

**Measurement of the Heavy Flavor Content of Jets in $p\bar{p}$
Collisions at $\sqrt{s} = 1.8$ TeV**

by

Phillip Todd Koehn

Submitted in Partial Fulfillment
of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor Paul L. Tipton
Department of Physics and Astronomy
The College
Arts and Sciences

University of Rochester
Rochester, New York

1996

Curriculum [REDACTED]

The author [REDACTED] He graduated with a Bachelor of Science degree in Physics from the University of Minnesota in 1990. He came to the University of Rochester in the fall of 1990 and began graduate studies in physics. He pursued his research in high energy physics under the direction of Professor Paul Tipton, and received the Master of Arts degree in Spring of 1992.

Acknowledgements

There are many people without whom, my thesis would not have been possible. First I must thank the members of the CDF collaboration and the Fermilab staffs for making the CDF experiment such a success. I am indebted to my advisor Paul Tipton for asking me to come and work with him. I am thankful for his guidance, friendship, and support. I am grateful to Richard who inspired this work, taught me how to do an analysis, and helped to keep it (and me) on track. I must thank Michelangelo for the many helpful discussions and insights into heavy flavor production. Thanks to Betsey and Steve for their support and for making life in Rochester enjoyable. I must also thank Fan, Brian, Kirsten, Rob, and Gordon, who made working in the CDF trailers and excursions to “Emilio’s” and “Pockets” so much fun. Thanks to Alan whose companionship and humor kept me sane (well, almost) at Rochester and at Fermilab. I would also like to thank my family for their love and encouragement. Finally, I am grateful to Ruth for her love and understanding throughout.

Abstract

Presented here is the first measurement of the bottom (b) and charm (c) flavor content of inclusive jets in collisions of protons and antiprotons at a center of mass energy of 1.8 TeV, using the Collider Detector (CDF) at the Fermilab Tevatron. Jets with displaced vertices are identified and the estimated proper decay-length distribution is fit to the expected distributions from b , c , and non-heavy flavor jets. The fraction of jets that contain at least one b or \bar{b} quark is approximately 0.02, while the fraction of jets containing at least one c or \bar{c} quark decreases from approximately 0.12 to 0.04 in the jet transverse energy range from 15 GeV to 135 GeV. Finally, the azimuthal correlations between b jets give information on the relative contributions of $b\bar{b}$ production processes.

Contents

1	Introduction	1
1.1	The Standard Model	1
1.2	Bottom and Charm Production	5
1.3	Overview of Analysis	14
2	Experimental Apparatus	19
2.1	The Tevatron	19
2.2	The Collider Detector at Fermilab	21
2.2.1	The Tracking Detectors	22
2.2.2	Calorimetry	27
2.2.3	The Trigger System	30
3	The Event Samples	32
3.1	The Data	32
3.1.1	Inclusive Jet Samples	32
3.1.2	The Inclusive Central Electron Sample	33
3.2	The Monte Carlo	38
3.2.1	Monte Carlo Event Generators	38
3.2.2	Detector Simulation	39

4 Tagging Heavy Flavor Jets	48
4.1 The JETVTX Algorithm	54
4.2 The Bottom and Charm Tagging Efficiency	60
4.3 Tagging efficiency scale factor.	63
5 The Determination and Fitting of the Estimated Proper Decay Length Distributions	75
5.1 Fit Method	75
5.2 Fit Results	79
5.3 Fit Consistency Check	82
6 Results	86
6.1 Fraction of Bottom and Charm per Jet	86
6.2 Cross Checks	89
6.2.1 Double Tags	89
6.2.2 D^* Production in Jets	94
6.2.3 Inclusive Bottom and Charm Quark Cross Section	100
6.3 $b\bar{b}$ Production Correlations	102
6.4 Systematic Uncertainty	107
7 Conclusions	117
Bibliography	119

List of Figures

1.1	The fundamental particles that make up the Standard Model. Quarks and gluons carry color (Red, Green, and Blue), while the leptons, the W^\pm , Z^0 , and γ are colorless. The quarks and leptons also have antiparticle counterparts of the same mass but of opposite charge. Interactions between the fermions (quarks and gluons) are mediated by the second group, the bosons (W^\pm , Z^0 , gluons, and γ).	3
1.2	Examples of Feynman diagrams depicting heavy flavor production at leading order ($2 \rightarrow 2$ processes).	9
1.3	Examples of Feynman diagrams depicting virtual corrections to leading order heavy flavor production.	10
1.4	Examples of Feynman diagrams depicting next - to - leading order heavy flavor production ($2 \rightarrow 3$ processes).	11
1.5	The inclusive b cross sections from CDF, using several decay modes. Also shown are NLO QCD calculations for different choices of renormalization/factorization scales and choice of b -quark mass. (Note: 90% of b quarks have P_T greater than P_T^{min} . The rapidity, y_b , of the b -quarks, is defined as $\tanh^{-1}(\frac{p_z}{E})$.)	13

1.6	Next-to-leading order QCD calculation of the Ratio of b jets to inclusive-jet as a function of E_T distributions, with jet cone size of $R = 0.4$ (described in Chapter 3), for different choices of renormalization and factorization scales.	15
1.7	Next-to-leading order QCD calculation of the Ratio of c jets to inclusive-jet E_T distributions, with jet cone size of $R = 0.4$ (described in Chapter 3), for different choices of renormalization and factorization scales.	16
2.1	Diagram of the Tevatron accelerator complex.	20
2.2	A side-view cross section of the CDF detector. The detector is forward-backward symmetric about the interaction region, located at the lower-right corner of the figure. The detector components are described in the text.	23
2.3	Transverse view of the CTC endplate illustrating the 9 superlayer geometry. The wire planes are tilted 45° relative to the radial to account for the lorentz angle of the ionization drift velocity.	24
2.4	Schematic of a single SVX barrel. The SVX detector is comprised of two barrels placed end to end, with their readouts at the outer ends of the barrels.	27
2.5	Diagram of a single central calorimeter wedge. Each wedge is comprised of 10 towers. A single tower spans $\Delta\eta \times \Delta\phi = 0.1 \times 15^\circ$. There are twelve wedges per arch in the central calorimetry system. There are a total of four arches.	29

3.1	The E_T spectrum for jets from the 50 GeV (upper plot) and the 20 GeV jet trigger samples (lower plot). Both distributions are normalized to unit area. The peak observed in both distributions is due to the trigger jet and the jets recoiling against them.	34
3.2	Signed impact parameter distribution for tracks in the jet_50 sample.	42
3.3	Quality Factor, Q , as a function of the number of obscured hits.	44
3.4	CTC track-finding efficiency as a function of Q for tracks in a mixture of jet data samples.	45
3.5	SVX tracking efficiencies as a function of track P_T and the density of the tracking environment. The upper plot shows the CTC-to-SVX linking efficiency as a function of track P_T . The lower plot shows the efficiency that tracks with three hits in the SVX pass the quality requirements (listed in the text) as a function of the number “nearby” tracks.	47
4.1	Simplified view of an event containing a secondary vertex, shown in the transverse $(r - \phi)$ plane.	49
4.2	Plot (a) is the distribution of the two-dimensional decay length, L_{xy} for a mixture B hadrons in jets with $E_T > 15$ GeV. Plot (b) is the P_T spectrum.	50
4.3	Plot (a) is the distribution of the two-dimensional decay length, L_{xy} for a mixture C hadrons in jets with $E_T > 15$ GeV. Plot (b) is the P_T spectrum.	51
4.4	Impact parameter significance for good quality SVX tracks from jet_50 data.	53
4.5	Impact parameter for good quality SVX tracks from b , c , and non-heavy flavor jet Monte Carlo.	54

4.6	Impact-parameter significance for good quality SVX tracks from b , c , and non-heavy flavor jet Monte Carlo.	55
4.7	Schematic showing a positive and negative two dimensional decay length, L_{xy} . Where L_{xy} is positive (negative) if the secondary vertex is reconstructed in the same (opposite) hemisphere as the jet.	57
4.8	The two-dimensional decay length, L_{xy} , for tags in b , c , and background jet Monte Carlo.	58
4.9	The two dimensional decay length, L_{xy} , for tagged jets in the jet_50 data sample. The vertical dashed line marks $L_{xy} = 0$	59
4.10	b and c jet tagging efficiencies as a function of E_T in the 50 GeV trigger jet Monte Carlo sample.	61
4.11	b and c jet tagging efficiencies as a function of E_T in the 20 GeV trigger jet Monte Carlo sample.	62
4.12	The jet_50 data and Monte Carlo taggable rate. The upper plot is the taggable rate as a function of jet E_T . The lower plot shows the associated scale factor.	66
4.13	The jet_20 data and Monte Carlo taggable rate. The upper plot is the taggable rate as a function of jet E_T . The lower plot shows the associated scale factor.	67
4.14	Scale factor calculated with weighted averages of the two tagging methods as a function of jet E_T . In the upper plot, the weighted average of the two tagging efficiencies in data (open points) is compared with Monte Carlo (solid points). The lower plot shows the scale factor. The solid line is the resultant constant fit, and the 2 dashed lines are the $\pm 1\sigma$ uncertainty in the fit.	71

4.15 The tagging efficiency as a function of jet E_T . A comparison of the two tagging methods for the inclusive electron data are shown in the upper plot. The lower plot shows the two methods in the $b\bar{b}$ Monte Carlo.	72
4.16 Comparison of tagged electron jet variables described in the text, Data (solid points) and Monte Carlo (histograms) are normalized to unit area.	73
5.1 The $c\tau$ distributions for HERWIG Monte Carlo b , c , and background tagged jets.	77
5.2 The $c\tau$ distribution in the jet_50 data (points) and fit (histogram) to the b , c , and background $c\tau$ distributions.	80
5.3 The $c\tau$ distribution in the jet_20 data (points) and fit (histogram) to the b , c , and background $c\tau$ distributions.	81
5.4 The three or more track $c\tau$ distribution in the jet_50 data (points) and fit (histogram) to the Monte Carlo b , c , and background.	84
6.1 The Fraction of b and c jets as a function of E_T . The data (points) are compared with HERWIG and PYTHIA predictions for b (lower contours) and c (upper contours). The open (solid) points indicate the jet_50 (jet_20) data. The uncertainties on data points are correlated. The inner error bars are the statistical uncertainty, the outer error bars are the statistical and the systematic uncertainties added in quadrature.	88

6.2	The next-to-leading order QCD calculation of the Ratio of b -jet to inclusive-jet E_T distributions, for different choices of renormalization and factorization scales ($\mu_R = \mu_F = \mu$) for jet cone sizes $R = 0.4$ and $R = 0.7$. The data points for $R = 0.4$ are shown with statistical uncertainty only.	90
6.3	Next-to-leading order QCD calculation of the Ratio of c -jet to inclusive-jet E_T distributions, for different choices of renormalization and factorization scales ($\mu_R = \mu_F = \mu$) for jet cone sizes $R = 0.4$ and $R = 0.7$. The data points for $R = 0.4$ are shown with statistical uncertainty only.	91
6.4	Schematic showing a double-tagged event. Double-tagged jet events are defined as two or more jets separately tagged in the same event.	92
6.5	The $c\tau$ distribution for jets in double-tagged events. The HERWIG Monte Carlo (histogram) $c\tau$ distribution for b jets is normalized to the data (points).	93
6.6	The $K\pi\pi - K\pi$ mass difference (ΔM). In the upper plot, the solid histogram is the distribution for D^* meson candidates, and the dashed histogram is background taken from the $K\pi$ (D^0) mass sideband. In the lower plot, the points are the D^* meson candidates, and the solid line is the fit. The total number of reconstructed D^* candidates is $143 \pm 28(\text{stat.}) \pm 20(\text{syst.})$	97
6.7	The inclusive b and c -quark cross sections measured with the $c\tau$ fit method compared with the NLO QCD predictions (for different choices of renormalization/factorization scales and choice of b -quark mass). The errors on the measured values for bottom and charm are the statistical and systematic uncertainties added in quadrature.	103

6.8 Examples of Feynman diagrams depicting heavy flavor production. Light quark-antiquark annihilation is shown in diagram (a), gluon-gluon fusion is depicted in diagrams (b) and (c). Diagrams (d) and (e) show examples of flavor excitation (initial state gluon splitting). Final state gluon splitting is depicted in diagrams (f) and (g).	104
6.9 The $\Delta\phi$ distribution for double-tagged $b\bar{b}$ Monte Carlo events. The figures show the $\Delta\phi$ distributions for the three production processes.	105
6.10 The $\Delta\phi$ distribution for double-tagged b data (points) events fit (histogram) to the sum of the three sub-processes from $b\bar{b}$ Monte Carlo.	106
6.11 The estimated proper lifetime distributions in jet_50 data tagged jets. The dashed histogram has the default SVX track χ^2 requirements. The solid histogram has the SVX χ^2 requirements degraded as described in the text. Both histograms are normalized to unit area.	110
6.12 Plots (a) and (b) are the bottom and charm Monte Carlo quark fragmentation (z) distributions. Plots (c) and (d) are the bottom and charm tagging efficiencies as a function of the fragmentation variable, z.	115

List of Tables

1.1	Coupling strength of forces described by the Standard Model.	4
2.1	Description of the charged particle tracking chambers.	26
2.2	Summary of CDF calorimetry subsystems.	28
3.1	Selection criteria for good electrons.	36
4.1	Monte Carlo b and c jet tagging efficiencies. The uncertainties are statistical only.	61
4.2	Summary of tags in data.	69
4.3	Summary of tags in $b\bar{b}$ Monte Carlo.	69
4.4	Comparison of tagging efficiencies in Monte Carlo and data. The uncertainty in the data efficiencies include both the error on F^B and the statistical error.	70
5.1	The $c\tau$ fit results for tagged jets. The uncertainties are from the fit error.	82
5.2	The $c\tau$ fit results for tagged jets containing three or more tracks. . .	84
5.3	The ratio of ≥ 2 - to - ≥ 3 track tagging efficiencies.	85

5.4	The b and c content of tagged jets obtained from the $c\tau$ fit of three or more tracks is used to predict the b and c content in tagged jets with two or more tracks.	85
6.1	The b and c -jet fractions integrated over the E_T range from 15 GeV to 135 GeV.	87
6.2	Monte Carlo b and c jet double-to-all tag ratios. The uncertainties are statistical only.	89
6.3	b jet Content of Double-Tagged Events. The uncertainties are statistical only.	90
6.4	Double-tag prediction of the number of tagged b jets. The uncertainties are statistical only.	94
6.5	Efficiencies for the acceptance calculation.	102
6.6	The relative fractions of each production mode determined by fitting the $\Delta\phi$ distribution of the double-tagged data, compared with HERWIG b Monte Carlo. The uncertainty for the data fit results is the fitter error. The uncertainty in the Monte Carlo fractions are the statistical error.	107
6.7	The tagging efficiencies for b and c tagged jets in Monte Carlo separated by production mode. Uncertainties are the statistical error. .	112
6.8	Summary of uncertainties (in %) in the F_b and F_c measurement. .	116

Chapter 1

Introduction

In recent years, high energy physicists have made substantial progress in understanding the microscopic physical universe: elementary particles and their interactions. Experimentally, elementary particles and interactions are studied by scattering particles at very high energies (equivalently, small distances), revealing for example, the quark and gluon structure of nucleons. The Standard Model [1, 2] is the theory that enumerates the fundamental constituents of matter and describes the forces through which they interact. Over the past two decades, predictions made by the Standard Model theory have been very successful in describing the experimental data.

1.1 The Standard Model

The Standard Model describes the interaction between the fundamental particles. The four forces through which these interactions occur are the strong force, the electromagnetic force, the weak force, and the gravitational force. Since gravity is extremely weak compared to the other forces, it may be neglected. In the

Standard Model, interactions are described by two gauge invariant theories; the Electroweak model that includes Quantum Electrodynamics (QED), and Quantum Chromodynamics (QCD) [3]. Gauge theories are a special class of quantum field theories where an invariance principle specifies the existence of interactions among particles. QCD, with a $SU(3)$ gauge symmetry, describes the strong interaction. The unification of the electromagnetic and weak interactions forms the $SU(2)_L \times U(1)$ Glashow-Weinberg-Salam electroweak theory.

The Standard Model contains the fundamental particles: quarks, leptons, and gauge bosons (see table 1.1). Quarks obey Fermi-Dirac statistics (fermions with spin = 1/2) and experience both the electroweak and strong interactions. There are six quarks whose existence has been experimentally verified; up (u), down (d), charm (c), strange (s), top (t), bottom (b). Quarks are organized in three families or doublets. Each quark also has a strong force color charge of Red (R), Green (G), or Blue (B). This is analogous the electric charge in the electromagnetic force. Leptons, also fermions, only participate in electroweak interactions, therefore carry no color charge. The leptons appear in three families: The electron (e), muon (μ), and tau (τ), paired with their associated neutrinos (ν_e, ν_μ, ν_τ). The e, μ , and τ all have mass and electric charge $-1|e|$. Neutrinos are electrically neutral and are consistent with being massless.

Gauge bosons, which obey Bose-Einstein statistics (bosons have integral spin), are particles that “mediate” or carry the forces described by the Standard Model. The strong, electromagnetic, and the weak forces are mediated by the gluon (g), photon (γ), vector bosons (W^\pm and Z^0), respectively. The coupling, α , is a dimensionless measure of the strength of an interaction, and is dependent upon the momentum transfer, q^2 , between the interacting particles. In Table 1.1 are listed the approximate asymptotic values of the coupling for the strong, electromagnetic, and weak forces. Because of the non-abelian nature of the strong force, gluons

Fermions			Bosons		
Quarks	Q/e	Weak Isospin			
$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$	$\begin{pmatrix} +\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} W^+ \\ Z^0 \\ W^- \end{pmatrix}$		
Leptons			8 Gluons	Higgs	
$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -\frac{1}{2} \\ +\frac{1}{2} \end{pmatrix}$	(g)	(h)	Photon (γ)

Figure 1.1: The fundamental particles that make up the Standard Model. Quarks and gluons carry color (Red, Green, and Blue), while the leptons, the W^\pm , Z^0 , and γ are colorless. The quarks and leptons also have antiparticle counterparts of the same mass but of opposite charge. Interactions between the fermions (quarks and gluons) are mediated by the second group, the bosons (W^\pm , Z^0 , gluons, and γ).

carry color charge and can therefore interact among themselves as well as quarks. This will have implications upon the short (large q^2) and long range (small q^2) behavior of the strong force (discussed below). In addition, there is a spin-zero Higgs boson. Not yet observed, the Higgs boson is required by the electroweak theory to explain the masses of fundamental particles.

Through the strong interaction, quarks form colorless bound states called hadrons. Two classes of hadrons are observed; mesons and baryons. Mesons are bosons consisting essentially of one quark and one antiquark, bound together through the exchange of gluons. For example, the π^+ meson is composed of $u\bar{d}$. Baryons are fermions consisting of three quarks, also bound together by gluons. Protons and neutrons are baryons composed of uud and udd quarks respectively. Other mesons and baryons are composed of different combinations of quark flavors.

QCD is a remarkable theory that exhibits two complimentary behaviors: asymp-

Table 1.1: Coupling strength of forces described by the Standard Model.

Interaction	Typical Coupling	Typical Range
Strong	$\alpha_s \sim 1$	$\sim 1 \text{ } F$
Electromagnetic	$\alpha \sim 10^{-2}$	∞
Weak	$\alpha_W \sim 10^{-6}$	$\sim 10^{-3} \text{ } F$

tic freedom and confinement. The coupling, α_s , is dependent upon the interaction momentum transfer scale (q^2). The q^2 dependence of α_s is approximately

$$\alpha_s(q^2) = \frac{12\pi}{(33 - 2N_f)\ln(q^2/\Lambda_{QCD}^2)} \quad (1.1)$$

N_f is the number of quark flavors, and Λ_{QCD} is a constant ($\sim 0.200 \text{ GeV}$) determined by experiment. The q^2 dependence of α_s causes the strength of the strong interaction to be very small ($\alpha_s(q^2) < 1$) at large q^2 (small distances $< 1 \text{ } F$). In the infinite q^2 limit, $\alpha_s \rightarrow 0$, quarks and gluons behave as if they were essentially free particles (asymptotic freedom). Conversely, when the momentum transfer is below $\sim (1 \text{ GeV})^2$, the coupling becomes large (~ 1). This means that for QCD, perturbation theory should work well at large q^2 (small distances), but break down at small q^2 (large distances) where $\alpha_s(q^2)$ is large. The scale beyond which perturbation theory roughly breaks down is Λ_{QCD} . It is believed that this is the reason why quarks and gluons are not observed in isolation, but are *confined* within hadrons.

Confinement implies that colored particles (quarks and gluons) manifest themselves as *jets* of hadrons. It is also believed that QCD has the property that the potential energy between two colored quarks increases linearly with the distance

between them. This means that it would take an infinite amount of energy to separate the participating quarks. Consequently, quarks are *confined* to hadronic color-neutral bound states. For example, in a high energy hadron-hadron collision, a quark from one hadron can scatter from a quark (or gluon) in the other. As the quark moves away from the quark (or antiquark) in the hadron to which it was bound, the energy in the color field between the quarks increases. Soon the energy will be large enough such that it is favorable to create quark-antiquark pairs out of the vacuum. These new quark-antiquark pairs then combine with each other and with the original interacting quarks producing color neutral (color singlet) hadrons in the final state. This process continues until the original interaction energy is dissipated, producing a multitude of hadrons called a *jet*. The direction of the jet will be approximately collinear with the quarks (gluons) that initiated it. Therefore experimentally, what is observed are jets of hadrons, and by studying jets one can learn about quarks and gluons. The process through which a quark or gluon becomes a jet of hadrons is called fragmentation.

1.2 Bottom and Charm Production

A measurement of jets containing bottom and charm quarks (heavy-flavor jets) is interesting both theoretically and experimentally. Such a measurement is a good test of perturbative QCD. In principle, systematic effects present in other QCD-data comparisons can be controlled or eliminated. By comparing the heavy-flavor jet P_T [4], as opposed to the heavy flavor quark P_T , infinities in the QCD calculation are controlled. By measuring the *ratio* of heavy-flavor jet production to all jets, uncertainties in parton-distribution functions (described below) and q^2 scales largely cancel. In addition, this measurement identifies heavy-flavor jets by the long lifetimes of bottom and charm hadrons, thus they are not dependent on

specific decay modes of these hadrons.

The signature of a heavy-quark jet is important to other physics processes. The work presented in this thesis helps identify that signature and explain possible backgrounds. For example, the top quark is detected through the production of a W boson and *b*-flavored jets (*b* jet). Signals for new physics such as an intermediate mass Higgs and supersymmetric partners to top may be detected by the presence of *b* jets in the final state. Charm quark structure functions may be studied in events containing a photon and a *c* jet. Strange quark structure functions may be probed in the production of a W boson and a *c* jet [5]. There are also interesting discrepancies between data and QCD theory in charm production. For example, at CDF, the prompt ψ' (an excited state of the $c\bar{c}$ hadron J/ψ) rate is higher than the theoretical prediction by a factor of ~ 50 at low P_T [6]. Present theoretical attempts (such as the color octet model) have failed to give an adequate explanation. Is the hadronic production of charm in agreement with QCD?

Bottom and charm quarks produced in high energy proton-antiproton ($p\bar{p}$) collisions are described by the QCD parton model [7]. As previously stated, the proton and antiproton are made up of point-like quarks and gluons which are sometimes referred to as partons. In the proton (antiproton), the partons are the light valence quarks uud ($\bar{u}\bar{u}\bar{d}$), gluons, and a sea of virtual quark-antiquark pairs. In a $p\bar{p}$ collision, a single parton in the proton will collide with a single parton in the antiproton. The remaining partons or “spectators” do not participate in the collision.

The QCD cross section for bottom and charm production is given by the short-distance partonic cross section convoluted with the momentum distributions of the

incoming partons:

$$\sigma_{p\bar{p}} = \sum_{i,j} \int dx_p dx_{\bar{p}} \sigma_{ij \rightarrow \hat{b}\bar{b}(c\bar{c})X} f_i^p(x_p, \mu) f_j^{\bar{p}}(x_{\bar{p}}, \mu) \quad (1.2)$$

The momentum distributions or “structure functions”, $f_{i(j)}^{p(\bar{p})}$, give the probability that a parton i (j) has a fraction x_p ($x_{\bar{p}}$) of the proton (antiproton) momenta. Structure functions are determined empirically in deep (large momentum transfer) inelastic scattering experiments where a lepton is scattered off a nucleon [8]. The evolution of the structure functions with the momentum transfer, q^2 , is calculated, for example, by using the Alterelli-Parisi equations [8]. In Equation 1.2 above, the structure functions are evaluated at the fragmentation/renormalization scale $q^2 = \mu^2$. This arbitrary scale, μ^2 , serves to separate the long and short distance effects of QCD. In heavy flavor production, a typical choice of scale is $\mu = \sqrt{m_{b(c)}^2 + P_{T,b(c)}^2}$. The short distance QCD partonic cross section, $\sigma_{ij \rightarrow \hat{b}\bar{b}(c\bar{c})X}$, is the cross section for the interaction where parton i interacts with j to produce $b\bar{b}$ ($c\bar{c}$). The term X refers to the rest of the partons in the event. The short distance cross section is calculated in a perturbation series in $\alpha_s(\mu^2)$. The final state $b\bar{b}$ ($c\bar{c}$) quarks will fragment and appear as jets in the detector. Quarks with masses less than $\Lambda_{QCD}(\sim 0.200)$ GeV are considered light (u, d, s), and will be symbolized generically as q (\bar{q}). Quarks with masses larger than Λ_{QCD} are considered heavy (c, b, t). The masses of the c , b , and t quarks are ~ 1.5 , ~ 4.5 , and 175 GeV/ c^2 [9] respectively. Hereafter, the term “heavy flavor” shall be used to refer to bottom and charm quarks (and antiquarks), and will be symbolized generically as Q (\bar{Q}).

The short distance QCD cross section for heavy flavor production is calculated as perturbative series expansion in α_s . Each term in the series is evaluated by computing the relevant matrix elements, represented by Feynman diagrams, for a given order. The lowest order (LO) term is of order α_s^2 , and the next-to-leading

order term is of order α_s^3 . A typical scale used to determine α_s is the heavy quark mass. The massive quark scale ensures that the perturbation calculation is well behaved.

The lowest order heavy flavor production is given by light quark-antiquark annihilation and gluon fusion:

$$q\bar{q} \rightarrow Q\bar{Q}$$

$$gg \rightarrow Q\bar{Q}$$

The Feynman diagrams for these processes are depicted in Figure 1.2. The theory of such events at lowest order perturbative QCD are well understood up to a choice in scale and choice of structure functions [10].

The next-to-leading order (NLO) heavy flavour production will include virtual corrections to the leading order processes and include the $2 \rightarrow 3$ processes:

$$q\bar{q} \rightarrow Q\bar{Q}g$$

$$gq(\bar{q}) \rightarrow Q\bar{Q}q(\bar{q})$$

$$g \rightarrow Q\bar{Q}\bar{q}$$

$$gg \rightarrow Q\bar{Q}g$$

Examples of virtual corrections to leading order are depicted in Figure 1.3. The $2 \rightarrow 3$ processes of gluon emission, initial-state gluons splitting (flavor excitation), and final-state gluon splitting is shown respectively in Figures 1.4a, 1.4b, 1.4c.

The contribution of NLO ($2 \rightarrow 3$) processes for heavy flavor production are not only expected to be non-negligible, but can be dominant at $p\bar{p}$ collider energies [11], (energies much higher than the heavy quark mass). At higher momentum transfers, the gluon structure functions are large and gluon-gluon scattering becomes

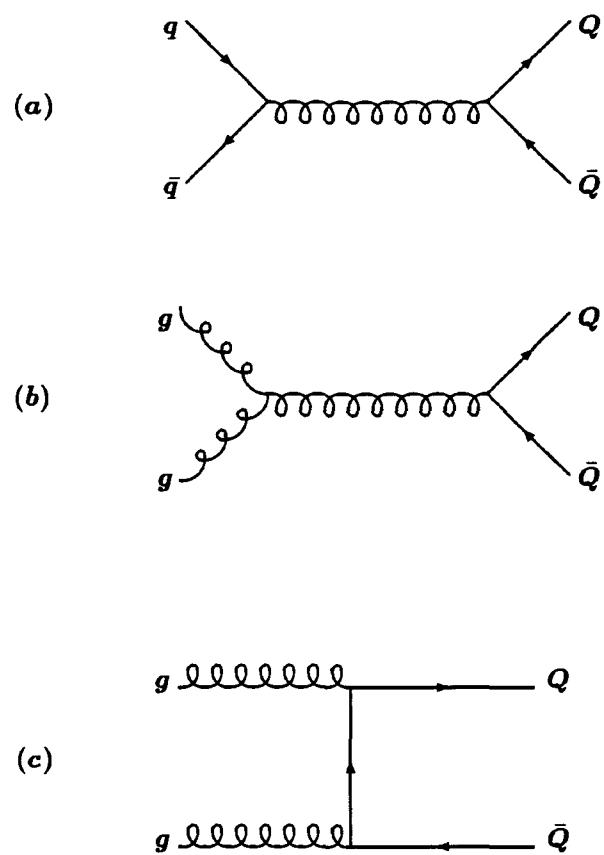


Figure 1.2: Examples of Feynman diagrams depicting heavy flavor production at leading order ($2 \rightarrow 2$ processes).

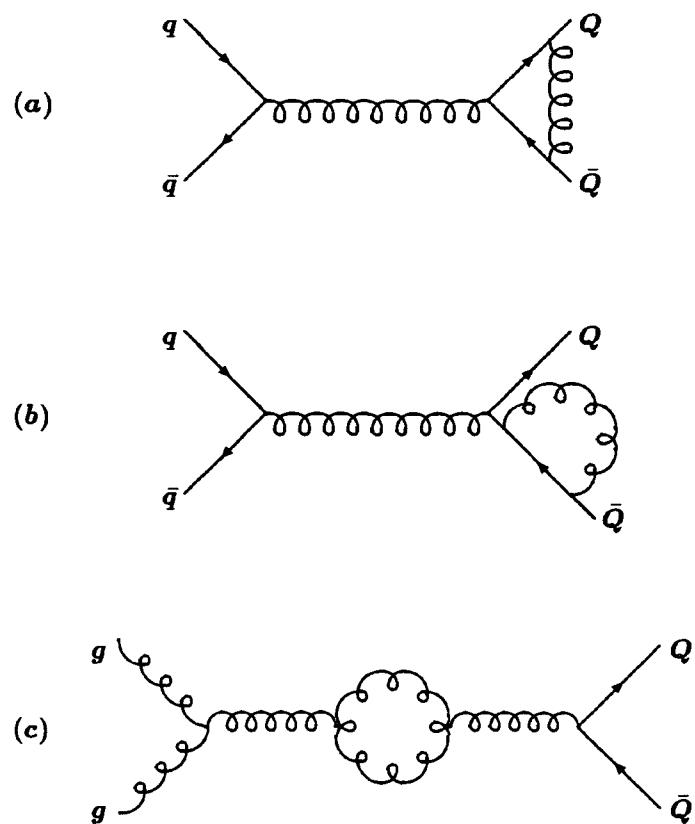


Figure 1.3: Examples of Feynman diagrams depicting virtual corrections to leading order heavy flavor production.

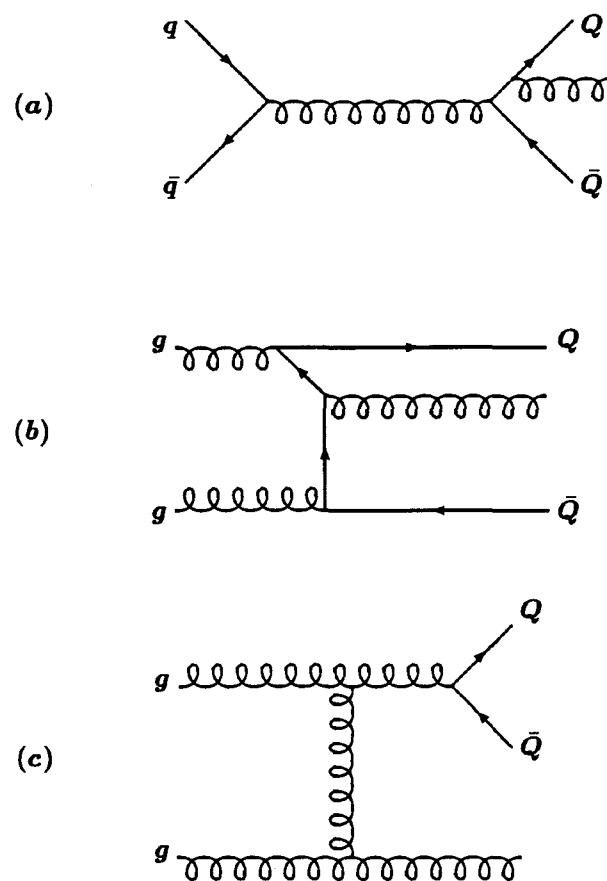


Figure 1.4: Examples of Feynman diagrams depicting next - to - leading order heavy flavor production ($2 \rightarrow 3$ processes).

important. This means that in a high energy regime, processes where a gluon fragments or “splits” into a heavy quark pair in the initial state (“flavor excitation”) or the final state can contribute a significant amount to the inclusive cross section. The QCD cross section for $gg \rightarrow gg$ is more than two orders of magnitude larger than $gg \rightarrow Q\bar{Q}$. Because of this, the estimated ratio of the $2 \rightarrow 3$ process of $gg \rightarrow gQ\bar{Q}$ to the $2 \rightarrow 2$ process of $gg \rightarrow Q\bar{Q}$ process can be as high as 5 - 10 [12].

The $O(\alpha_s^3)$ calculations for heavy flavor quark production have been performed [13] [14]. Figure 1.5 shows inclusive the b (\bar{b})-quark cross section ($p\bar{p} \rightarrow b(\bar{b})X$) measured at CDF [15] compared to the QCD prediction calculated to NLO. While the shape of the theoretical prediction is in good agreement with the data points, the overall normalization of the default theory curve is lower than the data by about a factor of two. Other NLO QCD comparisons with experimental measurements for the bottom cross section show similar results [16].

