Lang, J. McDonough, A. Milder, P. Riley, J. Ritchie, V. Vassilakopoulos, B. Ware, S. Worm; *College of William & Mary*— M. Eckause, D. Hancock, C. Hoff, J. Kane, Y. Kuang, R. Martin, R. Welsh, E. Wolin; *University of Richmond*—P. Rubin.

[39] S. Worm, Ph. D. Thesis, University of Texas, Austin, June 1995.

[40] R. D. Cousins *et al.*, Nucl. Instrum. Methods A **277**, 517 (1989).

[41] M. Hebert, talk presented at the AGS Users Meeting, BNL, June 15, 1995.

[42] AGS Proposal 865, May 1990.

[43] E865 is a collaboration of *Brookhaven National Laboratory*— D. Lazarus, L. Leipuner, H. Ma, P. Rehak; *Institute for Nuclear Research, Moscow*— G. S. Atoyan, V. V. Isakov, A. A. Poblaguyev, V. Postoev, A. Proskurjakov; *Joint Institute for Nuclear Research, Dubna*—B. Z. Zalikhanov; *University of New Mexico*— B. Bassalleck, S. Eilerts, H. Fisher, J. Lowe, D. Wolfe; *Paul Scherrer Institute*— W. Djordjadze, J. Egger, W. D. Herold, H. Kaspar, J. Missimer; *University of Basel*— G. Backendstoss, W. Menzel, H. Weyer; *University of Pittsburgh*— V. Atoyan, K. Baker (Hampton Univ.), E. Battiste, M. Gach, C. Felder, D. E. Krauss, M. Shubin, J. A. Thompson; *High Energy Physics Institute, Tbilisi State University*— N. Amaglobeli, Y. Bagaturia, D. Mazavia, T. Sachelashvili; *Yale University*— R. Appel, D. Bergman, S. Dhawan, H. Do, J. Lozano (Univ. of Connecticut), M. Zeller; *University of Zurich*—S. Pislak, P. Truoel.

[44] A. M. Diamant-Berger *et al.*, Phys. Lett. B**62**, 485 (1976).

[45] H. Ma, talk presented at the AGS Users Meeting, BNL, June 15, 1995.

[46] E787 is a collaboration of *Brookhaven National Laboratory*— S. Adler, M. S. Atiya, I.-H. Chiang, M. Diwan, J. S. Frank, J. S. Haggerty, S. H. Kettell, T. F. Kycia, K. K. Li, L. S. Littenberg, A. J. Stevens, R. C. Strand, C. H. Witzig; *INS, University of Tokyo* — K. Ino, T. Komatsubara, N. Muramatsu, H. Okuno, S. Sugimoto, K. Ukai; *KEK, National Laboratory for High Energy Physics* — T. Inagaki, S. Kabe, M. Kobayashi, Y. Kuno,

T.Sato, T. Shinkawa, Y. Yoshimura; *Osaka University* —Y. Kishi, T. Nakano; *Princeton University* — M. Ardebili, M. Convery, M. M. Ito, D. R. Marlow, R. McPherson, P. D. Meyers, F. C. Shoemaker, A. J. S. Smith, B. Stone; *TRIUMF* — M. Aoki, E. W. Blackmore, D. A. Bryman, P. Kitching, A. Konaka, J. A. Macdonald, J. Mildenberger, T. Numao, J.-M. Poutissou, R. Poutissou, G. Redlinger, J. Roy, M. Rozon, R. Soluk, A. S. Turcot.

[47] G. Redlinger, talk presented at the AGS Users Meeting, BNL, June 15, 1995; T. K. Komatsubara, talk presented at the AGS Users Meeting, BNL, June 15, 1995.

[48] M. S. Atiya *et al.*, Nucl. Instrum. Methods A **321**, 129 (1992).

An Overview of the Fermilab Fixed Target Program

P. Kasper

Fermi National Accelerator Laboratory
P.O. Box 500
Batavia, IL 60510

No paper was received from the author in time for printing.

RECENT RESULTS FROM THE CMD-2 DETECTOR AT THE VEPP-2M COLLIDER

E. P. Solodov

Budker Institute of Nuclear Physics, Novosibirsk, 630090, Russia

Representing the CMD-2 Collaboration

ABSTRACT

The general-purpose detector CMD-2 is taking data at the Novosibirsk VEPP-2M e^+e^- collider in the energy range 360–1400 MeV, with luminosity of $\approx 5.0 \times 10^{30}$ cm^{-2} s^{-1} for the ϕ resonance region. Data from ≈ 1500 nb^{-1} of integrated luminosity around 1.02 GeV and ≈ 500 nb^{-1} in the 600–1000 MeV range have been collected and preliminary analyses performed. We present progress in studies of the ϕ meson and $K_S K_L$ systems:

(a) measurement of the ϕ meson parameters;

(b) searches for ϕ rare decays. The new upper limits $B(\phi \to \eta'\gamma) < 2.4 \times 10^{-4}$, $B(\phi \to \pi^+\pi^-\pi^+\pi^-) < 1.0 \times 10^{-4}$, and $B(\phi \to f_0\gamma) < 8 \times 10^{-4}$ have been obtained;

(c) the study of the K_L interactions in the CsI calorimeter;

(d) with the help of 32,340 tagged K_S, the semi-rare decay of $K_S \to \pi^+\pi^-\gamma$ has been observed with a branching ratio of $(1.82 \pm 0.49) \times 10^{-3}$; and

(e) selection of events with $K_S K_L$ coupled decays and interactions. The regeneration cross section of the low momenta K_L was found to be $\sigma^{Be}_{reg} = 63 \pm 19$ mb.

Data from the 600–1000 MeV energy range are used for high-accuracy measurement of the e^+e^- annihilation cross section, and the preliminary analysis is presented in this paper.

©1995 by E. P. Solodov.

1 Introduction

The VEPP-2M collider at the Budker Institute of Nuclear Physics in Novosibirsk, Russia, shown in Fig. 1, covers the center-of-mass energy range from the two-pion threshold up to 1400 MeV (Ref. 1). Experiments at this collider yielded a number of important results in e^+e^- physics, including the most precise pion form factor measurements[2] and studies of the ϕ, ω, and ρ meson decays.[3,4] During 1988-92, it was upgraded to allow higher positron currents and injection of the electron and positron beams directly at the beam energy, rather than at lower energies and acceleration after injection. After installation of the new booster, VEPP-2M has peak luminosity $L \approx 5.0 \times 10^{30}$ cm^{-2}s^{-1} at 40 mA per beam at the ϕ center-of-mass energy.

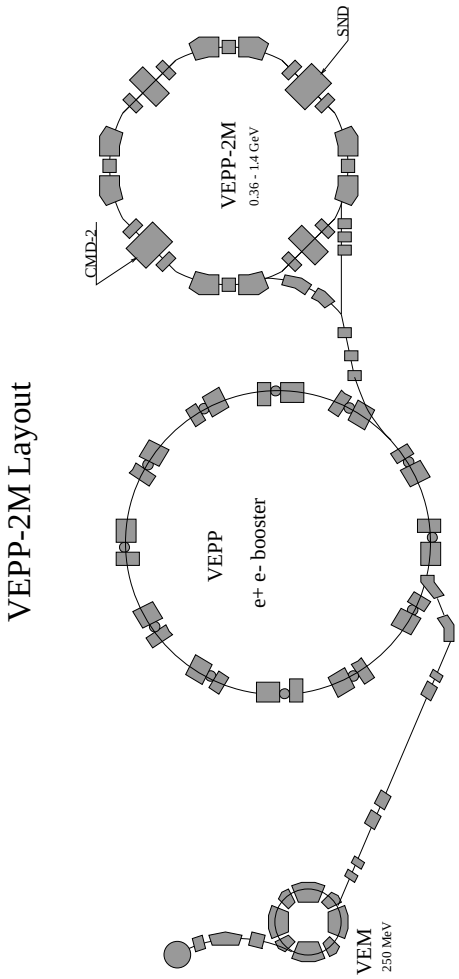


Fig. 1. The layout of the VEPP-2M collider at the Budker Institute of Nuclear Physics in Novosibirsk.

In this paper, we present some preliminary results from two energy regions: one is the relatively narrow region around the ϕ meson resonance, and the second is the region from 600 to 1000 MeV, essential for the measurements of the total hadron cross section in the VEPP-2M energy range.

2 Physics Motivation

The physics program of the CMD-2 detector is very rich; below we consider only some of the many aspects of the ϕ meson, and the total cross section of the e^+e^- annihilation into hadrons which will be studied.

As realized at the very early stages of the ϕ meson studies at colliding beam machines, K_SK_L pairs ($\approx 34\%$ of all ϕ decays) may be used as a new source for observing CP and CPT violation. These suggestions, including studies of quantum mechanical correlations, were discussed in papers[5,6] when the electron-positron collider at Novosibirsk VEPP-2M was under construction. The coupled decays of the K_SK_L mesons will allow the demonstration of the quantum mechanical correlations of the two particle decays (Einstein-Podolosky-Rosen paradox).[7]

The idea of constructing a more intensive source of ϕ mesons has been discussed by many authors.[8,9] The flux of events at these so-called "ϕ-factories," now under construction,[10,11] will provide an opportunity to make new precise measurements of a possible direct component in the decay of the $K_L \to \pi^+\pi^-,\ \pi^0\pi^0$ (ϵ'/ϵ), as well as the observation of the CP-violating three pion decays of the K_S for the first time. The study of the oscillations in the joint decay distributions could give information about real and imaginary parts of any CPT-violating amplitude.

At the VEPP-2M collider, which could be considered as a pre-ϕ-factory, with the CMD-2 detector, we have been proceeding step-by-step to prepare for work at the ϕ-factory which is now under construction. Studies of upgraded detectors and accelerators are in progress, including an intermediate $\approx 10^{32}$ luminosity collider for investigating the use of round beams, an important ingredient in the planned Novosibirsk ϕ-factory project.[12,13]

With the CMD-2 detector, the neutral kaons from ϕ decays are under study, and the coupled K_SK_L decays have been observed for the first time. The attempt to select the $K_L \to \pi^+\pi^-$ decays again emphasized problems with the semileptonic decay mode background as well as a high level of neutral kaon nuclear interactions, including regeneration of K_L into K_S. The opening of this kinematic region for the neutral kaon interactions study has been an additional argument for the construction of ϕ-factories, and we anticipate that the results obtained from the

data now in hand will be important in planning for ϕ-factory detectors and for physics strategies.

A possible problem with a measurement of ϵ'/ϵ at the level of 10^{-4}–10^{-5} would be an admixture of C = +1 into the final state, giving a component of $K_S K_S$ instead of the desired $K_S K_L$. Although efficient experimental cuts can reduce the effects of such an admixture,[14,15] a component as large as 5×10^{-5} of a C = +1 state would give a dominant contribution to the uncertainty of ϵ'/ϵ at the level of the planned ϕ-factory experiments.[15] The contamination from such a C-even K-$\overline{K}$ mode has been estimated by several authors as giving generally lower values,[8,16–18] but there are no experiments confirming these results.

The decay of the $\phi \to f_0\ \gamma$ with f_0 decaying to two kaons is too small to be seen at the VEPP-2M collider, and we hope to study the decay of $\phi \to f_0\ \gamma$ with a subsequent decay of the f_0 to two charged pions.[19] The two charged pion decay mode can be related to the two kaon decay and a limit on the C-even two kaon final state may be found. Estimates for the branching ratio of $\phi \to f_0\gamma$ range from very small to as high as 2.5 x 10^{-4} (Refs. 16–18).

The study of the f_0 is interesting by itself. The 20% decay probability into a two kaon final state seems puzzlingly high if f_0 is a member of the $S = 0$ and $I = 0$ meson nonet. Various explanations for this large coupling to kaons have been advanced,[16–18,20] including the idea that f_0 is really made of four quarks, with a "hidden strangeness" component: $(f_0 = s\overline{s}(u\ \overline{u}\ +\ d\ \overline{d})/\sqrt{2})$, or that it may be a K-$\overline{K}$ molecule. A limit from VEPP-2M will help to distinguish between these different possibilities.

With expected high luminosity at the ϕ-factories, rare decay modes of ϕ can be measured with high accuracy. For example, a measurement of the B($\phi \to \eta\prime\gamma$) would give important information about quark structure of light mesons and possible contributions from gluonium states (if any). Our new data obtained with the present statistics already improve upper limits for this process, as well as for $\phi \to \pi^+\pi^-\pi^+\pi^-$ and $\phi \to f_0\gamma$.

The data from the low-energy region from two pion production threshold up to 1400 MeV (maximum energy provided by VEPP-2M) are important both for the search of rare decays of the light vector mesons and for the calculation of the dispersion integral that relates the cross section of e^+e^- annihilation into hadrons to the value of the hadronic vacuum polarization. This value plays an important role in the interpretation of the fundamental Standard Model parameters and the evaluation of the anomalous magnetic moment of the muon,[21,22] which will be measured in the E821 experiment at BNL with an extremely high precision of

0.35 ppm. To evaluate the contribution of the hadronic vacuum polarization to the muon g-2 with the same accuracy, a systematic error in hadronic cross section should be less than 0.5%, because the total hadronic muon contribution to g-2 was recently re-evaluated and was found to be 72 ± 1.6 ppm (Ref. 23). The main part of this contribution comes from the energy range which is provided by the VEPP-2M collider. The data on the cross section $e^+e^- \to \pi^+\pi^-$ in the energy range between ρ and Φ mesons were obtained with the CMD-2 detector, and a preliminary analysis is presented in this paper.

3 The CMD-2 Detector

The CMD-2 detector has been described in more detail elsewhere.[6,24] The main systems of the detector are shown in Fig. 2.

The CsI barrel calorimeter with 6 x 6 x 15 cm^3 crystal size is placed outside a 0.4 r.l. superconducting solenoid with a 1 Tesla azimuthally symmetric magnetic field. The endcap calorimeter is made of $2.5 \times 2.5 \times 15$ cm^3 BGO crystals and was not installed for the data presented here. The drift chamber inside the solenoid has about 250 μ resolution transverse to the beam and 0.5–0.6 cm longitudinally. The muon range system uses the streamer tubes and has 1–3 cm spatial resolution.

The collected sample of the Bhabha events was used for the calibration and determination of the reconstruction efficiency in the drift chamber and in the calorimeter. A momentum resolution of 6–8% for 500 MeV/c charged particles and energy resolution of about 10% for gammas in the CsI calorimeter have been obtained.

The integrated luminosity collected in 1992–1993 at ϕ was mostly used for the detector study and software development. Not all detector systems were running properly, and data presented here are still preliminary—better detector understanding and better reconstruction programs available now will give results with less systematic errors.

About 7.2×10^7 triggers were recorded at the ϕ meson region. The total integrated luminosity, determined by selection of Bhabha events, was found to be 1500 nb^{-1}.

The largest part of the integrated luminosity ($\approx$ 1200 nb^{-1} in the 14 energy points around ϕ mass) has been collected during the 1993 summer runs and was used for studies of rare decay modes of ϕ, coupled decays in the K_SK_L system, and nuclear interactions of neutral kaons.

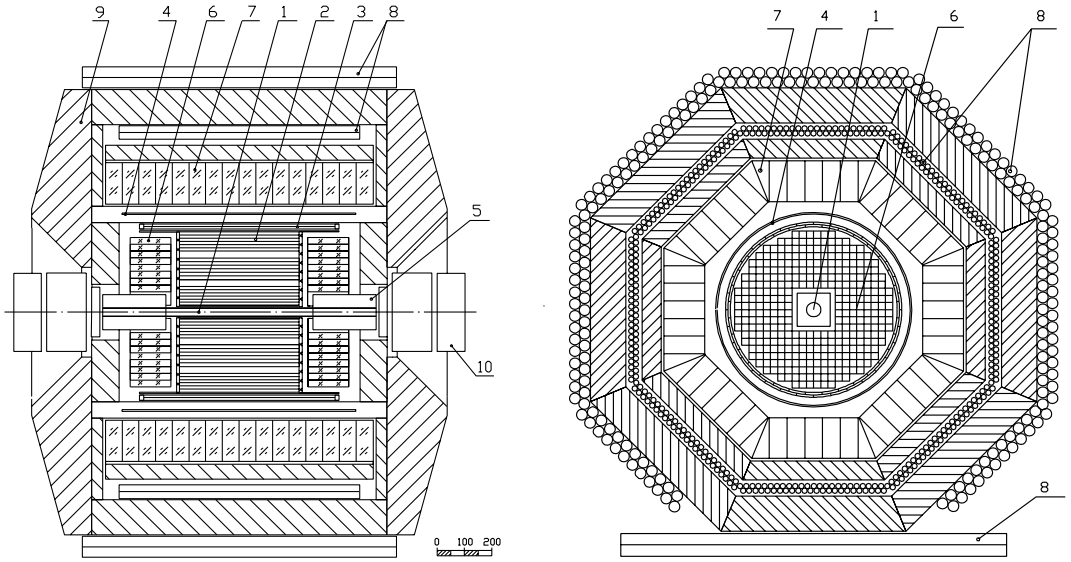


Fig. 2. Horizontal and vertical cross sections of the CMD-2 detector. (1) vacuum chamber; (2) drift chamber; (3) Z-chamber; (4) main solenoid; (5) compensating solenoid; (6) storage ring lenses; (7) calorimeter; (8) muon range system; and (9) magnet yoke.

The 1994–1995 runs were dedicated mostly to measurements of the total hadronic cross section at the energies below ϕ resonance. The integrated luminosity about ≈ 500 nb^{-1} was collected; the experimental conditions are described below.

4 ϕ Meson Parameters

The main branching ratios of the ϕ have been measured using ≈ 300 nb^{-1} of integrated luminosity, collected in 1992. It was the first time when four major decay modes of ϕ were measured in one experiment. The event selection and other details may be found in Ref. 25. The following results were obtained:

$m_\phi = 1019.380 \pm 0.034 \pm 0.048$ MeV,
$\Gamma_{tot} = 4.409 \pm 0.086 \pm 0.020$ MeV,
$\sigma(\phi \to K^+K^-) = 1993 \pm 65 \pm 82$ nb,
$\sigma(\phi \to K_SK_L) = 1360 \pm 25 \pm 49$ nb,
$\sigma(\phi \to 3\pi) = 656 \pm 24 \pm 30$ nb ,
$\sigma(\phi \to \eta\gamma) = 47.9 \pm 3.5 \pm 3.2$ nb, and
$\delta_{\omega-\phi} = (147 \pm 16)^o$.

The first error represents the uncertainty, and the second the systematic uncertainty. The relative phase of $\omega - \phi$ mixing in the three pion channel is in good agreement with the most precise measurement presented in Ref. 4, where $\delta_\phi = (155 \pm 15)^o$. The experimental cross sections of the ϕ production in the different modes, together with fit functions, are shown in Fig. 3.

All the major decay modes were simultaneously measured in one experiment; therefore, the branching ratios can be obtained as ratios of integrals over excitation curves independently of the width of the ϕ. The uncertainties due to the luminosity measurements are:

B$(\phi \to K^+K^-) = 49.1 \pm 1.2\%$,
B$(\phi \to K_SK_L) = 33.5 \pm 1.0\%$,
B$(\phi \to 3\pi) = 16.2 \pm 0.8\%$,
B$(\phi \to \eta\gamma) = 1.18 \pm 0.11\%$.

The electron width of the ϕ and its branching ratio to e^+e^- can also be calculated independently and were found to be

$\Gamma_{ee} = 1.27 \pm 0.05$ keV,
B($\phi \to$ee) $= (2.87 \pm 0.09) \times 10^{-4}$.

All results are consistent with Particle Data Group values.[29]

Here, we note that in all parameters, systematic errors dominate, and using all available statistics will improve these results only after systematic errors have been studied and there is better understanding of the detector.

5 Study of $\phi \to \eta\gamma$ and Search for $\eta\prime\gamma$

The decay of $\phi \to \eta\gamma$ was previously observed in neutral modes ($\eta \to \gamma\gamma$, $\eta \to 3\pi^0$) only. Detector CMD–2 gives the possibility to study $\phi \to \eta\gamma$ decay in the channel with charged particles, when η decays into $\pi^+\pi^-\pi^0$. So, after $\pi^0 \to \gamma\gamma$ decay, the final state consists of two charged pions and three photons. Two photons in the final state are from π^0; the third one has the maximum energy of all three—362 MeV at the ϕ meson peak.

We select $\eta\gamma$ events using the information about momenta and angles from the Drift Chamber for both charged particles and about an angle from the CsI calorimeter for a primary photon assuming that other photons are from π^0. The reconstructed invariant mass of three pions $M_{\pi^+\pi^-\pi^0}$ is the basic parameter we use to study the decay $\phi \to \eta\gamma$, and the distribution over it should have a peak around $M_\eta = 547.45$ MeV.

The distribution over $M_{\pi^+\pi^-\pi^0}$ for all 1993 ϕ meson data after some simple cuts is presented in Fig. 4. These distributions were used to get the numbers of $\eta\gamma$ events for the different beam energies.

The calculated cross section $\sigma_{e^+e^-\to\phi\to\eta\gamma}$ with a fit function is presented in Fig. 4. Using the electron width of ϕ from Ref. 29, the $Br(\phi \to \eta\gamma)$ was found to be:

$$Br(\phi \to \eta\gamma) \;=\; (1.12 \;\pm\; 0.06 \;\pm\; 0.15)\%.$$

This result is preliminary because the work on efficiency determination is not yet complete, and we hope to significantly decrease the systematic errors.

The decay $\phi \to \eta'\gamma$ was searched in the mode, where η' decays into $\pi^+\pi^-\eta$ and $\eta \to \gamma\gamma$. So in both $\eta\gamma$ and $\eta'\gamma$ final states, there are two charged particles

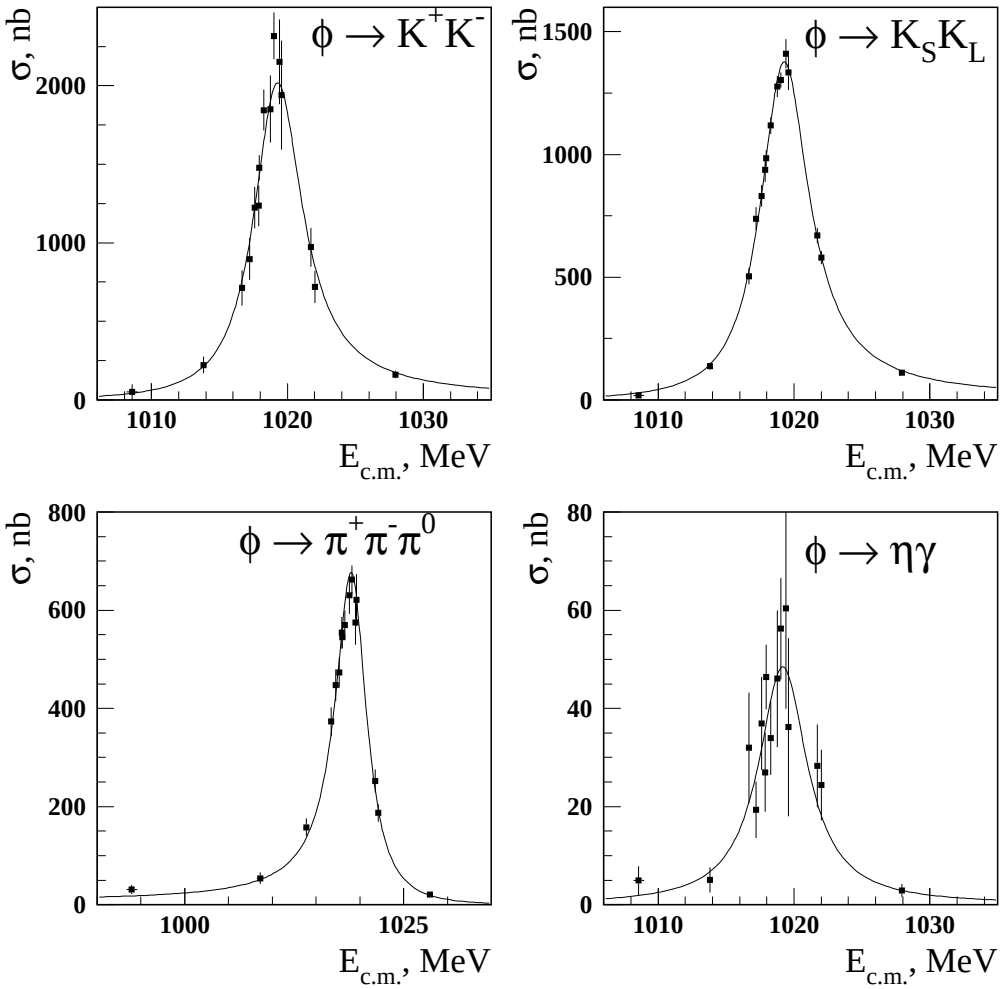


Fig. 3. The excitation curves for ϕ mesons in different channels.

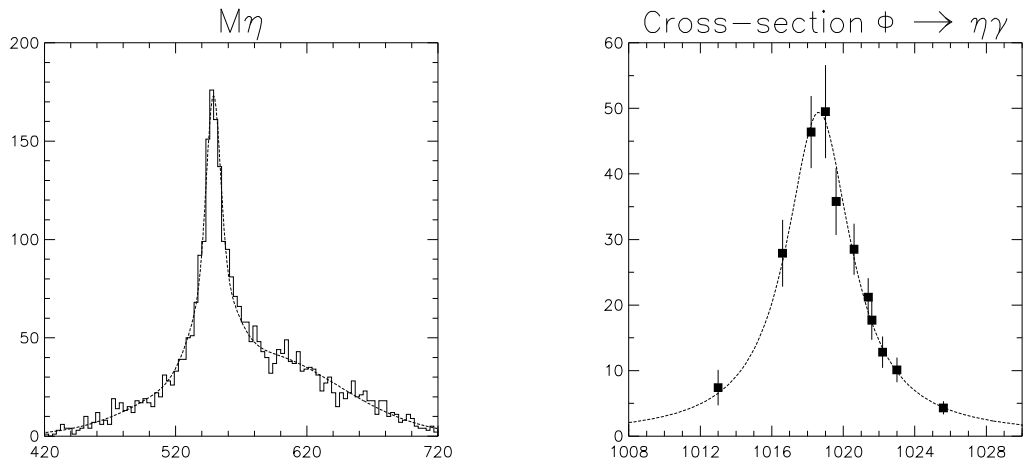


Fig. 4. The study of $\phi \to \eta\gamma$; invariant mass $M_{\pi^+\pi^-\pi^0}$ and $\phi \to \eta\gamma$ cross section.

and three photons. The events with all these detected particles were used for the constrained fit.

The scatter plot of the invariant masses for two soft photons M_{23} vs. the hardest photon energy ω_1 for the experimental data is presented in Fig. 5.

The decay into $\eta\gamma$ is the main background for $\eta'\gamma$. Removing the 481 $\eta\gamma$ events from the sample, the scatter plot of the invariant masses for the two hardest photons M_{12} vs. the weakest photon energy ω_3 was studied. For $\eta'\gamma$ events, M_{12} should be around η mass 547.5 MeV, while ω_3 is a monochromatic 60 MeV photon. Figure 5 presents the result of the 1992–1993 data together with simulation of $\phi \to \eta'\gamma$. We have one candidate $\eta'\gamma$ event, with one event estimated as the background. Using for the 90% C.L. upper limit $N_{\eta'\gamma} < 3$ and the ratio

$$\frac{Br(\phi \to \eta'\gamma)}{Br(\phi \to \eta\gamma)} = \frac{N_{\eta'\gamma}}{N_{\eta\gamma}} \cdot \frac{Br(\eta \to \pi^+\pi^-\pi^0)}{Br(\eta' \to \pi^+\pi^-\eta)} \cdot \frac{Br(\pi^0 \to \gamma\gamma)}{Br(\eta \to \gamma\gamma)} \cdot \frac{\varepsilon_{\eta\gamma}}{\varepsilon_{\eta'\gamma}} \quad ,$$

with the efficiencies obtained from the simulation, $\varepsilon_{\eta\gamma} = 14.4\%$ and $\varepsilon_{\eta'\gamma} = 6.4\%$, the following result has been obtained:

$$Br(\phi \to \eta'\gamma) < 2.4 \times 10^{-4} \quad .$$

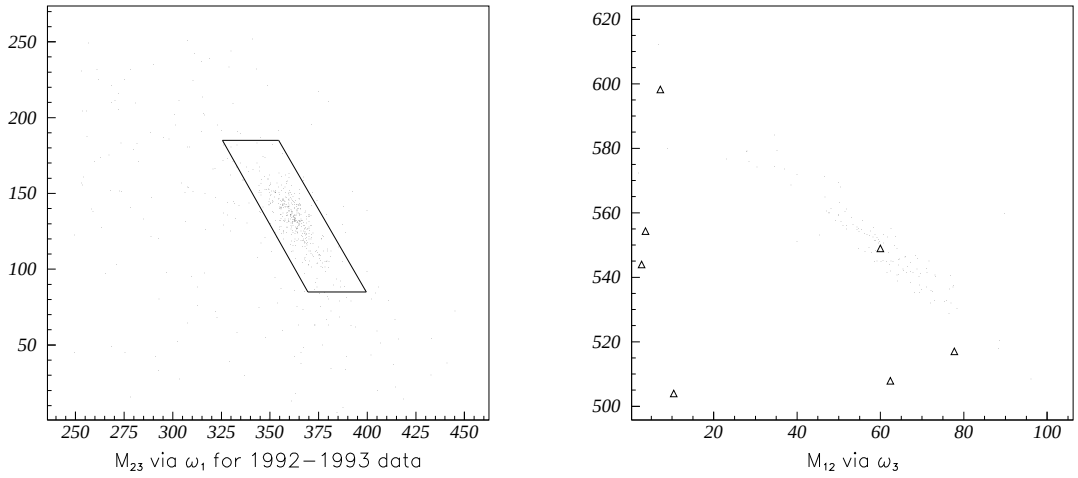


Fig. 5. The search for $\phi \to \eta'\gamma$ (1992–1993 data). Invariant mass M_{23} vs. ω_1 after the constrained fit (the box shows the $\eta\gamma$ cut); invariant mass M_{12} vs. ω_3 after constrained fits (dots are simulation, triangles the experiment).

6 Search for $\phi \to \pi^+\pi^-\pi^+\pi^-$

A sample of three- and four-track events was used to search for the process $\phi \to \pi^+\pi^-\pi^+\pi^-$. In this sample, tracks have to originate from the beam-beam interaction point within 0.5 cm in the r-ϕ plane and have at least nine hits in the Drift Chamber. The total charge has to be ± 1 for three-track events and zero for four-track events. To suppress background from the two-particle production and cosmic rays with some additional tracks, we reject events with at least two collinear tracks (mutual angle of any pair should be 0.16–3.0 rad in the r-ϕ plane).

Even with these selection criteria, we have a high background from the main channels of ϕ decaying into three-track events. So, in the search for the process $\phi \to \pi^+\pi^-\pi^+\pi^-$, only four-track events were used. The ratio of three- and four-track events at the energy points outside the ϕ meson region was used (along with simulation) for evaluation of a detection efficiency.

The scatter plot E_{tot} vs. P_{tot} for the selected four-track events is shown in Fig. 6. Here, P_{tot} is the magnitude of the total momentum of four charged particles, and E_{tot} is their total energy, assuming that all particles are pions. To extract

the number of events of the process $e^+e^- \to \pi^+\pi^-\pi^+\pi^-$, we apply a simple cut, shown in Fig. 6 as the box.

The extracted cross section vs. energy for the process $e^+e^- \to \pi^+\pi^-\pi^+\pi^-$ is also shown in Fig. 6. Only statistical errors are shown. A four-parameter function which contains linear background, amplitude, and phase of the process $\phi \to \pi^+\pi^-\pi^+\pi^-$ was used for the fitting. The result of the fit is shown in the plot by a smooth line. Using this fit and the uncertainty in the efficiency, one can get

$$\mathrm{B}r(\phi \to \pi^+\pi^-\pi^+\pi^-) < 1.0 \times 10^{-4} for C.L. = 90\% \quad .$$

This preliminary result is about nine times lower than the present upper limit for this process.[29] We plan to perform more simulations to improve both efficiency evaluation and data selection. Also, we plan to evaluate and apply radiative corrections which are about 5%.

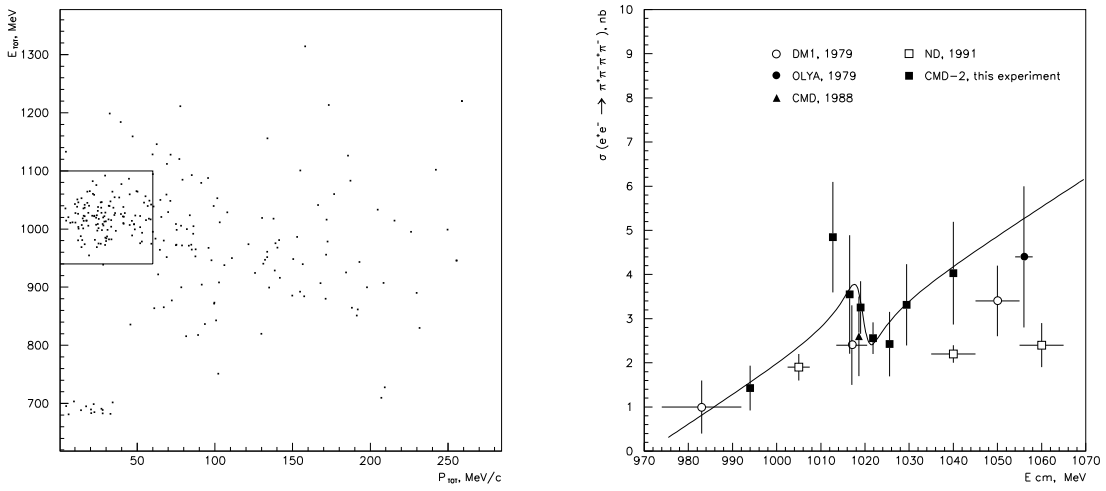


Fig. 6. The search for $\phi \to \pi^+\pi^-\pi^+\pi^-$. E_{tot} vs. P_{tot} for four-track candidates; cross section $e^+e^- \to \pi^+\pi^-\pi^+\pi^-$.

7 Search for $\phi \to f_0\gamma$

In order to extract the resonant contribution associated with the ϕ, two data sets were used. Energy points at $\mathrm{E}_{c.m.}$ from 1016–1023.2 MeV with the integrated luminosity 660 nb^{-1} were used for the ϕ region, and points at $\mathrm{E}_{c.m.}$ = 996, 1013,

1026, 1030, and 1040 MeV with the integrated luminosity 440 nb^{-1} were used for a background estimation (the "non-ϕ" region). The event candidates were selected by a requirement of only two charged tracks and only one photon with energy greater than 20 MeV in the detector. Total energy deposition was required to be less than 600 MeV, the average momentum of two charged tracks to be higher than 240 MeV, and the radial distance of the found vertex from the interaction region to be less than 0.15 cm. These cuts removed Bhabha events as well as charged and neutral kaons from ϕ decays. The requirement that the Z-coordinate of the vertex be within 10 cm at the detector center reduces cosmic ray background by a factor of two.

Each charged track was required to have a corresponding cluster in the calorimeter. This requirement reduced the number of pions by about 14% and helped avoiding nuclear interactions of the pions before the calorimeter with clusters in the wrong place. However, split clusters may still be present.

The main visible background for the studied process is $\phi \to \pi^+\pi^-\pi^0$ decay, when one of the photons from π^0 escapes detection. To reduce this background, a constrained fit was used. This fit required total energy and momentum conservation within detector resolutions for the three-body decay. For χ^2/d.f. less than three, only events with these requirements survived. But a three-pion background was still present, when one of the gammas from the π^0 had a very low energy and the event looked like a three-body decay. Figure 7 shows the spectrum of single gammas and the squared missing mass of two charged tracks (taken as pions) vs. detected gamma energy.

In the "ϕ region" data sample, a broad peak at 200–300 MeV in the gamma spectrum, also seen as a broad distribution on the scatter plot at $M^2_{\pi^0}$, represents background from the three pion decays. Points concentrated at zero mass and low energy represent events with one gamma. To reduce the three-pion and collinear-events background, the cuts $-15000 \le M^2_{inv} \le 15000$, $E_\gamma \le 140$ MeV and $\Delta\phi \ge 0.03$ rad were applied. The sample of events, selected with the above cuts, still contained about 30% of $e^+e^- \to \mu^+\mu^-\gamma$ events.

The visible cross section of the processes $e^+e^- \to \mu^+\mu^-\gamma + \pi^+\pi^-\gamma$ vs. energy is presented in Fig. 8(a). With the cuts listed above, the detection efficiency of these processes, obtained by simulation, was found to be 0.17, leaving 1.7 nb of the visible cross section. The observed 20% difference from the average experimental cross section is due to the losses of the low-energy gammas, not correctly described by a simulation. The curve shows a theoretical prediction of the cross section including the influence of ϕ on the photon propagator (vacuum polarization) and

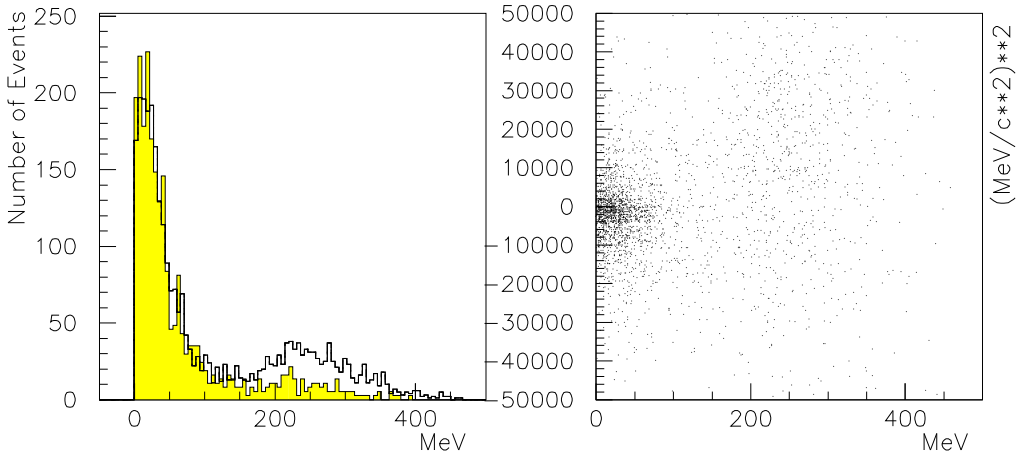


Fig. 7. Search for $\phi \to f_0\gamma$; (a) single gamma spectra (the shaded histogram represents the "non-ϕ" region), and (b) the squared missing mass vs. photon energy.

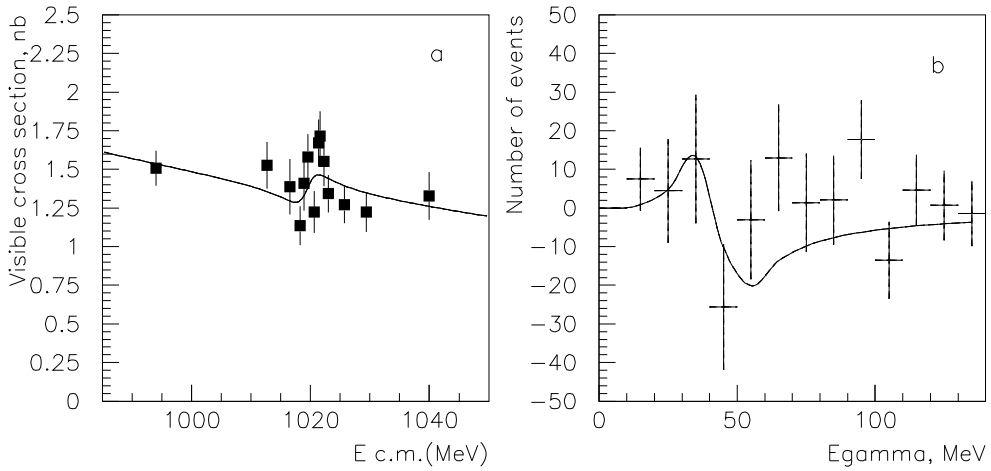


Fig. 8. $\phi \to f_0\gamma$ search. (a) Visible cross section for $e^+e^- \to \pi^+\pi^-\gamma + \mu^+\mu^-\gamma$ events. (b) Normalized difference in photon spectra. The curve is a prediction of the four-quark model with destructive interference and $B(\phi \to f_0\gamma) = 2.5 \times 10^{-4}$.

f_0 production, according to the Achasov four-quark model. The total branching ratio of $\phi \to \pi^+\pi^-\gamma$ could be extracted from this interference picture.

The signal from the decay of the $\phi \to f_0\gamma$ should be seen as a 30–40 MeV width structure at 45 MeV in the difference gamma spectra from the "ϕ" region and the "non-ϕ" region shown in Fig. 8(b), together with the theoretical prediction calculated by the Achasov model for the four-quark state. With the present statistics, only an upper limit can be set.

Taking into account the effective number of ϕ's 1.1×10^6 and all inefficiencies described above, the upper limit was found to be

$$B(\phi \to f_0\gamma) \leq 8.0 \times 10^{-4} \mathrm{at} C.L. = 90\% \quad .$$

8 The K_L Nuclear Interaction as a Tag for K_S Rare Decay Study

K_L candidates are selected by looking at calorimeter clusters opposite a two-track vertex with an effective mass consistent with that of the kaon. Figure 9(a) shows the space angle between the predicted missing momentum direction of K_L and the cluster in the calorimeter. Figure 9(b) shows the energy deposition of the presumed K_L, Fig. 9(c) shows the number of hit crystals, and Fig. 9(d) shows the probability for a K_L to interact, corrected for the interaction in surrounding materials. The clusters from K_L are very broad and in 25% of the cases are split into two to four pieces. Simulated distributions are shown shaded. Comparison of the data with the GEANT simulation (using the GHEISHA package) shows definite disagreement. The difference is due to completely incorrect cross sections for the low-energy kaons, used by the GHEISHA package.

Once the properties of the K_L clusters are understood, one can use the K_L cluster as a "tag" for K_S decays. Figure 10(a) shows the invariant mass distribution for the events with two charged tracks opposite to K_L clusters.

The constrained fit applied to the decay $K_S \to \pi^+\pi^-$ extracted 32,340 events, which were used for normalization. The rest is shown in Fig. 10(a) as the shaded histogram. In order to search for the $\pi^+\pi^-\gamma$ mode, the sample of events with an acollinearity angle less than 2.4 radians, $M_{inv} \leq 450$ MeV/c^2, and $E_\gamma \geq 50$ MeV was taken. Figure 10(b) shows the missing mass distribution for the two-track vertex, taking the K_S direction from the observed cluster, and the K_S momentum

Fig. 9. K_L interactions (dashed hits for simulation). (a) Space angle between calorimeter cluster and P_{mis}. (b) Energy deposition of K_L clusters. (c) Number of hit crystals (dashed hits for simulation). (d) Probability of K_L interactions in CsI.

from the known center-of-mass energy, and assuming $\phi \to K_S K_L$. 34.5 ± 8.0 events at zero mass, corresponding to the decay $K_S \to \pi^+\pi^-\gamma$, were found after fitting and background subtraction. The simulated ratio of the acceptances for $K_S \to \pi^+\pi^-$ and $K_S \to \pi^+\pi^-\gamma$ was found to be 2.49 ± 0.30, giving a branching ratio $\mathrm{Br}(K_S \to \pi^+\pi^-\gamma) = (1.82 \pm 0.49) \times 10^{-3}$.

Fig. 10. $K_S \to \pi^+\pi^-\gamma$ search. (a) Invariant mass for vertices opposite clusters. Events without K_S decaying to two pions are shown shaded. (b) Missing mass in the frame of K_S. Events with real gammas found are shown shaded.

9 The $K_S K_L$ Coupled Decay Study

The event candidates were selected from a sample in which two vertices, each of two opposite charged tracks, were seen. An example of this kind of event is shown in Fig. 11.

Figures 12(a) and 12(b) show scatter plots of the invariant mass of the two charged tracks, assuming that they are pions, vs. missing momentum for the vertex closest to the beam and to the other one. The concentration corresponding to K_S's dominates in Fig. 12(a) and is seen in Fig. 12(b). Two dimensional

Fig. 11. $\phi \to K_S K_L$ event with coupled decay.

cuts, $470 < M_{inv} < 525$ and $90 < P_{mis} < 130$ with an additional requirement to have another reconstructed vertex in the P_{mis} direction within detector resolution, select K_S's in any of the vertices, with K_L's remaining in the other. Figures 12(c) and 12(d) show a characteristic M_{inv} and P_{mis} broad distribution, expected from the main three-body K_L decays which are in good agreement with the simulation.

Figure 13(a) shows the distance from the point of origin to the decay point for selected K_S's. An exponential decay length is seen with the correct value 0.58 ± 0.03 cm convoluted with vertex position resolution 0.23 cm.

Figure 13(b) shows the vertex radius for the K_L decays. A loss of efficiency for these events is seen, since the K_L events should be approximately flat in this spatial region corresponding to the very early part of the K_L lifetime. A significant peak with 59 ± 16 events is also seen and is interpreted as the nuclear interactions of K_L at the 0.077 cm Be vacuum beam pipe.

The histogram in Fig. 13(d) shows the events consistent with two-pion decay at the K_L vertices, when the additional cut in M_{inv} was applied. With our resolution, the suppression of the semileptonic K_L decays by a factor of 20 was expected, and these events dominated at all radii (only two CP-violating K_L decays were expected with the present sample), except for the beam pipe, where 28 ± 6 extra events survived. We interpret these events as regeneration of K_L into K_S.

The rest of the peak events may be explained by Σ and Λ production, when two pions are detected and the recoil nucleon is unseen. With the applied M_{inv}

cut, about 10% of these events may be interpreted as pure two-pion decays and should be extracted from the candidates for regenerated events.

Fig. 12. $K_S K_L$ coupled decay study. Invariant mass vs. missing momentum for (a) first and (b) second vertex. (c) Invariant mass for K_L and K_S (shaded) after K_S selection. (d) Missing momentum for K_L and K_S (shaded) after K_S selection.

Figure 13(c) shows the projected angle difference between the missing momentum direction and a line connecting the K_L vertex with the beam position, for the events concentrated around the beam pipe. Dots with errors show the expected distribution for semileptonic decays of K_L, normalized to the expected number of these events. A peak at zero angle is seen, supporting the hypothesis of K_L into K_S regeneration.

Taking into account 10% corrections for the DC mylar window and 10% from nuclear interaction background, the regeneration cross section is found to be $\sigma_{reg}^{Be} = 63 \pm 19$ mb. A visible nuclear interaction cross section (excluding regeneration) is found to be $\sigma_{nucl}^{Be} = 60 \pm 18$ mb.

The obtained regeneration cross section is in agreement with the calculations, performed in Frascati,[30] that gave a value 40 mb for 114 MeV/c long-lived kaons. For the total nuclear interaction cross section of neutral kaons, one can obtain a value 549 $\pm$ 165 mb, taking into account the ratio 0.21 of hyperon production to all other inelastic processes[31] and 0.52 as the ratio of inelastic to elastic cross section.[30] It is also in good agreement with experimental data for higher momenta and the calculation of Ref. 30.

10 Study of $e^+e^- \to \pi^+\pi^-$

The data discussed in this talk were obtained by scanning the energy region between ϕ and ρ mesons with 10 MeV steps. About 1,000 pion pairs were sampled at every point. Most of the ~ 300 nb^{-1} integrated luminosity was collected during the 1994 run. In the 1995 run, this work was continued below the ρ peak down to 600 MeV. Using the resonance depolarization technique,[32] the beam energy at each point was measured with an accuracy of 10^{-4}. The detector trigger is described in Ref. 33. Events were recorded when:

- at least one track in the Drift Chamber was found by the tracking processor,[34] and
- the energy deposition in the CsI calorimeter was greater than 20 MeV.

About 20 million events were written onto magnetic tape. For off-line analysis, only collinear two-track events were selected. The cuts used for this selection are marked by arrows in Fig. 14. Events were used in a maximum likelihood function fit with the following global optimization parameters:

- number of electrons N_e,
- number of background events N_b,
- $\frac{N_\pi}{N_e+N_\mu}$, where N_π is the number of pions and N_μ is the number of muons.

The ratio N_μ/N_e was fixed from QED. The likelihood fit used information on the polar angle, longitudinal coordinate of the vertex, and energies deposited in the CsI calorimeter. As it is seen from Fig. 15, the experimental data are in good agreement with the fit.

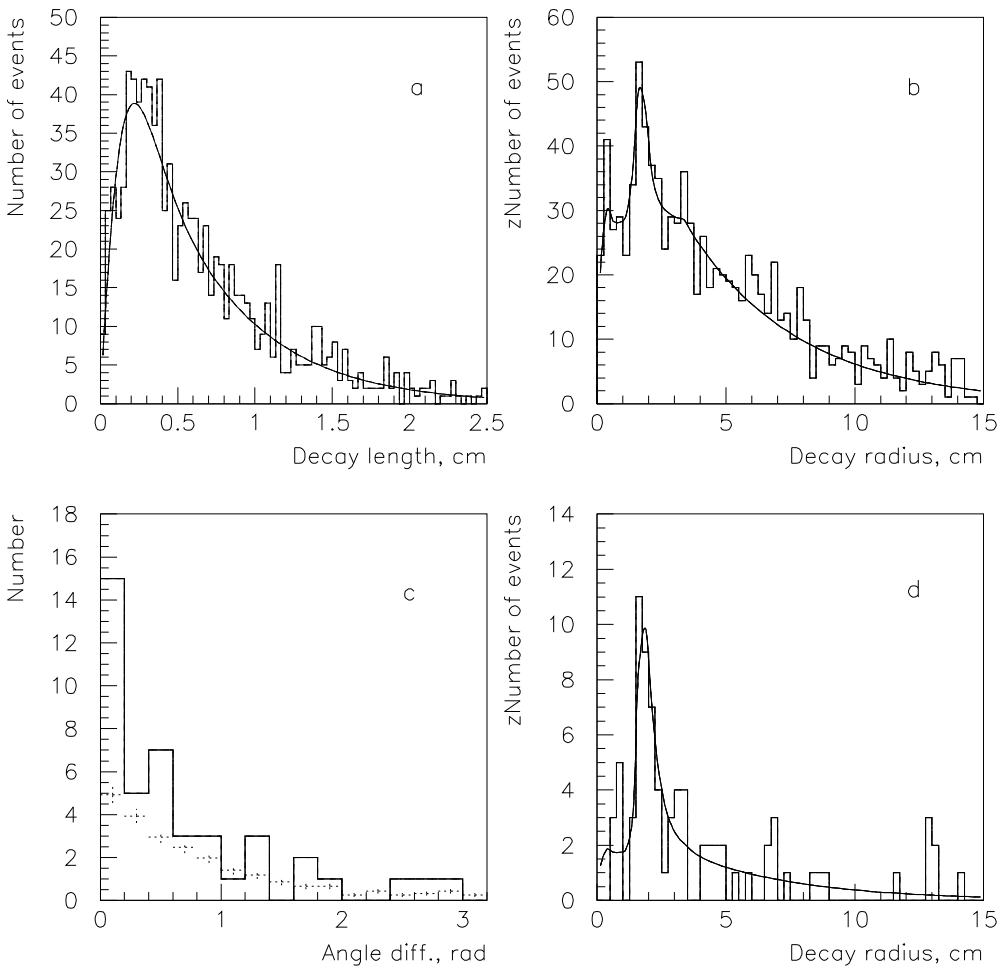


Fig. 13. $K_S K_L$ coupled decay study. (a) Decay length for K_S's. (b) Decay radius for K_L's. (c) Difference in angle of P_{mis} and the line from vertex to beam. (d) Decay radius for K_L's after the M_{inv} cut; crosses are the expectation from K_L semileptonic decays.

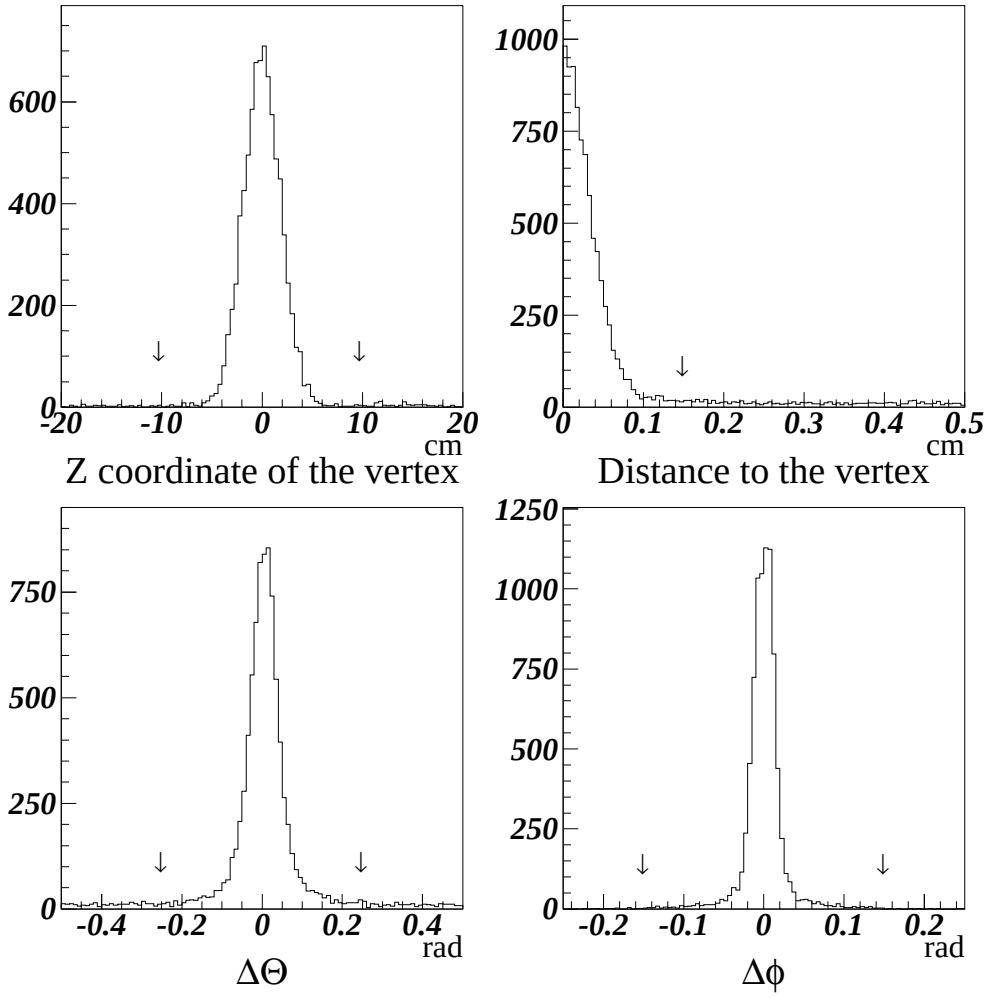


Fig. 14. Two-track event distributions and cuts imposed to select collinear events.

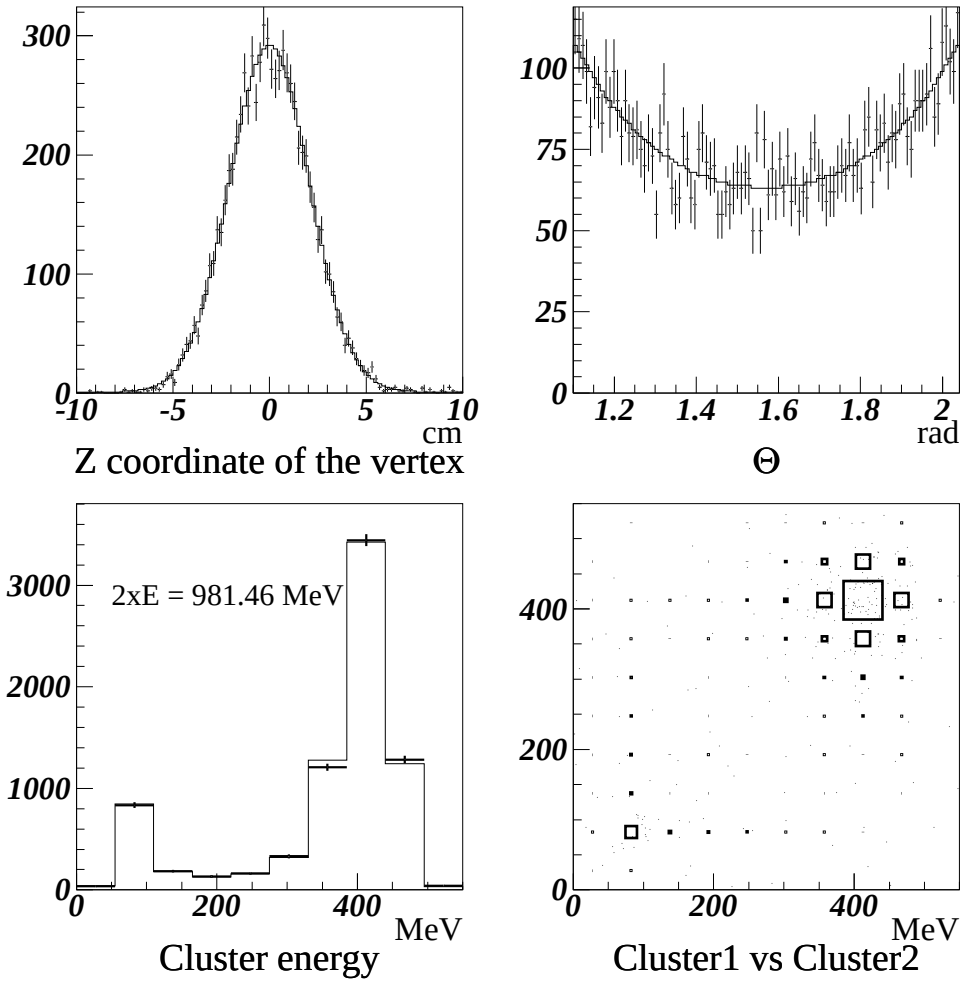


Fig. 15. The comparison of the experimental data with a fit.

The ratio $\frac{N_\pi}{N_e+N_\mu}$ allows us to express the $e^+e^- \to \pi^+\pi^-$ cross section in a simple way:

$$\sigma_\pi = \frac{N_\pi}{N_e + N_\mu} \cdot \frac{\epsilon_e\sigma_e(1+\delta_e) + \epsilon_\mu\sigma_\mu(1+\delta_\mu)}{\epsilon_\pi(1+\delta_\pi)},$$

where ϵ_e, σ_e, δ_e, ϵ_μ, σ_μ, δ_μ, ϵ_π, σ_π, and δ_π are detection efficiencies, cross sections, and radiative corrections for electrons, muons, and pions, respectively.

Pion form factor values are presented in Fig. 16 along with the results of the previous experiments. The statistical error of the $\pi^+\pi^-$ cross section at each energy point is less than 3%. At present, the total systematic error is estimated to be $\sim 1.5\%$. The main part of this error comes from the detector solid angle uncertainty $\sim 1\%$ and from the calculation of the radiative corrections for Bhabha events,[35] which are known with accuracy $\sim 1\%$.

The results from the OLYA detector have approximately the same statistical accuracy, whereas the systematical error in this energy range is about 4%.

Radiative corrections for all other channels of the e^+e^- annihilation into hadrons and muons were calculated with the accuracy about 0.2–0.5% (Refs. 36, 37) which is sufficient for the purposes of the experiment.

The Drift Chamber z-coordinate measurement can be calibrated by the Z chamber and thus improved significantly. Along with the more accurate calculations of the radiative corrections for Bhabha scattering events, this would decrease the systematic error to the level of 0.5%.

11 Conclusion

The next stage in this work is to process the data with improved detector resolution and to use all available particle identification information (drift chamber amplitudes, calorimeter energy deposition, and muon detector hits). The detector reconstruction efficiency is under intensive study and will reduce systematic errors for all results presented in this paper. Some other rare ϕ decay processes are under study.

The presence of regeneration and nuclear interaction background for the CP-violating decays of K_L will pose an additional background for ϕ-factory studies and should be under careful study.

Analysis of the collected data and new experimental runs are in progress, and we expect new results in the studies of ϕ, ω, ρ mesons and also in precision total hadronic cross-section measurements. The data taking at the ϕ is also planned with at least ten times more integrated luminosity before reconstruction of VEPP-

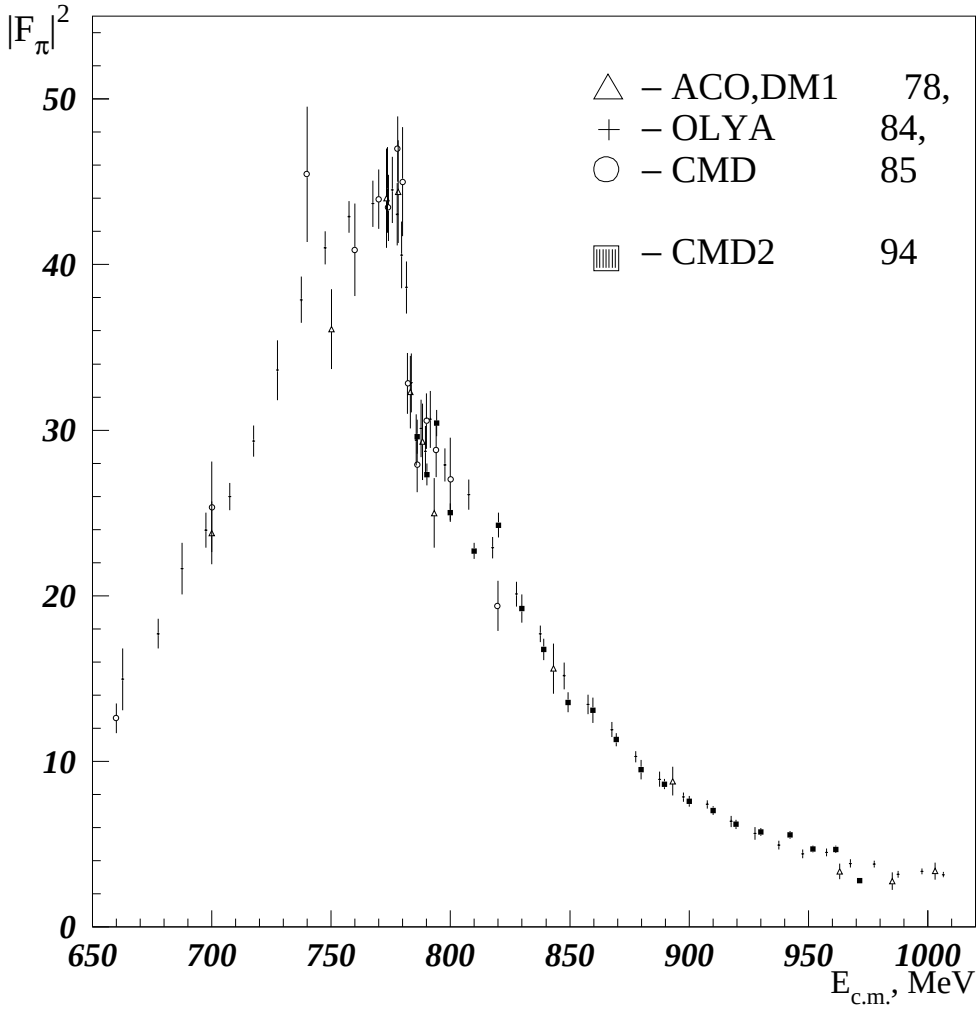


Fig. 16. The experimental data for $|F_\pi|^2$.

2M for round beam operation, which promises an additional factor of ten in the data sample.

Acknowledgments

I would like to thank R. Baldini and A. Michetti for their advice and useful discussions on the nuclear interaction study.

This work is supported in part by the U.S. Department of Energy, the U.S. National Science Foundation, and the International Science Foundation under grant RPT000.

References

[1] V. V. Anashin *et al.*, preprint INP **84-114** (Novosibirsk, 1984).

[2] L. M. Barkov *et al.*, Nucl. Phys. B **256**, 365 (1985).

[3] A. D. Bukin *et al.*, Sov. J. Nucl. Phys. **27**, 516 (1978).

[4] S. I. Dolinsky *et al.*, Phys. Rep. **202**, 99 (1991).

[5] V. N. Bayer, ZETFP **17**, 446 (1973).

[6] G. A. Aksenov *et al.*, preprint Budker INP 85-118 (Novosibirsk, 1985).

[7] P. Eberhard, contribution to the ϕ Factory Workshop at UCLA, April 1990; G. Ghirardi, *Proceedings of the Workshop on Physics and Detectors for DAΦNE* (Frascati, April 1991), p. 261.

[8] J. L. Rosner, I. Dunietz, and J. Hauser, Phys. Rev. D **35**, 2166 (1987).

[9] A partial list of other major contributions include: A. N. Skrinsky *et al.*, "Novosibirsk ϕ-factory project"; F. J. Botella, J. Bernabeu, and J. Roldan-Blois, CP Violation Conference, FTUV/89-35, IFIC/89-11, May 1989; G. Barbiellini and C. Santoni, CERN-EP/89-8 and CERN-PPE/90-124; J. A. Thompson, University of Pittsburgh, preprint PITT-90-09; contributions by J. A. Thompson, D. Cline, C. Buchanan, and others to the ϕ Factory Workshop at UCLA, April 1990; Y. Fukushima *et al.*, KEK preprint 89-159; summary talks by P. Franzini and M. Piccolo and others at the Workshop on Physics and Detectors for DAΦNE, Frascati, April 1991.

[10] A. N. Skrinsky, *Proceedings of the Workshop on Physics and Detectors for DAΦNE* (Frascati, April 1991), p. 67.

[11] G. Vignola, *Proceedings of the Workshop on Physics and Detectors for DAΦNE* (Frascati, April 1991), p. 11.

[12] A. N. Filippov *et al., Proceedings of the XVth International Conference on High Energy Accelerators*, V. II (Hamburg, Germany, 1991, World Scientific), p. 1145.

[13] S. Eidelman, E. Solodov, and J. Thompson, Nucl. Phys. B (Proceedings Supplement) A **24**, 174 (1991).

[14] D. Cocolicchio *et al.*, Phys. Lett. B **238**, 417 (1990).

[15] P. Franzini, *Proceedings of the Workshop on Physics and Detectors for DAΦNE* (Frascati, April 1991).

[16] N. N. Achasov and V. N. Ivanchenko, Nucl. Phys. B **315**, 465 (1989).

[17] S. Nussinov and T. N. Truong, Phys. Rev. Lett. **63**, 1349 (1989); A. A. Pivovarov, Soviet Physics—Levedev Institute Reports **9**, 12 (1990); N. Paver, contribution to ϕ Factory Workshop at UCLA, April 1990; J. L. Lucio and J. Pestieau, Phys. Rev. D **42**, 3253 (1990); and Oakes *et al.*, Phys. Rev. D **42** (1990).

[18] F. Close, plenary talk at the Workshop on Physics and Detectors for DAΦNE, Frascati, April 1991; and F. E. Close, N. Isgur, and S. Kumano, Nucl. Phys. B **389**, 513 (1993).

[19] S. Eidelman, J. A. Thompson, and C. H. Yang, *Proceedings of the Workshop on Physics and Detectors for DAΦNE* (Frascati, April 1991), p. 437.

[20] The summary talks of L. Maiani and R. Baldini Ferroli, of N. N. Achasov, and of M. Pennington at the Workshop on Physics and Detectors for DAΦNE, Frascati, April 1991.

[21] M. Vysotski, plenary talk, *Proceedings of the XXVII Int. Conf. on High-Energy Physics* (1993).

[22] Muon g-2 Design Report, AGS 821, BNL (July 1992).

[23] S. Eidelman and F. Jegerlehner, Z. Phys. C **67**, 585 (1995).

[24] E. V. Anashkin *et al.*, ICFA Instrumentation Bulletin **5**, 18 (1988).

[25] R. R. Akhmetshin *et al.*, preprint Budker INP 95-35, Novosibirsk, 1995.

[26] J. Lee-Franzini *et al.*, Phys. Lett. B **287**, 259 (1992).

[27] A. Bramon *et al.*, Phys. Lett. B **287**, 263 (1992).

[28] J. Lee-Franzini *et al.*, DAΦNE Physics Handbook, V.II (Frascati, 1992), p. 513.

[29] L. Montanet *et al.*, Phys. Rev. D **50**, 1173 (1994).

[30] A. Michetti, Thesis, Roma University, 1993.

[31] This value was extracted from the NUCRIN package.

[32] Y. S. Derbenev *et al.*, Part. Acc. **10**, 177 (1980).

[33] E. V. Anashkin *et al.*, Nucl. Instrum. Methods A **323**, 178 (1992).

[34] V. M. Aulchenko *et al.*, Nucl. Instrum. Methods A **252**, 299 (1986).

[35] F. A. Berends and R. Kleiss, Nucl. Phys. B **228**, 537 (1983).

[36] E. A. Kuraev and V. S. Fadin, Sov. J. Nucl. Phys. **41**, 466 (1985).

[37] S. I. Eidelman *et al.*, preprint Budker INP 95-34, Novosibirsk, 1995.

[38] F. A. Berends and G. J. Komen, Phys. Lett. B **63**, 432 (1976).

SEARCH FOR MILLI-CHARGED PARTICLES AT SLAC

W. G. J. Langeveld[⋆]
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309

Representing the mQ Collaboration

ABSTRACT

Particles with electric charge $q \equiv Qe \leq 10^{-3}\, e$ and masses in the range 1–1000 MeV/c^2 are not excluded by present experiments or by astrophysical or cosmological arguments. A beam dump experiment uniquely suited to the detection of such "milli-charged" particles has been carried out at SLAC, utilizing the short-duration pulses of the SLC electron beam to establish a tight coincidence window for the signal. The detector, a large scintillation counter sensitive to very small energy depositions, provided much greater sensitivity than previous searches. Analysis of the data leads to the exclusion of a substantial portion of the charge-mass plane. In this report, a preliminary mass-dependent upper limit is presented for the charge of milli-charged particles, ranging from $Q = 1.7 \times 10^{-5}$ at milli-charged particle mass 0.1 MeV/c^2 to $Q = 9.5 \times 10^{-4}$ at 100 MeV/c^2.

⋆ ⋆Work supported in part by the Department of Energy, contract DE-AC03-76SF00515.

©1995 by W. G. J. Langeveld.

1 Overview

Particles with a charge $q \equiv Qe \leq 10^{-3}\, e$, with e the absolute value of the charge of an electron, are called milli-charged (mQ) particles. In the following sections, I will describe an experiment[1] that was performed at the Stanford Linear Accelerator Center during 1994 and 1995 with the objective of searching for such particles, and report preliminary results.

In Sec. 2, I will briefly review what milli-charged particles are in more detail and why it makes sense to look for them. Some of the theoretical motivation[2,3] involves the concept of "shadow universes," but the main motivation for the experiment is the fact that milli-charged particles are not excluded experimentally in a large area of the mass vs. charge plane.[4–7]

Section 3 gives the presumed experimental signature for milli-charged particles and explains why SLAC is an ideal location to search for them. A discussion of possible backgrounds follows, and the experimental setup is described in detail.

Section 4 describes the analysis, and in Sec. 5, I present preliminary results.

2 What Are Milli-Charged Particles?

In general, milli-charged particles are particles with a charge $q \leq 10^{-3}\, e$. In the following, I will often use the abbreviation "mQ" to mean "milli-charged" or "milli-charged particle." The experiment I am about to describe in the following is, however, only sensitive to certain types of mQ particles.

For one thing, it is assumed that mQ's have only electromagnetic and gravitational interactions, i.e., they are not subject to weak or strong interactions. This restriction is more or less a natural consequence in a model due to Holdom, which I will discuss later.

For another, our experiment was only sensitive to mQ masses $\lesssim 100\, m_e$. And further, in order to reach our detector, mQ particles have to be stable or at least long-lived.

2.1 Why Look for Them?

The primary reason to look for mQ particles is because they have so far not been ruled out by experiment.[4–7] They are also not forbidden by established physical principles. In fact, charge quantization is poorly understood, and there is no a-priori reason to assume that all particles need to have a charge that is an integer multiple of $e/3$. The Standard Model can accommodate particles of arbitrary charge.

While it is true that most theories with supersymmetry require charge quantization at the fundamental level, there are mechanisms[3] by which mQ particles can be constructed (see the next subsection) without violating charge quantization.

Finally, if mQ particles exist, they might be a viable candidate for dark matter in the universe.

2.2 Shadow Universes

Holdom[3] has shown that mQ particles will arise naturally, and without violating charge quantization, in certain models involving what are called "shadow universes."

Imagine that there exist particles that to first order do not interact at all with the known matter in the universe, except gravitationally. Such particles might interact with each other in a similar way as the particles we all know. Such particles are generally known as "shadow particles," and the collection of them and their interactions make up a "shadow universe." Supposing there is a shadow-U(1) gauge group for these particles, it follows there is such a thing as a shadow photon (sometimes called a paraphoton).

Of course, it would be impossible to detect these particles if there were no interactions between regular particles and shadow particles at all. One way of introducing higher-order interactions is through postulating the existence of particles which carry both regular charge and shadow charge. If these dual-charge particles were light, then shadow matter would interact easily with regular matter, so in the model, they are heavy, with masses near the unification scale. Then one can have diagrams as in Fig. 1: a regular electron and positron annihilate into a regular photon, which turns into a loop of these dual-charge particles, which then turns into a shadow photon, which produces a pair of shadow particles. Of

Figure 1. A photon mixing with a shadow photon according to Holdom's mechanism.

course, if the mass of the intermediary is very large, the process would be very unlikely to occur.

The interesting thing about Holdom's model is that if there exist two types of such dual-charge particles, then there will be an interference between the two possible diagrams, and it turns out that the amplitude for the process involves the ratio of the two masses, and with suitably chosen values, the process would be more likely to occur.

One can treat the diagram of Fig. 1 in a simple manner. It so happens that in calculations, the regular photon coupled to the loop of intermediaries turning into a shadow photon coupled to the shadow particle pair acts like a regular photon coupling to the shadow-particle pair at reduced strength. In other words, the part of the diagram in Fig. 1 enclosed by the box can be replaced by a reduced coupling strength for that vertex, $q \equiv Qe$, with $Q \lesssim 10^{-3}$.

The basic production mechanism for milli-charged particles is then of the Bethe-Heitler type as shown in Fig. 2. Note the absence of the usual other diagram—it is of higher order in Q and therefore unimportant. Similarly, the detection mechanism (shown schematically in Fig. 3) is the excitation of an atom or molecule by the energy deposited by a milli-charged particle, and the subsequent deexcitation in the form of detectable photons. The amplitudes for both production and detection go as Qe, and therefore the total cross section for production and detection combined goes as Q^4.

Figure 3. Schematic representation of mQ detection.

2.3 Unexplored Regions

Figure 4 shows a graph of charge vs. mass and shaded regions of charge-mass combinations that have been excluded so far by other experiments.[4–7] Shown are regions excluded due to the most precise measurement of the Lamb Shift, and three regions excluded because of particle physics experiments: Mel Schwartz's SLAC Beam Dump experiment, the Fermilab E613 experiment, and the ASP Free

Figure 4. Currently excluded regions in the charge vs. mass plane.

Quark Search. The reinterpretation of the available data for the purpose of the graph was performed by S. Davidson *et al.*

Also shown are two regions excluded based on cosmological grounds. Masses below 1 MeV are excluded because of the effect the existence of mQ particles would have on nucleosynthesis: if mQ particles had a very small mass, they would have caused the universe to have cooled more rapidly, and nuclei would have started forming earlier, giving fewer free neutrons the chance to decay, which in turn would give rise to a higher He abundance than is currently observed.

The large triangular region at small charge and large mass was calculated with the assumption that the universe is not over-closed, i.e., $\Omega \leq 1$.

Neither of the cosmological bounds is especially firm.

Clearly, there exists a large region between charges of 10^{-3} and 10^{-6} and masses between 1 and 1000 MeV, where mQ particles have not been excluded.

3 The Experiment

The experiment reported[1] here ran in 1994 and 1995 at the Stanford Linear Accelerator Center. Essentially, the experiment can be categorized as a beam dump experiment, with the SLAC positron production target serving as a beam dump. The experiment was "parasitic" in the sense that it did not require dedicated beam time. Rather, whenever positrons were produced in the course of normal SLC operations, the experiment took data.

3.1 Signature

In such a beam dump experiment, where high-energy (29.5 GeV) electrons strike a target and produce light particles through a Bethe-Heitler-like mechanism, the particles will emerge near 0° with respect to the electron beam direction.

If the particles produced have milli-charge, they are expected to produce very little excitation and ionization in a detector. In a scintillator, they will produce only a single photon, leading to a single photoelectron detected in the photomultiplier.

For the same reason, mQ particles will travel through a large amount of material without losing much energy.

3.2 Why SLAC?

SLAC is ideal for a mQ particle search for several reasons. For one thing, beam time is essentially free because of the near-continuous use of the positron production target. For another, the pulsed nature of the SLC operation causes mQ particles to arrive at the detector at very precisely determined times. And finally, since the SLC pulse is so short, the time window within which mQ particles arrive at the detector is very small.

Overall, this leads to a very good signal-to-noise ratio. In addition, it is easy to estimate the background, simply by measuring it slightly before and slightly after the expected arrival time.

3.3 The Experimental Setup

The experiment consisted of a main scintillation detector 107 meters downstream of the positron production target, situated in a cylindrical 12-foot-diameter, 22-foot-deep pit which was excavated from the surrounding sandstone. In addition, a set of five scintillation panels, arranged in a cross, was located 85 meters downstream of the positron production target, lowered to beam level in one-foot-diameter holes drilled for the purpose.

A top view of the arrangement is depicted in Fig. 5. Shown are the accelerator itself, the positron extraction line, the positron production target, and the mQ experimental area. In the insert, the experimental area is shown in more detail: there are seven one-foot-diameter holes arranged in a line perpendicular to the beam line. These are called E1 through E7. Of importance to the experiment are only E1, E2, and E3. E1 is located on the beam line. Also in line with the beam line are holes P1 through P6. The main detector was located in P6. The trailer housed most of the electronics and the data acquisition system.

Figure 6 shows a head-on view of the five smaller scintillation panels. This setup was close enough to the positron production target that the highest energy muons were just barely able to reach these counters. The arrival time t_0 of muons was measured at E1, and the muon direction with respect to the nominal beam direction was determined using all five counters. They also provided a cross check for the luminosity determination and a measure of the alignment of the setup with the beam direction. Figure 7 shows a graph of muon flux versus deviation from the 0° beam line. For this particular measurement, the holes E4 and E5 were also used. The data points are fully consistent with a Gaussian shape, completely due to multiple scattering of the muons, as expected from EGS calculations.

A side view of the experimental setup is shown in Fig. 8. The electron beam strikes the six-radiation-length tungsten target, and the produced muons and mQ particles enter the sandstone after about 24 meters (80 feet). Muons will still reach the muon counters in the E-holes at 85 meters (280 feet, shown is only hole E1), and mQ particles are the only ones able to reach the main detector in hole P6 at 107 meters (350 feet).

The main detector is shown in Fig. 9. It consists of a set of four 8.25" x 8.25" x 54" blocks of plastic scintillator (Bicron 408), each with a shaped scintillator light guide glued to one end and equipped with an 8" hemispherical (Thorn-EMI)

Figure 5. Top view of the mQ experimental setup.

Figure 6. Head-on view of the four smaller scintillation panels used for muon detection.

phototube. Each block is wrapped in aluminum foil and plastic, and is surrounded by copper sheeting, serving as an isothermal surface during cooling. The four

Figure 7. Number of muons as a function of lateral distance from the muon beam axis.

assemblies are contained in a Lucite box, which in turn is situated in a copper container which could be cooled to $-20°$ C. Surrounding the copper box were a six-inch layer of thermal insulation and a four-inch layer of lead shielding. Each counter was also equipped with an LED for calibration purposes. A pneumatic device allowed a radioactive source to be put inside the copper box by way of a tube, again for calibration purposes.

The signal from each of the four phototubes, which were run at relatively low voltage, was split into two signals, one of which was fed into a low-noise $\times 40$ amplifier. Both signals were connected to an ADC, with one ADC channel measuring the full spectrum of pulse heights using the raw signal, while the other measured the low pulse-height spectrum from the amplifier in detail. The amplified signal was also fed into a discriminator and read out by a TDC, in order to measure the arrival time of the signal.

A schematic of the data acquisition is shown in Fig. 10. Analog signals from the experiment were digitized using CAMAC TDC's and ADC's. The CAMAC

Figure 8. Side view of the mQ experimental setup.

Figure 9. Schematic of the mQ main detector.

Figure 10. Representation of the mQ data acquisition system.

system was read out by an Amiga 3000T/040 personal computer (Millie), and the data were logged to a large hard disk. When the disk filled up, the data were offloaded through a fiber-optic Ethernet line to one of the SLAC central UNIX systems, and from there the data were written to tapes in the robotic tape silos.

The data acquisition system continuously updated a large number of on-line histograms, which were examined at the end of the run for any anomalies. Signals from the SLC control system were also read in through CAMAC, in order to measure certain machine parameters. In addition, certain records from the SLC controls database were transmitted from the main SLC computer to the data acquisition computer, and merged into the data stream, allowing monitoring of the beam parameters during off-line analysis.

The data acquisition computer was equipped with a modem and would send messages to the pager of the person on shift if something went wrong. A second Amiga computer (Mollie) was also equipped with a modem and monitored the status of the main computer. If the main computer went off-line for any reason, the second computer would page the person on shift. It was therefore possible to run the experiment without anyone in attendance. Typical runs lasted about 24 hours.

3.4 Backgrounds

The main background in the case of our large scintillator detector turned out to be scintillator luminescence, the source of which has not been completely pinned down.[8] The current hypothesis is that events that produce a lot of light in the scintillator (mostly cosmic rays) cause the scintillator to emit photons for a long time afterwards. This is a much larger source of noise than the thermionic noise from the tube itself, which was selected for its low noise characteristics to begin with. Another problem due to the large amount of light from cosmic rays is tube afterpulsing. Cooling the entire counter to about 5° C reduces scintillator luminescence somewhat. A hardware veto was installed to reduce noise from afterpulsing by almost a factor of ten, at the cost of introducing a tolerable ($\approx$ 30%) deadtime.

Cosmic rays themselves are not a source of background, because of the small coincidence window with the beam arrival time, and the fact that the pulse height recorded from any cosmic rays is so much larger than the single photoelectron scale expected for milli-charged particles.

Natural radioactivity from the surrounding sandstone is a source of background, but with four inches of lead shielding and the copper box surrounding

the detector, this background is small in general, and negligible in the SPE region (the typical signals from natural radioactivity are much larger).

Other potential backgrounds, such as neutrons, coming straight from the target or scattered from the atmosphere ("skyshine"), neutrinos, and gammas from muon radiation have been calculated and/or measured to be negligible.

One important feature of the experiment is that the background can be measured accurately: data obtained out-of-time with the beam-arrival time is a good measure for the background, and as a check, the experiment was repeated with the detector taken out of the beam.

4 Data Analysis

4.1 Luminosity

The experiment took data from August of 1994 to April of 1995, at an incident electron beam energy of 29.5 GeV. The number of electrons per beam pulse was of the order of 3.0×10^{10} under good beam conditions. In order to keep track of the integrated luminosity, the digitized current of a toroid in the electron beam line as well as the number of muons detected in muon counter M1 was recorded. The results presented here represent a (dead-time corrected) integrated luminosity corresponding to 1.03×10^{19} electrons on target.

4.2 Production and Acceptance

The number of mQ particles produced per electron on target was calculated using a Monte Carlo QED calculation written by M. Swartz. The calculation included the effects of electron showering in the six-radiation-length tungsten target. A numerical integration was then performed in order to determine the fraction of mQ's within the detector acceptance. It was further assumed that for small enough Q, the results scale with Q^2.

Figure 11. Typical time spectrum of muons arriving in M1.

4.3 Detection

The number of photoelectrons produced was calibrated by comparing the pulse-height spectra recorded for Co, Cs, and Am sources with EGS simulations. An additional calibration point comes from cosmic rays.

From the above, photoionization and δ-ray fractions were calculated that produce a single photon at $Q = 10^{-3}$. The results were then extrapolated to lower Q assuming a Q^2 dependence.

4.4 Arrival Time in the Detector

The arrival time t_0 of mQ's in the detector was arrived at in the following way. The arrival time of muons in muon counter M1 was 60 ns after the SLC reference signal used for this purpose. The time spectrum of muons arriving in M1 is shown in Fig. 11. The time of flight from M1 to the middle of the mQ detector is 90 ns at the speed of light.

Figure 13. Differences in arrival time of small and large signals. The sharp early peaks are due to large signals (without filter), and the asymmetric broad peaks are due to single photons. The three different sets correspond to the three different measurement locations.

A correction needs to be made to account for the differences in the lengths of the cables from M1 and the mQ detector to the trailer. These cable lengths were individually measured for M1 and each of the four mQ phototubes, and relative differences in the latter were taken into account. The overall correction is -33 ns.

An additional correction is needed to account for slewing: small signals arrive later than large signals. This effect was measured to be 16 ns. Figure 12 shows the setup used for the measurement: two small scintillators were used to make a cosmic ray trigger, and the light produced in one of the four main scintillators was measured by an eight-inch phototube. The measurement was done with the cosmic ray trigger in three different positions along the length of the detector. The set of measurements was repeated with an absorption filter in front of the eight-inch phototube, such that the light was attenuated to a single-photoelectron level. The results are shown in Fig. 13. The sharp peaks at early times are due to the measurements without a filter, where the full cosmic ray signal (tens of thousands of photons) arrived at the tube. The broad, later distributions are due to the measurements with a filter.

Adding the various contributions (see Fig. 14), one arrives at $t_0 = 133$ ns.

4.5 Time Window Size

The measurement of the time spectrum of single photons described above also shows that the resulting time spectrum is an asymmetric, broad distribution which starts at about $t_0 - 7$ ns and ends at about $t_0 + 33$ ns. Of the SPE events, 85% fall within this range.

4.6 Spectra

Figure 15 shows a density plot of pulse height as a function of the time, for the raw, unamplified pulse heights of one of the four counters. The cut-off at times above about 200 ns and pulse heights above channel 900 is due to the fact that large pulses that arrive late in the time window have part of their signal outside of the time window, resulting in a lower measured signal size. Aside from this effect, the spectrum looks the same for all times. A projection onto the pulse-height axis is shown in Fig. 16.

Figure 17 shows the density plot for the more interesting region of low pulse heights (measured using the $\times 40$ amplifier). The dark area between channel number 100 and 500 corresponds to the SPE peak. The darkness of each bin in Figs. 15 and 17 corresponds to the log of the number of counts in that bin. The horizontal dark band at the top of Fig. 17 is made up of overflows. The vertical dark band at about 20 ns corresponds to the ADC pedestal level. Aside from this,

Figure 14. Diagram of the determination of the mQ arrival time.

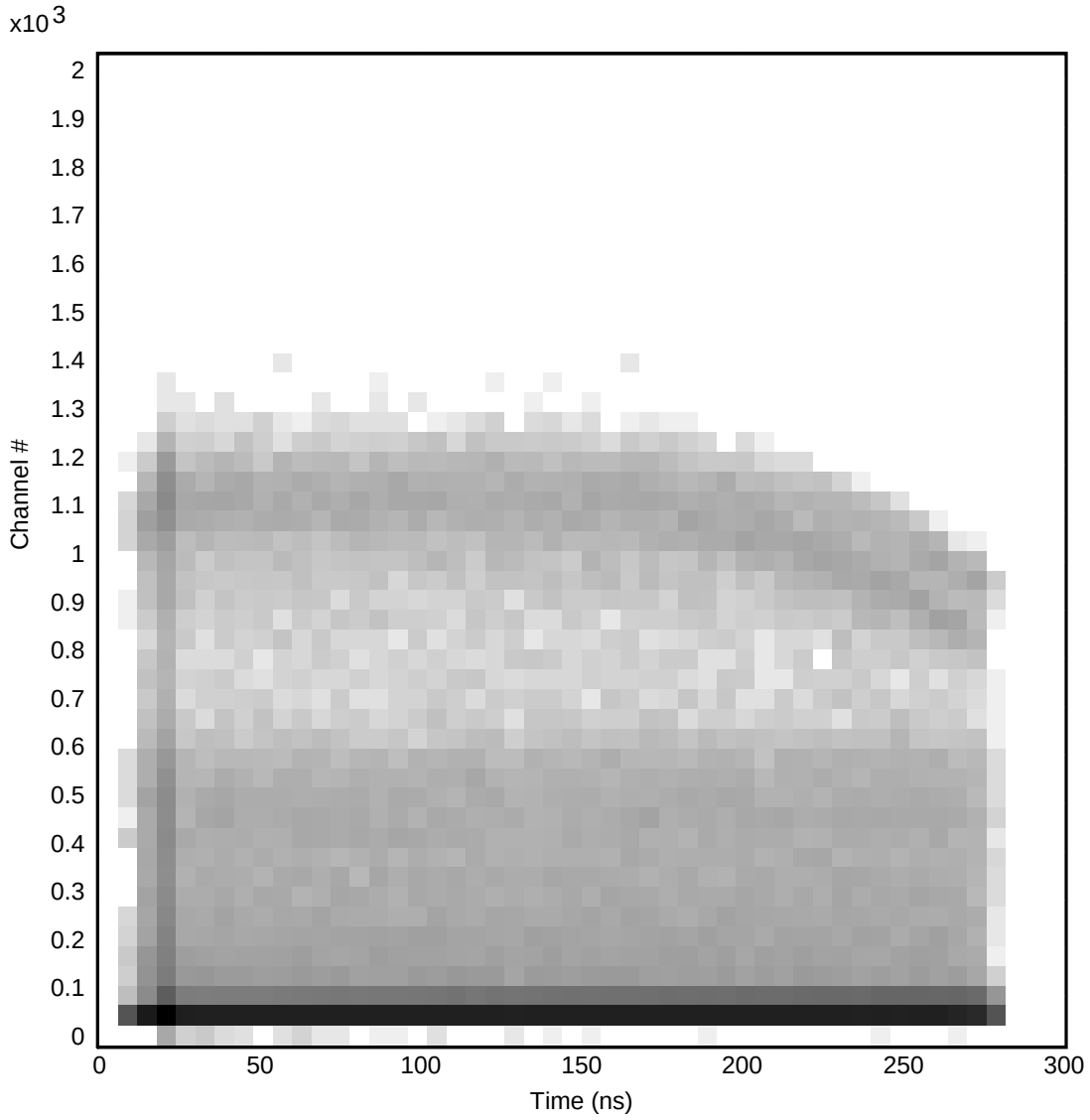


Figure 15. Density plot of the raw (unamplified) pulse height vs. time for one of the four counters.

there is no significant time dependence. A projection onto the pulse-height axis is shown in Fig. 18.

The time distribution of all events, summed over the four counters (after subtracting relative time offsets), is shown in Fig. 19. Note the offset vertical scale. A slight clustering of events near 150 ns is statistically insignificant. When a cut is made to allow only events with a pulse height between 2/3 and twice the nominal pedestal-subtracted SPE peak location (see Fig. 20), we arrive at

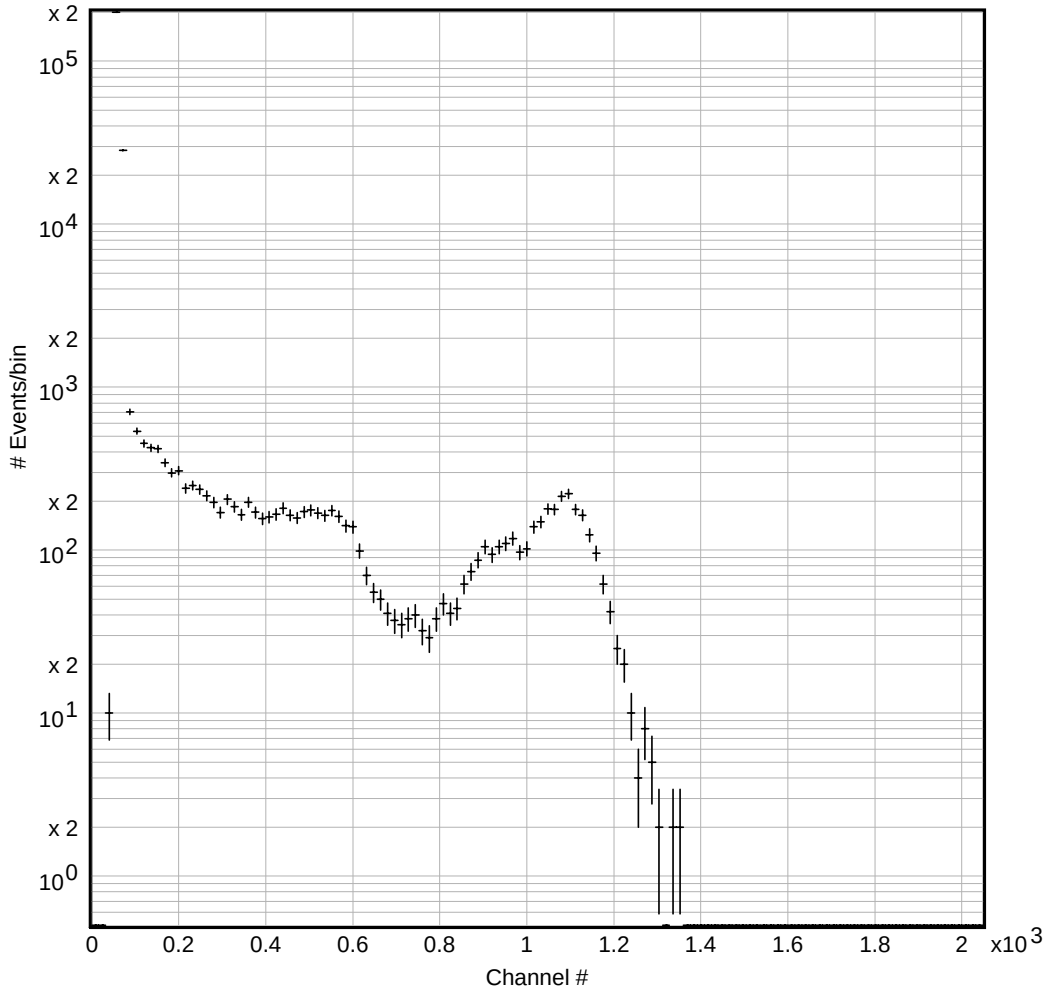


Figure 16. Number of events as a function of the raw (unamplified) pulse height of one of the four counters.

the time spectrum shown in Fig. 21. Note again the offset vertical scale. The same clustering around 150 ns survives this cut, which keeps 72.5% of the SPE events, but it is again not statistically significant. The overall upward slope of the spectrum is due to the asymmetric SPE cut: events with a pulse height below 2/3 SPE have a compensating downward-sloping time distribution. From other histograms, it is found that the spectrum of events that arrive early is shifted by a few channels toward a lower pulse height compared to events arriving late. One likely explanation for this effect is that the signals are capacitively coupled to the readout electronics and have an overshoot after the trailing edge of the signal. For

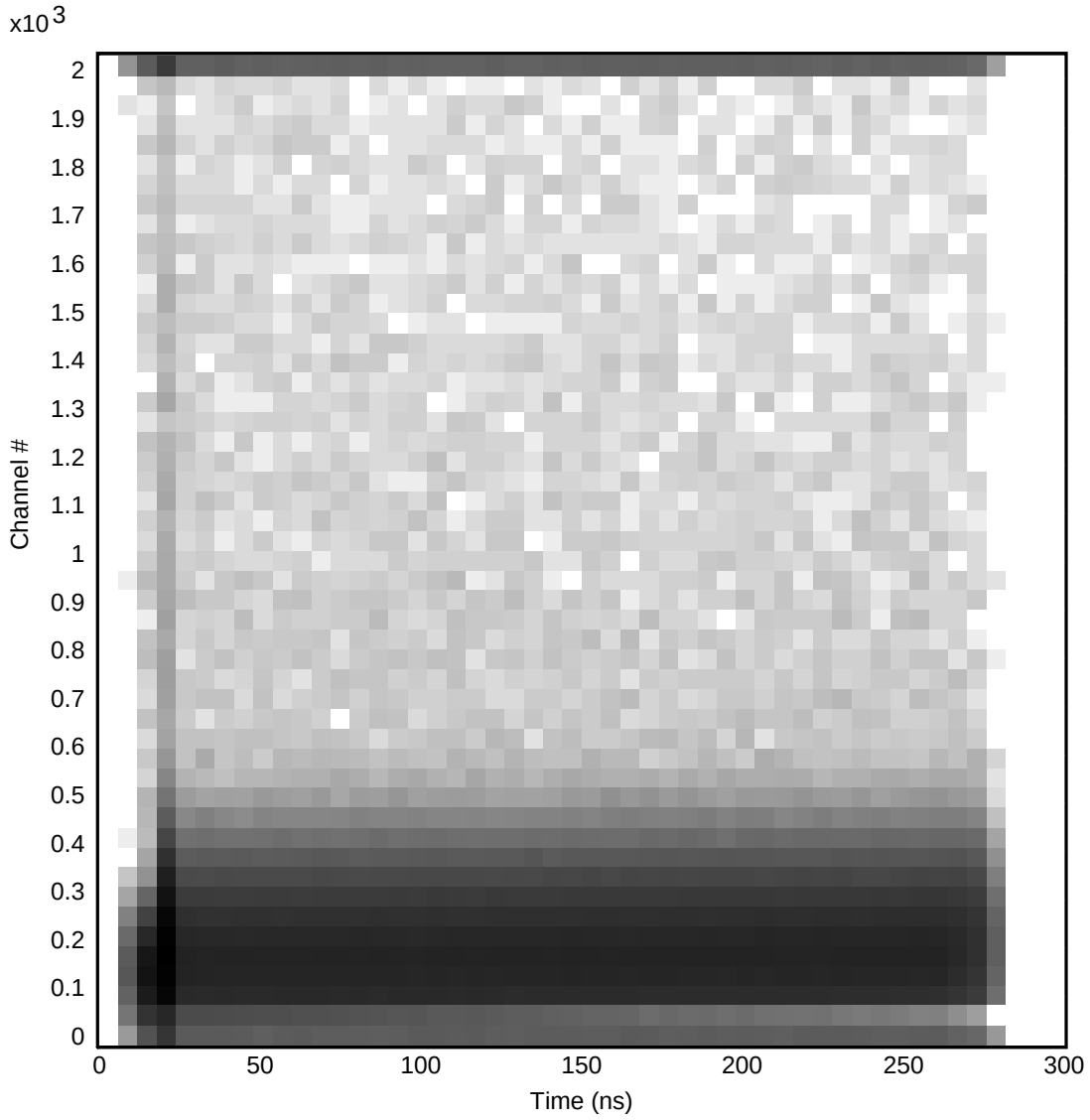


Figure 17. Density plot of pulse height vs. time, measured with the ×40 amplifier, for one of the four counters.

early arrival times, more of the overshoot takes place within the time interval the ADC integrates over than for later arrival times. This effect is linear with time.

The vertical lines in Fig. 21 at 126 and 166 ns outline the calculated time window within which mQ events would appear. A fit was made to the data outside this region, in order to estimate the background inside the signal region. A straight line was found to be a good fit ($\chi^2 = 1.02$ per degree of freedom).

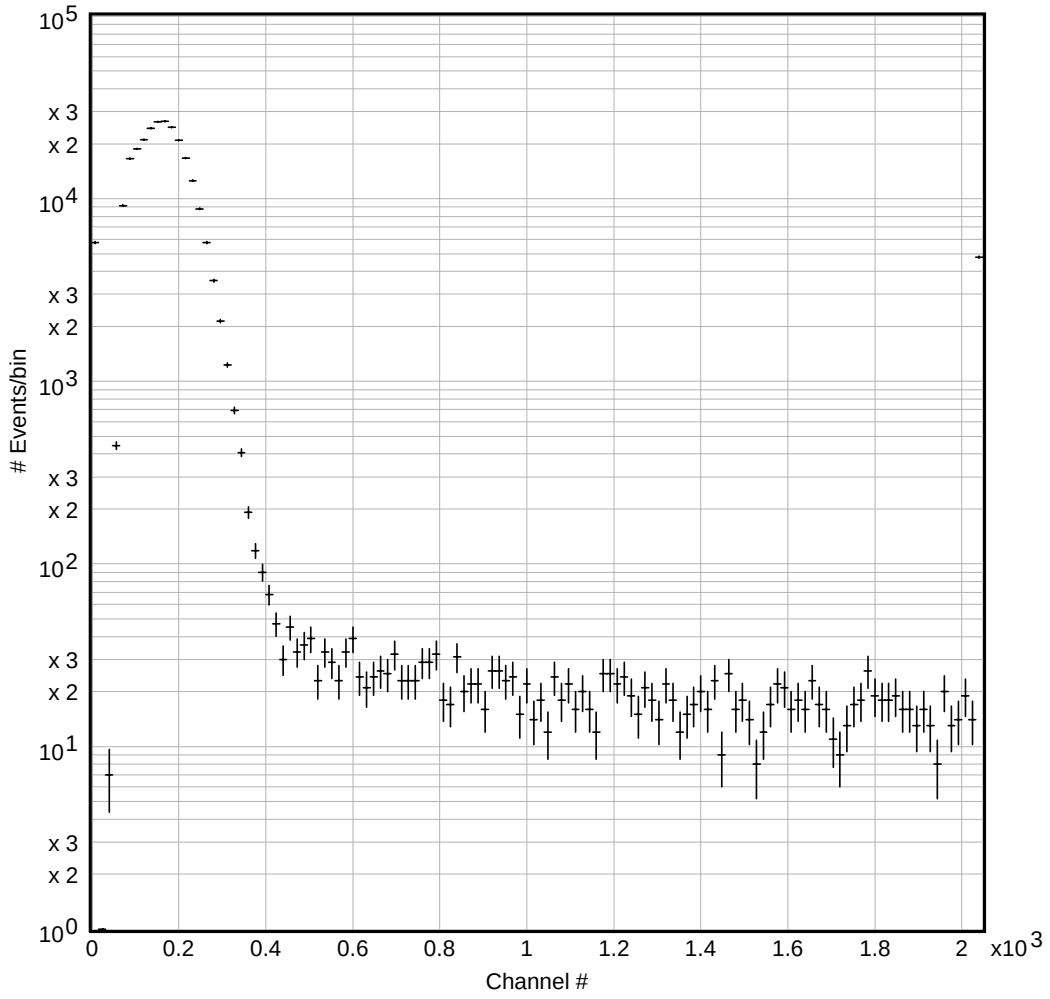


Figure 18. Number of events as a function of pulse height, measured with the ×40 amplifier, for one of the four counters.

There is no statistically significant excess above background in the signal region: the excess is less than one sigma above background in the 40 ns bin.

4.7 Checks

Various checks were made to ensure the significance of the experiment. The alignment of the detector with respect to the beam was checked repeatedly and in different ways. The detector was less than 2 cm off-center during most of the running period.

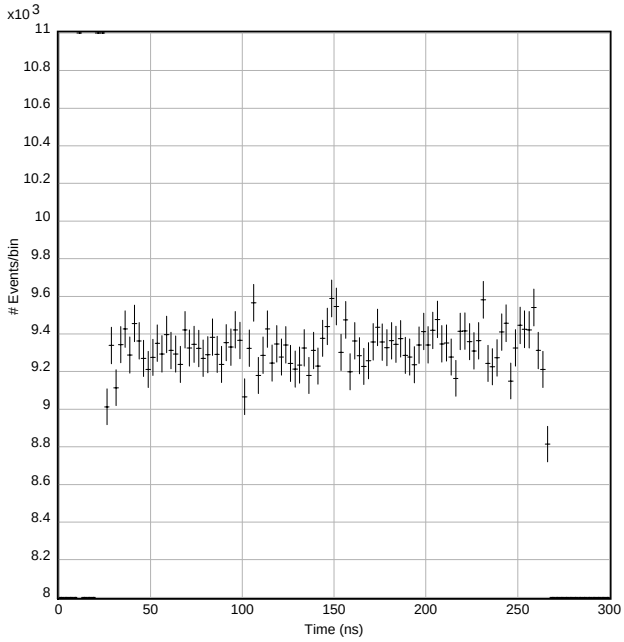


Figure 19. Time distribution for events of all pulse heights.

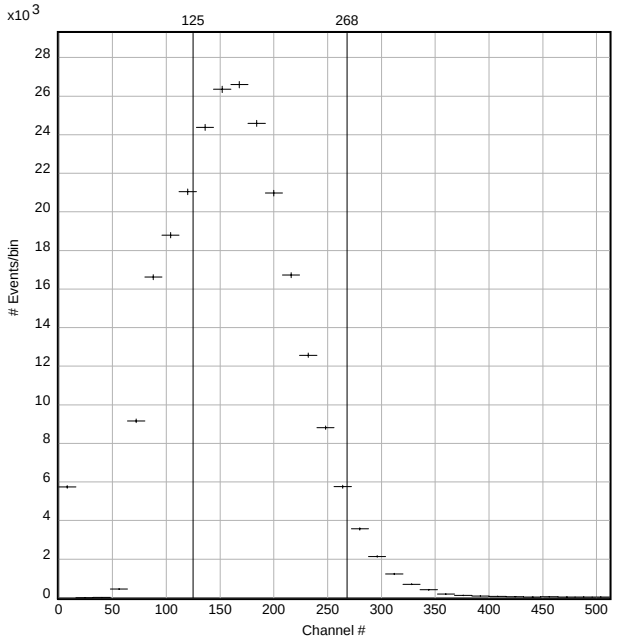


Figure 20. Single photoelectron peak. The vertical lines indicate the cuts on pulse height that were made in the analysis.

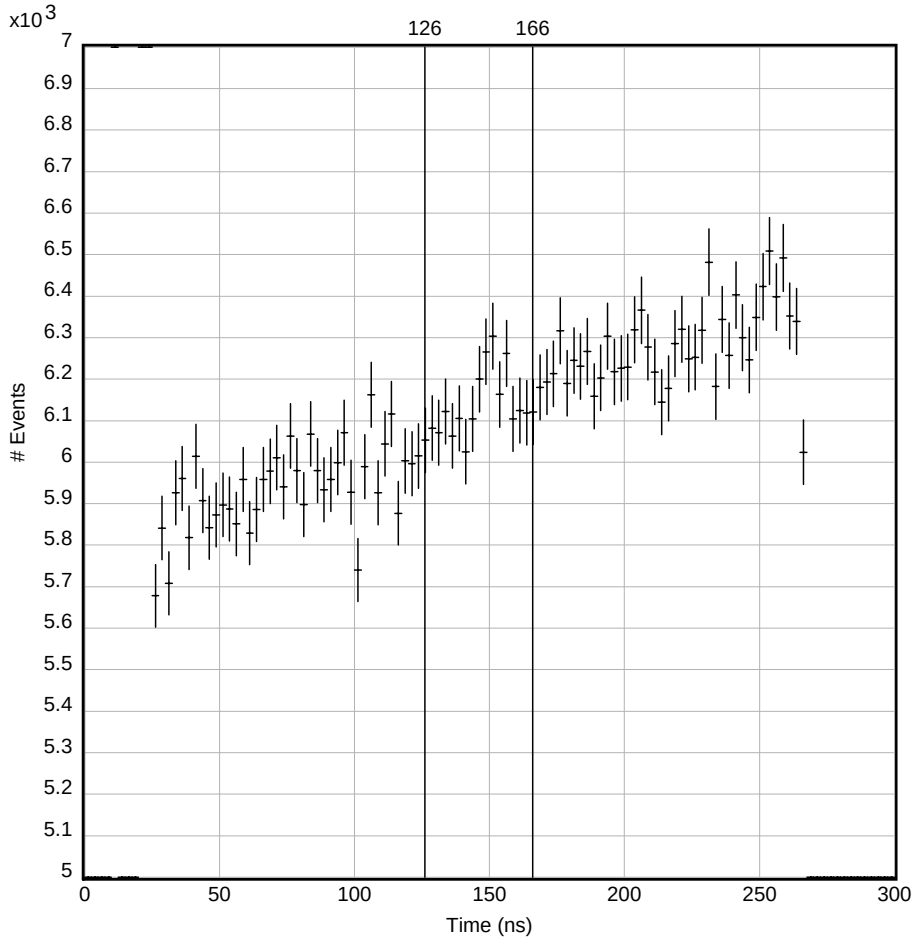


Figure 21. Time distribution for events in the SPE region.

The LED's mounted on the four scintillators of the detector were pulsed at regular intervals, and the resulting events were recorded. These measurements proved that the counter was alive, that the time measurements were stable, and that the timing resolution remained good during the runs included in this analysis.

Measurements of the average Americium source pulse height from "source" runs and the average position of the SPE peaks demonstrate that the counters were sensitive and stable during this time.

4.8 Efficiencies

As mentioned before, the SPE cut had an estimated efficiency of 72.5%. The limitation of the time window to a 40 ns interval around t_0 had an estimated efficiency of 85%.

4.9 Systematic Errors

In order to arrive at conservative estimates for upper limits, the following systematic errors were taken into account.

— The estimated systematic error in the current calculations of the production of mQ particles is a very conservative 50%.

— The estimated systematic error in the measurement of the luminosity is 10%.

— The estimated systematic error due to alignment uncertainties is 10%.

— The estimated systematic error due to uncertainties in the calculation of energy deposition is 10%.

— The estimated systematic error due to the calculated light yield is 20%.

— The estimated systematic error in the efficiency is 10%.

When added in quadrature, these contributions amount to an overall estimated systematic error of 57%. This number is of course dominated by the error in the calculations for mQ production. In the upper limit results that follow, the "worst case" production rate was assumed.

5 Preliminary Results and Conclusions

The upper limit for the charge of mQ particles allowed by our measurements was now determined, for four different mQ masses, as follows.

First, the estimated background was subtracted from the data, to obtain the signal. All events in the signal region were summed, and the error σ in the signal was calculated. If the total signal (after background subtraction) was less than zero, the signal was taken to be zero events. To this signal, 1.65 times σ was added in order to arrive at a 95% Confidence Level (CL) upper limit.

This procedure was repeated for various different values of t_0, in order to check what would happen if for some reason our measurement of t_0 were incorrect, or if mQ particles did not arrive at the expected time. As an example, the 95% CL upper limit for Q is plotted vs. t_0 in Fig. 22 for a mQ mass of 1 MeV/c^2. As

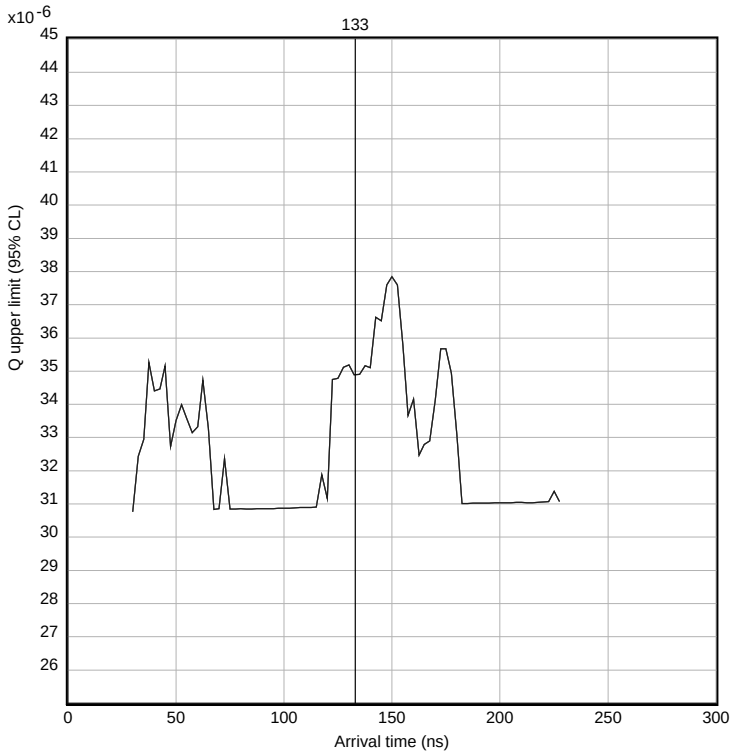


Figure 22. Upper limit on mQ charge for mQ mass 1 MeV, as a function of assumed arrival time.

expected, there is a small peak near 150 ns, but the upper limit is not significantly different from the upper limit calculated at the nominal value of $t_0 = 133$ ns, indicated by the vertical line in the figure.

The results of the experiment are summarized in Fig. 23. It shows the already excluded regions mentioned before, and in addition, the newly excluded region determined by this experiment. Specifically, the results for four different mQ mass values are:

m_{mQ} (MeV)	Q (95% CL)
0.1	$< 1.7 \times 10^{-5}$
1	$< 3.5 \times 10^{-5}$
10	$< 1.2 \times 10^{-4}$
100	$< 9.5 \times 10^{-4}$.

Figure 23. Preliminary results of this experiment. Shown is the newly excluded area, compared to the previously excluded regions.

Our detector is in principle sensitive to mQ masses above 100 MeV, but reliable limits have not yet been determined. We have also not yet computed an upper bound for the excluded region—our detector is not sensitive to large Q^2 ($Q \lesssim 0.2$), low-mass mQ's, because such particles would range out in the sandstone in front of the detector.

In conclusion, we have established preliminary upper limits for the fractional charge Q of mQ particles as a function of their mass. A significant fraction of the hitherto allowed region in the Q vs. mass plane was excluded.

References

[1] The members of the Milli-Charged Particle Search Experiment at SLAC are: R. Baggs, J. Ballam, S. Ecklund, C. Fertig, J. Jaros, K. Kase, A. Kulikov, W. Langeveld, B. Leonard, T. Marvin, T. Nakashima, W. R. Nelson, A. Odian, M. Pertsova, A. Prinz, G. Putallaz, and A. Weinstein.

[2] A. Y. Ignatiev, V. A. Kuzmin, and M. E. Shaposnikov, Phys. Lett. B **84**, 315 (1979).

[3] R. Holdom, Phys. Lett. B **166**, 196 (1986).

[4] M. I. Dobroliubov and A. Y. Ignatiev, Phys. Rev. Lett. **65**, 679 (1990).

[5] S. Davidson, B. Campbell, and D. Bailey, Phys. Rev. D **43**, 2314 (1991).

[6] R. N. Mohapatra and S. Nussinov, Int. J. of Mod. Phys. A **7**, 3817 (1992).

[7] S. Davidson and M. Peskin, Phys. Rev. D **49**, 2114 (1994).
[8] T. P. Marvin, SLAC-PUB-95-7042 (1995).

TEST OF QED AT CRITICAL FIELD STRENGTH

Christian Bula
Physics Department, Princeton University
Princeton, New Jersey 08544

Representing the E-144 Collaboration:
Princeton/Rochester/SLAC/Tennessee

ABSTRACT

In a new experiment at the Final Focus Test Beam at SLAC a low-emittance 46.6-GeV electron beam is brought into collision with terawatt pulses of 1054 nm or 527 nm wavelength from a Nd:glass laser. Peak laser intensities of 10^{18} W/cm^2 have been achieved corresponding to a value of 0.6 for the parameter $\eta = e\mathcal{E}/m\omega_0 c$. In this case an electron that crosses the center of the laser pulse has near-unit interaction probability. Results are presented for multiphoton Compton scattering in which an electron interacts with up to 4 laser photons, in agreement with theoretical calculations.

1 Introduction

The interaction of electrons with intense wave fields was first considered by Schott[1] which led to the introduction of the dimensionless measure of field strength

$$\eta = \frac{e\mathcal{E}_{\rm rms}}{m\omega_0 c} = \frac{e\mathcal{E}_{\rm rms}\lambda_0/2\pi}{mc^2} = \frac{e\sqrt{\langle A_\mu A^\mu\rangle}}{mc^2},$$

for a plane wave of laboratory frequency ω_0, wavelength λ_0, electric field E and four-vector potential A_μ. A field with $\eta = 1$ has a voltage drop of an electron rest mass per reduced laser wavelength $\lambda_0/2\pi$. In the average rest frame of an electron in a wave field the transverse motion has characteristic velocity $\beta^\star = v^\star/c$ related by $\gamma^\star\beta^\star = \eta$, where $\gamma = 1/\sqrt{1-\beta^2}$, so that parameter η is often called $v_{\rm osc}/c$ in weak fields. As η approaches and exceeds unity the classical radiation spectrum includes higher harmonics of the wave frequency ω_0 (multipole radiation). In the quantum view this corresponds to absorption of several wave photons before emission of a single photon of frequency ω:

$$e + n\omega_0 \rightarrow e' + \omega.$$

Only one observation of this effect has been reported: a weak signal of second-harmonic radiation in scattering of 1-keV electrons from a Q-switched Nd:YAG laser.[2] A closely related effect is higher-harmonic generation in a free-electron laser,[3] where η is often called k.

A quantum description of electrons in a strong wave field utilizes the Volkov solutions[4,5] to the Dirac equation, in which an electron is 'dressed' by continual absorption and re-emission of wave photons leading to an effective mass

$$\overline{m} = m\sqrt{1+\eta^2}.$$

The role of the effective mass in Compton scattering of electrons in a strong wave field was discussed by Sengupta[6] and others.[7–10] In nonuniform waves the effective energy $\overline{m}c^2$ is called the ponderomotive potential, which describes the forces on a charged particle as it enters or exits the wave.[11,12] Ponderomotive effects on electrons ejected from atoms in a wave field with $\eta \approx 1$ have recently been observed by Moore *et al.*.[13]

We report on an experiment in which 46.6-GeV electrons are scattered at the focus of an intense laser with wavelength $\lambda_0 = 1054$ (infrared) or 527 nm (green).

Under these conditions the photon energy in the rest frame of the electron beam is of order of the electron rest mass so that recoil effects are important. Absorption of a single photon corresponds to ordinary Compton scattering. However, at the laser intensities achieved ($I \approx 10^{18}$ W/cm^2, $\eta \approx 0.6$) the probability for multiphoton absorption is large and this effect was readily observed.

When n photons are absorbed by an electron of initial energy E_0 from a laser pulse with intensity parameter η and crossing angle θ_0 to the electron beam the minimum energy of the scattered electrons is

$$E_{\min} = E_0/[1 + 2nE_0\omega_0(1 + \cos\theta_0)/\overline{m}^2].$$

The higher effective mass of the electron in the wave field shifts the minimum scattered energy to slightly higher values. For ordinary Compton scattering ($n = 1$) the minimum scattered-electron energy is 25.6 GeV at $E_0 = 46.6$ GeV, $\eta = 0$, and $\theta_0 = 17°$. The spectrum of electrons scattered by absorption of more than one laser photon extends below 25.6 GeV permitting an identification of multiphoton (nonlinear) Compton scattering.

Figure 1 shows spectra of scattered electrons calculated according to ref.[10] for conditions representative of the present experiment with $\eta = 0.5$. The calculation includes the space-time profiles of the electron and laser beams and makes the adiabatic approximation that the rate based on infinite plane waves holds for the local value of η. The calculation also includes the effect of multiple Compton scattering in which an electron undergoes successive ordinary Compton scatters at different points as it traverses the laser focus. This process is physically distinct from nonlinear Compton scattering in which several photons are absorbed at a single point and a single high-energy photon is emitted. Figure 2a represents $n = 2$ nonlinear Compton scattering, while Fig. 2b represents two successive ordinary Compton scatters. Electron e' in Fig. 2b is real. The black circles indicate that the absorption of a wave photon by an electron in a Volkov state is not simply described by a vertex factor of charge e.

The curves in Fig. 1 are labeled by the highest number of photons that are absorbed in a single scattering event. Thus the dashed curve labeled $n = 1$ corresponds to ordinary Compton scattering, but extends below 25.6 GeV because of multiple ordinary Compton scattering. The curve labeled $n = 2$ also extends below the nominal minimum energy for nonlinear Compton scattering because additional ordinary Compton scatters also occur. The upper solid curve is the sum

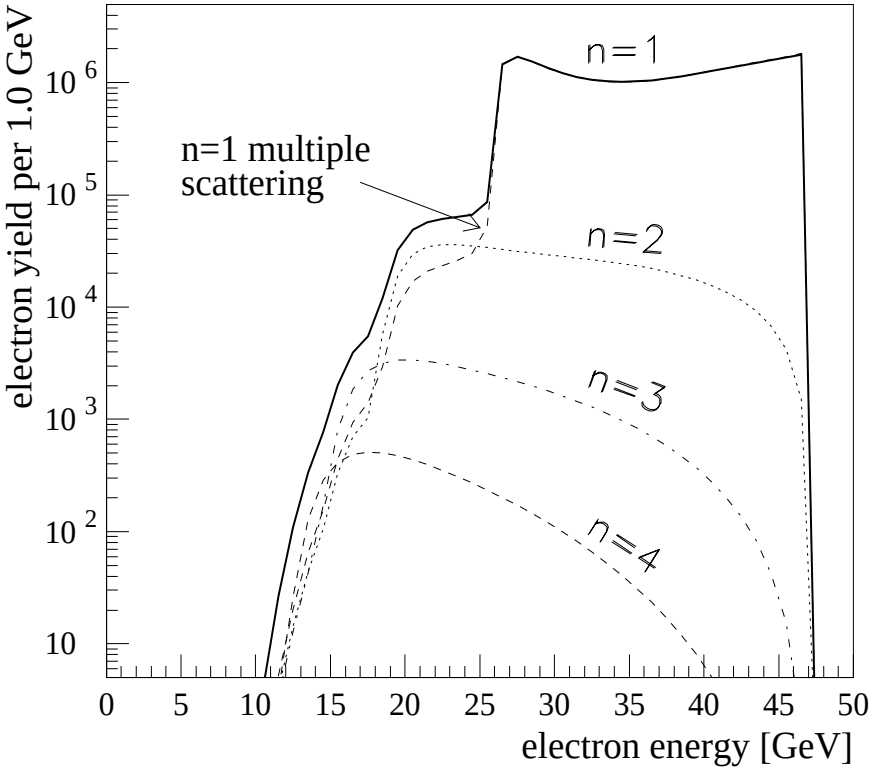


Figure 1: Calculated yield of scattered electrons from the collision of 5×10^9 46.6-GeV electrons with a circularly-polarized 1054-nm laser pulse with intensity parameter $\eta = 0.5$.

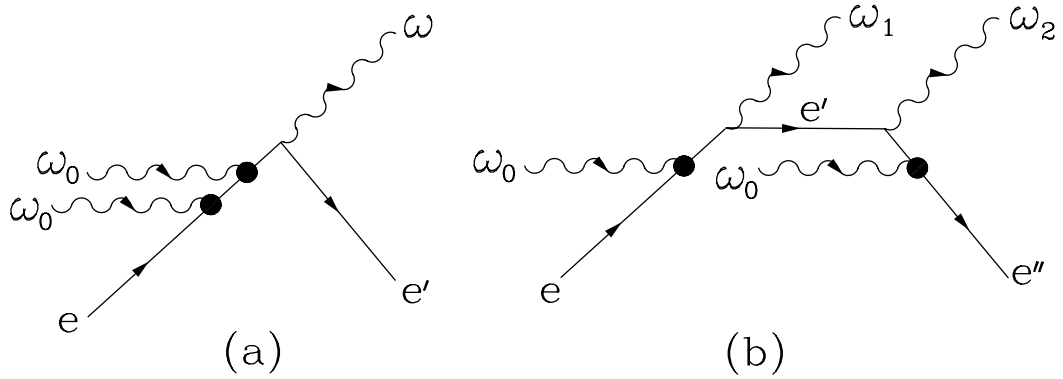


Figure 2: Diagrams representing (a) $n = 2$ nonlinear Compton scattering, and (b) double ordinary Compton scattering.

of all possible scatterings. Note that the simulated electron rates for $n = 2$ nonlinear Compton scattering and double ordinary Compton scattering are roughly equal in the energy range 20-25 GeV.

In quantum electrodynamics a natural measure of electromagnetic field strength is the so-called critical field for which the voltage drop across a Compton wavelength is an electron rest mass:

$$\mathcal{E}_{\rm crit} = \frac{m^2c^3}{e\hbar} = 1.3 \times 10^{16}\ {\rm V/cm} = 4.4 \times 10^{13}\ {\rm gauss}.$$

The critical field was first introduced by Sauter[14] as the characteristic field strength at which Klein's paradox[15] becomes important and was further interpreted by Heisenberg and Euler[16] as the field strength at which electron-positron pair creation becomes copious. For a particle in a strong wave field a useful dimensionless invariant is

$$\Upsilon = \frac{e\hbar}{m^3c^5}\sqrt{(F_{\mu\nu}p^\nu)^2} = \frac{\mathcal{E}^\star}{\mathcal{E}_{\rm crit}} \simeq \frac{2\gamma\mathcal{E}}{\mathcal{E}_{\rm crit}},$$

where $F_{\mu\nu}$ is the field tensor and p_ν is the particle's 4-vector; $\mathcal{E}^\star$ is the wave field in the particle's rest frame, and the final equality holds only if the particle is moving anticollinear to the wave with Lorentz boost γ. Static fields with values of Υ approaching one are thought to exist at the surface of neutron stars. The field at the surface of a nucleus has Υ less than one, but quasistatic fields with Υ exceeding unity arise in MeV heavy-ion collisions.