Recently, the calculation of heavy-flavor jet production has been performed [17] to NLO in QCD. A heavy flavor jet is defined as a jet containing at least one b (c) quark, or at least one anti- b (c) quark, or both the b (c) and the anti- b (c) quark in the same jet. Throughout the rest of this work, the two species of heavy flavor jets will be referred to as “ b jets” and “ c jets”. Figures 1.6 and 1.7 show the predicted production rates of heavy-flavor jets from reference [17]. In this calculation, the default value of the renormalization/factorization scale is $\mu_0 = \sqrt{m_{b(c)}^2 + E_{T,b(c)}^2}$, and the b - and c -quark masses used are $m_b = 4.75 \text{ GeV}/c^2$ and $m_c = 1.5 \text{ GeV}/c^2$. This calculation is potentially more precise for heavy-flavor *jet* variables compared with similar variables for heavy-flavor *quarks*. For example, at high momentum, large logarithmic singularities, caused by hard (high momentum) collinear gluon emission, appear in the perturbative expansion of the quark P_T distribution. These collinear logarithms need to be resummed. Such collinear logarithms do not occur in the calculation of the heavy flavor jet E_T distribution. This is because the jet

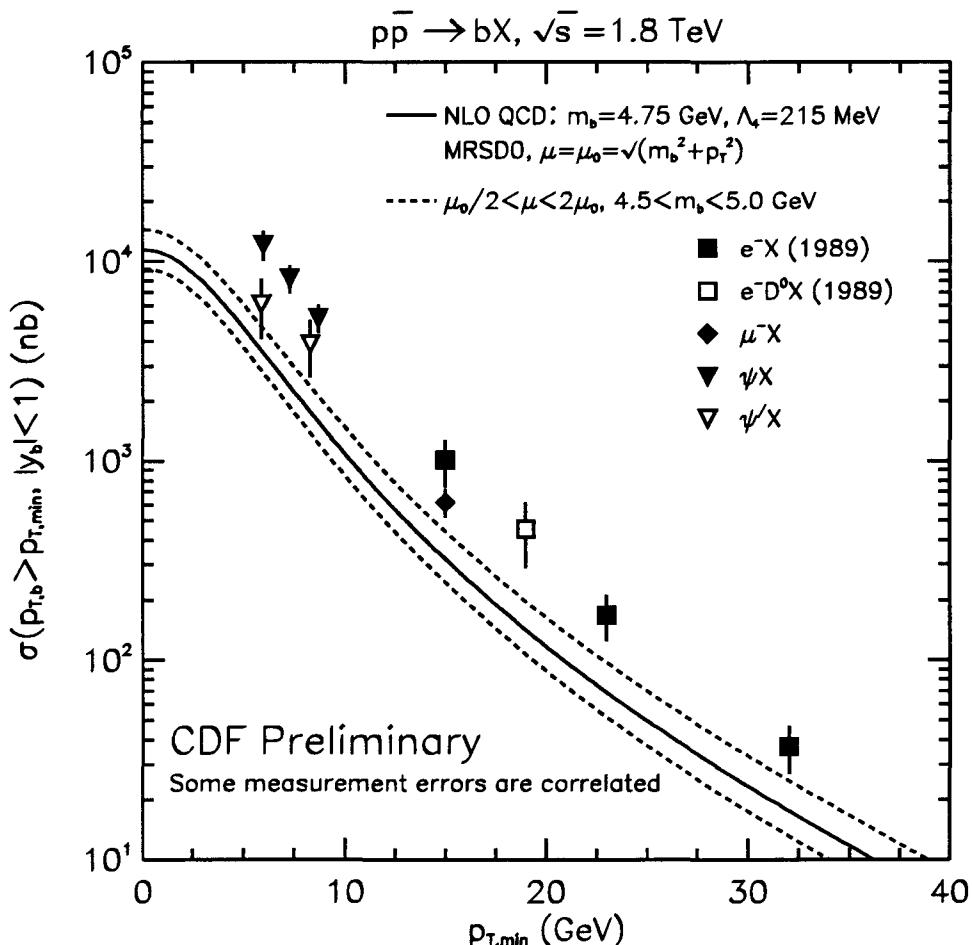


Figure 1.5: The inclusive b cross sections from CDF, using several decay modes. Also shown are NLO QCD calculations for different choices of renormalization/factorization scales and choice of b -quark mass. (Note: 90% of b quarks have P_T greater than P_T^{min} . The rapidity, y_b , of the b -quarks, is defined as $\tanh^{-1}(p_z/E)$.)

E_T does not depend on the energy sharing between the quark and the collinear gluon.

1.3 Overview of Analysis

In this physics analysis the heavy flavor content in an inclusive jet sample is extracted by identifying the decay of bottom and charm hadrons within the jet. This is done by searching for a secondary decay vertex displaced from the $p\bar{p}$ interaction point (called the primary vertex). By fitting the estimated proper decay-length or $c\tau$ (where c is the speed of light in vacuum, and τ is the proper lifetime) distribution in data to a sample of jets generated using a Monte Carlo simulation, the production rate of heavy-flavor jets is measured.

Bottom and charm quarks produced in high energy $p\bar{p}$ collisions will fragment to form jets containing B and C hadrons (along with a number of other hadrons). The average B -hadron proper decay length is $\sim 450 \mu\text{m}$ [18] and the average for a C hadron is $\sim 220 \mu\text{m}$ [19]. Since bottom and charm hadrons have long average $c\tau$, their decay vertices can be resolved from the primary vertex with the CDF detector. Bottom and charm hadrons in jets with energy greater than 15 GeV will travel as far as 1 to 2 mm from the primary vertex before they decay. In addition, because of the differences in their average proper decay lengths, the $c\tau$ distribution is used to distinguish b -flavored jets from c -flavored jets.

A secondary vertexing algorithm is used to reconstruct the heavy-flavor decay vertex within the jet by constraining charged particle tracks consistent with heavy flavor decay to a common vertex. The two-dimensional distance, the decay length L_{xy} , between the primary and secondary vertex is then determined. If a secondary vertex is found and if L_{xy} is large compared to its uncertainty, the jet is referred to as *tagged*. An estimate of the proper decay length is determined from L_{xy} .

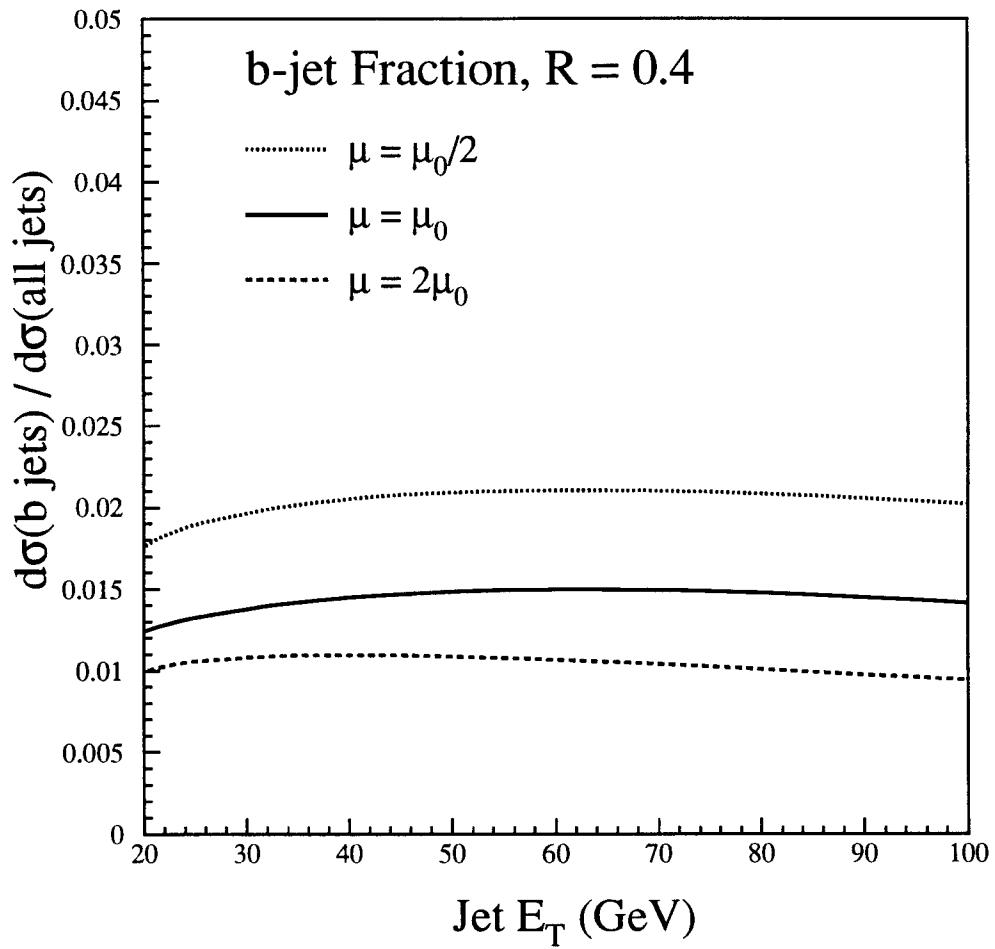


Figure 1.6: Next-to-leading order QCD calculation of the Ratio of b jets to inclusive-jet as a function of E_T distributions, with jet cone size of $R = 0.4$ (described in Chapter 3), for different choices of renormalization and factorization scales.

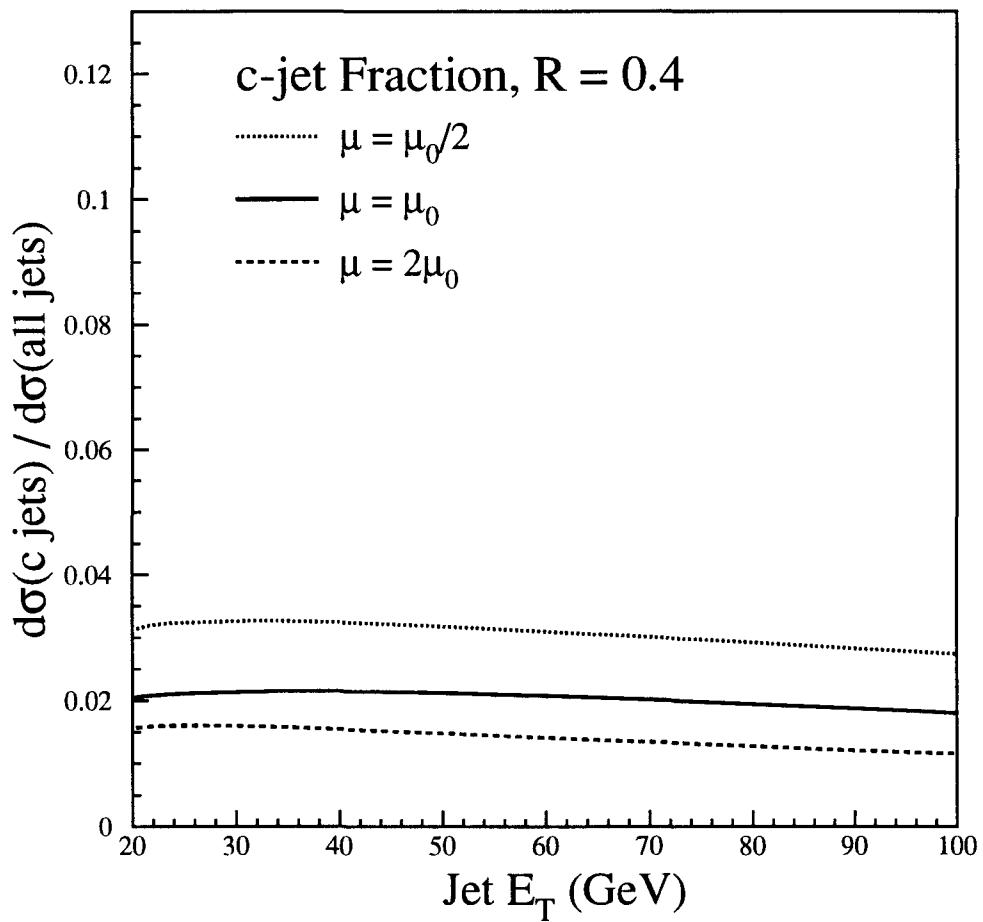


Figure 1.7: Next-to-leading order QCD calculation of the Ratio of c jets to inclusive-jet E_T distributions, with jet cone size of $R = 0.4$ (described in Chapter 3), for different choices of renormalization and factorization scales.

A control sample with a known b content, independent of the tagging algorithm, is used to test the tagging efficiency. B-flavored jets tagged in this sample serve also as a check on the shape of the estimated proper decay length distribution.

To obtain the fraction of tagged jets that come from b , c , and non-heavy flavor jet sources, the $c\tau$ distribution in the data is simultaneously fit to a sum of the expected $c\tau$ distributions from b , c , and non-heavy jets. Once the number of tagged b and c jets in the data are obtained from the fit, the fractions of b jet and the c jet fractions are determined by scaling the fit number of heavy-flavor jets by their respective tagging efficiency and dividing by all jets in the sample.

Several checks of these results have been performed using the data. Events with two separately tagged jets (double tags) are used to extract the b content. A Monte Carlo model predicts that the double tags found in data are primarily from $b\bar{b}$ production. By scaling the number of data double tags by the ratio of single tags to double tags obtained from a Monte Carlo $b\bar{b}$ simulation, the resulting b -jet fraction agrees well with that found with $c\tau$ fit method. The c -jet fraction is checked using the rate of fully reconstructed $D^{*\pm}$ (charm-flavored) mesons in jet data. The number of $D^{*\pm}$ candidates expected in the sample is calculated using the measured values for charm and bottom jet fraction from the $c\tau$ fit. The expected number is consistent with $D^{*\pm}$ rate measured in data. Finally, the inclusive bottom and charm-quark cross sections are determined using the $c\tau$ fit method. The results compare well with other CDF data and are consistent with QCD predictions to NLO.

A complementary analysis of double-tagged events is used to explore aspects of the $b\bar{b}$ production dynamics. The opening angle between tagged b jets is used to study the contributions of LO and NLO processes in $b\bar{b}$ production.

Finally, all systematic uncertainties in the measurement of the bottom and charm jet fraction caused by uncertainties in $c\tau$ distributions, fitting, and tagging efficiencies are studied.

Chapter 2

Experimental Apparatus

2.1 The Tevatron

The Fermilab Tevatron is a synchrotron storage ring 1 km in radius operating with 6 bunches of 900 GeV protons colliding with 6 bunches of 900 GeV antiprotons. The produced collisions have a total center-of-mass energy of 1.8 TeV. A diagram of the Fermilab Tevatron accelerator is shown in Figure 2.1.

At Fermilab, protons (p) are accelerated in a series of steps to attain an energy of 900 GeV. The protons start in a hydrogen glass bottle. They are then accelerated with two electrons attached to form H^- ions to 750 KeV in the Cockcroft Walton. After the electrons are stripped off, the protons are accelerated to 400 MeV in a 150 m linear accelerator (Linac). From the Linac, the protons are sent to a circular accelerator (a synchrotron called the “Booster”) of circumference 475 m, where they reach an energy of 8 GeV. The protons are then fed into the Main Ring (a synchrotron of 6300 m circumference), where they are accelerated to 150 GeV. Following the Main Ring, the protons are transferred into the superconducting accelerator called the Tevatron where they are accelerated to 900 GeV.

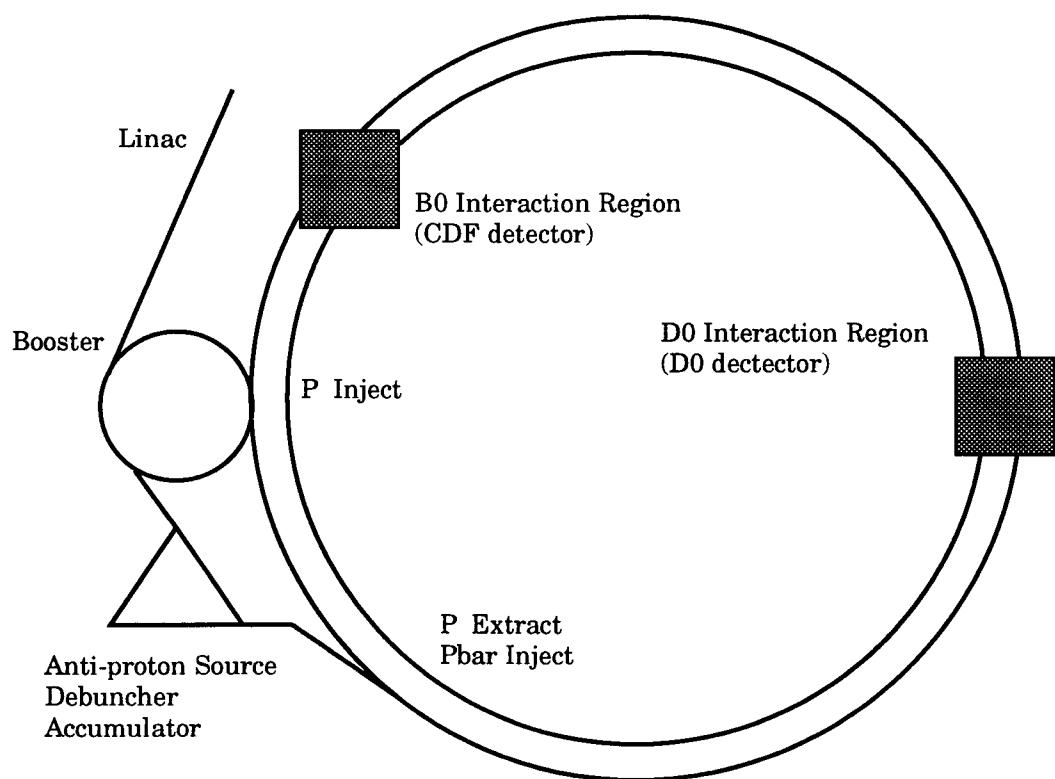


Figure 2.1: Diagram of the Tevatron accelerator complex.

Antiprotons (\bar{p}) are produced by diverting a sample of protons from the Booster, and accelerating them in the Main Ring up to an energy of 120 GeV. These protons are extracted and focussed onto a copper target to produce antiprotons. With momentum near 8 GeV, antiprotons are sent to the Debuncher and Accumulator for stochastic cooling to reduce the momentum spread. The antiprotons are “stacked” until $\sim 6 \times 10^{11}$ are collected, then are reinjected into the Main Ring until they reach 150 GeV, and are injected into the Tevatron. In counter-rotating helical orbits, both the p and \bar{p} bunches are accelerated in the same magnetic and RF fields. Quadrupole magnets are used to focus the bunches to a transverse diameter of $\sim 35 \mu\text{m}$. The entire process produces $\sim 10^{12}$ protons and $\sim 10^{11}$ antiprotons. More detailed information about the Tevatron can be found in references [20] [21]. The instantaneous luminosity (L) of the Tevatron can be obtained with the following equation:

$$L = \frac{N_p N_{\bar{p}} B f_0}{4\pi\sigma^2} \quad (2.1)$$

where N_p is the total number protons per bunch (1.2×10^{11}), $N_{\bar{p}}$ is the total number antiprotons per bunch (3.1×10^{10}), B is number of bunches of each type (6), f_0 is the frequency of bunch revolution (47.7 kHz), and σ^2 is the cross-sectional area of the bunches ($\sim 5 \times 10^{-5} \text{ cm}^2$). For the 1992-1993 Tevatron run (Run IA), the typical and best instantaneous luminosities are $5.4 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ and $9.2 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$, respectively [21].

2.2 The Collider Detector at Fermilab

The Collider Detector at Fermilab (CDF) is a general purpose detector designed to measure the energy, momentum, and identify electrons, muons, and photons [23]. It also identifies the sprays of particles called jets, which come from quarks and

gluons created from the proton anti-proton collisions. It is a cylindrical detector, approximately symmetric about the interaction point, with a central barrel region, plug regions capping the barrel at both ends, and two forward regions (see Fig.2.2). The detector uses electromagnetic (EM) and hadronic (HAD) calorimetry systems to sample the energy of particles. Charged-particle tracking chambers immersed in a magnetic field measure the momentum of particles. The CDF coordinate system is defined with the positive z axis pointing along the beamline in the direction of the protons, θ as the polar angle, and ϕ as the azimuthal angle. Pseudorapidity (η) is defined by the relation $\eta = -\ln(\tan(\theta/2))$.

2.2.1 The Tracking Detectors

The charged-particle tracking system at CDF consists of several tracking detectors within a 1.4 Tesla magnetic field oriented along the proton beam direction. The field is created by a 4.8 m long superconducting solenoid of radius 1.5 m.

The Central Tracking Chamber (CTC) is a large drift chamber that measures the transverse momentum of a charged particle by determining the curvature of its path in the magnetic field. The CTC is constructed of 84 drift layers organized into 9 superlayers: five axial layers are interspersed with four stereo layers. The drift wires of the five axial layers run parallel to the beamline, whereas the wires in the four stereo layers are at an angle of 3° to the beamline. Figure 2.3 shows a transverse view of the CTC endplate. The combination of the alternating axial and stereo layers provides 3 dimensional ($r - \phi - z$) tracking. For tracks with $|\eta| < 1.0$, the track finding efficiency is greater than 95%. The physical properties of the CTC are listed in Table 2.1.

A vertex time projection chamber (VTX) is primarily used to determine the z position of each event vertex. It is located within the inner region of the CTC. The

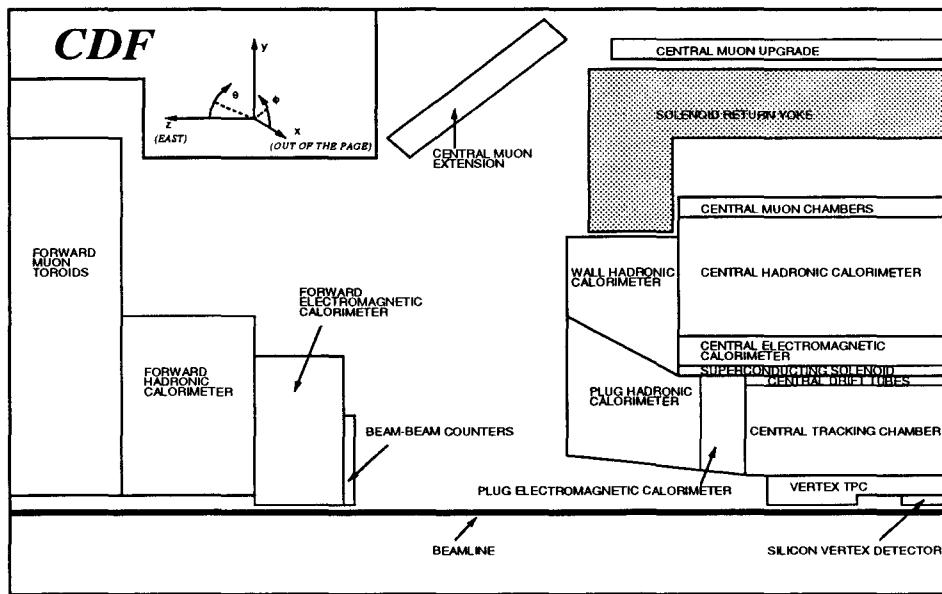


Figure 2.2: A side-view cross section of the CDF detector. The detector is forward-backward symmetric about the interaction region, located at the lower-right corner of the figure. The detector components are described in the text.

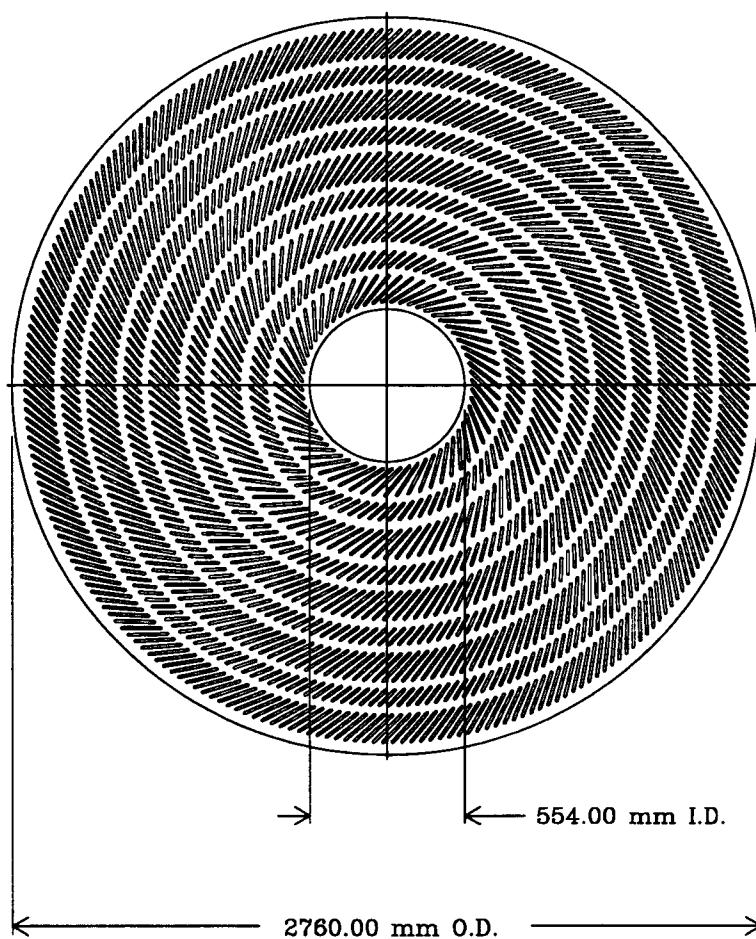


Figure 2.3: Transverse view of the CTC endplate illustrating the 9 superlayer geometry. The wire planes are tilted 45° relative to the radial to account for the lorentz angle of the ionization drift velocity.

VTX is comprised of 8 octagonal modules of time-projection chambers with sense wires running radially outward from the beamline. The VTX provides tracking information in the $r - z$ plane in a pseudorapidity range of $|\eta| < 3.5$. The uncertainty in the measurement of the z coordinate position of the primary event vertex is $\sigma_z = 1$ mm. Table 2.1 summarizes the physical characteristics of the VTX.

The silicon vertex detector (SVX) is a silicon strip microvertex detector that provides precise $r - \phi$ information for the reconstruction of charged particle tracks. It is located with its center on the nominal CDF interaction point, mounted within the VTX. The SVX consists of two barrels that are aligned end-to-end along the beam axis with a 2.15 cm gap between the two active regions at $z = 0$. Each barrel (Fig. 2.4) is composed of four concentric layers of silicon strip detectors. These detector layers are arranged in twelve 30° wedges. Each wedge holds four 25.5 cm long ladders, one per layer. The total number of ladders is $12(\text{wedges}) \times 4(\text{layers}) \times 2(\text{barrels}) = 96$. Each ladder has three 8.5 cm long single-sided silicon wafers, with readout strips running parallel to the beam line. The strips have a 60 μm pitch on the inner three layers, and 55 μm on the outer layer. The SVX single hit resolution is 13 μm , and has a hit efficiency of 96% per layer. The physical properties of the SVX are summarized in Table 2.1.

Combining SVX and CTC tracking information results in an impact parameter resolution that approaches 15 μm for high momentum tracks. The impact parameter is distance of closest approach of a track projected back to the primary $p\bar{p}$ interaction point. The momentum resolution of the SVX/CTC combination is $\delta P_T/P_T = \sqrt{(0.0009 P_T)^2 + (0.0066)^2}$ (GeV/c). Up to a few GeV/c, the P_T resolution in the SVX/CTC is dominated by multiple scattering in the silicon measuring planes and in the material between the CTC and the SVX. In the CTC, the intrinsic resolution (200 μm) dominates over the multiple scattering effects in the gas.

	Central tracking chamber (CTC)	Vertex tracking chamber (VTX)	Silicon vertex detector (SVX)
Polar Angle Coverage	$ \eta < 1.5$	$ \eta < 3.25$	$ \eta < 1.2$
Inner, Outer Radii (cm)	30.9, 132.0	8, 22	2.7, 7.9
Length (cm)	320	280	26
Layers	60 axial, 24 stereo	24	4
Wire/Strip Spacing	10 mm	6.3 mm	60 μm (inner 3 layers) 55 μm (outer layer)
Spacial Resolution	$200\mu\text{m}$ ($r - \phi$) 4 mm (r-z)	200-500 μm (r-z)	15 μm ($r - \phi$)
Momentum Resolution	$\delta P_T / P_T^2 = 0.002$		$\delta P_T / P_T^2 = 0.001$
Thickness	$\approx 0.015X_0$	$\approx 0.0045X_0$	$\approx 0.035X_0$

Table 2.1: Description of the charged particle tracking chambers.

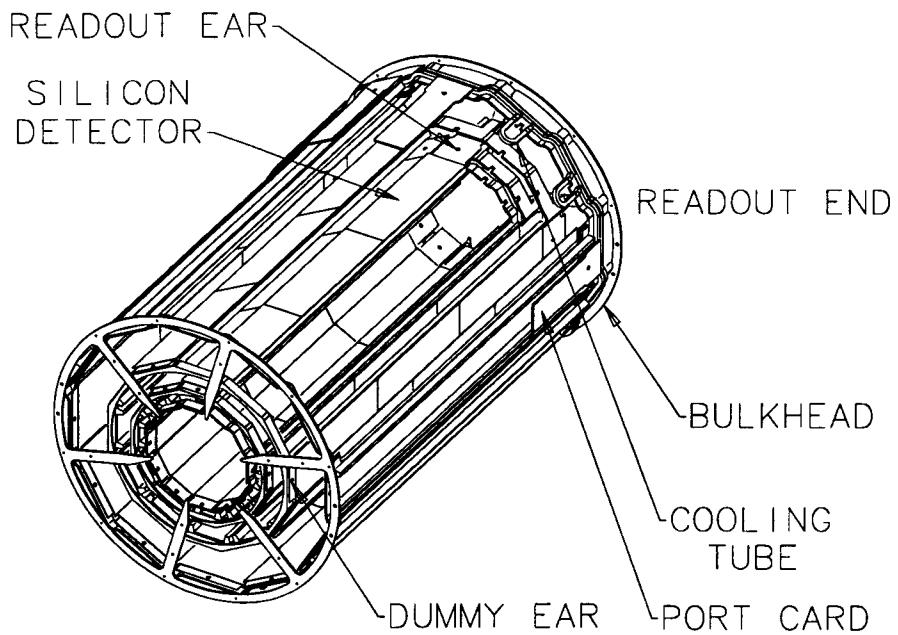


Figure 2.4: Schematic of a single SVX barrel. The SVX detector is comprised of two barrels placed end to end, with their readouts at the outer ends of the barrels.

2.2.2 Calorimetry

CDF employs electromagnetic (EM) and hadronic (HAD) calorimeters to sample the energy of particles and jets. The calorimetry surrounds the solenoid and tracking detector volume, covers a range 2π in azimuth and -4.2 to 4.2 in pseudorapidity. Segmented in η and ϕ , the calorimetry is arranged in a projective tower geometry pointing back to the interaction point. The calorimetry consists of three subsystems separated into regions of pseudorapidity: the central ($|\eta| < 1.1$), plug ($1.1 < |\eta| < 2.4$), and forward ($2.4 < |\eta| < 4.2$).

The central calorimeter is segmented into a series of wedges, each covering 15° in ϕ , containing an electromagnetic (CEM) section followed by a hadronic (CHA) section. A CEM wedge is shown in Figure 2.5. The electromagnetic section is comprised of alternating layers of a lead absorbing medium and a scintillator active medium. The hadronic section is comprised of alternating layers of iron and scin-

System	Coverage ($ \eta $)	Segmentation ($\Delta\eta \times \Delta\phi$)	Energy resolution (GeV)	Thickness
CEM	< 1.1	$0.1 \times 15^\circ$	$13.7\%/\sqrt{E_T} \oplus 2\%$	$18 X_0$
CHA	< 0.9	$0.1 \times 15^\circ$	$50\%/\sqrt{E_T} \oplus 3\%$	$4.5 \lambda_0$
WHA	$0.7 - 1.3$	$0.1 \times 15^\circ$	$75\%/\sqrt{E_T} \oplus 4\%$	$4.5 \lambda_0$
PEM	$1.1 - 2.4$	$0.1 \times 5^\circ$	$22\%/\sqrt{E_T} \oplus 2\%$	$18 - 21 X_0$
PHA	$1.3 - 2.4$	$0.1 \times 5^\circ$	$106\%/\sqrt{E_T} \oplus 6\%$	$5.7 \lambda_0$
FEM	$2.2 - 4.2$	$0.1 \times 5^\circ$	$26\%/\sqrt{E_T} \oplus 2\%$	$25 X_0$
FHA	$2.4 - 4.2$	$0.1 \times 5^\circ$	$137\%/\sqrt{E_T} \oplus 3\%$	$7.7 \lambda_0$

Table 2.2: Summary of CDF calorimetry subsystems.

tillator. Within the CEM, wire proportional strip chambers (CES) are embedded at approximately the position of maximum electromagnetic shower development (6 radiation lengths). The CES provides shower-position measurements in both the z and $r - \phi$ views. The physical properties of the calorimetry are summarized in Table 2.2.

The plug and forward calorimetry consists of electromagnetic (PEM, FEM) and hadronic (PHA, FHA) sections. The physical characteristics of the plug and forward are summarized in Table 2.2. The PEM is made up of 34 layers of lead absorber sheets sandwiched between conductive plastic proportional tube arrays. Each layer has a set of cathode pads and anodes that are read out. The cathode pads are arranged to form towers for the calorimeter segmentation. The PHA, located immediately behind the PEM, uses proportional tubes sandwiched between iron sheets. The forward calorimeters also use proportional chambers with cathode readout.

Used for the detection of muons, the central muon system (CMU) consists of four layers of drift chambers located outside the central calorimetry system. Positioned behind the CMU, is the central muon upgrade (CMP) comprised of an

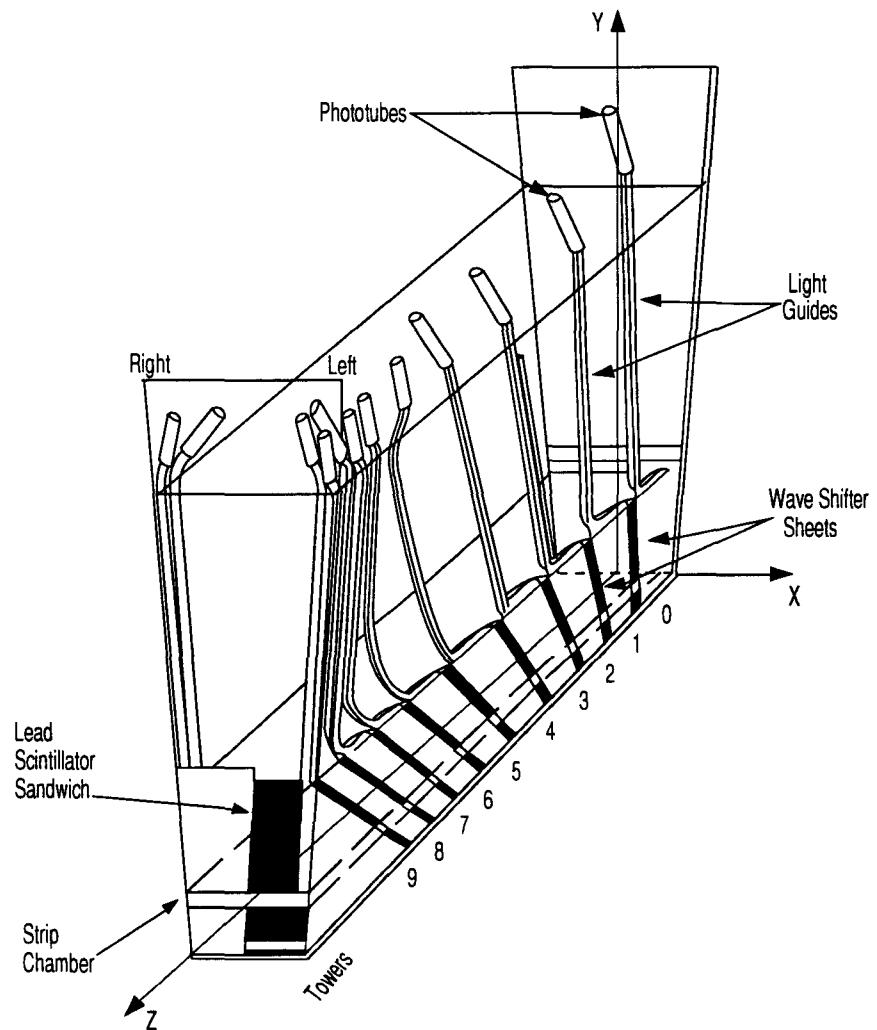


Figure 2.5: Diagram of a single central calorimeter wedge. Each wedge is comprised of 10 towers. A single tower spans $\Delta\eta \times \Delta\phi = 0.1 \times 15^\circ$. There are twelve wedges per arch in the central calorimetry system. There are a total of four arches.

additional 0.6 m of steel and an additional four layers of drift chambers. The central muon extension (CMX) is used for additional muon coverage and for triggering.

2.2.3 The Trigger System

The event Trigger for CDF is a 3 level system with the aim to maximize the number of interesting events written to tape. The trigger system is designed to select events with electrons, muons, or jets. With a bunch crossing rate of 280 kHz, at an average of ~ 1 interaction per crossing in Run IA, the trigger must reduce the output event rate to a few (3 - 5) Hz. This is the rate at which reconstructed events can be written to 8 mm tape. Trigger levels 1 and 2 are hardware triggers. Level 3 is implemented in software.