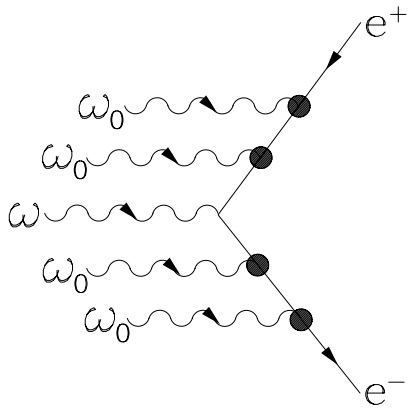


Figure 3: Diagram representing multiphoton pair creation.

Electron-positron creation can arise in the interactions of electrons with a wave in a two-step process in which a Compton-scattered photon collides with

wave photons to produce the pair. Weak-field pair creation by photons was first considered by Breit and Wheeler,[17] and Reiss[18] first discussed the strong-field case,

$$\omega + n\omega_0 \to e^+e^-,$$

in which several wave photons participate; see also refs.[8,10] Figure 3 represents the latter process for a case where an external photon and four wave photons combine to produce a pair.

The present experiment studies the basic interactions of electrons and photons in fields near the QED critical field strength. It is also relevant to the understanding of so-called beamstrahlung processes at future e^+e^-colliders where the fields surrounding the beam bunches approach $\mathcal{E}_{\rm crit}$,[19] and where the consequent pair creation will be a limiting background. The experiment provides a demonstration of the technology for e-γ and γ-γ collider options,[20] leading to measurements of the γWW coupling via the reaction $e\gamma \to W\nu$,[21] *etc.* Copious production of positrons in e-γ collisions can provide a low-emittance positron source due to the absence of final-state Coulomb scattering.[22]

The parameters η and Υ are not independent, and for electrons colliding head-on with a wave their relation is $\Upsilon/\eta = 2\gamma\hbar\omega_0/mc^2$. For GeV electrons interacting with a laser the ratio of Υ to η is near one, so experiments in these conditions probe nonlinear effects due to both multiphoton absorption and vacuum polarization.

2 Experimental Setup

2.1 Phase I

The experiment presented here is carried out in the Final Focus Test Beam at SLAC.[23] The setup for the first phase of the experiment is shown schematically in Fig. 4. The laser is focused at the interaction point, IP1, 10 m downstream of the Final Focus. A set of permanent dump magnets is used to direct the electron beam downwards to the dump and also serves as the analyzing magnet of our experiment.

Compton-scattered electrons are deflected away from the primary electron beam by the dump magnets and are detected in a Silicon-Tungsten calorimeter (ECAL),[29] sketched in Fig. 5a. Positrons were deflected to the opposite side of

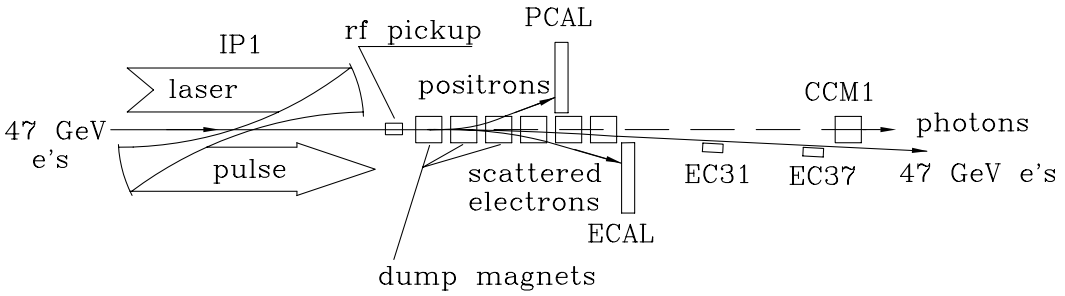


Figure 4: Sketch of experiment E-144 to detect scattered electrons and positrons produced in e-laser collisions at the SLAC Final Focus Test Beam.

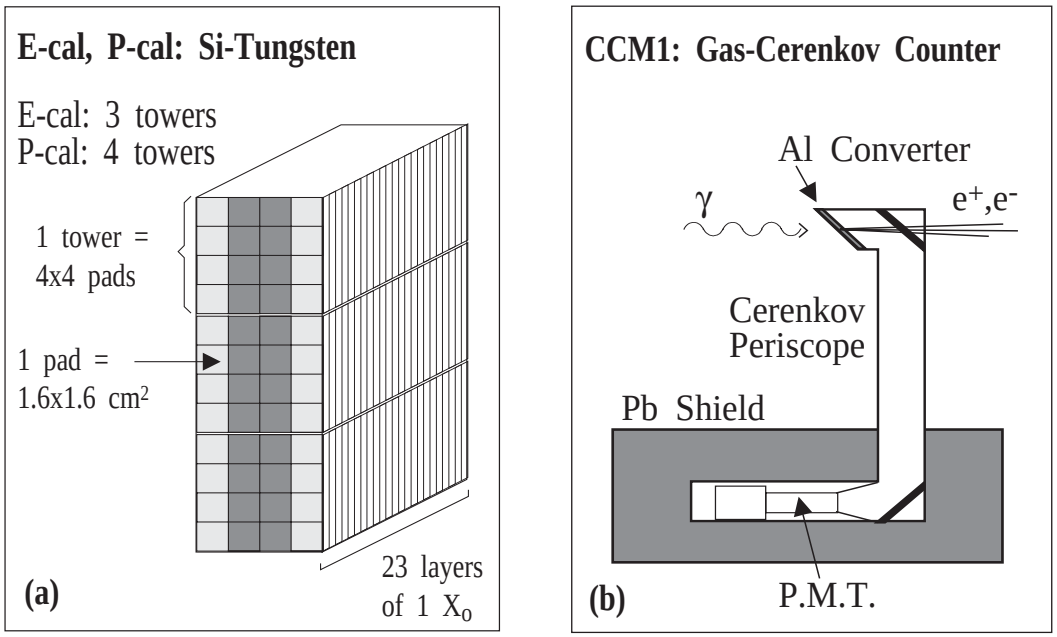


Figure 5: (a) The Silicon-Tungsten calorimeters ECAL and PCAL. (b) The gas Čerenkov monitor CCM1; monitors EC31 and EC37 are of similar construction.

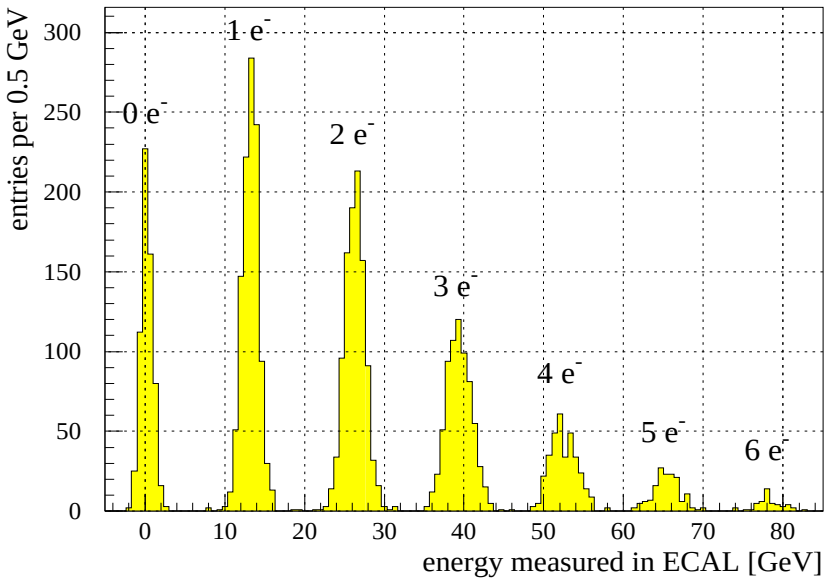


Figure 6: Energy measured by the calorimeter ECAL during a calibration run with 13 GeV electrons.

the electron beam where they could be detected in a similar calorimeter (PCAL). High-energy backscattered photons were detected by monitor CCM1 (Fig 5b) which observed Čerenkov light from the conversion of the photons in 0.2 radiation lengths of aluminum. Scattered electrons in the range 30-40 GeV were detected in Čerenkov monitors EC31 and EC37 of similar construction.

The Silicon-Tungsten calorimeters are segmented vertically and horizontally in 12 rows and 4 columns of 1.6 cm $\times$ 1.6 cm pads and in four longitudinal groups of 23 radiation lengths total thickness. The calorimeter energy resolution is $\sigma_E/E \approx 0.25/\sqrt{E(\mathrm{GeV})}$, whereas the size of the pads resulted in a momentum resolution of $\sigma_p/p \approx 0.04$. Both ECAL and PCAL were calibrated in parasitic running of the FFTB to the SLC program in which linac-halo electrons of energies between 5 and 25 GeV were transmitted by the FFTB when tuned to a lower energy. The number of such electrons varied between 1 and 100 per pulse, which provided an excellent calibration of the ECAL and PCAL over a wide dynamic range. Figure 6 shows the ECAL response to a 13 GeV test beam. The peaks corresponding to events with 0 to 6 electrons per beam bunch can easily be distinguished. The calibration runs also allowed a check of the field maps of the FFTB dump magnets that are used in our spectrometer.

2.2 Phase II and III

The setup of future phases of the experiment are sketched in Fig. 7 and Fig. 8. In the second phase a thin foil or wire will convert high-energy Compton photons to pairs that will be analyzed in a pair spectrometer based on CCD's. The CCD pair spectrometer, sketched in Fig. 9, will reconstruct the photon-energy spectrum with resolution sufficient to discern the effective mass $\overline{m}$.

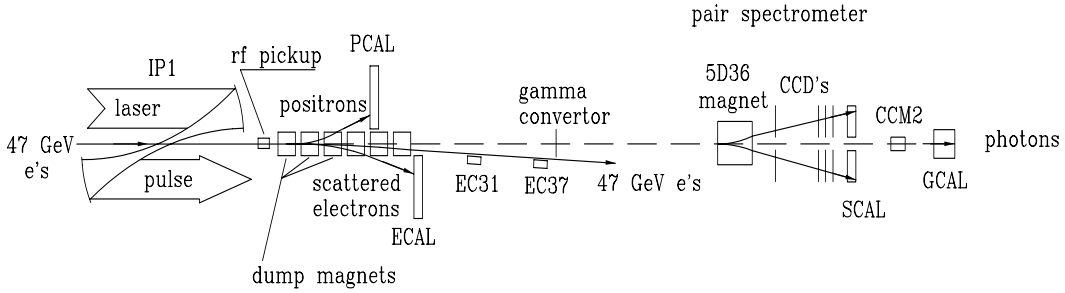


Figure 7: Sketch of the experiment with the addition of a pair spectrometer to analyze converted Compton photons.

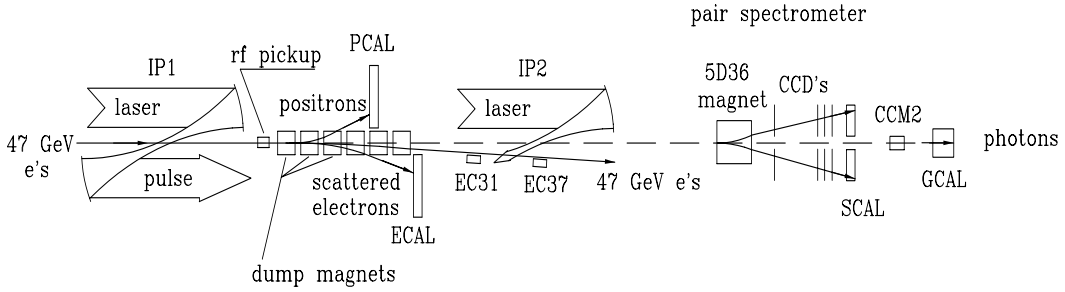


Figure 8: Sketch of the experiment with the addition of a second laser interaction point to study pair creation by light.

In a third phase (Fig. 8), part of the laser beam will collide with the high-energy Compton photons at a new interaction point, IP2, and the invariant mass of resulting pairs will be analyzed in the pair spectrometer free from backgrounds of electrons and positrons produced at IP1.

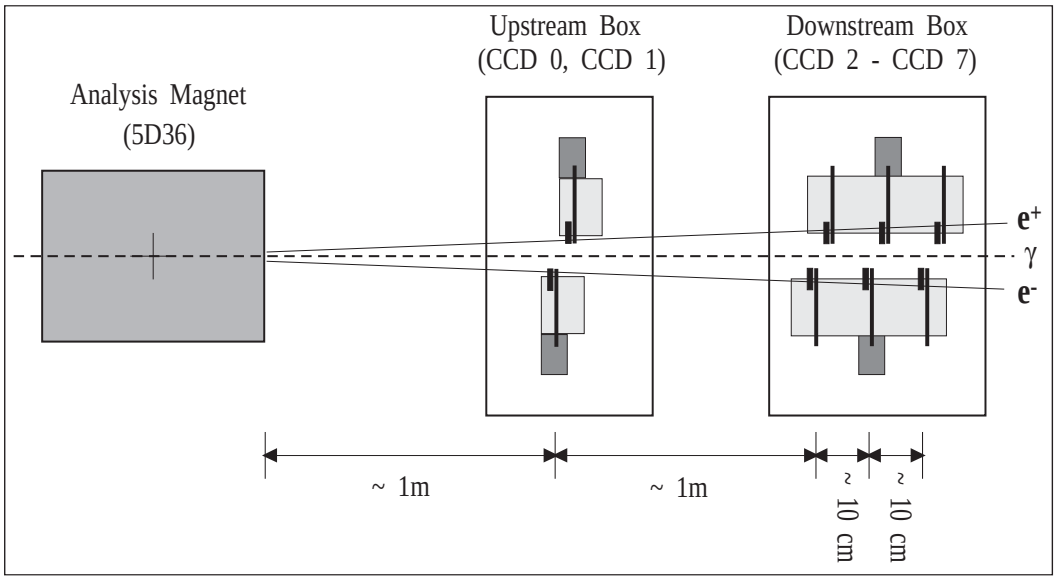


Figure 9: The CCD pair spectrometer.

3 The Laser System

The beam from a chirped-pulse-amplified terawatt Nd:glass laser system[24,25] is focused by off-axis-parabolic mirrors of 30 cm focal length with a 17° crossing angle onto the electron beam at IP1. The laser system is shown in Fig. 10 and delivered 1.5-ps (fwhm) wide pulses at 0.5 Hz of up to 1.2 J of infrared light, or 1 J of green light after frequency doubling in a KDP crystal. The relatively high repetition rate is achieved in a final laser amplifier with slab geometry.[26]

The laser-oscillator mode locker is synchronized to the 476-MHz drive of the SLAC linac klystrons via a rf/optical feedback system.[27] The observed jitter between the laser and linac pulses was 2 ps (rms).[28] The laser-pulse energy and area were measured for each shot. The laser pulse length was available for each shot during infrared running and as averages over short time intervals for green.

The peak focused laser intensity was obtained for infrared pulses of energy $U = 800$ mJ, focal area $A = 60$ μm^2 and pulse width $\Delta t = 1.5$ ps, for which $I = U/A\Delta t \approx 10^{18}$ W/cm^2 at $\lambda = 1054$nm, corresponding to a value of $\eta = 0.6$. Electrons that passed through the focus of the laser at peak intensity had a 25% probability of interacting.

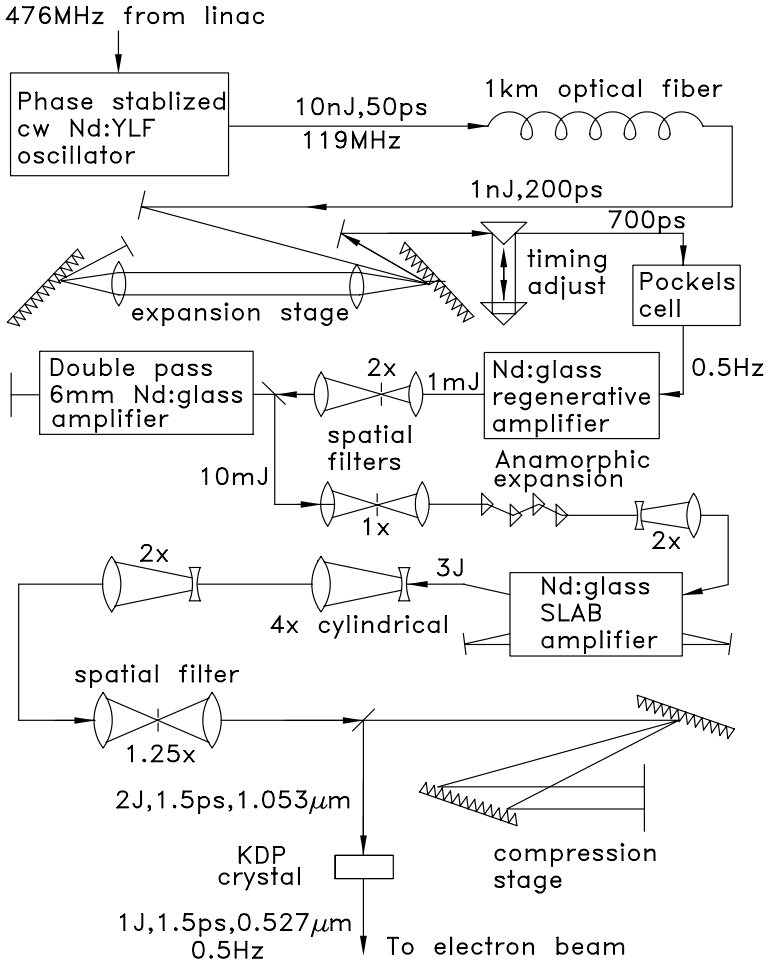


Figure 10: Sketch of the terawatt Nd:glass laser system.

4 Laser Pulse and Electron Bunch Overlap

The electron beam was operated at 10-30 Hz with an energy of 46.6 GeV and emittances $\varepsilon_x = 3 \times 10^{-10}$ m-rad and $\varepsilon_y = 3 \times 10^{-11}$ m-rad. The beam was tuned to a focus with $\sigma_x = 60$ μm and $\sigma_y = 70$ μm at the laser-electron interaction point. The electron bunch length was expanded to 3.6 ps (rms) to minimize the effect of the time jitter between the laser and electron pulses. Typical bunches contained 5×10^9 electrons. However, since the electron beam was significantly larger than the laser focal area only a small fraction of the electrons crossed through the peak field region.

The spatial and temporal overlap of the electron and laser beams was monitored by observing the Compton scattering rate in the ECAL and CCM1 detectors during horizontal (x), vertical (y) and time (t) scans of one beam across the other. Figure 11 shows results of a combined x-t scan. Figure 11a is derived from scattered photons and is dominated by ordinary Compton scattering. The slope of the data agrees with the 17° beam-crossing angle. Figure 11b is derived from electrons of energy less than 25.6 GeV where single Compton scattering does not contribute. The peak in Fig. 11b has a smaller space-time extent than that in Fig. 11a because the nonlinear process is more probable in the higher intensity regions of the laser beam.

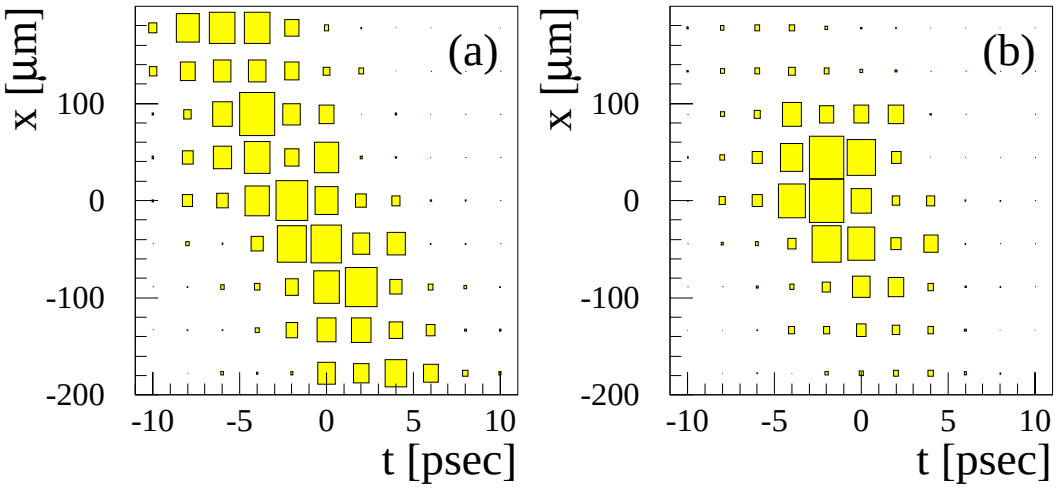


Figure 11: Observed rates of (a) ordinary and (b) nonlinear and multiple Compton scattering as a function of x and t offsets between the electron and laser beams. The area of each box is proportional to the signal size.

5 First Results

5.1 Electron Beam Polarization Measurement

In the commissioning of the present experiment in April 1994 a measurement was made of the longitudinal polarization of the electron beam. For this measurement, data were collected with circularly polarized green laser pulses of $\sim$ 3 mJ energy and $\sim$ 50 ps pulse width. To minimize the effect of shower spreading in the calorimeter, only the signal from the second longitudinal layer of ECAL (out of 23 layers) was used as a measure of the number of incident electrons.

The top row of ECAL was centered at $E = 25.6$ GeV, the electron energy corresponding to the zero crossing of the Compton asymmetry

$$A(E) \equiv \frac{N_+(E) - N_-(E)}{N_+(E) + N_-(E)}$$

where $N_+(E)$, $N_-(E)$ refer to the signal in layer 2 of ECAL for events with electron polarization along/against the momentum vector.

Figure 12 shows that the measured Compton asymmetries in the top 4 ECAL rows are in good agreement for the two data sets taken with the right and left circularly polarized laser.

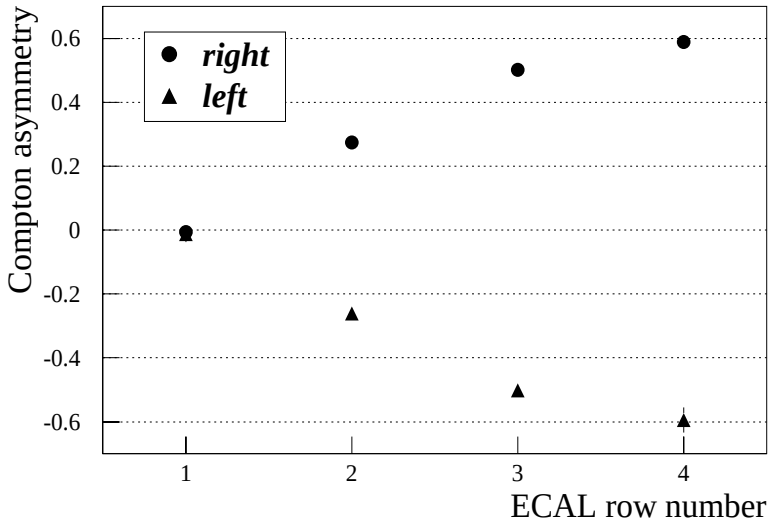


Figure 12: Measured Compton asymmetry in top 4 ECAL rows for right and left circularly polarized laser beams.

A fit of the observed asymmetries gave the result $P_e = 0.81^{+0.04}_{-0.01}$ for the longitudinal polarization of the electron beam,[30] in good agreement with measurements of the SLD collaboration. The upper error of 0.04 on the polarization is due to the uncertainty in the degree of circular polarization of the laser, and could readily be reduced to 0.01 in any future measurements.

5.2 Nonlinear Compton Scattering

Nonlinear effects in Compton scattering were investigated by detecting the scattered electrons. The ECAL sampled the scattered electrons in energy intervals about 1.5 GeV wide. The highest energy sampled was 30 GeV, but the maximum sampled energy could be reduced by lowering the entire calorimeter away from the beam. When positioned with maximum energy below 25.6 GeV, only electrons from nonlinear scattering were detected.

An ECAL channel saturated at 12 TeV, while at peak laser intensity some 10^7 Compton scatters occur per pulse. Hence the ECAL could not be used to study ordinary Compton scattering for laser intensities higher than about 0.001 of peak. Shower cross-talk between calorimeter pads and backsplash from ordinary Compton-scattered electrons that hit components of the beamline limited the dynamic range of ECAL to about 100:1. Because of this and the rapidly decreasing electron yield at lower energies only data from the top 4 calorimeter rows were used in the analysis. Thus the complete mapping of the nonlinear Compton spectrum required data collection at several laser intensities and positions of the ECAL. Figure 13 summarizes the data collection strategy for runs with the infrared laser beam. The accessible range of the scattered electron energy versus the laser intensity is shown as the white area. In the dark shaded area some of the ECAL channels would saturate, while the light shaded area corresponds to signals in ECAL pads dominated by cross-talk and background.

Data were collected with circularly polarized beams at laser pulse energies between 14 and 800 mJ at $\lambda_0 = 1054$ nm, and between 7 and 320 mJ at 527 nm. The energy measured in the calorimeter pads, each of which accepted a limited momentum bite, gave the spectrum of electrons scattered in that pulse. Corrections were applied for shower cross-talk between calorimeter pads, and for backgrounds from high energy Compton scattered electrons that hit beamline components. Two methods were used to estimate the corrections, based on shower spread in-

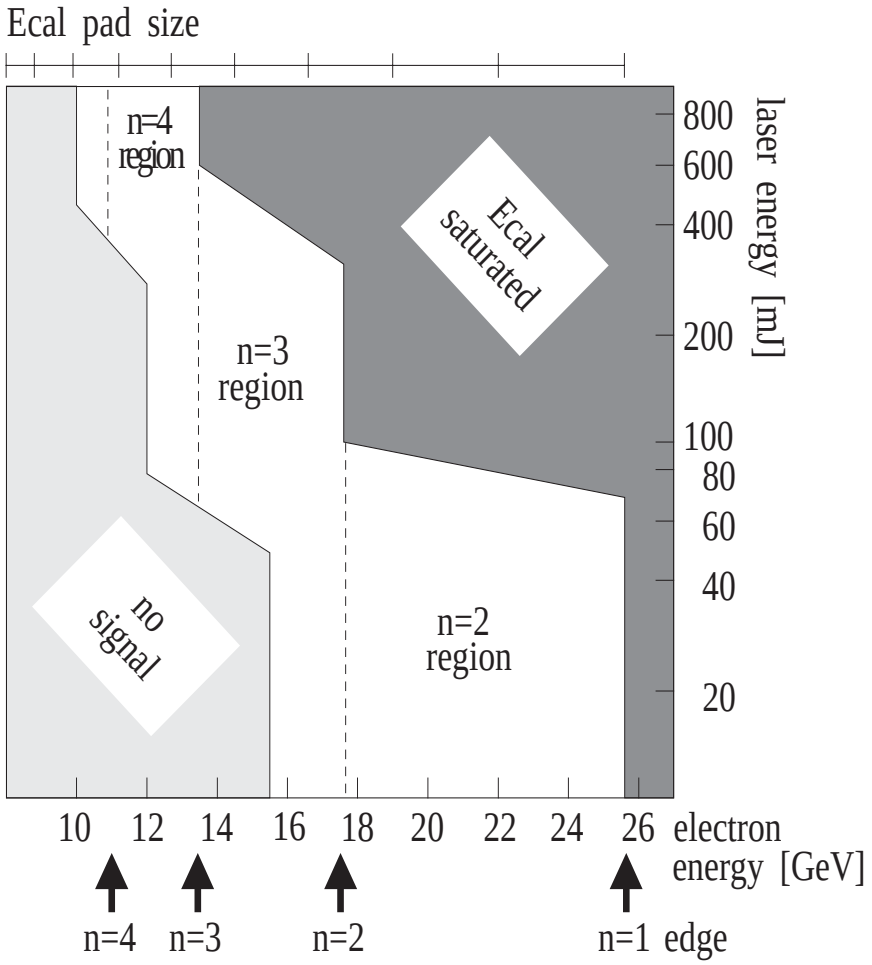


Figure 13: Data collection strategy for the infrared laser beam. The size of an ECAL pad is shown at the top of the figure. The minimum energy of an electron scattered off n laser photons is indicated at the bottom.

formation from calibration runs and on signal in calorimeter channels outside the acceptance for Compton scattering. The average of the two methods is used, and the difference is taken as a contribution to the systematic uncertainty.

Because of the time jitter between the electron and laser pulses the interaction flux was not readily determined from beam measurements alone. Instead, we use the rate of Compton-scattered photons, N_γ, measured by CCM1 as a normalization. To first order the normalized rate equals the normalized cross section:

$$\frac{1}{N_\gamma}\frac{dN}{dE} \approx \frac{1}{\sigma}\frac{d\sigma}{dE},$$

where σ is the total cross section which is close to the ordinary Compton cross section, $\sigma_C = 1.9 \times 10^{-25}$ cm^2 for infrared and 3.0×10^{-25} cm^2 for green.

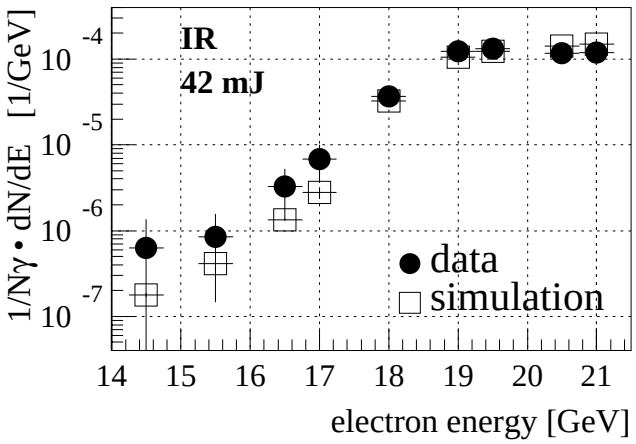


Figure 14: Energy spectra of scattered electrons as observed in the ECAL calorimeter for infrared laser pulses of 42 mJ energy.

In Fig. 14, the rate of scattered electrons normalized to the Compton γ-ray rate is plotted against the electron energy, for infrared laser pulses with a nominal energy of 42 mJ. The open squares represent a simulation of each pulse using the corresponding laser and electron beam parameters at the collision point. The simulation includes both nonlinear and multiple ordinary Compton scatterings. Only energies below the minimum for single Compton scattering are shown. The plateau at 19-21 GeV corresponds to two-photon scatters, and the fall-off at 17-18 GeV is evidence for the two-photon kinematic limit at 17.6 GeV as smeared by the spatial resolution of the calorimeter.

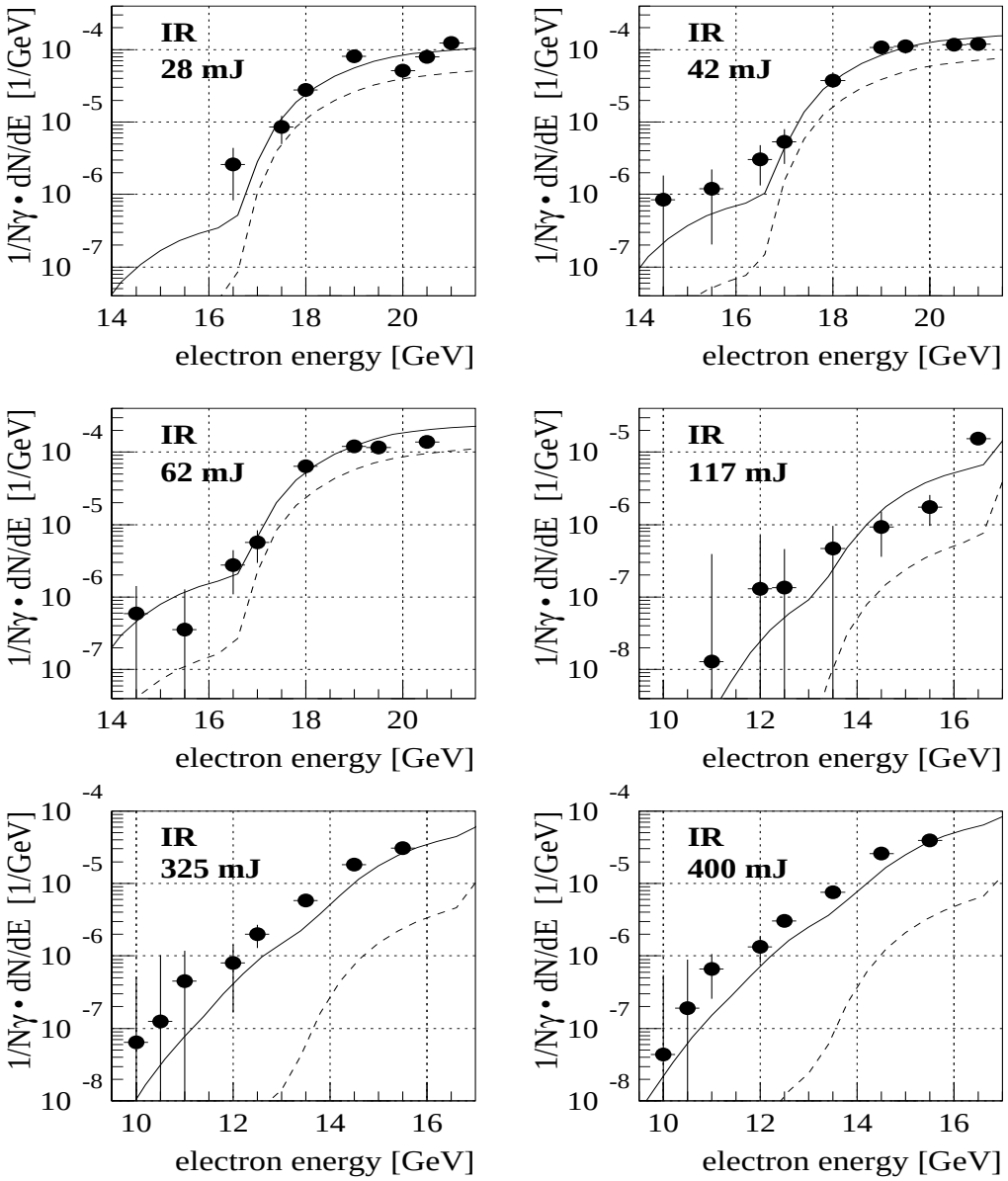


Figure 15: Energy spectra of scattered electrons for infrared laser pulses with circular polarization and nominal energies between 28 mJ and 400 mJ. The data (filled-in circles) has been scaled to standard values of the interaction geometry. The solid line represents the simulation and the dashed line shows the simulated contribution for multiple ordinary Compton scattering only.

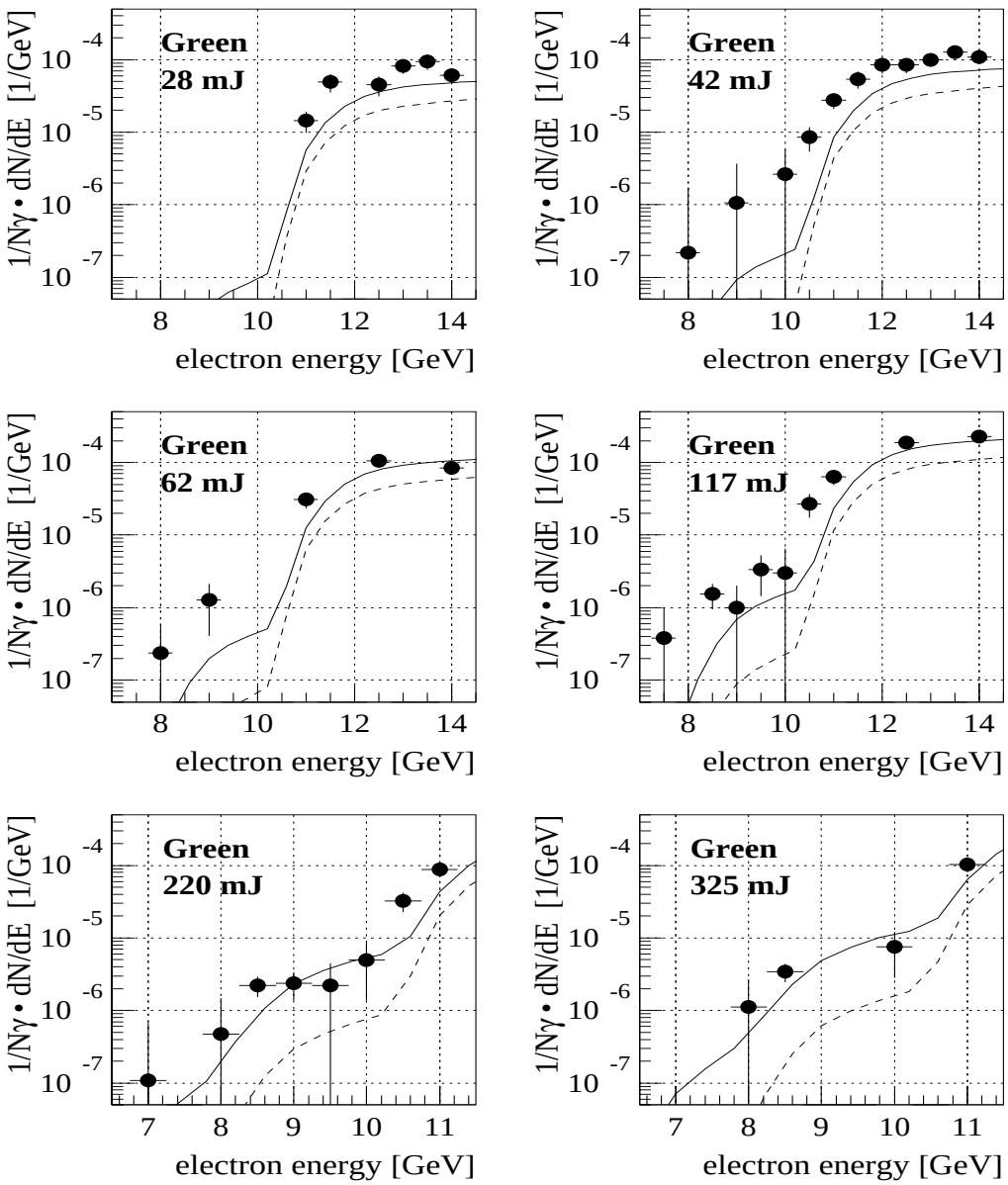


Figure 16: Energy spectra of scattered electrons for green laser pulses with circular polarization and nominal energies between 28 mJ and 325 mJ. The data (filled-in circles) has been scaled to standard values of the interaction geometry. The solid line represents the full simulation and the dashed line shows the simulated contribution for multiple ordinary Compton scattering only.

To compensate for small variations in the beam parameters during the run, the data in Figs. 15-17 have been scaled by the ratio of the simulated rates at observed and standard values of electron and laser beam-spot dimensions. Figure 15 shows results from infrared data at 6 laser energies differing by more than an order of magnitude. The full simulation is shown as the solid curve. The rate calculated for multiple ordinary Compton scattering is shown as the dashed curve which clearly cannot account for the observations. The kinematic limit for $n = 3$ scattering at 13.5 GeV cannot be resolved in the data, but the expected effect is only a very small shoulder in the rate. The two last plots at laser pulse energies of 325 and 400 mJ show proof of $n = 4$ scattering in the momentum range of 11-13 GeV. Figure 16 shows similar results for green data at 6 laser energies between 28 and 325 mJ. The $n = 2$ plateau at 12-14 GeV as well as the $n = 2$ kinematic limit at 10.9 GeV can be discerned in the data. The data points between 8 GeV and 10 GeV in the plots with the highest laser intensities are evidence of $n = 3$ scattering in green data.

In Fig. 17 we illustrate the rise in the normalized nonlinear rate with laser intensity. As the rates are normalized to the total Compton-scattering photon signal which is primarily ordinary Compton scattering, data at electron energies dominated by order n should vary with laser pulse intensity as I^{n-1}. The shaded bands shown for each electron momentum represent the simulation including an uncertainty in laser intensity of $\Delta I/I = 0.3$ for infrared and $\Delta I/I = ^{+0.5}_{-0.3}$ for green laser pulses. The $n = 2$ and $n = 3$ data sets in Fig. 17a and the $n = 2$ set in Fig. 17b agree reasonably well with expectations for the slopes as well as the magnitudes of the rates. For the lowest electron momenta shown in Figs. 17a and 17b, only the data at the highest laser intensities represents a signal well above background and therefore the observed slope does not agree well with expectations.

The error bars shown in Fig. 14 represent statistical uncertainty in the number of scattered electrons and the systematic uncertainty in the correction for backgrounds in the calorimeter. In Figs. 15-17 the error bars also include uncertainties in the scaling to standard beam conditions.

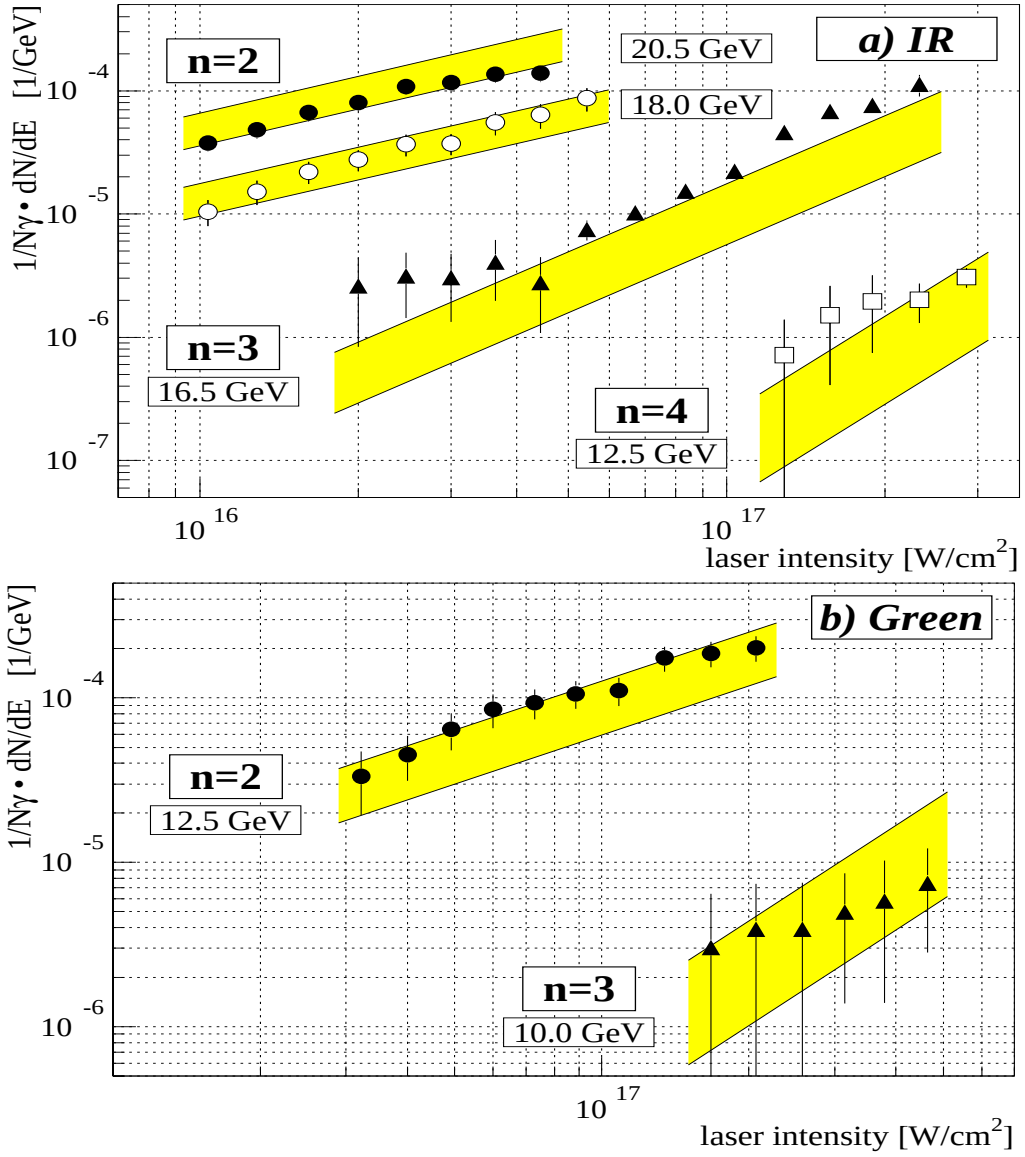


Figure 17: The normalized rate of scattered electrons of energies corresponding to n = 2, 3 and 4 laser photons per interaction versus the intensity of the laser field at the interaction point for (a) infrared and (b) green laser pulses. The bands represent a simulation of the experiment including an uncertainty in laser intensity of $\Delta I/I = 0.3$ for infrared and $\Delta I/I = ^{+0.5}_{-0.3}$ for green.

6 Conclusion

We measured the longitudinal polarization of the electron beam at the FFTB and found good agreement with measurements of the SLD collaboration.

We observed at two different laser wavelengths a clear signal for nonlinear Compton scattering in the spectrum of the scattered electrons. At the highest laser intensities achieved, up to 4 laser photons were absorbed in a single scattering event. The dependence of the scattered electron rate on electron momentum and laser intensity agree within experimental uncertainty with theory[10] over a wide range of laser pulse energies.

References

[1] G.A. Schott, *Electromagnetic Radiation*, (Cambridge University Press, 1912).

[2] T.J. Englert and E.A. Rinehart, Phys. Rev. A **28**, 1539 (1983).

[3] See, for example, M. Billardon *et al.*, J. de Phys., Colloq. **44**, C1 (1983).

[4] D.M. Volkov, Z. Phys. **94**, 250 (1935).

[5] See also V.B. Berestetskii *et al.*, *Quantum Electrodynamics*, 2nd ed., (Pergamon Press, 1982), §40.

[6] N.D. Sengupta, Bull. Math. Soc. (Calcutta) **44**, 175 (1952).

[7] L.S. Brown and T.W.B. Kibble, Phys. Rev. **133**, A705 (1964).

[8] A.I. Nikishov and V.I. Ritus, Sov. Phys. JETP **19**, 529, 1191 (1964); **20**, 757 (1965); see also §101 of ref.[5]

[9] I.I. Gol'dman, Sov. Phys. JETP **19**, 954 (1964); Phys. Lett. **8**, 103 (1964).

[10] N.B. Narozhny *et al.*, Sov. Phys. JETP **20**, 622 (1965).

[11] T.W.B. Kibble, Phys. Rev. Lett. **16**, 1054 (1966); Phys. Rev. **150**,1060 (1966).

[12] J.H. Eberly and A. Sleeper, Phys. Rev. **176**, 1570 (1968).

[13] C.I. Moore *et al.*, Phys. Rev. Lett **74**, 2439 (1995).

[14] F. Sauter, Z. Phys. **69**, 742 (1931). See also §129, prob. 2 of ref.[5]

[15] O. Klein, Z. Phys. **53**, 157 (1929).

[16] W. Heisenberg and H. Euler, Z. Phys. **98**, 718 (1936).

[17] G. Breit and J.A. Wheeler, Phys. Rev. **46**, 1087 (1934).

[18] H.R. Reiss, J. Math. Phys. **3**, 59 (1962). Phys. Rev. Lett. **26**, 1072 (1971).

[19] M. Bell and J.S. Bell, Part. Acc. **24**, 1 (1988); R. Blankenbecler and S.D. Drell, Phys. Rev. Lett. **61**, 2324 (1988); M. Jacob and T.T. Wu, Nucl. Phys. **B303**, 373, 389 (1989); **B327**, 285 (1989); P. Chen and K. Yokoya, Phys. Rev. Lett. **61**, 1101 (1988); V.N. Baier *et al.*, Nucl. Phys. **B328**, 387 (1989); P. Chen and V.L. Telnov, Phys. Rev. Lett. **63**, 1796 (1989); R. Blankenbecler *et al.*, Phys. Rev. D **40**, 2462 (1989).

[20] I.F. Ginzburg *et al.*, Nucl. Instr. and Meth. **205**, 47 (1983).

[21] K.O. Mikaelian, Phys. Rev. D **17**, 750 (1978); **30**, 1115 (1984); I.F. Ginzburg *et al.*, Nucl. Phys. **B228**, 285 (1983).

[22] P. Chen and R.B. Palmer, SLAC-PUB-5966 (Nov. 1992).

[23] V. Balakin *et al.*, Phys. Rev. Lett. **74**, 2479, (1995).

[24] D. Strickland and G. Mourou, Opt. Comm. **55**, 447 (1985).

[25] C. Bamber *et al.*, U. Rochester preprint UR-1428 (June 1995).

[26] W.S. Martin and J.P. Chernoch, U.S. Patent 3633126 (1972).

[27] M.J.W. Rodwell *et al.*, Opt. Lett. **11**, 638 (1986).

[28] T. Kotseroglou *et al.*, U. of Rochester preprint UR-xxxx, in preparation.

[29] S.C. Berridge *et al.*, IEEE Trans. Nucl. Sci. **37**, 1191 (1990).

[30] M.B. Woods *et al.*, E144 Internal Note (Oct. 17, 1994).

Heavy Quark Physics from SLD

Robert Messner*
Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

Representing the SLD Collaboration

ABSTRACT

This report covers preliminary measurements from SLD on heavy quark production at the Z^0, using 150,000 hadronic Z^0 decays accumulated during the 1993—1995 runs. A measurement of R_b with a lifetime double tag is presented. The high electron beam polarization of the SLC is employed in the direct measurement of the parity-violating parameters A_b and A_c by use of the left-right forward-backward asymmetry. The lifetimes of B^+ and B^0 mesons have been measured by two analyses. The first identifies semileptonic decays of B mesons with high (p,p_t) leptons; the second analysis isolates a sample of B meson decays with a two-dimensional impact parameter tag and reconstructs the decay length and charge using a topological vertex reconstruction method.

*This work supported by the Department of Energy, contract DE-AC03-76SF00515 (SLAC).

©1995 by Robert Messner.

Introduction

The Stanford Linear Collider (SLC) together with the SLC Large Detector (SLD) provides an excellent facility for the study of the decay of the Z^0. The significant developments over the past year have been the most fundamental: the SLC has been able to deliver increased luminosity to the SLD detector at higher electron polarization. This progress is illustrated in Fig. 1, which shows the increase in the average polarization of the electron beam over time. During the 1993 running period, the SLD collected ~50,000 Z^0 decays with a mean electron beam polarization of (63 ± 1)%. In 1994-1995, SLD recorded an additional ~100,000 decays with a mean longitudinal polarization of (77 ± 1)%. Combined with the analysis advantage of a small, stable beam spot and the superior three-dimensional resolution of the SLD vertex detector, this data set has provided accurate measurements of fundamental electroweak parameters. We will cover preliminary results for a set of topics SLD has studied involving heavy quark physics. These include a measurement of R_b, heavy quark asymmetries, and a measurement of B meson lifetimes utilizing two analysis methods.

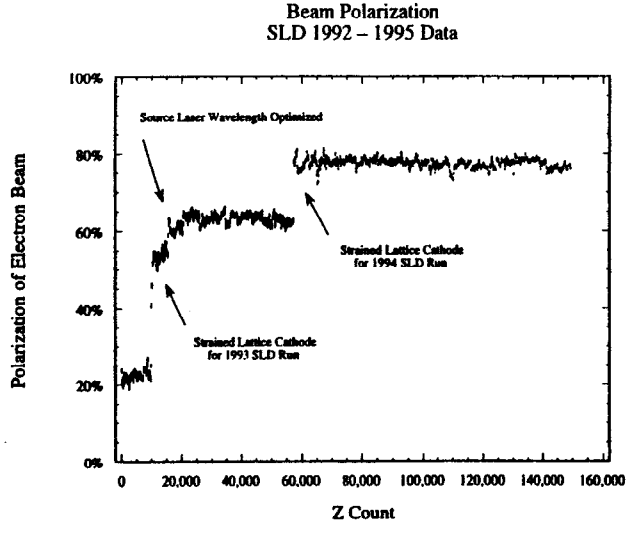


Figure 1. A summary of the beam polarization and the Z^0 decays recorded by the SLD.

A side view of the SLD is given in Fig. 2; the interaction between the SLC and the SLD and the production and measurement of the electron beam polarization are covered in the talk by T. Schalk in these proceedings.

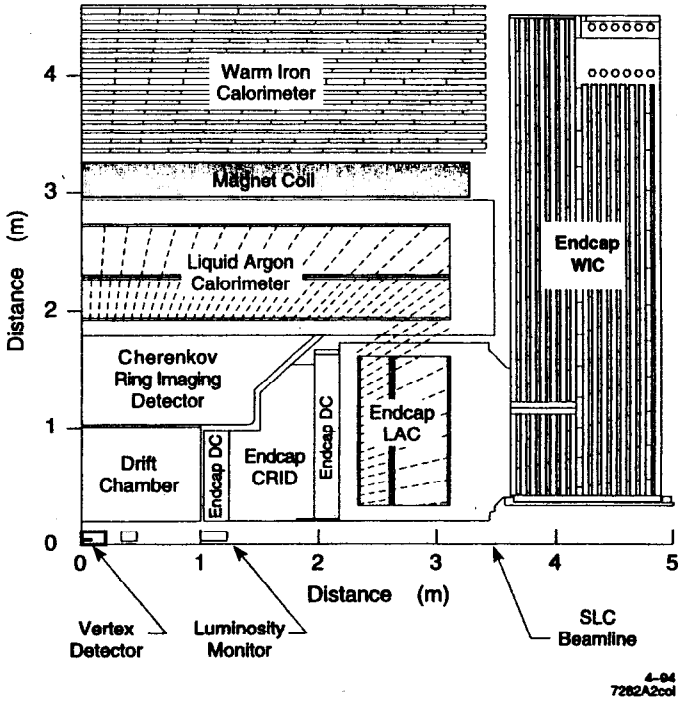


Figure 2. A quadrant of the SLD detector.

The SLD Detector

A detailed description of the SLD detector can be found in Ref. [1]. Working from the interaction point (IP) outward, the detector incorporates a vertex detector (VXD) for precise position measurements close to the interaction region, a central drift chamber (CDC) for charged particle tracking and momentum measurements, a Cherenkov ring imaging detector (CRID) for particle identification, a liquid argon calorimeter (LAC) for measuring energy flow and providing electron identification, the solenoid coil, and the warm iron calorimeter (WIC), which provides the magnetic field flux return and muon identification, and serves as a tail catcher for hadronic showers extending beyond the LAC.

Together, the VXD and the CDC provide the core of the SLD tracking measurements used in the analyses to be discussed later.[2] The VXD consists of 480 charge-coupled devices (CCD's) surrounding a 1 mm thick beryllium beam pipe with an inner radius of 25 mm. Each CCD is an array of 375 x 578 square pixels 22 μm on a side. The

CCD's are arranged in four concentric cylinders at radii ranging from 2.9 cm to 4.1 cm. The inner (outer) cylinder covers a range of polar angles defined by $\cos\theta < 0.85$ (0.75). The CDC is a cylinder 1.8 m long with an inner radius of 0.2 m and an outer radius of 1.0 m. Six hundred and forty drift cells are arranged in ten superlayers covering radii from 24 cm to 96 cm. Each cell in a superlayer has eight sense wires spaced radially by 5 mm. An individual sense wire provides a measurement of the drift distance with a spatial resolution averaging 70 μm over the entire drift cell. Tracks are reconstructed at polar angles in the range $\cos\theta < 0.85$.

The capabilities of the VXD are summarized in Figures 3 and 4. Figure 3 displays the measured miss distance between the two tracks of muon pairs obtained from decays of the $Z^0 \rightarrow \mu^+\mu^-$. Since the two muons originate from a common point, this is a good measure of the intrinsic resolution for an individual track's impact parameter relative to its origin; the width of the distribution is divided by $\sqrt{2}$ because there are two tracks used to make the measurement. Correcting for this, the resolution in the r-ϕ plane is found to be 11.2 microns, and in the r-z plane 37 microns.

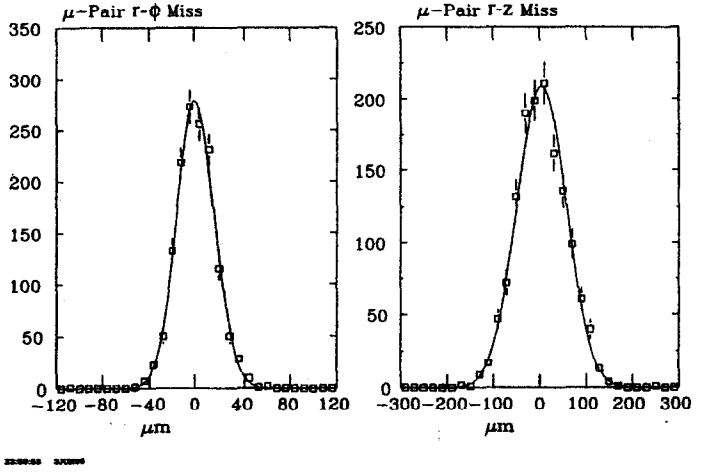


Figure 3. The miss distance between muons in the r-ϕ and r-z planes.

Figure 4 demonstrates the accuracy of the beam spot determination. It plots the impact parameter for individual muon tracks, again from $Z^0 \rightarrow \mu^+\mu^-$ pairs, relative to the position of the IP in the r-ϕ plane. The IP position is determined independently from hadronic events. The width of the distribution is 12.9 microns. Unfolding the intrinsic impact parameter resolution of the tracks, we find the accuracy for the IP spot determination to be 6.4 microns.

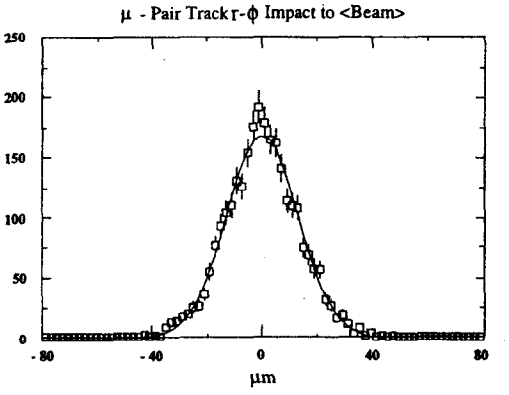


Figure 4. The impact parameter for muons relative to the IP.

Couplings of Fermions to the Z^0

The Standard Model (SM) makes definite predictions for a number of attributes of Z^0 decays, making possible precision tests by comparing these with the experimental values. Possible new physics beyond the SM may make itself manifest through radiative corrections to the SM couplings. An elegant and thorough review of the status of world measurements and the implications for the SM is given in the lectures by M. Swartz in these proceedings (see also Ref. [3]). We concentrate here on couplings of the Z^0 to heavy quarks. Tests using measurements of heavy quark production, especially that of the b quark, are primarily sensitive to corrections at the $Z^0 \rightarrow f\bar{f}$ vertex. This is distinct from measurements of $\sin^2\theta_w$, such as the electron asymmetry or tau polarization asymmetry, which are tests sensitive to corrections that involve the Z^0 propagator. The two types of measurements complement one another in the nature of their probing of the SM.

R_b: Lifetime Double Tag

A quantity of particular current interest is the branching fraction of the Z^0 into b quarks. Significant top mass corrections are expected to affect the cross section for $Z^0 \rightarrow b\bar{b}$. The best experimental quantity to compare with predictions is R_b, the ratio $\Gamma(Z^0 \rightarrow b\bar{b}) / \Gamma(Z^0 \rightarrow \text{hadrons})$. In this ratio, most corrections excepting vertex corrections cancel, so that the theoretical ambiguities are relatively small. Moreover, because the vertex corrections are isolated, the predicted value of R_b becomes a function of the top mass, again with little uncertainty. The ratio has the additional advantage that experimental uncertainties tend to cancel as well. LEP has made measurements of R_b using several techniques, and their current average is 3 σ high compared to the SM. (See the talk by D. Strom in these proceedings for a current review of the LEP measurements.)

The SLD measurement of R_b employs a lifetime double-tag technique similar to the ALEPH measurement.[4, 5] The primary difference lies in the intrinsic resolution of the vertex detector, including its three-dimensional information, and the size of the interaction region. The information of the IP position is utilized in the following manner. After a selection of Z^0 decays has been made, each event is divided into two hemispheres using the direction of the highest momentum jet as the axis. Track parameters are computed using only the information from the VXD and CDC systems; a second analysis of each track is made by adding the beam spot position as an extra hit on the track. A variable, χ, is defined which represents the difference in the square root of the chi-squared of the fit track for the two computations. This is equivalent to the normalized three-dimensional impact parameter to the primary vertex for each track. Tracks that originate close to the IP will have a small value of the impact parameter. Tracks that originate further from the IP, such as those originating from decays of heavy quarks, will have a large value. An additional refinement is made by assigning a sign to χ depending upon whether the point of closest approach of the track to the axis of the highest momentum jet is in the same hemisphere (χ positive) or the opposite hemisphere (χ negative) as the track itself.

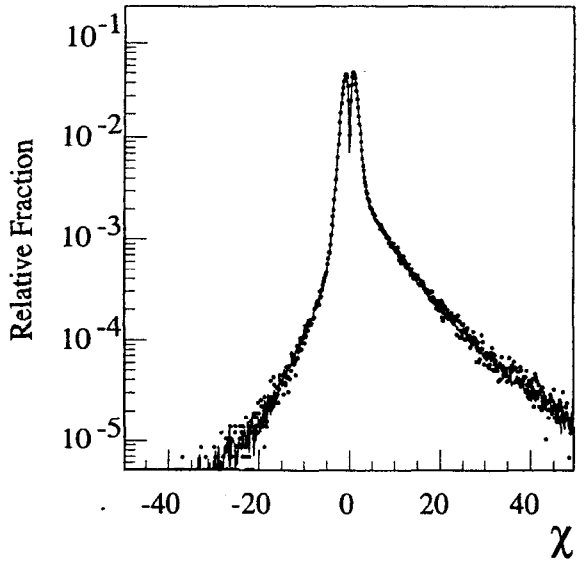


Figure 5. The χ distribution for data and Monte Carlo.