The lowest level trigger, level 1, requires calorimeter energy summed in ($\Delta\eta = 0.2 \times \Delta\phi = 15^\circ$) above a threshold, a transverse energy imbalance in the calorimetry, or hits in the muon chambers. The trigger takes 2 μ s to make the decision whether to accept the event. Since the beam crossings occur every 3.5 μ s, there is no dead-time incurred. The output event rate delivered to level 2 is a few kHz.

At level 2, the largest number of events are rejected. A hardware cluster finder forms clusters of calorimetry towers. For each cluster, E_T , $\langle \eta \rangle$, and $\langle \phi \rangle$ are determined. High momentum tracks are reconstructed in $r - \phi$ from hits in the CTC by the Central Fast Tracker. The momentum resolution of the tracks is $\delta P_T/P_T \approx 0.035 \times P_T$. Found tracks can be matched to clusters in the central electromagnetic calorimetry to form an electron candidate, or to track segments in the CMU, CMP, and CMX to form a muon candidate. The trigger decision time is 20 μ s, and the output event rate is ~ 12 Hz.

Level 3 is a software trigger that performs an almost full reconstruction of the event. It differs from the “offline” reconstruction in that it uses two dimensional

tracking instead of three. Level 3 consists of commercial processors (Silicon graphics “farm”) running at 1000 MIPS. The events are written out to 8 mm tape for “offline” processing at a rate of 5 - 10 Hz.

The instantaneous luminosity at CDF is monitored by counting minimum bias events with the beam-beam counters (BBC). A minimum bias event is a very weak or “soft” $p\bar{p}$ interaction that generates a spray of particles from the interaction point at small angles relative to the beam axis. These events make up the vast majority of interactions at the Tevatron. The BBC consists of a plane of scintillators mounted in front of the FMU that covers the forward and back regions of $0.32^\circ < \theta < 4.47^\circ$. The BBC is triggered when there is a coincidence of hits in the BBC with the crossing of particle bunches in the CDF detector. The instantaneous luminosity measured at CDF is found using the expression:

$$L(t) = \frac{N_{BBC}(t)}{\sigma_{BBC}} \quad (2.2)$$

$N_{BBC}(t)$ is the number of BBC hits at time t . The BBC cross section, σ_{BBC} , can be expressed as

$$\sigma_{BBC} = \sigma_{tot} \cdot \frac{N_{BBC}}{N_{tot}} \quad (2.3)$$

Where σ_{tot} is the total elastic and inelastic $p\bar{p}$ cross section, N_{BBC} is the total number of BBC triggered events, and N_{tot} is the total number of events. Integrating equation 2.2 over time gives the total luminosity delivered. The integrated luminosity for the 1992 run (Run IA) is 19.3 pb^{-1} with an uncertainty of 3.3% [22].

Chapter 3

The Event Samples

The data used in this analysis are inclusive jet samples and a sample of inclusive electrons. Monte Carlo samples use a detector simulation that models the geometry of the CDF detector, and uses a parameterization of the response observed in data to charged particles passing through the tracking chambers and jet energy deposition in the calorimetry, rather than a derivation of the response from first principles. The generation of Monte Carlo will be described below. Both the data and Monte Carlo must pass the same selection criteria and are processed with the same analysis code.

3.1 The Data

3.1.1 Inclusive Jet Samples

At CDF, a jet will show up as energy deposited in a “cluster” of calorimetry towers. The CDF clustering algorithm uses a cone of a fixed radius in $\eta - \phi$ space ($\Delta R = \sqrt{\eta^2 + \phi^2}$) to define a jet [24]. Event samples triggered on jets with nominal E_T thresholds of 20 and 50 GeV, corresponding to integrated luminosities of

0.038 pb⁻¹ and 0.655 pb⁻¹ (the jet_20 and jet_50 trigger samples, respectively). Jet energies are corrected as a function of P_T and η for calorimeter nonlinearities, uninstrumented regions, energy lost outside of the clustering cone, and energy gained in the underlying event (soft interactions of spectator quarks and gluons) [25]. In addition, all jets must have corrected $E_T \geq 15.0$ GeV, $|\eta| < 2.0$, and $|Z_{vertex}^{event}| < 30.0$ cm. After these requirements there are 429214 jets in ~ 220000 events in the jet_20 sample, and 272039 jets in ~ 120000 events in the jet_50 sample. Figure 3.1 shows the corrected E_T spectrum of the inclusive jet data samples.

3.1.2 The Inclusive Central Electron Sample

The Inclusive Central Electron sample (the ICE sample) is made up of low P_T electrons passing the selection criteria listed in Table 3.1. The following variables are used to select *good* electrons. The E/P requirement insures that the momentum (P) measured in tracking chambers balances the energy (E) measured in the electromagnetic calorimetry. Lateral energy sharing (LSHR) of calorimetry towers from an electron cluster must be consistent with test beam electron showers. This requirement eliminates clusters which are too wide, such as multiple particle backgrounds that could possibly fake an electron. The ratio of hadronic energy to electromagnetic energy, HAD/EM is used to select electrons. Electron showers are nearly fully contained in the electromagnetic (EM) calorimetry, and hadronic showers tend to deposit energy in both the electromagnetic and hadronic (HAD) sections. This variable is dependent on the physics because electrons from W boson decay, for example, will be more isolated than electrons from semileptonic bottom decay where there will be energy deposited in the hadronic section from the b jet. The δx and δz are the separations (in the $r - \phi$ and z view respectively) between the extrapolated CTC track position and electron shower position measured in the

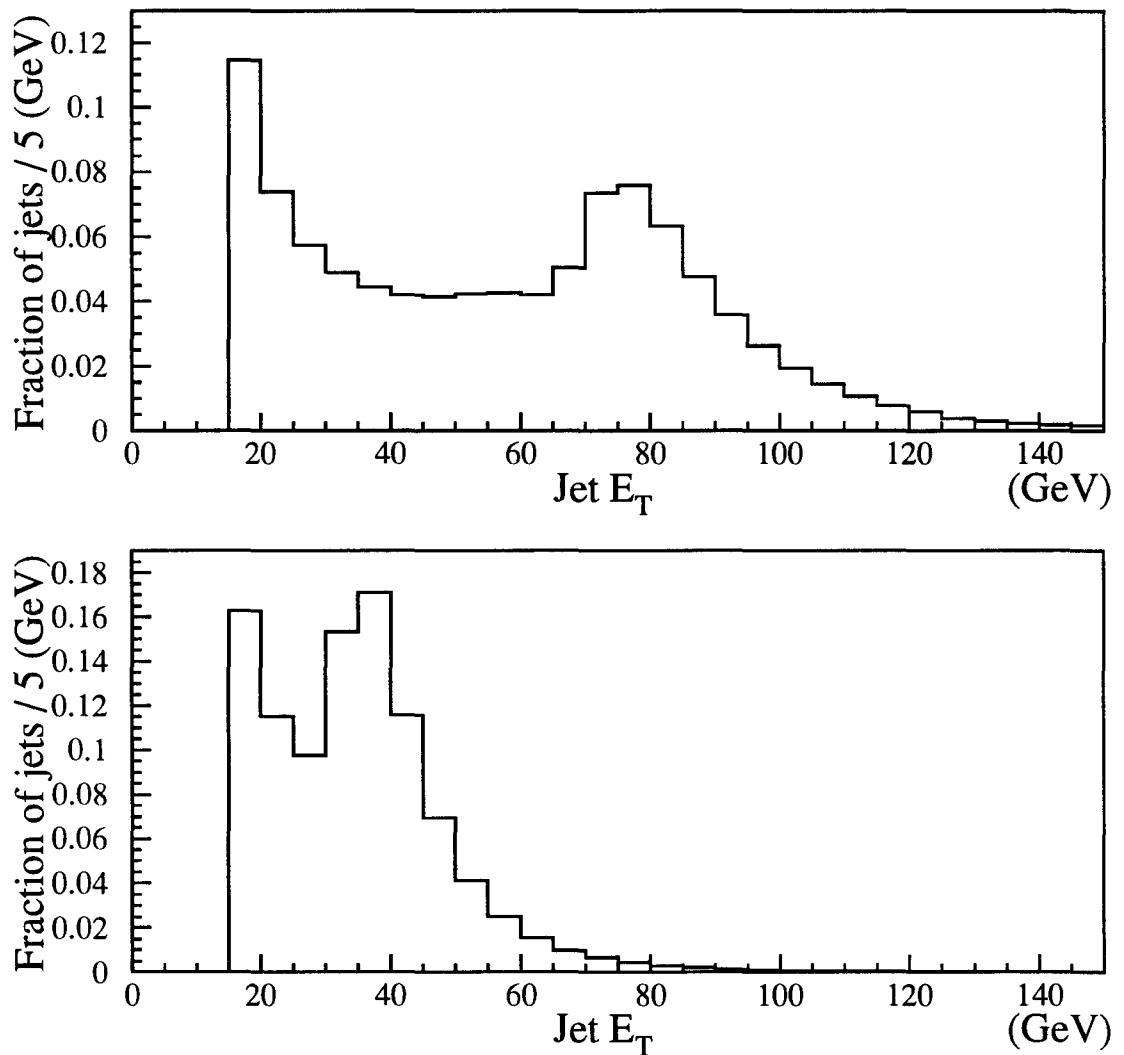


Figure 3.1: The E_T spectrum for jets from the 50 GeV (upper plot) and the 20 GeV jet trigger samples (lower plot). Both distributions are normalized to unit area. The peak observed in both distributions is due to the trigger jet and the jets recoiling against them.

CES. These “matching” variables are used to eliminate overlapping backgrounds from charged and neutral hadrons. The electron shower shape measured in the CES must be consistent with test beam electrons. A χ^2 test is used to compare the pulse height shape in the z -view (χ_{strip}) and the $r - \phi$ view (χ_{wire}). When a W boson decays to an electron or muon, the final state includes a charged lepton with high transverse momentum (P_T), and an imbalance of transverse energy (E_T) from the undetected neutrino called the *missing E_T* or \cancel{E}_T . Therefore, the \cancel{E}_T requirement removes electrons from the decay of W bosons. The z coordinate position is chosen such that the event vertex is near the center of the CDF detector. A fiducial requirement removes electrons near the detector boundaries. Finally, electrons produced by photons that convert as they traverse the detector material are removed. These electrons are identified by searching for an additional oppositely signed charged track near the electron track that extrapolates to a common tangent point. Electrons are labeled as conversions if they pass the following requirements: The separation from the nearby track in $r - \phi$ is less than 0.2 cm and the difference in $\cot(\theta)$ for the two tracks is less than 0.06.

The inclusive electron data sample is used to test the b -tagging algorithm because the sample contains a large fraction from semileptonic b -quark decay. The fraction of primary electrons coming from semileptonic b decay (F^B) in this sample is determined by looking for the muon from the semileptonic decay of the charmed meson (D) produced from the b -decay. In this search, the decay mode used is $\bar{B}_x \rightarrow e^- \bar{\nu}_e DX$, followed by $D \rightarrow \mu^+ \bar{\nu}_\mu X$, where the B_x is any flavor B meson or baryon, and D includes D^0, D^+, D_s mesons (The charge conjugate modes are implied). The charge of the muon from the charmed meson decay will be opposite to that of the electron from the B hadron decay. The invariant mass of the e and μ is kinematically limited to be less than the B mass. These two characteristics are used to extract a signal for this decay.

Good Electron Candidate Requirements	
$E_T^{\text{electron}} > 8.0$	GeV
$E/P < 1.5$	
$\text{LSHR} < 0.2$	
$\text{Had}/\text{Em} < 0.05$	
$\delta z < 3.0$	cm
$\delta x < 1.5$	cm
$\chi_{\text{strip}} < 10.0$	
$\chi_{\text{wire}} < 10.0$	
$E_T < 20.0$	GeV
$Z_{\text{vertex}}^{\text{event}} < 30.0$	cm
electron in fiducial region	
conversions removed	

Table 3.1: Selection criteria for good electrons.

Electrons are identified using the requirements listed above in Table 3.1. The muon identification requirements are listed below:

- $P_T^\mu > 2$ GeV/c.
- Hadronic energy in the muon tower < 5.0 GeV.
- Electromagnetic energy in the muon tower < 2.0 GeV.
- CMU-CMP matching requirements:

$$|\Delta X|/\sigma_X < 3.0$$

$$|\Delta Z|/\sigma_Z < 3.0$$

- Muon found in CMU or CMP.

The invariant mass is formed for all same-sign and opposite signed electron-muon ($e - \mu$) pairs. To insure that the $e - \mu$ pair is from the same B hadron, the

invariant mass is required to be less than $5 \text{ GeV}/c^2$. Because, one would expect no charge correlation in $e - \mu$ pairs from background, the amount of signal is determined by the observed number of excess opposite signed $e - \mu$ pairs. The total number of opposite sign $e - \mu$ pairs is 75 ± 8.7 , the number of same sign pairs is 18 ± 4.2 , resulting in an excess of 57 ± 9.6 . The uncertainty in these numbers are the statistical error.

Using the total number of excess opposite sign $e - \mu$ pairs, F^B is calculated with the expression:

$$F^B = \frac{N_{e-\mu}}{N_{tot}^e Br(D \rightarrow \mu X) \epsilon_l} \quad (3.1)$$

Where $N_{e-\mu}$ is the total number of excess $e - \mu$ pairs, $N_{tot}^e (= 18975)$ is the total number of electrons passing the cuts listed in Table 3.1, ϵ_l is muon reconstruction efficiency, and $Br(D \rightarrow \mu X) (= 0.103 \pm 0.003)$ [26] is the average branching ratio of charmed mesons decaying semileptonically. With ISAJET Monte Carlo (described below) using the above decay mode, the reconstruction efficiency is $\epsilon_l = 0.080 \pm 0.012(stat)$. From equation 3.1, the fraction of electrons from semileptonic b -decay is 0.372 ± 0.081 . The error is the statistical and systematic uncertainties added in quadrature.

The major contributions to the systematic uncertainty include the uncertainty in $Br(D \rightarrow \mu X)$, the reconstruction efficiency, and the background subtraction. $Br(D \rightarrow \mu X)$ contains an unknown mixture of D mesons whose decay constitutes the signal for this measurement (although the branching ratios for the individual modes are reasonably well measured). In addition, there is a contribution from the $D_s \rightarrow \mu X$ mode for which the branching ratio is not measured. Uncertainty in the tracking efficiency will affect the reconstruction efficiency of the muon. Furthermore there may be some local charge correlation not due to the cascade B decay, such as kaon and pion decay in flight. This contribution is believed to be small.

3.2 The Monte Carlo

3.2.1 Monte Carlo Event Generators

Multi-process Shower Monte Carlo event generators are used to simulate the physics of $p\bar{p}$ collisions in order to understand the characteristics of interesting physical processes, and to separate them from background events. These programs use QCD and electroweak theory to generate complete $p\bar{p}$ collisions that in principle may be treated in the same way as data for the purpose of analysis. The processes generated with these Monte Carlos are accurate only to leading order and use QCD cascade approximations to simulate higher orders.

The Monte Carlo event generators used in this analysis are HERWIG [27], PYTHIA [28], and ISAJET [29]. HERWIG (version 5.6) is a multi-purpose coherent, parton shower Monte Carlo, with cluster hadronization and an underlying event model based on data. This generator is used to create $b\bar{b}X$, $c\bar{c}X$, non-heavy flavor jet events with 20 GeV and 50 GeV jet trigger simulation. These samples are used to model the data jet samples in which the bottom and charm quark content is measured. To test the ability of the jet-tagging algorithm to identify jets with secondary vertices from B -hadron decay, the inclusive electron data sample is compared to a Monte Carlo sample of $b\bar{b}X$ jet events (the B hadron is forced to decay semileptonically 100% of the time). PYTHIA (version 5.6) uses JETSET [30] (version 7.3) to provide coherent final state showers, uses string hadronization and decays, and its underlying event model is based upon multiple parton scattering. This Monte Carlo is also used to produce samples of $b\bar{b}X$, $c\bar{c}X$, and non-heavy flavor jet events. ISAJET (version 7.0) is an incoherent parton shower program that uses independent hadronization and an underlying event model. This generator is used to produce events with the decay $B_x \rightarrow e^-\bar{\nu}_e DX, D \rightarrow \mu^+\bar{\nu}_\mu$ to

obtain the muon reconstruction efficiency described in the measurement of F^B above. ISAJET is also used to generate $b\bar{b}X$ and $c\bar{c}X$ jet events events forcing the decay: $D^{*\pm} \rightarrow D^0\pi^\pm$, $D^0 \rightarrow K^\mp\pi^\pm$. This Monte Carlo sample will be used to find the reconstruction efficiency of $D^{*\pm}$'s in jets in a cross check of the overall charm content. These “generator-level” physics events are passed on to software that simulates the CDF detector response to such events.

3.2.2 Detector Simulation

In this analysis, b and c jets are identified by searching for secondary decay vertices formed with good quality (described below) SVX tracks. Furthermore, the relative fractions of b and c jets are determined by using the shape of the estimated proper decay length distribution ($c\tau$) as a distinguishing variable. The impact parameter (the distance of closest approach between a track and the primary vertex) of the daughter tracks from a decaying object is directly proportional its proper decay length. Therefore an accurate modeling of the dense tracking environment (and subsequently the tracking variables such as impact parameter) in jet events is essential. The simulation of the CDF Detector consists of a jet and tracking model. To produce and cluster jets in the CDF detector, a model based upon the detector response measured from test beam data is used. The tracking simulation takes the resolutions of variables and the tracking efficiencies measured for good quality SVX tracks found in jet data, and applies them to the Monte Carlo charged particles.

The CDF Fast Simulation Package called QFL [31] (version 3.59) is used in conjunction with generator-level Monte Carlo events to produce and cluster jets in the CDF detector. QFL uses a parameterization of the detector response to jet energy deposition in the calorimeter, rather than a derivation of the response from

first principles. Stable particles from the generated event are transported through a cylindrical tracking chamber and charged particles are bent in the magnetic field. After exiting the tracking chamber, particles pass through the solenoid or the CTC endplate and enter the electromagnetic and hadronic calorimeters. The energy deposition in the calorimetry for generated particles of a given energy was parameterized with test-beam data.

The tracking simulation begins with a list from the Monte Carlo generated event containing the coordinates of the event vertex and momentum 3-vectors for all jets, B (C) hadrons, K_S^0 and Λ 's, their decay products, and particles from the underlying event. Using this information, one can calculate the five variables that fully describe the helical path of a charged particle (called a track) at CDF:

- $\cot(\theta)$: θ is the polar angle of the track.
- C : $1/(\text{radius of curvature})$ of the helical path of a charged particle traversing a magnetic field in (m^{-1}).
- ϕ : azimuthal angle of the track.
- Z_0 : z coordinate position of the track.
- d_0 : the distance of closest approach between the track and the primary event vertex.

For example, the transverse momentum, P_T , can be derived with the relation:

$$P_T = \frac{Q * 0.29979 * B}{|C|} \quad (3.2)$$

where B is the magnetic field of 1.4116 Tesla and Q is the charge of the particle. The uncertainties of $\cot(\theta)$, C , ϕ , and z are measured for *good* quality SVX tracks in jet data. For a SVX track to be of good quality, the following is required:

- $P_T \geq 2.0 \text{ GeV}/c$
- The minimum number of SVX layers with a hit (cluster of readout strips) is three.
- At least two of the SVX hits must have less than four strips in the cluster, no dead strips in the cluster, no hits shared with another track, and a cluster charge profile consistent with that of only one track.
- The maximum SVX χ^2 per degree of freedom is < 6 .
- The maximum track Z_0 relative to the primary vertex Z is 3.0 cm.

The Monte Carlo track variables (with the exception of the impact parameter) are smeared with a gaussian probability using the measured data uncertainty for σ of the distribution.

The impact parameter, d , for each Monte Carlo track is convoluted with a resolution function taken from the data. The d for each Monte Carlo track in the event is calculated with the relation:

$$\vec{d} = \frac{\vec{P}_T^{Track}}{|\vec{P}_T^{Track}|} \times \vec{L}_{xy} \quad (3.3)$$

\vec{L}_{xy} is the decay length of the parent hadron and \vec{P}_T^{Track} is the momentum of the daughter track in the plane transverse to the beam axis. The \vec{P}_T^{Track} is taken from the generator level Monte Carlo. The decay length is found with the expression $\vec{L}_{xy} = c\tau * \vec{P}_T^{Hadron}/M$. The parent hadron mass, M , is obtained from the particle data book [32]. The $c\tau$ of the parent hadron is generated at random with an exponential decay distribution where the decay constant is the lifetime of that specific hadron also taken from the particle data book. \vec{P}_T^{Hadron} is the generator-level Monte Carlo transverse momentum for the parent hadron. Figure 3.2 shows

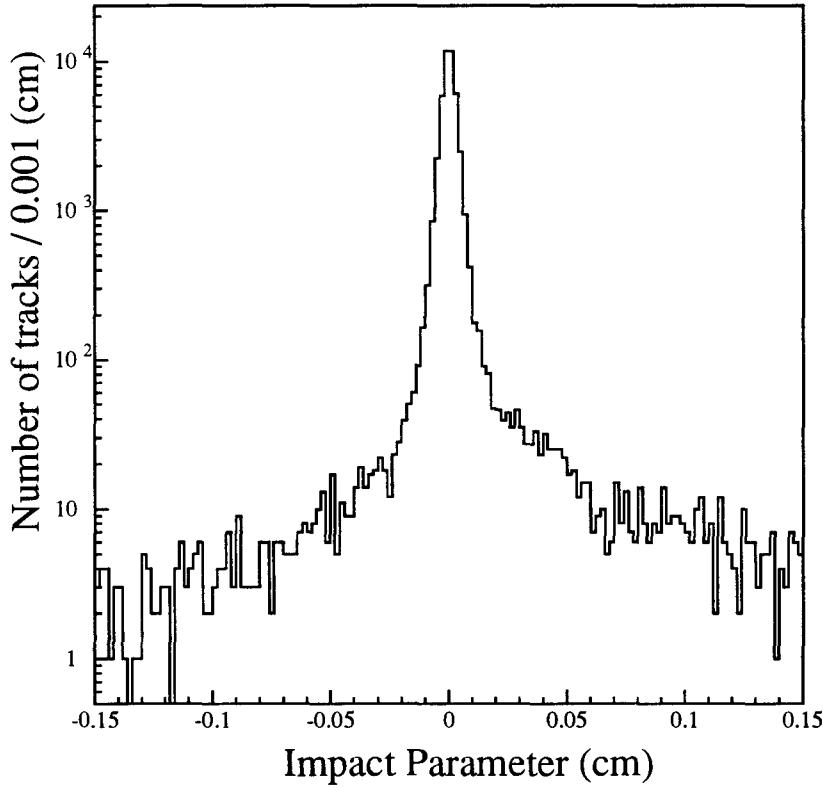


Figure 3.2: Signed impact parameter distribution for tracks in the jet_50 sample.

the impact parameter distribution for SVX tracks in jet data. The distribution shows that impact parameters from tracks associated with jets may be positive or negative. The sign of the impact parameter is determined by the angle it makes with the jet axis. If the angle is acute, the track originated from the same side of the event as the jet and the impact parameter is positive. If the angle is obtuse, the track originated from the opposite side of the event from the jet and the impact parameter is negative. Tracks in jets arising from zero lifetime objects have equal numbers of positive and negative impact parameters, whereas tracks from the decay of long lived objects inside a jet have positive impact parameters. Therefore the shape of the negative impact parameter distribution is a measure of the resolution, while the shape of the positive impact parameter distribution

contains both resolution and lifetime effects. The excess of positive over negative impact parameters shown in Figure 3.2 indicates the presence of tracks from long-lived objects. From the jet data, a list containing impact parameters, their uncertainty, and P_T , is made for good quality SVX tracks. The list is created such that it reflects the shape of the resolution function; it is made up of equal numbers of positive and negative impact parameters. For each Monte Carlo seed track, an impact parameter is randomly chosen from the data track list and added to the calculated impact parameter. However, this occurs only if the difference in P_T between the data and Monte Carlo track is less than 30% of the Monte Carlo P_T . The process continues until a track with a suitable P_T is found. The uncertainty in the impact parameter from the data track chosen from the list is simply assigned to the Monte Carlo track. These chosen elements of the data track list are used only once per simulation. This convolution method takes into account possible correlations from the impact parameter uncertainty and transverse momentum.

The CTC tracking efficiency model starts by checking the seed track to see whether it is within the CTC geometric acceptance. If not, the track is discarded. The CTC tracking efficiency, estimated in jet data samples, is parameterized as a function of the density of the tracking environment [33]. A “quality factor” variable, Q , inversely proportional to the track density, is used to characterize the tracking environment. Q is a function of the number obscured CTC wire hits due to nearby tracks that “overlap” the seed track. Figure 3.3 shows the variable Q as a function of obscured hits for Monte Carlo seed tracks embedded inside real data jet events. The CTC track-finding efficiency is estimated by first embedding Monte Carlo tracks into data jet events composed of a mixture of 20, 50, 70, and 100 GeV jet E_T thresholds. Next the event is processed with the standard CDF offline track reconstruction software. The efficiency with which the embedded track is found is parameterized as a function of the Q variable. Figure 3.4 shows the track finding

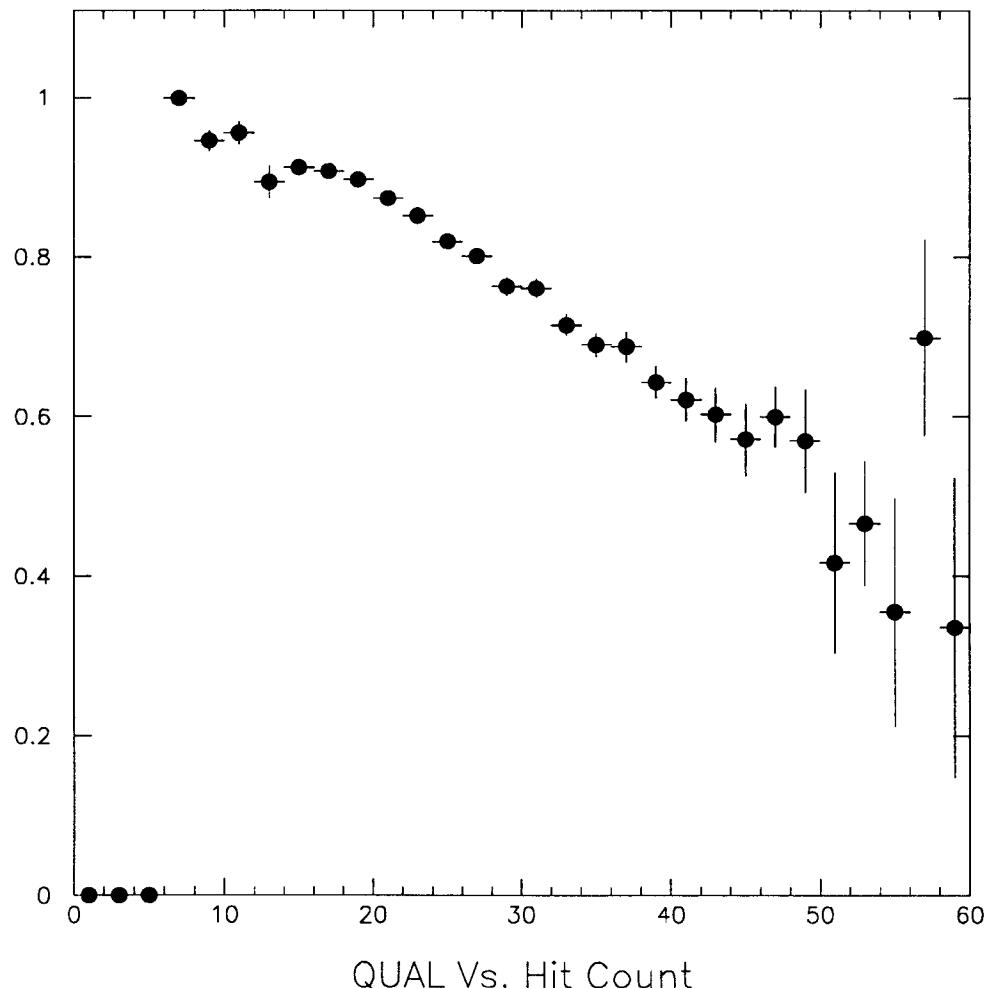


Figure 3.3: Quality Factor, Q , as a function of the number of obscured hits.

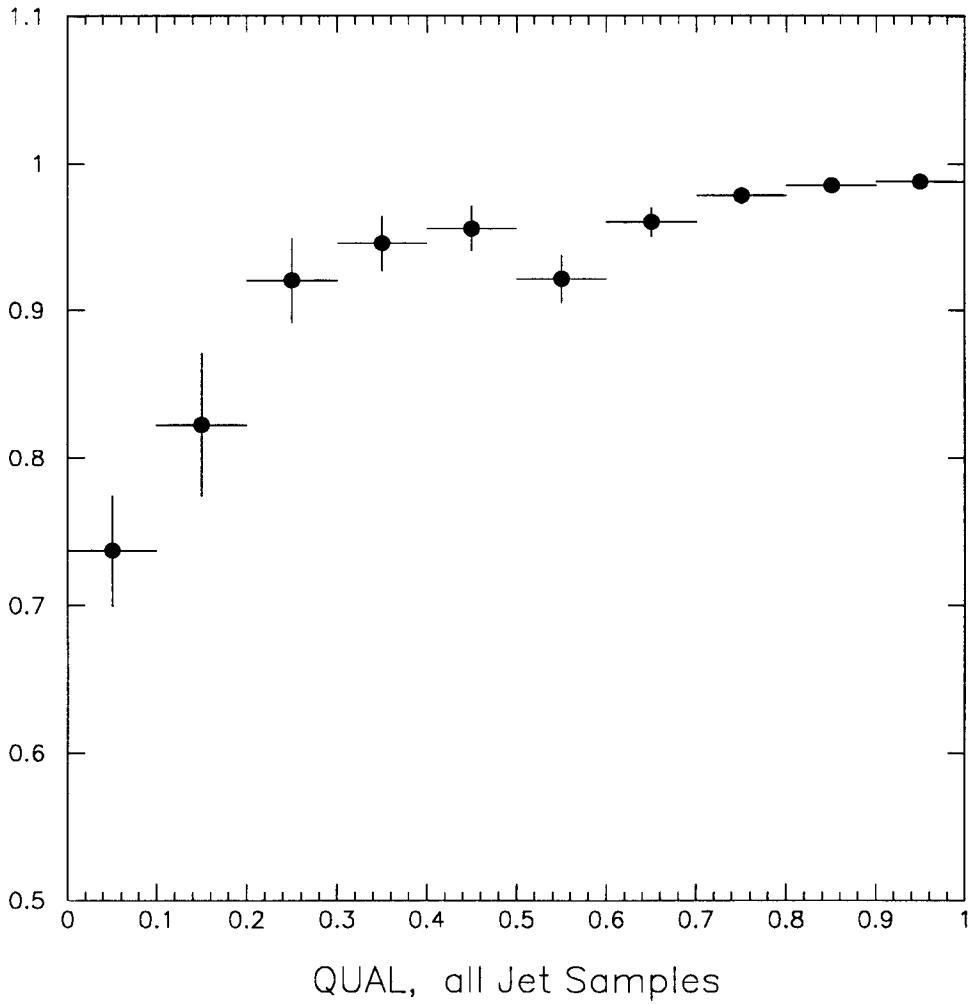


Figure 3.4: CTC track-finding efficiency as a function of Q for tracks in a mixture of jet data samples.

efficiency in the CTC as a function of Q . The integrated efficiency of finding a track with $P_T > 2.0$ GeV/c in the CTC is 96%, and is relatively flat as a function of track momenta. Seed tracks within the CTC fiducial region are then removed at random to according to the CTC tracking efficiency distribution as a function of track Q .

The CTC-to-SVX linking efficiency is measured in jet data. This efficiency is the rate at which a found CTC track, that projects into the SVX fiducial region, is reconstructed as a SVX track with hits on three or more layers of silicon. The upper plot of Figure 3.5 shows the CTC-to-SVX linking efficiency is weakly depen-

dent upon the track P_T . This distribution is parameterized as a function of P_T and used as a probability distribution to discard the Monte Carlo tracks at the rate observed in data. The overall CTC-to-SVX linking efficiency is 83%. The lower plot of Figure 3.5 shows the efficiency of SVX tracks passing good quality criteria (described earlier in this Chapter) as a function of the density of the track environment. The track density is measured by counting the number of nearby tracks in the event that fall within a cone of half angle 8° centered on the seed track axis. The integrated efficiency with which an SVX track passes the good quality criteria is 81%. Most of the SVX tracks that fail the quality requirements are poorly reconstructed. The combined CTC, CTC/SVX linking, and SVX tracking efficiency is nearly 70%.

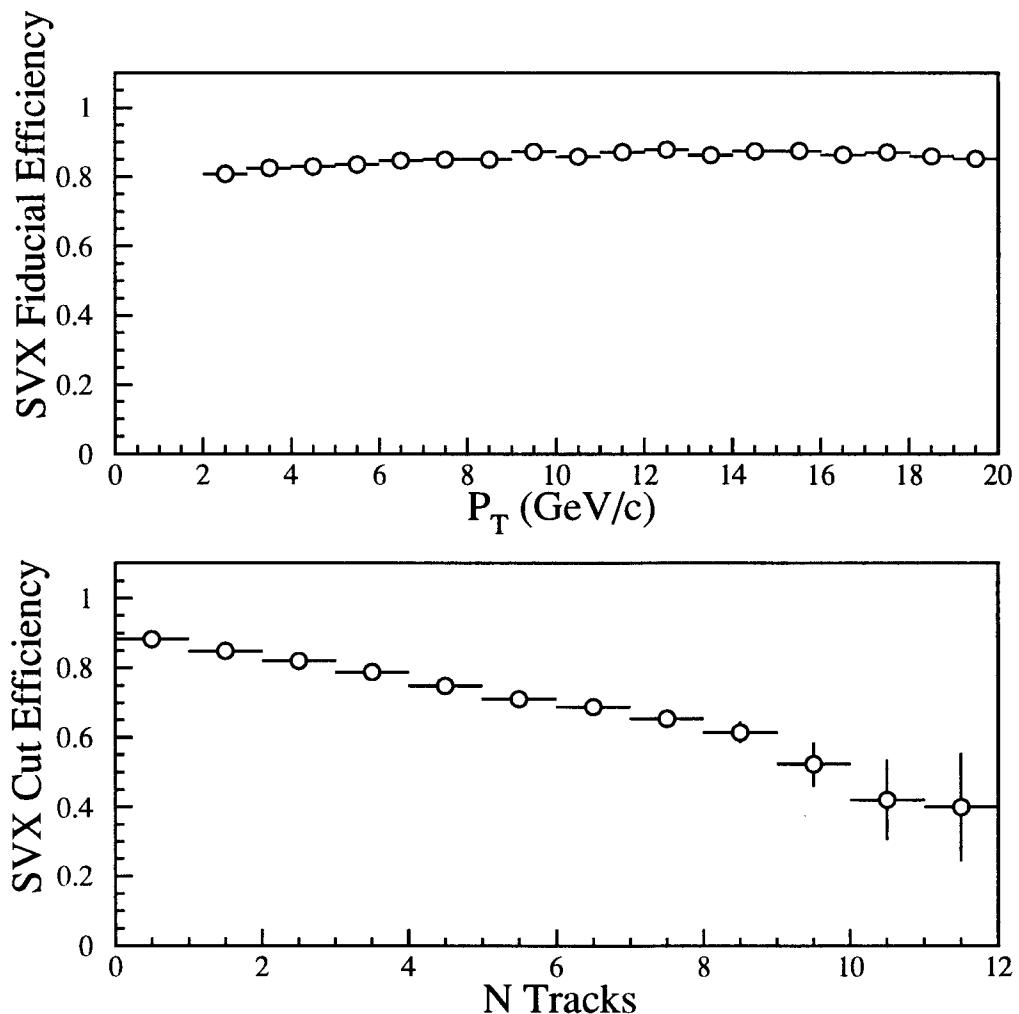


Figure 3.5: SVX tracking efficiencies as a function of track P_T and the density of the tracking environment. The upper plot shows the CTC-to-SVX linking efficiency as a function of track P_T . The lower plot shows the efficiency that tracks with three hits in the SVX pass the quality requirements (listed in the text) as a function of the number “nearby” tracks.

Chapter 4

Tagging Heavy Flavor Jets

The secondary-vertex algorithm (JETVTEX) is designed to find decay vertices displaced from the $p\bar{p}$ collision point, called the primary vertex. Bottom (b) and charm (c) quarks are long lived and hadronize, forming B and C hadrons along with a number of other hadrons (pions, kaons, etc...). Figure 4.1 is an illustration of a typical B or C -hadron decay as it might look in the CDF detector. The distance between the primary vertex and the B and C -hadron decay (secondary) vertex is the decay length (L). The projection of L in the $r - \phi$ plane is called the two-dimensional (2d) decay length (L_{xy}). The proper decay length of the B hadron is $c\tau \simeq 450 \mu\text{m}$ [18]. For the charm hadrons D^\pm and D^0 , the average $c\tau$'s are $\simeq 314 \mu\text{m}$ and $124 \mu\text{m}$ respectively [19]. Because of the B and C hadron's long lifetimes and large boost, they can travel a measurable distance from the primary vertex before decaying. The upper plots in Figures 4.2 and 4.3 show the distributions for 2d decay length of B and C hadrons. The lower plots in these two Figures show the average P_T for Monte Carlo generated B and C hadrons, in jets with $E_T > 15 \text{ GeV}$, are large in comparison to the b and c quark masses.

Tracks from B and C -hadron decay are measurably *displaced* (large impact

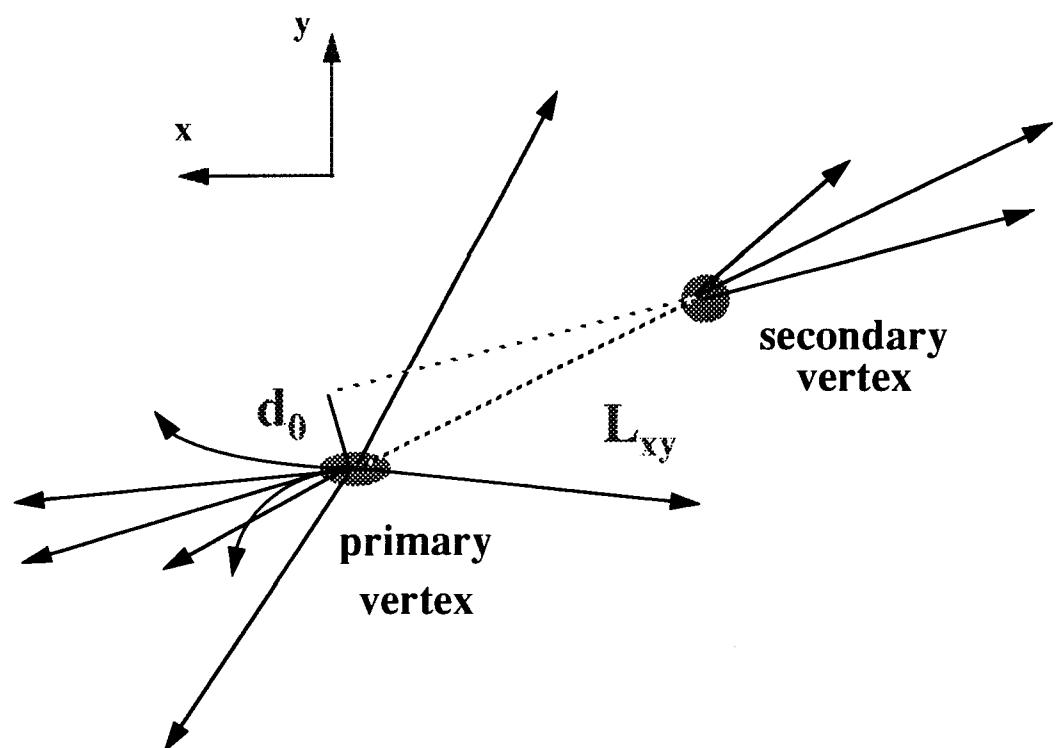


Figure 4.1: Simplified view of an event containing a secondary vertex, shown in the transverse $(r - \phi)$ plane.

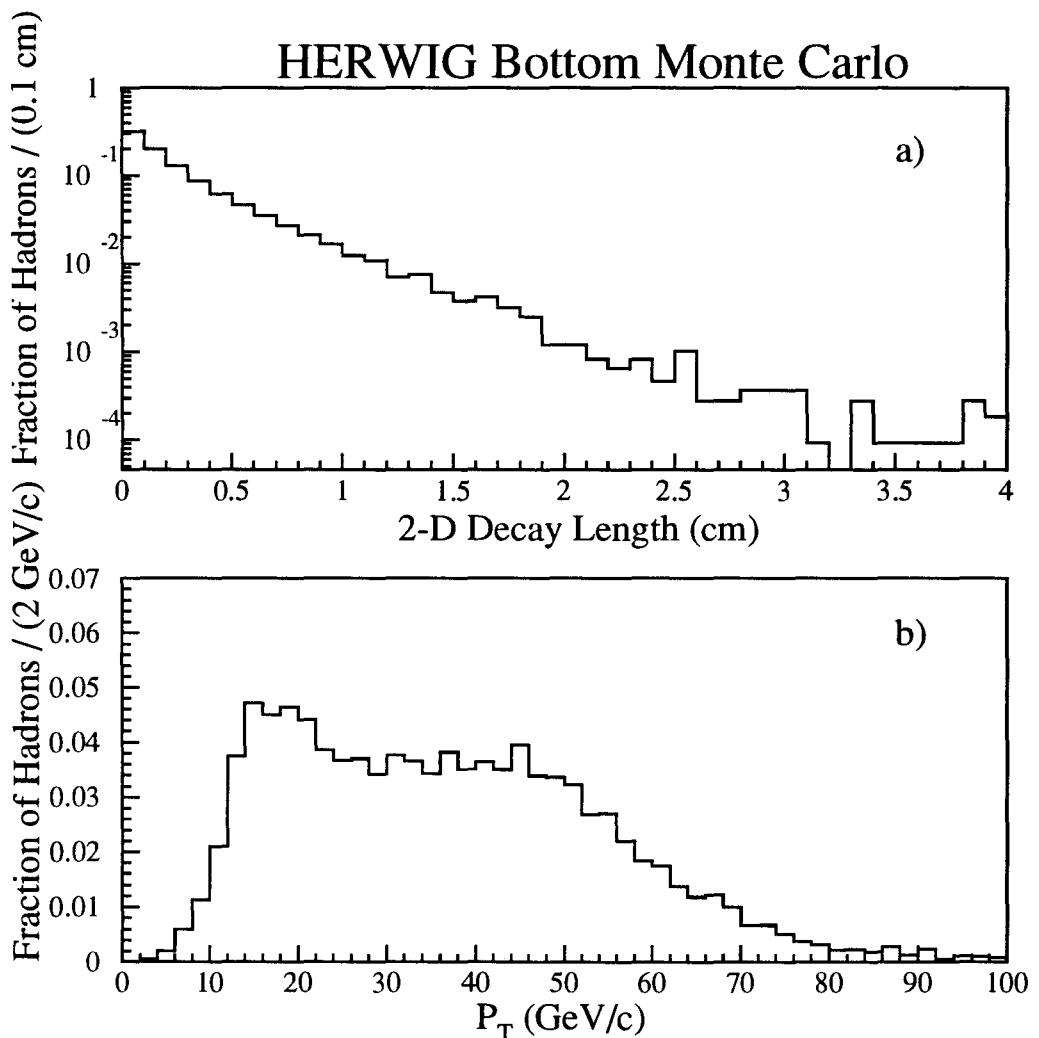


Figure 4.2: Plot (a) is the distribution of the two-dimensional decay length, L_{xy} for a mixture B hadrons in jets with $E_T > 15$ GeV. Plot (b) is the P_T spectrum.

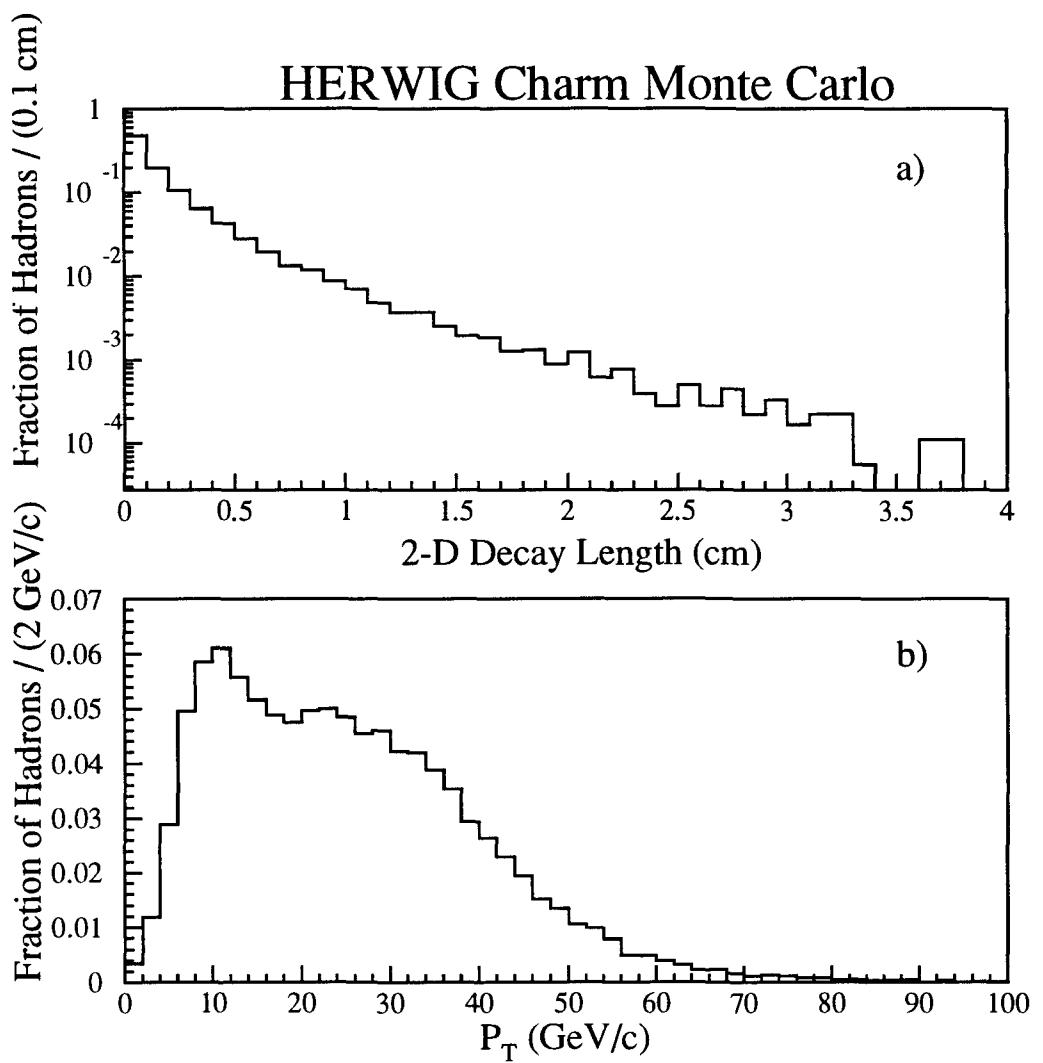


Figure 4.3: Plot (a) is the distribution of the two-dimensional decay length, L_{xy} for a mixture C hadrons in jets with $E_T > 15$ GeV. Plot (b) is the P_T spectrum.

parameter) from the primary vertex. Tracks from the decay of short-lived particles tend not to be displaced. The ability to identify these displaced tracks depends upon the resolution for determining the trajectory of these tracks and the position of the primary vertex. The primary vertex has a gaussian distribution with $\sigma \sim 30$ cm along the beam axis and $\sigma \sim 35 \mu\text{m}$ in the plane transverse to the beam axis. The CDF detector axis and the beam axis are not parallel and have a relative slope of $\sim 5 \mu\text{m}/\text{cm}$ in the horizontal plane and $\sim -3 \mu\text{m}/\text{cm}$ in the vertical plane. The difference between the detector axis and the beam axis at the nominal interaction point, $z = 0$, shifts between 200 and 1200 μm (400 and -1000 μm) in the horizontal (vertical) plane. Because of changing Tevatron conditions, the slope and displacement drifted during the course of data taking, yet are measured on a run-by-run basis to accuracies of $\sim 0.4 \mu\text{m}/\text{cm}$ for the slope and $\sim 10 \mu\text{m}$ for the displacement.

The position of the primary vertex is found for each event using a weighted fit of SVX tracks and the z position determined by the VTX. Corrections are made to account for the beam offset and slope. Displaced tracks are removed from the fit with an iterative procedure. The uncertainty in the primary vertex position varies from 6 to 36 μm in the plane transverse to beam axis.

Good quality SVX tracks (defined in Chapter 3.2.2) that are displaced from the primary vertex are used as input to the secondary vertex finding algorithm. To remove poorly measured tracks, photon conversions, long-lived K_S^0 's and Λ 's, the maximum impact parameter for each track must be less than 1500 μm . A displaced track is required to have a large impact parameter, d , compared to its estimated uncertainty, σ_d , such that the significance, $\frac{|d|}{\sigma_d}$, is larger than 3. The uncertainty in the impact parameter ranges from 50 μm for charged tracks with P_T of 1 GeV/c, and asymptotically to 15 μm for 10 GeV/c tracks. Figure 3.2 (from Section 3.2.2) and Figure 4.4 show the signed impact parameter and its significance for SVX

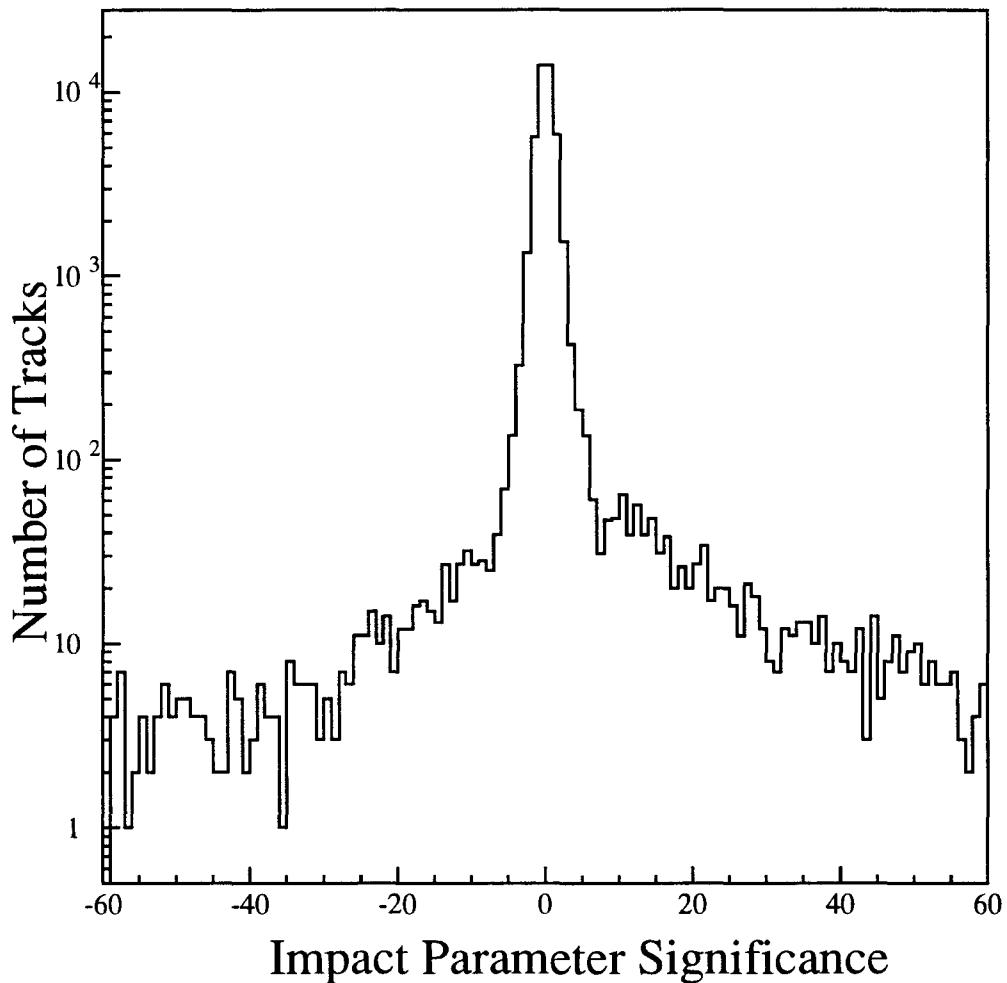


Figure 4.4: Impact parameter significance for good quality SVX tracks from jet_50 data.

tracks in the jet_50 data sample.

Figures 4.5 and 4.6 show the impact parameter and its significance for SVX tracks from bottom, charm, and non-heavy flavor HERWIG Monte Carlo jets. The impact parameter distribution is plotted with an arbitrary sign in order to better illustrate the relative displacements for bottom, charm, and background tracks. Tracks from b jets are more displaced than tracks from c or background jets. The fraction of tracks that pass the $\frac{|d|}{\sigma_d} > 3$ requirement is 0.23 for b , 0.11 for c , and

HERWIG Monte Carlo

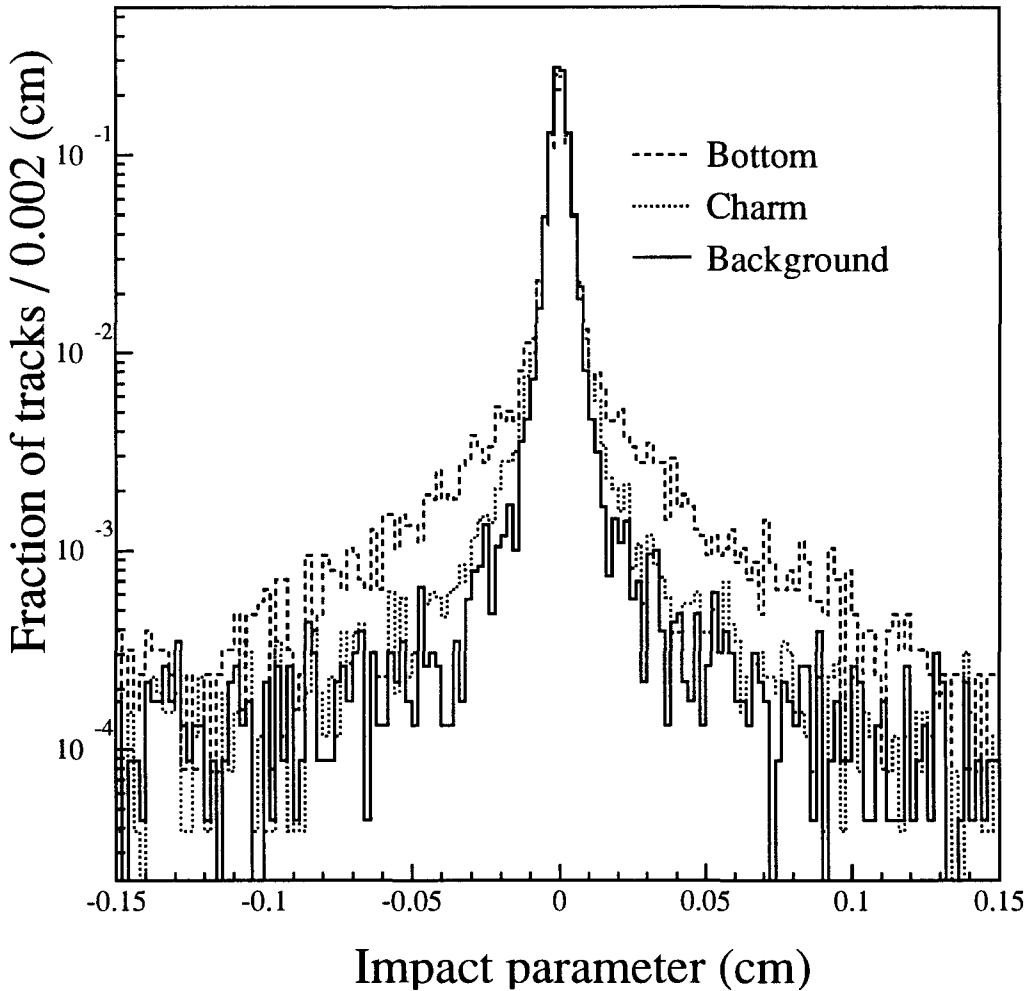


Figure 4.5: Impact parameter for good quality SVX tracks from b , c , and non-heavy flavor jet Monte Carlo.

0.07 for background jets. It should be noted that the d and $\frac{d}{\sigma_d}$ distributions for background have long-extending tails caused by mismeasured tracks.

4.1 The JETVTX Algorithm

JETVTX, the algorithm that searches for secondary vertices in jets, uses good quality, displaced SVX tracks that have been associated to jets which have $E_T > 15$

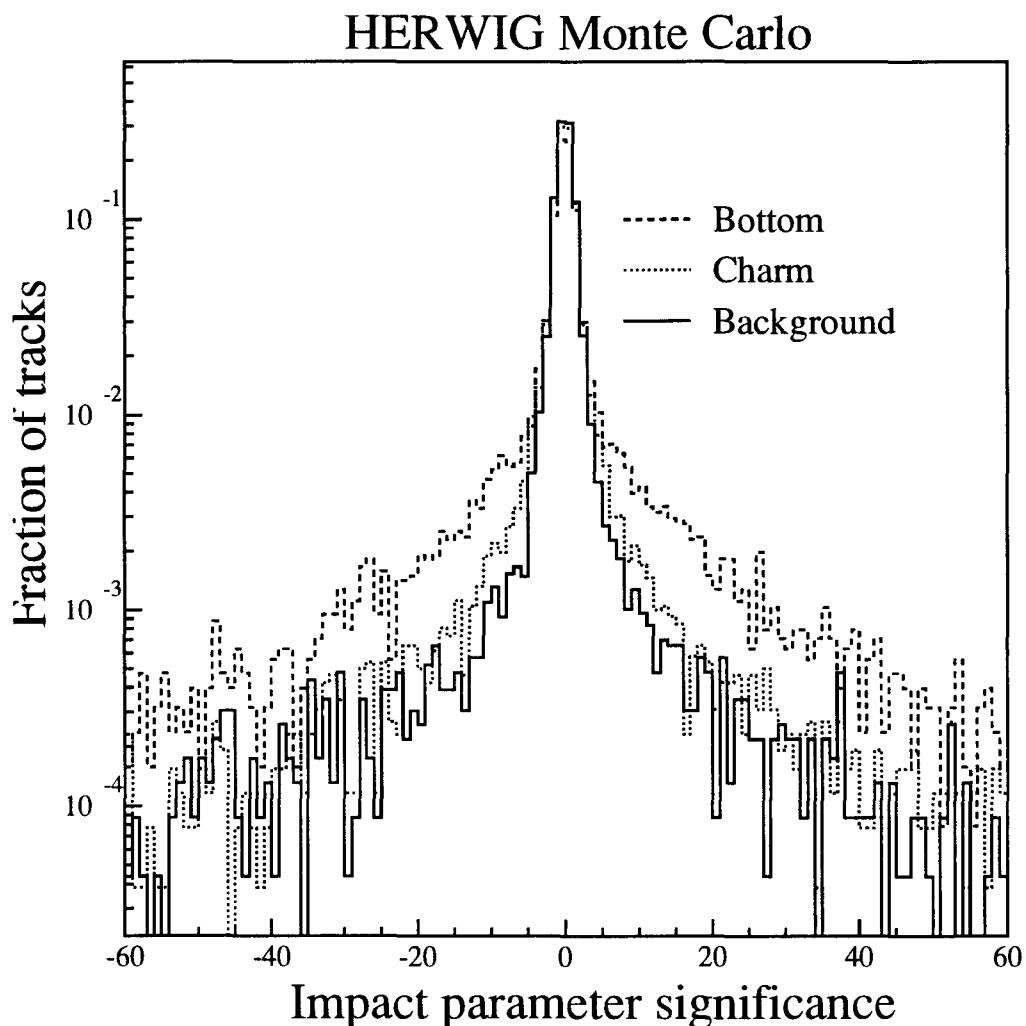


Figure 4.6: Impact-parameter significance for good quality SVX tracks from b , c , and non-heavy flavor jet Monte Carlo.

GeV and $|\eta| < 2$. The association is performed by requiring the cosine of the angle between the jet axis and the track in the $r - \phi$ plane to be larger than 0.80. Next, tracks associated to jets must pass the following selection criteria:

- $P_T \geq 2.0 \text{ GeV}/c$
- Tracks must have hits on at least two stereo superlayers and two axial superlayers of the CTC.
- They must be good quality SVX tracks, as described in Section 3.2.2.
- The impact parameter is less than $1500 \mu\text{m}$.
- The track must be displaced, $\frac{|d|}{\sigma_d} > 3$.

When a jet has at least two tracks satisfying the above criteria attached to it, the tracks are constrained to a common vertex with a two-dimensional fit in the $(r - \phi)$ plane. Because the vertex resolution is much larger (over two orders of magnitude) in the z-direction compared to the resolution in the $r - \phi$ plane, a two-dimensional rather than a three-dimensional fit is used. During the fit, the individual χ^2 contributed by each track to the overall χ^2 of the vertex fit is calculated. In an iterative process, tracks contributing a χ^2 larger than 50 to the vertex fit are removed, and the fit is attempted with the remaining tracks. This procedure continues until tracks are no longer removed or there are only two tracks left to form a common vertex. Tracks in the secondary vertex are uniquely assigned to that jet. Jets with secondary vertices successfully found are considered tagged.

For each common vertex formed, L_{xy} and $\sigma_{L_{xy}}$ ($\sim 130 \mu\text{m}$) are calculated. The decay length is signed. The sign is taken from the vector dot product of the direction of L_{xy} and the vector sum of the momenta of the tracks in the secondary vertex. As shown in Figure 4.7, L_{xy} is positive (negative) if the secondary vertex

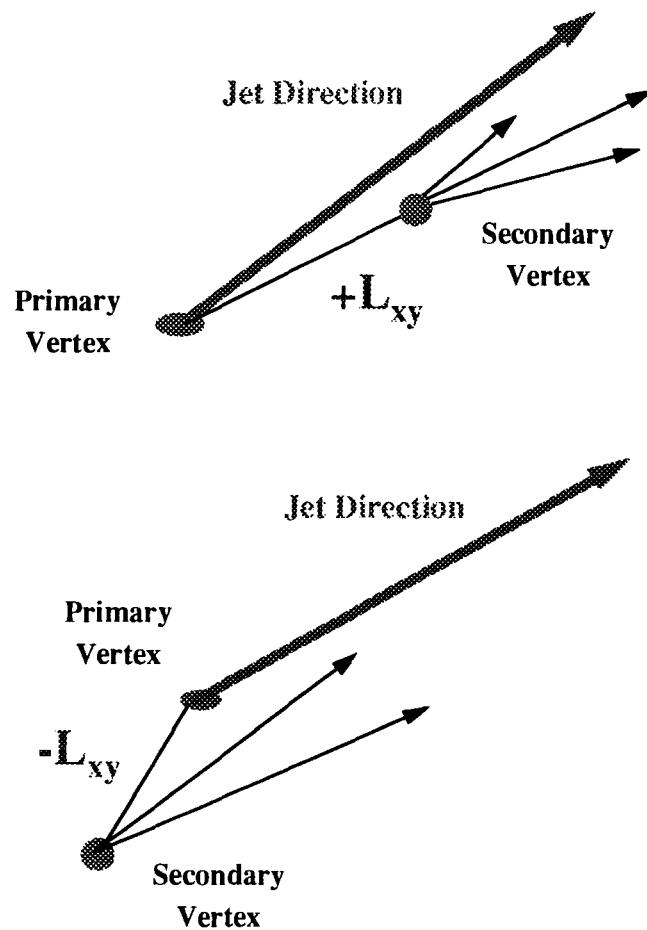


Figure 4.7: Schematic showing a positive and negative two dimensional decay length, L_{xy} . Where L_{xy} is positive (negative) if the secondary vertex is reconstructed in the same (opposite) hemisphere as the jet.

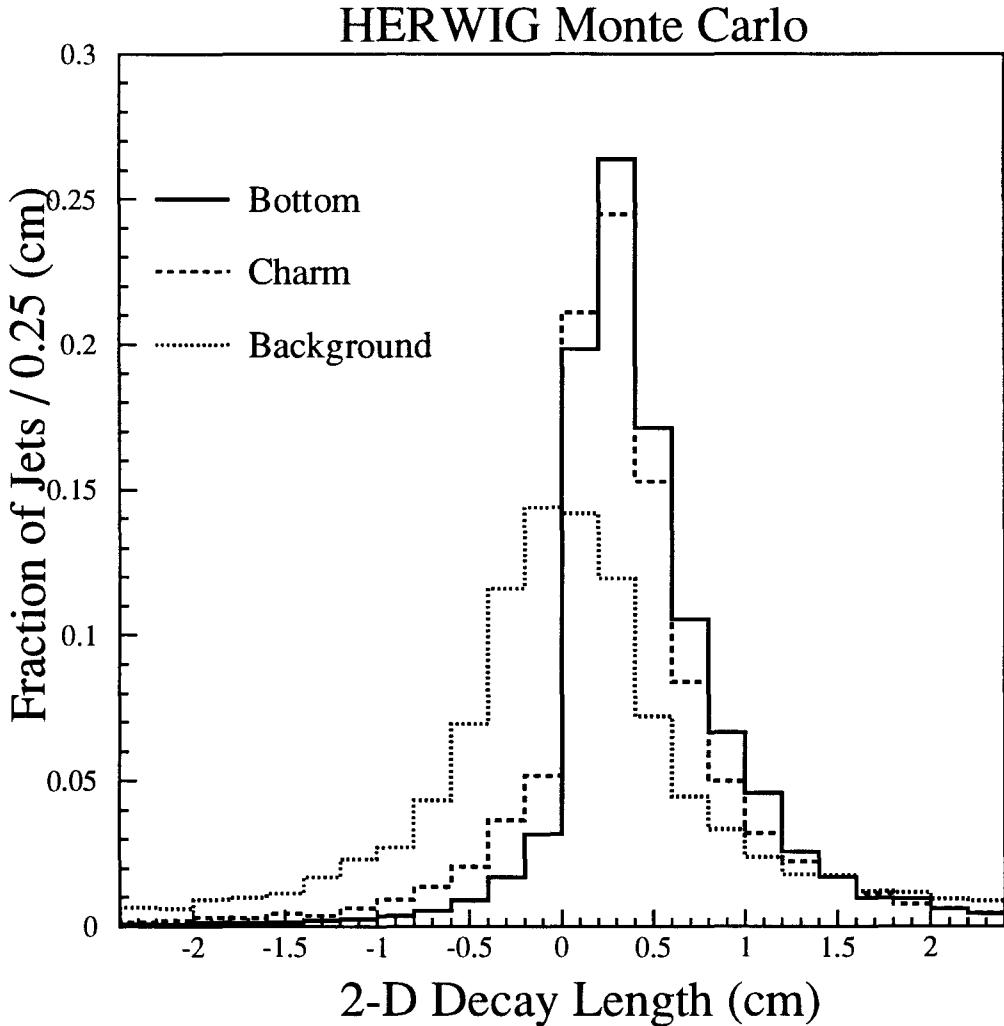


Figure 4.8: The two-dimensional decay length, L_{xy} , for tags in b , c , and background jet Monte Carlo.

is reconstructed in the same (opposite) hemisphere as the jet. Positive (negative) tagged jets have $+L_{xy}$ ($-L_{xy}$). Figure 4.8 illustrates that jets containing bottom and charm decay will have predominantly positive L_{xy} distributions. The L_{xy} distribution for non-heavy flavor jets, also shown in Figure 4.8, is symmetric about $L_{xy} = 0$. Tagged jets containing many poorly-measured tracks and no true secondary vertex, are equally likely to have a positive or negative L_{xy} . The majority of the secondary vertices of these mistagged jets are made up of only two tracks.

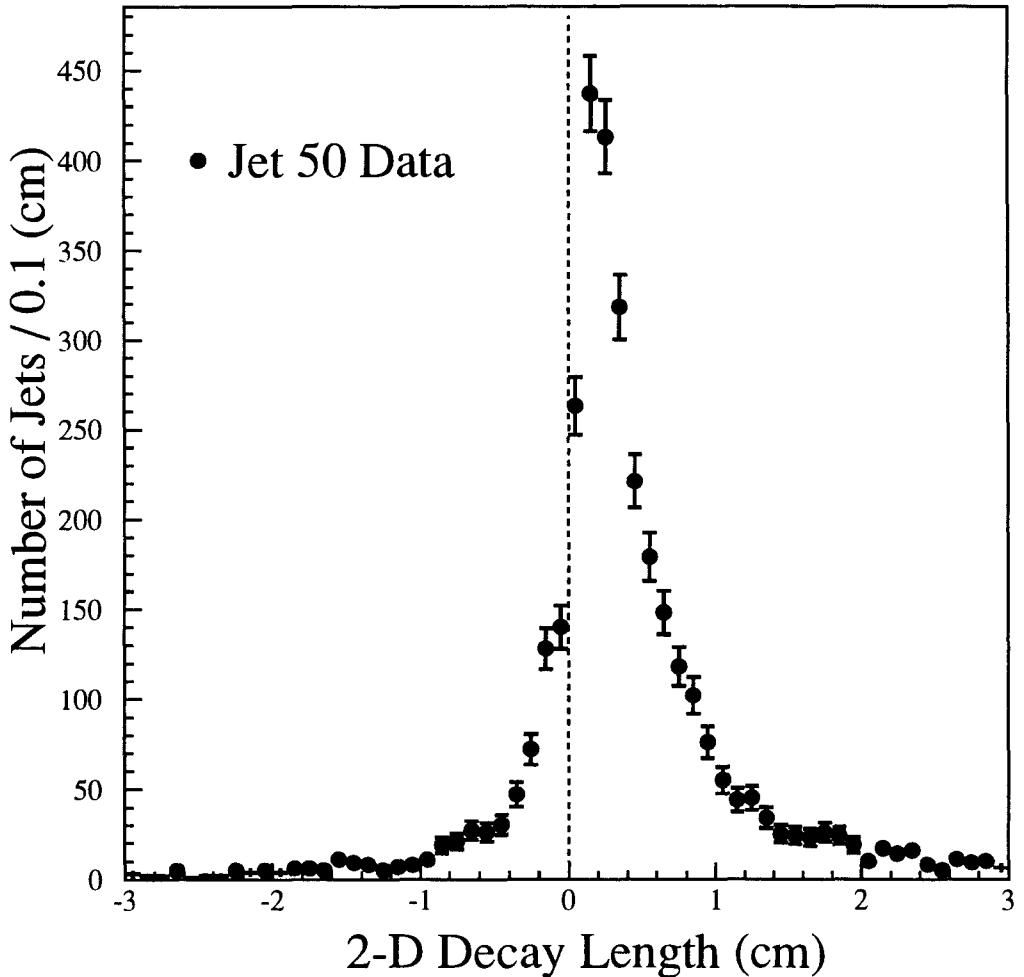


Figure 4.9: The two dimensional decay length, L_{xy} , for tagged jets in the jet_50 data sample. The vertical dashed line marks $L_{xy} = 0$.