A plot of χ compared with Monte Carlo is shown in Fig. 5. The dip at the origin is due to the lack of phase space for tracks to point exactly at the IP in three dimensions. In the region of positive χ, a tail can be seen at the larger values; these tracks are predominately from particles decaying far from the IP. Decays of particles containing heavy quarks dominate this region. The probability that an individual track originated from the primary vertex is then computed using the shape of the distribution of negative χ, reflected about the origin, for the template. The next step is to look at the ensemble of tracks in a particular hemisphere. Only tracks with positive χ are kept. A joint probability, that the ensemble of tracks in a hemisphere is consistent with coming from the IP, is then formed using a Poisson χ^2 probability distribution.

Hemispheres with a low joint probability represent a data sample enriched in b decays. A cut on the joint probability is then made; b decays are isolated by requiring that events have a joint probability below the value of the cut. Figure 6 shows a plot of the purity, Π_b, of the resultant b sample and the efficiency, ε, for the different quark species as a function of the value used for the cut. The precise value of the cut is determined by that value which minimizes the total error of the resultant data sample for measuring R_b.

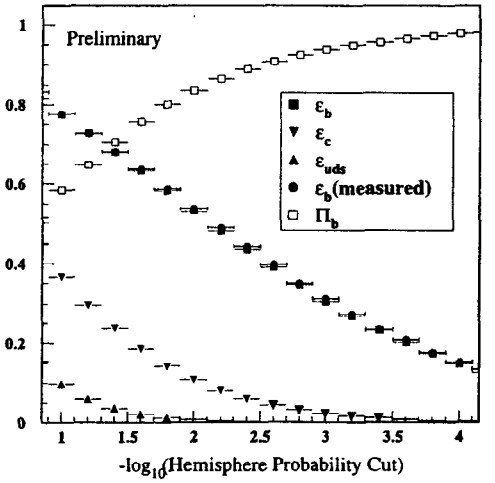


Figure 6. The purity, Π_b, and efficiency, ε_b, of the b sample as a function of the cut on the log of the hemisphere joint probability.

Knowledge of the efficiency for tagging b events is crucial for determining R_b. Note that Fig. 6 has two values plotted for ε_b. One is determined purely by Monte Carlo, while the other has been determined from the data by comparing the number of hemispheres tagged as containing a b quark with the number of events for which both hemispheres are so tagged. One would like the efficiency to be determined completely by the data if so possible. In a perfect world, if there were no background so that the b tags represented a pure b quark sample and the tagging of either of the two hemispheres was independent of the other, R_b and the efficiency, ε_b, could be obtained by solving the following two simultaneous equations:

$$n_{hemi} = 2\varepsilon_b R_b N_{event} \text{ and } n_{double} = \varepsilon_b^2 R_b N_{event}.$$

In practice, however, non-b related backgrounds are present and their effect must be included. Similarly, the tagging efficiency of the two hemispheres has a nonzero degree of correlation between them due to a number of causes; a simple example would be the acceptance of the detector, while others include the effects of gluon radiation or small errors in the determination of the IP position. The correlation likewise must be accounted for. The resultant equations are more complicated, but R_b and ε_b can still be extracted with input from the Monte Carlo for the values of ε_{uds}, ε_c, and the correlations between hemisphere efficiencies, and assuming a SM value of R_c =0.171.

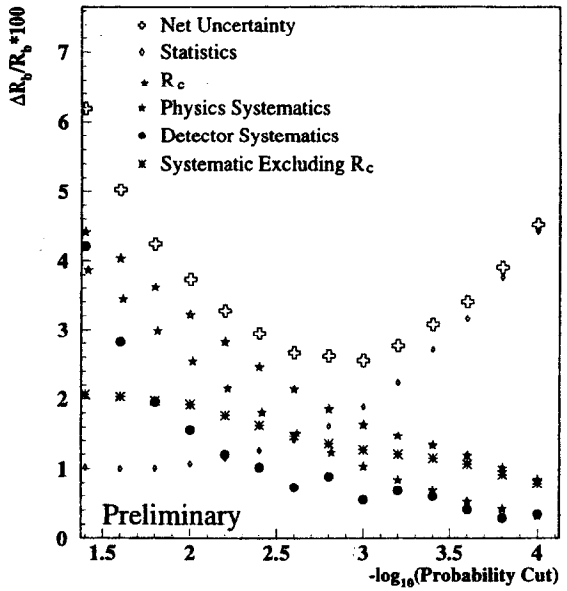


Figure 7. The uncertainty in R_b as a function of the cut on the log of the hemisphere joint probability.

The systematic and statistical errors for the value of R_b extracted from the data set are shown in Fig. 7. The choice of the value for the cut on the joint probability is set at the minimum in the curve for the net uncertainty in the value of R_b. This represents a balance between the statistics of the final sample and the systematic errors of the analysis. The variation in the result for R_b as the cut is changed is shown in Fig. 8; the flatness of this curve gives us confidence in the stability of the result.

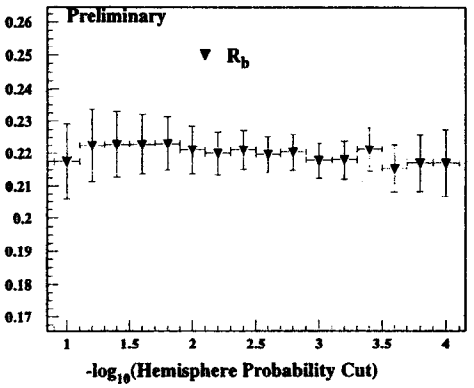


Figure 8 . The dependence of the result for R_b upon the value of the cut on the log of the hemisphere joint probability.

The minimum combined systematic and statistical uncertainty occurs at a hemisphere joint probability cut of $10^{-3.0}$. The efficiency ε_b for this cut measured from data is 31.3 ± 0.6 %, with a b-tagging purity Π_b of 94%. The preliminary result for R_b is

$$R_b = 0.2171 \pm 0.0040 \text{ (stat)} \pm 0.0037 \text{ (sys)} \pm 0.0023 \text{ } (R_c).$$

The value is consistent with the expectation of the SM. The SLD result is not yet precise enough either to certify or to de-certify the 3σ discrepancy currently observed by the LEP experiments.

Table 1

Systematic Source	$\delta R_b/R_b$	Systematic Source	$\delta R_b/R_b$
Correlation Systematics		Charm Systematics	
B lifetime	0.03%	D lifetime	0.15%
B-decay multiplicity	0.14%	D-decay multiplicity	0.87%
B → D model	0.02%	c → D fractions	0.57%
b fragmentation	0.30%	c fragmentation	0.56%
Λ_b fraction	0.31%	$R_c = 0.171 \pm 0.017$	1.05%
MC statistics	0.41%	MC statistics	0.26%
Light quark systematics		Detector systematics	
K, Λ production	0.04%	V^0 rejection	0.85%
g → cc̄, bb̄	0.19%	Tracking efficiency	0.41%
MC statistics	0.11%	Impact resolutions	0.35%
		Beam position tails	0.29%

A breakdown of the sources of the systematic error is given in Table 1. The largest contributors are seen to lie in the charm sector. Note that the systematic error due to the uncertainty in the value for R_c has been isolated. Improvements in techniques to remove charm decays will be a key to reducing systematic errors. Further gains in this analysis will require the development of an efficient, ultra-high purity b-tag to improve systematics in hand with a larger sample of data.

Z^0 Asymmetries at the SLD

In the SM, fermions have vector and axial vector couplings v_f and a_f to the Z^0. The asymmetries at the Z-pole depend on the vector-axial vector interference term. Conventionally, one defines for each fermion the quantity

$$A_f = \frac{2v_f a_f}{v_f^2 + a_f^2}.$$

The asymmetries one has access to experimentally involve different combinations of the A_f relevant to a particular process (see M. Swartz, these proceedings.) The advantage of the high polarization of the incident electron beam at the SLC is that it enables us to separate the electron coupling from the fermion coupling to the Z^0.[6] As an example, one can look at the expression for the left-right forward-backward asymmetry $\tilde{A}_{FB}^f$. This is written below for the b quark:

$$\tilde{A}_{FB}^b = \frac{[\sigma_L(F) - \sigma_L(B)] - [\sigma_R(F) - \sigma_R(B)]}{[\sigma_L(F) + \sigma_L(B) + \sigma_R(B) + \sigma_R(F)]} = \tfrac{3}{4} P_e A_b.$$

F (B) denotes an outgoing fermion that goes in the same (opposite) direction as the incident electron. σ_L (σ_R) is the cross section for left (right) polarized electrons colliding with (unpolarized) positrons. Besides providing a means to measure A_b directly, the use of $\tilde{A}_{FB}^b$ has the practical effect of enhancing the analytical power of the events used for the measurement relative to a measurement of the unpolarized forward-backward asymmetry A_{FB}^b. This is because the unpolarized forward-backward asymmetry has the quantity A_e rather than P_e multiplying the final state fermion term:

$$A_{FB}^b = \tfrac{3}{4} A_e A_b.$$

Since A_e is small, the relative statistical sensitivity of the two methods, $(P_e / A_e)^2$, becomes quite large for a highly polarized beam; for the SLD, the enhancement we gain by eliminating the electron coupling is on the order of 25.

The expression for $\tilde{A}_{FB}^f$ above is obtained by integrating the Born-level differential cross section:

$$\frac{\partial \sigma_f}{\partial \cos\theta} \propto (1 - A_e P_e)\left(\frac{1 + \cos^2\theta}{2} + A_f\left(\frac{A_e - P_e}{1 - A_e P_e}\right)\cos\theta\right).$$

Here, θ is the angle between the incoming e^- and the outgoing fermion f. The direction of the thrust axis can be used to determine the direction of the outgoing fermion. The sign of P_e in the differential cross section can also be manipulated to give the left-right forward-backward asymmetry as a function of $\cos\theta$:

$$\tilde{A}_{FB}^f(\cos\theta) = |P_e| A_f \frac{2\cos\theta}{1 + \cos^2\theta}.$$

One needs, in addition to providing a polarized incident electron beam, to select a sample of events enriched in a particular quark flavor and to tag the direction of the quark versus the antiquark to determine if an event is a forward or a backward event.

The majority of the analyses to follow use the full Born-level differential cross section in making the fits. The use of the full form for the cross section in fitting the data allows the analysis to include corrections and detector effects in a more thorough and straightforward manner, while maintaining the ability to eliminate the dependence on the electron coupling as discussed above.

Measurement of A_c Using D^{*+} and D^+

This is an update of the measurement made with the 1993 data.[7] In order to tag $c\bar{c}$ events and separate them from $b\bar{b}$ and uds backgrounds, kinematic and vertex analyses are used to find events containing one of two $D^{*+} \to \pi_s^+ D^0$ decay chains, with $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$. In the latter case, the π^0 is not reconstructed. The π_s^+ in the D^{*+} decay is known as the spectator pion and carries the sign of the charm quark. Vertexing techniques also cleanly isolate the mode $D^+ \to K^-\pi^+\pi^+$. Both analyses benefit from the three-dimensional VXD information and the precise knowledge of the IP location.

In the D^{*+} kinematic analysis, D^0 candidates are formed by cutting on the D^0 invariant mass spectrum. For example, for the $K\pi$ mode, the central value is 1.765 GeV/c^2 < m($K^-\pi^+$) < 1.965 GeV/c^2, and, for the "satellite" peak, we require 1.50 GeV/c^2 < m($K^-\pi^+$) < 1.70 GeV/c^2. A cut is made on the opening angle between the direction of the D^0 candidate in the lab frame and the K in the rest frame of the D^0. After this cut, the remaining candidates are combined with a slow pion having the correct charge and p_π > 1.0 GeV/c. Finally, we take advantage of the fact that $c\bar{c}$ events are produced at high x_{D^*}, defined as $2 \cdot E_{D^*}/E_{cm}$, relative to background; x_{D^*} is required to be ≥ 0.4.

For the D^{*+} vertex analysis, the emphasis is on the fact that D^0's in $c\bar{c}$ events have a significant three-dimensional decay length and are produced at the IP. The cuts on the opening angle and p_π are eliminated, and the x_{D^*} cut is reduced to $x_{D^*} \geq 0.2$. D^0 tracks

are required to have a good three-dimensional vertex fit, with a three-dimensional decay length distinct from the IP by 2.5 σ_L, where σ_L is the error on L. To assure that the direction of the D^0 points to the IP, the two-dimensional impact parameter of the D^0 to the IP is required to be < 20 μm. Together, the vertex and impact parameter cuts strongly reject combinatoric background and D^{*+} from beauty cascades.

A vertex-style analysis is used to isolate $D^+ \rightarrow K^-\pi^+\pi^+$ in $c\bar{c}$ events. A cut is made on the opening angle between the direction of the D^+ candidate in the lab frame and the K in the rest frame of the D^+; x_{D^+} is required to be ≥ 0.4. The three-dimensional decay length measured from the IP is required to be at least 3 • σ_L. Finally, the angle between a line drawn from the D^+ vertex to the IP and the direction of the D must be < 5 mr in the r-φ plane and < 20 mr in the r-z plane.

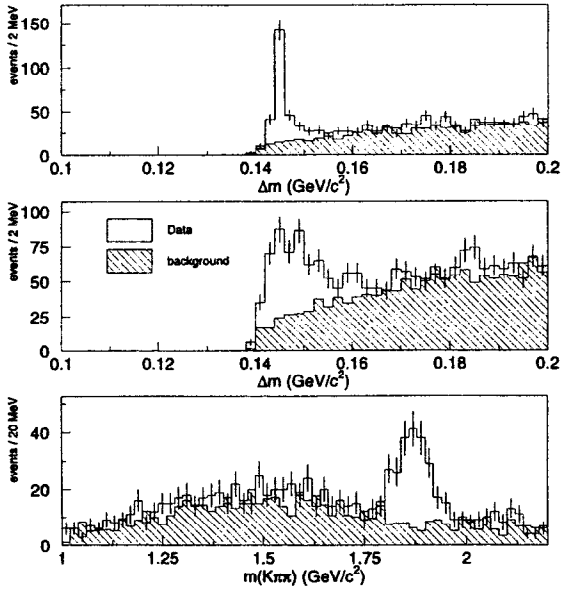


Figure 9. Distributions of the D^{*+} - D^0 mass difference for $D^0 \rightarrow K^-\pi^+$(top) and $D^0 \rightarrow K^-\pi^+\pi^0$ (middle); and the $D^+ \rightarrow K^-\pi^+\pi^+$ mass (bottom).

Plots of the D^{*+} - D^0 mass differences and the $D^+ \rightarrow K^-\pi^+\pi^+$ mass are shown in Fig. 9. A clean signal is extracted in each case. The signal region in the plots of the D^{*+} - D^0 mass difference is taken to be $\Delta m < 0.15$ GeV/c², and the sideband region used is 0.16 GeV/c² < Δm < 0.20 GeV/c². The union of the two analyses is used to determine the asymmetry. The signal region in the mass plot for $D^+ \rightarrow K^-\pi^+\pi^+$ is 1.8 GeV/c² < m($K^-\pi^+\pi^+$) < 1.94 GeV/c², while the sideband regions are 1.64 GeV/c² < m < 1.74 GeV/c² and 2.0 GeV/c² < m < 2.1 GeV/c². The raw asymmetry is plotted in Fig. 10 as a product of the charge of the charm meson and $\cos\theta_D$, where θ_D is the polar angle of the D meson momentum, separately for left- and right-handed electron polarization.

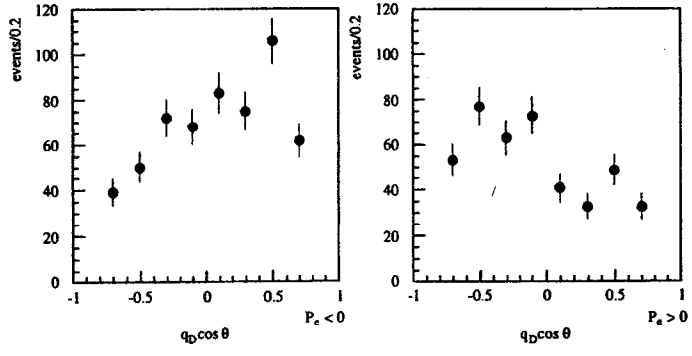


Figure 10. The raw asymmetry for the events in the signal mass region.

A maximum likelihood fit for the combined sample is used to extract A_C, taking into account the information on $D^{(*)}$ momentum-dependent fractions of $c\bar{c}$, $b\bar{b}$ signal, and background components. A_c^D and A_b^D are the asymmetries from $D^{(*)+}$ decays in tagged $c\bar{c}$ and $b\bar{b}$ events, respectively:

$$\ln L = \sum_{i=1}^{n} \ln\{ P_c^j(x_D^i)\left(\frac{1-P_eA_e}{2(A_e-P_e)}(1+y_i^2)+y_iA_c^D\Delta^c\right) + P_b^j(x_D^i)\left(\frac{1-P_eA_e}{2(A_e-P_e)}(1+y_i^2)+y_iA_b^D\Delta^b\right) + P_{RCBG}^j(x_D^i)\left((1+y_i^2)+2A_{RCBG}y^i\right)\} .$$

The preliminary result obtained is:

$$A_C = 0.64 \pm 0.11 \text{ (stat)} \pm 0.06 \text{ (sys)}.$$

The dominant systematics are related to the random combinatoric background (RCBG), as shown in Table 2; this is largely due to limited statistics in the sideband regions. Thus, the systematic errors can be expected to be reduced with a larger data sample.

Table 2

Systematic source	δA_c
RCBG fraction and acceptance	0.039
RCBG asymmetry	0.028
$c \rightarrow D/b \rightarrow D$ fraction	0.011
$b \rightarrow D$ asymmetry	0.022
b, c fragmentation	0.019
Beam polarization	0.007
QCD correction	0.007
Total systematic	0.058

Measurement of A_b, A_c with Leptons

This analysis selects $Z^0 \rightarrow c\bar{c}$ and $b\bar{b}$ events with semileptonic decays. The analysis is an updated measurement based on the analysis of the 1993 data to include all 1993–1995 data.[8] Electrons are identified in the LAC by requiring agreement between the track momentum and the calorimeter electromagnetic energy, little or no calorimeter hadronic energy, and a reasonable front/back electromagnetic energy ratio. Muons are identified by comparing hits in the WIC with the extrapolated track, taking track extrapolation errors and multiple scattering into account. The lepton charge provides the quark anti-quark determination, and the quark direction is obtained from the jet nearest the lepton.

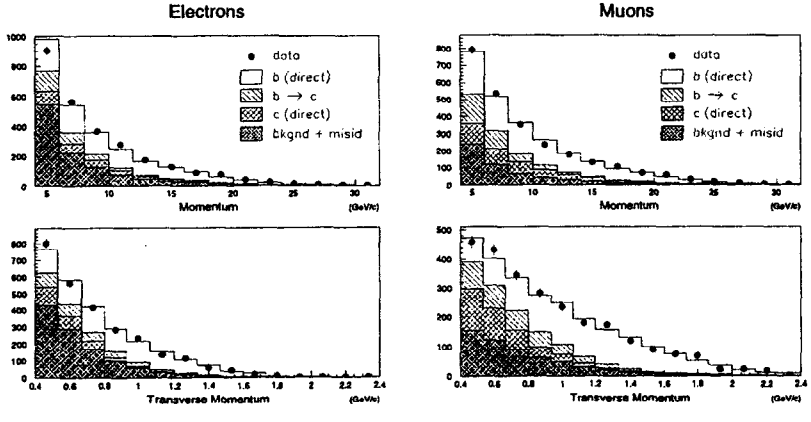


Figure 11. Distributions of momentum and transverse momentum with respect to the nearest jet axis for identified electrons (left) and muons (right) in the data (points) compared to the Monte Carlo prediction (histograms) for various sources.

A_b and A_c are simultaneously extracted from the sample of leptons using a maximum likelihood fit of the identified leptons to the theoretical cross section, employing the distributions of the lepton momentum and transverse momentum. The Monte Carlo is used to estimate the composition of the lepton sample, determining the contributions to the observed asymmetry from all lepton sources and backgrounds. Fig. 11 shows a comparison of the lepton momentum and transverse momentum distributions between data and Monte Carlo separately for electrons and muons.

The preliminary result combining all our 1993 through 1995 muon and electron data is given below.

$$A_b = 0.87 \pm 0.07 \text{ (stat)} \pm 0.08 \text{ (syst)}$$
$$A_c = 0.44 \pm 0.11 \text{ (stat)} \pm 0.13 \text{ (syst)}.$$

A summary of the systematic errors and their sources are given in Table 3.

Table 3

Systematic source	δA_b	δA_c	Systematic source	δA_b	δA_c
Lepton mis-id rate	0.020	0.026	b-fragmentation	0.004	0.016
Background asymmetry	0.010	0.026	c-fragmentation	0.010	0.026
Jet axis simulation	0.043	0.030	$Br(b \rightarrow \bar{c} \rightarrow l)$	0.003	0.030
MC weights	0.032	0.032	$Br(b \rightarrow \tau \rightarrow l)$	0.002	0.015
Tracking efficiency	0.012	0.009	$Br(c \rightarrow l)$	0.003	0.023
$R_b = 0.218 \pm 0.002$	-0.006	0.006	$b \rightarrow l$ model	0.008	0.008
$R_c = 0.171 \pm 0.014$	0.006	-0.037	$c \rightarrow l$ model	0.037	0.042
$\bar{\chi} = 0.120 \pm 0.010$	0.017	0.000	Beam polarization	0.011	0.006
$Br(b \rightarrow l) = 10.80 \pm 0.78\%$	-0.016	0.030	QCD correction	0.008	0.040
$Br(b \rightarrow c \rightarrow l) = 9.3 \pm 1.6\%$	0.011	-0.075	Total Systematic	0.078	0.132

Note that the systematic errors have begun to dominate the overall errors. In particular, systematic errors in semileptonic branching ratios produce a large contribution. Better knowledge in this area will lead to improvements in the final result.

Self-Calibrated A_b Measurement Using a Lifetime Tag and Momentum-Weighted Track Charge

The analysis utilizes an impact parameter tag to select an enriched sample of $Z \rightarrow b\bar{b}$ events.[9] The direction of the primary b quark is determined by use of the net momentum-weighted track charge, a method first suggested by Feynman and Field, to assign the charge of the b quark.[10] An improved calibration technique reduces the model dependence involved in determining the analysis power of the momentum-weighted track charge method.[11, 12]

B events from the hadronic decay sample are tagged using tracking information from both the CDC and theVXD. The tracks were projected onto the plane perpendicular to the beam axis and the impact parameter measured relative to the IP. The distribution of the normalized two-dimensional impact parameter is shown in Fig. 12.

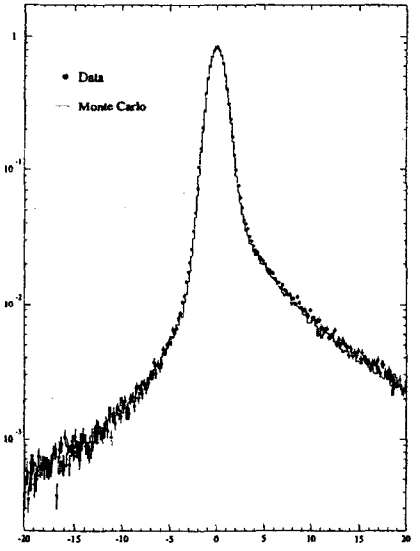


Figure 12. Distribution of the normalized impact parameter, d/σ_d.

B events are required to have ≥ 3 tracks at 3σ in the two-dimensional impact parameter. This requirement is 61% efficient for tagging b events, giving a sample that is 89% pure. The event composition versus the number of tagging tracks at 3σ is summarized in Fig. 13.

Next, the event is divided into two hemispheres along the thrust axis. The hemisphere momentum-weighted track charge difference is formed:

$$Q_{diff} = -\sum_{tracks} q_i \cdot sgn(\vec{p}_i \cdot \hat{T}) \left| (\vec{p}_i \cdot \hat{T}) \right|^K .$$

Here, $\hat{T}$ is the thrust axis while q_i and p_i are the particle's charge and momentum, respectively. When Q_{diff} is > 0, $\hat{T}$ is taken to be the b quark direction. K is chosen so as to maximize the expression's sensitivity to the b quark direction and is set to the value of 0.5 for this analysis. Fig. 14 shows a comparison of the $|Q_{diff}|$ distribution between data and Monte Carlo.

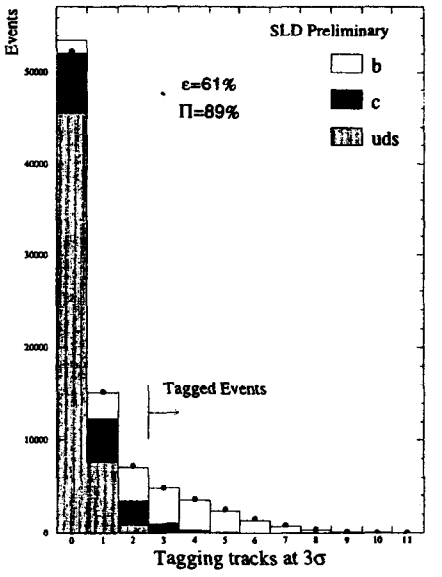


Figure 13. Event composition versus the number of tagging tracks at 3σ.

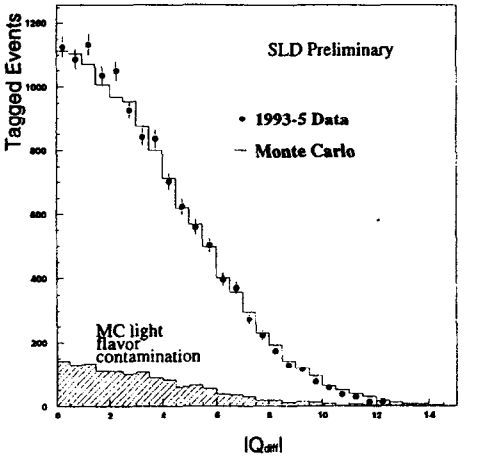


Figure 14. Comparison of the momentum-weighted charge $|Q_{diff}|$ between data and Monte Carlo.

In order to extract A_b, a maximum likelihood fit is made to the b sample using the following probability distribution for each event *i*:

$$P_i = 1+\cos^2\theta_i + 2\cos\theta_i \frac{A_e - P_e^i}{1 - A_e P_e^i} \begin{pmatrix} f_b^i A_b (2P^i_{correct,b} - 1)(1 - \Delta^b_{QCD}(\cos\theta_i)) + \\ f_c^i A_c (2P^i_{correct,c} - 1)(1 - \Delta^c_{QCD}(\cos\theta_i)) + \\ (1 - f_b^i - f_c^i) A_{bkgnd} (2P^i_{correct,bkgnd} - 1) \end{pmatrix}.$$

The terms in the expression for the probability distribution include corrections, Δ_{QCD}, for QCD effects that cause the direction of the final jet to differ from the initial direction of the quark, and the estimated asymmetry, A_{bkgnd}, from $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ decays of the Z. A_e is the asymmetry in the electron coupling to the Z, P_e^i is the polarization of the electron beam when the particular event was recorded, and $f^i_{b(c)}$ is the probability that the particular event was a $Z \rightarrow b\bar{b}$ ($c\bar{c}$) decay. $f^i_{b(c)}$ is parameterized as a function of the number of tracks missing the origin by 3σ. $P^i_{correct,b}$ and $P^i_{correct,c}$ are the probabilities that the weighted momentum method has made the correct sign determination. Fig. 15 demonstrates that the sign determination is effective, showing the angular distribution of the signed thrust axis separately for left- and right–handed electron polarization.

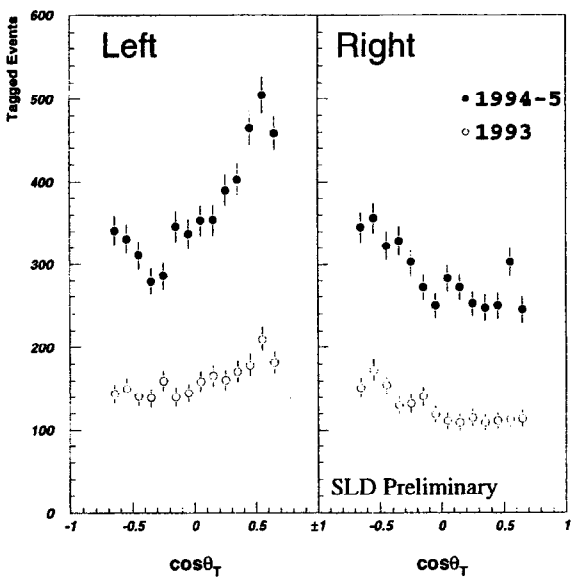


Figure 15. Distribution of the signed thrust axis in the 1993—1995 data sample.

$P^i_{correct,b}$ and $P^i_{correct,c}$ are evaluated as a function of $|Q_{diff}|$. Both can be estimated from Monte Carlo, but $P^i_{correct,b}$ can also be obtained from the data by using the information contained in the distribution of the difference of the hemisphere charges, Q_{diff}, defined above and in the distribution of the sum of the hemisphere charges, Q_s, defined below:

$$Q_s = \sum_{tracks} q_i \left|(\vec{p}_i \cdot \hat{T})\right|^K.$$

The relevant quantities are found in the widths of the distributions of Q_s and Q_{diff}, σ_s and σ_{diff}. Q_s and Q_{diff} are observables that are equivalent to the sum and difference of the momentum-weighted charges in the b-quark hemisphere, Q_b, and in the $\bar{b}$-quark hemisphere, $Q_{\bar{b}}$:

$$Q_s = Q_b + Q_{\bar{b}}$$
$$Q_{diff} = Q_b - Q_{\bar{b}}.$$

$P^i_{correct,b}$ is the fraction of time that $Q_{diff} < 0$ as a function of $|Q_{diff}|$. With the assumption that the two hemisphere distributions are Gaussian and uncorrelated, the probability that the determination of the b-quark direction is correct is given by

$$P_{correct}(|Q_{diff}|) = \frac{1}{(1+e^{-\alpha|Q_{diff}|})},$$

where α is a function of σ_s and σ_{diff}:

$$\alpha = \frac{2\sqrt{2\left(\frac{\sigma_{diff}}{\sigma_s} - 1\right)}}{\sigma_s}.$$

This expression is modified by hemisphere charge correlations due to overall charge conservation in the hadronization process and tracks which migrate from one hemisphere to the other. The result is a distortion of the joint probability of the two hemispheres from a circular Gaussian to a Gaussian ellipsoid, as illustrated in Fig. 16.

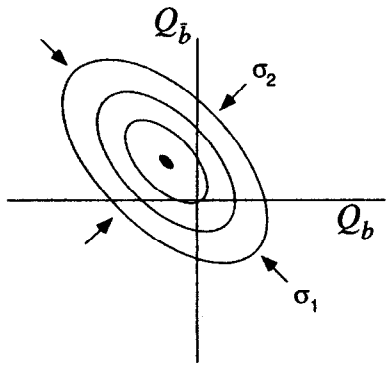


Figure 16. Effect of interhemisphere correlations on the momentum-weighted charge distributions.

The effect can be parameterized and a new expression for α obtained which accounts for the correlation:

$$\alpha = \frac{2\sqrt{2\left(\frac{\sigma_{diff}}{(1+\lambda)\sigma_s} - 1\right)}}{(1+\lambda)\sigma_s}.$$

A comparison between the values for α, σ_s, and σ_{diff} for data and Monte Carlo are shown below. Note that λ is obtained from the Monte Carlo only.

Table 4

	Data	Monte Carlo
σ_s	3.669 ± 0.023	3.791 ± 0.010
σ_{diff}	4.205 ± 0.027	4.345 ± 0.011
λ	-	0.029
α	0.253 ± 0.013	0.245 ± 0.005

The value for A_b obtained from the maximum likelihood fit is given below:

$$A_b = 0.843 \pm 0.046 \text{ (stat)} \pm 0.051 \text{ (sys)}.$$

The value is consistent with that expected for the SM. As can be seen from the final result, the measurement is limited by systematic errors. The systematics are dominated by the limitations of the self-calibration technique, first in the statistics of the data used to calculate α, and second by our ability to calculate the effect of hemisphere charge correlations. A conservative estimate of the systematics due to our knowledge of the effects of charge correlations has been arrived at by varying the fragmentation models employed in the Monte Carlo analysis; these include the JETSET string fragmentation model, the HERWIG 5.7 generator, and the independent fragmentation model. The range of variation among the models has been used for the range of error.

The systematic errors are summarized below. The uncertainty due to the α_b calibration statistics will improve with larger data samples. Increased statistics and improved b selection criteria will also reduce the uncertainty due to the b-tag flavor composition. The issue of the hemisphere charge correlation will require further study before improvements can be made in the final result.

Table 5

Systematic Source	$\delta A_b/A_b$
α_b calibration statistics	3.4 %
P(Q) shape	1.0 %
cosθ dependence of α_b	1.5 %
hemisphere charge correlation	3.7 %
light flavor subtraction	0.2 %
$c\bar{c}$ analyzing power	0.2 %
b-tag flavor composition	2.6 %
$A_c = 0.67 \pm 0.07$	1.0 %
$A_{bkgd} = 0 \pm 0.50$	0.6 %
beam polarization	0.8 %
QCD correction	0.9 %
Total Systematic	6.2%

A New Measurement of A_b with Tagged $K^{\pm}$

The analysis employs the fact that in the decay sequence $\bar{B} \rightarrow D \rightarrow K^-$, the identity of the b quark is given by the charge of the final state kaon.[13] Charged kaons are identified using the gas radiator of the SLD CRID. The analysis proceeds by selecting b events from the hadronic Z-decay sample using the two-dimensional impact parameter tag of the previous section. Events are then divided into hemispheres; in each hemisphere, tracks in the momentum range $3 < p < 20$ GeV/c with an impact parameter in the $r\phi$ plane $> 1.5\,\sigma$ are selected.

These tracks are subjected to particle identification criteria which correspond to a K: π efficiency ratio of ~ 12 : 1. An event kaon charge sum is formed for the two event hemispheres:

$$Q_\varepsilon = \sum^{Hemi-1} Q_K - \sum^{Hemi-2} Q_K .$$

If Q_ε is less than zero, then the direction of hemisphere 1 is used for the direction of the b quark. Monte Carlo studies show that of the b events tagged by the impact parameter b tag, 30% will have a value of Q_ε that is nonzero, and hence have the

direction of the b quark determined. Furthermore, of these events, 71% have the b quark direction assigned correctly.

The operation of the CRID detector is covered in the talk by R. Plano in these proceedings (see also Ref. [14]). As an example of the quality of the CRID particle ID information, Fig. 17 shows the Cherenkov radius vs. particle momentum for a small sample of tracks in the detector. The tagging efficiency of the CRID detector is calibrated by the use of τ decays to find the probability that a pion could be misidentified as a kaon. One- and three-prong τ decays provide a sample of pions (electrons and muons are not distinguished from pions) with a small, but well-known, K admixture. This study shows that 75% of the final candidates are kaons.

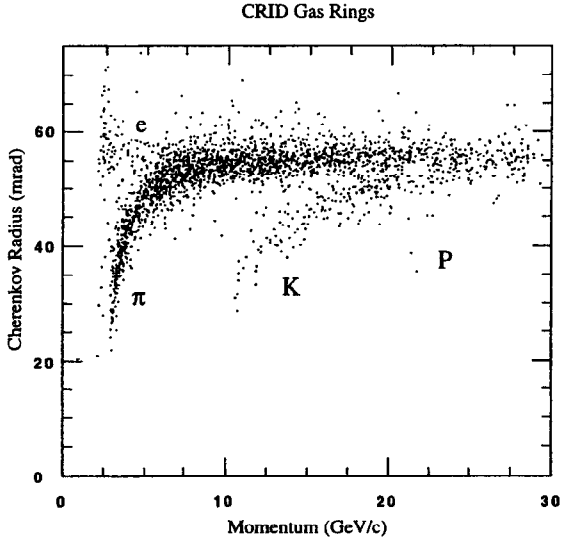


Figure 17. The measured Cherenkov radius versus particle momentum.

Figure 18 shows a comparison of the number of kaons per event for data and Monte Carlo. As a consistency check, one can compute the fraction of events with both hemispheres signed by kaons that have opposite sign. This is a good test of how well we understand the b quark correct signing probability. For data, this fraction is 62.4 ± 2.9%, which can be compared to the value for Monte Carlo, 61.9 ± 1.5%. With a larger data sample, the number of single and double hemisphere tags can be used to calibrate the b quark direction tagging efficiency in a fashion similar to the analysis that extracted the b quark identification efficiency in the measurement of R_b.

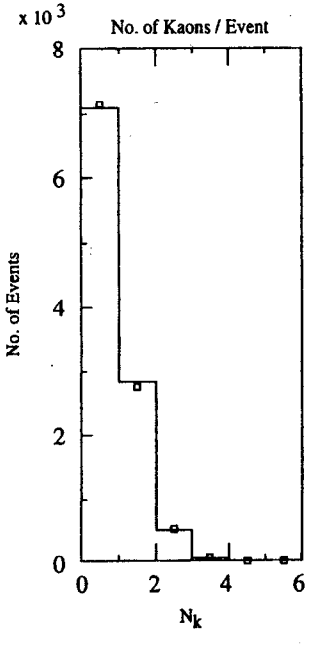


Figure 18. The number of kaons per event for data (points) and Monte Carlo (histogram).

The background subtracted asymmetry $\tilde{A}_{data}^{corr}$ is formed as a function of cosθ. Monte Carlo b events are processed through the same analysis to form $\tilde{A}_{MC}^{b}$. The value for the A_b measurement is obtained by scaling $\tilde{A}_{MC}^{b}$ to fit the data; the result is shown in Fig. 19. The fit effectively includes QCD corrections as in the JETSET MC. The preliminary K-tag result for A_b is

$$A_b = 0.91 \pm 0.09 \text{ (stat)} \pm 0.09 \text{ (sys)}.$$

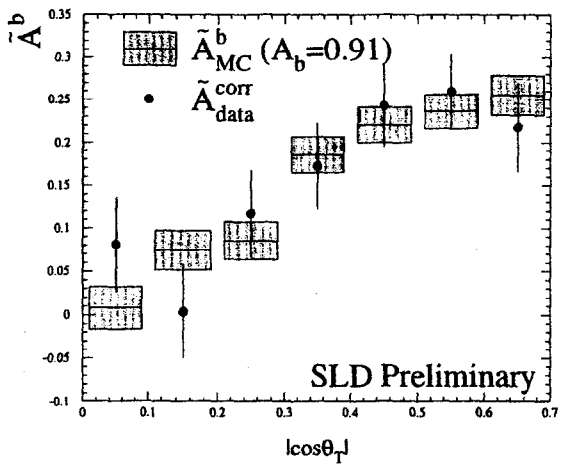


Figure 19. Distribution of the signed left-right forward-backward asymmetry.

A preliminary estimate of the systematics is given in Table 6. It should be noted that the majority of the detector and physics systematics associated with the uncertainty of the b event analyzing power can be understood with a calibration from the double-hemisphere charge comparison when statistics are sufficient for the task.

Table 6

Systematic Source	$\delta A_b/A_b$
kaon mis-id.	0.053
tracking efficiency	0.019
MC statistics	0.030
B production/mixing	0.040
B → D model	0.011
B vertex K yield	0.041
charm decay K yield	0.030
uds K production	0.011
b, c fragmentation	0.007
tag composition	0.002
$A_c = 0.666 \pm 0.070$	-0.014
beam polarization	0.007
QCD correction	0.009
Total Systematic	0.094

$Z b\bar{b}$ Coupling Parity-Violation versus $\delta \sin^2\theta_W^{eff}$

The preliminary SLD measurements presented here have been combined with a simultaneous fit to A_b and A_c, taking into account the systematic correlations between measurements. The combined SLD results are $A_b = 0.858 \pm 0.054$ and $A_c = 0.577 \pm 0.097$, with a 12.3% A_b, A_c correlation. The SM predictions are $A_b = 0.935$ and $A_c = 0.666$. These results can be compared with the average LEP measurements: $A_b = 0.884 \pm 0.032$ and $A_c = 0.642 \pm 0.053$, which are derived from the LEP $A_{FB}^{0,b}$ and $A_{FB}^{0,c}$ results[15] using $A_e = 0.1506 \pm 0.0028$ from a combination of the SLD A_{LR} and LEP A_l.

Fig. 20 shows the complementary nature of a direct measurement of A_b to the electroweak measurements of $\sin^2\theta_W^{eff}$ and A_{FB}^b. The plot is made according to the scheme proposed by Takeuchi et al.[16], which is discussed more fully in the review by M. Swartz in these proceedings. The deviations from the SM can be represented as a cross-section-like variable, ξ_b, and a parity-violation-like variable, ζ_b, in addition to $\delta\sin^2\theta_W^{eff}$. The allowed ζ_b versus $\delta\sin^2\theta_W^{eff}$ bands for a number of current experimental results are shown in Fig. 20. The SM point at (0,0) is defined by m_t = 180 GeV/c^2, m_H = 300 GeV/c^2, $\alpha_s = 0.117$, and α_{em} = 1/128.96. The thin horizontal band around (0,0) corresponds to the SM m_t, m_H variations as indicated on the plot. The 68% and 90% CL contours for the best fit to all measurements are also shown.

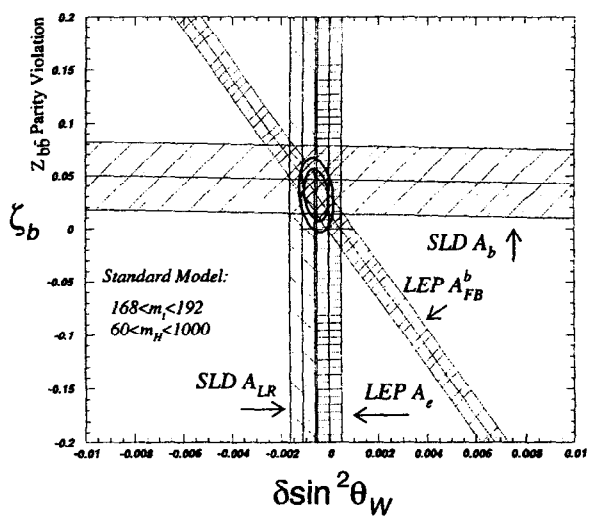


Figure 20. $Z b\bar{b}$ coupling parity violation versus $\delta\sin^2\theta_W^{eff}$.

Measurements of the B^+ and B^0 Meson Lifetimes

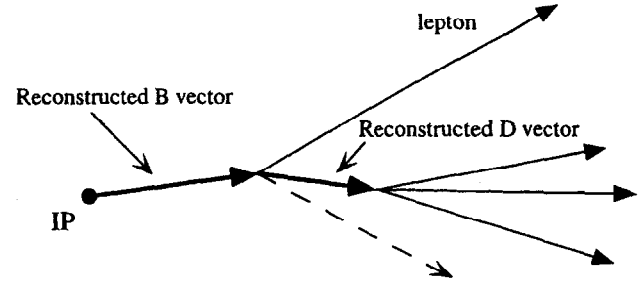
In the naive version of the spectator model of heavy quark decays, the decay of the heavy quark proceeds independently of the accompanying light quarks. This means that different hadrons containing the same heavy quark should have the same lifetime. In the case of charm decays, this model did not work very well; experimental data gave a hierarchy of lifetimes instead, providing evidence for nonspectator processes. However, since these corrections scale as $1/m_Q^2$, the decays of b quarks should be less influenced by such effects and the lifetimes are expected to differ by less than 10%.[17] Consequently, B hadron lifetimes are expected to be useful for developing a better understanding of the fundamental heavy hadron decay mechanism. Moreover, the decay widths are important in that they normalize the rate determinations from which the CKM element V_{cb} is extracted. There are also more pedestrian reasons for interest in the B lifetimes; they are fundamental in determining the effect of cuts on decay lengths, and therefore their values feed into measurements of other heavy quark parameters.

SLD employs two complementary methods to extract these quantities.[18] Method 1 takes a semi-exclusive approach, restricting the sample to B hadron semileptonic decays. The advantage of this is that the vertices one works with are relatively clean. Method 2 takes a more inclusive, topological approach which searches for vertices in three-dimensional coordinate space. This method has the advantage of almost an order of magnitude increase in the number of decays available for analysis.

Method 1: Semileptonic Analysis

A sketch of the topology of the decays we are looking for is shown in Fig. 21. The task is to bring together the separate pieces of the decay sequence. The topology of the B semileptonic decay modes provides a lookup table correlating the allowed net charge and number of prongs at both the B vertex and the D vertex with the charge of the original B particle.

High momentum electrons and muons are identified which have a high transverse momentum relative to the nearest jet axis. A search is made for a two-, three-, or four-prong D candidate vertex. A B vertex is formed by intersecting the D vertex momentum vector with the lepton. The D candidate decay length needs to be relatively large, > 4 σ, relative to the IP, and the mass is required to be less than 1.9 GeV/c^2. An attempt is made to find a primary track which can be attached to the B vertex (this is done in order to attach slow transition pions from D* decays.) Cuts are made on the resultant B vertex, requiring a mass > 1.4 GeV/c^2 (including the D), and a decay length > 800 microns. Finally, a set of B → D linking cuts are made which depend on the topology of the candidate found. If a 1-prong B decay has been found, the distance of closest approach between the B vertex and the D momentum vector is required to be less than (130,100,70) microns for (two-,three-,four)-prong D vertices. For a two-prong B decay, the three-dimensional impact parameter between the D vector and the B vertex must be less than 200 microns.



	B vertex	D vertex
B^0	1 prong 2 prongs	3 prongs 2 or 4 prongs
B^+	1 prong	2 or 4 prongs

Figure 20. The decay topology in the semileptonic analysis.

The charge constraints at each vertex can be relaxed to provide a test of how well the kinematic selection is proceeding. Fig. 22 shows the number of 2-prong B vertices and two-, three-, and four-prong D vertices as a function of the net charge. The agreement between data and Monte Carlo is quite good. For example, the charge distribution for two-prong B vertices shows that the track associated with the lepton most often has the opposite charge, as expected. The clear excess at zero net charge, even after the size of the charge ± 2 wings is taken into account, is evidence that the selection criteria are functioning well.

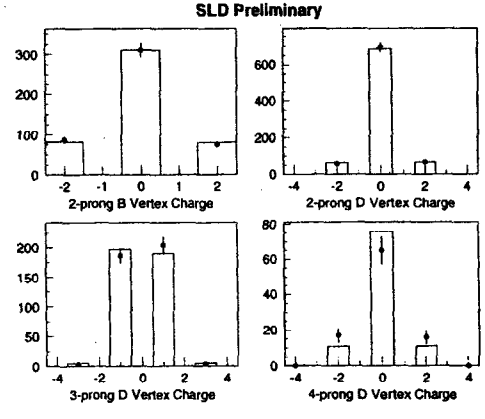


Figure 22. Comparison of the charge assignment between the data (points) and Monte Carlo (histogram) for the semileptonic analysis.

Similarly, a test of the quality of the data sample can be made by plotting the left-right forward-backward asymmetry for the final sample of B candidates. This is shown in Fig. 23 separately for the charged and the neutral B sample. The distribution shows good agreement between Monte Carlo and data asymmetries. The dilution of the asymmetry seen for the neutral B case is a result of B^0 -$\bar{B}^0$ mixing. If the charge assignment were completely random, the charged and neutral cases would show the same asymmetry.

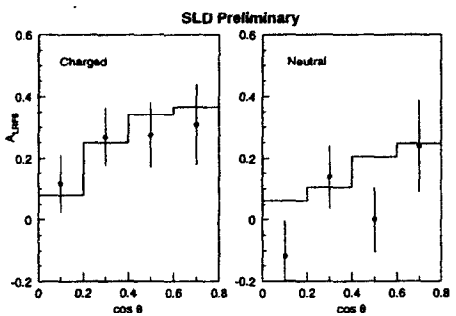


Figure 23. The left-right forward-backward asymmetry for charged and neutral decays in the semileptonic data for data (points) and Monte Carlo (histogram).

The purity and composition of the B sample for the semileptonic case is summarized in Table 7.

Table 7

	Neutral	Charged
$b\bar{b}$	98.7%	95.3%
$c\bar{c}$	1.1%	3.8%
uds	0.2%	0.9%
Bu	15.8%	70.1%
Bd	65.4%	18.4%
$B\bar{s}$	13.6%	4.1%
B baryons	3.9%	2.3%

The final sample consists of 428 neutral and 548 charged decays. The lifetime is extracted from the decay length distribution of the secondary vertices using a binned maximum likelihood technique. Simultaneous maximum likelihood fits were made to the charged,to neutral lifetime ratio R,and either the charged or the neutral lifetime. Plots of the decay lengths for neutral and charged events comparing data and the MC best fit are shown in Fig. 24. The preliminary results for the lifetime ratio and the charged and neutral lifetimes are:

$$R = \tau^{\pm}/\tau^0 = 0.94^{+0.14}_{-0.12}\ (\text{stat}) \pm 0.07\ (\text{sys})$$

$$\tau_{B^0} = 1.60^{+0.15}_{-0.14}\ (\text{stat}) \pm 0.10\ (\text{sys})\ \text{ps}$$

$$\tau_{B^{\pm}} = 1.49^{+0.11}_{-0.10}\ (\text{stat}) \pm 0.05\ (\text{sys})\ \text{ps}.$$

The systematic errors will be described below (see also Table 9).

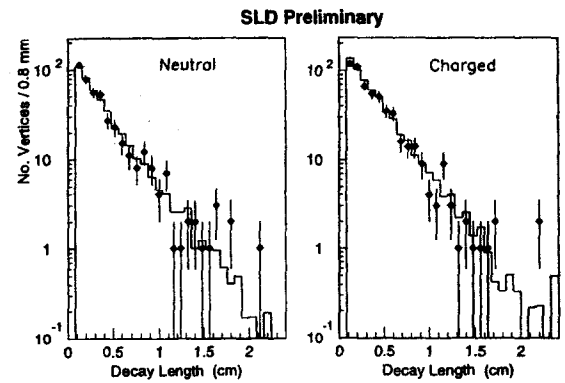


Figure 24. Decay length distributions for charged and neutral decays in the semileptonic analysis for data (points) and the best fit Monte Carlo (histogram).

Method 2: Topological Approach

Initially, a sample of $b\bar{b}$ events is selected with the two-dimensional impact parameter tag described previously. In this analysis the strategy is to search for vertices in three-dimensional coordinate space. Each track i is associated with a function $f_i(r)$ representing a Gaussian probability "tube" for the track trajectory. The goal is to find those locations in space with a maximum summed probability; a vertex function which can serve this purpose is defined for a point in space r:

$$V(r) = \sum_i f_i(r) - \frac{\sum_i f_i^2(r)}{\sum_i f_i(r)}.$$

The projections of the track functions and, separately, the vertex function onto the xy plane are shown in Fig. 25 for a sample event. A close look at the vertex function shows two peaks. One is at the location of the IP, and the second isolates a secondary-decay vertex.

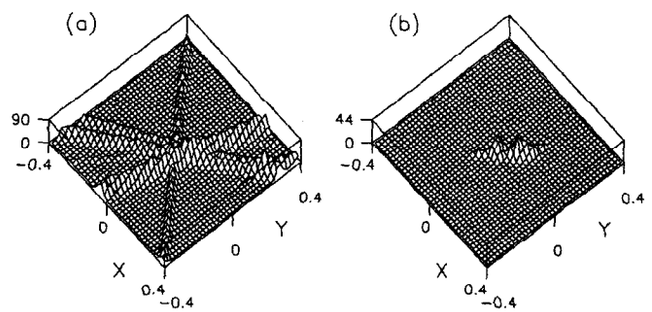


Figure 25. (a) The track and (b) the vertex functions projected onto the xy plane.

The secondary vertex is used as a seed location. Additional tracks are added to the original vertex; here, optimization of the charge reconstruction as indicated by the Monte Carlo is the primary concern. A vertex axis is formed by drawing a line between the IP and the seed vertex, and the distance between the IP and the seed vertex, D, determined. The transverse impact parameter, T, and the corresponding distance along the vertex axis, L, are calculated for all candidate tracks. Those with small T and large L/D are likely to be associated with the seed location. Good quality tracks with $T < 0.1$ cm and $L/D > 0.3$ are added to the initial set of tracks forming the secondary vertex.

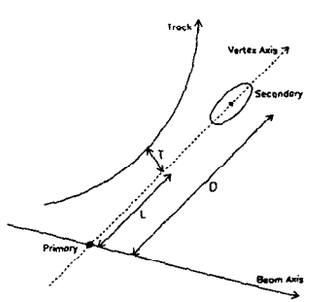


Figure 26. Impact parameter of a track with respect to the seed vertex axis.

Two cuts on the fully reconstructed secondary vertex are made. The first requires that the mass of the fully reconstructed secondary vertex be greater than 2 GeV/c^2. The quark composition of the secondary vertex is shown as a function of the vertex mass for both neutral and charged vertices in Fig. 27. After the mass cut, the neutral sample is 99.3% $b\bar{b}$ and the charged sample 99.0% $b\bar{b}$. Note that a clear K^0 peak can be seen in the neutral distribution. The second cut demands that the minimum decay length of the secondary vertex be greater than 1 millimeter. This is to minimize confusion with tracks originating from the IP.

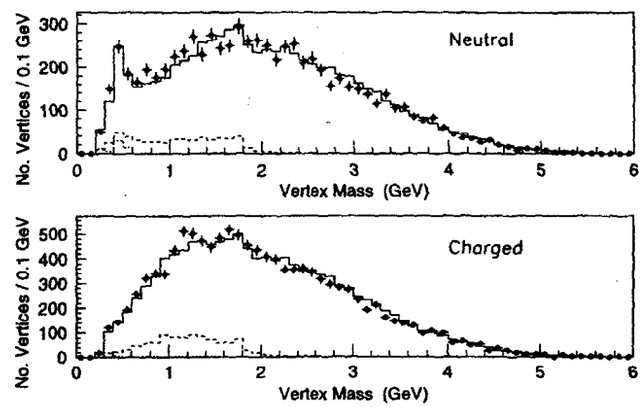


Figure 27. Mass of reconstructed secondary vertices for charged and neutral decays in the topological analysis. The dashed (dotted) histograms represent the charm (uds) contribution from Monte Carlo.

The neutral sample consists of the hemispheres with secondary vertex charge equal to zero, while the charged sample consists of those with a secondary vertex charge equal to ±1, ±2, or ±3. Fig. 28 illustrates the sample composition as a function of the charge of the reconstructed vertices. The distribution of the reconstructed charge shows good agreement between data and Monte Carlo. The information is summarized in Table 8.

Table 8

	Neutral (Q= 0)	Charged (Q = ±1, ±2, ±3)
Bu	22.2%	56.2%
Bd	55.5%	29.8%
Bs	15.3%	8.2%
B baryons	6.3%	4.8%

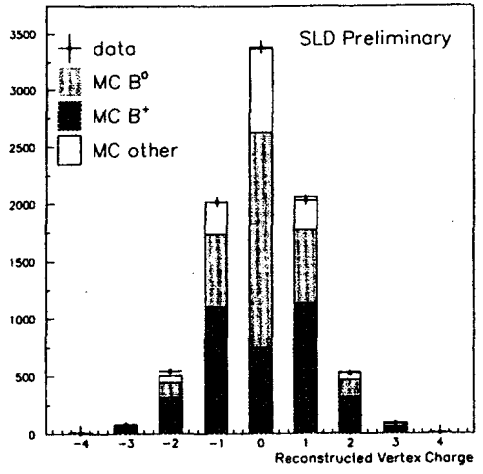


Figure 28. Distribution of charges reconstructed by the topological analysis.

A further test can be made for the charged sample. Fig. 29 shows the distribution of cosθ between the event thrust axis (used as the best estimator of the original b quark direction) and the positron direction signed by the product of the electron polarization and the reconstructed vertex charge. Neutral B decays with the wrong charge assignment would cause a dilution of the observed asymmetry and flatten the distribution as a function of cosθ. Again, there is good agreement between data and Monte Carlo.

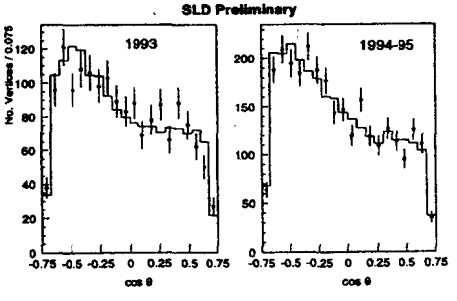


Figure 29. Distribution of the event thrust axis with respect to the positron beam, signed by the product of electron polarization and reconstructed charge for data (points) and Monte Carlo (histogram).

The final sample contains 3382 neutral and 5303 charged decays. Simultaneous maximum likelihood fits were made to the lifetime ratio R and either to the charged or the neutral lifetime. Plots of the decay lengths for neutral and charged events comparing data and the MC best fit are shown in Fig. 30. The preliminary results for the lifetime ratio and the charged and neutral lifetimes are:

$$R = \tau^{\pm}/\tau^{0} = 1.08^{+0.09}_{-0.08}\ (\text{stat}) \pm 0.10\ (\text{sys})$$

$$\tau_{B^0} = 1.55 \pm 0.07\ (\text{stat}) \pm 0.12\ (\text{sys})\ \text{ps}$$

$$\tau_{B^{\pm}} = 1.67 \pm 0.06\ (\text{stat}) \pm 0.09\ (\text{sys})\ \text{ps}.$$

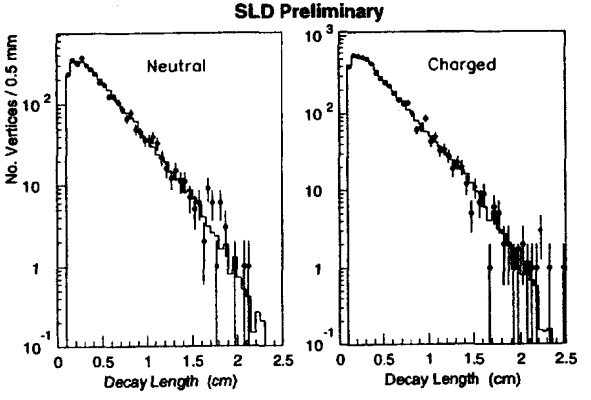


Figure 30. Decay length distributions for neutral and charged decays in the topological analysis for data (points) and the best fit Monte Carlo (histogram).

The systematic errors for both the semileptonic and the topological analyses are summarized in Table 9. The main contribution to the systematic error due to detector modeling originates from the uncertainty in the track reconstruction efficiency. Contributions to the systematic error due to physics modeling include the uncertainties in the b quark fragmentation and the B meson decay model, as well as the sensitivity to assumptions concerning B_S and B baryon production and lifetimes. The largest contribution to the systematic error arises from uncertainties in the fitting procedure and also from Monte Carlo statistics in the semileptonic analysis. The fitting uncertainties were conservatively estimated by varying the bin size used in the decay length fit distributions, and by modifying the cuts on the minimum and/or maximum decay lengths used in the fit.

Table 9

Systematic error	Semileptonic $\Delta\tau_{B^0}$ (ps)	Semileptonic $\Delta\tau_{B^+}$ (ps)	Semileptonic $\Delta(\tau^+/\tau^0)$	Topological $\Delta\tau_{B^0}$ (ps)	Topological $\Delta\tau_{B^+}$ (ps)	Topological $\Delta(\tau^+/\tau^0)$
Detector Modeling						
Charge assignment	0.004	0.016	0.014	0.040	0.040	0.040
Lepton ID	0.001	0.006	0.002	-	-	-
Physics Modeling						
b fragmentation	0.056	0.026	0.022	0.032	0.032	<0.005
B decay charm	0.009	0.006	0.007	<0.005	0.020	0.020
BR(B → D**/νX)	0.011	0.004	0.006	-	-	-
B decay multiplicity	0.011	0.015	0.016	0.010	0.030	0.030
B_s fraction	0.006	0.004	0.005	0.005	0.019	0.010
B baryon fraction	0.014	0.004	0.007	0.021	0.014	0.006
B_s lifetime	0.029	0.001	0.017	0.053	0.013	0.051
B baryon lifetime	0.005	0.002	0.003	0.016	0.004	0.012
B → D spectrum	-	-	-	0.025	0.006	0.019
Charm hadron τ	0.011	0.001	0.007	-	-	-
Monte Carlo and Fitting						
Fit systematics	0.060	0.010	0.040	0.087	0.057	0.056
MC statistics	0.042	0.030	0.039	0.021	0.018	0.027
TOTAL	0.100	0.047	0.068	0.122	0.091	0.100

Summary

We have presented preliminary SLD results on heavy quark physics which are competitive with the current world measurements. The future holds promise for further progress in the precision and scope of the SLD analyses, with advances coming on a number of fronts. The schedule for SLD data running calls for an additional 500,000 Z decays at ~ 80% electron polarization, providing a factor of four increase in statistics over the present data sample. An upgraded CCD vertex detector has been installed prior to the January 1996 SLD run.[19] The new detector provides improved solid angle coverage ($|\cos\theta|_{max} = 0.75 \rightarrow |\cos\theta|_{max} = 0.85$), overlapping three-layer CCD coverage, and an increased lever arm, all of which result in significantly improved resolution and efficiency. Finally, new analysis techniques promise to take better advantage of the analyzing power provided by the small IP spot size, the vertex detector, and the electron polarization.[20] Taken together, these developments should open new, exciting possibilities for physics at SLD.

Acknowledgments

We thank the personnel of the SLAC Accelerator Department for their outstanding achievements in bringing micrometer-size beams into collision at the SLD with remarkable reliability.

REFERENCES

1) G. Agnew *et al.*, SLD Design Report, SLAC-0273 (1984).

2) VXD: G. Agnew *et al.*, SLAC-PUB-5906 (1992); CDC: M. Hildreth *et al.*, SLAC-PUB-6656 (1994).

3) M. E. Peskin and T. Takeuchi, *Phys. Rev.* D46, 381 (1992).
M. Swartz, "Tests of the Electroweak Standard Model," in Proceedings of the XVIth International Symposium on Lepton and Photon Interactions, Ithaca,NY, 1993, edited by P. Drell and D. Rubin, pp. 381-424.

4) SLD Collab.: K. Abe *et al.*, SLAC-PUB-7004, contribution to the International Europhysics Conference on High Energy Physics (HEP95).

5) ALEPH Collab.: D. Buskulic *et al.*, Phys. Lett. **B313**, 535 (1993).

6) A. Blondel, B. W. Lynn, F. M. Renard, and C. Verzegnassi, Nucl. Phys. **B304**, 438 (1988).

7) SLD Collab.: K. Abe *et al.*, SLAC-PUB-6681, submit. to *Phys. Rev. Lett.*

8) SLD Collab.: K. Abe *et al.*, *Phys. Rev. Lett.* **74**, 2895 (1995).

9) SLD Collab.: K. Abe *et al.*, SLAC-PUB-6569, 1995, submit. to *Phys. Rev. D.*

10) R. D. Field and R. P. Feynman, Nucl. Phys. **B136**, 1 (1978).

11) SLD Collab.: K. Abe *et al.*, *Phys. Rev. Lett.* **74**, 2890 (1995).

12) SLD Collab.: K. Abe *et al.*, SLAC-PUB-6979, contribution to the International Europhysics Conference on High Energy Physics (HEP95).

13) SLD Collab.: K. Abe *et al.*, contribution (EPS0251) to the International Europhysics Conference on High Energy Physics (HEP95).

14) K. Abe *et al.*, "Results from the SLD Barrel CRID Detector," published in IEEE *Trans. Nucl. Sci.* 41: 862 (1994).

15) A. Olchevski, "Precision Tests of the Standard Model," in *Proceedings of the International EuroPhysics Conference on High Energy Physics* (Brussels, 1995).

16) T. Takeuchi, A. Grant, and J. Rosner, Fermilab-Conf-94/279-T.

17) I. I. Bigi *et al.*, in *B Decays*, ed. by S. Stone, World Scientific (1994).

18) SLD Collab.: K. Abe *et al.*, SLAC-PUB-6972, "Preliminary Measurements of B^0 and B^+ Lifetimes at SLD," contributed paper to LP'95: The International Symposium on Lepton-Photon Interactions (Beijing, China, August 10-15, 1995).

19) J. Brau, SLAC-PUB-7070, "VXD3: the SLD Vertex Detector Upgrade Based on a 307 MPixel CCD System," presented at the Nuclear Science Symposium (NNS) (San Francisco, CA, October 21-28, 1995).

20) For example, see E. Etzion, SLAC-PUB-7170, "Measurement of R_b at SLD," paper presented at the Moriond Electroweak Conference, 1996.

Heavy Quark Physics from LEP

P. J. Dornan
Imperial College of Science Technology and Medicine
London SW7 2BZ, UK

Representing the LEP Collaborations

ABSTRACT

A review of some of the latest results on heavy flavor physics from the LEP Collaborations is presented. The emphasis is on B physics, particularly new results and those where discrepancies with theory are emerging. A brief description is given of the many techniques which have been developed to permit these analyses.

CONTENTS

1. Introduction
2. Tools and Techniques
3. Electroweak B Physics
4. b Quark Fragmentation Function
5. Excited B States
6. Λ_b Polarization
7. Decays
8. Exclusive Reconstruction and Mass of the Λ_b
9. Lifetimes
10. Time-Dependent Mixing
11. Summary and Outlook

©1995 by P. J. Dornan.

1 Introduction

Heavy flavor physics has become one of the great successes of the LEP experiments due to the high efficiency with which the Z decays to a pair of b quarks can be tagged. Consequently, in this review I shall concentrate on results in the b sector even though there are now a number of interesting charm results emerging.

By the end of 1994, each of the LEP experiments ALEPH, DELPHI, L3, and OPAL had recorded between three and four million hadronic Z decays, and hence, about 700 K $Z \to b\bar{b}$ per experiment. However, not all have been recorded with the full detector, and in particular, L3 has only recently introduced a silicon microvertex detector. These detectors take time to run, and so far no results have been presented with this detector, and hence, there will be few L3 results in this talk.

Although the LEP experiments cannot hope to compete on the sheer numbers of B events which are available at the $\Upsilon(4S)$ to CLEO, the higher energy does yield a number of advantages. At LEP, the b's result from Z decay, and this gives a unique opportunity to test the electroweak standard model in the quark sector where the sensitivity to the electroweak parameters can be much greater than for the leptonic Z decays. Of course, longitudinal polarization of the beams gives even greater opportunity for stringent electroweak tests in the b sector, but this looks as though it will continue to remain a dream at LEP.

At the high energy, the B_s, Λ_b, and excited B states are produced as well as the $B^{\pm}$ and B^0, and all are produced with sufficient energy to travel long distances compared to the precision which can be achieved with the silicon vertex detectors. This enables the measurement of accurate lifetimes and direct measurements of the oscillations resulting from neutral B mixing. The high energies cause the B states to be produced in jets with additional fragmentation particles. Although this is frequently a cause of signal dilution and increased background, it does enable an understanding of heavy quark fragmentation and helps to isolate those particles which come from a common B parent.

The results presented in this review will be based on those presented at the Pisa International Heavy Flavour Symposium, the Beauty '95 Workshop at Oxford, and the submissions currently available for the Brussels and Beijing conferences. However, where preliminary results have since been finalized, the final reference is given.

2 Tools and Techniques

A substantial array of tools have been developed by the LEP experiments to undertake bottom physics. The most significant are briefly reviewed below.

2.1 Tagging via Displaced Vertices

This is the great breakthrough which has made b physics so productive. It requires a silicon microstrip vertex detector with a point resolution of ~10 μm giving an impact parameter resolution of ~40 μm. Secondary vertex resolution along the direction of flight of the B is ~300 μm, and the flight paths are 1-2 mm. The ideal detector has readout in both the $r\phi$ and z coordinates, but although all experiments now have such detectors, only ALEPH, which has had a double-sided detector since 1991, has used this for the present results. DELPHI has a three-layer vertex detector. This helps with redundancy and pattern recognition but has little effect on the vertex resolution.

The major problem associated with "lifetime" tagging comes from the charm background, particularly the proportion of charged D's as their long lifetime can give decay lengths comparable to those from the B states. Evaluation of charm contamination depends upon knowledge of both the production of the various charm states in $Z \to c\bar{c}$ decays and the topological decay rates for these states. Whilst these are known adequately for most investigations, they remain a serious problem for precision measurements such as the measurement of R_b. The other failing of the lifetime tag is the fact that it gives no information on whether it was a b or a $\bar{b}$ quark which caused the tag. For asymmetry, mixing, and branching ratio measurements, this is crucial, and therefore in these cases, the lifetime tag must be supplemented with a measure of the quark sign.

2.2 Tagging via High p_t Leptons

This was the first method to be employed before the vertex detectors were installed. Approximately 20% of b decays are semileptonic to either an electron or a muon, and due to the high mass of the B and its hard fragmentation, about half of these give rise to a high-momentum lepton ($\geq$ 3 GeV/c) with high momentum transfer ($\geq$ 1 GeV/c) with respect to the B direction.

As prompt leptons constitute only about 1% of the charged tracks, the detectors must have good electron and muon identification, and, particularly good ability to minimize hadron background in the lepton sample. These were major design criteria for the four

experiments and have been successfully achieved. Consequently, lepton tagging remains a powerful tool, particularly when associated with an identified hadron.

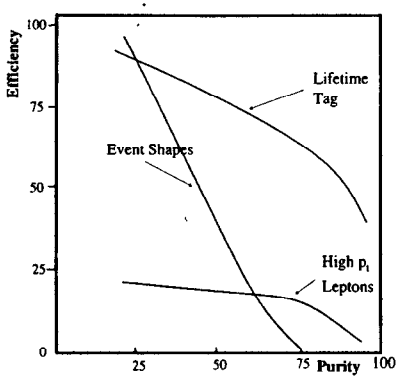


Fig. 1. Typical Efficiency vs Purity Curves for the three tagging techniques.

The major problem is the low efficiency which arises from the basic branching ratio. Additional difficulties arise from charm production as the charm semileptonic decay rate is comparable with that from beauty, although as the transverse momentum from charm is much smaller, it can be significantly reduced with the p_t cut. In principle, the lepton gives the sign of the b quark, negative lepton from b, positive from $\bar{b}$; however, this is changed if the B's mix before decaying or the observed lepton results from the semileptonic decay of the charm quark from the b decay. These dilution factors must be understood and corrected for where this is important. In many analyses, lack of complete knowledge of both the b-decay spectrum and the b-fragmentation function limits the accuracy.

2.3 Tagging with Event Shapes

This method is used more rarely and now virtually never on its own. Due to the high mass of the B hadrons, the $Z \rightarrow b\bar{b}$ events are slightly more spherical than those from the lighter quarks, and this fact can be used to achieve some discrimination. As the method uses all decay modes, a high efficiency can be achieved, but this is at the expense of low purity. Estimates of the efficiency and purity depend on the simulation of the nonperturbative fragmentation, and so it is hard to achieve an absolute measurement with known systematics. Nevertheless, the method is not without merit; perhaps the most significant being that, unlike the above two approaches, charm events are more *uds*-like than b-like. It can, hence, be of significant value in forming a charm tag when coupled with a lifetime or lepton signature.

Typical efficiency/purity plots for these three basic b-tags are shown in Fig. 1.

2.4 D^* Reconstruction

In the multiparticle B jets, combinatorial background usually makes it extremely difficult to identify the decaying hadrons in the B decay products. However, the very low Q value for the decay $D^{*+} \rightarrow D^0\pi^+$ enables a D^* sample to be extracted with low background. D^* states, of course, are also produced in charm events, but partial discrimination can be achieved on the basis of the momentum spectrum, as those from charm production are primary D^*'s and have a substantially harder spectrum than those cascading down from the B decay. However, when an identified D^{*+} is combined with a negative lepton on the same side of the event, this forms a very strong and clean signature for the decay $B \rightarrow D^*\ell\nu$ or $B \rightarrow D^*\pi\ell\nu$. Statistics are such that $D^{*\pm}\ell^{\mp}$ tags are now a very useful signature.

2.5 Hadron Identification

The Cabibbo favored decay chain $b \rightarrow c \rightarrow s$ requires that one of the final hadrons should be a kaon. Identification of this kaon significantly reduces combinatorial background when examining the decay products. All experiments use dE/dx from their main tracking chambers, but here DELPHI have a significant advantage due to their RICH counters which give good π/K separation up to 20 GeV.

2.6 E_{miss}—Neutrino Energy

Many of the investigations at LEP use semileptonic decays to identify particular states for analysis. Clean signals can then be achieved, but as the semileptonic decays involve a missing neutrino, a full kinematic reconstruction in the B rest frame is impossible. However, if the detector is hermetic and the calorimeters have sufficient directional and energy resolution, a good estimate can be made of the missing neutrino energy. As energy is not shared equally between the two halves of an event, this must frequently be estimated with an algorithm such as the following:

$$E_\nu = E_{beam} - E_{vis} + \frac{M^2_{same} - M^2_{opp}}{4E_{beam}}.$$

This is used by ALEPH and gives them a resolution on the neutrino energy of about 2.5 GeV. It makes possible a number of the following analyses based on semileptonic decays.

2.7 Estimation of the *B* Momentum

This is a procedure sometimes referred to as inclusive *B* reconstruction. As true exclusive reconstruction is impossible for all but a very few *B* decays, alternative techniques have been evolved by which tracks in a hadronic *b* jet are associated with either the decaying *B* or the primary vertex on the basis of quantities such as vertex information, rapidity along the jet axis, etc. The exact procedure varies from collaboration to collaboration and depends upon the particular study under consideration. The technique has proved to be particularly effective when searching for excited *B* states because similar errors occur for both *B* and B^*, and hence, cancel in the difference which gives the signal.

2.8 Tag Calibration Using Two Hemispheres

This is a powerful and increasingly useful method. When events are divided by a plane perpendicular to the thrust axis, usually each hemisphere will contain one of the primary quarks. With the increase in statistics, it has become possible to use this to calibrate tags by comparing the number of tags in single hemispheres with the number of events which have tags in both hemispheres. The relations are

$$N_1 = \varepsilon_b R_b + \varepsilon_c R_c + \varepsilon_{uds}(1 - R_b - R_c)$$
$$N_2 = C\varepsilon_b^2 R_b + \varepsilon_c^2 R_c + \varepsilon_{uds}^2(1 - R_b - R_c)$$

The ε's define the efficiencies of the tag for *b*, *c*, and *uds* hemispheres, and *C* represents a correlation between the hemispheres as the two *b's* will not always be on opposite sides. It is assumed that *C*, ε_c, and ε_{uds} are small.

R_b can be eliminated, and hence, the tagging procedure can be calibrated without further recourse to models or Monte Carlo except for the small parameters. Alternatively, the tagging efficiency can be eliminated to determine R_b, and this is the basis of all competitive measurements of R_b. The main problems result from the charm efficiency and the correlation term, as these cannot be eliminated and must be obtained in another way, usually from Monte Carlo.

2.9 Jet Charge

Jet charge is defined by

$$Q_{jet} = \frac{\sum_i q_i x_i^\kappa}{\sum_i x_i^\kappa},$$

with *x* a kinematical quantity such as *p*, p_{long}, or *y* relative to the jet axis, and κ is a weighting factor usually chosen between 0.5 and 1.

This provides an alternative method to determine whether a *B* jet contains a *b* or a $\bar{b}$. Although not as clean as the lepton method, it has the major advantage that it is not restricted to semileptonic decays, and it therefore complements the lifetime tagging technique. Moreover, unlike general jet charge algorithms which are heavily dependent on Monte Carlo for their efficiency estimation, the lifetime tag can be used to establish the important charge separation parameter for the *b*'s directly from the data. As in so many cases, the major problem then becomes an understanding of the background from charm events.

3 Electroweak *B* Physics

The ability to tag $Z \to b\bar{b}$ events with high efficiency and purity has made it possible to perform tests of the Standard Model in the quark sector with a precision comparable to that in the lepton sector. The *b* sector, however, allows the test of electroweak vertex corrections due to the high coupling of the *b* to top with the strength of these corrections being proportional to m_{top}^2. As a result of the high top mass, these corrections are of the order of 1% for $R_b = \Gamma_{b\bar{b}}/\Gamma_{had}$, and hence, just within the realm of experimental measurement. The other major electroweak measurement which is performed with *b* quarks at LEP is the forward-backward asymmetry in *Z* decay. This provides the single most sensitive measurement from the LEP experiments for the measurement of $\sin^2 \vartheta_W^{eff}$.

3.1 *Z* Decay Width to *b* Quarks, R_b

Measurements of R_b must aim at an accuracy of ~1%, and therefore, cannot rely upon Monte Carlo for efficiency estimations. For these measurements, the use of double hemisphere tags to eliminate the basic efficiency of the *b* tag is vital.

The best methods rely upon the lifetime tag, and these can now reach individual accuracies approaching the 1% level. Difficulties result from the charm background and the correlations between the hemispheres which may result, e.g., from hard gluon production.

Both of these must be taken from Monte Carlo. Results using these procedures have been given by the three experiments with vertex detectors; all are currently systematic limited, but there is certainly the capability to reduce the errors by understanding these problems. The LEP electroweak working group is investigating the problem and is establishing principles by which each experiment quotes its results so that a realistic attempt at combining them may be made.

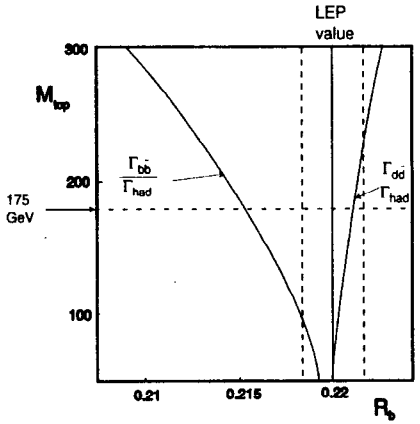


Fig. 2. The current LEP value of R_b with predictions for both R_b and R_d in the Standard Model as a function of the top mass.

The second significant technique is to use the lepton tag. This suffers from lower statistics as well as similar problems with charm background and hemisphere correlations. The analyses fit variables related to the single and dilepton p and p_t spectra. Such analyses, referred to as "global analyses," do not just give the electroweak parameters such as the b and c widths and asymmetries, but also other quantities of interest for b physics such as the $b \to \ell$ and $b \to c \to \ell$ branching ratios, the mean energy fraction taken by the b and c hadrons in the fragmentation and the integrated mixing parameter, χ. The accuracy for R_b is, however, roughly a factor of two worse than that from the lifetime methods.

Methods which are becoming increasingly popular use more than one tag. These can certainly improve the tagging rate, and hence, the statistical accuracy; however, the trade-off with systematics is not always simple to establish.

In the review by Karlsson[1] at the recent Pisa meeting, the average LEP value for R_b was given as 0.2196 ± 0.0019 with R_c fixed at the Standard Model value of 0.171. This gives a discrepancy at the 2—3 σ level with the expected value of 0.2155 for a top mass of 175 GeV as shown in Fig. 2. However, one must be careful in quoting confidence limits as measurements are now systematically limited,which makes the combination of results from different techniques, and even more, from different collaborations, hazardous as such errors are highly unlikely to be Gaussian.

Precision measurements at this level are difficult but the LEP community has been making significant efforts to understand and overcome the problems both with the individual measurements and how to combine them. Nevertheless, R_b probes unexplored areas of the Standard Model, so its measurement is currently both exciting and challenging, and conceivably, it is giving the first indication of a deviation from the Standard Model.

3.2 The b Forward-Backward Asymmetry

The forward-backward asymmetry is given in terms of the vector and axial couplings of the electron and the produced fermion by

$$A_{FB}^{f\bar{f}} = \frac{3}{4} \frac{4 v_e a_e v_f a_f}{\left(v_e^2 + a_e^2\right)\left(v_f^2 + a_f^2\right)}.$$

The most sensitive asymmetry measurement to determine $\sin^2 \vartheta_W^{eff}$ in unpolarized Z decay results from $b\bar{b}$ production. This can be seen from

$$\frac{\partial A_{FB}^{f\bar{f}}}{\partial \sin^2 \theta_W} = 4\left(-2I_3^f - Q_f + 8Q_f \sin^2 \theta_W\right)$$

$$\sim 3 \quad \text{for} \quad Z \to b\bar{b}$$

$$\sim 0.6 \quad \text{for} \quad Z \to \mu^+\mu^-.$$

The lepton tag still provides the basis for most measurements of this asymmetry; however, the high efficiency of the lifetime tag can now be employed when coupled to jet charge measurements to determine the direction of the b quark. The accuracy of the two methods is comparable, and as there is little correlation between the samples, they can be combined with comparative ease. It is also relatively simple to combine results from the LEP experiments to achieve an overall value as statistical uncertainties still dominate, and this gives the single most accurate technique at LEP for measurement of $\sin^2 \vartheta_W^{eff}$.

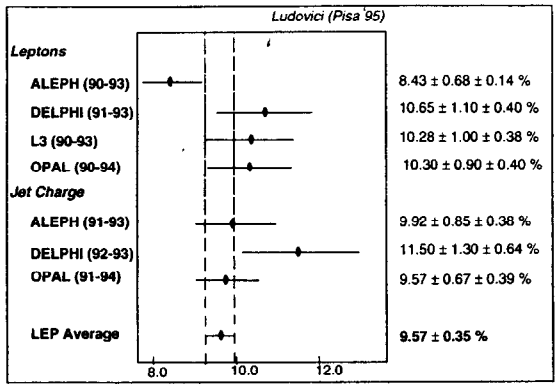


Fig. 3. Summary of the LEP $A_{fb}(b\bar{b})$ measurements from Ludovici's review at Pisa 1995.

The asymmetry values, as summarized by Ludovici[2] at Pisa, are given in Fig.3. The average at the Z pole is 0.0957 ± 0.0035, which leads to a value of $\sin^2 \vartheta_W^{eff}$ of 0.23182 ± 0.00064.

4 *b* Quark Fragmentation Function

One of the parameters given by the global lepton analyses described above is the mean B energy in the fragmentation. However, the analyses assume that the fragmentation follows the model of Petersen et al.,[3] although there is no direct confirmation of this.

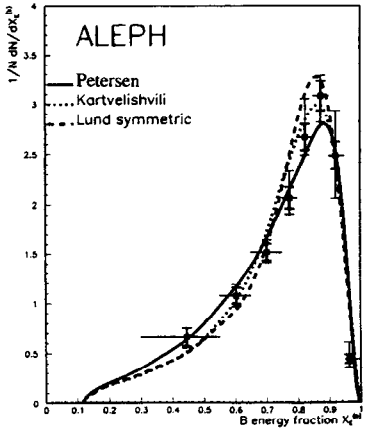


Fig. 4. The measured B energy fraction with model predictions.

Recently, ALEPH[4] have made use of the ability to reconstruct the missing neutrino energy to reconstruct the semileptonic $D^*\ell\nu$, $D^+\ell^-\nu$, and $D^0X\ell^-\nu$ decay modes to determine the $x = E_b/E_{beam}$ distribution. This is compared in Fig. 4 with predictions from the JETSET Monte Carlo for fragmentation schemes by Kartvelishvili[5] and the Lund symmetric procedure[6], usually used for the light quarks in JETSET simulations, as well as the Petersen prescription. The measurements yield a mean value for $\langle x_b \rangle$ of $0.715 \pm 0.007 \pm 0.013$. However, after adjusting the model parameters the data is still inadequate to distinguish these schemes although they verify that there is no significant discrepancy with the Petersen procedure after the ε parameter is correctly chosen.

In a similar analysis, OPAL[7] obtain a value of $\langle x_b \rangle = 0.695 \pm 0.006 \pm 0.008$.

5 Excited B States

There has been a major attempt recently to establish what percentage of B hadrons are formed in the s-wave states, B or B^*, or in one of the p-wave states, generically termed B^{**}. The predicted B^{**} states are given in Table 1. Heavy quark effective theory predicts that two will be narrow ($\Gamma \sim 10$ MeV) whilst two will be broad ($\Gamma \sim 100$ MeV).

State	J^P	Width	Decays to
B_2^*	2^+	narrow	$B\pi$, $B^*\pi \to B\pi\gamma$
B_1	1^+	narrow	$B^*\pi \to B\pi\gamma$
B_1^*	1^+	wide	$B^*\pi \to B\pi\gamma$
B_0^*	0^+	wide	$B\pi$

Table 1. B^{**} states predicted by heavy quark effective theory.

The standard inclusive analyses use various types of inclusive B reconstruction, with the inclusive *"B"* paired with either a low energy γ (for the B^*) or a low energy π or K for the B^{**} or B_s^{**}. All states should have low Q values, given by $Q = M(\text{"}B\text{"}X) - M(\text{"}B\text{"}) - M(X)$, with X either a γ, π, or K. This is plotted, and as errors due to inadequate B reconstruction mainly cancel in the difference, evidence for these excited states is revealed by a peak in the Q spectrum. In the B^{**} analysis, the resolution of approximately 45 GeV is such that the loss of the soft γ in the decays $B^{**} \to B^*\pi \to B\pi\gamma$ does not significantly affect the signal.

5.1 The B*

The B^* was first observed by L3 (Ref. 8) and makes full use of their BGO calorimeter which has considerably superior energy resolution, particularly for low-energy photons, than the electromagnetic calorimeters of the other LEP experiments. In the absence of an operational vertex detector, L3 tags the B using a high p_t muon, determines the B^* direction from the direction of the jet containing the muon, and sets the magnitude of the B^*

momentum to the mean value expected from the fragmentation spectrum of 37 GeV. This allows the associated photon momentum to be transferred into the "B*" rest frame, where a peak indicates the two-body decay of the B^*. This is shown in Fig. 5, the peak is at 46.3 ± 1.9 MeV, and they estimate the vector to pseudoscalar production ratio to be

$$\frac{N_{B^*}}{N_{B^*}+N_B} = 0.76 \pm 0.10,$$

which is very close to the simple 3:1 prediction.

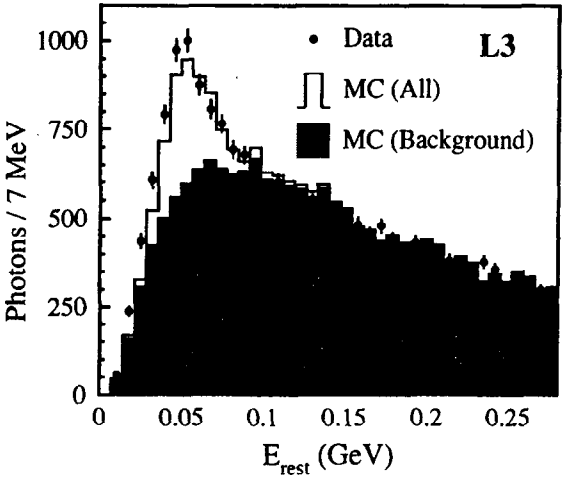


Fig. 5. B* Production in L3. The photon energy in the "B*" rest frame.

ALEPH[9] and DELPHI[10] also report similar analyses, but, as their calorimeters are inferior at these low energies, they use electron-positron conversions to detect the low energy photons. However, both use their vertex detectors to obtain a higher tagging efficiency and have more comprehensive inclusive B algorithms which permit estimates of both the B^* direction and energy. The results are quite consistent with the L3 ones.

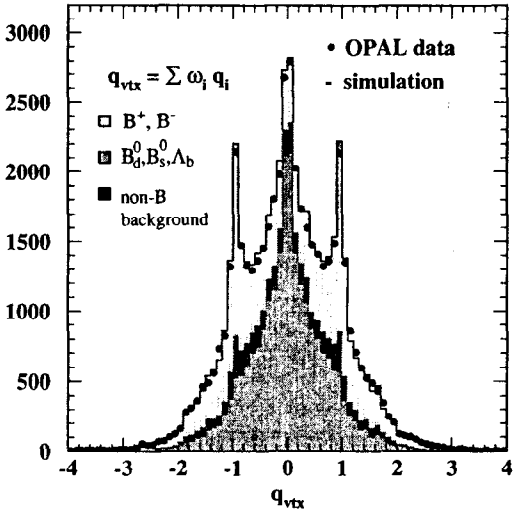


Fig. 6. The inclusive B reconstruction procedure used by OPAL using charge weighting.

5.2 The B^{**}

Inclusive evidence for the B^{**} states has come from ALEPH,[9] DELPHI,[11] and OPAL.[12] In ALEPH and DELPHI, the inclusive B reconstruction depends upon the rapidity along the jet axis coupled with vertex or impact parameter cuts. OPAL have evolved a different procedure which enables them to also have some estimate of the charge of the B state. They use the charged tracks to determine secondary vertices in jets and then weight each track in the jet with the probability, ω_i, that it came from the secondary, and then sum these weights multiplied by the charge of the relevant track. In this way, they obtain an estimate of both the charge of the secondary, $q_{vtx} = \Sigma \omega_i q_i$, and the momentum vector of the

charged decay products. To this, they add 70% of the observed neutral energy in a cone around the P direction to give the estimate of P_B. The results of the procedure for the charge separation are shown in Fig. 6; the composition of the sample following a cut $|q_{vtx}| > 0.6$ is 54% charged B's and 33% neutral B's, whilst for $|q_{vtx}| < 0.6$ it is 24% charged B's and 67% neutrals. Agreement of the procedure between data and Monte Carlo is good, as can be seen from Fig. 6.

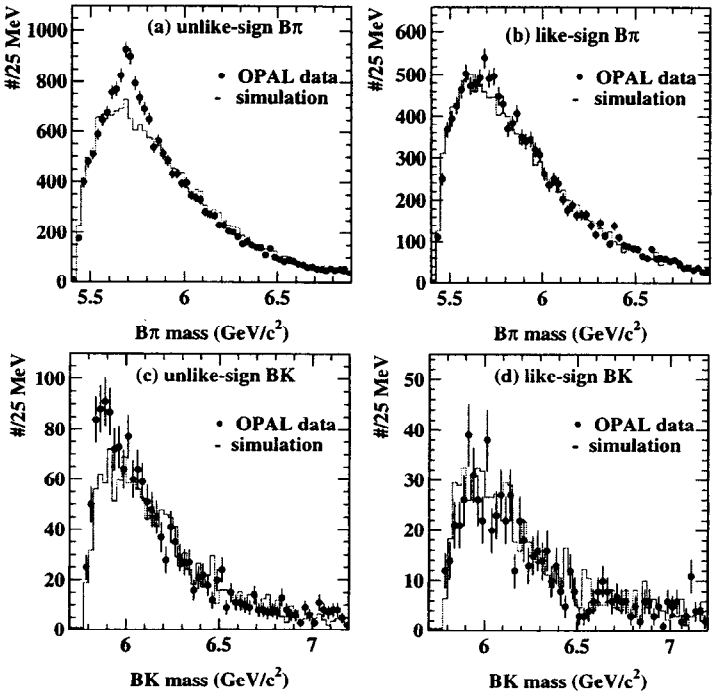


Fig. 7. OPAL results on inclusive B** production. Unlike and like sign Bπ combinations are shown in (a) and (b), unlike and like sign BK combinations in (c) an (d).

Following this procedure, OPAL pair a primary π or K with like or unlike charge to that of the B. The results are shown in Fig. 7, where good peaks are observed in the $B\pi(K)$ plots when the B and $\pi(K)$ have opposite signs. No peaks are observed when the B is paired with a π or K of like sign. They also have a peak (not shown) when a neutral B is paired with a charged π but no peak when paired with a kaon. All these observations are as expected from the production of excited B^{**}states.

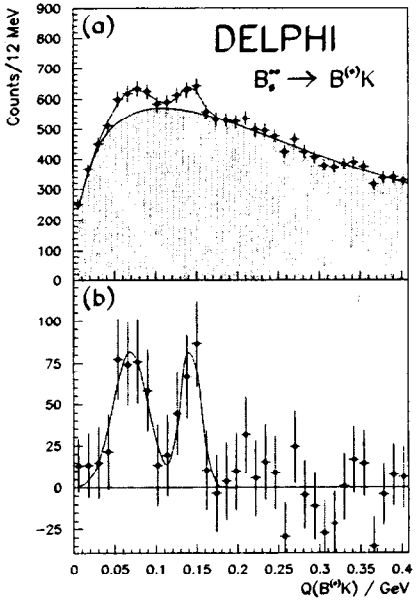


Fig. 8. DELPHI results for the Q value for BK combinations showing evidence for two narrow B_s^{**} states. Full results are given in (a), background subtracted in (b).

ALEPH and DELPHI plot the Q value for the supposed B^{**} decay but make no attempt to search for charge correlations. DELPHI uses information from both *dE/dx* and their RICH counters to distinguish π's from kaons, and this plot yields some evidence for the production of two B_s^{**} states attributed to the two narrow ones. The results are shown in Fig. 8. A two Gaussian fit to these peaks yields mean values of $70 \pm 4 \pm 8$ MeV and $142 \pm 4 \pm 8$ MeV with widths of 21 ± 6 and 13 ± 6 MeV. However, as the widths of the Gaussian fits to these two peaks are less than the resolution, this preliminary result must await further confirmation.

All experiments indicate that the proportion of b quarks fragmenting to B^{**} is approximately 30%.

5.3 Exclusive B^{**} Reconstruction

Whilst the inclusive technique demonstrates unambiguously the production of B^{**} states, the mass resolution of typically 45 MeV is such that it is unlikely to yield information on the relative production rates of the different states. ALEPH[13] have preliminary results on exclusive B^{**} production using a substantial sample (435) of exclusively reconstructed B mesons decaying to a variety of final states. They pair the reconstructed B mesons with the pion which has the maximum p_{long} along the corresponding jet axis and then compare right- and wrong-sign mass combinations to look for the signal. The mass resolution is approximately 5 MeV, much better than for the inclusive B analysis. The right-sign $B\pi$ mass plot is shown in Fig. 9; an enhancement is clearly visible with respect to the background determined from the wrong-sign pairs. A single Gaussian does not fit this well; two give a much better fit which is consistent with the two narrow states expected from heavy quark theory. The two Gaussian fit yields masses of 5585^{+74}_{-34} MeV and 5703 ± 14 MeV with widths of 28^{+18}_{-14} and 42^{+41}_{-18} MeV. However, this cannot yet be considered a confirmation; in particular, statistics do not allow any information to be extracted on the spin-parity of the decaying B^{**}.

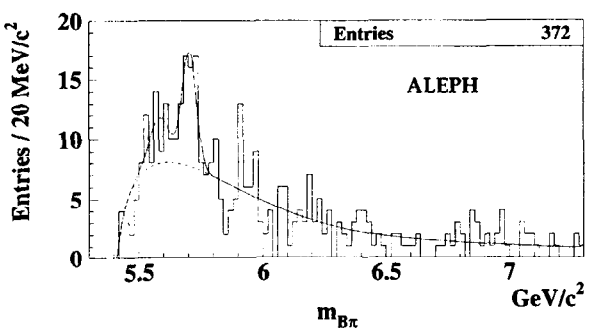


Fig. 9. Right sign Bπ combinations from exclusive B decays reconstructed by ALEPH.

The overall rate of B^{**} production from this exclusive analysis is also found to be $30 \pm 6\%$ in agreement with the inclusive analyses.

5.4 Evidence for Σ_b and Σ_b^* Baryons

DELPHI[14] have taken this analysis further and produced the first preliminary evidence for Σ_b and Σ_b^* baryons at LEP. They extend the basic B^{**} analysis by looking for inclusively reconstructed B hadrons which are enriched in baryons, by identifying protons, reconstructed lambdas, and neutral hadron showers in the B jets. Examination of the Q value distribution for these states yields a 9σ enhancement which can best be fitted by two Gaussians with means of $33 \pm 3 \pm 8$ and $89 \pm 3 \pm 8$ MeV with both widths fixed at the expected resolution. It is shown in Fig. 10(a). These two peaks are ascribed to the Σ_b and Σ_b^* respectively, and from the size of the peaks they find that $4.8 \pm 0.6 \pm 1.5\%$ of all b's produced in Z decay fragment to a Σ_b or Σ_b^*, which is approximately half of the expected b-baryon production. Repeating the analysis with an antibaryon cut shows no signal, see Fig. 10(b).

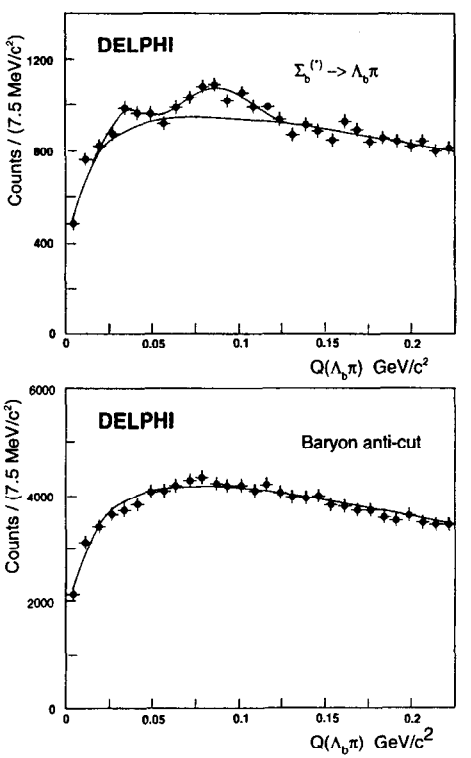


Fig. 10. DELPHI Q value distributions for (a) inclusive "Λ_b"π combinations showing evidence for the production of Σ_b and Σ_b^*, (b) with an antibaryon cut.

6 Λ_b Polarization

If, as indicated from the above DELPHI results, there is a significant amount of $\Sigma_b^{(*)}$ production, then the high polarization of the b quark resulting from the Z decay (~–94%) is likely to be lost by the time the lowest baryon, the Λ_b, is produced. Any Λ_b polarization is reflected in the energy spectra of both the charged lepton and the neutrino. ALEPH have used this to measure the Λ_b polarization using a variable, y, equal to the ratio of the mean energies of charged leptons and neutrinos. This is particularly sensitive to the polarization but demands a good estimate of the missing neutrino energy. The relation to the polarization is given by

$$y = \frac{< E_\ell >}{< E_\nu >} = \frac{7 - P_{\Lambda_b}}{6 + 2P_{\Lambda_b}} + O\left(\frac{m_{\Lambda_c}^2}{m_b^2}\right).$$

The procedure adopted is then to compare the measured value of y with that from a Monte Carlo in which the Λ_b is produced unpolarized

$$R(y) = \frac{y_{data}}{y_{MC}}.$$

The relationship of $R(y)$ to P_{Λ_b} is shown in Fig. 11 together with the ALEPH value. Their result is

$$P_{\Lambda_b} = -0.28^{+0.23+0.13}_{-0.20-0.12},$$

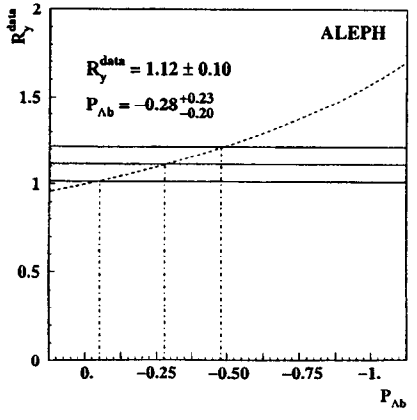


Fig. 11. The ALEPH value of $R(y)$ and its relationship to the Λ_b polarization.

which suggests that much of the b quark polarization is indeed lost in the formation of the Λ_b, consistent with substantial $\Sigma_b^{(*)}$ production.

7 Decays

7.1 Inclusive Semileptonic

The rate of semileptonic B decay is the cause of two problems. The values which have been measured at both the Z and the $\Upsilon(4S)$, typically between 10 and 12%, are lower than theoretical expectations and the value measured at the Z is systematically higher than that measured at the $\Upsilon(4S)$. In fact, the reverse would be expected due to the production of the Λ_b at LEP, which is expected to have a lower semileptonic branching ratio than the B meson in keeping with its lower lifetime.

Until recently, most LEP measurements have relied upon the overall "global" fits to the single and dilepton spectra, and values have been produced in conjunction with measurements of R_b, R_c, etc. However, this summer ALEPH[16] has produced two new analyses dedicated purely to the task of measuring the primary $b \to \ell$ branching ratio and the cascade $b \to c \to \ell$ rates.

Both of these use the "lifetime" tag in one hemisphere to select a pure sample of $Z \to b\bar{b}$ events and then examine the leptons in the opposite hemisphere. In the first approach, hemispheres opposite to the tag which contain either a single lepton or two oppositely charged leptons are selected and an overall likelihood fit made to both the numbers of single and dilepton hemispheres and to the p_t spectra. The numbers are sensitive to the absolute tagging efficiency, which is determined on the data using single- and double-tag information, whilst the fit to the spectra is affected more by modeling uncertainties than the absolute efficiency. The two aspects of the fit are therefore complementary. The fit to the final single-lepton p_t spectrum is shown in Fig. 12, and the analysis yields the preliminary result

$$Br(b \to \ell\nu X) = 11.34 \pm 0.13 \pm 0.27^{+0.41}_{-0.27}\%$$

$$Br(b \to c \to \ell\nu X) = 7.86 \pm 0.19^{+0.46+0.39}_{-0.36-0.47}\%$$

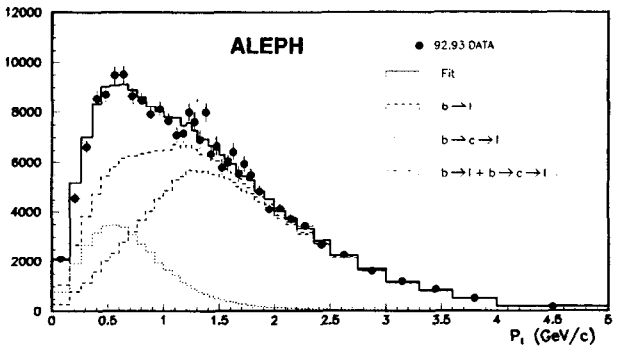


Fig. 12. The fit to the lepton p_t spectrum in B decay from the ALEPH analysis using models for the decay distribution.

with the errors respectively statistical, systematic, and modeling. It is clear that whilst this method has good statistical precision, the modeling of the lepton spectra seriously limits the accuracy.

To attempt to overcome this, the second approach builds on techniques adopted earlier at ARGUS and CLEO to minimize the model dependence. Two samples are prepared; the first uses the "lifetime" tag with a very hard cut to establish a sample of hemispheres with a very high b purity containing a single lepton. A second sample of opposite side dileptons is then prepared in which one of the leptons has a high p_t and is used as a tag lepton; this, after corrections for mixing, etc., yields the sign of the decaying b quark. To improve the statistical precision, the dilepton sample is augmented by a single-lepton sample with an opposite hemisphere jet charge identification and a soft lifetime cut. The contribution of primary and cascade b decays to both of these samples can be simply estimated with no reference to models although the necessary cuts which have to be applied to the samples for the lepton identification, etc., imply that the model dependence is not totally zero.

The overall p_t dependence is shown in Fig. 13. It is clear that the statistical precision is inferior to the first method but the preliminary value obtained

$$Br(b \to \ell\nu X) = 11.01 \pm 0.23 \pm 0.28 \pm 0.11\%$$
$$Br(b \to c \to \ell\nu X) = 8.30 \pm 0.31 \pm 0.42 \pm 0.12\%$$

shows the much reduced model dependence.

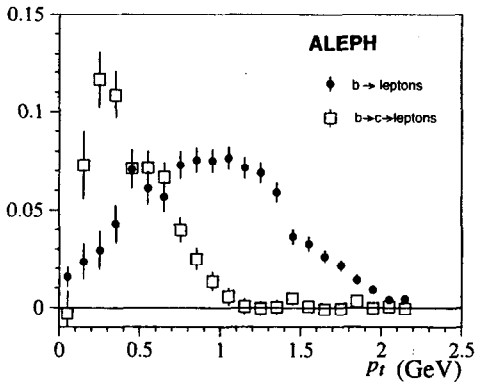


Fig. 13. The results of ALEPH for the model independent fit of the lepton p_t spectrum in B decay.

These two measurements are consistent with each other, consistent with the earlier measurements at the Z, and higher than the latest $\Upsilon(4S)$ measurements. At the Pisa conference, Schmitt,[17] in his summary, combined the LEP measurements to yield $Br(b \to \ell\nu X)$ =11.25 ± 0.24%, whereas when he takes the latest $\Upsilon(4S)$ value and predicts the expected value at LEP after correcting for Λ_b production, he obtains 10.0 ± 0.4%, a discrepancy between 2 and 3σ. Whilst not that strong statistically, the fact that this discrepancy has remained for so long suggests that there could be a systematic flaw in one of the analyses.

7.2 Measurement of $|V_{cb}|$

Effective heavy quark theory can be used to extract a value of the CKM-matrix element V_{cb} from an analysis of the decay $B^0 \to D^{*-}\ell^+\nu$. In the heavy quark limit for zero recoil of the D^*, the normal three form factors reduce to a single one, $F(\omega)$. This is a function of the q^2 to the lepton-neutrino system and normally written in terms of a variable ω defined by

$$\omega = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}$$

such that

$$\text{as } q^2 \to q^2_{\max}, \qquad \omega \to 1,$$

and in the heavy quark limit, $F(\omega) \equiv 1$.

This method has already been used by CLEO with substantially more events than are available to the LEP experiments. However, for this particular measurement, the boost given to the B state at LEP is of value because, as a result, the D^* products have a substantial momentum in the apparatus, whereas at the $\Upsilon(4S)$, the two B's are produced virtually at rest and the pion from the D^* decay in the limit of zero recoil is very soft and suffers from reconstruction difficulties. In practice, ω is determined using the decay kinematics and an extrapolation made to $\omega = 1$ using the linear form

$$F(\omega) = F(1)\left[1 + a^2(1-\omega)\right].$$

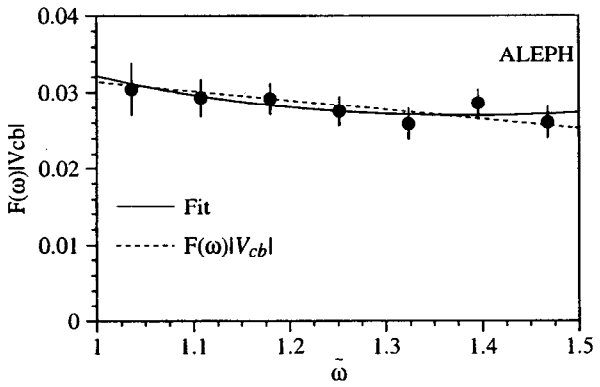


Fig. 14. Linear extrapolation of $F(\omega)|V_{cb}|$ as a function of ω.

This yields $F(1)|V_{cb}|$. However, as the resolution function varies with ω, the linear function is modified for the fit to the data. Such an analysis has been made by ALEPH[18] and their result is shown in Fig. 14, where $F(\omega)|V_{cb}|$ is plotted against ω. In this plot, the underlying linear extrapolation, which is totally adequate for the experimental precision, is shown by the dashed line. Extrapolation to $\omega = 1$ making due allowance for the resolution yields the value of $F(1)|V_{cb}|$. This extrapolation yields

$$F(1)\,|V_{cb}| = (31.4 \pm 2.3 \pm 2.5) \times 10^{-3}$$
$$a^2 = 0.39 \pm 0.21 \pm 0.12.$$

In the limit of infinitely heavy quarks, $F(1)$ is expected to equal one, but for finite mass b quarks, there are corrections which are the source of some controversy. Using Neubert's[19] value for $F(1) = 0.91 \pm 0.04$, the resulting value of $|V_{cb}|$ is

$$|V_{cb}| = (34.5 \pm 2.5 \pm 2.7 \pm 1.5_{theory}) \times 10^{-3}.$$

The analysis also yields a competitive value for the $D^* l\nu$ branching ratio

$$Br(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}) = (5.18 \pm 0.30 \pm 0.62)\%.$$

7.3 The Branching Ratios for $b \to \tau X \nu$ and $b \to \tau \nu$

ALEPH[20] and L3 (Ref. 21) have made the only measurements so far of the $Br(b \to \tau X \nu_\tau)$; ALEPH has also obtained the first upper limit for the exclusive branching ratio $Br(b \to \tau \nu_\tau)$.

The ALEPH analysis takes advantage of their lifetime b tag to first select a pure sample of B decays. They then eliminate b hemispheres in which a lepton is identified and fit the missing energy spectrum in each nonleptonic hemisphere for the τ component. As there are two neutrinos produced in $b \to \tau$ decays, there is more missing energy in the hemisphere than for all other decays, particularly after the removal of the majority of the semileptonic decays to electrons and muons.

In Fig. 15, the hemisphere missing energy plot for this analysis is shown with a clear contribution from the inclusive $b \to \tau X \nu$ decay. The value obtained by ALEPH for the branching ratio is

$$Br(b \to \tau X \overline{\nu}_\tau) = 2.75 \pm 0.30 \pm 0.37\%.$$

The exclusive decay rate, $b \to \tau \overline{\nu}$, is characterized by an even greater hemisphere missing energy, and an upper limit can be established by examining the spectrum above 30 GeV. With current statistics, no signal is observed and a 90% confidence level upper limit of

$$Br(b \to \tau \overline{\nu}_\tau) < 1.8 \times 10^{-3} \text{ at } 90\% \text{ C.L.}$$

is found.

Both values are consistent with predictions based on the Standard Model, 0.023 and 5×10^{-5} respectively. They are of particular interest because the rate could be strongly enhanced by charged Higgs intermediaries in the MSSM and an enhanced

$b \to \tau$ rate had also been considered a possible explanation for the low semileptonic branching ratio. From the measured inclusive rate, a limit

$$\frac{\tan\beta}{m_{H^\pm}} < 0.52 \text{ GeV}^{-1}$$

can be set at 90% C.L.

The L3 value for $Br(b \to \tau X \nu_\tau)$ of $2.4 \pm 0.7 \pm 0.8\%$ agrees with the ALEPH value.

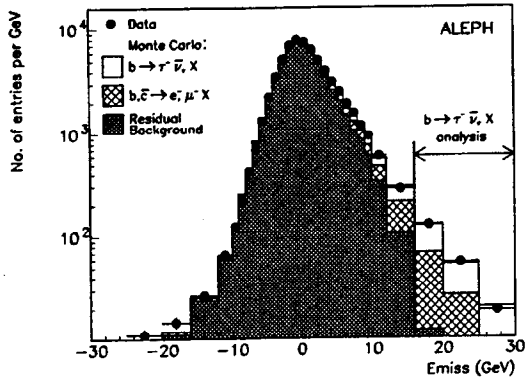


Fig. 15. The hemisphere missing energy spectrum for the ALEPH $b \to \tau X \nu$ analysis.

8 Exclusive Reconstruction and Mass of the Λ_b

Attempts at exclusive reconstruction of the Λ_b have proved to be more difficult than imagined. There are new preliminary results from ALEPH[22] and OPAL,[23] with candidates for the reconstruction $\Lambda_b \to \Lambda_c^+ \pi^-$. Unfortunately, this channel is more subject to misinterpretation from backgrounds from other B decays than the cleanest one, $\Lambda_b \to J/\psi\,\phi$. At the present time, none of the LEP experiments has any candidates for this decay, but OPAL claims one candidate for $\Lambda_b \to \Lambda_c^+ \pi^-$, whilst ALEPH has five candidates with $P_{\Lambda_b} > 20$ GeV and four for $P_{\Lambda_b} > 30$ GeV where potential backgrounds are small. This mass plot is shown in Fig. 16. ALEPH claim a significance of 2.5σ for the four events with $P_{\Lambda_b} > 30$ GeV, and for these they quote a mass of

$$M(\Lambda_b) = 5621 \pm 17 \pm 15 \text{ MeV}.$$

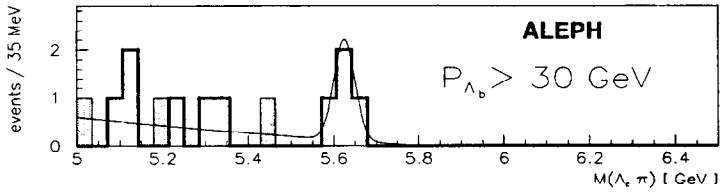


Fig. 16. The Λ_b signal in the ALEPH $\Lambda_c \pi$ mass plot.

9 Lifetimes

Most methods use the semileptonic decays which are isolated most cleanly in the data. Early methods which relied on the impact parameter of the lepton as an estimator for the lifetime have in general given way to methods involving vertexing of both the charm and the bottom vertex. Nevertheless, the impact parameter method still proves useful when vertexing proves difficult, such as with the Λ_b or for inclusive measurements when the exact nature of the final states is poorly known. In the latter case, the weak sensitivity of the impact parameter to the actual momentum of the B state is advantageous.

9.1 Inclusive

The inclusive lifetime, $< \tau_b >$, is given by

$$< \tau_b > = \sum f_i \tau_i$$

with f the fraction of decays for the particular analysis channel. Hence, it need not necessarily be the same in all analyses.

There are two new measurements, a final one from ALEPH[24] using, for the first time, the three-dimensional impact parameter and a new preliminary one from DELPHI[25] using an inclusive vertexing technique. For the former, there is only a low sensitivity to the B momentum, and this correction is taken from Monte Carlo in which the models for semileptonic B decay have been optimized. For the latter, the B momentum is estimated from the visible momentum of the tracks which are used for

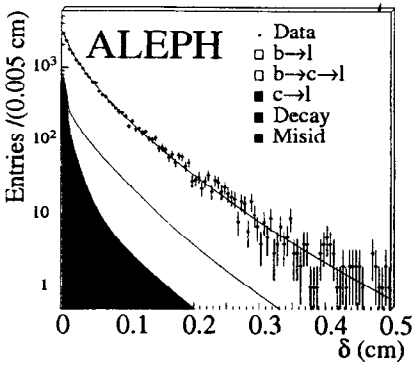


Fig. 17. The ALEPH impact parameter distribution for the mean B lifetime.

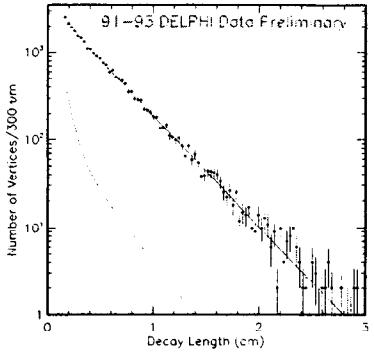


Fig.18. The DELPHI decay length distribution for the mean B lifetime.

the vertexing and this is scaled to the estimated B momentum on the basis of the Monte Carlo. The results are:

$$\langle\tau_b\rangle = 1.533 \pm 0.013 \pm 0.022 \text{ ps} \qquad \text{ALEPH,}$$

$$\langle\tau_b\rangle = 1.600 \pm 0.010 \pm 0.028 \text{ ps} \qquad \text{DELPHI.}$$

Both results are limited by systematics, but they are barely consistent; any discrepancy must be due to either the different event samples or an underestimation of the effects of the unknown B momentum. The plots of the ALEPH impact parameter distribution and the DELPHI decay length distribution with the fits are shown in Figs. 17 and 18, respectively. They show the high statistical quality of the data now available.

9.2 Exclusive

Most measurements of exclusive lifetimes depend upon a partial reconstruction of a semileptonic decay mode which serves both to identify the b state from which the final state originates and to establish both the decay path length and an estimate of the momentum of the B state. The usual procedure is to select events which characterize the particular state in question, such as same side $D_s^{\pm}\ell^{\mp}$ for the B_s. The charm decay products are first identified using the relevant invariant mass. These give the charm momentum vector and are vertexed to give the charm decay point. The lepton and the charm momentum vector are then vertexed to give the B decay point. The magnitude of the B momentum is estimated using the missing energy corrected to take account of non-two jet topologies as described in Sec. 2.6 for the neutrino energy. The B direction is determined from the e^+e^- interaction point and the B decay point.

The techniques for the momentum and direction determination, and the cuts to select the sample, vary considerably from experiment to experiment depending upon the nature of the vertex detector and the hermeticity of the overall detector. The latest exclusive lifetimes are given in Fig. 19, which is taken from the summary by Rizzo[26] at the Beauty '95 conference. Many numbers are still preliminary.

9.2.1 B^+/B^0 Lifetime

Separation of a pure sample of either B^+ or B^0 mesons without full exclusive reconstruction is difficult, and although there are now first results from ALEPH[27] using exclusive reconstruction, the statistics are still too limited to enable measurement at a level which could challenge any predictions. In the absence of a full reconstruction, the traditional method is to rely upon the semileptonic decays and use the fact that the easiest charm state to identify is the D^0 after its decay into $K\pi$, $K\pi\pi$, $K\pi\pi\pi$. Furthermore, it is relatively simple to establish whether this D^0 has resulted from the decay of a charged D^* as this yields a slow charged pion and a very clean peak in the $M(D^0\pi) - M(D^0)$ mass plot. The method then relies upon the fact that in charged B decay, a neutral charm state is produced, whereas in neutral B decay, a charged charm state is produced. Hence, an identified D^{*+} with a negative lepton is assumed to have originated from a B^0 whilst a D^0, not identified as the daughter of a charged D^*, is assumed to have originated from a $B^{\pm}$.

The main difficulty in the method results from the difficulty in knowing the charged B background in the neutral B sample and vice-versa. Potential causes of this result from failure to successfully reconstruct the slow charged pion from the D^* decay, background under the D^* signal resulting from combinatorial association of an unrelated π with a D^0 from a charged B, and the poorly known production rate for higher D^*'s (D^{**}) in semileptonic B decay. The latter is the most serious, and the difficulty is compounded by the even less well-known relative production ratios of the possible D^{**} states. Some of these can decay to $D^*\pi$, others to $D\pi$, and one to both. The overall percentage of decays to D^{**} states is believed to be about 30%. This leads

to a systematic error, but determining the magnitude of this is difficult as the error depends upon the ratio of the lifetimes, and hence, the uncertainty from cross population of the channels can only be estimated when the lifetimes are well-known! The present results from LEP using this method show a surprising tendency to give a B^+/B^0 lifetime ratio of one. The results are:

ALEPH[27]	$0.98 \pm 0.08 \pm 0.02$
DELPHI[28]	$1.00 \pm 0.16 \pm 0.10$
OPAL[29]	$0.99 \pm 0.14 \pm 0.05$.

With such uncertainties, the fact that the results are all so close to one is surprising, and even more so when one compares the mean B^+/B^0 lifetime using this method which is 1.60, 1.61, and 1.53 for the three experiments respectively.

In addition to the ALEPH exclusive reconstruction, another method which has been used to determine these lifetimes at LEP is a topological one from DELPHI[30] in which they attempt to determine the number of charged tracks emanating from the B decay. This requires an excellent understanding of the performance of the vertex detector as the efficiency has to be established from the simulation. The results are given in Fig. 19. The best LEP average values given by Rizzo at Beauty '95 are

$$\tau(B^+) = 1.63 \pm 0.06 \text{ ps}$$
$$\tau(B^0) = 1.56 \pm 0.06 \text{ ps}.$$

The LEP measurements are, thus, consistent with the difference between the B^+ and B^0 lifetimes, which is expected at the ~5% level. The current error quoted is about 6%, but because of the difficulties associated with fully separating the two charge states, it is probably unlikely that LEP measurements will be capable of measuring a difference with much greater precision. Full reconstruction would seem to be by far the most successful approach, but this will only be possible with an adequate sample at the hadron colliders.

9.2.2 B_s Lifetime

The semileptonic decay of the B_s to $D_s X \ell \nu$ resulting in same side $D_s^{\pm}\ell^{\mp}$ gives the best result, but statistics are limited. Such measurements have been performed by all three experiments with the average dominated by a recent ALEPH[31] result with an error ~0.16 ps.

DELPHI[32] and ALEPH[33] have also used other methods which depend primarily on the observation of just a D_s in the event. Such techniques produce a much larger sample of events but suffer from high and poorly known backgrounds which lead to substantial uncertainties. The current results are given in Fig. 19, and the average LEP value is

$$\tau(B_s) = 1.59 \pm 0.11 \text{ ps}.$$

The average is extremely close to the values for the B^+ and B^0 as expected.

9.2.3 Λ_b Lifetime

Present evidence on the B meson decays suggests that the spectator model with small QCD corrections is valid. Such models also predict that the Λ_b lifetime, effectively the lifetime of the lowest lying B baryon, should be within 10% of the meson lifetimes. Measurements of this quantity come only from the LEP experiments, and they are now giving a precision which can, and does, seriously challenge this prediction.

All measurements rely upon the correlation of a baryon with an appropriately charged same-side lepton. The decay chain used is

$$\Lambda_b^0 \rightarrow \Lambda_c^+ \ell^- \bar{\nu}$$
$$\Lambda_c^+ \rightarrow \Lambda X, pX \quad ,$$

and the signal is then isolated using a same-side correlation of $\Lambda\ell^-$, $p\ell^-$, or $\Lambda_c^+\ell^-$ with the Λ_c^+ decaying to $pK^-\pi^+$, $\Lambda\pi^+\pi^+\pi^-$, or pK^0. The actual methods used differ between the collaborations. All use a vertexing procedure for the $\Lambda_c^+\ell^-$ events, but whilst OPAL[34] and DELPHI[35] use a similar technique for the statistically superior $\Lambda\ell^-$ and $p\ell^-$ samples, ALEPH[36] rely upon the lepton impact parameter rather than vertex the Λ with the lepton. The results from the three experiments are remarkably consistent, all showing a value about 20% lower than the B meson lifetimes. The actual values are shown in Fig. 19, and the OPAL data are shown in Fig. 20. The LEP average is

$$\tau(\Lambda_b) = 1.20 \pm 0.07 \text{ ps}.$$

This is 25 ± 8% less than the average B^+/B^0 lifetime and suggests that additional corrections are necessary in the spectator model to satisfactorily describe the bottom baryons.

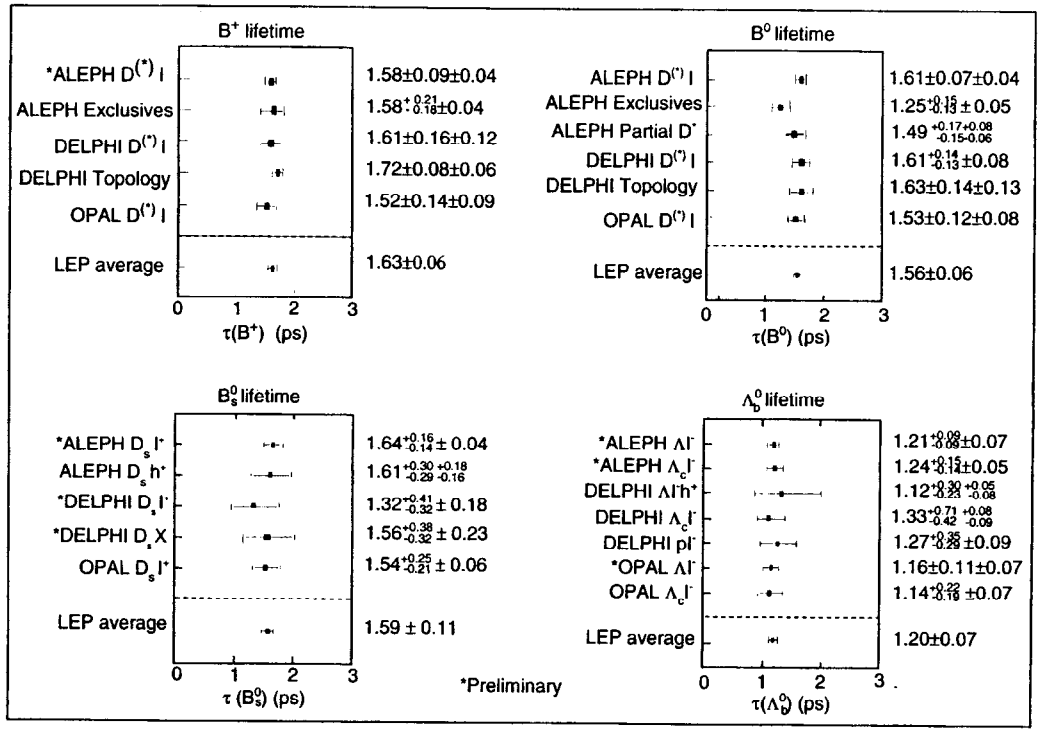


Fig. 19. Exclusive B lifetime results from Rizzo's review at Beauty'95.

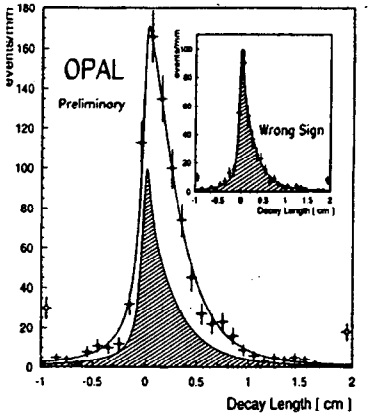


Fig. 20. The decay length distribution for the OPAL Λ_b lifetime measurement.