Figure 4.9 shows the L_{xy} for tagged jets in the jet_50 data sample. The distribution shows more positive tags than negative tags, indicating the presence of real heavy flavor.

To better identify tagged jets that contain bottom or charm decay, the following requirements must be satisfied:

- To remove possible conversions from an interaction occurring in the inner-most layer of the SVX, L_{xy} must be less than 2.5 cm.

- A decay length significance of $|L_{xy}|/\sigma_{L_{xy}} > 3$ is efficient at removing non-heavy flavor jets with secondary vertices constructed from mismeasured tracks.
- Tagged jets with two-track vertices consistent with coming from $K_s^0 \rightarrow \pi^+\pi^-$ and $\Lambda^0 \rightarrow p^+\pi^-$ decay are removed. This is done by forming the two-track invariant mass (assuming the two tracks are both pions or one pion and the other track a proton), and removing the tagged jets with invariant masses falling within 3σ uncertainty of the world average value [32] for the K_s^0 and Λ^0 masses.

4.2 The Bottom and Charm Tagging Efficiency

In order to measure the bottom and charm fraction of inclusive jets, the efficiency with which JETVTX identifies or tags b and c jets must be evaluated. HERWIG generated $b\bar{b}X$ and $c\bar{c}X$ jet events are used in this evaluation.

The tagging efficiency, $\epsilon_{tag}^{b(c)}$, is found with the following equation:

$$\epsilon_{tag}^{b(c)} = \frac{\text{tagged } b(c) \text{ jets}}{\text{all } b(c) \text{ jets}} \quad (4.1)$$

Both positive and negatively tagged jets are used. The tagging efficiency is corrected to account for differences in tagging between data and the Monte Carlo simulation. This correction will be discussed in the next section. The corrected b and c jet tagging efficiencies are listed in Table 4.1. Figures 4.10 and 4.11 show that the tagging efficiency increases as a function of E_T . In addition, the tagging efficiency for b jets is larger than that for c jets. Thus the overall tagging efficiency for the Monte Carlo 50 GeV jet trigger sample is larger than the 20 GeV. Figure 4.10 and Figure 4.11 show that the tagging efficiency increases as a function of E_T . Thus the overall tagging efficiency for the 50 GeV jet trigger

	20 GeV Trigger	50 GeV Trigger
ϵ_{tag}^b	0.1550 ± 0.0050	0.2217 ± 0.0035
ϵ_{tag}^c	0.0334 ± 0.0017	0.0560 ± 0.0013

Table 4.1: Monte Carlo b and c jet tagging efficiencies. The uncertainties are statistical only.

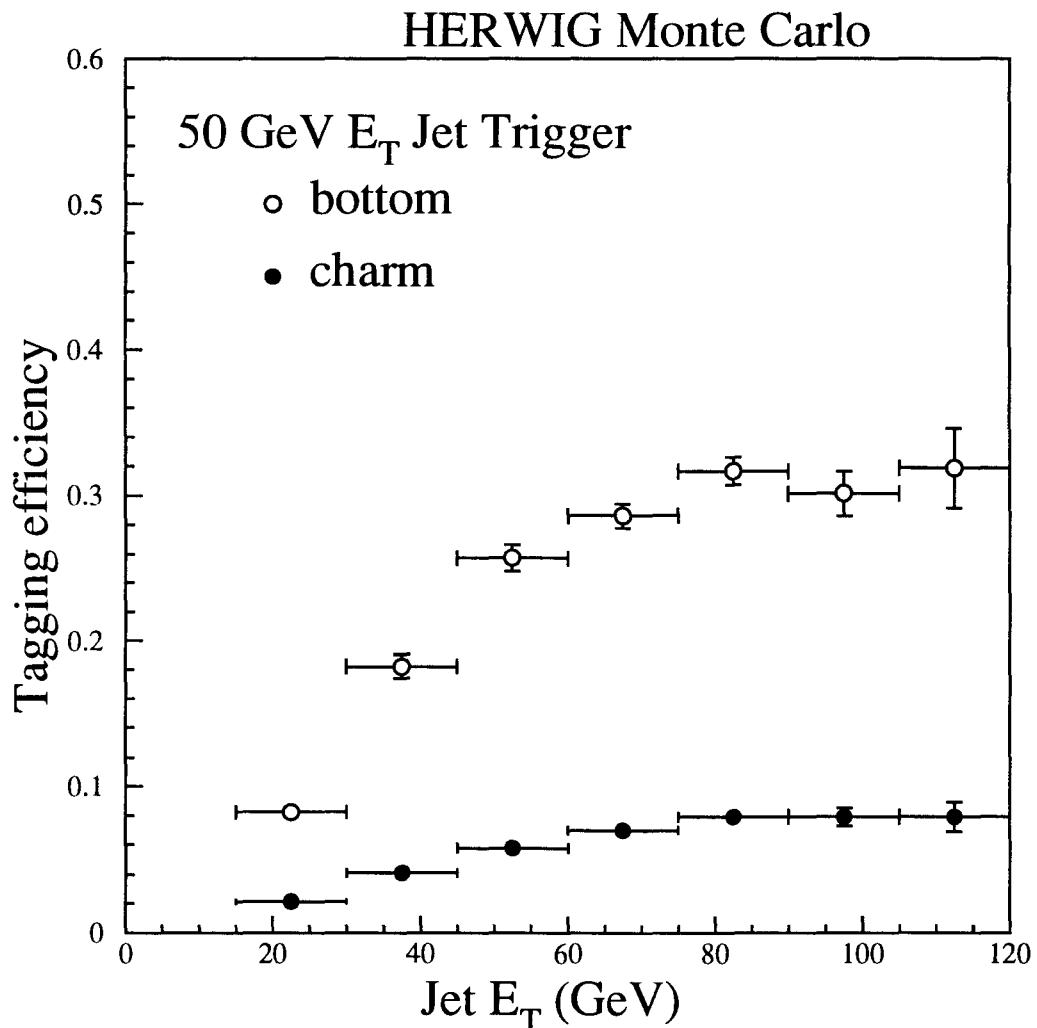


Figure 4.10: b and c jet tagging efficiencies as a function of E_T in the 50 GeV trigger jet Monte Carlo sample.

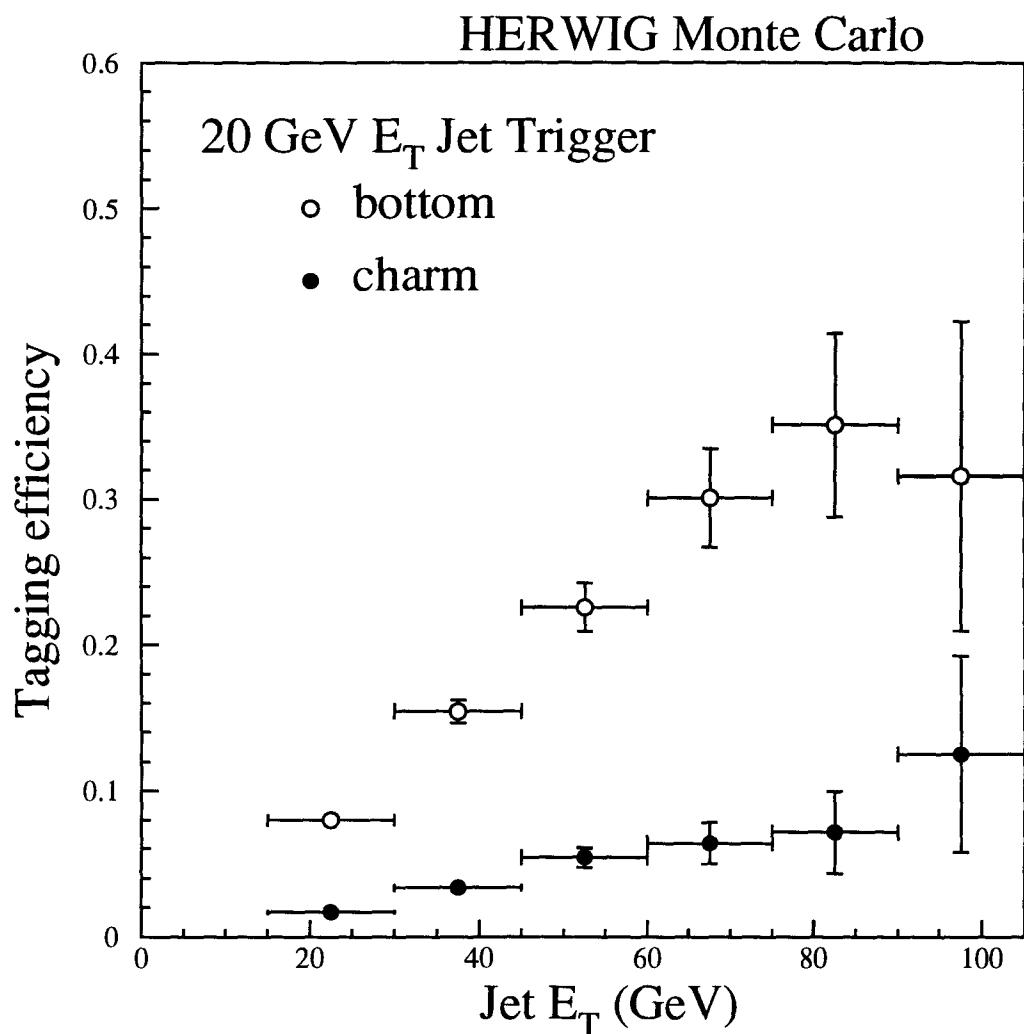


Figure 4.11: b and c jet tagging efficiencies as a function of E_T in the 20 GeV trigger jet Monte Carlo sample.

sample is expected to be larger than the 20 GeV. The higher b jet tag rate may be traced to three effects. First, it can be shown that the impact parameter of a track is proportional to the proper decay length of the parent particle. The impact parameter distribution for tracks from B -hadron decay is wider than tracks from C -hadron decay (see Figure 4.5). Tracks with larger impact parameters are more likely to be displaced. The impact parameter significance distribution (Figure 4.6) shows that b jets have a larger fraction of displaced tracks, $\frac{|d|}{\sigma_d} > 3$, than c or non-heavy flavor jets. As described in Chapter 4.1, only good quality SVX tracks that are displaced are input to the JETVTX secondary vertex finding algorithm. Therefore, tracks from b decay will be accepted by the vertex finding algorithm more often than tracks from c decay. Subsequently, b jets will be more likely to be tagged than c jets. This has the largest effect on the difference between the tagging efficiencies. The inclusion of sequential $b \rightarrow c$ decay in the calculation of d for tracks from b decay also tends to increase the d of the track making it more likely to be displaced. Lastly, the average track multiplicity is higher in b than in c decay. Jets with higher track multiplicity are tagged more often.

4.3 Tagging efficiency scale factor.

To test for differences in tagging b jets in data and Monte Carlo, JETVTX's performance is checked in a large inclusive electron data sample, and compared with similar events generated with Monte Carlo. As seen previously in Chapter 3, the fraction of electrons coming from semileptonic b decay in this sample has been measured to be 0.372 ± 0.081 . For this reason, the tagged jets associated with an electron in this sample are almost entirely b jets [34]. A scale factor is used to correct the differences in tagging efficiency between data and Monte Carlo. If the Monte Carlo model did a perfect job in simulating the physical event and

detector effects, the scale factor would be unity. The jet data is used explicitly to characterize the resolution of the tracking parameters in the Monte Carlo model. Moreover, SVX and CTC tracking efficiencies, measured in jet data as a function of the density of the track environment, are also implemented in the Monte Carlo model. Still, a scale factor is needed to correct the difference between data and Monte Carlo tagging efficiencies. The scale factor is defined in equation 4.2 below:

$$F^{scale} = \frac{\text{data Tagging Efficiency}}{\text{Monte Carlo Tagging Efficiency}} \quad (4.2)$$

The sources of the difference may include correlations between tracking variables, properties of b (c) fragmentation, and features of b (c) decay that the Monte Carlo does not model accurately.

The overall tagging efficiency (ϵ) can be separated into two terms:

$$\epsilon = \epsilon_{taggable} * \epsilon_{tag} \quad (4.3)$$

The first term, the taggable rate, is defined to be the number of taggable jets divided by the total number of jets. A taggable jet is associated with at least two good quality SVX tracks as listed in Chapter 3.2.2. The second term, the tagging efficiency, is defined as the number JETVTX tagged jets divided by the number of taggable jets. By defining the overall tagging in this manner, the full scale factor can be determined in the same way:

$$F^{scale} = F_{taggable}^{scale} * F_{tag}^{scale} \quad (4.4)$$

$F_{taggable}^{scale}$ corrects the differences in taggable jet production rates between the generic jet Monte Carlo and the inclusive jet data sample [36]. The taggable scale factor is defined as the data taggable rate divided by the taggable rate found in Monte

Carlo. F_{tag}^{scale} corrects the differences in tagging efficiencies between the $b\bar{b}$ Monte Carlo and the inclusive electron data sample. The scale factor for the tagging efficiency is defined as the data tagging efficiency divided by the Monte Carlo tagging efficiency.

In the calculation of the scale factor for the taggable rate, the Run IA jet_20 and jet_50 data samples are used. For Monte Carlo inclusive jet event simulation, HERWIG $b\bar{b}X$, $c\bar{c}X$, and non-heavy flavor jet events with 20 and 50 GeV E_T trigger filter are used. The detector simulation is described in Chapter 3.

The upper plot in Figure 4.12 shows the taggable rate in the Run IA jet_50 sample compared with that found in HERWIG generic jet Monte Carlo. The lower plot in the same figure is the corresponding scale factor as a function of jet E_T . The distribution is flat and consistent with the value of 1.0 for jets with E_T between 15 to 90 GeV. Above 90 GeV, the scale factor slowly falls to approximately 0.8 at 200 GeV. The integrated taggable rate for the jet_50 data is $0.462 \pm 0.001(stat)$. In Monte Carlo, the taggable rate is $0.491 \pm 0.002(stat)$. This gives a scale factor of $0.941 \pm 0.003(stat)$. The Monte Carlo model overestimates the taggable jet production rate. Figure 4.13 shows the taggable rate plots for the Run IA jet_20 sample and generic jet Monte Carlo. In data, the integrated taggable rate is 0.385 ± 0.001 . In Monte Carlo, it is 0.305 ± 0.002 . This gives a scale factor of 1.26 ± 0.01 . In this case, the Monte Carlo model underestimates the taggable jet production rate by 26%.

For the calculation of the tagging efficiency scale factor, two methods are employed to find the tagging efficiency (ϵ_{tag}): single tagging and double tagging. Both methods (to be described below) measure the efficiency for tagging b's that decay semileptonically in the inclusive electron sample.

The samples used in this calculation are the Run IA inclusive electron data and HERWIG Monte Carlo was used to generate $b\bar{b}X$ jet events. Each Monte Carlo

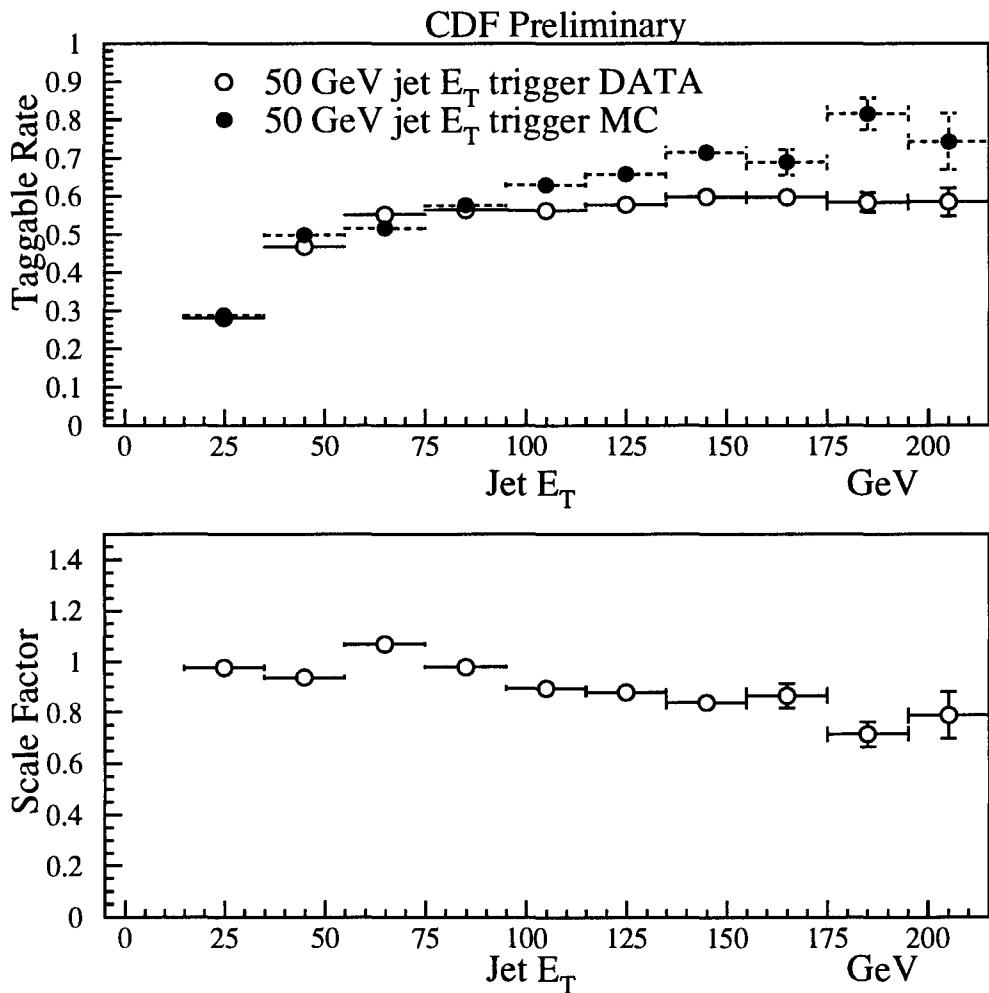


Figure 4.12: The jet_50 data and Monte Carlo taggable rate. The upper plot is the taggable rate as a function of jet E_T . The lower plot shows the associated scale factor.

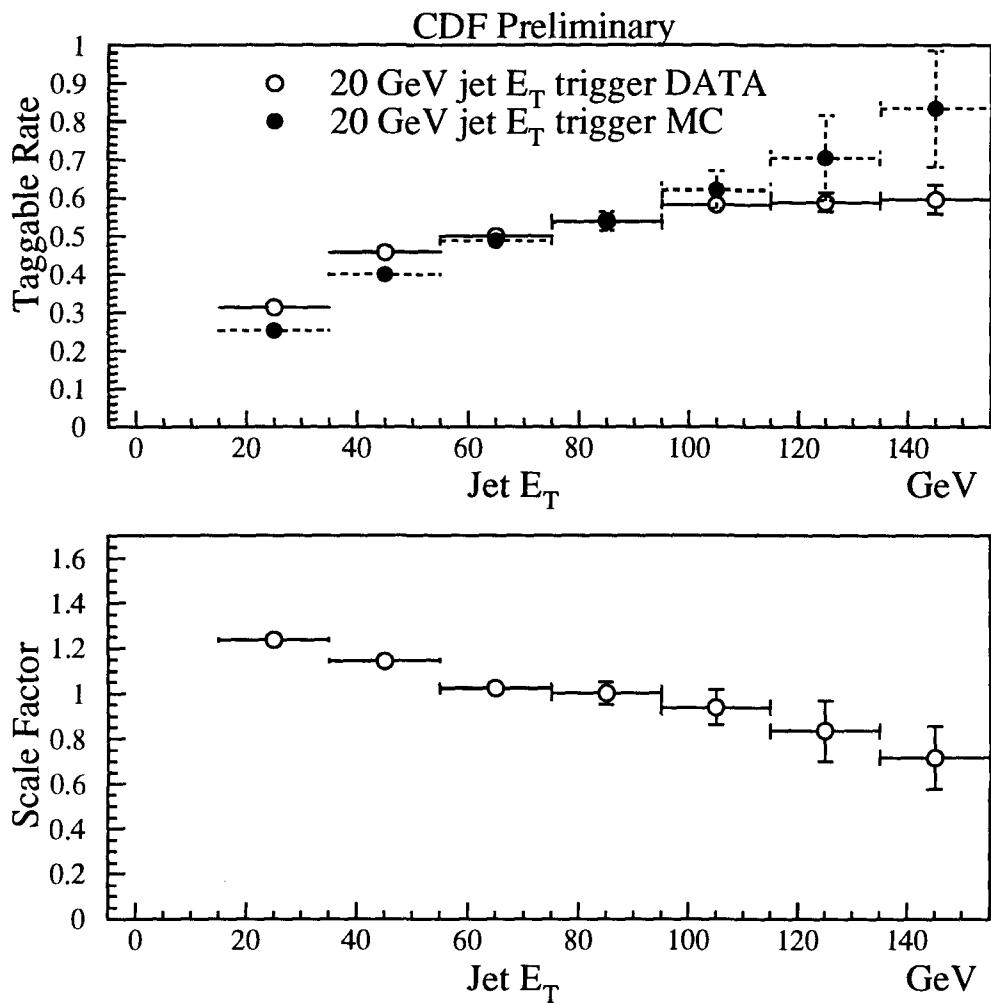


Figure 4.13: The jet_20 data and Monte Carlo taggable rate. The upper plot is the taggable rate as a function of jet E_T . The lower plot shows the associated scale factor.

event must contain a B hadron decaying semileptonically, to a C hadron and an electron ($B_x \rightarrow D_x e X$). The detector simulation was described earlier in the text. In both the inclusive electron data and the $b\bar{b}$ Monte Carlo sample good electrons are associated to jets (electron jet) by demanding that the electron to be within a cosine cone of 0.95 centered on the jet axis. Table 3.1 (in chapter 3) lists the selection criteria for good electrons. Next, an additional jet is required in the event not associated (away jet) with the electron that is a distance $\Delta R > 2.5$ from the electron-jet. Finally, all jets in all data and Monte Carlo samples described above, must have $E_T > 15$ GeV and $|\eta_{detector}| < 2.0$.

The single tag efficiency (ϵ_{tag}^{single}) is defined as the number of tagged electron jets, divided by the total number of taggable electron jets. In data, this efficiency must be corrected for the b -fraction of the sample, and for mistagged electron jets [37]. The expression for the single tag efficiency in data is:

$$\epsilon_{tag}^{single} = \frac{N_{ele-tag} - N_{ele-tag}^{pred} * (1 - F^B)}{N_{ele-jets} * F^B} \quad (4.5)$$

The numerator is the corrected number of tagged electron jets, where $N_{ele-tag}$ is number of observed tagged electron jets, and $N_{ele-tag}^{pred}$ is the predicted number of mistagged electron jets. In this calculation, only positive signed tags are used. F^B (0.372 ± 0.081) is the fraction of electrons in the Run IA inclusive electron data sample that come from semileptonic B decay. The denominator is total number of electron jets where the electron comes from semileptonic B decay.

The double tag efficiency (ϵ_{tag}^{double}) is the total number of tagged electron jets, provided that the away-jet in the event has been tagged, divided by the total number of away-tags. The double tag efficiency must be corrected for the number of mistagged away jets. The double-tag efficiency in data is calculated with the

Inclusive Electron data	
sample	Taggable jets
electron jets	28701
tagged electron jets	3268
tagged away jets	453
electron+away tags	130
predicted electron jet tags	474
corrected tagged electron jets	2969 ± 86
predicted away tags	163
corrected away-tags	350 ± 31

Table 4.2: Summary of tags in data.

HERWIG $b\bar{b}$ Monte Carlo	
sample	Taggable jets
good electron jets	2136
tagged electron jets	641
away tags	129
electron+away tags	38

Table 4.3: Summary of tags in $b\bar{b}$ Monte Carlo.

relation:

$$\epsilon_{tag}^{double} = \frac{N_{ele+away-tag}}{N_{away} - N_{away}^{pred} * (1 - F^B)} \quad (4.6)$$

Listed in Table 4.2 are the observed tags in the inclusive electron data. Table 4.3 contains the tags in the $b\bar{b}$ Monte Carlo. The integrated tagging efficiencies for data and Monte Carlo are summarized in Table 4.4. The uncertainty in the data single tag efficiency is dominated by the error in F^B , whereas the uncertainty in the double tagging efficiency is dominated by statistics. Since both the single tag

Tagging efficiencies per taggable jet		
Tagging Method	$b\bar{b}$ Monte Carlo	data
single	0.300 ± 0.010	0.280 ± 0.060
double	0.294 ± 0.040	0.287 ± 0.036
weighted average	0.300 ± 0.010	0.285 ± 0.031

Table 4.4: Comparison of tagging efficiencies in Monte Carlo and data. The uncertainty in the data efficiencies include both the error on F^B and the statistical error.

and the double tag methods measure the efficiency to tag semileptonic b decay, the weighted average of the two is calculated. In Monte Carlo, the weighted average of 0.300 ± 0.010 is in good agreement with the value of 0.285 ± 0.031 measured in data.

In Figure 4.15 the single and double tagging efficiencies are plotted as a function of jet E_T . For both data and Monte Carlo, the two tagging methods are in agreement. The upper plot in Figure 4.14 shows the weighted averages of the single and double tag efficiencies for data and Monte Carlo as a function of jet E_T . The lower plot is the scale factor as a function of jet E_T . The scale factor is consistent with being independent of jet E_T . Fitting the distribution to a constant value, the resulting scale factor is 0.96 ± 0.08 . The error in the scale factor is the uncertainty in the fit. In Figure 4.16 several variables of the tagged electron jets are plotted: the estimated proper decay length (will be described in the following chapter), the tagged jet invariant mass formed with the tracks in the tag (Mass), the combined P_T of those tracks, and the number of tracks in the tagged jet (NTRK). The Monte Carlo distributions agree with what is observed in the data. This gives us confidence that the Monte Carlo model simulates the tagged jets properly.

Equation 4.4 is used to calculate the full scale factor. The taggable rate in the

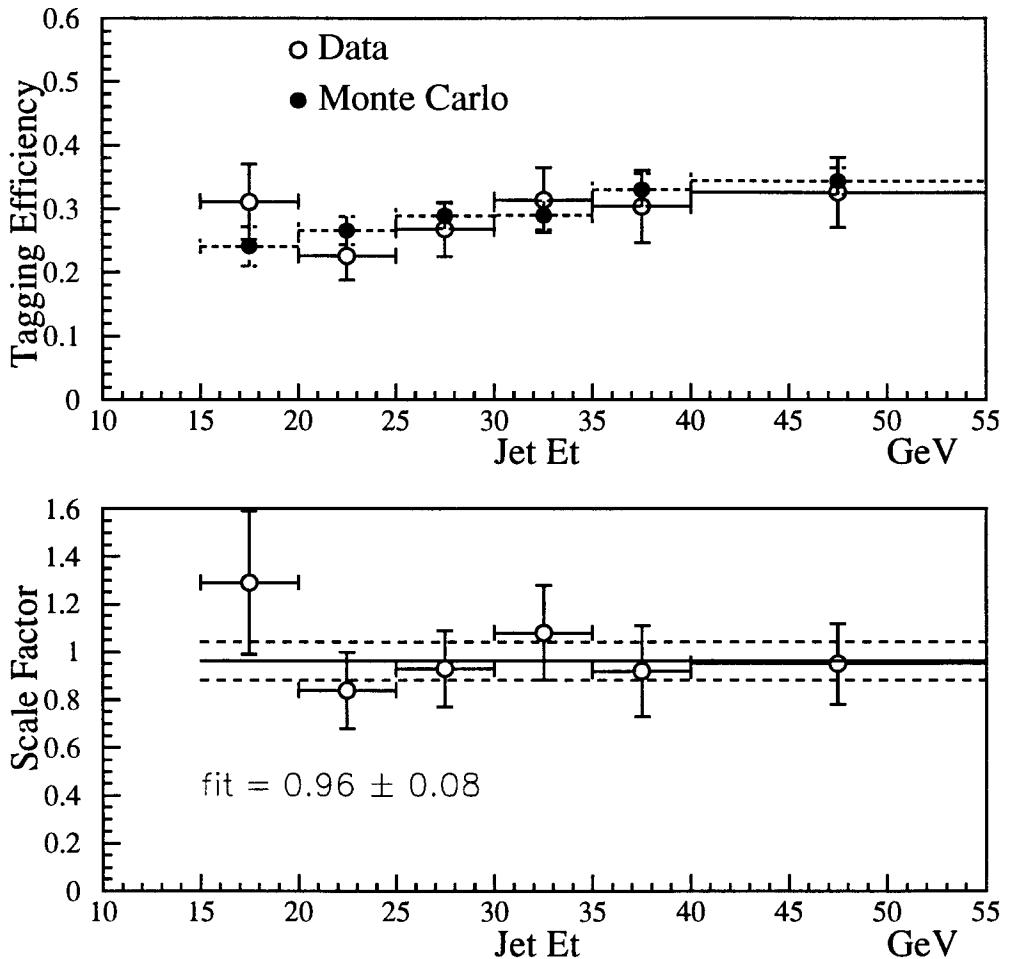


Figure 4.14: Scale factor calculated with weighted averages of the two tagging methods as a function of jet E_T . In the upper plot, the weighted average of the two tagging efficiencies in data (open points) is compared with Monte Carlo (solid points). The lower plot shows the scale factor. The solid line is the resultant constant fit, and the 2 dashed lines are the $\pm 1\sigma$ uncertainty in the fit.

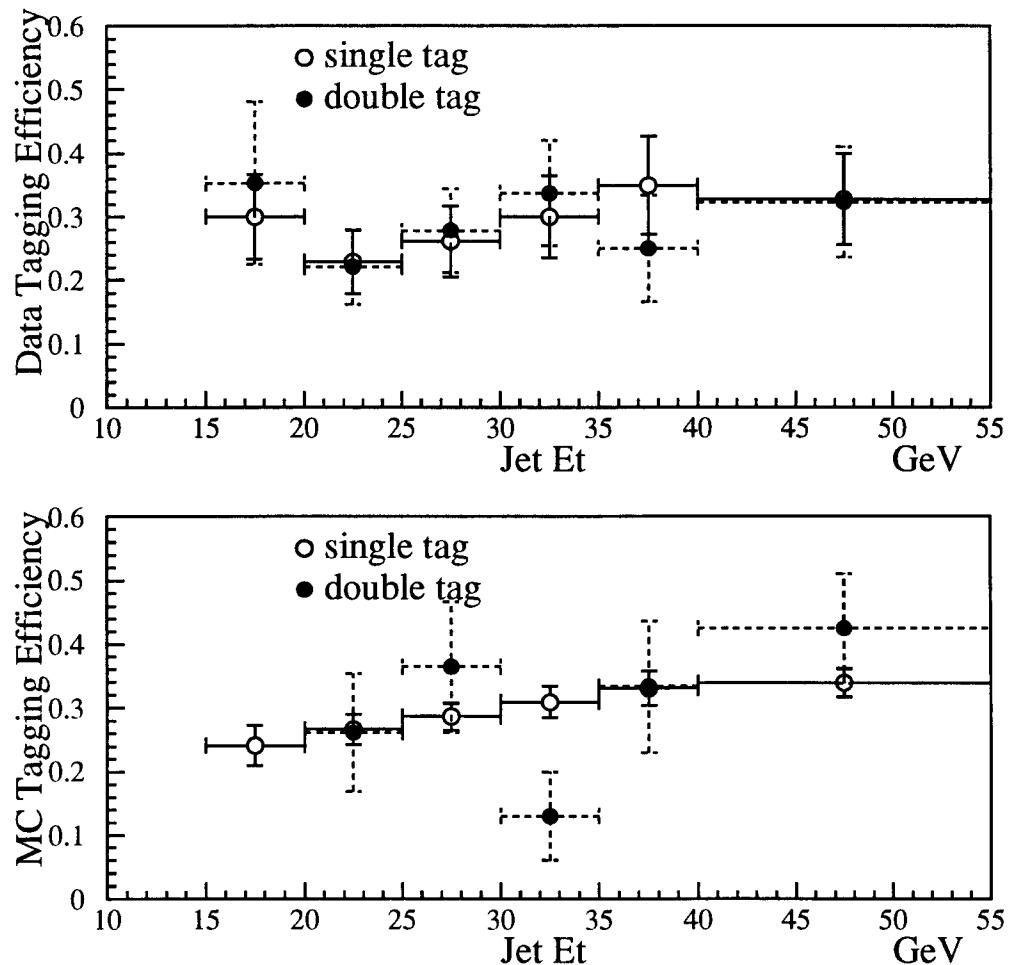


Figure 4.15: The tagging efficiency as a function of jet E_T . A comparison of the two tagging methods for the inclusive electron data are shown in the upper plot. The lower plot shows the two methods in the $b\bar{b}$ Monte Carlo.

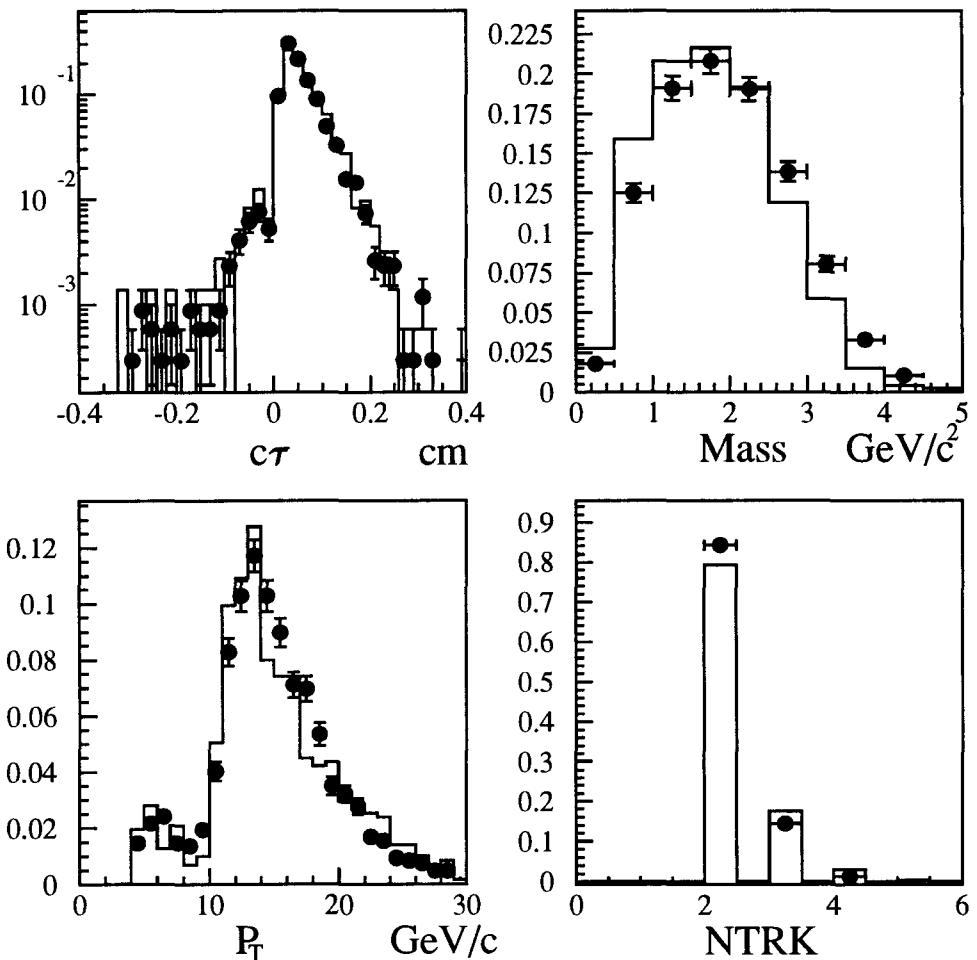


Figure 4.16: Comparison of tagged electron jet variables described in the text, Data (solid points) and Monte Carlo (histograms) are normalized to unit area.

jet_50 sample has been corrected so $F_{taggable}^{scale} = 1$. Hence the full scale factor is $F^{scale} = F_{tag}^{scale} = 0.96 \pm 0.08$. For the jet_20 sample, $F^{scale} = 1.21 \pm 0.09$.