9.2.4 Ξ_b Lifetime

A preliminary measurement has been made of the Ξ_b lifetime by ALEPH[37] using same-side correlation of $\Xi^{\pm}\ell^{\mp}$ to identify the events and a similar impact parameter procedure to the one used for the Λ_b. The value obtained, $\tau(\Xi_b) = 1.15^{+0.35}_{-0.25} \pm 0.20$ ps, still has substantial uncertainty but again shows a low value with respect to the B meson lifetimes. The method used cannot distinguish the Ξ_b^0 from the Ξ_b^-. An earlier preliminary DELPHI[38] measurement had given $1.5^{+0.7}_{-0.4} \pm 0.3$ ps.

10 Time-Dependent Mixing

The use of the silicon vertex detectors has enabled the observation and measurement of the $B^0 - \overline{B}^0$ oscillations resulting from the second-order weak process. The results are quite beautiful.

For a produced B^0, the decaying particle is either a B^0 or $\overline{B}^0$ given by:

$$\begin{matrix} B^0 \\ \overline{B}^0 \end{matrix} = \tfrac{1}{2} e^{-\Gamma t}\left[1 \begin{matrix} + \\ - \end{matrix} \cos \Delta m t\right],$$

where Δm is the mass difference between the two B^0 states and Γ is the decay constant. This is usually assumed to be the same for both B^0 states, although there are suggestions that for the B_s^0, there could be a lifetime difference approaching 10% between the two mass eigenstates. The quantity $x = \Delta m / \Gamma$ can be expressed, for the B_d^0, by

$$x_d = \tau_{B_d} \frac{G_F^2}{6\pi^2} m_{top}^2 f_{B_d}^2 B_{B_d} f_1\left(\frac{m_t^2}{m_W^2}\right)\left|V_{td}^* V_{tb}\right|^2.$$

This contains the important matrix element V_{td}, which defines one of the sides of the conventional unitarity triangle describing CP violation in the B system, but as the structure and bag constants are poorly known, a measurement of x_d cannot yield an accurate measurement of V_{td}. However, whilst the absolute values of these constants are not well-known, many uncertainties drop out when one considers the ratio of mixing in the B_s and B_d systems. The analogous relation for x_s is

$$x_s = \tau_{B_s} \frac{G_F^2}{6\pi^2} m_{top}^2 f_{B_s}^2 B_{B_s} f_1\left(\frac{m_t^2}{m_W^2}\right)\left|V_{ts}^* V_{tb}\right|^2,$$

and predictions for the ratio give

$$\frac{x_s}{x_d} = (1.34 \pm 0.15)\left|\frac{V_{ts}}{V_{td}}\right|^2.$$

Hence, as one expects that $|V_{ts}| \cong |V_{cb}|$, measurement of x_d and x_s enable a measurement of V_{td}. For the B_d, the oscillation time is comparable with the lifetime, and so x_d can be determined from the integrated mixing parameter, χ, which can be expressed in terms of x_d by

$$\chi = \frac{x^2}{2(1+x^2)}.$$

This has been measured from the like-sign dilepton rate at the $\Upsilon(4S)$ where only the B_d is produced, but where time-dependent measurements are not practicable.

For the B_s, however, the oscillation rate is predicted to be considerably higher, and so integrated measurements have no sensitivity to x_s. This is in agreement with current integrated measurements made at LEP which are all consistent with the maximum χ value of 0.5 corresponding to infinite x_s.

To measure the oscillations, it is necessary to measure the decay time for a $B^0(\overline{B}^0)$ state to decay by means of a channel which reveals whether it was a B^0 or $\overline{B}^0$ at decay, and it is also necessary to use a tag which identifies whether it was produced as a B^0 or $\overline{B}^0$. The situation is described in Fig. 21. The identified event is split into two halves usually with respect to the thrust axis, and then on the probe side, it is necessary to determine the B^0 or $\overline{B}^0$ nature and estimate the proper time for the decay from the reconstructed B momentum and decay distance. On the opposite side, the tag side, some property is used to identify the nature of the B on that side which, after corrections for dilutions from mixing, backgrounds, etc., tags the nature of the B produced on the probe side.

There are now an increasing number of signatures, particularly for the probe side. Those for which results are currently available use either a D^*, a D^* and lepton, or just a lepton on the probe side with either a lepton or a measure of the jet charge on the tag

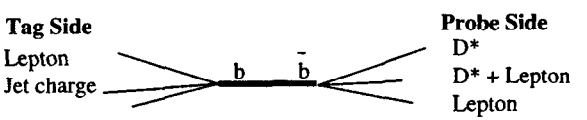


Fig.21. Possible arrangements for the measurements of B oscillations.

side. The D^* and D^*-lepton procedures allow better vertexing for the proper time, but as the B_s does not decay to a D^*, these measurements only give information on x_d. When a lepton is used on the probe side, the decay can come from either a B_d or a B_S, and so these methods have the potential to also give information on Δm_s.

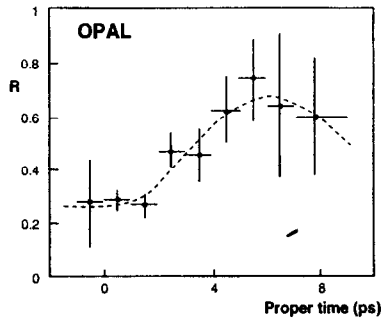


Fig. 22. The proper time distribution for the proportion of like sign objects in the OPAL D*-leptonvs jet charge Δm_d analysis.

10.1 Measurement of Δm_d

All three experiments have now produced excellent results for B_d oscillations and obtained values for x_d which, when averaged over the experiments, yield a more accurate measurement than that obtained from the integrated measurements at the $\Upsilon(4S)$. The quality of the data can be seen in sample results from OPAL,[39] ALEPH,[40] and DELPHI[41] shown in Figs. 22, 23, and 24. All three plots use different analyses; for the OPAL data, Fig. 22, the probe is a D^* with a lepton and the tag is a measure of the jet charge; for ALEPH, Fig. 23, leptons are used for both probe and tag, and for the DELPHI analysis, the probe is a lepton and the tag the jet charge. The values from these and other measurements are given in Fig. 25 which is taken from Stocchi's[42] summary at Beauty '95. The mean value is 0.456 ± 0.020 ps^{-1} to be compared with 0.428 ± 0.050 from the integrated measurement at the $\Upsilon(4S)$. The LEP value corresponds to an x_d value of 0.711 ± 0.044 for the above B_d^0 lifetime.

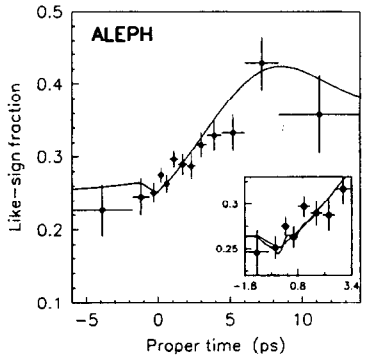


Fig. 23. Like sign fraction as a function of proper time for the ALEPH lepton-lepton Δm_d analysis.

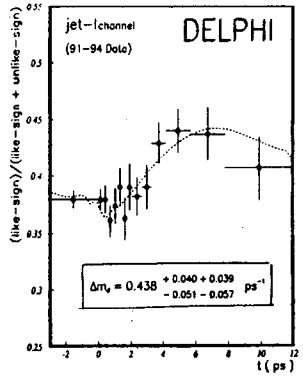


Fig. 24. The proper time distribution for the proportion of like sign objects in the DELPHI lepton vs jet charge Δm_d analysis.

10.2 Lower Limit on Δm_s

The real prize for LEP1, however, would be a definitive measurement of x_s. So far, only lower limits have been given, but these are now reaching values which are in the range expected by theoretical estimation from x_d. These suggest that x_s could be anywhere between a factor of ten to 60 times greater than x_d. The determination of the lower limit at a particular confidence level, and hence, the combination of results from the LEP experiments, is, however, a matter of considerable topical concern and discussion.

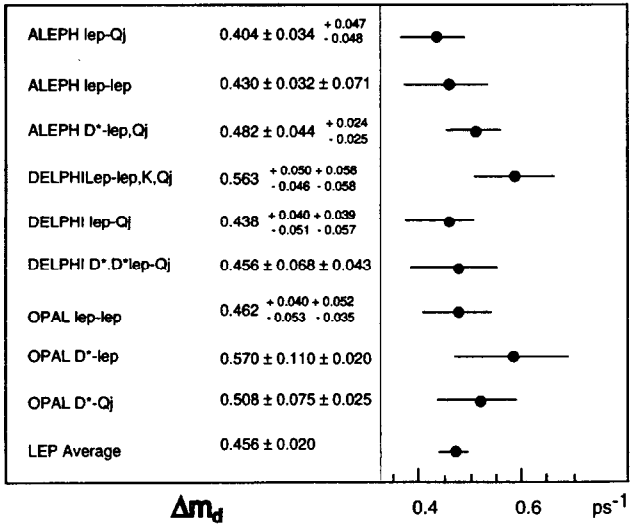


Fig. 25. Current values of Δm_d from the LEP experiments taken from Stocchi's review at Beauty'95.

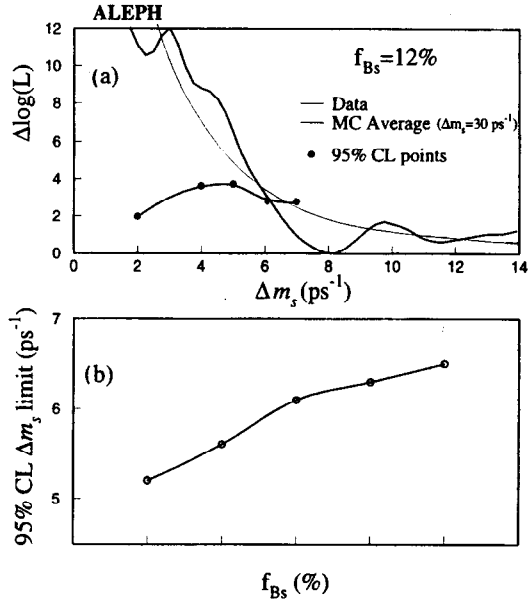


Fig. 26. ALEPH results on the lower limit for Δm_S using the lepton-jet charge technique. The difference in log likelihood is shown in (a) as a function of Δm_s whilst in (b) the sensitivity of the limit to the assumed proportion of B_S mesons is shown.

The method which is most commonly employed is to make an unbinned maximum likelihood fit to the like-sign proper time distributions which have a B_s component and then plot $\Delta\mathcal{L} = (\log\mathcal{L} - \log\mathcal{L}_{min})$ vs Δm_s. If the estimated errors are both correct and Gaussian, then the 95% C.L. is given when $\Delta\mathcal{L} = 2$. Problematic is the estimation of the systematic uncertainty. One technique is to use many MC samples for different values of Δm_s in which the experimental uncertainties are parameterized, and a limit taken so that 95% of the samples yield a lower value. This technique can also show if in the real data, a statistical fluctuation had artificially helped to yield a higher limit than could be expected from the sensitivity of the detector and the statistics of the measurement. This is referred to as the "luck" factor, and again there are mixed opinions about whether it is appropriate to quote a high limit which has primarily resulted from a lucky fluctuation.

With these caveats, the present results are quoted in Table 2, although more can be expected for the summer conferences. The best current value comes from the

ALEPH[43] measurement using a lepton on the probe side and jet charge on the tag side; the result is shown in Fig. 26(a). However, when comparing results, it is crucial to know the proportion, f_s, of B_s assumed in the event sample. The best DELPHI[41] result also results from a lepton-jet charge technique, but as they assume a lower value for fs, the results are quite similar. The sensitivity of the limit to the assumed value of f_s for the ALEPH result is shown in Fig. 26(b). An OPAL update is expected imminently.

	f_s	Technique	Δm_s limit, ps^{-1}
ALEPH	0.12	lep-lep	> 5.6
	0.12	lep-Q_{jet}	> 6.1
	0.12	K-lep	> 4.0
DELPHI	0.10	lep-Q_{jet}	>4.2
	0.10	$D_s$lep-Q_{jet}	>1.5
OPAL	0.12	lep-lep	>2.2

Table 2. Lower limits on Δm_s.

So far, no satisfactory method of combining the results has been established. This will probably be necessary if LEP is to have any chance of determining x_s. It is amusing to speculate that many of the likelihoods presently minimize at a Δm_s value around eight. Such a value is well within the capabilities of the LEP detectors so there can be a real hope that by combining results and using all the data, including that from 1995, x_s could be the last major measurement from LEP1.

11 Summary and Outlook

Primarily due to the success of the silicon vertex detectors at tagging the long-lived B states, heavy flavor physics has become a major part of the LEP program. Many results are still statistically limited, and so a doubling or tripling of present statistics would add considerably to our knowledge; they would also help many measurements currently systematically limited as understanding of poorly known branching ratios and decay distributions would continue to improve. However, 1995 is likely to be the last year at which LEP will take any substantial data at the Z, and in 1995, no more than a million hadronic Z decays per experiment can be hoped for.

Nevertheless, important questions remain to be answered. With the increased data, it should become possible to fix finally the semileptonic b-branching ratio; it is almost certainly lower than theoreticians would wish, but a measurement to a few percent would give them a goal. It is also intriguing why this measurement is always found to be higher at LEP than at the $\Upsilon(4S)$. In a similar area, the ratio of the B baryon to the B-meson lifetime would welcome further improvement, although it is now clearly lower than simple predictions. In both of these, we are now aware that the data is not in perfect agreement with expectations, and therefore, further advances in the theory are necessary.

Decays of the Z to b quarks gives a unique opportunity to investigate the Standard Model in the quark sector; in particular, an accurate measurement of the basic Z to b width R_b provides one of the most sensitive tests to radiative corrections in the Standard Model and to possible non-Standard Model effects. The current 2-3σ discrepancy is enticing and the cause of a major effort by the experiments to understand their systematics and learn how to reliably combine results. It is reasonable to expect a further decrease on the overall uncertainty on R_b over the next year.

However, the greatest aim for the LEP program over this final year must be a definitive measurement of x_s, the mixing parameter in the B_s system. With the latest data and an understanding of how to combine results, it should certainly be possible to measure this up to about 12 and maybe to 15, which would be well within the expected range. A definitive measurement would be of major significance to B phenomenology over the next ten years, as, if not measured at LEP, this may have to wait for the LHC.

The final year of heavy flavor physics at the Z could still provide great excitement.

References

[1] M. Karlsson, 6th International Symposium on Heavy Flavour Physics, Pisa, June 1995.

[2] L. Ludovici, 6th International Symposium on Heavy Flavour Physics, Pisa, June 1995.

[3] C. Petersen, D. Schlatter, I. Schmitt, and P. Zerwas, Phys Rev. D **27**, 105 (1983).

[4] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **357**, 699 (1995).

[5] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov, Phys. Lett. B **78**, 615 (1978).

[6] B. Andersson, G. Gustavson, and B. Söderberg, Z. Phys. C **20**, 317 (1983); M. G. Bowler, Z. Phys. C **11**, 169 (1981); D. A. Morris, Nucl. Phys. B **313**, 634 (1989).

[7] OPAL Collaboration, G. Alexander *et al.*, "A study of b quark fragmentation into B^0 and B^+ mesons at LEP," CERN-PPE/95-122.

[8] L3 Collaboration, M. Acciarri *et al.*, Phys. Lett. B **345**, 589 (1995).

[9] ALEPH Collaboration, D. Buskulic *et al.*, "Production of excited beauty states in Z decays," CERN-PPE/95-108.

[10] DELPHI Collaboration, P. Abreu *et al.*, "B^* observation and properties," CERN-PPE/95-53.

[11] DELPHI Collaboration, P. Abreu *et al.*, Phys. Lett. B **345**, 598 (1995); DELPHI Collaboration, P. Abreu *et al.*, "Observation of orbitally excited B and B_s mesons," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[12] OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **66**, 19 (1995).

[13] ALEPH Collaboration, D. Buskulic *et al.*, "Resonant structure and flavour-tagging in the $B\pi$ system using fully reconstructed B decays," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[14] DELPHI Collaboration, P. Abreu *et al.*, "First evidence for Σ_b and Σ_b^* baryons," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[15] ALEPH Collaboration, D. Buskulic *et al.*, "Measurement of the Λ_b polarisation in Z decays," CERN-PPE/95-156.

[16] ALEPH Collaboration, D. Buskulic *et al.*, "Measurement of the semileptonic b branching ratios from inclusive leptons in Z decays," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[17] M. Schmitt, 6th International Symposium on Heavy Flavour Physics, Pisa, June 1995.

[18] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **359**, 236 (1995).

[19] M. Neubert, CERN-TH/95-107 (1995).

[20] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **343**, 444 (1995).

[21] L3 Collaboration, A. Adam *et al.*, Phys. Lett. B **332**, 201 (1994).

[22] ALEPH Collaboration, D. Buskulic *et al.*, "Λ_b exclusive decay," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[23] OPAL Collaboration, G. Alexander *et al.*, "Search for exclusive Λ_b decays with the OPAL detector at LEP," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[24] ALEPH Collaboration, D. Buskulic *et al.*, "A precise measurement of the average b hadron lifetime," CERN PPE/95-121.

[25] DELPHI Collaboration, P. Abreu *et al.*, "Updated precision measurement of the average lifetime of B hadrons," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[26] G. Rizzo, Third International Workshop on B-Physics at Hadron Machines, Oxford, July 1995.

[27] ALEPH Collaboration, D. Buskulic *et al.*, "Measurement of the $\overline{B}^0$ and B^- meson lifetimes," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[28] DELPHI Collaboration, P. Abreu *et al.*, "A measurement of B^+ and B^0 lifetimes using $\overline{D}\ell^+$ events," CERN-PPE/95-60.

[29] OPAL Collaboration, R. Akers *et al.*, "Improved measurements of the B^0 and B^+ meson lifetimes," CERN-PPE/95-19.

[30] DELPHI Collaboration, P. Abreu *et al.*, "Lifetimes of charged and neutral B hadrons using event topology," CERN-PPE/95-059.

[31] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **361**, 221 (1995).

[32] DELPHI Collaboration, P. Abreu *et al.*, Z. Phys C **61**, 407 (1994).

[33] ALEPH Collaboration, D. Buskulic *et al.*, "Measurement of the D_s meson production in Z decays and of the B_s lifetime," CERN-PPE/95-092.

[34] OPAL Collaboration, R. Akers *et al.*, "Measurement of the average b-baryon lifetime and the product branching ratio $f(b \to \Lambda_b) \cdot BR(\Lambda_b \to \Lambda \ell^- \overline{\nu} X)$," CERN-PPE/95-90.

[35] DELPHI Collaboration, P. Abreu *et al.*, "Lifetime and production rate of beauty baryons from Z decays," CERN-PPE/95-54.

[36] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **357**, 685 (1995).

[37] ALEPH Collaboration, D. Buskulic *et al.*, "Evidence for strange b baryon production in Z decays," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[38] DELPHI Collaboration, P. Abreu *et al.*, "Production of strange B-baryons decaying into $\Xi^{\pm}\ell^{\pm}$ pairs at LEP," CERN-PPE/95-29.

[39] OPAL Collaboration, R. Akers *et al.*, Phys. Lett. B **336**, 411 (1994).

[40] ALEPH Collaboration, D. Buskulic *et al.*, "Measurement of the $B_d - \overline{B}_d$ oscillation frequency," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[41] DELPHI Collaboration, P. Abreu *et al.*, "Study of $B^0 - \overline{B}^0$ oscillations," submitted to the EPS HEP95 Conference, Brussels, July 1995.

[42] A. Stocchi, 3rd International Workshop on B-physics at Hadron Machines, Oxford, July 1995.

[43] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **356**, 409 (1995).

NEW RESULTS ON CLEO'S HEAVY QUARKS — BOTTOM AND CHARM

Scott Menary
University of California, Santa Barbara
menary@charm.physics.ucsb.edu

Representing the CLEO Collaboration

ABSTRACT

While the top quark is confined to virtual reality for CLEO, the increased luminosity of the Cornell Electron Storage Ring (CESR) and the improved photon detection capabilities of the CLEO II detector have allowed for a rich program in the physics of CLEO's "heavy" quarks — bottom and charm. I will describe new results in the B meson sector including the first observation of exclusive $b \to u\ell\nu$ decays, upper limits on gluonic penguin decay rates, and precise measurements of semileptonic and hadronic $b \to c$ branching fractions. The charmed hadron results that are discussed include the observation of isospin violation in D_s^{*+} decays, an update on measurements of the D_s^+ decay constant, and the observation of a new excited Ξ_c charmed baryon. These measurements have had a large impact on our understanding of heavy quark physics.

1 Introduction

The central goal of heavy flavor physics below the top quark threshold is to measure the elements of the Cabibbo-Kobayashi-Maskawa (CKM) Matrix, since it is the Standard Model prescription for $\mathcal{CP}$ Violation. This can be accomplished through measuring a multitude of b hadron decay rates and time-dependent decay asymmetries. Besides being interesting in and of itself, charmed hadron physics is also relevant to B decays through being a laboratory where many of the questions regarding QCD effects in heavy flavor decay can be addressed. After a short description of the experimental considerations of doing physics at the $\Upsilon(4S)$, I will describe a host of B meson and charmed hadron measurements done by CLEO in the last half-year.[1] I will also try to give some feel for the implications of these measurements towards our understanding of heavy flavor physics.

2 CLEO and CESR

At CESR, the highest luminosity collider in the world, the electron and positron beam energies are set to perform physics in the region of the Υ resonances, the system of $b\bar{b}$ bound states. The hadronic e^+e^- cross-section in this center-of-mass region is shown in Fig. 1. All the Υ resonances lower in mass than the $\Upsilon(4S)$ are

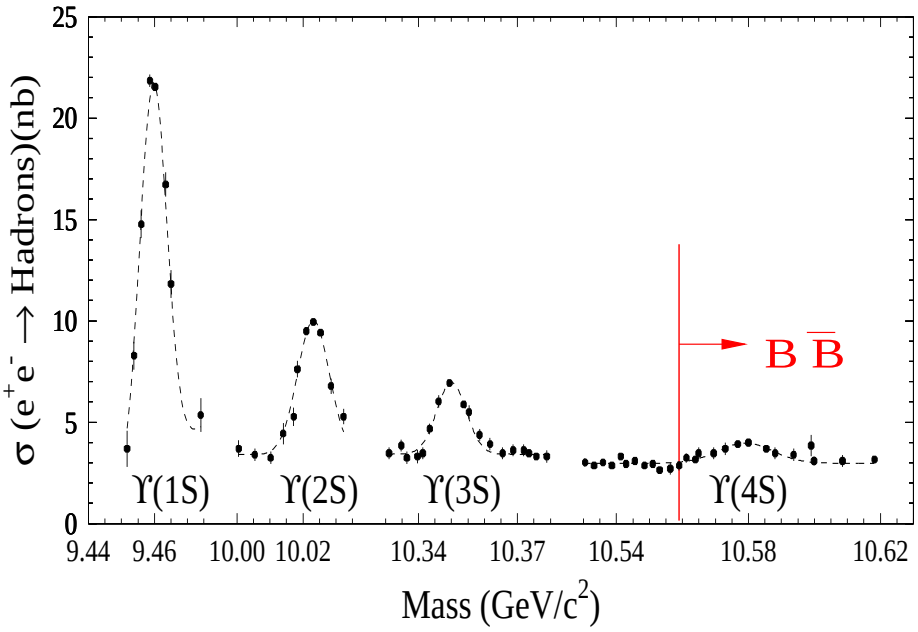


Figure 1: The e^+e^- cross-section in the Υ region.

below threshold for producing a $B\bar{B}$ pair. The cross-section at the $\Upsilon(4S)$ is about a nanobarn above the "continuum" cross-section of ~ 3.4 nb, and $c\bar{c}$ production constitutes about a nanobarn of the continuum. Hence, every fb^{-1} of data taken at the $\Upsilon(4S)$ contains about 10^6 $B\bar{B}$ and $c\bar{c}$ pairs. Further, the b quark decays essentially 100% of the time to a c quark giving another million charmed particles per fb^{-1}. The CLEO data sample to date consists of 3.3 fb^{-1} of data taken at the $\Upsilon(4S)$ resonance (so-called "on-resonance" data) and about half as much taken at an energy below the $B\bar{B}$ threshold (referred to as "continuum" or "off-resonance" data). The results discussed in this paper are based on about two-thirds of this data.

The CLEO II detector, shown in a cutaway r-z view in Fig. 2, measures both charged and neutral particles with excellent efficiency and resolution.[2]

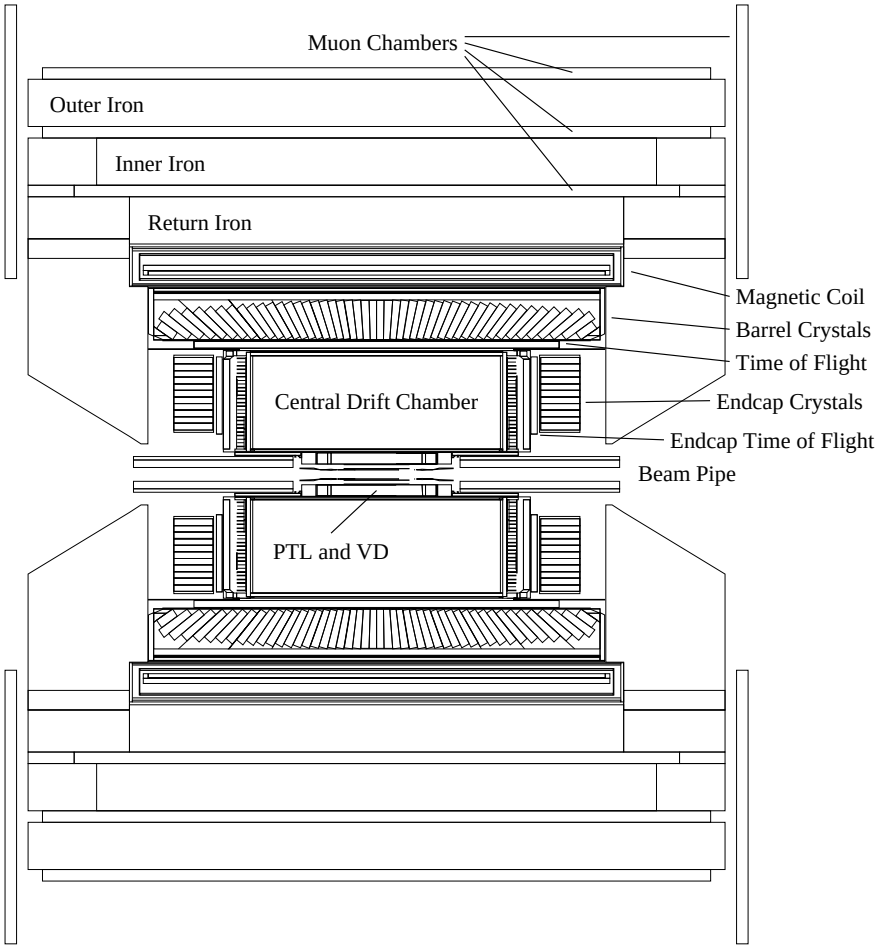


Figure 2: An r-z view of the CLEO II detector. For scale, the Time of Flight (TOF) system is at a radius of ~1 meter.

The tracking detectors, the TOF system, and the CsI calorimeter are located inside the 1.5 Tesla magnetic field. Analyses requiring good lepton identification or maximal photon energy resolution generally restrict the candidate particles to having been detected in the "barrel" part of the detector (i.e., $\cos\theta < 0.71$ where θ is the angle between the particle's momentum and the beam axis). Most analyses have a minimum electron momentum requirement of ~ 600 MeV while the detector is only efficient for muons above about 1 GeV. Photons of energy down to 30 MeV are used while the charged tracking efficiency turns on at around 50 MeV and plateaus in the mid- to high-$90^{\rm th}$ percentile, depending on particle type and track quality requirements. Particle identification is done using dE/dx measurements in the main drift chamber and the TOF system. This provides greater than 2σ K-π separation up to about 1 GeV, with much better separation than this from dE/dx up to about 700 MeV, and around 1.8σ separation at 2.4 GeV, the momentum region of interest for charmless hadronic B decay searches.

There are several unique aspects of doing physics at the $\Upsilon(4S)$ which make it especially good for studying B mesons and charmed hadrons. First, considering B meson production, since the $\Upsilon(4S)$ is just above the threshold for producing $B\bar{B}$, there are no B^*, B_s, or Λ_b hadrons produced, nor are there any extra particles produced along with the $B\bar{B}$ pair. This leads to the very powerful constraint that the B meson energy is equal to the beam energy. This can be used to select B meson candidates by requiring that ΔE, the difference between the measured sum of the charged and neutral energies of the daughters of the B candidate and the beam energy, be close to zero. Also, the B mass resolution is greatly improved by using the beam-constrained mass (M_B), defined by:

$$M_B^2 = E_{beam}^2 - \left(\sum_i \vec{p_i}\right)^2 , \qquad (1)$$

where $\vec{p_i}$ is the momentum of the i-th daughter of the B candidate. The M_B resolution of about 2.6 MeV is determined by the beam energy spread and is a factor of ten better than the resolution in invariant mass obtained from simply summing the four-momenta of the B daughters.

Since the B mesons are produced almost at rest (the average B momentum is $\sim$ 320 MeV), their decay products are uniformly distributed throughout the volume of the detector leading to events that tend to be "spherical" in shape, as illustrated in Fig. 3. Continuum $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) events are more

jet-like in structure, as shown in Fig. 4. Event shapes are utilized to distinguish $B\bar{B}$ events from continuum events. Charmed hadrons produced in the continuum can often be the largest source of background in B physics analyses at the $\Upsilon(4S)$ (the $B \to K^*\gamma$ analysis is an example of such) which is why some amount of data is taken below $B\bar{B}$ threshold so as to be able to study the characteristics of these events.

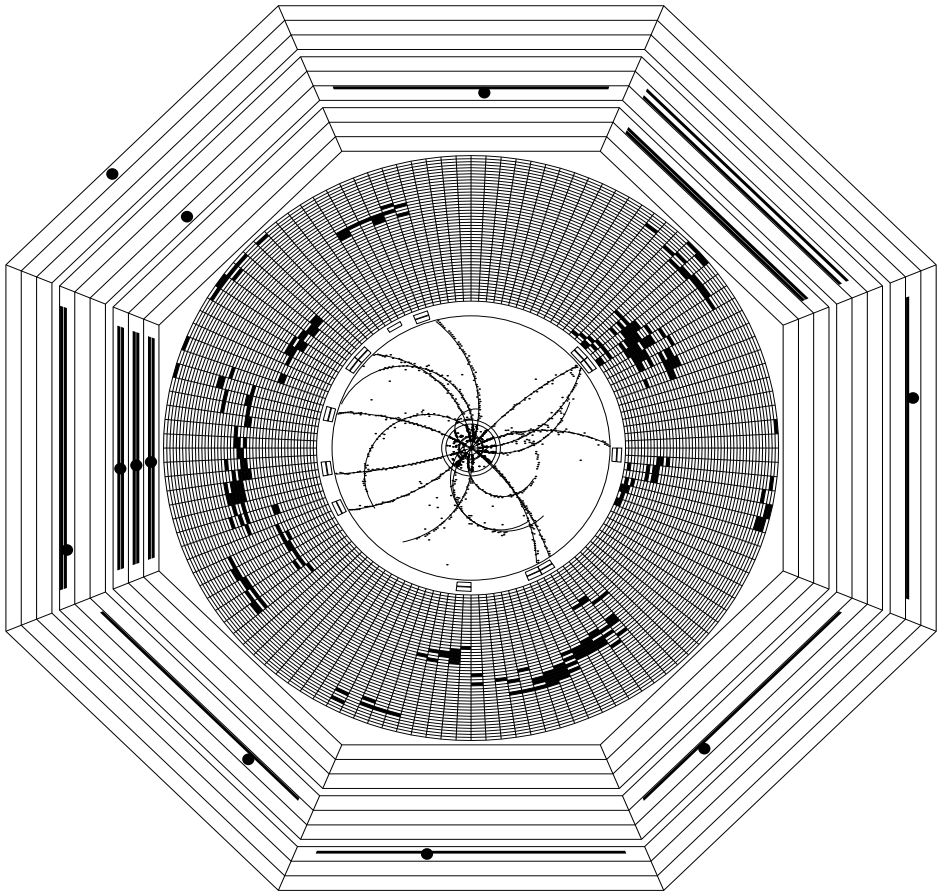


Figure 3: An r-ϕ view of a $B\bar{B}$ event in the CLEO II detector. See Figure 4 for a description of the display.

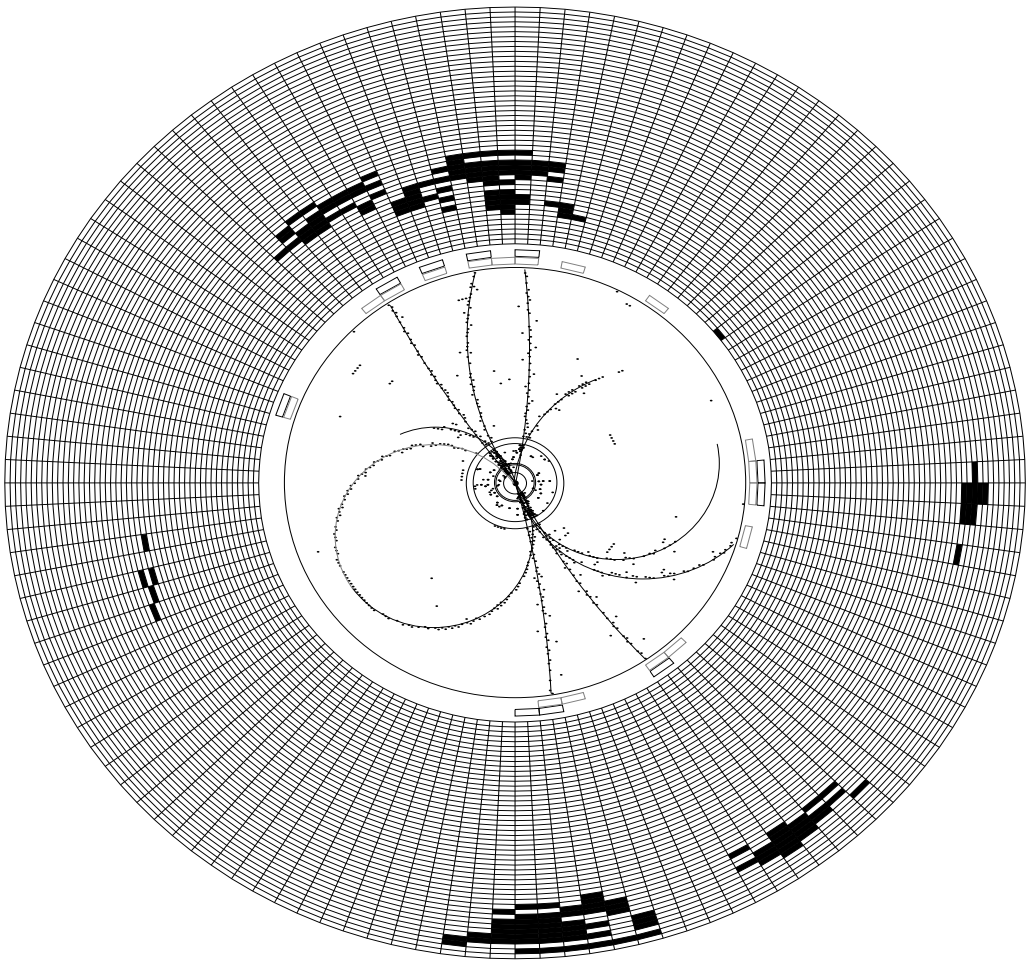


Figure 4: A fully reconstructed, off-resonance $e^+e^- \to \Lambda_c^- \pi^+ \pi^- \Lambda_c^{*+}(2593)$ event where the Λ_c^{*+} decayed to $\Lambda_c^+ \pi^+ \pi^-$ and both Λ_c's decayed to $pK\pi$. Starting from the center are shown the hits and reconstructed tracks in the tracking devices. Outside of this is a thin annulus where hit TOF counters are shown as empty rectangles. Beyond this is a representation of the calorimeter which gives information in the z view as well. The crystals are shown as boxes where the inner radius is furthest from the viewer (in z) and the outer radius is closest. A crystal is blackened where an energy deposition above some threshold was measured. The muon counters (not shown here but in Figure 3) in the return yoke of the magnet are outside of the calorimeter.

3 Charmed Hadrons

In $e^+e^- \to c\bar{c}$ events at 10.58 GeV center-of-mass, the charmed hadron carries most of the charmed quark's energy (which is the beam energy). Conversely, the absolute kinematic cutoff for charmed hadrons from B decay is $m_B/2 \approx 2.5$ GeV. This is illustrated by the inclusive D_s^+ momentum spectrum in Fig. 5 where there is a clear demarcation in momentum between D_s^+ mesons produced in B decay and those from the continuum. Since the combinatorial background generally

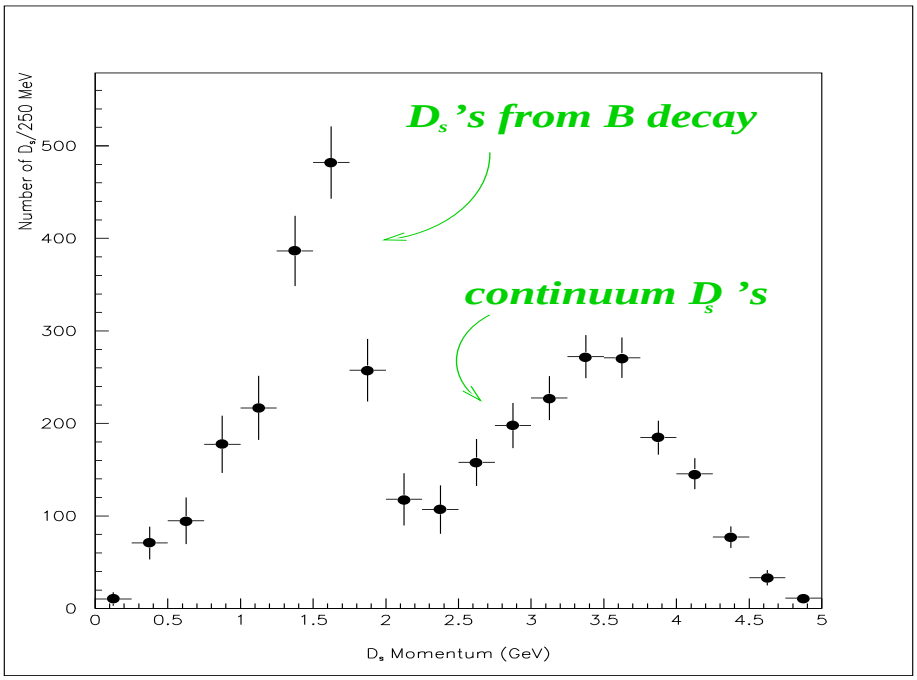


Figure 5: The D_s^+ momentum spectrum.

falls sharply with increasing momentum, most charm analyses require a minimum charmed hadron candidate momentum of 2.5 to 3 GeV.*

The discovery of the large lifetime difference between the D^+ and D^0 mesons was the first clue that charmed meson decays are much more complicated than the simple spectator picture would predict. In some senses, charmed baryons are even more interesting because, for example, the W-exchange diagram is not helicity suppressed. The fact that the Λ_c lifetime is about half that of the D^0 supports the view that interesting things are happening in charmed baryon hadronic decays.

*Another variable that is often used is $x \equiv p/p_{max}$ where $p_{max} = \sqrt{E_{beam}^2 - m_{hadron}^2}$. So, requiring $x > 0.5$ or 0.6 is equivalent to the momentum cut mentioned in the text.

3.1 Charmed Baryons — The Year of the Ξ_c

Last year CLEO published many new Λ_c results including: the observation of a new excited Λ_c,[3] measurements of semileptonic Λ_c decays,[4] and the observation of new hadronic decay channels involving the η.[5] This year has seen mostly new results on the Ξ_c, the csq baryon (where $q = u$ or d).

3.1.1 New Ξ_c Decay Modes

The CLEO result[6] of $\tau(\Xi_c^+)/\tau(\Xi_c^0) = 2.46 \pm 0.70^{+0.33}_{-0.23}$ is another indication that it is going to take the observation of many different Ξ_c hadronic decay modes to untangle the relative strengths of the various diagrams involved in Ξ_c decay.† Simple spectator decay results in Ξ_c^+ decays to Ξ^0, and CLEO reported previously on the observation of such decay modes.[7] Other hyperons, such as the Σ^+ or Λ, are produced through more exotic decay mechanisms like internal W emission.

New CLEO results on Ξ_c^+ decays are given in Table 1. The Λ is reconstructed in its decay to $p\pi^-$ while the $p\pi^0$ channel is used to reconstruct the Σ^+. The $\Sigma^+K^-\pi^+$ final state is found to be roughly 50% two-body $\Sigma^+\bar{K}^{*0}$ and 50% nonresonant.

Decay Mode	x_p cut	Events	$\mathcal{E}$ (%)	$\mathcal{B}/\mathcal{B}(\Xi_c^+ \to \Xi^-\pi^+\pi^+)$
$\Sigma^+K^-\pi^+$	0.5	119 ± 23	10.4	$1.18 \pm 0.26 \pm 0.17$
$\Sigma^+\bar{K}^{*0}$	0.5	61 ± 17	9.8	$0.92 \pm 0.27 \pm 0.14$
$\Lambda K^-\pi^+\pi^+$	0.6	61 ± 15	11.5	$0.58 \pm 0.16 \pm 0.07$
$\Xi^-\pi^+\pi^+$	0.5	131 ± 14	10.6	1.0

Table 1: Summary of results on new Ξ_c^+ decay modes. The $\Sigma^+K^-\pi^+$ mode includes both resonant and non-resonant contributions. The efficiencies ($\mathcal{E}$) do not include branching fractions to the observed final states.

3.1.2 Observation of an Excited Ξ_c State

There are two Ξ_c states in which the sq diquark is in an $S = 1$ state — the Ξ_c' and Ξ_c^* with $J^P = \frac{1}{2}^+$ and $\frac{3}{2}^+$, respectively. The Ξ_c' is predicted to be below threshold

†It will be standard throughout this paper that the first error given on a result is statistical and the second is the systematic error. Also, when a hadron's charge is given, the charge conjugate hadron (or decay chain) is implicitly included unless otherwise stated.

for decaying to $\Xi_c\pi$ and so decays radiatively. The Ξ_c^* was expected to be just above threshold for the pionic decay giving hope that the width would be rather narrow, à la $D^* \to D\pi$. The signals for the Ξ_c^+ decay channels chosen for the Ξ_c^* search are shown in Fig. 6. These decay channels were selected based on the

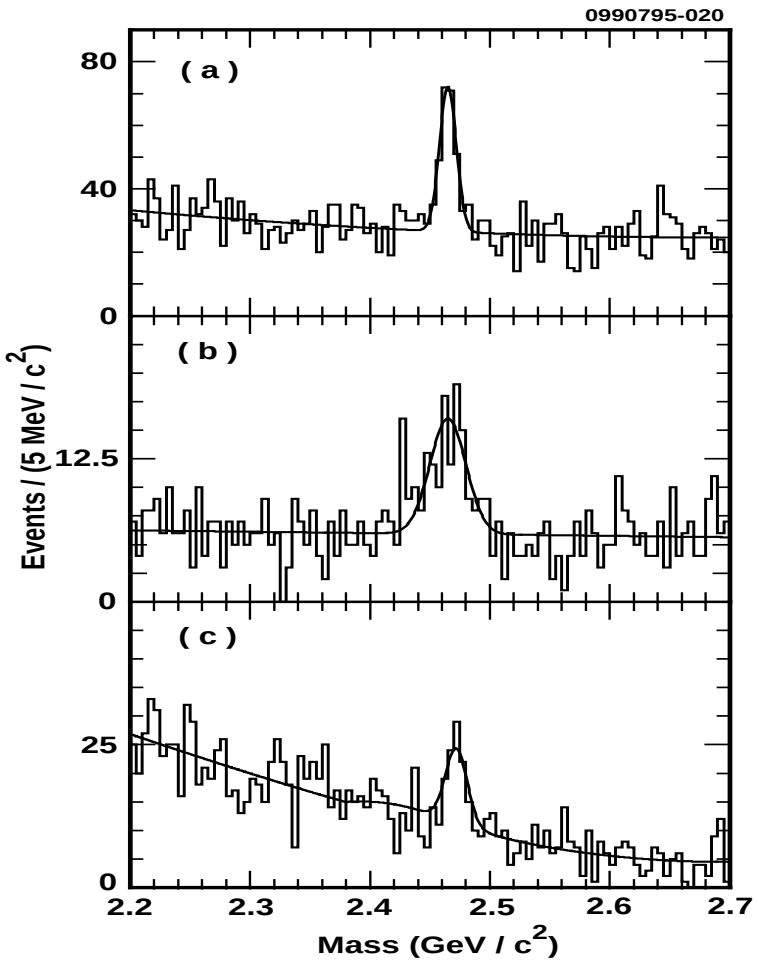


Figure 6: The Ξ_c^+ samples used in the Ξ_c^{*0} analysis. The Ξ_c^+ final states are: (a) $\Xi^-\pi^+\pi^+$, (b) $\Xi^0\pi^+\pi^0$, and (c) $\Sigma^+\bar{K}^{*0}$.

desire to have good signal to background before the addition of the extra pion. Note that the "new" $\Sigma^+\bar{K}^{*0}$ mode is one of the channels. Also, the $\Xi^0\pi^+\pi^0$,$\Xi^0 \to \Lambda\pi^0, \Lambda \to p\pi^-$ decay chain is rather tricky since it involves the reconstruction of two detached vertices for the long-lived Ξ^0 and Λ, the first of which is a decay to two neutral particles.

The Ξ_c^+ candidates are then combined with each remaining π^- track and the mass difference, $M(\Xi_c^+\pi^-) - M(\Xi_c^+)$, is calculated. The result is shown in Fig. 7 where a clear peak at threshold is evident. The signal function used to fit the

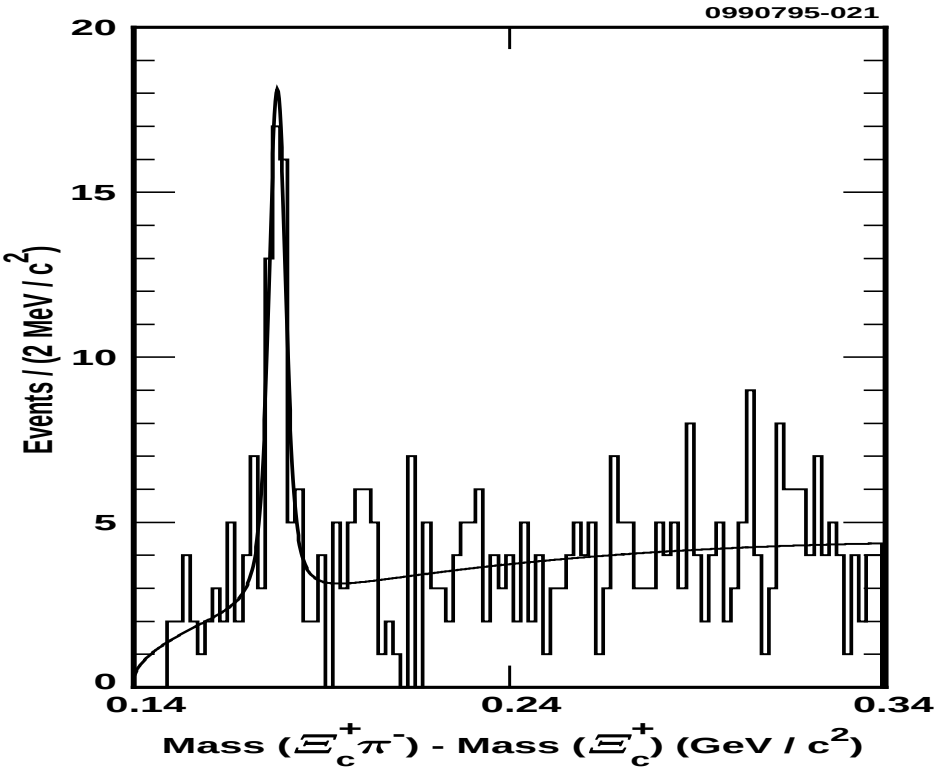


Figure 7: The $M(\Xi_c^+\pi^-) - M(\Xi_c^+)$ data (histogram) and fit.

peak consisted of a Breit-Wigner convoluted with a Gaussian resolution function ($\sigma = 1.6$ MeV). The mass difference is found to be $178.2 \pm 0.5 \pm 1.0$ MeV, and the 90% confidence level upper limit on the width is 5.5 MeV. The mass and width of this new state are consistent with the theoretical predictions for the $J^P = \frac{3}{2}^+$ Ξ_c^{*0}. It is found that a rather large fraction, $(27 \pm 6 \pm 6)\%$, of Ξ_c^+ baryons come from Ξ_c^{*0} decay.

3.2 Charmed Mesons — The Year of the D_s^+

This year saw a number of new and updated D_s^+ results from CLEO. A few of them are summarized in this section.

3.2.1 Observation of the Isospin-Violating Decay $D_s^{*+} \to D_s^+\pi^0$

The D_s^{*+} has been observed exclusively in the $D_s^+\gamma$ final state. The $D_s^{*+} \to D_s^+\pi^0$ decay chain is kinematically allowed but is forbidden by isospin conservation. It was recently suggested by Cho and Wise[8] that the isospin-violating decay could occur via the scenario where the D_s^{*+} decays to a D_s^+ and a virtual η (through

its $s\bar{s}$ component) which then mixes into a π^0. This is illustrated in Fig. 8. The

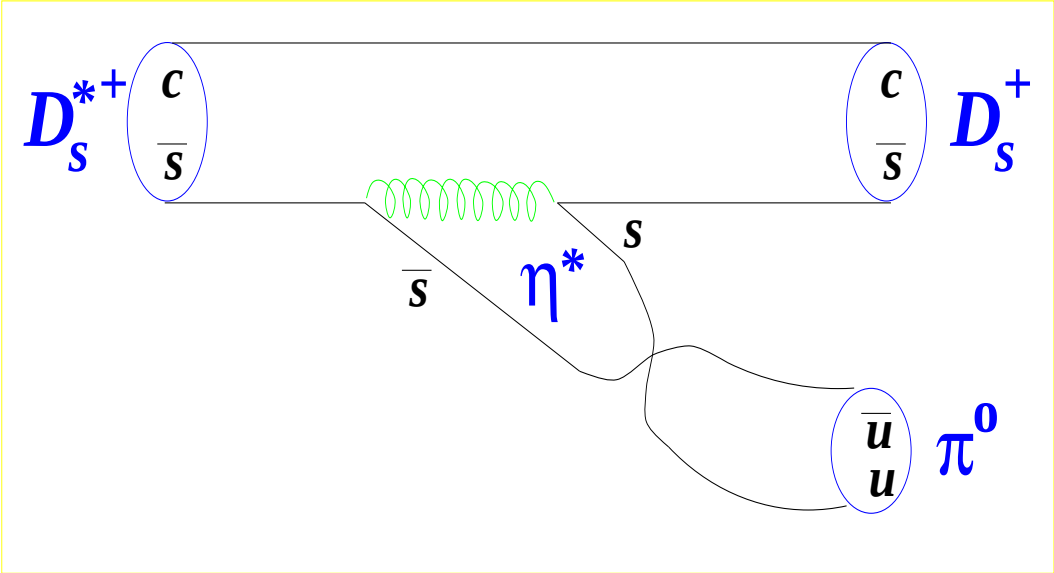


Figure 8: The diagram describing $D_s^{*+} \to D_s^+\pi^0$ decay.

decay amplitude is proportional to the light quark masses in the combination $(m_d - m_u)/[m_s - (m_d + m_u)/2]$, which is zero in the limit of equal u and d quark masses (i.e., absolute isospin conservation). The prediction from Ref. 8 is that $R_0 \equiv \Gamma(D_s^{*+} \to D_s^+\pi^0)/\Gamma(D_s^{*+} \to D_s^+\gamma) =\sim 0.01 - 0.10$.

The CLEO analysis of this decay mode starts with the "standard" $D_s^+ \to \phi\pi^+, \phi \to K^+K^-$ decay chain which offers the best efficiency and signal-to-background (mostly because of the narrowness of the ϕ) of the many D_s^+ decay channels. The resulting peak after combining the D_s^+ candidates with each π^0 in the event is shown in Fig. 9. The significance of the peak, which contains $14.7^{+4.6}_{-4.0}$ events, is found to be greater than 3.9 standard deviations.

The major background is combinatorics since there are many random π^0's in an event. Two "physics" backgrounds were found to be negligible. There is no contribution to the D_s^+ mass region from the $D^{*+} \to D^+\pi^0, D^+ \to K^-\pi^+\pi^+$ decay chain, where one of the pions is misidentified as a kaon, because of the requirement that the K^- and "false" K^+ reconstruct at the ϕ mass. Also, the possible background from $D_s^{*+} \to D_s^+\gamma$ plus random photons was found to be negligible, both from a Monte Carlo study and in the data using $D_s^{*+} \to D_s^+\gamma$ events.

The resulting value for R_0 is $0.062^{+0.020}_{-0.018} \pm 0.22$. If it is assumed that the two branching fractions sum to one, then the individual branching fractions are $\mathcal{B}(D_s^{*+} \to D_s^+\pi^0) = 0.058^{+0.018}_{-0.016} \pm 0.020$ and $\mathcal{B}(D_s^{*+} \to D_s^+\gamma) = 0.942^{+0.016}_{-0.018} \pm 0.020$. The width of the signal is entirely consistent with being due to detector resolution.

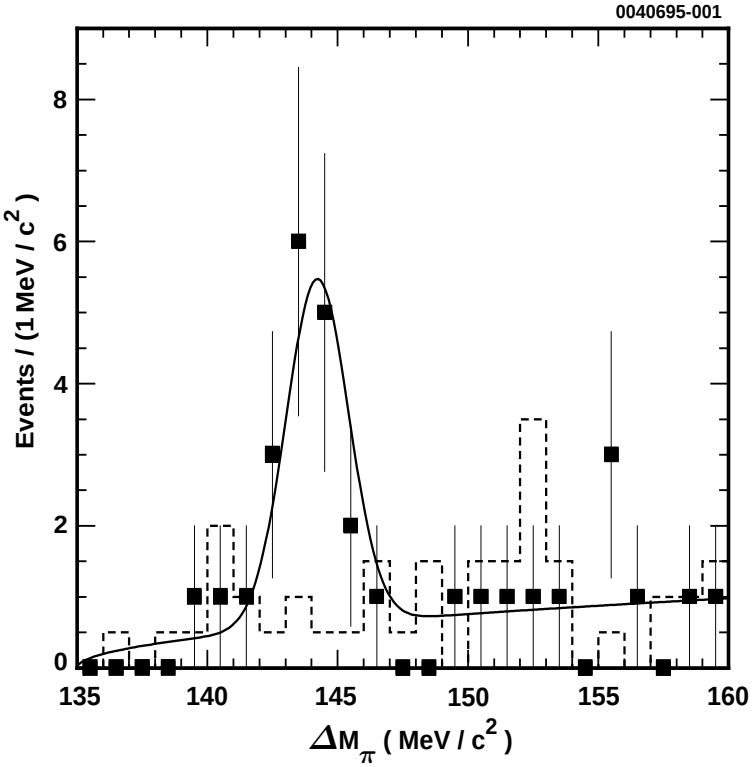


Figure 9: The $\Delta M_\pi \equiv M(D_s^+\pi^0) - M(D_s^+)$ mass spectrum. The points are the data after all cuts, the solid line is the fit to the data, and the dashed line is an estimate of the background using D_s^+ and π^0 mass sidebands.

There are two other interesting results which follow from the observation of this isospin-violating decay mode. Conservation of spin and parity in the strong decay of the D_s^{*+} to two pseudoscalars means it must have natural spin-parity (0^+, 1^-, 2^+ ...) and the radiative decay rules out $0^\pm$. Thus, the most likely spin-parity assignment is $J^P = 1^-$, the same as for the D^{*+} and D^{*0}. Finally, even though this result is statistics limited, the mass difference resolution for the hadronic channel is so much better than for the radiative decay that the value for the mass difference measured using this decay is as precise as the published number from CLEO of $M(D_s^{*+}) - M(D_s^+) = 144.22 \pm 0.47 \pm 0.37$ MeV using the $D_s^{*+} \to D_s^+\gamma$ channel. The

result for the hadronic channel is $M(D_s^{*+}) - M(D_s^+) = 143.76 \pm 0.39 \pm 0.40$ MeV, and the average of the two (they are statistically and systematically uncorrelated) is 143.97 ± 0.41 MeV.

3.2.2 Update on $D_s^+ \to \mu^+\nu$ and f_{D_s}, the D_s^+ Decay Constant

Decay constants are a measure of the nonperturbative physics associated with quarks binding into mesons and are a source of great activity for those doing Lattice Gauge, QCD Sum Rules, and Quark Model calculations. Decay constants are important because they are often the largest source of uncertainty in extracting parameters of the Standard Model from measurements. For example, in $B^0 - \bar{B}^0$ mixing, the mixing parameter is given by:

$$x_d = \Delta M/\Gamma \propto V_{tb}^2 V_{td}^2 f_B^2 B_B m_t^2 F(m_t/m_W)^2$$

where F is a slowly varying function of m_t/m_W. With the present precision on m_t, the largest source of uncertainty in the extraction of V_{td} from measurements of x_d is the product of f_B, the B decay constant, and $\sqrt{B_B}$, where B_B is a parameter describing the degree to which the box diagrams dominate mixing. As another example, a calculation of the expected rate for the decay $B \to D^{*+}D^{*-}$, which is a mode with similar "$\mathcal{CP}$ reach" to the famous ψK final state,[9] requires knowledge of the D meson decay constant, f_D.

To see why there have been measurements of the D_s^+ decay constant, f_{D_s}, and not of f_D or f_B, consider the decay rate for the weak annihilation of a $Q\bar{q}$ pseudoscalar meson, M, into $\ell\nu$.

$$\Gamma(M^+ \to \ell^+\nu) = \frac{1}{8\pi} G_F^2 f_M^2 m_\ell^2 M_M \Big(1 - \frac{m_\ell^2}{M_M^2}\Big)^2 |V_{Qq}|^2 \qquad (2)$$

where M_M and m_ℓ are the masses of the meson and lepton, respectively, f_M is the pseudoscalar decay constant, and V_{Qq} is the relevant CKM matrix element. Helicity suppression is evident in the factor of m_ℓ^2. The B^+ annihilation rates are predicted to be small (with a branching fraction of $\sim 10^{-5}$ for the least helicity suppressed channel $B^+ \to \tau^+\nu$) because the relevant CKM matrix element is V_{ub}. The charm annihilation rates are not small but the D_s^+ rates are Cabibbo favored over the D^+ (i.e., by about $|V_{cs}/V_{cd}|^2$) making the D_s^+ leptonic decay the best bet experimentally.

Last year CLEO published[10] a measurement of $\Gamma(D_s^+ \to \mu^+\nu)/\Gamma(D_s^+ \to \phi\pi^+)$. The basic technique involved using the $D_s^{*+} \to D_s^+\gamma$ decay chain and the missing

momentum and energy in an event to calculate the neutrino's momentum. A key point to this analysis is that the $\mu\nu$ channel dominates over the $e\nu$ channel because of helicity suppression, but the backgrounds are essentially independent of lepton flavor. Hence, the analysis is performed for both electrons and muons, and whatever remains in the electron analysis is directly subtracted from the $\mu\nu$ sample. The signal is then seen as a peak in the mass difference plot. The results for the updated measurement are shown in Fig. 10. The primary differences

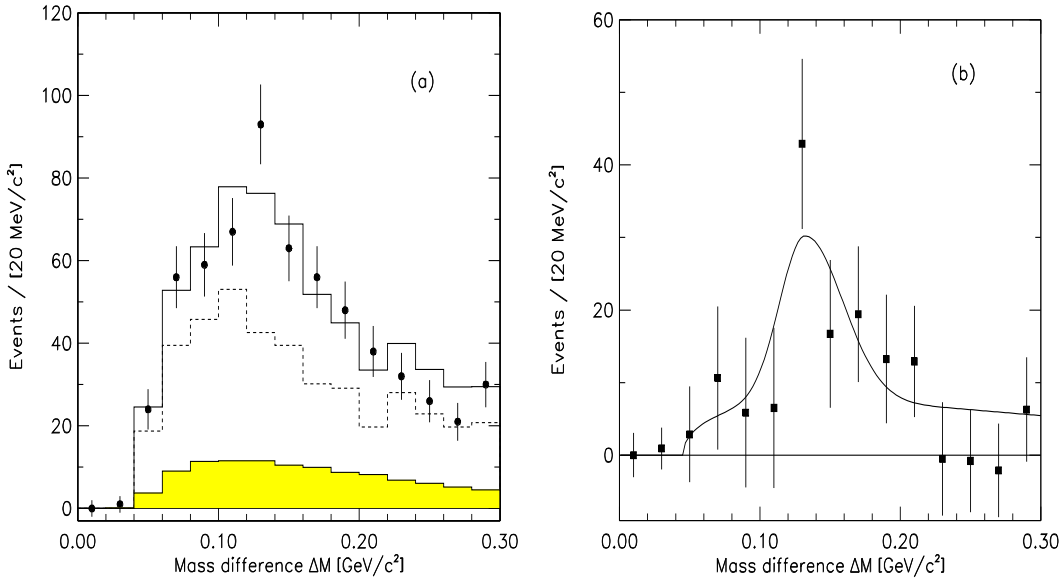


Figure 10: The $\Delta M \equiv M(D_s^+\gamma) - M(D_s^+)$ distributions in the $D_s^+ \to \mu^+\nu$ analysis. Figure (a) includes the muon data (solid points), the electron data (dashed histogram) and excess muon fakes over lepton fakes (shaded histogram). The solid histogram is the result of a fit. Figure (b) shows the distribution after electron and fake subtraction and the resultant fit.

between this update and the published result are the addition of almost 50% more data and much improved measurements of the probability that a hadron is misidentified as a lepton (i.e., the lepton fake rates).

The result is $\Gamma(D_s^+ \to \mu^+\nu)/\Gamma(D_s^+ \to \phi\pi^+) = 0.184 \pm 0.038 \pm 0.038$. Using Eq. (2) along with $\mathcal{B}(D_s^+ \to \phi\pi^+) = (3.5 \pm 0.4)\%$ and $\tau_{D_s} = (4.67 \pm 0.17) \times 10^{-13}$ s (from Ref. 11) gives $f_{D_s} = 284 \pm 30 \pm 30 \pm 16$ MeV. The hope is that f_{D_s} can be utilized to calibrate the various theoretical techniques being used to calculate f_B and f_D. A comparison of this result with the theoretical predictions is given in the next section.

3.2.3 A Detour into B Physics: f_{D_s} from $B \to D_s^{(*)+}\bar{D}^{(*)}$ Decays

In the dominant process leading to two-body decays of the type[‡] $B \to D_s^{(*)+}D^{(*)}$, shown in Fig. 11, the $D_s^{(*)+}$ is produced from the fragmentation of the W^+.

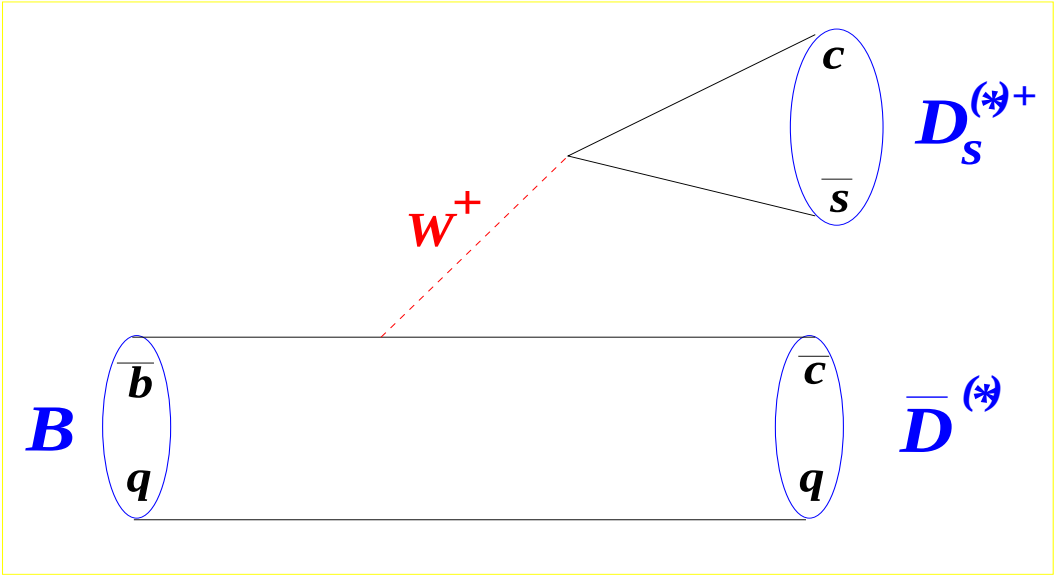


Figure 11: The spectator diagram for $B \to D_s^{(*)+}\bar{D}^{(*)}$ decay.

Assuming that the decay products of the W^+ do not interact with the final states formed at the $b \to c$ vertex (in analogy to semileptonic decays), then the amplitude for these decays can be "factorized" into a product of hadronic currents. The branching fraction for $B \to D_s^+\bar{D}$ decays, for example, is then given by:

$$\mathcal{B}(B \to D_s^+\bar{D}) = KG_F^2 a_1^2 \left|V_{cb}V_{cs}^*\right|^2 f_{D_s}^2 F^2(q^2 = m_{D_s}^2)\tau_B$$

where K is a kinematic factor, and the form factor, F, which is a measure of the probability that the $\bar{c}$ and q quarks will bind to form a meson, is a function of q^2 (= mass^2 of the virtual W). The factorization parameter a_1, which is essentially a QCD correction factor, is the relevant one for external spectator decays like Fig. 11 and can, in principle, be extracted from the measured $B \to \bar{D}\pi$ branching fractions.

[‡]In $B \to D_s^{(*)+}\bar{D}^{(*)}$, $\bar{D}$ is a generic representation of the $\bar{c}q$ mesons, the D^- and $\bar{D}^0$, while the symbol $(*)$ implies that the branching fraction for both the nonexcited and excited states of the meson were separately measured.

The value of f_{D_s} measured from $D_s^+ \to \mu^+\nu$ decays could be used to test the factorization hypothesis in hadronic B decays. Conversely, factorization can be assumed and the measured nonleptonic rates used to extract f_{D_s}, as will be done here. The $B^+ \to D_s^{(*)+}\bar{D}^{(*)0}$ beam-constrained mass plots from CLEO[12] are shown in Fig. 12. The four B^0 modes are also measured and the "average" branching fractions[§] are found to be:

$$\begin{aligned}
\mathcal{B}(B \to D_s^+ \bar{D}) &= (1.10 \pm 0.17 \pm 0.28 \pm 0.13)\% \\
\mathcal{B}(B \to D_s^{*+} \bar{D}) &= (0.89 \pm 0.21 \pm 0.20 \pm 0.10)\% \\
\mathcal{B}(B \to D_s^+ \bar{D}^*) &= (1.12 \pm 0.21 \pm 0.26 \pm 0.13)\% \\
\mathcal{B}(B \to D_s^{*+} \bar{D}^*) &= (2.41 \pm 0.45 \pm 0.51 \pm 0.29)\%
\end{aligned}$$

The D_s^+ decay constant can be extracted from these branching fractions through the use of ratios in which many of the experimental and theoretical errors cancel. Experimental systematic errors are reduced by using CLEO numbers for both the numerator and denominator. The ratio

$$\frac{\Gamma(B \to \bar{D}^* D_s^+)}{d\Gamma(B \to \bar{D}^* e^+ \nu_e)/dq^2|_{q^2=m_{D_s}^2}} = 6\pi^2 a_1^2 \delta f_{D_s}^2 |V_{cs}|^2$$

(δ is calculable) has the advantage that the uncertainty in the form factor cancels although a_1 is still present. Conversely, comparing the hadronic rates where the W^+ decays to $u\bar{d}$ as opposed to $c\bar{s}$ gives, for example,

$$\frac{\Gamma(B \to \bar{D} D_s^+)}{\Gamma(B^0 \to D^- \pi^+)} = K \frac{f_{D_s}^2}{f_\pi^2} \frac{|V_{cs}|^2}{|V_{ud}|^2} \frac{F(q^2 = m_{D_s}^2)^2}{F(q^2 = m_\pi^2)^2}$$

where K is a calculable kinematic factor. The QCD correction factor a_1 cancels (there is some debate on this as discussed in Ref. 12) but model dependence is introduced through the parameterization of the form factor, F, since it is sampled at a different q^2 in the numerator and denominator.

The CLEO results are given in Table 2 along with the theoretical predictions. Both the experimental and theoretical uncertainties need to be reduced before a meaningful comparison can be done. It should be noted, however, that the

[§]That is, $\mathcal{B}(B \to D_s^{(*)+}\bar{D}^{(*)})$ is the average of the $B^0 \to D_s^{(*)+}D^{(*)-}$ and $B^+ \to D_s^{(*)+}D^{(*)0}$ branching fractions. See Ref. 12 for a discussion and justification of this.

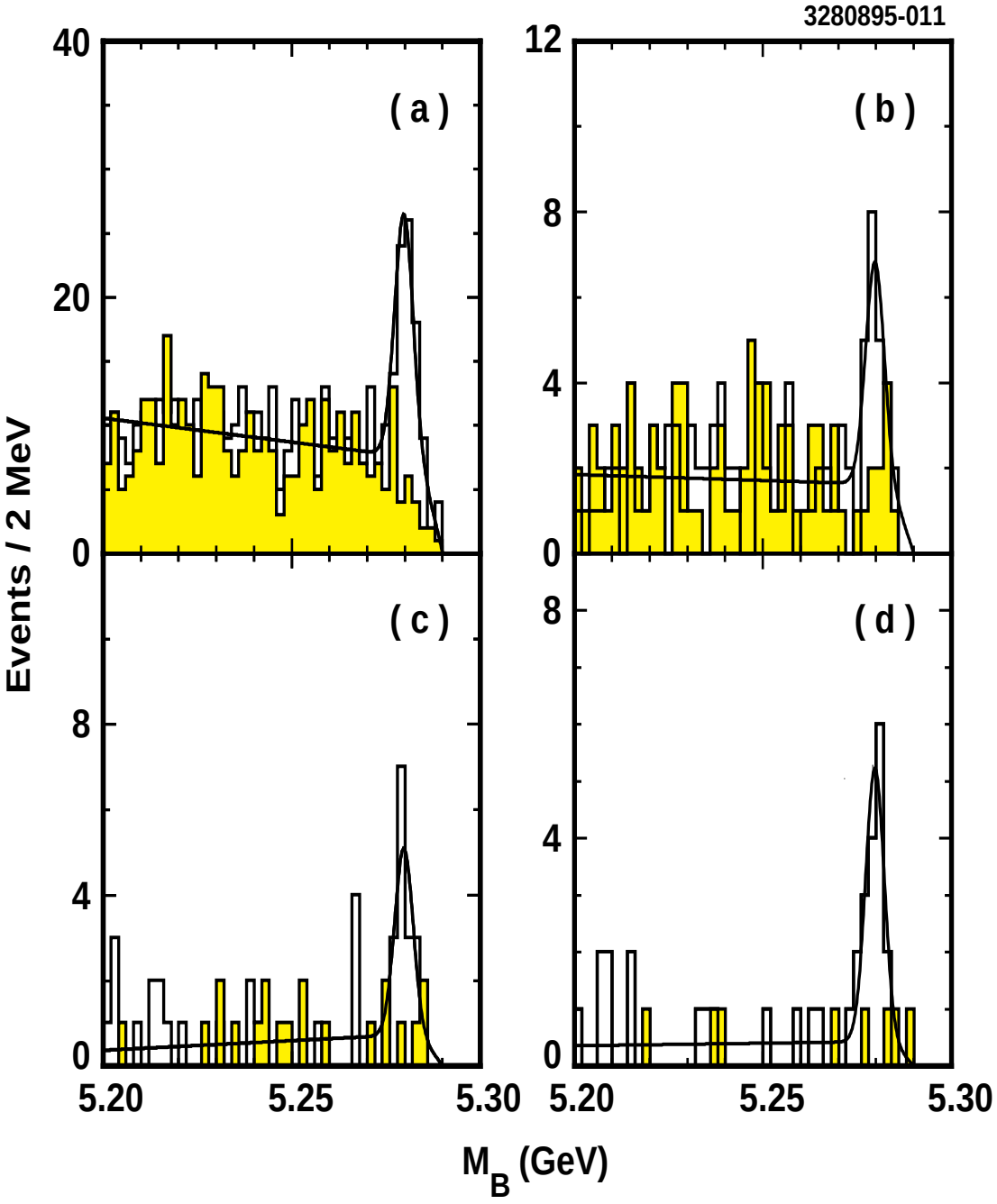


Figure 12: The beam-constrained mass spectra for the B^+ decay modes: (a) $D_s^+\bar{D}^0$, (b) $D_s^{*+}\bar{D}^0$, (c) $D_s^+\bar{D}^{*0}$, and (d) $D_s^{*+}\bar{D}^{*0}$. The solid histogram is the data within the ΔE signal window ($|\Delta E| < 25$ MeV) while the filled histogram is the data in ΔE sidebands. The curves are the results of fits where the fitting function consists of a Gaussian to describe the signal and a background function which is linear for $M_B < 5.282$ GeV and parabolic, with a kinematic cutoff, for larger values of M_B.

CLEO II Results	f_{D_s} (MeV)
$D_s^+ \to \mu^+\nu_\mu$	$284 \pm 30 \pm 30$
$\Gamma(B \to D_s^{(*)+}\bar{D}^*)/d\Gamma(B \to \bar{D}^* e^+\nu_e)/dq^2\|_{q^2=m^2_{D_s^{(*)}}}$	$281 \pm 22 \pm 32$
$\sum\Gamma(B \to D_s^{(*)+}\bar{D}^{(*)})/\sum\Gamma(B^0 \to \pi^+/\rho^+ D^{(*)-})$	$288 \pm 13 \pm 38$
Theoretical Predictions	f_{D_s} (MeV)
Lattice QCD[13]	235 ± 15
QCD Sum Rules[13]	260 ± 50
Quark Models[14]	$\sim 290 \pm 20$

Table 2: Determinations of f_{D_s}. A $\mathcal{B}(D_s^+ \to \phi\pi^+) = (3.5 \pm 0.4)\%$ is common to all experimental results and its uncertainty is not included in the systematic error. A value of $a_1 = 1.07 \pm 0.04 \pm 0.06$ was taken from the CLEO paper[15] on hadronic B decays to final states not including the D_s^+.

theoretical uncertainties in ratios like f_{D_s}/f_D are generally considered to be under better control and, therefore, a reasonably precise value for f_D could be obtained from a measurement of f_{D_s}.

3.2.4 Semileptonic D_s^+ Decays

Semileptonic decays are particularly simple to treat theoretically because there is only the one diagram and there are no final state interactions.[16] In b quark physics, a precise value for V_{cb} can be arrived at through measurements of $B \to D^*\ell\nu$ decays and the inclusive lepton momentum spectrum. This is possible because of theoretical advances in understanding the form factors in such heavy-to-heavy (i.e., $b \to c$) transitions. However, such techniques are not applicable in heavy-to-light transitions like $b \to u$, thereby complicating the extraction of V_{ub}. Charmed semileptonic decays could be the key to this problem because the CKM matrix elements, V_{cs} and V_{cd}, are known, allowing the heavy-to-light form factors to be probed experimentally. It was perhaps somewhat surprising, given the initial assumptions about the simplicity of semileptonic decays, that the original calculations of $\Gamma(D \to \bar{K}^*\ell\nu)/\Gamma(D \to \bar{K}\ell\nu)$ were about a factor of two higher than the experimental value of 0.56 ± 0.05. Some postdictions do a better job, but it would be very interesting to measure this same vector-to-pseudoscalar ratio for the D_s^+,

the "other" charmed meson system.

CLEO has published results on $D_s^+ \to \phi\ell\nu$ decays.[17] The CLEO II detector is well suited for observing $D_s^+ \to \eta\ell\nu$ decays because of the power of the CsI calorimeter in reconstructing the two photons from the $\eta \to \gamma\gamma$ decay chain. The η' is reconstructed through its decay to $\eta\pi^+\pi^-$. The analysis uses both electrons and muons, although the muon results are corrected for phase space and the final results are then given as semielectronic branching fractions. There are sufficient statistics in the $\phi\ell\nu$ and $\eta\ell\nu$ samples to also perform the analysis by using the $D_s^{*+} \to D_s^+\gamma$ mode (the so-called "tagged" analysis). This requirement reduces the backgrounds and allows for the minimum electron momentum to be lowered to 0.7 GeV from the 1 GeV requirement used in the "untagged" analysis. There is actually little correlation between the tagged and untagged results, and they are combined in the $\eta\ell\nu$ case. The mass plots for the untagged analysis are shown in Fig. 13. The results are given in Table 3. Many of the systematic errors cancel in the ratios.

$\Gamma(D_s^+ \to \eta e^+\nu)/\Gamma(D_s^+ \to \phi e^+\nu)$	$1.24 \pm 0.12 \pm 0.15$
$\Gamma(D_s^+ \to \eta' e^+\nu)/\Gamma(D_s^+ \to \phi e^+\nu)$	$0.43 \pm 0.11 \pm 0.07$
$\Gamma(D_s^+ \to \eta' e^+\nu)/\Gamma(D_s^+ \to \eta e^+\nu)$	$0.35 \pm 0.09 \pm 0.07$
$\Gamma(D_s^+ \to \phi e^+\nu)/\Gamma(D_s^+ \to (\eta+\eta') e^+\nu)$	$0.60 \pm 0.06 \pm 0.06$

Table 3: Summary of measurements for $D_s^+ \to X e^+\nu$.

The vector-to-pseudoscalar ratio of $0.60 \pm 0.06 \pm 0.06$ for the D_s^+ agrees nicely with the value for the D meson and with the theoretical predictions. There is some non-negligible dependence in the calculations on the value of the $\eta - \eta'$ mixing angle. The $\Gamma(D_s^+ \to \eta' e^+\nu)/\Gamma(D_s^+ \to \eta e^+\nu)$ ratio is also interesting since it is predicted to be equal to $\Gamma(D_s^+ \to \eta'\rho^+)/\Gamma(D_s^+ \to \eta\rho^+)$ from factorization.[18] However, this ratio of hadronic rates is found by CLEO[19] to be 1.20 ± 0.35, not in very good agreement with the factorization expectation.

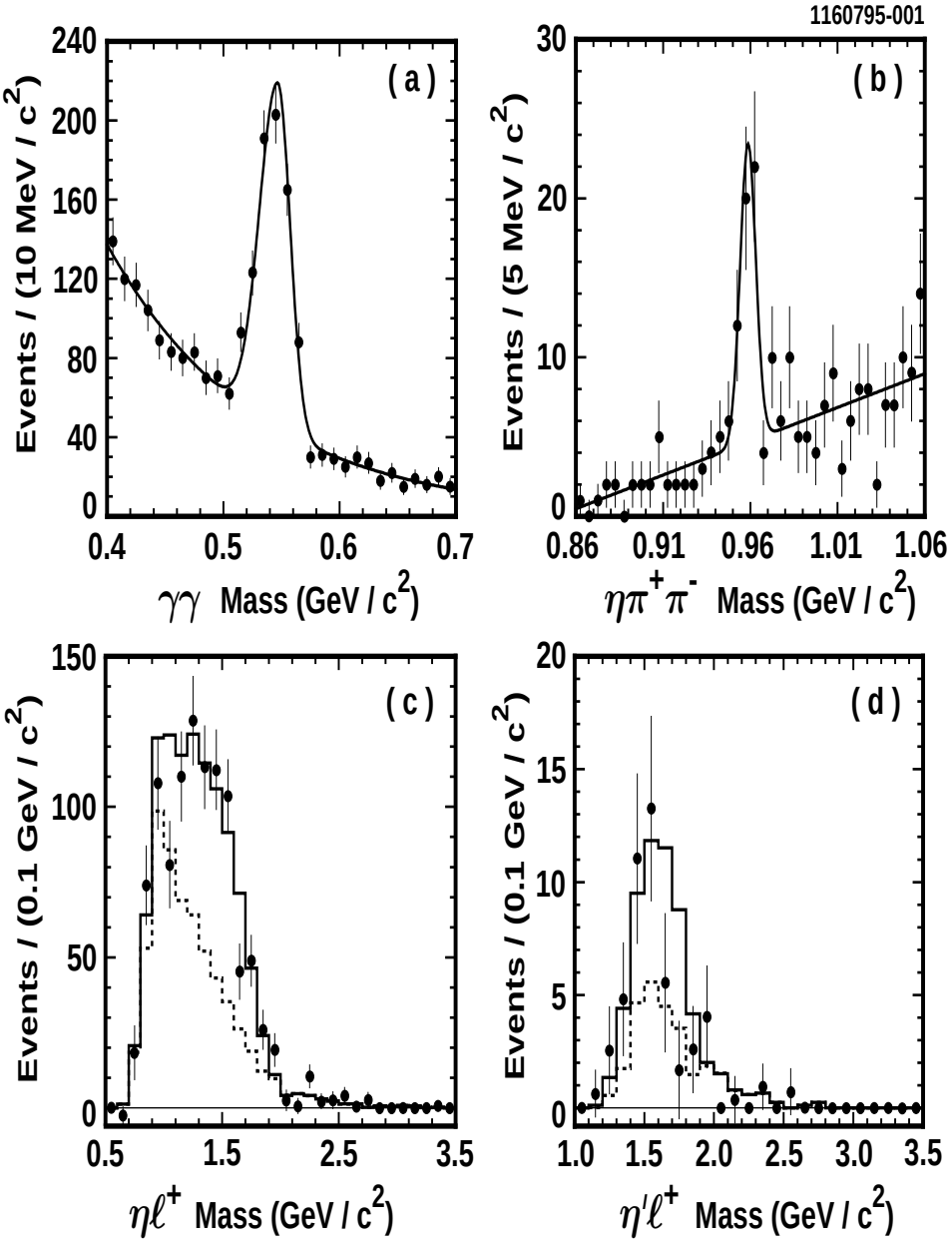


Figure 13: Invariant mass plots for the non-D_s^{*+} tagged $D_s^+ \to \eta\ell\nu$ and $D_s^+ \to \eta'\ell\nu$ analysis. The η and η' signals are shown in (a) and (b), respectively. The solid curves are the fits to each spectrum. The $\eta\ell^+$ and $\eta'\ell^+$ mass spectra are shown in (c) and (d), respectively. The solid histogram shows the simulated signal plus the predicted background while the dotted histogram is the predicted background alone.

4 Bottom Mesons

The basic B meson decay diagrams are given in Fig. 15. The spectator diagram dominates the rate although the color suppressed (or, internal W decay) diagram is a significant player as evidenced by the large $B \to \psi^{(\prime)} X$ branching fractions.[20] Neither the annihilation nor the W-exchange process has been observed, and their total rates are expected to be small. CLEO observed the inclusive branching fraction for the radiative penguin to be around 2×10^{-4} while there is no unambiguous evidence for gluonic penguin decays (see Sec. 4.2.1). There is some question as to whether even the dominant decay processes are understood, as will be discussed in the next session. The following sections will be concerned with rare decays — in particular, a search for gluonic penguin decays and the first observation of an exclusive semileptonic $b \to u$ decay channel.

4.1 $b \to c$ Decays and the Charm Count

Assuming that the diagram of Fig. 14 dominates[¶] and that the $b \to u$ contribution

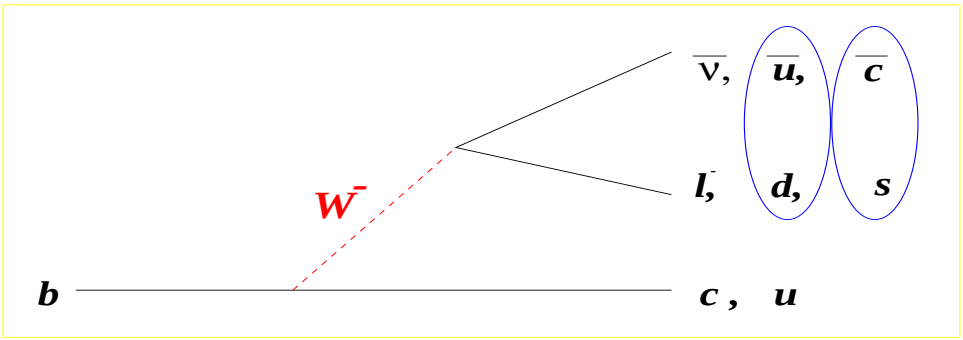


Figure 14: The dominant b quark decay diagram.

is negligible, then the total B meson decay rate, Γ, is given by:

$$\Gamma = \Gamma(b \to c\ell^-\bar{\nu}) + \Gamma(b \to c\bar{u}d) + \Gamma(b \to c\bar{c}s).$$

In terms of branching fractions, this becomes:

$$\mathcal{B}(B \to X\ell\nu) = 1 - \mathcal{B}(b \to c\bar{u}d) - \mathcal{B}(b \to c\bar{c}s). \quad (3)$$

[¶]The Cabibbo-suppressed channels like $c\bar{u}s$ are implicitly included here as are the internal W decay channels.

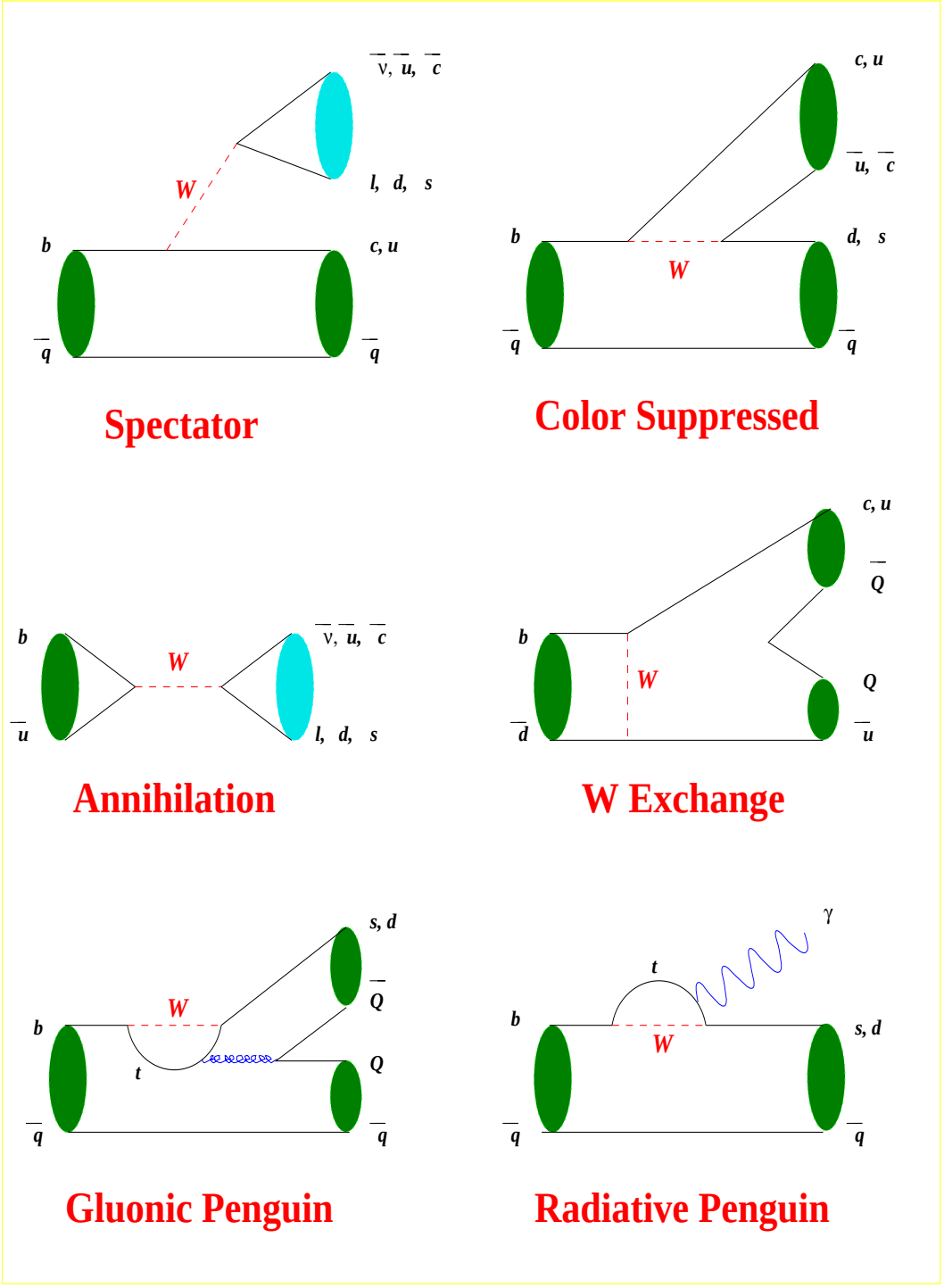


Figure 15: B meson decay diagrams.

If all lepton and quark masses were small relative to m_B, and there were no complications due to strong interactions, then $\mathcal{B}(B \to Xe^+\nu) \equiv \mathcal{B}_e$ would be 1/9. Phase space corrections due to the large τ^+ lepton and charm quark masses raises $\mathcal{B}_e$ to about 17%. Bigi and collaborators claimed[21] that, even including all QCD corrections, the lower limit on the semielectronic B branching fraction is 12.5%. Yet measurements from CLEO, ARGUS, and LEP consistently find $\mathcal{B}_e$ to be less than 11%, hence the moniker "baffling" for this branching fraction.

The measurements have predominantly come from fitting the momentum spectrum of leptons from B decay. Part of the difficulty in extracting $\mathcal{B}_e$ from the inclusive lepton momentum spectrum is that leptons in B meson decays come not only from the primary $b \to c\ell\nu$ decay mechanism but also from the subsequent semileptonic decay of the charm quark. Thus a large model dependence is introduced when trying to extract $\mathcal{B}_e$ since fitting the spectrum requires a functional form for both components.

A way around this difficulty is to tag the flavor of the decaying b quark. In the B meson rest frame (essentially the lab frame for CLEO), the leptons from the secondary charm quark decay are generally "soft", with only about 3% having momentum greater than 1.4 GeV. In $\Upsilon(4S)$ decays there are no other particles produced along with the B and $\bar{B}$ mesons. Therefore, using a high momentum lepton to tag the charge of one of them essentially tags the charge of both (with a calculable correction due to mixing in $B^0\bar{B}^0$ events). This is illustrated in Fig. 16. In this example, the ℓ^+ tag indicates that a $\bar{b}$ decayed semileptonically. If there

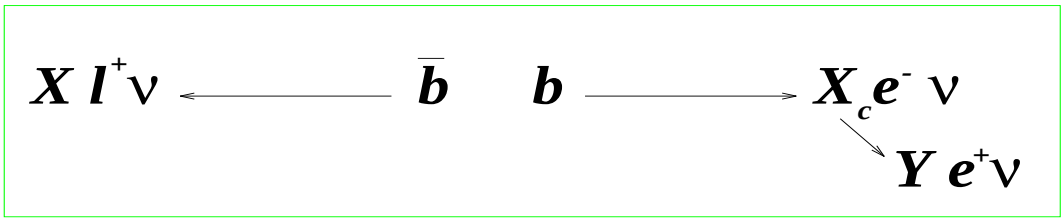


Figure 16: Charge correlations in the lepton tagged measurement of the semielectronic B branching fraction.

is a second lepton in the event, then this lepton's charge distinguishes whether it came from the decay of the b quark or from the semileptonic decay of the charmed hadron produced in the weak $b \to c$ decay.

This technique was first proposed by ARGUS, and CLEO added a momentum-dependent cut on the angle between the tag lepton and the second lepton to

eliminate events where the two leptons come from the same B. Both electrons and muons are used by CLEO as tags. Only electrons are used as the second lepton since the detector is efficient down to electron momenta of 600 MeV whereas the muon fake rates get large for muons of momentum below 1.4 GeV. Using this technique allows for the separation of the primary and cascade lepton momentum spectra, as shown in Fig. 17. The $B \to Xe^+\nu$ branching fraction is found to be $(10.49 \pm 0.17 \pm 0.43)\%$. Some small model dependence comes in when accounting for the undetected fraction of the spectrum below 600 MeV. This was estimated to comprise $(6.1 \pm 0.5)\%$ of the total.

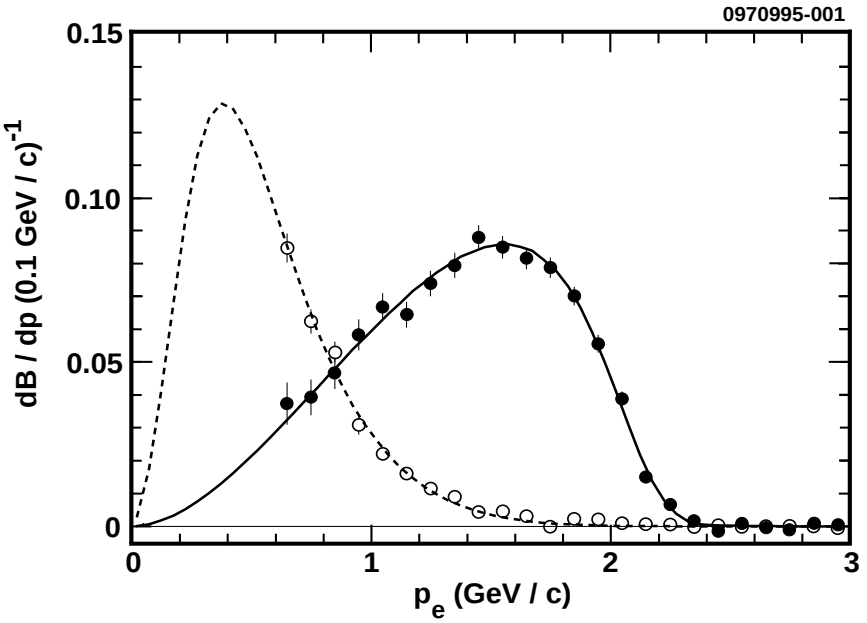


Figure 17: The spectra of electrons from $B \to Xe\nu$ (filled circles) and $b \to c \to Ye\nu$ (open circles) after continuum and fake subtraction, and the mixing correction. The curve is an example of one of the fits used to estimate the uncertainty in the extrapolation from 600 MeV to the origin.

This model-independent measurement strongly supports the conclusion that the semielectronic B branching fraction is significantly below 12.5%. It is clear from Eq. (3) that for $\mathcal{B}(B \to X\ell\nu)$ to go down, $\mathcal{B}(b \to c\bar{u}d) + \mathcal{B}(b \to c\bar{c}s)$ must go up. Recent calculations[22] have shown that higher order perturbative corrections lower the value of $\mathcal{B}_e$ considerably but at the price of simultaneously raising the

contribution from $\mathcal{B}(b \to c\bar{c}s)$ decays. Since this final state contains two charm quarks, this prediction has the experimentally observable result of raising n_c, the number of charm quarks per B decay, from the value of around 1.15 associated with the larger predicted $\mathcal{B}_e$ to > 1.3. The number of charm quarks per B decay can be calculated using the measured inclusive branching fractions. Naively this is given by:

$$\begin{aligned} n_c &= \mathcal{B}(B \to D^0 X) + \mathcal{B}(B \to D^+ X) + \mathcal{B}(B \to D_s X) \\ &+ \mathcal{B}(B \to \Lambda_c X) + \mathcal{B}(B \to \Xi_c^+ X) + \mathcal{B}(B \to \Xi_c^0 X) \\ &+ 2\mathcal{B}(B \to \psi X) + 2\mathcal{B}(B \to \psi' X) \\ &+ 2\mathcal{B}(B \to \chi_{c1} X) + 2\mathcal{B}(B \to \chi_{c2} X) + 2\mathcal{B}(B \to \eta_c X) \end{aligned} \quad (4)$$

where some assumptions must be made about how baryons and D_s^+ mesons are produced in B decay. CLEO has released new, more precise values for all of these branching fractions in the last year except for the D^0 and D^+ final states where the new measurements should be available in early 1996. The experimental value of n_c is around the 1.15 with an uncertainty of about 0.05. The dominant uncertainties in many of these branching fractions are systematic, some of which are correlated, so it is difficult to get a precise value for the uncertainty in n_c. Even given this, however, the data do not support an enhanced $\mathcal{B}(b \to c\bar{c}s)$ component so $\mathcal{B}_e$ will remain "baffling" for a while yet.

4.2 Rare Hadronic B Decays

CLEO has investigated a slew of rare hadronic B decay channels including the $\pi\pi$, $K\pi$, and KK final states (with both charged and neutral kaons and pions) as well as the $\pi\rho$, $K\rho$, $K^*\pi$, $K\phi$, $K^*\phi$, and $\phi\phi$ final states.[23] There is still no significant signal in any particular channel although the upper limits on the branching fractions for some, e.g. $\pi^+\pi^-$, $\pi^+\pi^0$, and $\pi^+\rho^-$, are encroaching on the theoretical predictions. The beam-constrained mass plots for the $\pi\pi$, $K\pi$, and KK channels are shown in Fig. 18.

The combined $\pi^+\pi^-$ and $K^+\pi^-$ signal is now > 4 standard deviations from 0 although the K/π separation of $\sim 1.8\sigma$ is still not enough to resolve the two states with the present statistics. This is illustrated in Fig. 19 which shows the results of the maximum likelihood fit to the two charged track final state.

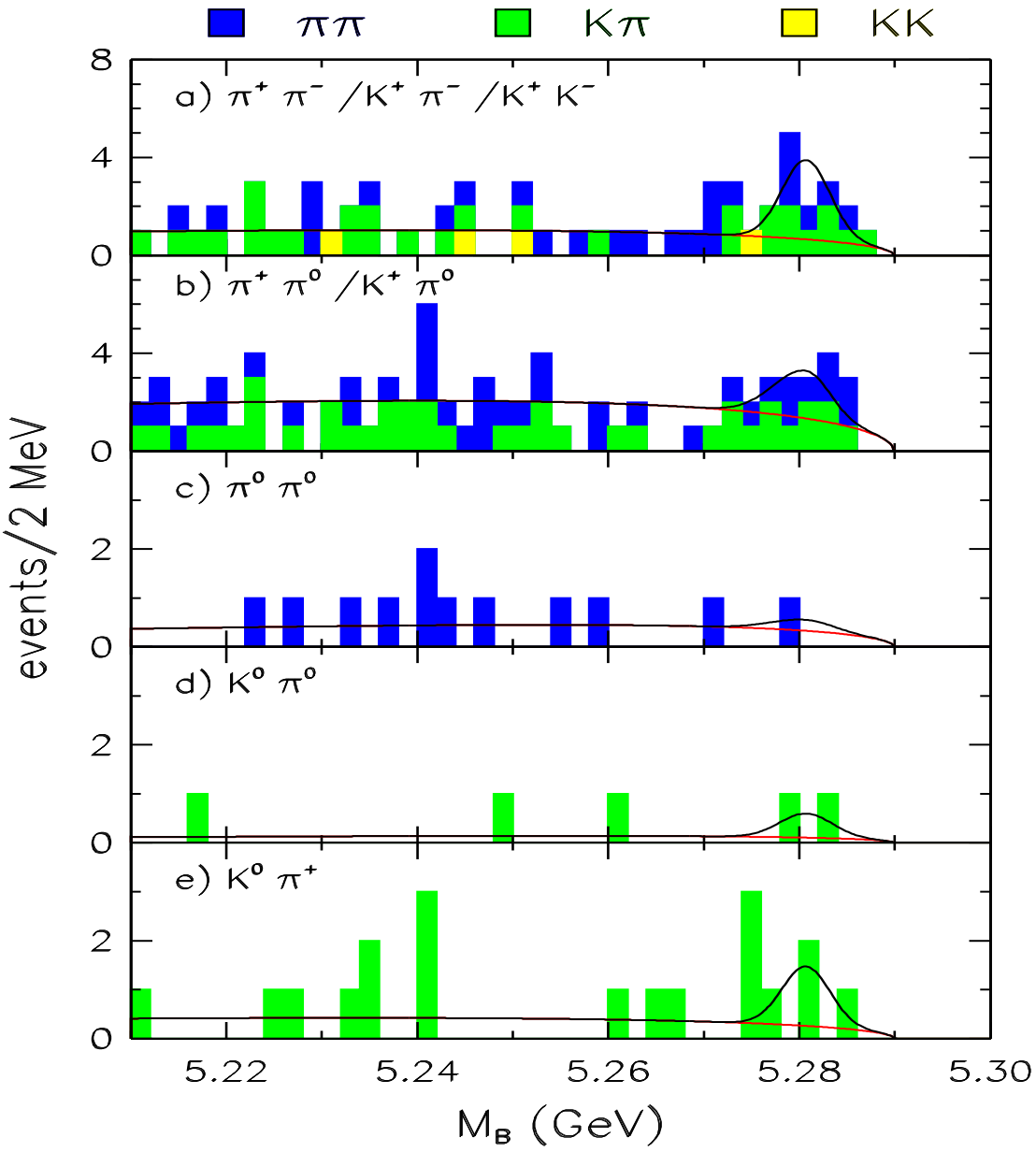


Figure 18: The beam-constrained mass spectra for the $\pi\pi$, $K\pi$, and KK channels. The modes in (a) and (b) are sorted by dE/dx according to the most likely hypothesis.

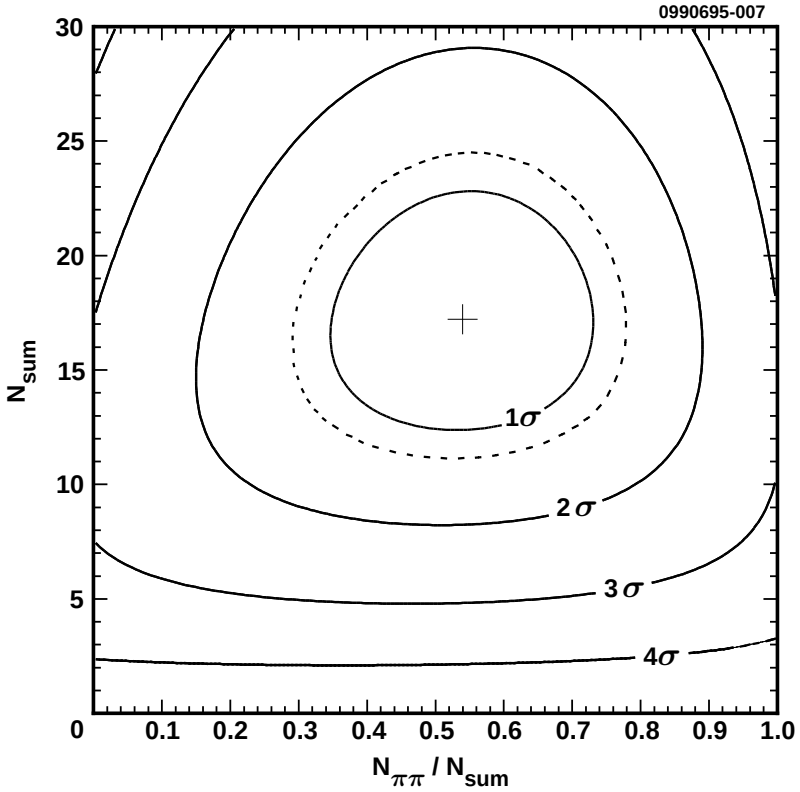


Figure 19: The central value (+) of the likelihood fit to $N_{sum} \equiv N_{\pi\pi} + N_{K\pi}$ and $N_{\pi\pi}/N_{sum}$ for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$. The solid curves are the $n\sigma$ contours and the dotted curve is the 1.28σ contour.

4.2.1 Penguin Pollution and $B \to X_s\phi$

The $B^0 \to \pi^+\pi^-$ decay channel is of great interest because it is one of the modes that can be used at an asymmetric "B Factory" to measure an angle of the unitarity triangle. There is one caveat to this, however, and that is so-called "Penguin Pollution," which is illustrated in Fig. 20. Only for the top diagram in Fig. 20 can an angle of the unitarity triangle be cleanly extracted from measurements of time-dependent decay asymmetries in $B^0 \to \pi^+\pi^-$ decays. If the contribution to the rate from the bottom diagram in Fig. 20, the gluonic penguin diagram, is non-negligible, then a more complicated isospin analysis of the full $\pi\pi$ system must be done to get to the desired information on the CKM matrix.[24]

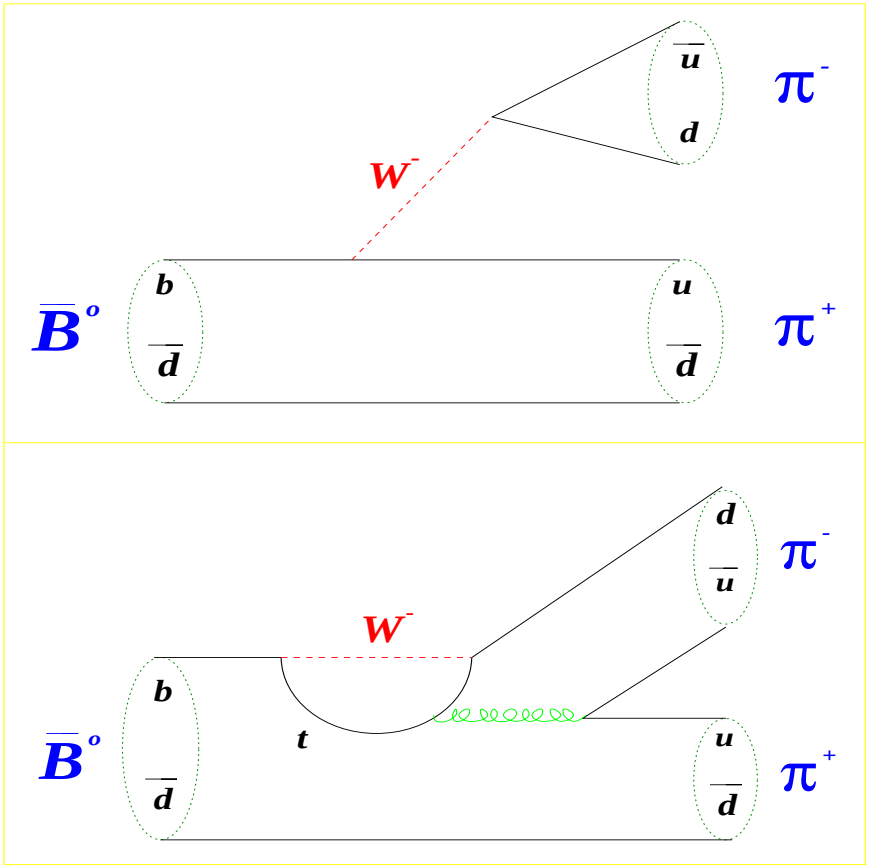


Figure 20: The tree level (top) and gluonic penguin (bottom) diagrams for the $B \to \pi^+\pi^-$ decay channel.

It is extremely difficult to calculate the relative strength of the gluonic penguin in decays like $B^0 \to \pi^+\pi^-$, and so it would be desirable to observe a decay which proceeds solely through this mechanism. Assuming that weak annihilation processes producing the $s\bar{s}s$ final state (e.g., $b\bar{u} \to W^- \to s\bar{u}$ with $s\bar{s}$ popping) are negligible, then $b \to s\bar{s}s$ decay proceeds through a penguin decay process. The gluonic penguin, illustrated in Fig. 21, is expected to dominate over the electromagnetic penguin so the observation of decays of the form $B \to X_s\phi$ (where X_s can be one particle or a system of particles with net strangeness 1) would be an unambiguous signature for the gluonic penguin. Theoretical understanding of such rates would lead to some confidence that the contributions of these diagrams to other processes could be reliably calculated.

The upper limits from CLEO on some exclusive $B \to X_s\phi$ branching fractions are given in Table 4. While these branching fractions indicate that this process is not dominating the rate, there is uncertainty in how the $s\bar{s}s\bar{q}$ final state hadronizes.

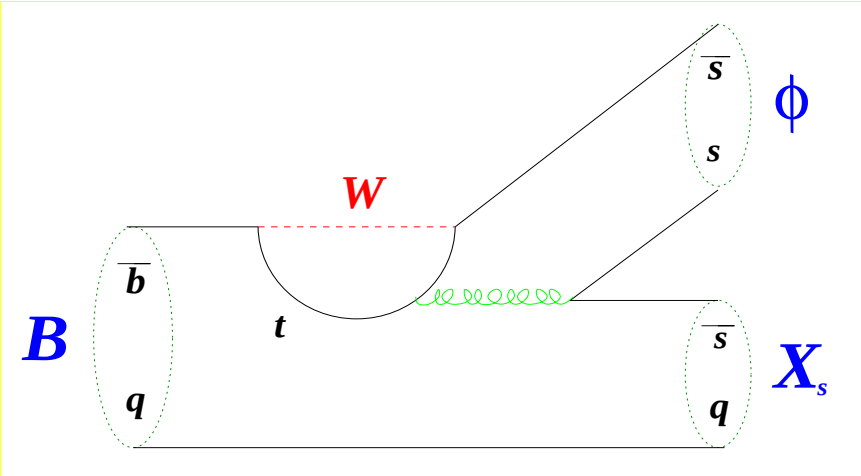


Figure 21: The pure penguin $B \to X_s\phi$ decay.

This motivates a measurement of the inclusive $b \to sg \to s\bar{s}s$ rate since it may be calculable from first principles and should be at least an order of magnitude higher than the rate for any exclusive channel. Some schemes even have very large rates for this, "solving" the baffling semielectronic branching fraction problem without raising n_c (see Sec. 4.1).

Decay Mode	CLEO II Upper Limit
$B^0 \to K^0_S\phi$	$< 4.4 \times 10^{-5}$
$B^+ \to K^+\phi$	$< 1.2 \times 10^{-5}$
$B^0 \to K^{*0}\phi$	$< 4.3 \times 10^{-5}$
$B^+ \to K^{*+}\phi$	$< 7.0 \times 10^{-5}$

Table 4: Upper limits on exclusive $B \to X_s\phi$ branching fractions.

Two techniques are used to search for $B \to X_s\phi$. One method involves searching for ϕ mesons from B decay with momenta beyond the endpoint for having originated from the standard tree level B decays. The continuum-subtracted ϕ momentum spectrum is the left-hand plot in Fig. 22. There is no evidence of an excess in the signal region of $0.4 < x < 0.5$ and the 90% confidence level upper limit is found to be $\mathcal{B}(B \to X_s\phi) < 2.2 \times 10^{-4}$ for $2.0 < p_\phi < 2.6$ GeV. The other technique, the so-called "B Reconstruction" technique, is a slight variation of the method that was found to be very powerful in the measurement of the inclusive $b \to s\gamma$ rate.[25] The basic idea is to combine a K^+K^- pair which reconstruct to

within ± 20 MeV of the ϕ mass, a kaon (either a $K^{\pm}$ or a $K^0_S \to \pi^+\pi^-$), and zero to four pions (with at most one π^0), and to keep the "best" (i.e., most likely B) combination based on beam-constrained mass, ΔE, particle ID, and K^0_S or π^0 mass, if they are used. The goal here is not to fully reconstruct these exclusive final states but to reject continuum background which is much less likely than the signal to satisfy these criteria. The results of this analysis are shown in the right-hand plots in Fig. 22 where again there is no evidence for a signal above the continuum. The 90% confidence level upper limit for this technique is found to be $\mathcal{B}(B \to X_s\phi) < 1.1 \times 10^{-4}$. The two methods are complementary since, even though the first one is not as powerful at suppressing background, it is much less sensitive to the details of how the $s\bar{q}s\bar{s}$ final state hadronizes. It is clear from these measurements that the gluonic penguin rate is not anomalously high although a rate large enough to complicate the extraction of CKM parameters from measurements at an asymmetric B Factory is not ruled out.

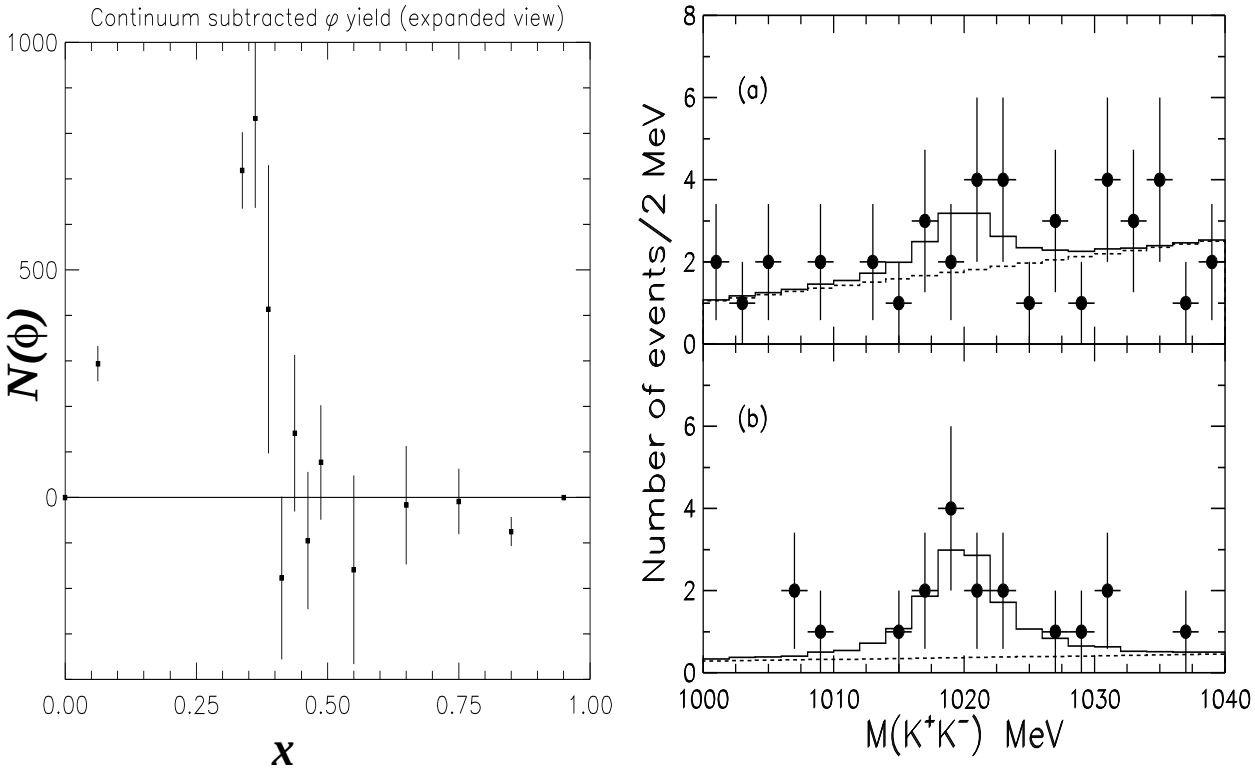


Figure 22: Searches for $B \to X_s\phi$. The left-hand plot is the continuum subtracted ϕ yield as a function of x. The "signal" region is $0.4 < x < 0.5$. The right-hand plots are the K^+K^- invariant mass for the "B Reconstruction" candidates (defined in the text) for (a) on-resonance data and (b) continuum data.

4.3 "Neutrino Reconstruction" and the Observation of Exclusive Semileptonic $b \to u\ell\nu$ Decays

The measurement of V_{ub} is perhaps the most important activity in B physics today. ARGUS and CLEO established that $V_{ub} > 0$ by observing leptons with momentum beyond the endpoint for $b \to c$ transitions.[26,27] However, there are large theoretical uncertainties as to the fraction of the rate encompassed by these measurements thus complicating the extraction of V_{ub}. The ratio $|V_{ub}/V_{cb}|$ is presently known to be between 0.07 and 0.11, with theoretical uncertainties determining the size of this range. There is some hope that measurements of exclusive semileptonic $b \to u\ell\nu$ decays could lead to a more precise determination of $|V_{ub}/V_{cb}|$ because the form factors in such heavy-to-light transitions can be measured experimentally using semileptonic charm decays (see Sec. 3.2.4).

The problem in reconstructing an exclusive $b \to u\ell\nu$ decay is, of course, that the neutrino is not detected. However, the excellent hermiticity of the CLEO II detector (coverage of $\sim 95\%$ and $\sim 98\%$ of 4π for tracks and photons, respectively) allows for the neutrino four-momentum to be "reconstructed" by using the "missing" energy and momentum in an event. Specifically,

$$\begin{aligned} E_\nu = E_{miss} &\equiv 2E_{beam} - \sum_i E_i & (\sigma \sim 260 \text{ MeV}) \\ \vec{p}_\nu = \vec{p}_{miss} &\equiv -\sum_i \vec{p}_i & (\sigma \sim 110 \text{ MeV}) \end{aligned}$$

where the index i runs over all charged tracks and showers in the calorimeter not associated with charged tracks. The resolutions given are for events with no K_L^0 mesons or extra neutrinos. For a real neutrino, $M_{miss}^2 = E_{miss}^2 - |\vec{p}_{miss}|^2$ should be consistent with zero. The criterion $(M_{miss}^2/2E_{miss}) < 300$ MeV is used because the resolution in M_{miss}^2 varies roughly like $2E_{miss}\sigma_{E_{miss}}$. The neutrino energy is set to $|\vec{p}_{miss}|$ because of the better momentum resolution.

Further requirements are then imposed on events to suppress background. Only one charged lepton is allowed per event since another charged lepton immediately implies there being another neutrino. Also, the total charge of an event is required to be zero to ensure that a charged track has not been missed. For a candidate event, this "reconstructed" neutrino can then be used to fully reconstruct a B meson, and the standard beam-constrained mass and ΔE variables can be examined.

The method outlined to this point could be used to reconstruct any exclusive semileptonic B decay mode. It is useful to perform the analysis on a $b \to c\ell\nu$ channel that has been well measured using the standard missing mass techniques as a test of the "neutrino reconstruction" technique itself and as a measure of the systematic error. The results from an analysis of the $B^0 \to D^{*-}\ell^+\nu$ decay mode are shown in Fig. 23 where there is good agreement between the data and the

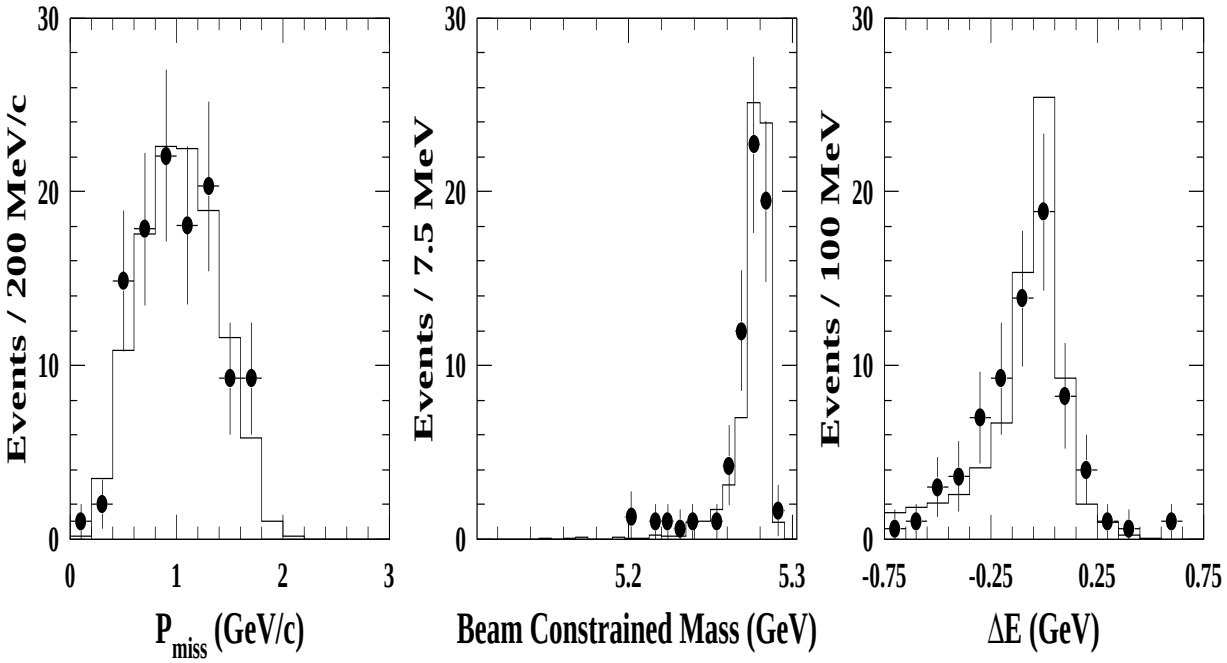


Figure 23: The "Neutrino Reconstruction" distributions in an analysis of the $B^0 \to D^{*-}\ell^+\nu$ decay chain. The points with error bars are the data and the histograms are from the Monte Carlo simulation.

Monte Carlo for the various kinematic distributions. The neutrino reconstruction result of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu) = (4.65 \pm 0.65)\%$ agrees with the CLEO published number[28] of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu) = (4.49 \pm 0.32 \pm 0.32)\%$.

Five modes are searched for in the $b \to u$ analysis — the pseudoscalar modes $\pi^-\ell^+\nu$ and $\pi^0\ell^+\nu$, and the vector modes $\rho^-\ell^+\nu$, $\rho^0\ell^+\nu$, and $\omega\ell^+\nu$. Both electrons and muons are used and, to suppress the $b \to c\ell\nu$ backgrounds, the leptons are required to have momenta greater than 1.5(2) GeV in the pseudoscalar(vector) modes. When extracting the yields, the ΔE-m_B distributions for the five modes are fit simultaneously. This allows for the utilization of the isospin constraints $\Gamma(B^0 \to \pi^-\ell^+\nu)$=$2\Gamma(B^+ \to \pi^0\ell^+\nu)$ and $\Gamma(B^0 \to \rho^-\ell^+\nu)$=$2\Gamma(B^+ \to \rho^0\ell^+\nu) \approx$ $2\Gamma(B^+ \to \omega\ell^+\nu)$. This method also allows for the feed-across from the vector to

pseudoscalar modes to be handled in a consistent manner.

A further complication in the vector modes are nonresonant $\pi^-\pi^0$, $\pi^-\pi^+$, and $\pi^-\pi^+\pi^0$ contributions, for which neither the rate nor the shape is known. This is dealt with in three different ways: (1) fit for the rates with just a simple cut in $\pi\pi(3\pi)$ mass around the $\rho(\omega)$, (2) fit for the rates after subtracting from the $\pi\pi(3\pi)$ mass in the $\rho(\omega)$ peak regions sideband samples in $\pi\pi(3\pi)$ mass, and (3) include the $\pi\pi(3\pi)$ distributions in the fit where additional assumptions must be made about the shapes of the nonresonant and background components. To illustrate the results from method (2), the beam-constrained mass plots for the ΔE signal region are shown in Fig. 24.