Chapter 5

The Determination and Fitting of the Estimated Proper Decay Length Distributions

5.1 Fit Method

A common method for determining the proper decay length for a decaying particle uses the measured three-dimensional (3-d) decay length (L_{xyz}):

$$c\tau = L_{xyz} \cdot \frac{M}{P} \quad (5.1)$$

M is the mass and P is the magnitude of the 3-vector momentum of the particle before decay. The ratio $\frac{M}{P}$ gives the lorentz factor of $\frac{1}{\beta\gamma}$. Because of the poor z-resolution of the secondary vertex, the 2-d decay length (L_{xy}) is used in this

analysis to determine the proper decay length:

$$c\tau = L_{xy} \cdot \frac{M}{P_T} \quad (5.2)$$

P_T is the transverse momentum of the particle before decay. The ratio $\frac{M}{P_T}$ gives the lorentz factor of $\frac{1}{\beta\gamma}$ in the plane transverse to the beam axis.

The secondary-vertex finder estimates the two-dimensional decay length of a long-lived object using displaced tracks within a jet. However, the object is usually not fully reconstructed because not all tracks from the decay are found. Neutral particles such as neutrinos and π^0 's will not be detected in the tracking chambers. Some charged particle tracks will not be successfully attached to the secondary vertex because they are poorly measured. Moreover, since the combined tracking efficiency is $\sim 70\%$ (see Chapter 3), a significant part of the time tracks will be completely missed. The invariant mass (M_V) and transverse momentum (P_T), determined with the tracks that make up the secondary vertex, can be related to the M and P_T of the long-lived object before decay by a correction factor determined from Monte Carlo:

$$K = \frac{M/P_T}{M_V/P_{TV}} \quad (5.3)$$

Using equation 5.3, the estimated proper decay length for a tagged jet can be expressed as:

$$c\tau = L_{xy} \cdot \frac{M_V}{P_{TV}} \cdot K \quad (5.4)$$

Correction factors have been found for both B (K_b) and C hadron (K_c) decays using Monte Carlo simulation. The distributions of K_b and K_c are independent of P_{TV} above 15 GeV/c. Fit to a constant value, the correction factors are similar:

- Bottom Correction factor $K_b = 1.39 \pm 0.02$

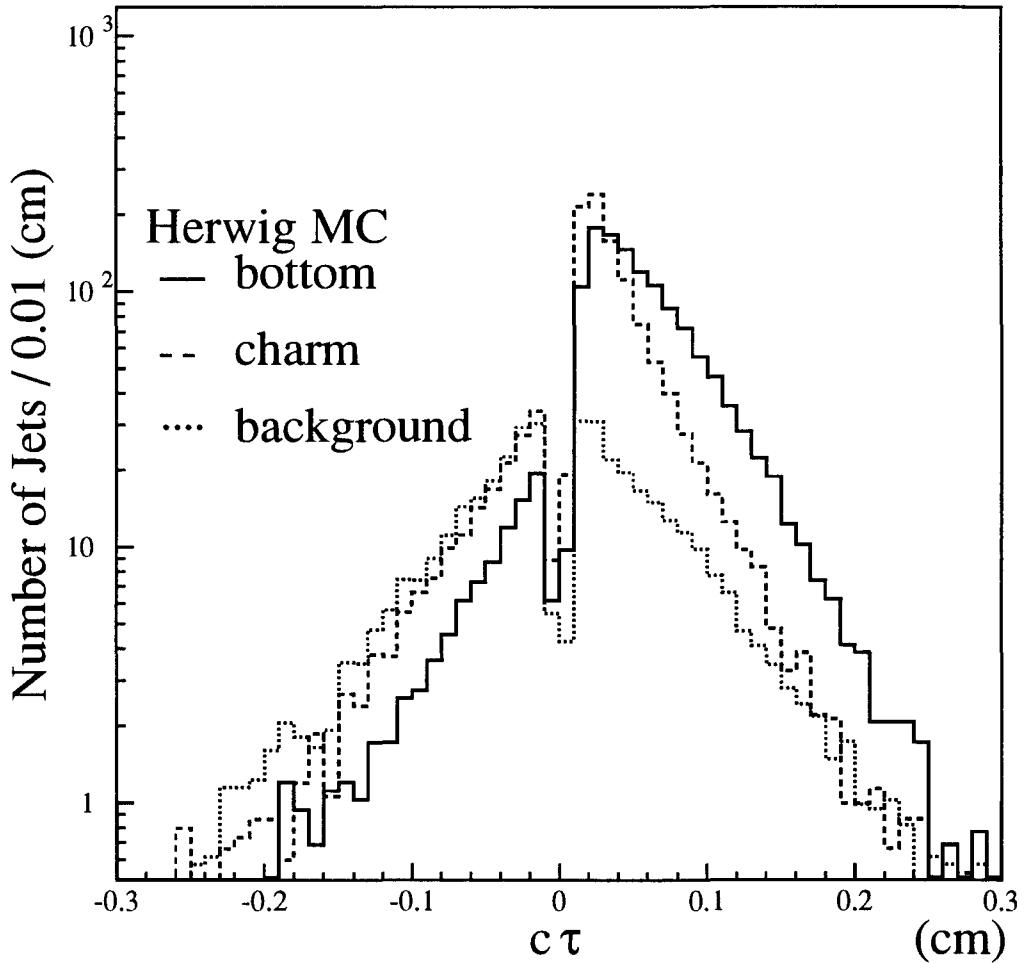


Figure 5.1: The $c\tau$ distributions for HERWIG Monte Carlo b , c , and background tagged jets.

- Charm Correction factor $K_c = 1.19 \pm 0.01$

Where the uncertainties are the error in the fit. The average value of the correction factor, $K_{avg} = 1.29$, is used for data and all Monte Carlo $c\tau$ distributions. Using the average value of the correction factor does not significantly limit $c\tau$ as a discriminant between b and c quarks.

The Monte Carlo generated $c\tau$ distributions for c , b , and background tagged jets are shown in Figure 5.1. The contents of each Monte Carlo distribution are

described below:

- The $c\tau$ distribution for b jets contains the decay products from B hadrons formed by a fragmenting b quark. B jets may also contain the decay products of C hadrons from the sequential decay of B hadrons.
- The $c\tau c$ jet distribution contains the decay products from C hadrons formed by a fragmenting c quark only. The decay products from C hadrons produced by the sequential decay of B hadrons are not included.
- The background $c\tau$ distribution contains non-heavy flavor jets only. There are no b or c jets included.

When normalized to equal areas, the distinguishing qualities of the three distributions may be observed. Both the b and c distributions show a large positive decay length tag asymmetry about $c\tau = 0$. This is expected for jets from heavy flavor decay. Because of its shorter lifetime, the $c\tau$ distribution for c is more central (the distribution peaks closer to zero) than the b distribution. The c distribution has a larger fraction of negative decay length tags than b . The ratio of the number of positive decay length tags to the number of negative tags is $\sim \frac{13}{1}$ for b , compared with $\sim \frac{6}{1}$ for c . The larger number of negative signed tags for c is due to lifetime effects and impact parameter resolution smearing. The background distribution is symmetric about zero, and has long extending tails. The slope of the positive side of the $c\tau$ distribution for background jets is less steep than that for either of the b or c distributions. Ordinarily, this would indicate a longer lifetime for background. Recall that the tagging algorithm requires tracks with large impact parameters. In Chapter 4 it was shown that the negative signed impact parameter distribution for data jets is used as the impact parameter resolution function to smear Monte Carlo generated tracks. In addition, this distribution was shown to

have long extending tails. The majority of the secondary vertices in tagged (or rather "mistagged") background jets are composed of two poorly measured tracks with large impact parameters from the tails of the resolution function. Therefore secondary vertices made up of these tracks give larger $c\tau$'s producing a shallower slope in the background shape. The slopes on the negative side of the $c\tau$ distributions for b , c , and background jets are approximately the same as the positive side slope of the background $c\tau$. This also illustrates the effects of resolution smearing. The narrow depression centered about $c\tau$ of zero in all three distributions is caused by the requirement that the decay length significance, $\frac{|L_{xy}|}{\sigma_{L_{xy}}}$, must be greater than 3. This requirement is efficient in removing mistags (See Chapter 4).

The $c\tau$ distributions for tagged jets in the jet_50 and the jet_20 data samples are shown in Figures 5.2 and 5.3. These distributions show a positive decay length tag asymmetry that indicates the presence of real heavy flavor. The ratio of the number of positive decay length tags to the number of negative tags is $\sim \frac{5}{1}$. These distributions also contain a narrow depression at $c\tau = 0$ due to the decay significance requirement. The $c\tau$ from data tagged jets is fit to a sum of three components: the Monte Carlo generated $c\tau$ distributions of tagged b , c , and background tagged jets. The fit is a χ^2 minimization performed with the software package MNFIT (which uses MINUIT for minimization) [38]. The only constraint in the fit is that the sum of the normalizations for each component is unity.

5.2 Fit Results

The fit for the jet_50 sample is shown in Figure 5.2. Of 3006 tagged jets, the fractions (with the fit uncertainty) are:

- $(45.2 \pm 3.5)\% b$ jets

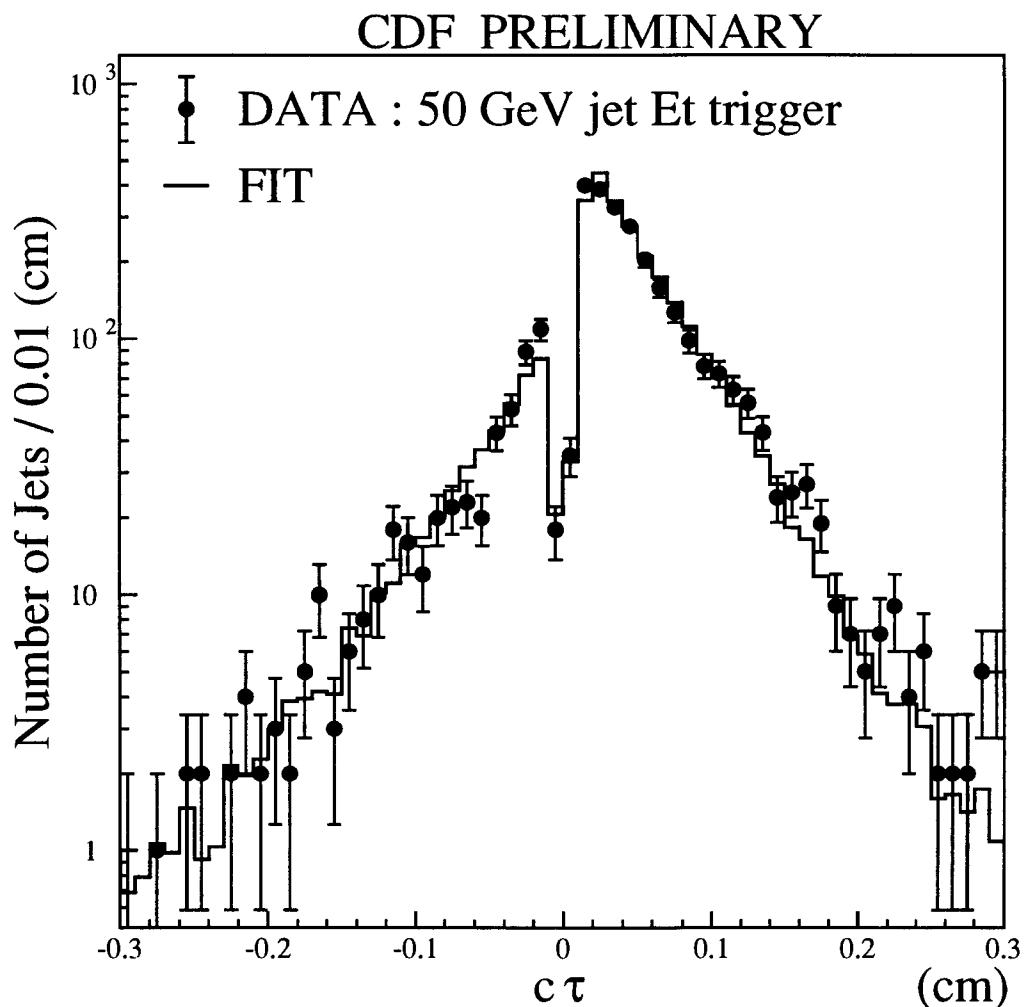


Figure 5.2: The $c\tau$ distribution in the jet_50 data (points) and fit (histogram) to the b , c , and background $c\tau$ distributions.

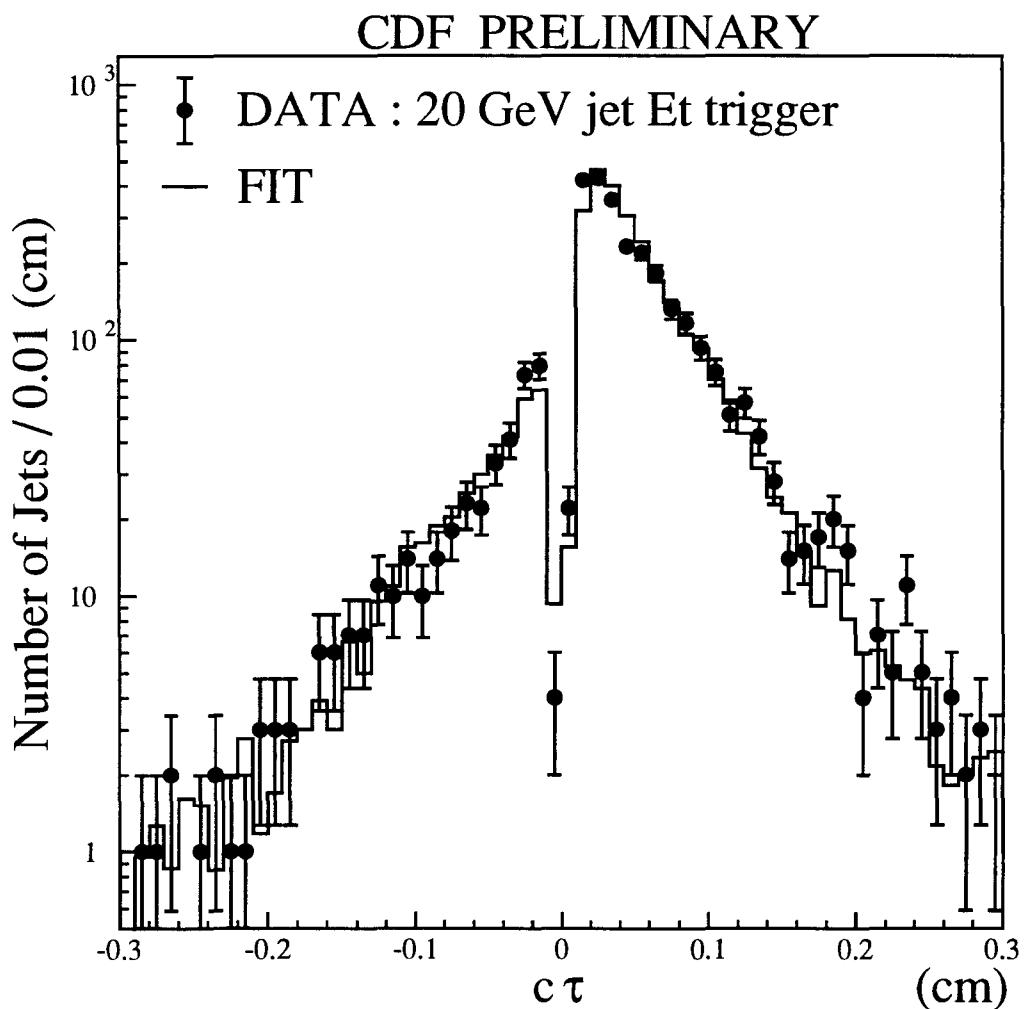


Figure 5.3: The $c\tau$ distribution in the jet_20 data (points) and fit (histogram) to the b , c , and background $c\tau$ distributions.

Content of Tagged Jets			
data jets	<i>b</i> jets	<i>c</i> jets	Background jets
jet_50 3006	1359 ± 105	1223 ± 117	424 ± 57
jet_20 3008	1363 ± 132	1290 ± 135	355 ± 51

Table 5.1: The $c\tau$ fit results for tagged jets. The uncertainties are from the fit error.

- $(40.7 \pm 3.9)\%$ *c* jets
- $(14.1 \pm 1.9)\%$ Background jets

The resulting $c\tau$ fit for the jet_20 sample is shown in Figure 5.3. Of the 3008 tagged jets:

- $(45.3 \pm 4.4)\%$ *b* jets
- $(42.9 \pm 4.5)\%$ *c* jets
- $(11.8 \pm 1.7)\%$ Background jets

5.3 Fit Consistency Check

One can plot the $c\tau$ distribution for tags that use only 2 tracks to tags with 3 or more tracks. Comparing the plus-to-minus ratio of tags in the two track only ($\sim \frac{4}{1}$) $c\tau$ distribution to the three or more track ($\sim \frac{18}{1}$) $c\tau$, the 3 track ratio is much higher, indicating that the background is greatly reduced. This is because background tagged jets or mistags are dominated by secondary vertices that contain only two tracks. The three or more track *b* (*c*) fit can be used as a consistency check of the results of the standard two or more track $c\tau$ fit. Using

the relation below, the number of two or more track tags from the number of fit three or more track tags can be predicted:

$$Expected \quad b(c) \text{ tags} = N_{b(c)\text{tag}}^{\geq 3 \text{ track}} * \frac{\epsilon_{b(c)}^{\geq 2 \text{ track}}}{\epsilon_{b(c)}^{\geq 3 \text{ track}}} \quad (5.5)$$

$N_{b(c)\text{tag}}^{\geq 3 \text{ track}}$ is the b (c) fraction of tags with three or more tracks in the secondary vertex obtained with the data $c\tau$ fit. The tagging efficiency, obtained from Monte Carlo, for two (three) or more tracks in the tag is $\epsilon_{b(c)}^{\geq 2 \text{ track}}(\epsilon_{b(c)}^{\geq 3 \text{ track}})$. Table 5.3 lists the tagging efficiency ratios of $\frac{\epsilon_{b(c)}^{\geq 2 \text{ track}}}{\epsilon_{b(c)}^{\geq 3 \text{ track}}}$. Figure 5.4 shows the data $c\tau$ distribution for secondary vertices with three or more tracks, fit to b , c , and background Monte Carlo. The b , c , and background jet content of tagged jets in the jet_50 and jet_20 data samples with three or more tracks in the secondary vertex are listed below.

For the jet_50 sample:

- $79 \pm 7\%$ b jets
- $20 \pm 7\%$ c jets
- $1 \pm 1\%$ Background jets

For the jet_20 sample:

- $77 \pm 13\%$ b jets
- $23 \pm 13\%$ c jets
- $0 + 3\%$ Background jets

The content of the tags is very different than for the two-track tags. There is much more b , about half as much c , and very little background. the number of tags of each type are shown in table 5.2. The uncertainties in the jet content

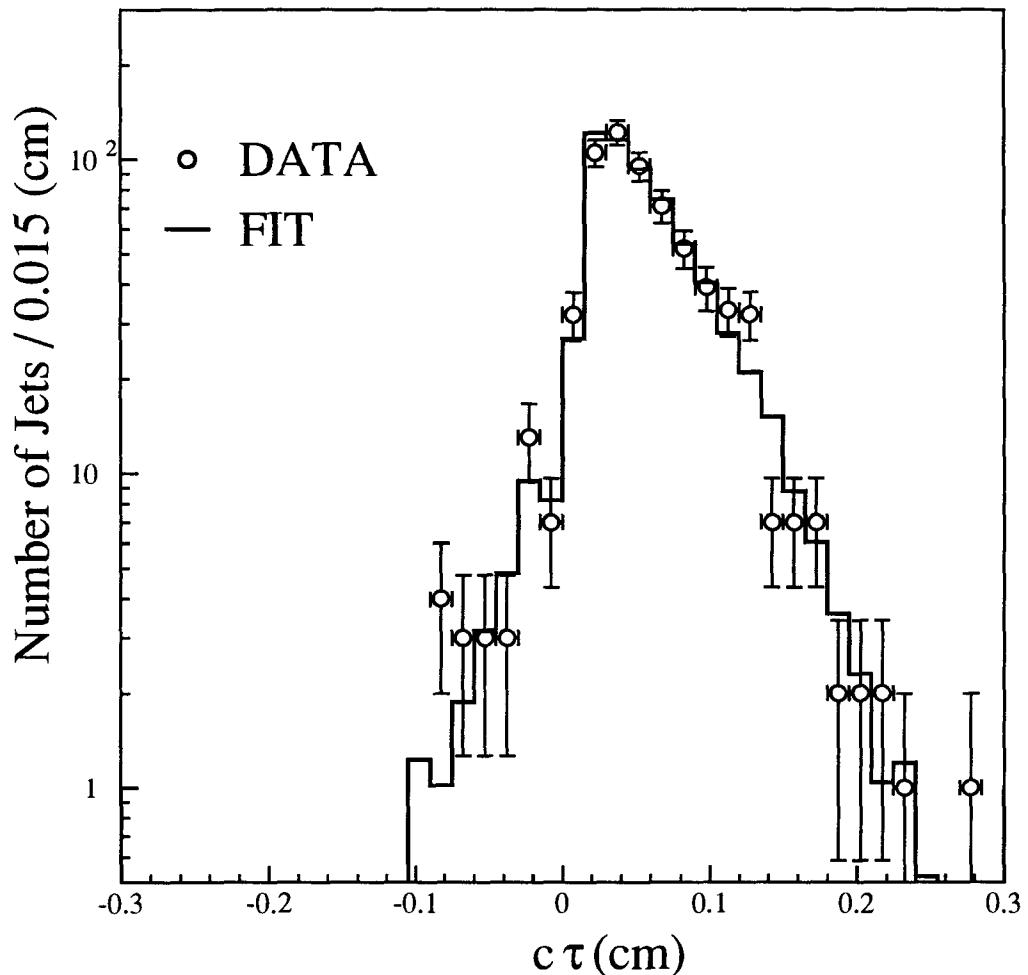


Figure 5.4: The three or more track $c\tau$ distribution in the jet_50 data (points) and fit (histogram) to the Monte Carlo b , c , and background.

B, C, and Background Content of ≥ 3 track Tagged Jets				
Data Tags	Bottom	Charm	Background	
jet_50: 645	510 ± 45	129 ± 45		6 ± 6
jet_20: 458	353 ± 60	105 ± 60		$0 + 3$

Table 5.2: The $c\tau$ fit results for tagged jets containing three or more tracks.

The Ratio of MC Tagging Efficiencies: $\epsilon^{\geq 2 \text{ track}} - \text{to} - \epsilon^{\geq 3 \text{ track}}$		
Monte Carlo Sample	Bottom Ratio	Charm Ratio
jet_50:	2.58 ± 0.07	6.62 ± 0.07
jet_20:	3.87 ± 0.07	7.37 ± 0.16

Table 5.3: The ratio of ≥ 2 - to - ≥ 3 track tagging efficiencies.

	Prediction From ≥ 3 Track Fit	Standard ≥ 2 Track Fit
jet_50 b jets	1316 ± 121	1359 ± 105
jet_50 c jets	854 ± 298	1223 ± 117
jet_20 b jets	1385 ± 233	1363 ± 132
jet_20 c jets	774 ± 443	1290 ± 135

Table 5.4: The b and c content of tagged jets obtained from the $c\tau$ fit of three or more tracks is used to predict the b and c content in tagged jets with two or more tracks.

are due to the fitter uncertainty. With the numbers taken from Table 5.2, and using equation 5.5, the predicted b and c content of the total number of two or more tagged jets is listed in table 5.4. The prediction of the b content agrees well with the standard two or more track $c\tau$ fit results (see Table 5.1). For c jets, the prediction is consistent within uncertainty.

Chapter 6

Results

6.1 Fraction of Bottom and Charm per Jet

The fraction of all jets that come from heavy flavor can be determined using the expression:

$$F_{b(c)} = \frac{N_{\text{Jets}}^{\text{Tagged } b(c)}}{\epsilon_{\text{tag}}^{b(c)} N_{\text{Jets}}^{\text{Total}}} \quad (6.1)$$

$N_{\text{Jets}}^{\text{Tagged } b(c)}$ is the number of tagged b (c) jets determined from the $c\tau$ fit, and $N_{\text{Jets}}^{\text{Total}}$ is the total number of jets in the data sample. The tagging efficiency, $\epsilon_{\text{tag}}^{b(c)}$, determined using Monte Carlo, is defined as the total number of b (c) jets that are tagged, divided by the number of jets that contain at least one b (c) quark. The tagging efficiencies (described in Chapter 4.2) are used to correct the number of b and c tagged jets obtained from the $c\tau$ fit. Using equation 6.1, the absolute b and c fraction per jet is measured. The data results and the generator-level HERWIG and PYTHIA Monte Carlo predictions are listed in Table 6.1. The first uncertainty in the heavy-flavor jet fractions in data is the statistical error and the second is the systematic uncertainty (described below in Section 6.4). The uncertainty in

sample	$F_b(10^{-2})$	$F_c(10^{-2})$
JET_50 data	$2.25 \pm 0.18^{+0.29}_{-0.34}$	$8.0 \pm 0.7^{+1.0}_{-2.1}$
JET_20 data	$2.05 \pm 0.20^{+0.39}_{-0.41}$	$9.0 \pm 0.9^{+1.7}_{-2.6}$
HERWIG	1.88 ± 0.06	3.66 ± 0.08
PYTHIA	2.66 ± 0.08	4.44 ± 0.10

Table 6.1: The b and c -jet fractions integrated over the E_T range from 15 GeV to 135 GeV.

the Monte Carlo fractions are due to the finite amount of statistics in the samples. The HERWIG prediction uses KMRS [39] and MRSA [40] structure functions, $\Lambda_{QCD} = 200$ MeV, and b and c -quark masses of 4.9 and 1.6 GeV/ c^2 . PYTHIA uses CTEQ [41] structure functions, $\Lambda_{QCD} = 213$ MeV, and b and c -quark masses of 4.9 and 1.6 GeV/ c^2 . The difference in c jet content between the jet_20 and jet_50 samples is due to the E_T dependence of F_c and the different E_T spectra of the two data samples. The HERWIG and PYTHIA predictions agree with the measured value of F_b . The measured F_c is about a factor of two higher than the HERWIG and PYTHIA predictions.

The measurement of $F_{b(c)}$ as a function of jet E_T is obtained by performing the $c\tau$ fit separately for each E_T bin, and also determining the Monte Carlo tagging efficiencies for that bin. The resulting distributions of $F_{b(c)}$ for both data and Monte Carlo, are given in Figure 6.1. The F_b and F_b distributions for data compared with the NLO QCD calculation [17] are shown in Figures 6.2 and 6.3. The HERWIG, PYTHIA, and NLO QCD F_b distributions are in agreement with the measured F_b ; the distributions are approximately independent of E_T . The measured F_c distribution falls from about 0.12 at 15 GeV, to 0.04 at 120 GeV. The predicted F_c distributions from HERWIG and PYTHIA are roughly flat at 0.04 over the same E_T range. The NLO QCD calculation gives values of F_c between 0.015 to 0.03

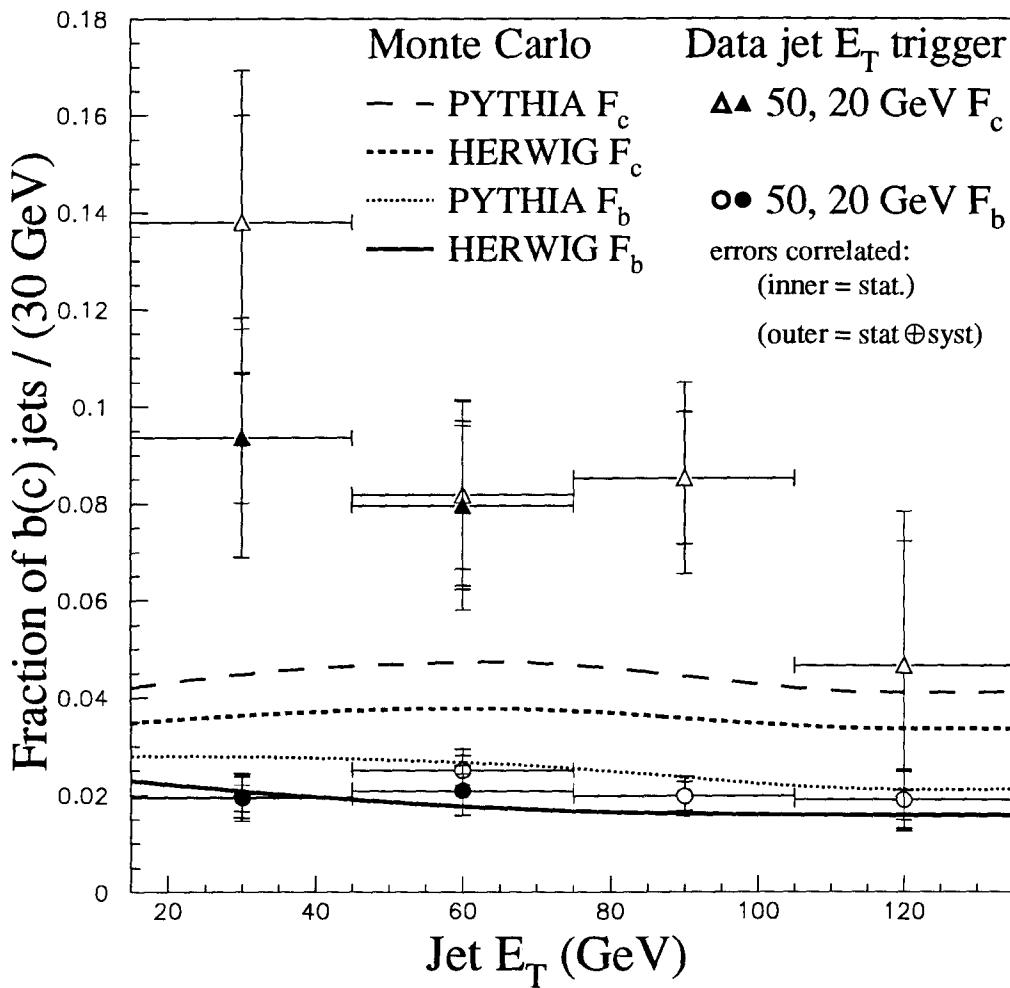


Figure 6.1: The Fraction of b and c jets as a function of E_T . The data (points) are compared with HERWIG and PYTHIA predictions for b (lower contours) and c (upper contours). The open (solid) points indicate the jet_50 (jet_20) data. The uncertainties on data points are correlated. The inner error bars are the statistical uncertainty, the outer error bars are the statistical and the systematic uncertainties added in quadrature.

	50 GeV Trigger	20 GeV Trigger
<i>b</i> jets	0.0455 ± 0.0015	0.0345 ± 0.0030
<i>c</i> jets	0.0070 ± 0.0006	0.0075 ± 0.0015

Table 6.2: Monte Carlo *b* and *c* jet double-to-all tag ratios. The uncertainties are statistical only.

depending upon the choice of factorization/renormalization scale. These values are also independent of E_T in the range 15 GeV to 100 GeV.

6.2 Cross Checks

6.2.1 Double Tags

Double-tagged jet events, when two or more jets separately are tagged in the same event (see Figure 6.4), may be used for an independent check on the $c\tau$ method of obtaining the *b* content of tagged jets. From a Monte Carlo simulation the ratio of the total number of jets in double-tagged events to the total number of tags is determined. As listed in Table 6.2, this ratio (shown with statistical uncertainties) is many times larger for *b* jets than for *c* jets. Since the $c\tau$ fit results indicate that the number of *b* and *c* tagged jets are almost equal, the majority of double tags result from *b* jets. Figure 6.5 shows the $c\tau$ distribution for jets in double-tagged data events compared with *b* jets in double-tagged events from HERWIG Monte Carlo.

In the jet_50 data sample, there are $78.0 \pm 6.2(stat)$ tagged jets in 39 double-tagged events observed. In the jet_20 data sample, there are $48.0 \pm 4.9(stat)$ tagged jets in 24 double-tagged events. Using the ratios in Table 6.2 to correct for a small *c* content, the number of *b* jets in data double-tagged events is listed in Table 6.3.

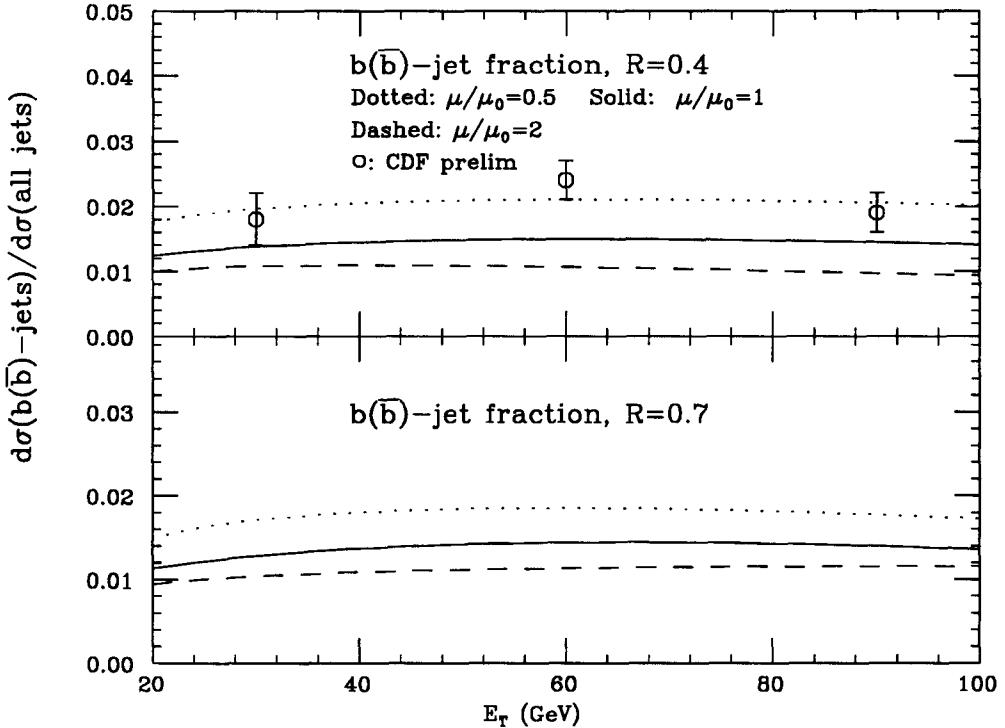


Figure 6.2: The next-to-leading order QCD calculation of the Ratio of b -jet to inclusive-jet E_T distributions, for different choices of renormalization and factorization scales ($\mu_R = \mu_F = \mu$) for jet cone sizes $R = 0.4$ and $R = 0.7$. The data points for $R = 0.4$ are shown with statistical uncertainty only.

	jet_50	jet_20
Data jets	78 ± 6.2	48 ± 4.9
b jets	67.6 ± 9.2	39.4 ± 9.3

Table 6.3: b jet Content of Double-Tagged Events. The uncertainties are statistical only.