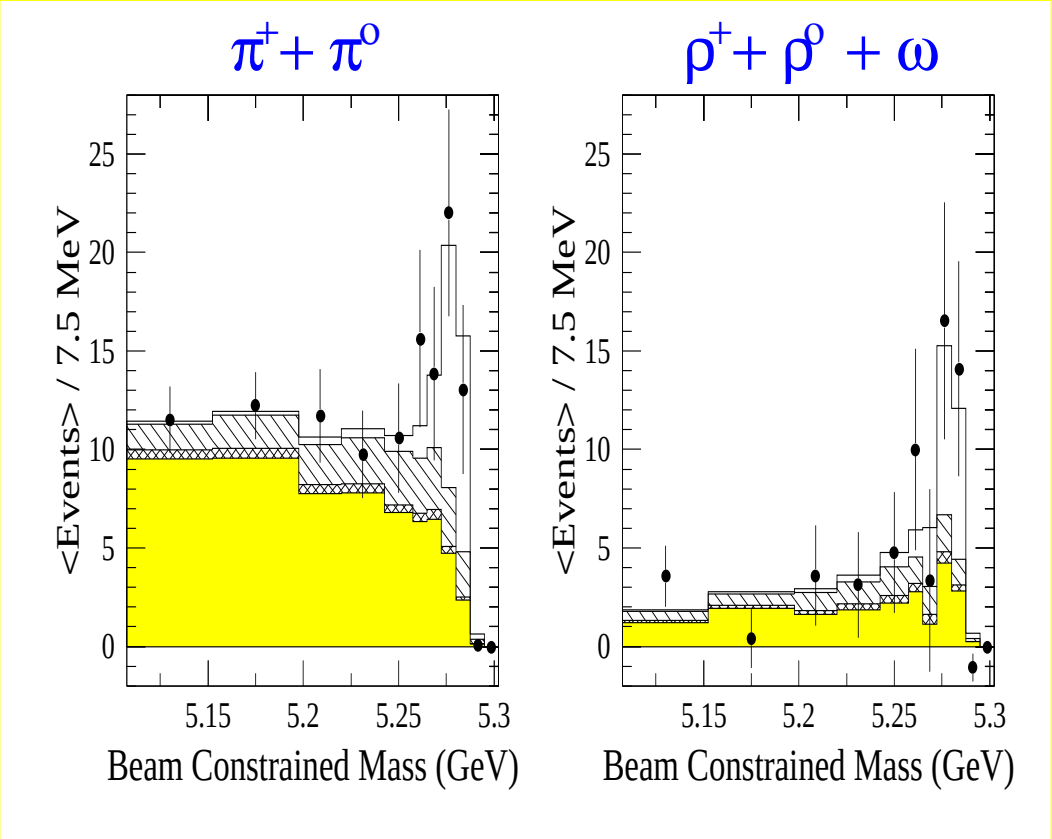


Figure 24: Beam-constrained mass distributions for the combined $\pi^-\ell^+\nu$ and $\pi^-\ell^+\nu$ (left) channels, and the sum of the vector modes (right) for the analysis with $\pi\pi/3\pi$ mass sideband subtractions. The points with error bars are the continuum- and fake-subtracted data. The histograms are the signal (hollow), the contribution from $b \to c\ell\nu$ (shaded), feed-down from higher mass $b \to u\ell\nu$ (cross-hatched), and signal mode cross-feed (hatched).

The yields in the vector modes from the different methods are about equal, indicating that there is very little nonresonant contribution to the rate.

Several other distributions were examined to check that these signals are consistent with having come from $b \to u\ell\nu$ decays. The lepton momentum spectra for the ΔE–M_B signal region are shown in Fig. 25 where, needless to say, the lepton momentum cuts have been removed in these plots. The lepton momentum

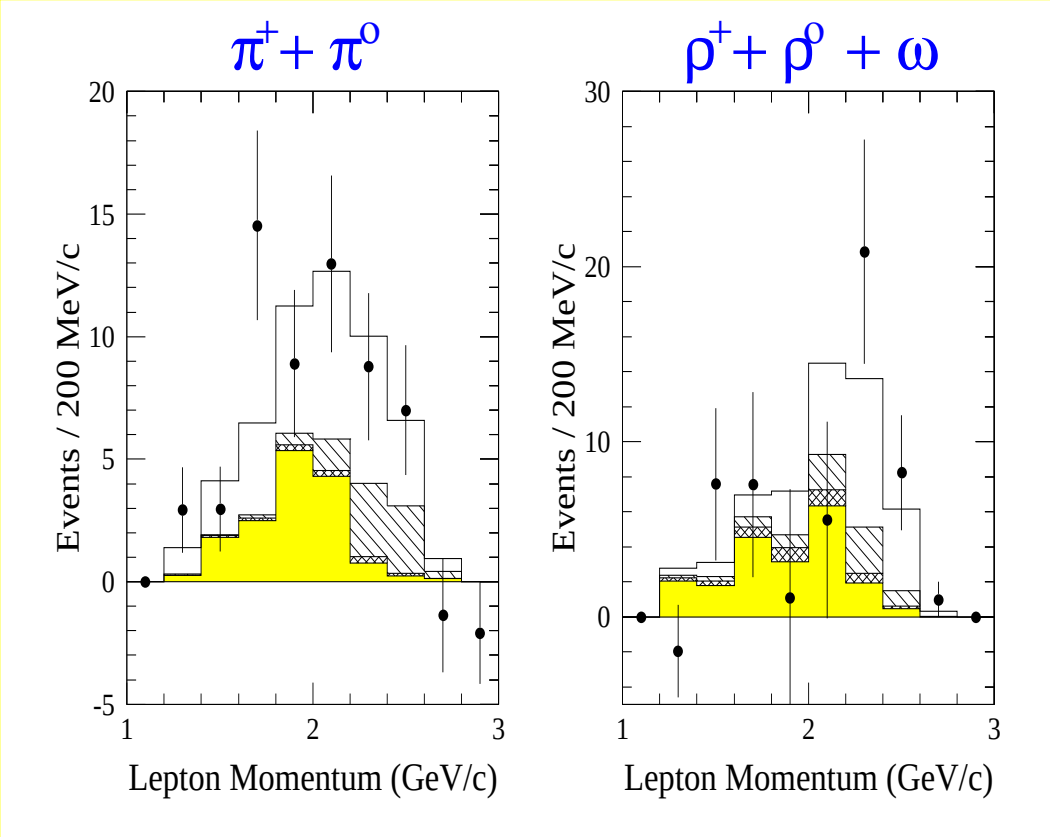


Figure 25: Lepton momentum spectra for $B \to \pi\ell\nu$ (left) and $B \to \rho\ell\nu$ (right). The various components have the same meanings as in Fig. 24 (except that here the points are also sideband subtracted) where the normalizations are predictions using the results of the fit to the mass spectra.

spectra are quite "hard," with many leptons beyond the $b \to c$ endpoint, indicative of having originated from $b \to u$ transitions. The exact shapes of the lepton momentum spectra depend on the poorly known $b \to u$ form factors.

There is form factor independent information in the distribution of the angle defined in Fig. 26. Because of the $V - A$ structure of the weak interaction, the signal in the pseudoscalar case follows a $\sin^2\theta_{\pi\ell}$ distribution. For the vector

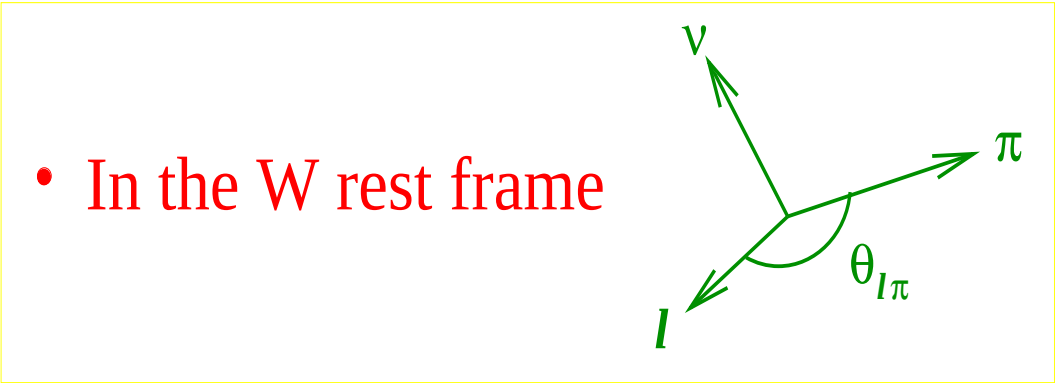


Figure 26: Definition of $\theta_{\pi\ell}$ in $B \to \pi\ell\nu$ decays. The angle for the vector meson case, $\theta_{\rho\ell}$, follows by replacing the pion with a vector meson.

modes, the details of the angular distributions depend on the form factor but a forward peaking in $\cos\theta_{\rho\ell}$ is expected. The distributions, shown in Fig. 27, are quite consistent with the $b \to u$ nature of the decays.

The final branching fractions are model dependent because the efficiencies for the various modes depend on the form factors used in the Monte Carlo. Specifically, different form factors concentrate the rate in different regions of q^2, and the efficiency is a function of q^2. Results obtained for the WSB and ISGW[‖] models are:[29,30]

$$\begin{aligned} \mathcal{B}(B^0 \to \pi^-\ell^+\nu) &= (1.34 \pm 0.35 \pm 0.28) \times 10^{-4} \quad \text{ISGW} \\ &= (1.63 \pm 0.46 \pm 0.34) \times 10^{-4} \quad \text{WSB} \end{aligned}$$

$$\begin{aligned} \mathcal{B}(B^0 \to \rho^-\ell^+\nu) &= (2.28 \pm 0.36 \pm 0.59^{+0.00}_{-0.46}) \times 10^{-4} \quad \text{ISGW} \\ &= (3.88 \pm 0.54 \pm 1.01^{+0.00}_{-0.78}) \times 10^{-4} \quad \text{WSB} \end{aligned}$$

where the third error for the vector mode is the uncertainty due to the non-resonant contribution.

There is some potential to discriminate between models using the ratio of rates. The results are given in Table 5 where the ISGW prediction appears to be inconsistent with the data. More studies of the model dependence need to be performed before a value of V_{ub} can be extracted from this analysis.

[‖] This is the 'original' ISGW. Results including the so-called ISGW2 model will be included in the final analysis.

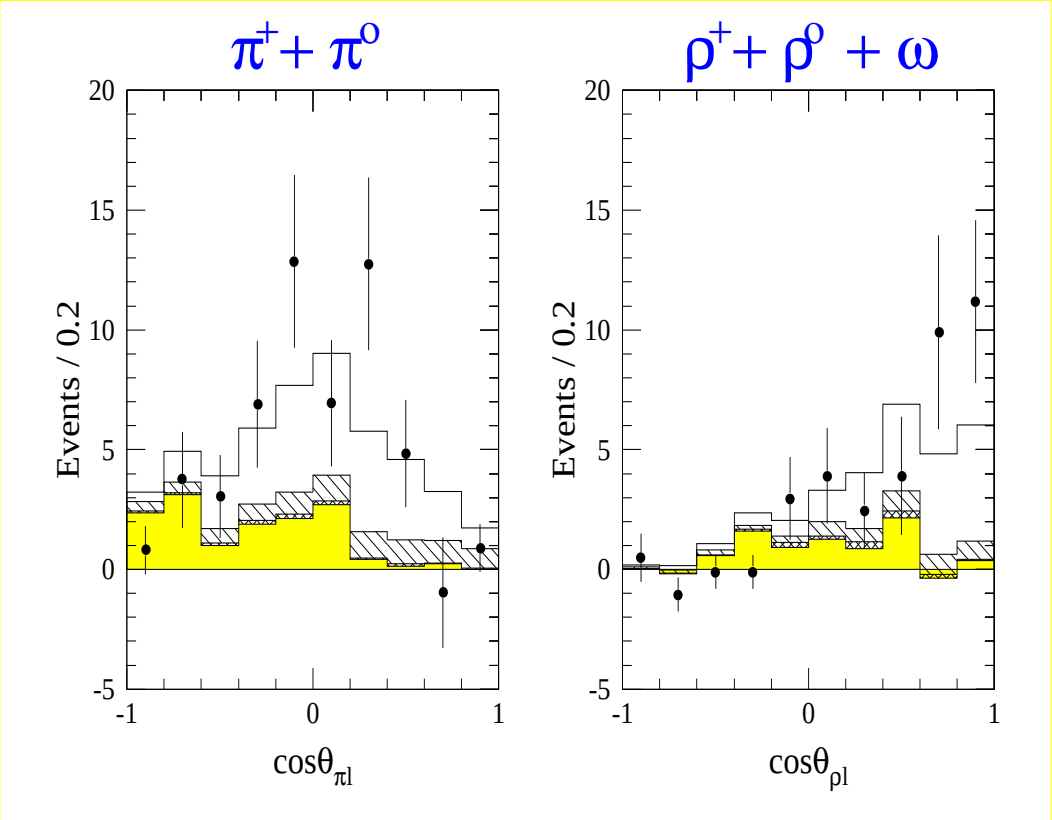


Figure 27: The $\cos\theta_{\pi\ell}$ (left) and $\cos\theta_{\rho\ell}$ (right) distributions where the various components have the same meaning as for Fig. 25.

Model	Prediction	CLEO II
ISGW	4.0	$1.70^{+0.81}_{-0.50} \pm 0.58^{+0.00}_{-0.34}$
WSB	3.0–4.3	$2.39 \pm^{+0.81}_{-0.50} \pm 0.58^{+0.00}_{-0.34}$

Table 5: Predictions and experimental results for the ratio of the $b \to u\ell\nu$ rates $\Gamma(B^0 \to \rho^-\ell^+\nu)/\Gamma(B^0 \to \pi^-\ell^+\nu)$.

5 Summary and the Future

The power of the CLEO II detector to reconstruct charged and neutral particles with high efficiency and good resolution, coupled with the increased luminosity of CESR, has allowed CLEO to probe deeply into our understanding of heavy

flavor physics. The first observation of exclusive $b \to u\ell\nu$ decays bodes well for our ability to measure V_{ub}, which is crucial to testing the entire CKM description of $\mathcal{CP}$ Violation. Meanwhile, in the $b \to c$ sector, the combination of precise measurements of the "baffling" semileptonic and hadronic branching fractions may be pointing towards a problem in the theoretical treatment of heavy flavor decays. The large CLEO charm sample has allowed for the observation of rare phenomena in the D_s^+ system like the observation of isospin violation and purely leptonic decays. Meanwhile, charmed baryons continue to be a laboratory for testing our understanding of the relative contributions of various quark level decay diagrams.

There are a series of improvements planned which will ensure the increased productivity of CLEO. CESR upgrades[31] will result in a doubling of the luminosity in 1995 with plans to exceed a luminosity of 10^{33} cm^{-2}s^{-1}, i.e. in the range of a B Factory, towards the end of this century. As for CLEO, a silicon vertex detector will be installed in the fall of 1995. This will not only improve tracking in general but, in particular, will improve the resolution on the crucial $D^{*+} - D^0$ mass difference by a factor of two or three. As well, the ability to reconstruct detached vertices will open up much of the physics associated with the long-lived D^+. The CLEO III era[32] will begin in 1997 with the installation of a new silicon vertex detector and a new drift chamber. Also planned is the installation of a Ring Imaging Cherenkov Counter (RICH) for much improved particle identification. The CsI calorimeter could be said to have revolutionized the physics potential of CLEO by making available a vast number of new channels involving photons (and, hence, the π^0), and the RICH could have a similarly profound impact on the physics reach of CLEO. Some of the physics gains that are made available to CLEO through better particle identification include:

- separating $B \to K^-\pi$ from $B \to \pi^-\pi$. This is crucial if one wants to use the integrated rates to extract angles of the unitarity triangle.
- separating $B^- \to D^0K^-$ from $B^- \to D^0\pi^-$. $B \to D^0K$ decays offer an intriguing way to measure an angle of the unitarity triangle[33] but the signal is swamped by the CKM favored $D^0\pi$ channel.
- separating $B \to \rho\gamma$ from $B \to K^*\gamma$. Measuring the ratio of these two rates was once touted as the best way to get to $|V_{ts}/V_{td}|$ and measuring either rate will give information pertinent to $b \to u$ decays.

- separating $D^+ \to \rho^0 \ell^+ \nu$ from $D^+ \to \bar{K}^{*0} \ell \nu$. The hope is that the form factors measured in this Cabibbo-suppressed semileptonic D^+ decay can be used, in lieu of theoretical models, in calculations involving $b \to u$ transitions.

These improvements will ensure that CLEO remains a "top" player in heavy quark physics well into the next century.

References

[1] The papers describing these results in much more detail, and giving the relevant references to other results and to theoretical predictions, can be accessed from the Cornell Wilson Lab home page on the World Wide Web. The URL is `http://w4.lns.cornell.edu/public/PUBLIC.html`. The relevant links are **CLNS reports**, which are essentially the papers as submitted to the journals, and **conference papers**, where the papers submitted to the 1995 summer conferences can be found.

[2] CLEO Collaboration, Y. Kubota *et al.*, Nucl. Inst. Methods, **A320**, 66 (1992).

[3] CLEO Collaboration, K. Edwards *et al.*, Phys. Rev. Lett. **74**, 3331 (1995).

[4] CLEO Collaboration, T. Miao *et al.*, Phys. Lett. **B323**, 219 (1994); G. Crawford *et al.*, Phys. Rev. Lett. **75**, 624 (1995).

[5] CLEO Collaboration, R. Ammar *et al.*, Phys. Rev. Lett. **74**, 3534 (1995).

[6] CLEO Collaboration, J. Alexander *et al.*, Phys. Rev. Lett. **74**, 3113 (1995).

[7] CLEO Collaboration, A. Bean *et al.*, CLEO CONF 93-8, (1993).

[8] P. Cho and M. B. Wise, Phys. Rev. **D49**, 6228 (1994), and (presumably) Mark Wise's contribution to these proceedings.

[9] The BaBar Technical Design Report, SLAC-R-95-457, (1995).

[10] CLEO Collaboration, D. Acosta *et al.*, Phys. Rev. **D49**, 5690 (1994).

[11] Particle Data Group, L. Montanet *et al.*, Phys. Rev. **D50**, S1 (1994).

[12] CLEO Collaboration, D. Gibaut *et al.*, CLNS 95/1354, submitted to Phys. Rev. **D**.

[13] G. Martinelli, "$D-$ and $B-$Meson Phenomenology from Lattice QCD," CERN-TH/95-116, (1995).

[14] S. Capstick and S. Godfrey, Phys. Rev. **D41**, 2856 (1990).

[15] CLEO Collaboration, M. S. Alam *et al.*, Phys. Rev. **D50**, 43 (1994).

[16] For a review, see J. D. Richman and P. R. Burchat, "Leptonic and Semileptonic Decays of Charm and Bottom Hadrons," UCSB-HEP-95-08, (1995).

[17] CLEO Collaboration, P. Avery *et al.*, Phys. Lett. **B337**, 405 (1994); F. Butler *et al.*, Phys. Lett. **B324**, 255 (1994).

[18] A. N. Kamal *et al.*, Phys. Rev. **D49**, 1330 (1994).

[19] CLEO Collaboration, P. Avery *et al.*, Phys. Rev. Lett. **68**, 1279 (1992).

[20] CLEO Collaboration, R. Balest *et al.*, Phys. Rev. **D52**, 2661 (1995).

[21] I. Bigi *et al.*, Phys. Lett. **B323**, 408 (1994).

[22] E. Bagan *et al.*, Phys. Lett. **B342**, 362 (1995).

[23] CLEO Collaboration, D. M. Asner *et al.*, CLNS95/1338, submitted to Phys. Rev. **D**.

[24] D. London, "CP Violation in the B System: What's Old, What's New?," **hep-ph/9406412**, (1994).

[25] CLEO Collaboration, M. S. Alam *et al.*, Phys. Rev. Lett. **74**, 2885 (1995).

[26] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. **B234**, 409 (1990); **B255**, 297 (1991).

[27] CLEO Collaboration, R. Fulton *et al.*, Phys. Rev. Lett. **64**, 16 (1990); J. Bartelt *et al.*, Phys. Rev. Lett. **71**, 4111 (1993).

[28] CLEO Collaboration, B. Barish *et al.*, Phys. Rev. **D51**, 1014 (1995).

[29] M. Wirbel *et al.*, Zeit. Phys. **C29**, 637 (1985).

[30] N. Isgur *et al.*, Phys. Rev. **D39**, 799 (1989).

[31] "The CESR/CLEO Upgrade Project," CLNS93/1265, (1993).

[32] "The CLEO III Detector: Design and Physics Goals," CLNS94/1277, (1994).

[33] M. Gronau and D. Wyler, Phys. Lett **B265**, 172 (1991).

Heavy Quark Physics from the Tevatron

P. Sphicas

Massachusetts Institute of Technology
Laboratory for Nuclear Science
Cambridge, MA 02139

No paper was received from the author in time for printing.

RECENT ADVANCES IN HEAVY QUARK THEORY

Mark Wise*
California Institute of Technology, Pasadena, California 91125

ABSTRACT

Some recent developments in heavy quark theory are reviewed. Particular emphasis is given to inclusive weak decays of hadrons containing a b quark. The isospin violating hadronic decay $D_s^* \to D_s \pi^0$ is also discussed.

1 Introduction

In this lecture, I describe some of the developments in heavy quark theory that have occurred recently. Those aspects of heavy quark theory that impact the determination of parameters in the Standard Model like $|V_{ub}|, |V_{cb}|$, and m_b are the most important. A precise determination of m_b may play a role in testing ideas about unification of the strong, weak, and electromagnetic interactions. Many unified theories predict that $m_b = m_\tau$ at the unification scale. (Later, I will distinguish between various definitions of the heavy quark mass, e.g., the pole mass or the $\overline{MS}$ mass. It is the $\overline{MS}$ mass that is approximately equal to the tau mass at the GUT scale.) In the Standard Model, the couplings of the W bosons to the quarks are given in terms of the elements of the Cabibbo-Kobayashi-Maskawa matrix, V_{ij}, which arises from diagonalizing the quark mass matrices. In the minimal Standard Model (i.e., one Higgs doublet), it is this matrix that is responsible

*Work supported in part by the U.S. Dept. of Energy under Grant No. DE-FG03-92-ER40701.

©1995 by Mark Wise.

for the CP nonconservation observed in weak kaon decays. (The limit on the electric dipole moment of the neutron means that the QCD vacuum angle is too small, $\bar{\theta} < 10^{-9}$, to have a measurable impact on weak decays.) A precise determination of the elements $|V_{ub}|$ and $|V_{cb}|$ will play an important role in testing this picture for the origin of CP violation and will constrain extensions of the Standard Model that make predictions for the form of the quark mass matrices.

While most of this lecture is directed towards aspects of heavy quark theory that impact the determination of $|V_{ub}|, |V_{cb}|$, and m_b, I will also spend some time discussing the implications of a recent measurement of the branching ratio for the isospin violating decay $D_s^* \to D_s \pi^0$. Heavy quark theory is still a rapidly developing field of study, and in this lecture, I only review a small part of the recent activity in this subject area.

2 Inclusive Semileptonic B and Λ_b Decay

The theory of inclusive B decay has developed rapidly over the last few years.[1–3] In this lecture, I consider inclusive semileptonic B decay in some detail and then generalize this discussion to other cases. The inclusive B semileptonic decay rate is equal to the b-quark decay rate with corrections suppressed by powers of Λ_{QCD}/m_b. Over the past few years, it has been shown how to express these nonperturbative QCD corrections in terms of the matrix elements of local operators in the heavy quark effective theory. The method used involves an operator product expansion and a transition to the heavy quark effective theory.

As far as hadronic physics is concerned, the basic quantity needed in inclusive semileptonic B decay is the second-rank hadronic tensor

$$W^{\mu\nu} = (2\pi)^3 \sum_X \delta^4(p_B - q - p_X) \langle B(v)| J_j^{\mu\dagger} |X\rangle \langle X| J_j^{\nu} |B(v)\rangle. \qquad (1)$$

In Eq. (1), J_j^μ is the weak current

$$J_j^\mu = \bar{q}_j \gamma^\mu \frac{(1-\gamma_5)}{2} b, \qquad j = u, c,$$

for either $b \to c$ or $b \to u$ transitions, v is the four velocity of the B meson, $p_B = m_B v$, and the sum goes over all possible final

hadronic states X. The tensor $W_{\mu\nu}$ can be expanded in terms of scalars, $W_a(q^2, v\cdot q),\ a = 1, ... 5$, as follows:

$$W^{\mu\nu} = -g^{\mu\nu}W_1 + v^\mu v^\nu W_2 - i\varepsilon^{\mu\nu\alpha\beta}v_\alpha q_\beta W_3$$

$$+ q^\mu q^\nu W_4 + (q^\mu v^\nu + q^\nu v^\mu)W_5. \quad (2)$$

The form factors W_4 and W_5 give effects proportional to the lepton mass and can be neglected in $B \to Xe\bar{\nu}_e$ decay (they are important for $B \to X\tau\bar{\nu}_\tau$ decay).[4] In terms of the scalar form factors W_a, the inclusive semileptonic $B \to Xe\bar{\nu}_e$ differential decay rate is

$$\frac{d\Gamma}{dq^2 dE_e dE_\nu} = \frac{|V_{jb}|^2 G_F^2}{2\pi^3}[W_1 q^2 + W_2(2E_e E_\nu - \frac{1}{2}q^2) + W_3 q^2(E_e - E_\nu)]$$

$$\times\, \theta(E_\nu - q^2/4E_e). \quad (3)$$

Here, E_e and E_ν are the electron and neutrino energies in the B rest frame. The limit over neutrino energies given by the theta function comes from

$$q^2 = (p_e + p_{\bar{\nu}})^2 = 2E_e E_\nu(1 - \cos\theta_{e\nu}) < 4E_e E_\nu. \quad (4)$$

The form factors W_a are proportional to the discontinuity across a cut in the analogous form factors for the time-ordered product of currents

$$T^{\mu\nu} = -i\int d^4x e^{-iq\cdot x}\langle B(v)|\{T(J_j^{\mu\dagger}(x)J_j^\nu(0)\}|B(v)\rangle$$

$$= -g^{\mu\nu}T_1 + v^\mu v^\nu T_2 - i\varepsilon^{\mu\nu\alpha\beta}v_\alpha q_\beta T_3 + q^\mu q^\nu T_4 + (q^\mu v^\nu + q^\nu v^\mu)T_5, \quad (5)$$

where $T_a = T_a(q^2, v\cdot q), a = 1, ..., 5$. Viewing q^2 as fixed, T_a has cuts in the complex $v\cdot q$ plane along the real axis. The discontinuity across the cut associated with $B \to Xe\bar{\nu}_e$ semileptonic decay gives the W_a. (There are other cuts not associated with this process. For example, along the positive real axis, there is a cut corresponding to $\nu_e B \to Xe$, where X contains two b quarks. This cut arises from $J_j^{\mu\dagger}$ acting on the B meson to produce X.) It is possible to express weighted averages (over $v\cdot q$) of the form factors W_a as contour integrals of the analytically continued $T_a(q^2, v\cdot q)$, where, for the most part, the contour is not close to the cuts.

We perform an operator product expansion on the time-ordered product of the two currents. Because this expansion holds at the operator level, we can identify the operators and their coefficients by taking matrix elements of the time-ordered product between b-quark "states" and comparing that with b-quark matrix elements of local operators. The momentum of the incoming b quark is written as $p_b = m_b v + k$ and the residual momentum, k, is expanded in, with higher powers of k being associated with higher dimensional operators in the heavy quark effective theory. The leading operators encountered are $\bar{b}\gamma_\lambda b$ and $\bar{b}\gamma_\lambda\gamma_5 b$. In a B meson, the second of these has a zero matrix element because of the parity invariance of the strong interactions. The first operator has a known forward matrix element because it is the conserved b-quark number current

$$\langle B(v)|\bar{b}\gamma_\lambda b|B(v)\rangle = 2v_\lambda. \tag{6}$$

(Hadronic B-meson states are normalized to $2v_0$ instead of $2m_B v_0$.) At this level, the operator-produced expansion reproduces the b-quark decay rate.

At zeroth order in the residual momentum k, there is no reason to make a transition to the heavy quark effective theory. However, at linear order in k, it is convenient, for keeping track of the m_b dependence of matrix elements, to make the transition to the heavy quark effective theory (HQET) defining

$$b(x) = e^{-im_b v\cdot x} h_v^{(b)}(x), \tag{7}$$

where the HQET b-quark field, $h_v^{(b)}(x)$, satisfies

$$\not{v} h_v^{(b)}(x) = h_v^{(b)}(x). \tag{8}$$

The operators that are encountered at linear order have dimension four, and the only ones are $\bar{h}_v^{(b)} iD^\tau \gamma_5 h_v^{(b)}$ and $\bar{h}_v^{(b)} iD^\tau h_v^{(b)}$, where D^τ denotes a covariant derivative. But these operators have a zero forward matrix element. For example, the first vanishes by parity, while for the second, Lorentz invariance implies that

$$\langle B(v)|\bar{h}_v^{(b)} iD^\tau h_v^{(b)}|B(v)\rangle = Y v^\tau. \tag{9}$$

Contracting the above with v^τ and using $v^2 = 1$ gives

$$\langle B(v)|\bar{h}_v^{(b)} iv \cdot D h_v^{(b)}|B(v)\rangle = Y, \tag{10}$$

and the equation of motion in HQET,

$$iv \cdot D h_v^{(b)} = 0, \tag{11}$$

implies that $Y = 0$. This means that there are no Λ_{QCD}/m_b corrections to the b-quark decay picture! Nonperturbative strong interaction corrections first arise at order $(\Lambda_{QCD}/m_b)^2$ and are parametrized by the two matrix elements[2,3]

$$\lambda_1 = \frac{1}{2}\langle B(v)|\bar{h}_v^{(b)}(iD)^2 h_v^{(b)}|B(v)\rangle, \tag{12a}$$

and

$$\lambda_2 = \frac{1}{6}\langle B(v)|\frac{g}{2}\bar{h}_v^{(b)}\sigma_{\alpha\beta}G^{\alpha\beta}h_v^{(b)}|B(v)\rangle. \tag{12b}$$

The first of these is related to the kinetic energy of the b quark in the B meson, and the second is related to the chromomagnetic energy arising from the b-quark spin. λ_2 is determined by the $B^* - B$ mass splitting to be

$$\lambda_2 = m_b\left(\frac{m_{B^*} - m_B}{2}\right) = 0.12\ GeV^2, \tag{13}$$

and we expect (at the order of magnitude level) $\lambda_1 \approx -\lambda_2$.

The operator product expansion gives the decay rate in terms of quark kinematics with the phase space set by the heavy quark pole masses m_b and m_c. However, we can reexpress the differential decay rate in terms of hadron masses using

$$m_B = m_b + \bar{\Lambda} - \frac{\lambda_1 + 3\lambda_2}{2m_b}, \tag{14a}$$

$$m_D = m_c + \bar{\Lambda} - \frac{\lambda_1 + 3\lambda_2}{2m_c}. \tag{14b}$$

The differential decay rate depends on $\bar{\Lambda}$ and λ_1 (λ_2 is fixed by experiment) and may be used to determine these quantities. At the present time, such an analysis[5] (including perturbative QCD corrections at order α_s) for semileptonic $B \to X_c e\bar{\nu}_e$ decay gives the lower bound

$\bar{\Lambda} > \left[0.33 - 0.07\left(\frac{\lambda_1}{0.1\ GeV^2}\right)\right] GeV$. It is $\bar{\Lambda}$ that determines the pole mass m_b. The pole mass is not a physical quantity,[6] and the perturbative expression for the $\overline{MS}$ mass $\overline{m_b}(m_b)$ in terms of the pole mass, m_b, is not Borel summable giving rise to what is sometimes called a "renormalon ambiguity" in the pole mass. However, when the differential semileptonic decay rate is expressed in terms of hadron masses and $\bar{\Lambda}$, the perturbative QCD corrections to the decay rate are also not Borel summable. If $\bar{\Lambda}$ (or equivalently, the b-quark pole mass) extracted from the differential semileptonic decay rate is used to get the $\overline{MS}$ mass, these ambiguities cancel so one can arrive at a meaningful prediction for the $\overline{MS}$ b-quark mass. The basic lesson here is that it is fine to introduce unphysical quantities like the heavy quark pole mass or $\bar{\Lambda}$ as long as one works consistently to a given order of QCD perturbation theory. Since the final relations one considers always involve relations between physically measurable quantities, any "renormalon ambiguities" resulting from the bad behavior of the QCD perturbation series at large orders will cancel out.[7]

The method I have outlined for semileptonic B decay has been extended to polarized Λ_b decay and to the rare decay, $B \to X_s\gamma$ (Ref. 8). The latter may play a particularly important role[9] in extracting the parameter $\bar{\Lambda}$. Study of the exclusive decay $\Lambda_b \to \Lambda_c e\bar{\nu}_e$ can also lead to a determination of $\bar{\Lambda}$ (Ref. 10).

We have seen that the electron spectrum in semileptonic B decay, $d\Gamma/dE_e$, can be predicted, including nonperturbative strong interactions effects, in terms of $\bar{\Lambda}$ and two matrix elements λ_1 and λ_2. Over most of the phase space, this description is adequate with λ_1 and λ_2 giving only modest corrections $\sim 5\%$. However, for extracting $|V_{ub}|$, it is necessary to focus on the endpoint region of electron energies $(m_B^2 - m_D^2)/2m_B < E_e < (m_B^2 - m_\pi^2)/2m_B$, where $b \to c$ transitions are forbidden kinematically. For electron energies very near their maximal value, only low-mass final hadronic states are allowed and a description in terms of the operator product expansion is inappropriate. For $B \to X_u e\bar{\nu}_e$ decay, the nonperturbative QCD corrections proportional to λ_1 and λ_2 are singular at $E_e^{max} = m_b/2$. They must be averaged over a region of electron energies ΔE_e before a comparison with experiment

can be made. It is sufficient to stop the operator product expansion at dimension five operators provided $\Delta E_e \gg \Lambda_{QCD}$. This is too large a region of electron energies to be useful for getting $|V_{ub}|$ from the endpoint region of the electron spectrum in semileptonic B decay. If a particular infinite class of operators is included, the resolution with which the electron spectrum can be examined near maximal electron energies is improved to $\Delta E_e \sim \Lambda_{QCD}$. This may be small enough to allow a comparison with experimental data in the endpoint region. However, there is now a loss of predictive power because an infinite number of nonperturbative matrix elements are needed to characterize the electron energy spectrum. Fortunately, it has been shown that the same infinite class of operators (occurring in the same linear combination) determines the photon energy spectrum for the inclusive decays $B \to X_s\gamma$ near maximal photon energy. In principle, experimental information on $B \to X_s\gamma$ can be used to predict the electron spectrum in B decay, in a region near the maximal electron energy[11] that may be small enough to allow a model-independent extraction of $|V_{ub}|$. A comparison between exclusive B and D decays can also lead to a model-independent determination of $|V_{ub}|$ (Ref. 12).

In addition to the nonperturbative QCD corrections suppressed by powers of Λ_{QCD}/m_b, there are perturbative α_s corrections to the b-quark semileptonic decay rate that must be included to make an accurate prediction for the B or Λ_b semileptonic decay rate and the electron energy spectrum. These have been calculated at order[13] $(\alpha_s(m_b)/\pi)$, and recently, the corrections of order $(\alpha_s(m_b)/\pi)^2$ that are proportional to the QCD beta function (these are tagged by computing the part of the order $(\alpha_s(m_b)/\pi)^2$ correction proportional to the number of light quark flavors) have also been computed.[14] Typically, these are the most important order $(\alpha_s(m_b)/\pi)^2$ corrections and Brodsky, Lepage, and Mackenzie[15] have advocated choosing the argument of α_s in the leading perturbative QCD correction to remove this "two loop" correction.

For definiteness, consider the case of $b \to u$ transitions. Then

$$\Gamma(B \to X_u e \bar{\nu}_e) = \frac{G_F^2 |V_{ub}|^2 m_B^5}{192\pi^3}\Bigg[1 - 2.41\frac{\bar{\alpha}_s(m_b)}{\pi}$$

$$- 2.98\beta_0\left(\frac{\bar{\alpha}_s(m_b)}{\pi}\right)^2 - \frac{5\bar{\Lambda}}{m_B} + ...\Bigg], \tag{15}$$

where

$$\beta_0 = 11 - \frac{2}{3}n_f,$$

with n_f the number of light quark flavors. If the subtraction point used for the strong coupling in the order α_s term is changed from m_b to μ_{BLM} (μ_{BLM} is chosen so that the two-loop term proportional to β_0 is removed), one finds $\mu_{BLM} \simeq 0.08\ m_b$. The two-loop term proportional to β_0 can be reduced in comparison to the order α_s term if one eliminates $\bar{\Lambda}$ in favor of a physically measurable quantity characteristic of these decays. For example, the average final hadronic mass squared is[5]

$$< m_{X_u}^2 > = m_B^2\left[0.20\frac{\alpha_s(m_b)}{\pi} + 0.35\beta_0\left(\frac{\alpha_s(m_b)}{\pi}\right)^2 + \frac{7}{10}\frac{\bar{\Lambda}}{m_B^2} + ...\right], \tag{16}$$

and reexpressing the semileptonic decay rate in terms of this quantity gives

$$\Gamma(B \to X_u \ell \bar{\nu}_e) = \frac{G_F^2 |V_{ub}|^2 m_B^5}{192\pi^3}\Bigg[1 - 0.98\frac{\alpha_s(m_b)}{\pi}$$

$$- 0.48\beta_0\left(\frac{\alpha_s(m_b)}{\pi}\right)^2 - 7.14\frac{\langle m_{X_u}^2\rangle}{m_B^2}\Bigg]. \tag{17}$$

Now the BLM scale is considerably larger, $\mu_{BLM} = 0.38\ m_b$. Note that expressing the decay rate in terms of a physical quantity free of renormalon ambiguities does not guarantee a reasonably large BLM scale. For example, if one expresses the semileptonic decay rate in terms of the $\overline{MS}$ quark mass $\bar{m}_b(m_b)$, the BLM scale is still quite low.[14]

Perturbative QCD corrections to the electron spectrum (like the nonperturbative ones) become large in the endpoint region, and careful consideration of their effects is necessary for an extraction of $|V_{ub}|$ from the endpoint region of the electron spectrum.[16]

3 Inclusive Nonleptonic B and Λ_b Decay

The ideas I have outlined for inclusive semileptonic decay of hadrons containing a b quark have also been applied to nonleptonic decays.[17] Now there is no analog to $v\cdot q$ to analytically continue. Nonetheless, we expect to be able to express the total decay rate as b-quark decay plus nonperturbative QCD corrections given by forward matrix elements of local operators. This is because the energy release in B decay is large enough that threshold effects which spoil the applicability of local duality are probably negligibly small. (A similar argument is used to compare $R(s) = \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \mu^+\mu^-)$ with data at a fixed, but large, s.) The general structure of the nonperturbative QCD corrections to the nonleptonic decay rate is (schematically)

$$\Gamma = \Gamma_0\Bigg[1 + \frac{c}{m_b^2}\langle B(v)|\bar{h}_v^{(b)}(iD)^2 h_v^{(b)}|B(v)\rangle + \frac{d}{m_b^2}\langle B(v)|\frac{g}{2}\bar{h}_v^{(b)}\sigma_{\alpha\beta}G^{\alpha\beta}h_v^{(b)}|B(v)\rangle + \frac{e_\Gamma}{m_b^3}\langle B(v)|\bar{h}_v^{(b)}\Gamma q\bar{q}\Gamma h_v^{(b)}|B(v)\rangle + ...\Bigg], \tag{18}$$

where the ellipsis denote terms of order higher than $1/m_b^3$, and Γ_0 is the b-quark decay rate. A similar formula holds for the nonleptonic Λ_b decay rate. An interesting aspect of the order $1/m_b^3$ corrections is that they correspond to contributions to the nonleptonic decay rate where the phase space (at the quark level) is basically two body, and so the coefficients, e_Γ, are enhanced by a factor of 16 π^2 over c and d. Phenomenological models suggest that the contributions of the four quark operators are the most important for lifetime differences between hadrons containing a b quark.

Experimentally, the Λ_b lifetime is about 20% shorter than the B lifetime. This is a smaller lifetime than can be accommodated by quark model estimates of the matrix elements of the four quark operators. The charm quark mass dependence of the perturbative order α_s corrections to the nonleptonic decay rate increase the $b \to c\bar{c}s$ contribution[18] leading to an expected charm multiplicity of $n_c \sim 1.3$. Experimentally, the charm multiplicity is only $n_c = 1.17 \pm 0.04$.

At the present time, it is difficult to interpret these conflicts between the theory of inclusive nonleptonic decay and the experimental data. One possibility is that the matrix elements of the four quark operators are unusually large and the experimental value of the charm multiplicity (which relies on absolute branching ratios) is mismeasured. Another possibility is that both the charm multiplicity and the Λ_b lifetime are correctly measured and one has an unusually large violation of local duality in inclusive nonleptonic B decay. In any case, it seems prudent given these problems to use the semileptonic decay width for precision extractions of $|V_{cb}|$ from experiment rather than the B lifetime. Inclusive semileptonic B decay should give a determination of $|V_{cb}|$ with a precision comparable to the extraction of $|V_{cb}|$ from exclusive $B \to D^* e \bar{\nu}_e$ decay.[19] Some recent work that uses dispersion relations[20] to reduce uncertainties associated with the extrapolation of the Isgur–Wise function to zero recoil may improve the accuracy of extractions of $|V_{cb}|$ from this exclusive decay.

4 Decays of D^* Mesons

The ground state multiplet of charm mesons has spin of the light degrees of freedom, $s_\ell = 1/2$, and negative parity. This gives a doublet of mesons with total spins zero and one. The heavier, spin-one mesons decay to the lower mass spin-zero mesons by emission of either a photon or a pion. The measured branching ratios are shown in the table below.

Decay Mode	**Branching Ratio %**
$D^{*0} \to D^0 \pi^0$	$63.6 \pm 2.3 \pm 3.3$
$D^{*0} \to D^0 \gamma$	$36.4 \pm 2.3 \pm 3.3$
$D^{*+} \to D^0 \pi^+$	$68.1 \pm 1.0 \pm 1.3$
$D^{*+} \to D^+ \pi^0$	$30.8 \pm 0.4 \pm 0.8$
$D^{*+} \to D^+ \gamma$	$1.1 \pm 1.4 \pm 1.6$

In the nonrelativistic constituent quark model, the invariant matrix elements for radiative D^* decay are determined in terms of the *constituent* quark masses

$$M(D^{*0} \to D^0\gamma) \propto \left[\frac{2}{3m_c} + \frac{2}{3m_u}\right], \tag{19a}$$

$$M(D^{*+} \to D^+\gamma) \propto \left[\frac{2}{3m_c} - \frac{1}{3m_d}\right], \tag{19b}$$

$$M(D_s^{*0} \to D_s^0\gamma) \propto \left[\frac{2}{3m_c} - \frac{1}{3m_s}\right]. \tag{19c}$$

For $m_u = m_d = 350\ MeV$, $m_s = 550\ MeV$, and $m_c = 1.7\ GeV$, these are in the ratio

$$M(D^{*0} \to D^0\gamma) : M(D^{*+} \to D^+\gamma) : M(D_s^{*0} \to D_s^0\gamma)$$

$$= 1 \quad : \quad -0.25 \quad : \quad -0.1. \tag{20}$$

Presumably, the smallness of the radiative D^{*+} decay rate is due to the cancellation between down and charm quark magnetic moments in Eq. (19b). Notice that this cancellation is even stronger for the D_s^* decay because the constituent strange quark is heavier than the down quark. But how can we verify experimentally that this decay rate is very small? After all, the D_s^* is too narrow for its width to be measured. The answer is through measurement of the $D_s^* \to D_s\pi^0$ branching ratio. At leading order in chiral perturbation theory,[21] $D_s^* \to D_s\pi^0$ decay arises from isospin violating $\eta - \pi^0$ mixing which gives the rate

$$\Gamma(D_s^* \to D_s\pi^0) = \frac{g^2}{48\pi f^2}\left[\frac{m_u - m_d}{m_s - (m_u + m_d)/2}\right]^2 |\vec{p}_\pi|^3. \tag{21}$$

The factor in square brackets is $1/43.7$ (since this is greater than α/π, electromagnetic contributions to isospin violation can be neglected). In Eq. (21), g is the $D^*D\pi$ coupling. Equation (21) implies that

$$Br(D_s^* \to D_s\pi) = \frac{\Gamma(D_s^* \to D_s\pi^0)}{\Gamma(D_s^* \to D_s\gamma)}$$

$$= [8 \times 10^{-5}/Br(D^{*+} \to D^{*+}\gamma)]\left[\frac{\Gamma(D^{*+} \to D^{*+}\gamma)}{\Gamma(D_s^* \to D_s\gamma)}\right]. \tag{22}$$

Here, we have used the theoretical expression for the $D^{*+} \to D\pi$ decay rate to eliminate g. We expect $Br(D^{*+} \to D^+\gamma)$ to be around 1%. Then the branching ratio for $D_s^* \to D_s\pi^0$ should be significantly greater than 10^{-2} if the constituent quark model suppression of the $D_s^* \to D_s\gamma$ amplitude occurs in nature. The recent CLEO measurement,[22] $Br(D_s \to D_s\pi^0) = 0.062^{+0.020}_{-0.018} \pm 0.022$, indicates that, at least at some level, this suppression does occur.

References

[1] M. Shifman and M. Voloshin, Sov. J. Nucl. Phys. **41**, 120 (1985).

[2] J. Chay *et al.*, Phys. Lett. B **247**, 399 (1990); I. I. Bigi *et al.*, Phys. Lett. B **293**, 430 (1992); Phys. Lett. B **297**(E), 477 (1993); I. I. Bigi *et al.*, Phys. Rev. Lett. **71**, 496 (1993).

[3] A. V. Manohar and M. B. Wise, Phys. Rev. D **49**, 1310 (1994); B. Blok *et al.*, Phys. Rev. D **49**, 3356 (1994); T. Mannel, Nucl. Phys. B **413**, 396 (1994).

[4] A. F. Falk *et al.*, Phys. Lett. B **326**, 145 (1994); L. Koyrakh, Phys. Rev. D **49**, 3379 (1994); S. Balk *et al.*, Z. Phys. C **64**, 37 (1994).

[5] A. F. Falk *et al.*, UT PT-95-11 (hep-ph/9507284), 1995 (unpublished); UTPT-95-24 (hep-ph/9511454), 1995 (unpublished).

[6] M. Beneke and V. M. Braun, Nucl. Phys. B **426**, 301 (1994); I. Bigi *et al.*, Phys. Rev. D **50**, 2234 (1994).

[7] M. Beneke *et al.*, Phys. Rev. Lett. **73**, 3058 (1994); M. Luke *et al.*, Phys. Rev. D **51**, 4924 (1994); M. Neubert and C. Sachrajda, Nucl. Phys. B **438**, 235 (1995).

[8] A. F. Falk *et al.*, Phys. Rev. D **49**, 3367 (1994); M. Neubert, Phys. Rev. D **49**, 4623 (1994); I. Bigi *et al.*, in the Fermilab Meeting, *Proceedings of the Annual Meeting of the DPF of the APS*, edited by C. Albright *et al.* (World Scientific, Singapore, 1993), p. 610.

[9] A. Kapustin and Z. Ligeti, Phys. Lett. B **355**, 318 (1995); R. D. Dikeman *et al.,* TPT-MINN-95/9-T hep-ph/9505397, 1995 (unpublished).

[10] H. Georgi *et al.*, Phys. Lett. B **252**, 456 (1990).

[11] M. Neubert, Phys. Rev. D **49**, 3392 (1994); ibid 4623; I. I. Bigi *et al.*, Int. J. Mod. Phys. A **9**, 2467 (1994).

[12] For a recent discussion, see Z. Ligeti and M. B. Wise, CALT-68-2029, hep-ph/9512225, 1995 (unpublished).

[13] M. Jezabek and J. H. Kühn, Nucl. Phys. B **314**, 1 (1989).

[14] M. Luke *et al.*, Phys. Lett. B **343**, 329 (1995); Phys. Lett. B **345**, 301 (1995).

[15] S. J. Brodsky, P. Lepage, and P. B. Mackenzie, Phys. Rev. D **28**, 228 (1983).

[16] A. F. Falk *et al.*, Phys. Rev. D **49**, 4553 (1994); G. P. Korchemsky and G. Sterman, Phys. Lett. B **340**, 108 (1994).

[17] For a recent review, see I. I. Bigi, UND-HEP-PH-BIG02, hep-ph/9508408, 1995 (unpublished).

[18] Q. Hokin and X-Y. Pham, Ann. of Phys. **155**, 202 (1984); E. Bagen *et al.*, Phys. Lett. B **351**, 546 (1995).

[19] For a review, see M. Neubert, Phys. Rev. D **45**, 250 (1994).

[20] C. G. Boyd *et al.*, UCSD-PTH-95-11, hep-ph/9508211, 1995 (unpublished).

[21] P. Cho and M. B. Wise, Phys. Rev. D **49**, 6228 (1994).

[22] J. Gronberg *et al.*, (CLEO Collaboration) Phys. Rev. Lett. **75**, 3232 (1995).

THE SLAC SUMMER INSTITUTE ON PARTICLE PHYSICS
1973 - 1994

1973 DEEP INELASTIC ELECTROPRODUCTION

"Hadron Production in Deep Inelastic Processes"	J.D. Bjorken
"Hadron Production in the Collision of Virtual Photons with Nucleons–an Experimental Review"	M. Perl
"Phenomenological Analysis"	F.J. Gilman
"Asymptotic Behavior and Short Distance Singularities"	Y. Frishman

WEAK CURRENTS AND INTERACTIONS

"Raw Notes of Lectures on Current Algebra and PCAC"	S. Drell
"Experimental Phenomenology"	M. Schwartz S. Wojcicki
"Gauge-Theories of Weak and Electromagnetic Interaction"	T. Appelquist
"How to Transform Current Quarks to Constituent Quarks and Try to Predict Hadron Decays"	J. Primack K. Kugler

1974 THE STRONG INTERACTIONS

"Diffractive Processes"	D.W.G.S. Leith
"Amplitude Structure in Two- and Quasi-Two-Body Processes"	M. Davier
"Resonances: Experimental Review"	R.J. Cashmore
"Resonances: A Quark View of Hadron Spectroscopy and Transitions"	F.J. Gilman
"Lectures on Inclusive Hadronic Processes"	D. Sivers
"Large Momentum Transfer Processes"	R. Blankenbecler
"Hadron Dynamics"	H.D.I. Abarbanel

1975 DEEP HADRONIC STRUCTURE AND THE NEW PARTICLES

"Leptons as a Probe of Hadronic Structure"	F.J. Gilman
"Lepton Scattering as a Probe of Hadronic Structure"	E.D. Bloom
"High $p_\perp$ Dynamics"	J.D. Bjorken
"Hadronic Collision and Hadronic Structure (An Experimental Review)"	M. Davier
"The New Spectroscopy"	H. Harari
"The New Spectroscopy (An Experimental Review)"	G.H. Trilling

1976 WEAK INTERACTIONS AT HIGH ENERGY AND THE PRODUCTION OF NEW PARTICLES

"Weak Interaction Theory and Neutral Currents"	J.D. Bjorken
"Weak Interactions at High Energy"	S.G. Wojcicki
"ψ Spectroscopy"	G.J. Feldman
"A New Lepton?"	G.J. Feldman
"Lectures on the New Particles"	J.D. Jackson
"New Particle Production"	D. Hitlin

1977 QUARK SPECTROSCOPY AND HADRON DYNAMICS

"Quarks and Leptons"	H. Harari
"Quark Confinement"	S.D. Drell
"Hadron Spectroscopy"	F.J. Gilman
"Lectures on the Quark Model, Ordinary Mesons, Charmed Mesons, and Heavy Leptons"	M.L. Perl
"Quarks and Particle Production"	K.C. Moffeit

1978 WEAK INTERACTIONS–PRESENT AND FUTURE

"Accelerator Neutrino Experiments"	D.H. Perkins
"Weak Interactions at High Energies"	J. Ellis
"Gauge Theories of the Weak Interactions"	H. Quinn
"Weak Decays"	S. Wojcicki

1979 QUANTUM CHROMODYNAMICS

"Lepton Nucleon Scattering"	W.G. Atwood
"Massive Lepton Pair Production"	R. Stroynowski
"QCD Phenomenology of the Large P_T Processes"	R. Stroynowski
"Perturbative Quantum Chromodynamics"	S.J. Brodsky
"Elements of Quantum Chromodynamics"	J.D. Bjorken

1980 THE WEAK INTERACTIONS

"Gauge Theories of Weak Interactions"	M.J. Veltman
"Neutrinos and Neutrino Interactions"	F.J. Sciulli
"Weak Decays of Strange and Heavy Quarks"	D. Hitlin
"From the Standard Model to Composite Quarks and Leptons"	H. Harari
"Physics of Particle Detectors"	D.M. Ritson J. Jaros J. Marx H.A. Gordon R.S. Gilmore W.B. Atwood

1981 THE STRONG INTERACTIONS

"Quark-Antiquark Bound State Spectroscopy and QCD"	E.D. Bloom
"Meson Spectroscopy: Quark States and Glueballs"	M.S. Chanowitz
"Quantum Chromodynamics and Hadronic Interactions at Short Distances"	S.J. Brodsky
"Untangling Jets from Hadronic Final States"	G. Fox
"Heavy Flavor Production from Photons and Hadrons"	C.A. Heusch
"Design Constraints and Limitations in e^+e^- Storage Rings"	J. LeDuff
"Linear Colliders: A Preview"	H. Wiedemann

1982 PHYSICS AT VERY HIGH ENERGIES

"Expectations for Old and New Physics at High Energy Colliders"	R.N. Cahn
"Beyond the Standard Model in Lepton Scattering and Beta Decay"	M. Strovink
"Grand Unification, Proton Decay, and Neutrino Oscillations"	H.H. Williams
"e^+e^- Interactions at Very High Energy: Searching Beyond the Standard Model"	J. Dorfan
"The Hierarchy Gauge Susskind Problem, Technicolor, Supersymmetry, and All That"	L. Susskind
"Composite Models for Quarks and Leptons"	H. Harari
"Electron-Proton Colliding Beams: The Physics Programme and Machine"	B.H. Wiik

1983 DYNAMICS AND SPECTROSCOPY AT HIGH ENERGY

"Jets in QCD: A Theorist's Perspective"	S. Ellis
"Jets in e^+e^- Annihilation"	R. Hollebeek
"Jet Production in High Energy Hadron Collisions"	R.F. Schwitters
"Aspects of the Dynamics of Heavy-Quark Systems"	M.E. Peskin
"Heavy Particle Spectroscopy and Dynamics B's to Z's"	M.G.D. Gilchriese
"... And for Our Next Spectroscopy?"	J. Ellis
"Review of the Physics and Technology of Charged Particle Detectors"	A.H. Walenta
"A Review of the Physics and Technology of High-Energy Calorimeter Devices"	P.M. Mockett

1984 THE SIXTH QUARK

"The Last Hurrah for Quarkonium Physics: The Top System"	E. Eichten
"Developments in Solid State Vertex Detectors"	C.J.S. Damerell
"Experimental Methods of Heavy Quark Detection"	T. Himel
"Production and Uses of Heavy Quarks"	G.L. Kane
"Weak Interactions of Quarks and Leptons: Experimental Status"	S. Wojcicki
"The Experimental Method of Ring-Imaging Cherenkov (RICH) Counters"	T. Ekelöf
"Weak Interactions of Quarks and Leptons (Theory)"	H. Harari

PIEF-FEST

"Pief"	S. Drell
"Accelerator Physics"	R.R. Wilson
"High Energy Theory"	T.D. Lee
"Science and Technology Policies for the 1980s"	F. Press
"Inclusive Lepton-Hadron Experiments"	J. Steinberger
"Forty-Five Years of e^+e^- Annihilation Physics: 1956 to 2001"	B. Richter
"We Need More Piefs"	J. Wiesner

1985 SUPERSYMMETRY

"Introduction to Supersymmetry"	J. Polchinski
"Signatures of Supersymmetry in e^+e^- Collisions"	D. Burke
"Properties of Supersymmetric Particles & Processes"	R.M. Barnett
"Supersymmetry: Experimental Signatures at Hadron Colliders"	P. Darriulat
"Superworld/Hyperworlds"	M.E. Peskin
"Very High Energy Colliders"	B. Richter
"Some Issues Involved in Designing a 1 TeV (c.m.) $e^{\pm}$ Linear Collider Using Conventional Technology"	G.A. Loew
"Wake Field Accelerators"	P.B. Wilson
"Plasma Accelerators"	R.D. Ruth & P. Chen
"Collider Scaling and Cost Estimation"	R.B. Palmer

1986 PROBING THE STANDARD MODEL

"Electroweak Interactions–Standard and Beyond"	H. Harari
"CP and Other Experimental Probes of Electroweak Phenomena"	B. Winstein
"Perturbative QCD: K-Factors"	R. Field
"Experimental Tests of Quantum Chromodynamics"	J. Dorfan
"Phenomenology of Heavy Quark Systems"	F. Gilman
"Heavy Quark Spectroscopy and Decay"	R. Schindler
"Some Aspects of Computing in High Energy Physics"	P. Kunz
"Data Acquisition for High Energy Physics Experiments"	M. Breidenbach

DRELL-A-BRATION

"Topics in Arms Control"	M. Goldberger
"Compositeness Below 1 TeV"	R. Jaffe
"Our Changing View of Hadron Structure"	J. Bjorken

1987 LOOKING BEYOND THE Z

"Theory of e^+e^- Collisions at Very High Energy"	M.E. Peskin
"Prospects for Physics at e^+e^- Linear Colliders"	G.J. Feldman
"Physics with Polarized Electron Beams"	M.L. Swartz
"Electron-Proton Physics at HERA"	G. Wolf
"Hadron Colliders Beyond the Z^0"	C. Quigg
"Physics at Hadron Colliders (Experimental View)"	J.L. Siegrist
"The Dialogue Between Particle Physics and Cosmology"	B. Sadoulet
"Requirements for Detectors at SSC"	M.G.D. Gilchriese

1988 PROBING THE WEAK INTERACTION: CP VIOLATION AND RARE DECAYS

"Results on b-Decay in e^+e^- Collisions, with Emphasis on CP Violation"	K. Berkelman
"Precious Rarities–On Rare Decays of K, D and B Mesons"	I. Bigi
"Superconducting Detectors for Monopoles and Weakly Interacting Particles"	B. Cabrera
"Cosmic Relics from the Big Bang"	L. Hall
"Double Beta Decay"	M.S. Witherell
"Neutrino Masses and Mixings"	L. Wolfenstein
"The Bottom Quark: A Key to 'Beyond Standard' Physics"	H. Harari
"b-Physics in Fixed Target Experiments"	J. Sandweiss
"Experimental Searches for Rare Decays"	A.J.S. Smith

1989 PHYSICS AT THE 100 GeV MASS SCALE

"Heavy Quarks-Experimental"	R. Hollebeek
"The Theory of Heavy Flavour Production"	R.K. Ellis
"Precision Experiments in Electroweak Interactions (Experimental)"	M. Swartz
"Theory of Precision Electroweak Measurements"	M. Peskin
"Applications of QCD to Hadron-Hadron Collisions: Experimental"	M.D. Shapiro
"Applications of QCD to Hadron-Hadron Collisions: Theoretical"	I. Hinchcliffe
"W^+W^- Interactions and the Search for the Higgs Boson"	M.E. Levi
"Electroweak Symmetry Breaking: Higgs/Whatever"	M.S. Chanowitz
"Electron-Positron Storage Rings as Heavy Quark Factories"	R.H. Siemann
"Prospects for Next-Generation e^+e^- Linear Colliders"	R.D. Ruth
"Current Prospects for Hadron Colliders"	H. Edwards
"Hadron Colliders Beyond the SSC"	M. Tigner

1990 GAUGE BOSONS AND HEAVY QUARKS

"Z Decays and Tests of the Standard Model"	R. Cahn
"Future Experiments at LEP I and LEP II"	D. Treille
"Experimental Studies of Electroweak Gauge Bosons"	D. Burke
"Top Quark Topics"	G. Kane
"The Next Linear Collider"	R. Ruth
"Theoretical Aspects of B-Physics"	J. Bjorken
"Fundamental Aspects of B-Physics"	P. Drell
"B-Factory Storage Ring Design"	J. Rees

1991 LEPTON-HADRON SCATTERING

"A Historical Review of Lepton-Photon Scattering"	R. Taylor
"Uses of Superconductivity in Particle Accelerators"	R. Siemann
"Theoretical Aspects of Lepton-Hadron Scattering"	S. Drell
"Lepton-Hadron Scattering from Scaling Violation through HERA"	F. Sciulli
"Experimental Studies on QCD"	T. Hansl-Kozanecka
"HERA: The New Frontier"	J. Feltesse
"Topics in Calorimetry for High Energy Physics"	R. Hollebeek
"High Energy Lepton Hadron Scattering as a Probe of QCD"	R. Peccei

1992 THE THIRD FAMILY AND THE PHYSICS OF FLAVOR

"Predicting Quark and Neutrino Masses and Mixings"	L.J. Hall
"Experimental Techniques for B Physics"	N.A. Roe
"Experimental Aspects of CP Violation"	H.N. Nelson
"CP Violation"	Y. Nir
"Trigger and Data Acquisition at High-Rate Colliders"	A.J. Lankford
"Vertex Detectors"	V. Luth
"Tau Physics"	M.L. Perl
"Experimental Searches for the Top Quark"	R. Hollebeek
"Top-ology"	M. Peskin

1993 SPIN STRUCTURE IN HIGH ENERGY PROCESSES

"Spin, Mass, and Symmetry"	M.E. Peskin
"Physics with Polarized Z^0s"	P. Rowson
"Spin and Precision Electroweak Physics"	W.J. Marciano
"Polarized Electron Sources"	R. Prepost
"Polarization Phenomena in Quantum Chromodynamics"	S.J. Brodsky
"Polarized Lepton-Nucleon Scattering"	E. Hughes
"Polarized Targets in High Energy Physics"	G.D. Cates, Jr.
"Spin Dynamics in Storage Rings and Linear Accelerators"	J. Irwin
"Spin Formalism and Applications to New Physics Searches"	H.E. Haber

1994 PARTICLE PHYSICS, ASTROPHYSICS & COSMOLOGY

"Introduction to the Big Bang"	M. Turner
"Gravity Wave Experiments"	R. Weiss K. Thorne
"Neutrino Masses and Mixing"	S. Mishra
"Laser Technologies & Astrophysics"	R. Byer
"Dark Matter and Large Scale Structure"	J. Primack
"Searches for Baryonic Dark Matter via Microlensing"	C. Alcock
"Low-Noise Detectors in Astrophysics"	B. Sadoulet
"High-Energy Astronomy & Gravity"	E. Bloom
"Dark Matter Particle Candidates"	M. Srednicki

AUTHOR INDEX

Bula, C. 495

Chivukula, R. S. 375

Damerell, C. J. S. 103
Dornan, P. J. 525

Enomoto, R. 351

Fujii, K. 099

Hadley, N. J. 309
Hall, L. J. 261
Hewett, J. L. 187
Hughes, R. E. 291

Kasper, P. 461
Kühn, J. H. 001

Lang, K. 443
Langeveld, W. G. J. 479

Mallik, U. 385
Meier, K. 403
Menary, S. 547
Messner, R. 507

Peskin, M. 327
Plano, R. 427

Schalk, T. 329
Sinervo, P. K. 065
Solodov, E. P. 463
Sphicas, P. 567
Strom, D. 331
Stuart, L. M. 413
Swartz, M. 101

Wise, M. 569
Wood, D. 359

Ypsilantis, T. 225

PRIVATE ROAD
FOOT TRAFFIC
ONLY