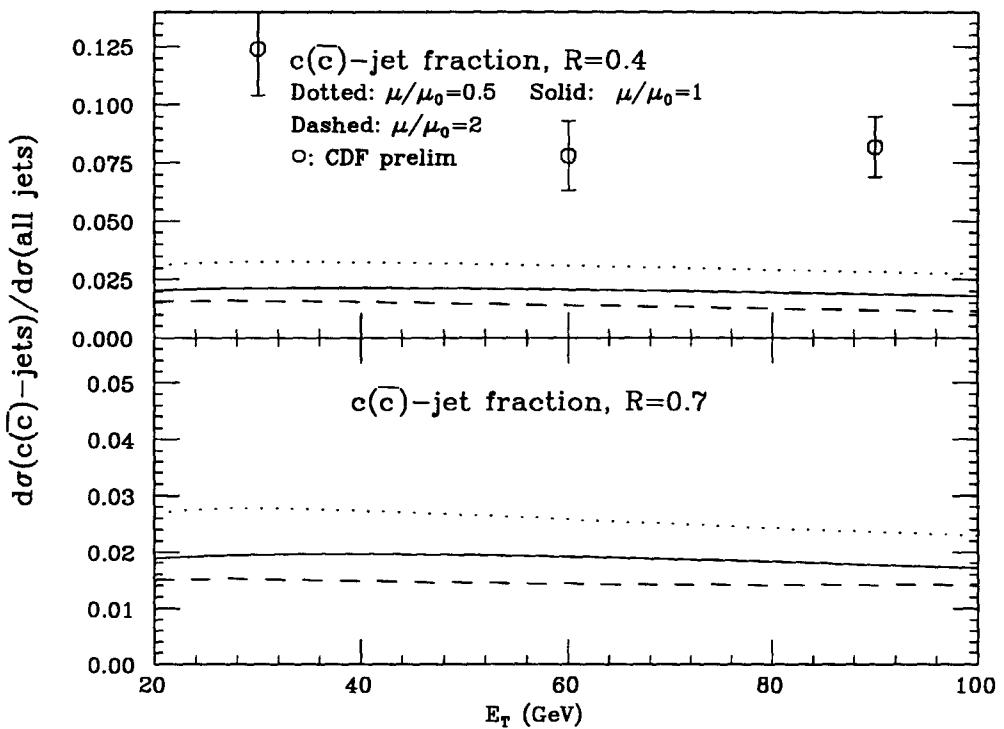


Figure 6.3: Next-to-leading order QCD calculation of the Ratio of c -jet to inclusive-jet E_T distributions, for different choices of renormalization and factorization scales ($\mu_R = \mu_F = \mu$) for jet cone sizes $R = 0.4$ and $R = 0.7$. The data points for $R = 0.4$ are shown with statistical uncertainty only.

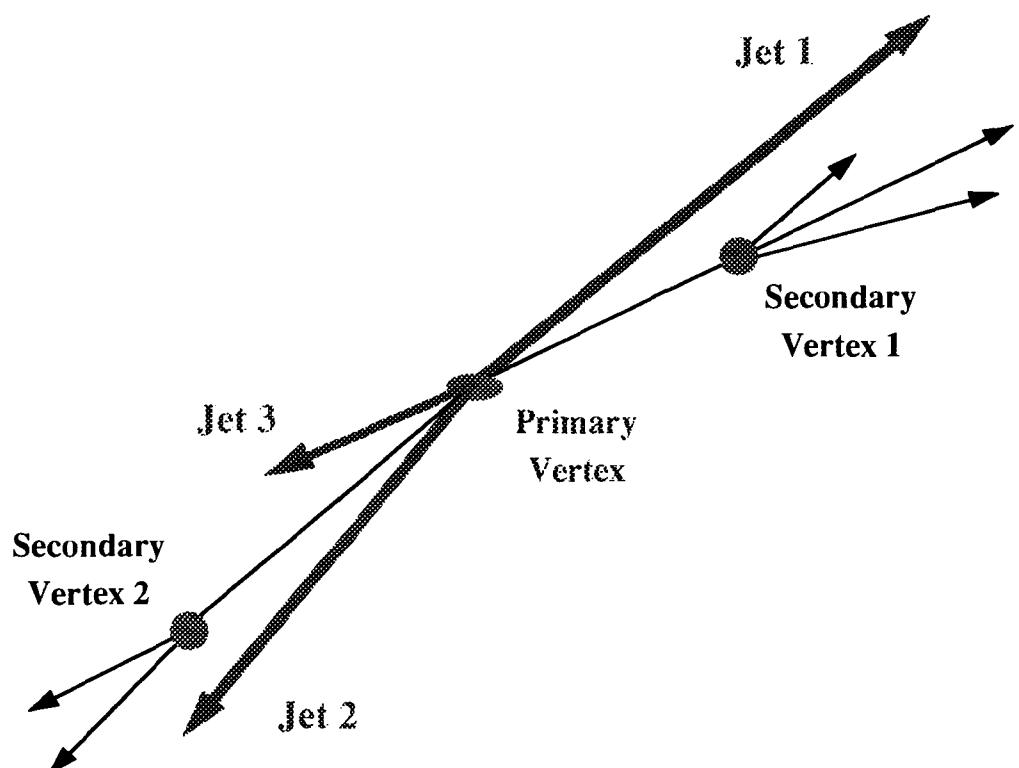


Figure 6.4: Schematic showing a double-tagged event. Double-tagged jet events are defined as two or more jets separately tagged in the same event.

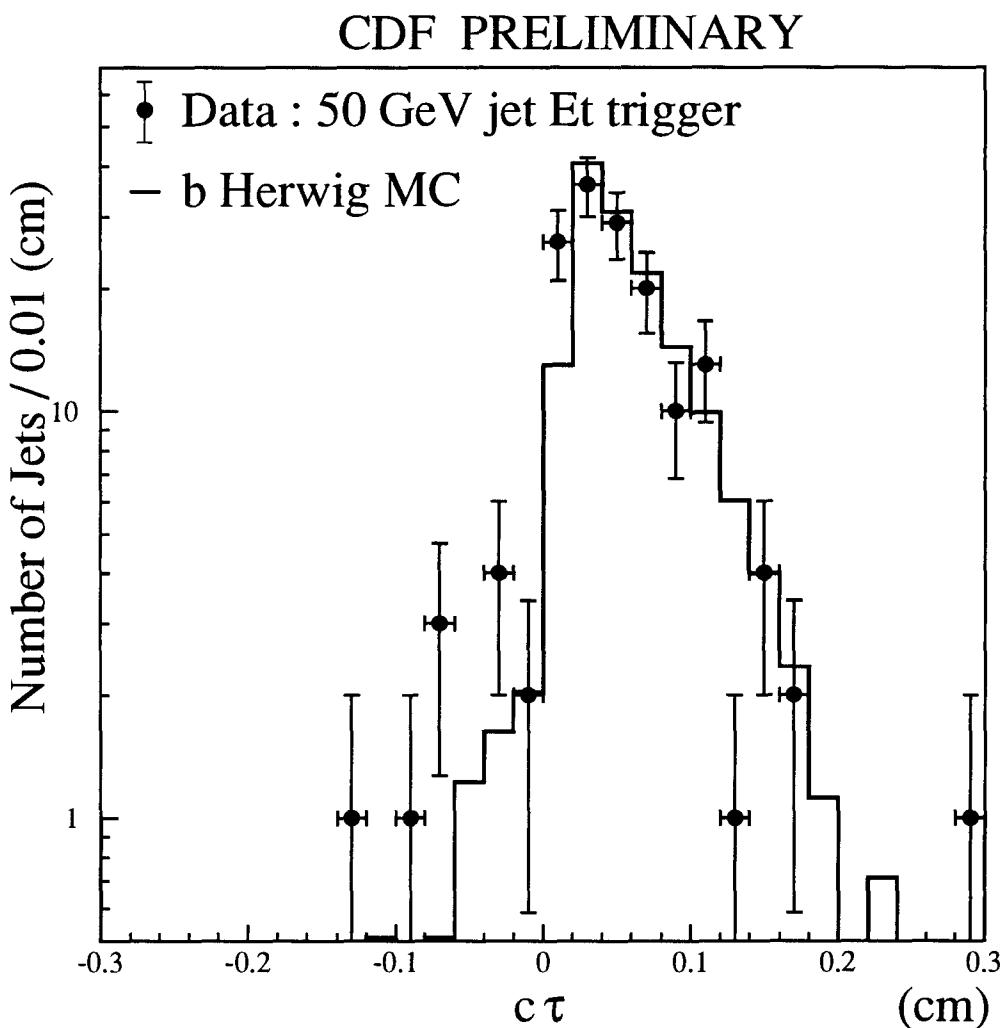


Figure 6.5: The $c\tau$ distribution for jets in double-tagged events. The HERWIG Monte Carlo (histogram) $c\tau$ distribution for b jets is normalized to the data (points).

Data sample	Double-tag prediction	$c\tau$ fit
jet_50	1485 ± 208	1359 ± 105
jet_20	1142 ± 288	1363 ± 132

Table 6.4: Double-tag prediction of the number of tagged b jets. The uncertainties are statistical only.

To predict the total number of tagged b jets in the data, the number of b jets in the double-tagged data events in Table 6.3 is scaled by the double-to-all tag ratios determined from Monte Carlo (Table 6.2). The predictions are shown in Table 6.4. For both data samples, the predicted numbers of b jets are in good agreement with the $c\tau$ fit results (see Table 5.1).

To predict the value of F_b , an expression similar to equation 6.1 is used with the predicted number of b jets from Table 6.4. At an average jet E_T of 58 GeV, the predicted value of F_b in the jet_50 sample is $(2.2 \pm 0.3(stat)) \times 10^{-2}$. This is in good agreement with the absolute measurement of $F_b = (2.25 \pm 0.18(stat)^{+0.29}_{-0.34}(syst)) \times 10^{-2}$ and with the F_b distribution in Figure 6.1 determined using the $c\tau$ fit method. In the jet_20 sample, the predicted b jet fraction is $F_b = (1.7 \pm 0.5(stat)) \times 10^{-2}$ at an average jet E_T of 35 GeV. Again, this is in agreement with the integrated measurement of $F_b = (2.05 \pm 0.20(stat)^{+0.39}_{-0.41}(syst)) \times 10^{-2}$ and with the F_b distribution resulting from the $c\tau$ fit.

6.2.2 D^* Production in Jets

The measurement of the production of the charm meson $D^{*\pm}$ in jet events [42] is an independent check on the c -jet content. An ISAJET Monte Carlo simulation predicts that the observed $D^{*\pm}$'s are predominately produced from direct charm and not from bottom decay. The charm content in jets is checked by comparing the

number of fully reconstructed $D^{*\pm}$ candidates in jets, to the number of expected candidates using measured values of the F_b and F_c from $c\tau$ fit method.

$D^{*\pm}$ mesons are reconstructed in jet events using the exclusive decay channel $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K^-\pi^+$ (the conjugate modes of this decay are implied). The decay channel $D^{*+} \rightarrow D^0\pi^+$ is used because of its very low Q value ($m_{D^*} - m_{D^0} - m_\pi = 6 \text{ MeV}/c^2$), reducing the phase space for background. The π^+ in this decay (this “soft”, low P_T , pion will hereafter be denoted as π_s) takes a relatively small fraction of the D^{*+} momentum, providing a convenient signature to identify the decay channel.

A standard [43] method is used to search for $D^{*\pm}$ candidates. First, all possible charged track combinations of $K^\mp\pi^\pm\pi_s^\pm$ are found (assuming all tracks are either K or π), keeping the correct charge correlation (“right sign”) between the K and the π_s . Next, the 2-track ($D^0 \rightarrow K\pi$) invariant mass is formed. The D^0 candidate must have a mass within a $30 \text{ MeV}/c^2$ window centered the world average D^0 mass of $1864.6 \pm 0.5 \text{ MeV}/c^2$ [32]. The invariant mass is formed with the D^0 ($K\pi$) candidate and the soft pion (π_s). Finally, $D^{*\pm}$ candidates will be evident as an excess of events over combinatoric background in the distribution of the invariant mass difference (ΔM) between the $K\pi\pi_s$ and the $K\pi$ combinations. The combinatoric background for ΔM is obtained with “right sign” combinations with a $\text{Mass}(K^-\pi^+) > 2.0 \text{ GeV}/c^2$ (well above the true D^0 mass), normalized to the ΔM distribution above the nominal $D^{*\pm} - D^0$ mass difference. A combinatoric background distribution composed of “wrong sign” (incorrect charge correlation between the K and the π_s) $K\pi\pi_s$ combinations is very similar in shape.

The following kinematic requirements must be satisfied for all tracks used in $D^{*\pm}$ candidates; Charged tracks must have at least four hits on two or more CTC axial superlayers and at least two hits on two or more CTC stereo superlayers. If the track is linked to hits in the SVX, only two hits in the SVX are required. All

tracks must have $\Delta Z = |Z^{CTC} - Z^{PRIMARY}| < 5$ cm. D^* candidates must pass the following criteria:

- $P_T^K > 1$ GeV/c
- $P_T^\pi > 0.7$ GeV/c
- $P_T^{\pi_s} > 0.4$ GeV/c
- $|M(K\pi) - 1864.6| < 30.0$ MeV/ c^2
- $P_T^{D^*} > 6.0$ GeV/c

All D^* candidates must be associated with a jet. This is accomplished by specifying that the cosine of the angle between the 3-momenta of the D^* candidate and the nearest jet to be greater than 0.8. All jets must have $E_T \geq 15.0$ GeV, detector $|\eta| < 2.0$, $|Z_{vertex}^{event}| < 30.0$ cm, and cone radius $R = 0.4$ ($R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$).

The upper plot in Figure 6.6 shows the ΔM distribution for “right sign” $K\pi\pi_s$ combinations. The solid histogram denotes the D^* candidates and the dashed histogram is background composed of combinations with mass of the $(K\pi)$ greater than 2.0 GeV/ c^2 , normalized to the ΔM distribution above 0.151 GeV/ c^2 . A significant excess is observed over background at the world average $D^{*\pm} - D^0$ mass difference of 145.42 ± 0.05 MeV/ c^2 [32]. The lower plot in the same Figure shows a minimum χ^2 fit (solid line) to the ΔM distribution for D^* candidates (points). The ΔM distribution is fit to the following function:

$$A * (\Delta M - M_\pi)^B + P1 * \exp(-(\Delta M - P2)^2 / 2 * P3^2) \quad (6.2)$$

The first term of the function describes the background shape of the distribution. Parameters A and B are free-floating constants determined by the fit, and M_π is the pion mass. The second term in the function is a gaussian distribution describing

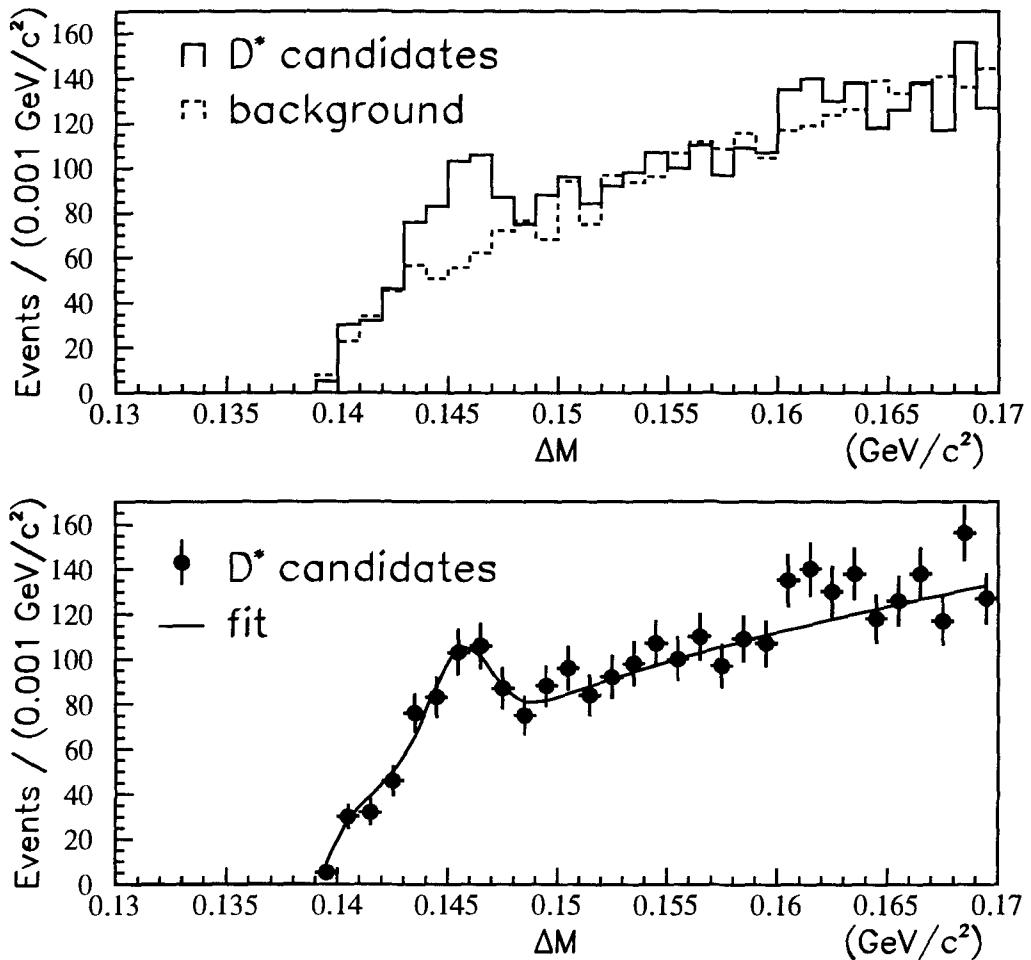


Figure 6.6: The $K\pi\pi - K\pi$ mass difference (ΔM). In the upper plot, the solid histogram is the distribution for D^* meson candidates, and the dashed histogram is background taken from the $K\pi$ (D^0) mass sideband. In the lower plot, the points are the D^* meson candidates, and the solid line is the fit. The total number of reconstructed D^* candidates is $143 \pm 28(\text{stat.}) \pm 20(\text{syst.})$.

the shape of the signal. The amplitude ($P1$), mean ($P2$), and width ($P3$) of the gaussian are also free-floating parameters determined by the fit. The total number of reconstructed $D^{*\pm}$'s is taken as the area of the gaussian distribution returned from the fit. This is a reasonable approach given the shape of the background distribution. The total number of reconstructed $D^{*\pm}$ candidates obtained from the fit is $143 \pm 28(\text{stat.}) \pm 20(\text{syst.})$. The statistical error is the uncertainty returned by the fit. The systematic uncertainty is described below. The mean of the gaussian distribution returned by the fit is $145.5 \pm 0.3 \text{ MeV}/c^2$. This value agrees with the world average $D^{*\pm} - D^0$ mass difference.

The $\pm 14\%$ systematic uncertainty on the total number of $D^{*\pm}$ candidates is estimated by taking the difference between the area of the gaussian fit and two alternative methods for obtaining the total number of reconstructed $D^{*\pm}$. In the first method, the total number of $D^{*\pm}$ candidates is taken as the excess number of events above background in the signal window of $|\Delta M - 145.4| < 3 \text{ MeV}/c^2$. The number of background events is found by integrating the background function of $A * (\Delta M - M_\pi)^B$ over the signal window. For the second method, the total number of $D^{*\pm}$ candidates is determined by performing a straight subtraction of the number of background events from the total number of events in the signal window (See Figure 6.6).

Using the measured values of F_b and F_c , the number of *For good* quality SVX tracks in jet data, expected $D^{*\pm}$ candidates from b and c jets is determined with the following equations:

$$N_b^{D^*} = F_b N_{jets} Br(B \rightarrow D^{*\pm} X) Br(D^{*\pm} \rightarrow D^0 \pi^\pm) Br(D^0 \rightarrow K^\mp \pi^\pm) \epsilon_{RE}^b \quad (6.3)$$

$$N_c^{D^*} = F_c N_{jets} P(c \rightarrow D^{*\pm} X) * Br(D^{*\pm} \rightarrow D^0 \pi^\pm) Br(D^0 \rightarrow K^\mp \pi^\pm) \epsilon_{RE}^c \quad (6.4)$$

The total number of jets in the sample is N_{jets} . $P(c \rightarrow D^{*\pm} X)$ is the probability that a charm quark fragments to $D^{*\pm} X$. From an ALEPH measurement[44], the value of $P(c \rightarrow D^{*\pm} X)Br(D^{*\pm} \rightarrow D^0\pi^\pm)Br(D^0 \rightarrow K^\mp\pi^\pm)$ is $(6.5 \pm 0.5(stat.) \pm 0.2(syst.)) \times 10^{-3}$. The following branching ratios are taken from the Particle Data Book [32]:

- $Br(B \rightarrow D^{*\pm} X) = (23.1 \pm 3.3)\%$
- $Br(D^{*\pm} \rightarrow D^0\pi^\pm) = (68.3 \pm 1.4)\%$
- $Br(D^0 \rightarrow K^\mp\pi^\pm) = (3.83 \pm 0.12)\%$

The reconstruction efficiency (ϵ_{RE}), obtained using ISAJET [29] Monte Carlo of $b\bar{b}$ and $c\bar{c}$ events, is defined as the total number of reconstructed $D^{*\pm}$ passing the criteria listed above, divided by the total number of produced $D^{*\pm}$ associated with jets. For each event the sequential decay $D^{*\pm} \rightarrow D^0\pi^\pm, D^0 \rightarrow K^\mp\pi^\pm$ is forced to occur 100% of the time. The detector simulation is described in Chapter 3. The reconstruction efficiency for $D^{*\pm}$ produced from $b\bar{b}$ events is $(19.6 \pm 0.6)\%$ and $(33.0 \pm 0.5)\%$ from $c\bar{c}$ events. The $b\bar{b}$ reconstruction efficiency is lower than $c\bar{c}$ because $D^{*\pm}$'s from the decay of B hadrons tend have a softer energy distributions than $D^{*\pm}$'s produced from $c\bar{c}$. Scaling branching ratios by the reconstruction efficiencies, and assuming equal production rates, the ratio of $D^{*\pm}$ produced by charm to that produced by bottom is approximately 1.6-to-1. Furthermore, assuming a c -to- b ratio in jets from HERWIG of 1.9-to-1, then the ratio of $D^{*\pm}$ produced by charm to that produced by bottom is at least 3-to-1. Thus fully reconstructing $D^{*\pm}$ is primarily a check on the charm content.

The number of expected $D^{*\pm}$ candidates from b and c jets are obtained from equations 6.3 and 6.4:

- Number of $D^{*\pm}$ from c jets is $83 \pm 11(stat)^{+16}_{-24}(syst)$

- Number of $D^{*\pm}$ from b jets is $10 \pm 2(stat) \pm 2(syst)$
- Expected $D^{*\pm}$ candidate Total = $93 \pm 11(stat)^{+19}_{-26}(syst)$

The systematic uncertainty in the expected values comes from the measured values of F_b and F_c in the jet_20 sample. The total number of expected $D^{*\pm}$ candidates is consistent with the total number reconstructed. It is important to note that the expected number of $D^{*\pm}$ is not greater than the total reconstructed. If the measured value of F_c is high by a factor of two, the expected number of $D^{*\pm}$ candidates would also be larger than the total reconstructed by approximately the same factor.

Finally, $D^{*\pm}$ candidates are not successfully reconstructed in the jet_50 sample using this method because the signal-to-background ratio is poor compared to the jet_20 sample. This is mainly due to the increased combinatoric background caused by the higher average track multiplicity per jet in the jet_50 sample. In the jet_20 sample, the ratio of signal-to-background is 1-to-3. With the measured values of F_c and F_b for the jet_50 sample in Table 6.1, and using equations 6.3 and 6.4, 45 $D^{*\pm}$ candidates are expected in the jet_50 sample. A total of 748 background events is estimated by counting the number of “wrong sign” combinations in the signal window of the ΔM distribution. This results in an estimated signal-to-background ratio of nearly 1-to-17, more than 5 times worse than the jet_20 sample.

6.2.3 Inclusive Bottom and Charm Quark Cross Section

As an additional cross check of the F_b and F_c measurements, the inclusive cross sections for bottom and charm quark production are calculated using the results of the $c\tau$ fit method. The results are compared with CDF measurements for bottom and theoretical predictions for both bottom and charm. In this measurement, the tagged jets are required to pass the trigger threshold (“trigger jet”) of 50 GeV in

the jet_50 sample. This is done to reduce systematic uncertainties in the calculation of the trigger acceptance. The cross section for $p\bar{p} \rightarrow bX$ (cX) is measured for quarks with $|y| < 1.0$ and $P_T > P_T^{min}$ (90% of all quarks have a P_T greater than P_T^{min}). From the Monte Carlo simulation, b quarks are found have a $P_T^{min} = 35.0^{+3}_{-1}$ GeV/c, and c quarks have a $P_T^{min} = 25.0 \pm 2$ GeV/c. The values for P_T^{min} are less than the jet E_T trigger threshold because the dominate $b\bar{b}$ ($c\bar{c}$) production mode is gluon splitting. From HERWIG Monte Carlo, about 50% of gluons that fragment to $b\bar{b}$ or $c\bar{c}$ have both quarks inside the same jet, each quark sharing approximately equal amounts of the total energy. Thus in a 50 GeV $g \rightarrow b\bar{b}$ jet, both the b and \bar{b} quarks would have an energy of ~ 25 GeV.

The equation below is used to calculate the inclusive cross section for b and c quarks:

$$\sigma_{b(c)} = \frac{N_{fit}^{b(c)} * 1/2}{\mathcal{L} * \epsilon_{acc}^{b(c)}} \quad (6.5)$$

- $N_{fit}^{b(c)} * 1/2$ is the number of b or c -trigger jets from the $c\tau$ fit. $N_{fit}^b = 595 \pm 62$ and $N_{fit}^c = 509 \pm 72$. The factor of 1/2 is to count only a b (c) quark or \bar{b} (\bar{c}) quark, but not both.
- $\mathcal{L} = (13100/20) \text{ nb}^{-1}$ is the luminosity of the jet_50 sample, corrected for prescaling. The uncertainty on the luminosity is 3.3% [22].

The acceptance, $\epsilon_{acc}^{b(c)}$, is determined using the following relation:

$$\epsilon_{acc}^{b(c)} = \epsilon_Q^{QJ} * \epsilon_{QJ}^{Trig \; QJ} * \epsilon_{Trig}^{Tag \; QJ} \quad (6.6)$$

- ϵ_Q^{QJ} : The efficiency that a b (c) quark, with $P_T > P_T^{min}$ and $|y| < 1$, is associated with a jet with $E_T > 50.0$ GeV.
- $\epsilon_{QJ}^{Trig \; QJ}$: The efficiency that b (c) jets pass the jet_50 trigger threshold.

Monte Carlo	ϵ_Q^{QJ}	$\epsilon_{QJ}^{Trig\;QJ}$	$\epsilon_{Trig\;QJ}^{Tag}$
<i>b</i> jets	0.089 ± 0.009	0.657 ± 0.008	0.404 ± 0.009
<i>c</i> jets	0.031 ± 0.003	0.644 ± 0.005	0.092 ± 0.004

Table 6.5: Efficiencies for the acceptance calculation.

- $\epsilon_{Trig\;QJ}^{Tag}$: The tagging efficiency of the trigger jets.

These efficiencies are listed in Table 6.5. The inclusive cross section for $p\bar{p} \rightarrow bX$ is $\sigma_b = 19 \pm 2(stat.) \pm 4(syst.)$ nb and for $p\bar{p} \rightarrow cX$ is $\sigma_c = 211 \pm 30(stat.)^{+34}_{-59}(syst.)$ nb (the systematic uncertainties will be described in the following section). The cross section measured for *b* is consistent with other CDF results, as shown in Chapter 1.2 (Figure 1.5). In addition, the *b* measurement is consistent with the NLO theory prediction shown in Figure 6.7. The inclusive *c* cross section, shown in the same Figure, is higher than the theoretical prediction.

6.3 $b\bar{b}$ Production Correlations

As a by-product of the analysis discussed thus far, the double-tagged events, a very pure $b\bar{b}$ sample (see Section 6.2.1), can be used to explore other aspects of the $b\bar{b}$ production dynamics. Heavy flavor production can be represented approximately by the superposition of three classes of processes: direct production via light quark-antiquark annihilation or gluon-gluon fusion, final state gluon splitting, and flavor excitation. Figure 6.8 shows examples of possible Feynman diagrams for these processes.

The azimuthal angle between the two tagged jets, $\Delta\phi$, is used to study $b\bar{b}$ correlations [47]. Figure 6.9 shows the $\Delta\phi$ distribution between the tagged *b* jets in

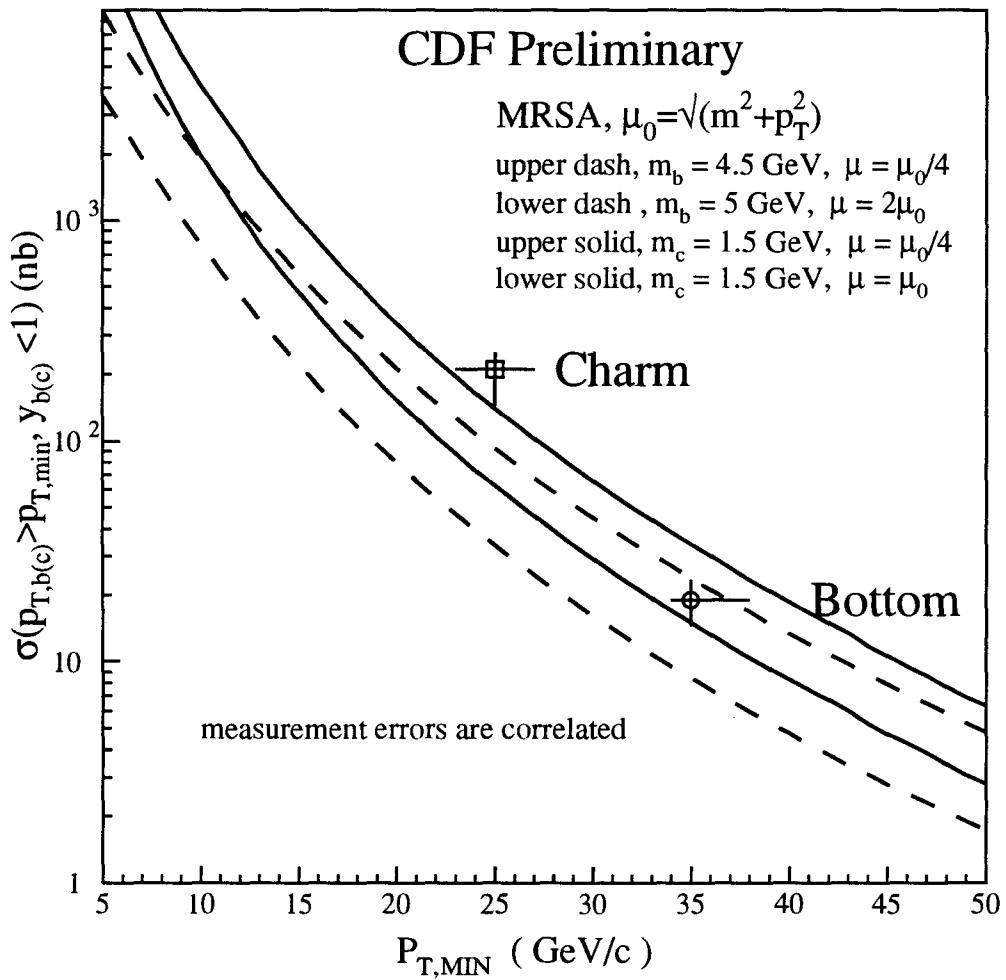


Figure 6.7: The inclusive b and c -quark cross sections measured with the $c\tau$ fit method compared with the NLO QCD predictions (for different choices of renormalization/factorization scales and choice of b -quark mass). The errors on the measured values for bottom and charm are the statistical and systematic uncertainties added in quadrature.

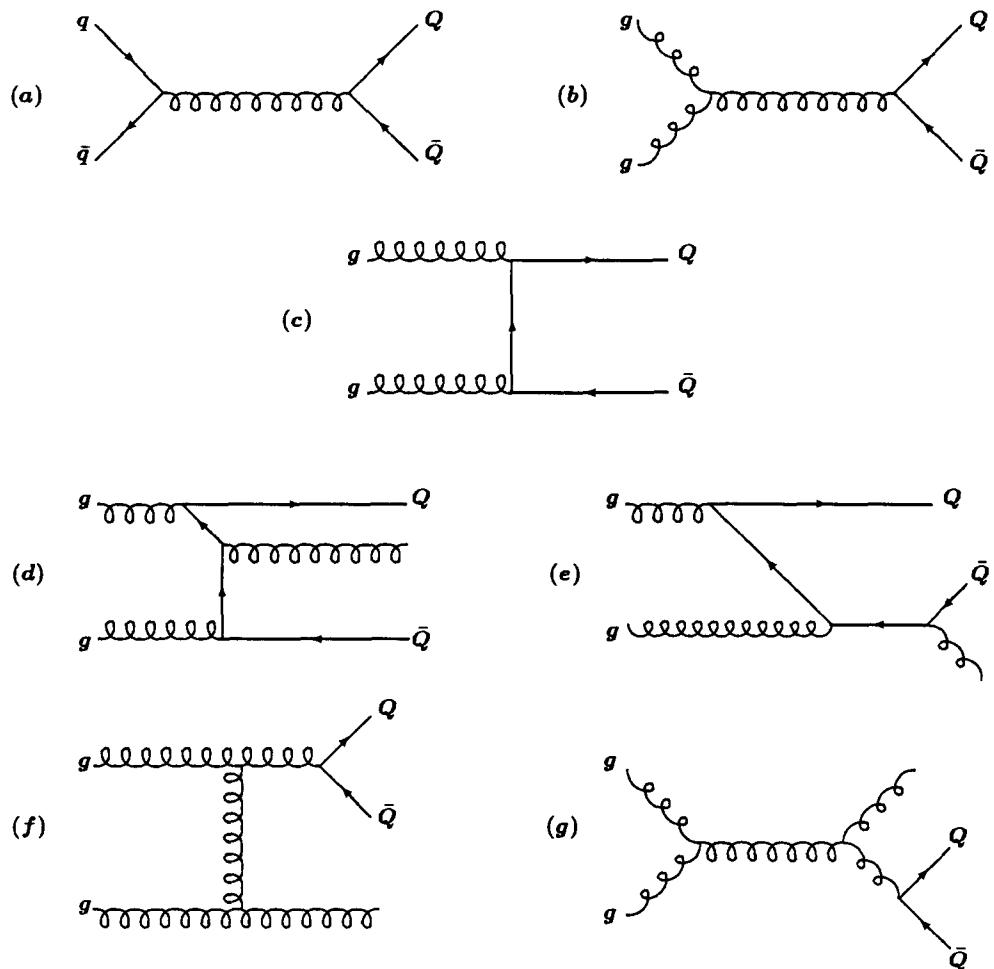


Figure 6.8: Examples of Feynman diagrams depicting heavy flavor production. Light quark-antiquark annihilation is shown in diagram (a), gluon-gluon fusion is depicted in diagrams (b) and (c). Diagrams (d) and (e) show examples of flavor excitation (initial state gluon splitting). Final state gluon splitting is depicted in diagrams (f) and (g).

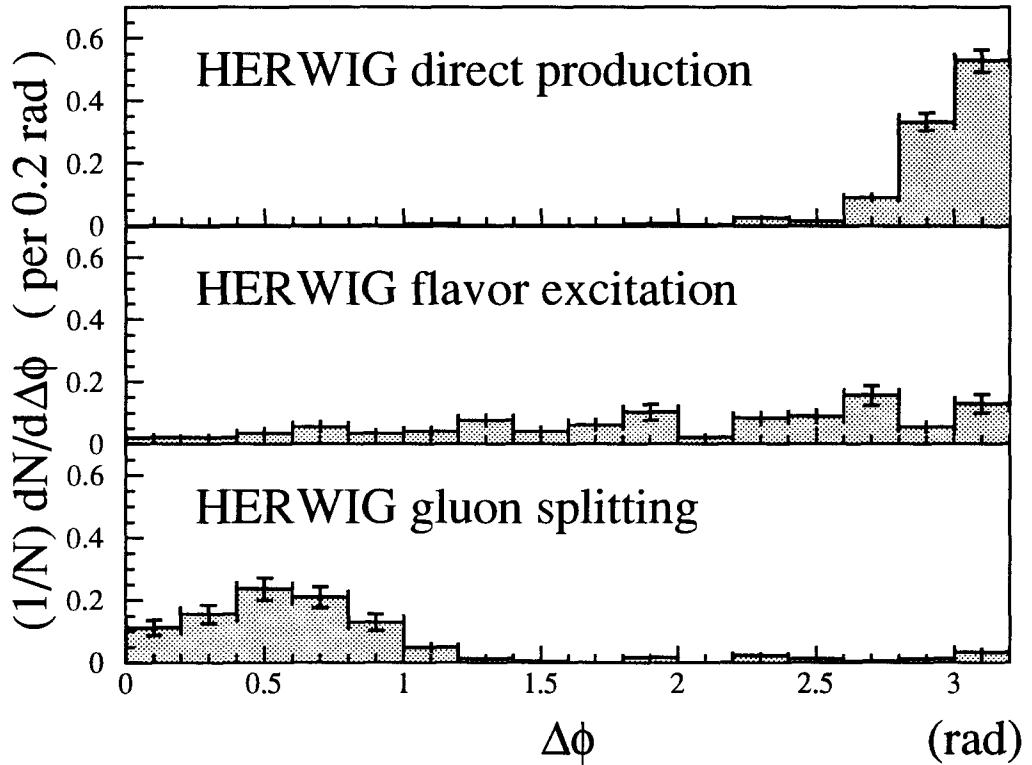


Figure 6.9: The $\Delta\phi$ distribution for double-tagged $b\bar{b}$ Monte Carlo events. The figures show the $\Delta\phi$ distributions for the three production processes.

the Monte Carlo sample of double-tagged events, separated by production process. The distributions in $\Delta\phi$ for each process are distinct. In direct production, the jets produced are dominantly back-to-back, while jets from final state gluon splitting tend to be collinear. B jets produced through the flavor excitation mode are more uniformly distributed in $\Delta\phi$.

By fitting the $\Delta\phi$ distribution in jet_50 data to the sum of the three sub-process distributions from b Monte Carlo, the relative contributions of each subprocess is

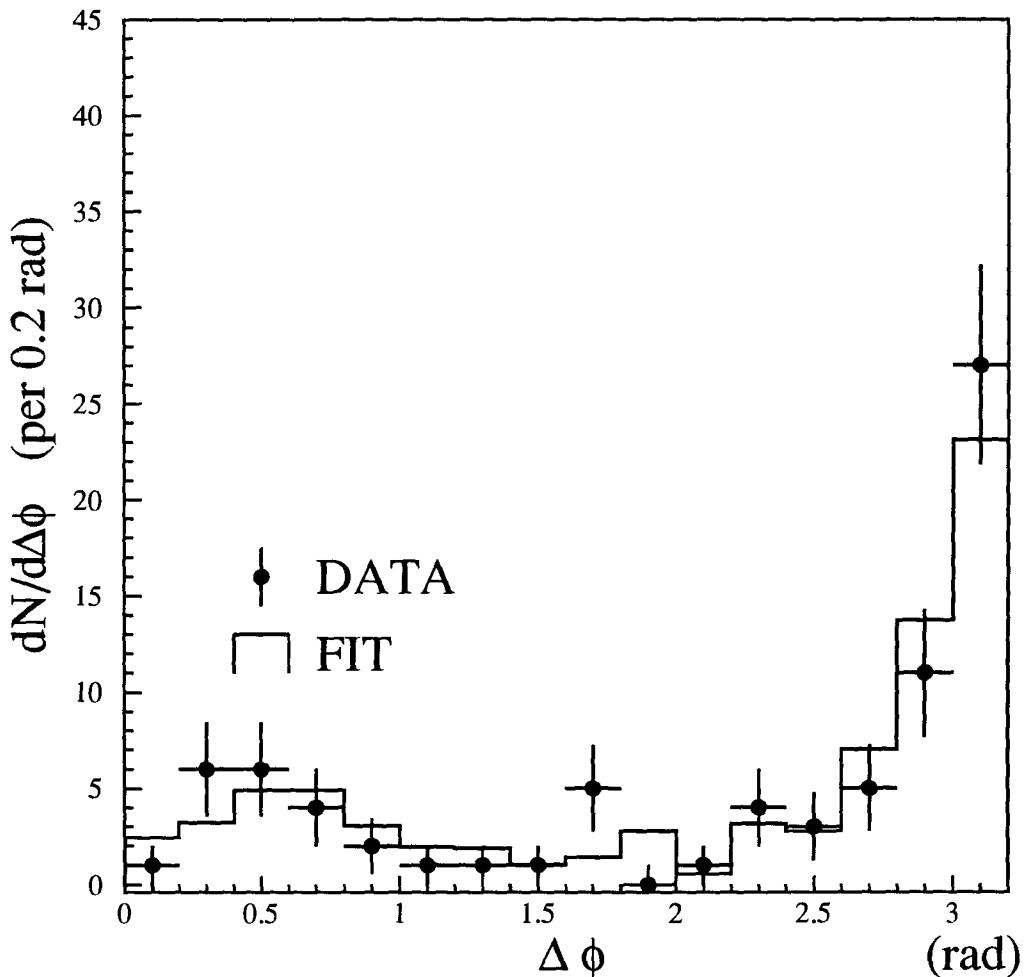


Figure 6.10: The $\Delta\phi$ distribution for double-tagged b data (points) events fit (histogram) to the sum of the three sub-processes from $b\bar{b}$ Monte Carlo.

determined. The fit is shown in Figure 6.10, and the resultant fractions of each production mode are listed in Table 6.3. The fractions are in good agreement with the respective contributions from HERWIG. The same level of agreement is seen in the jet.20 sample.

	Direct	Flavor excitation	Gluon splitting
Fit results	0.59 ± 0.11	0.17 ± 0.11	0.24 ± 0.08
b Monte Carlo	0.55 ± 0.02	0.20 ± 0.02	0.25 ± 0.02

Table 6.6: The relative fractions of each production mode determined by fitting the $\Delta\phi$ distribution of the double-tagged data, compared with HERWIG b Monte Carlo. The uncertainty for the data fit results is the fitter error. The uncertainty in the Monte Carlo fractions are the statistical error.

6.4 Systematic Uncertainty

The systematic uncertainties in the measurement of F_b and F_c are estimated by measuring various effects (integrated over all E_T and as a function of E_T) that influence the tagging efficiencies and the $c\tau$ fitting. The effects and their impact upon the b and c jet fractions are described below. The overall uncertainties are listed in Table 6.4. The values reported in this section are percentages relative to the measured values of F_b and F_c . The systematic uncertainties binned in jet E_T are shown in Figure 6.1.

The systematic uncertainty due to the variance in the overall tagging efficiency as a function of jet E_T , is estimated by combining the uncertainties in $F_{taggable}^{scale}$ and F_{tag}^{scale} . The systematic uncertainty from $F_{taggable}^{scale}$ is determined by taking the maximum difference per E_T bin in the scale factors calculated for b , c , and background (non-heavy flavor) jet Monte Carlo. Because most of the jets in the inclusive jet sample are not from heavy flavor, this is a conservative estimate. The overall E_T -weighted uncertainty in the taggable rate in the jet_50 and jet_20 Monte Carlo is $\pm 4\%$ and $\pm 14\%$ respectively. In Figure 4.14 the tagging efficiency scale factor, F_{tag}^{scale} , is independent of jet E_T . Therefore it is reasonable to fit the distribution to a constant factor. The $\pm 1\sigma$ error in the fit of $\pm 8\%$ is taken as the measure of the systematic uncertainty in the tagging efficiency scale factor.

The full systematic uncertainty is determined by adding the contributions from the taggable scale factor in quadrature with the tagging efficiency scale factor. In the jet_50 Monte Carlo, the full systematic uncertainty is $\pm 9\%$, and for the jet_20, the uncertainty is $\pm 16\%$. There is not a charm enriched data sample analogous to the inclusive electron sample with which the charm tagging efficiency can be checked. Therefore a reasonable assumption is made that both the scale factor and its associated systematic uncertainty can be used for charm tagging as well.

It is assumed that the $c\tau$ distribution observed in data tagged jets is due to a combination of b , c , and (non-heavy flavor) background. The Monte Carlo model of the SVX/CTC tracking is relied upon to determine the b and c tagging efficiencies used in obtaining F_b and F_c . Since the Monte Carlo can also be used to determine the background tagging efficiency, all of the information is available to predict the number of jets before tagging, as a consistency check. This prediction can be compared to the observed number of jets. Any discrepancy provides information on the validity of the original assumptions. The following relation is used to predict the total number of jets before tagging:

$$N_{\text{calc.}}^{\text{predict}} = \text{Tagged Jets} * \left(\frac{\%b}{\epsilon_{\text{tag}}^b} + \frac{\%c}{\epsilon_{\text{tag}}^c} + \frac{\%BKG}{\epsilon_{\text{tag}}^{\text{BKG}}} \right) \quad (6.7)$$

where ϵ_{tag}^b and ϵ_{tag}^c are the bottom and charm tagging efficiencies (see Chapter 4.2) and $\epsilon_{\text{tag}}^{\text{BKG}}$ is the background tagging efficiency. The total number of jets predicted is $\sim 15\%$ lower (but within uncertainty) than what is observed in data. However, because the background tagging efficiency is smaller by more than a factor of twenty compared to the b and c jet tagging efficiencies, equation 6.7 is primarily a check on the background tagging efficiency rather than a check on the knowledge of b or c production. A systematic uncertainty can be set on the model using the above relation as a constraint in the fit. Within uncertainties, the fit is constrained

to yield the observed number of jets before tagging. The constraint lowers the c jet fit fraction by 23% of its value and lowers the b jet fit fraction by 6.4%. As a conservative measure, a negative systematic is set for F_b at -6.4% and for F_c -23.0% . The effect of the constraint is greatly diminished for both the b and c fits as a function of jet E_T . The larger uncertainties in the data points, from statistically smaller samples binned in E_T , permits more tolerance in the fit. The average uncertainty as a function of jet E_T is $\sim 1\%$ for F_b and $\sim 6\%$ for F_c .

It is assumed that the background $c\tau$ distribution is symmetric about $c\tau = 0$. This may not be the case and it could change the b and c jet fractions on the $c\tau$ fit. The sources of background in data come from: mistags, K_S^0 , Λ^0 's, and interactions in the silicon (the innermost layer of the SVX). Since K_S^0 and Λ^0 's are explicitly removed, and L_{XY} is required to be less than 2.5 cm (this excludes the innermost layer of the SVX), most of the background comes from mistags. Tagged jets in data are used to set a bound on the ratio of positive to negative decay length tags (\pm ratio) by loosening or inverting the SVX χ^2 requirement on one or more of the tracks in the tagged jet (recall that good quality SVX tracks have a $\chi^2 < 6$). Degrading the SVX track χ^2 reduces the heavy flavor content and increases the number of mistags in the data tagged jets. The combination of at least one track with $\chi^2 \geq 3$ and an additional track with $\chi^2 \geq 5$, puts a bound on the \pm ratio of $1.2 \pm 0.2(stat.)$. The Monte Carlo background has a \pm ratio of $1.02 \pm 0.02(stat.)$. This effect is illustrated in Figure 6.11. The $c\tau$ fit is performed using the standard b and c shapes, and a background $c\tau$ distribution forced to have an asymmetry determined from the data. Because the limiting asymmetry measured in data is $1.2 \pm 0.2(stat.)$, the Monte Carlo background asymmetry is varied from 1.0 to 1.4. As the asymmetry is increased, the resulting fit fractions for b and c decrease from their default fit values by 5.0% and 7.2% respectively. A negative systematic uncertainty for F_b is set at -7.2% and for $F_c -5.0\%$.

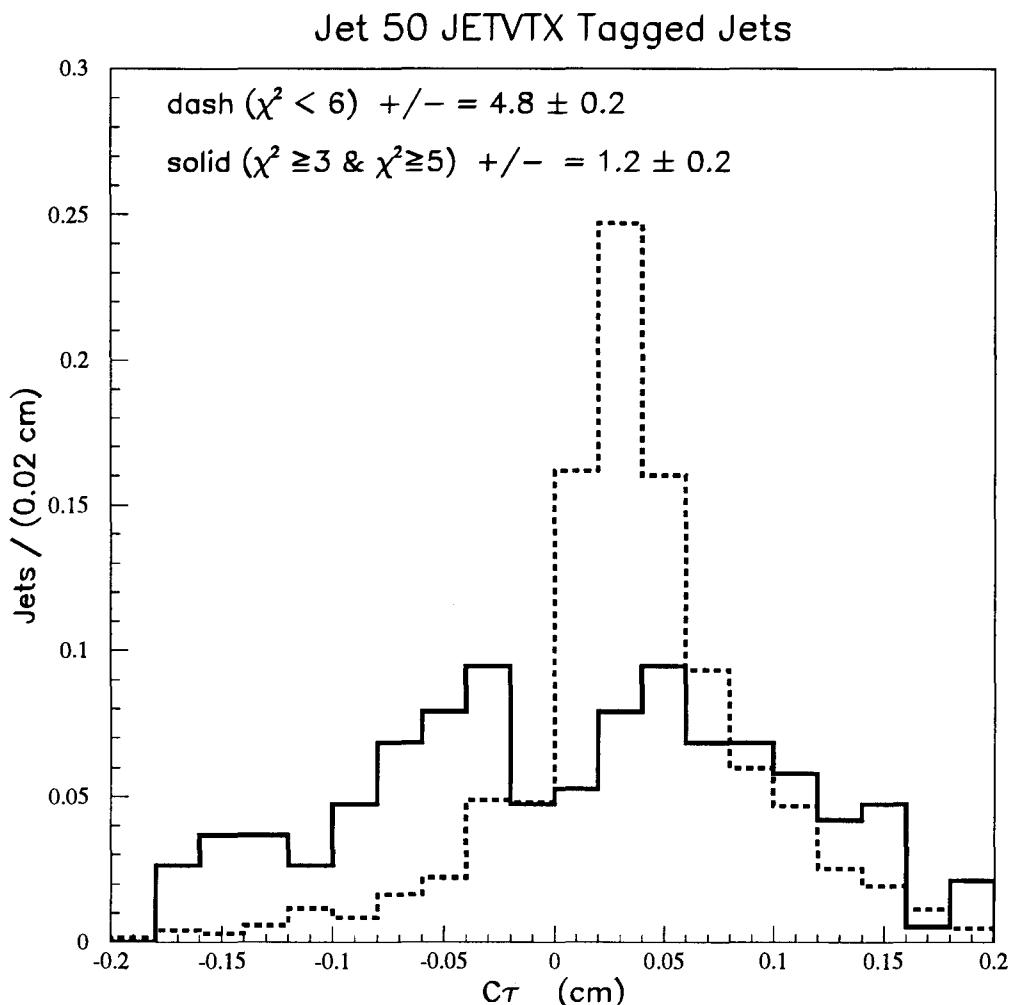


Figure 6.11: The estimated proper lifetime distributions in jet_50 data tagged jets. The dashed histogram has the default SVX track χ^2 requirements. The solid histogram has the SVX χ^2 requirements degraded as described in the text. Both histograms are normalized to unit area.

Uncertainties in the data CTC and SVX tracking efficiencies employed in the Monte Carlo simulation will mainly influence the tagging efficiencies. An increase (decrease) in tracking efficiency will in turn increase (decrease) the tagging efficiency. The SVX and CTC tracking efficiencies are described in Chapter 3.2.2. The individual tracking efficiencies are varied simultaneously so the combined tracking efficiency changes by $\pm 1\sigma$ uncertainty. The observed the change in the overall F_b measurement is $\pm 4.9\%$, the change in F_c is $^{+6.8\%}_{-2.6\%}$.

The vector-to-pseudoscalar production ratio for charmed mesons (the production ratio of spin=1 D^* to spin=0 D mesons) will influence the charm-tagging efficiency and the shape of the $c\tau$ distribution. The probability of producing a charmed vector meson is 0.75 based upon naive spin counting. A CLEO [45] measurement of this probability is $0.85 \pm 0.11(stat.) \pm 0.17(syst.)$, compared with a value of 0.70 used by the HERWIG Monte Carlo. To estimate the systematic uncertainty, the fraction of D^* produced is varied from 0.5 to 1.0. For each fraction, the charm tagging efficiency was calculated and the $c\tau$ distributions refit. As the amount of D^* is increased, the amount of D^0 increases as well, because of the dominant D^* to D^0 decays. This has the effect of decreasing the overall charm tagging efficiency (largely due to the shorter lifetime of the D^0), which in turn would increase the charm fraction per jet. This also introduces an effect on the slope of the charm $c\tau$ distribution, and thus also affects the charm fraction of the $c\tau$ fit. These two effects compensate for charm, but only one effect, the $c\tau$ shape, is relevant for bottom. As a result, the systematic uncertainty in the charm fraction per jet is slightly smaller than for bottom, F_b is $^{+5.8\%}_{-3.4\%}$, and the uncertainty for F_c is $^{+2.0\%}_{-3.1\%}$.

The rate at which the bottom meson decays into D^{*0} and $D^{*\pm}$ versus D^0 and D^\pm in turn affects the relative amount of D^0 versus D^\pm . The systematic uncertainty due to this effect may be set by observing the change in F_b when the charm meson

Production Mode Tagging Efficiencies			
Monte Carlo	DIRECT	FLAVOR EXCITATION	GLUON SPLITTING
<i>b</i> jets	0.276 ± 0.009	0.243 ± 0.007	0.216 ± 0.005
<i>c</i> jets	0.056 ± 0.004	0.051 ± 0.002	0.055 ± 0.002

Table 6.7: The tagging efficiencies for b and c tagged jets in Monte Carlo separated by production mode. Uncertainties are the statistical error.

in the sequential $b \rightarrow c$ decay is either 100% D^0 or 100% D^\pm . The change in F_b is small and the resulting systematic is $\pm^{3.4\%}_{1.5\%}$ (Note that the systematic would be much smaller if the measured branching ratio uncertainties [32] are used).

Double-tagged jet events provide information on the uncertainty in the relative fractions of production modes. As shown earlier in Section 6.3, heavy-flavor production in jets can be separated into three sub-processes: direct production, gluon splitting, and flavor excitation. Because the Monte Carlo b and c jets produced through these modes have different tagging efficiencies, the relative amount of each sub-process will affect the b and c jet fractions. Therefore, if one can estimate how much the relative amounts differ in Monte Carlo from what is found in data, a systematic uncertainty can be set. Table 6.7 lists the fractions and tagging efficiencies for b and c jets separated by sub-process.

It is assumed that the differences in the amount (see Table 6.3) of each production mode observed in the double tags between the data fit and the b Monte Carlo can be used to set the systematic uncertainty in the production mode fractions in all tagged b jets. The effect of the difference on the overall tagging efficiency, and subsequently F_b , is estimated with the following method: First, the maximum deviation between the double-tag fit of $\Delta\phi$ and the HERWIG generated fraction of a particular production mode is determined. This maximum deviation is in turn

applied to the relative fraction of that same mode in all tagged Monte Carlo b jets. The fractions of the other two modes are permitted to float while constraining the sum of all three fractions to unity. The uncertainty in F_b is determined to be $^{+3.0\%}_{-4.0\%}$. Assuming the same method can be used for c jet production, the uncertainty in F_c is determined to be $^{+3.4\%}_{-2.2\%}$. If this assumption is not made for c , a systematic can be set by noting the differences in efficiency for each c production mode. A worst case systematic uncertainty can be determined by assuming that only one of the three production sub-processes contributes to the observed tags. This approach indicates that the systematic in F_c due to the production mode uncertainty is $\pm 7\%$. However, a systematic uncertainty determined in this way does not depend on any assumptions about how production mechanisms differ between double-tagged events and all tagged jets. (Using this same approach for F_b , the systematic uncertainty would go from $^{+3.0\%}_{-4.0\%}$ to $\pm 12\%$).

Obviously, large changes in the lifetimes for bottom and charm used in the Monte Carlo will affect their respective tagging efficiencies and $c\tau$ fits, which in turn will affect the value of $F_{b(c)}$. The lifetimes of the B and C hadrons in the Monte Carlo simulations were raised and lowered by 1σ about their world average values [32] to observe the change in $F_{b(c)}$. Because the errors on the measured lifetimes are small, on the order of a few percent, F_b and F_c change very little. The resulting change in F_b is $^{+2.4\%}_{-1.5\%}$ and for F_c $^{+2.5\%}_{-3.8\%}$.

Monte Carlo statistical uncertainties come from the finite statistics used in determining the b and c tagging efficiencies. Chapter 4.2 describes the efficiencies used in determining F_b and F_c , and the resulting systematic uncertainties are $\pm 1.5\%$ for F_b and $\pm 2.5\%$ for F_c .

There is also a systematic due to the uncertainty in the b and c quark fragmentation. This uncertainty was estimated by measuring the change in the bottom and charm tagging efficiencies as the energy-momentum variable (z) distribution is var-

ied by $\pm 1\sigma$ in Monte Carlo from the measured values[46]. The energy-momentum variable z is defined as:

$$z = \frac{E^{hadron} + P_{||}^{hadron}}{E^{quark} + P_{||}^{quark}}$$

Where $P_{||}^{hadron}$ is the projection of the hadron momentum in the direction of the quark. The z distributions for bottom and charm Monte Carlo can be seen in Figure 6.12. Their mean z values are consistent with their measured values. The bottom and charm tagging efficiencies (also shown in Figure 6.12) vary slowly as a function of z in the region of their respective mean values, consequently the resulting systematic for F_b is $^{+1.1\%}_{-0.6\%}$ and for F_c is $^{+2.0\%}_{-1.3\%}$.

A summary of the systematics is listed in Table 6.4. In the jet_50 sample, the resulting total systematic uncertainty for F_b is $^{+13\%}_{-15\%}$ and for F_c is $^{+13\%}_{-26\%}$. In the jet_20 sample, the resulting total systematic uncertainty for F_b is $^{+19\%}_{-20\%}$ and for F_c is $^{+19\%}_{-29\%}$. The result systematic for σ_b is $\pm 19\%$ and for σ_c is $^{+16\%}_{-28\%}$. The increase in the size of the systematics for the cross sections (evaluated in the jet_50 sample) relative to that for the b and c fractions is due to the additional uncertainty in the Monte Carlo statistics, and the uncertainty in the integrated luminosity [22].

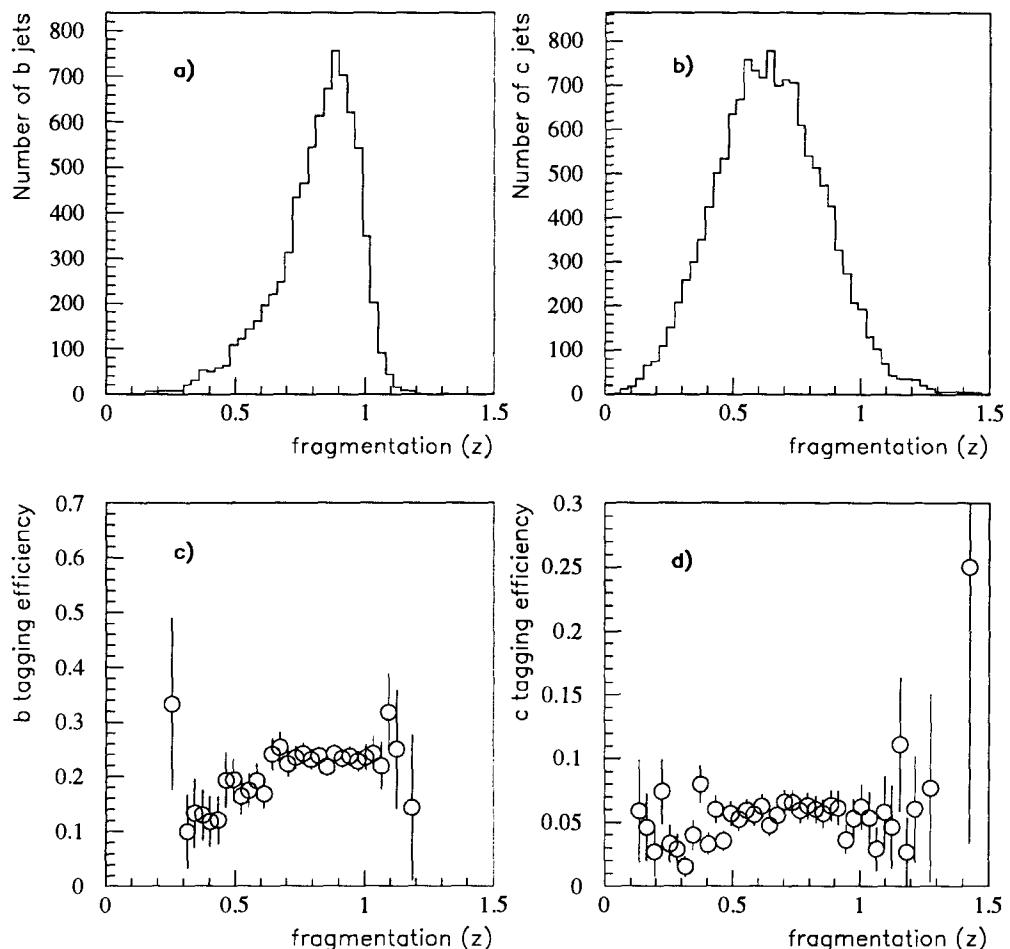


Figure 6.12: Plots (a) and (b) are the bottom and charm Monte Carlo quark fragmentation (z) distributions. Plots (c) and (d) are the bottom and charm tagging efficiencies as a function of the fragmentation variable, z .

Uncertainty	ΔF_b	ΔF_c
Tag efficiency JET_50(JET_20)	$\pm 9.0(16.0)$	$\pm 9.0(16.0)$
Total jet fit constraint	-6.4	-23.0
Background $c\tau$ shape	-7.2	-5.0
Tracking efficiency	± 4.9	$+6.8/ - 2.6$
D^*/D production	$+5.8/ - 3.4$	$+2.0/ - 3.1$
D^*/D in $b \rightarrow c$	$+3.4/ - 1.5$
Production mode	$+3.0/ - 4.0$	$+3.4/ - 2.2$
Lifetime measurements	$+2.4/ - 1.5$	$+2.5/ - 3.8$
Monte Carlo statistics	± 1.5	± 2.5
Fragmentation	$+1.1/ - 0.6$	$+2.0/ - 1.3$
Total JET_50	$+13/ - 15$	$+13/ - 26$
Total JET_20	$+19/ - 20$	$+19/ - 29$

Table 6.8: Summary of uncertainties (in %) in the F_b and F_c measurement.

Chapter 7

Conclusions

In this thesis, a secondary-vertex finding algorithm is used to identify a sample of jets with vertices displaced from the $p\bar{p}$ interaction point. By fitting the estimated proper decay-length distribution in data to a sample of jets generated using a Monte Carlo simulation, the fraction of jets that contain at least 1 b (\bar{b}) quark, and the fraction of jets that contain at least 1 c (\bar{c}) quark, F_b and F_c respectively, are measured.

The absolute rate and the E_T dependence (over a jet E_T range from 15 to 135 GeV) of F_b agrees well with NLO QCD calculations and with the HERWIG and PYTHIA Monte Carlo predictions. Moreover, the F_b is cross checked by making an independent measurement using a $b\bar{b}$ enriched data sample: double-tagged events. The resultant F_b agrees well with that measured using the estimated decay length fit.

The measured value of the F_c , with relatively large systematic errors, is approximately a factor of two higher than the predictions from the NLO QCD calculation and the HERWIG and PYTHIA Monte Carlos. F_c is checked with the data rate of $D^{*\pm}$ mesons in jet events. The number of expected $D^{*\pm}$'s are determined using

F_c and F_b from the $c\tau$ fit method. The Monte Carlo predicts that the majority of $D^{*\pm}$'s come from $c\bar{c}$ production. The expected number of $D^{*\pm}$'s is consistent with the total number of independently reconstructed $D^{*\pm}$. This would imply that the theoretical calculations and the Monte Carlo predictions underestimate the value of F_c .

With the OPAL detector at LEP, the multiplicity of heavy quark pairs from gluon splitting in Z^0 hadronic decay events produced in e^+e^- annihilation events has been measured [48]. Two methods are employed in this measurement. First, gluon splitting to heavy quarks is studied in the analysis of $D^{*\pm}$ production in Z^0 hadronic decay events at low $D^{*\pm}$ energies. In a second method, gluon splitting to heavy quarks is studied in an analysis of 3-jet events with a prompt lepton in the lowest energy jet (assumed to be the gluonic jet). The LEP measurements for $g \rightarrow c\bar{c}$ indicate that the data production rate is consistently higher than the rate predicted by HERWIG by a factor of two. This is interesting in that the c -jet fraction (where $c\bar{c}$ is believed to be mainly produced by gluon splitting) presented in this work is approximately a factor of two higher than the HERWIG prediction. However, there is reasonably good agreement between LEP results and the ARIADNE and JETSET Monte Carlo predictions.

Finally, an analysis of $b\bar{b}$ production dynamics was performed in double-tagged jet events. It was shown that leading order and next-to-leading order $b\bar{b}$ production modes may be distinguished with the distribution of azimuthal correlations between tagged jets in double-tagged events. By fitting the distribution in data to the sum of the expected distributions of the production modes from $b\bar{b}$ Monte Carlo, the relative fractions of each mode can be measured. It was shown that the HERWIG predictions of the contributions of each production mode is in good agreement with results found in double-tagged $b\bar{b}$ data events.

Bibliography

- [1] S.L. Glashow Nucl. Phys. **22**, 579 (1961);
 S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
 A Salam, in *Elementary Particle Theory: Relativistic Groups and Analyticity (Nobel Symposium No. 8)*, edited by N. Svartholm (Almqvist and Wiksell, Sweden, 1968), p.367.
- [2] S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970);
 M. Kobayashi and M. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [3] *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions* by Chris Quigg (The Benjamin/Cummings Publishing Company, Inc. 1983).
- [4] In the polar coordinate system used at CDF: z is along the proton beam axis, ϕ is the azimuthal angle, and θ is the polar angle. The transverse momentum (energy) of a particle is $P_T(E_T) = P(E) \sin \theta$, where θ is the polar angle of the energy cluster with respect to z.

- [5] U.Bauer, F.HAlzen, S.Keller, and M.Mangano, Phys. Lett. **B318**, 544 (1993).
- [6] F.Abe *et al.* FERMILAB-CONF-96/156-E.
- [7] *Photon Hadron Interactions* by R.P.Feynman (W.A. Benjamin, 1972).
- [8] *The Structure of the Proton* by R.G.Roberts (Cambridge University Press, 1990).
- [9] F.Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).
- [10] B.L. Combridge, Nucl. Phys. **B151**, 429 (1979);
J.Babcock, D.Silvers, and S. Wolfram, Phys. Rev. **D18**, 162 (1978).
- [11] V.Barger and R.J.N.Phillips, Phys. Rev. D **31**, 215 (1984).
- [12] F.Halzen and P.Hoyer, Phys. Lett. **154B**, 324 (1984).
- [13] P.Nason, S.Dawson, and R.K.Ellis, Nucl. Phys. **B303**, 607 (1988);
B327, 49 (1988); W.Beenakker, H.Kuijf, W.L.van Neerven and J.Smith, Phys. Rev. **D40** 54 (1989); W.Beenakker *et al.*, Nucl. Phys. Rev. **B351** 507 (1991).

- [14] M. Mangano, P.Nason and G.Ridolfi, Nucl. Phys. **B373**, 295 (1992).
- [15] F. Abe, *et al.*, Phys. Rev. Lett. **71**, 2396 (1993) and references therein.
- [16] S. Frixione, ETH-TH-96-23, May (1996).
- [17] S. Frixione and M.L.Mangano, CERN-TH-96-085, May (1996).
- [18] F. Abe *et al.*, Phys. Rev. Lett. **72**, 3456 (1994).
- [19] S. Malvezzi, hep-ph/9507391, June (1995).
- [20] S.D.Holmes, FERMILAB-Conf-87/160 (unpublished).
- [21] <http://www-fermi3.fnal.gov>.
- [22] P.Derwent, CDF Internal Publication 2361 (unpublished).
- [23] F.Abe *et al.*, Nucl. Inst. Meth. **A 271**, 378 (1988).
- [24] F.Abe *et al.*, Phys.Rev.D**45**, 1448 (1992).

- [25] F. Abe *et al.*, Phys. Rev. Lett. **69**, 2896 (1992).
- [26] The Particle Data Group, Phys. Rev. D **45**, Part 2 (1992).
- [27] G. Marchesini and B.R. Webber, Nucl. Phys. **B310**, 461 (1988);
G. Marchesini *et al.*, Comput. Phys. Comm. **67**, 465 (1992).
- [28] H.-U. Bengtsson and T. Sjöstrand, Comput. Phys. Commun. **46**, 43 (1987).
- [29] F. Paige and S.D. Protopopescu, BNL Report No. 38034, 1986.
- [30] T. Sjöstrand, CERN-TH-6488-92, May (1992).
- [31] A. Roodman, S. Behrands, M. Shapiro, CDF Internal Publication 1344
(unpublished).
- [32] Particle Data Group, R.M. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996).
- [33] A. Caner, A. Mukherjee and A. Yagil, CDF Internal Publication 2363
(unpublished).
- [34] The b content is much larger than the c content in the inclusive electron sample. In Reference [35] it is shown that the b to c rate in the this sample is about

4 to 1 before tagging. Since the tagging efficiency for b is about 3 times that of c in the inclusive jet sample for jets of similar E_T , the ratio of b to c after tagging is $\approx 12/1$. Therefore, the majority of tags in the inclusive electron sample come from bottom.

- [35] F. Abe *et al.*, Phys. Rev. Lett. **71**, 500 (1993).
- [36] The taggable production rate varies with jet E_T . Because the inclusive jets and the jets from the inclusive electrons have different E_T spectra, the $b\bar{b}$ Monte Carlo and the inclusive electron data sample are not used for this part of the full scale factor.
- [37] F. Abe *et al.*, Phys. Rev. D**50**, 2966 (1994).
- [38] I.C. Brock, MNFIT: A Fitting and Plotting Package Using MINUIT Version 4.02 June 2, 1994.
- [39] J. Kwiecinski, A.D. Martin, W.J. Stirling, and R.G. Roberts, Phys. Rev. D**42**, 3645 (1990); A.D. Martin, W.J. Stirling, and R.G. Roberts, Phys. Rev. D**43**, 3648 (1991).
- [40] A.D. Martin, W.J. Stirling, and R.G. Roberts, Phys. Rev. D**50**, 6734 (1994).

- [41] H.L.Lai *et al.*, hep-ph/9410404.
- [42] G. Arnison *et al.*, Phys. Lett. **147B**, 222 (1984);
M.Ikeda, Ph.D. Thesis, University of California Riverside, 1989;
F. Abe *et al.*, Phys. Rev. Lett. **64**, 4 (1990).
- [43] S. Nussinov, Phys. Rev. Lett., **35**, 1672 (1975).
- [44] ALEPH Collab., D.Buskulic *et al.*, Z. Phys. C62, 1-14 (1994).
- [45] D.Bortoletto *et al.*, Phys. Rev. D **37**, 1719 (1988).
- [46] Chrin *et al.*, Z. Phys. **C 36**, 165 (1987).
- [47] F. Abe *et al.* Phys. Rev. D**53**, 1051 (1996);
C.Albajar *et al.*, CERN-PPE/93-153 August (1993).
- [48] R.Akers *et al.*, Phys. Rev. B**353**, 595 (1995) and references therein.