Fermilab

FERMILAB-THESIS-1998-17

A STUDY OF W DECAY CHARGE ASYMMETRY USING HADRONIC TAU DECAYS IN PROTON-ANTIPROTON COLLISIONS AT $\sqrt{S} = 1.8$ TEV

BY EDWARD WILLIAM KUNS

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Professor Terence Watts

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

January, 1998

© 1998
Edward William Kuns
ALL RIGHTS RESERVED

ABSTRACT OF THE DISSERTATION

A Study of W Decay Charge Asymmetry Using Hadronic Tau Decays in Proton-Antiproton Collisions at $\sqrt{s} = 1.8$ TeV

by Edward William Kuns

Dissertation Director: Professor Terence Watts

This dissertation presents a measurement of the tau charge asymmetry in events where the taus are produced by W decays. This charge asymmetry appears as different rapidity distributions for positive and negative taus. Two competing effects generate tau charge asymmetry. The production mechanism for the W gauge boson generates a charge asymmetry which is a function of the ratio of parton distribution functions, $d(x)/u(x)$, measured at $x \approx M_W/\sqrt{s}$. This is the dominant effect for tau charge asymmetry at small rapidity. At higher rapidity, however, the competing charge asymmetry from parity violation in W decay to taus becomes dominant.

This tau asymmetry measurement is consistent with the Standard Model with a χ^2 per degree of freedom equal to 2.5 for 4 degrees of freedom when the asymmetry measurement is folded about $y = 0$, taking advantage of the CP symmetry of the underlying physics, and 8.9 for 8 degrees of freedom when it is not.

This measurement introduces some methods and variables of interest to future analyses using hadronic decay modes of taus. This work was done using the CDF detector in $\overline{\mathrm{p}}\mathrm{p}$ collisions at $\sqrt{s} = 1.8$ TeV at Fermilab's Tevatron accelerator.

Acknowledgements

No physics analysis occurs in a vacuum; this is especially true in experimental high energy physics where the scale of the apparatus means one must rely on the technical expertise of many others. I gratefully acknowledge the support and work of my fellow CDF collaborators (listed in Appendix D) and the Fermilab staff. Without their hard work, the Tevatron and the CDF detector would not exist, let alone be the successful and highly tuned instruments they are.

I received insightful and helpful physics guidance from many people. I would like to thank, in particular, my advisor Terry Watts, who remained a confident source of enthusiasm. John Conway provided invaluable help with physics analysis, and I am also deeply grateful for the help from Tom Devlin and Cal Loomis. I owe a debt of gratitude to the CDF Electroweak physics group conveners, especially Young Kee Kim, Larry Noduleman, Rob Wagner, and Debby Errede. Thanks also to Henry Frisch for suggesting this topic for study.

I could not have completed my study without the essential administrative help of Ruth Bennett and Nancy DeHaan at Rutgers and Carol Piccolo at CDF. Thanks are also due to the people who kept the computers running and to Jim Patrick, without whom our experiment might have collected many fewer events. I will always remember the camaraderie of my fellow Aces and of "Ace mom and dad": Dee and Steve Hahn.

I owe my continued sanity to many circles of friends. Early in my tenure at Fermilab, I spent many a noon or weekend afternoon with the "lunch group": Alma Karas, Moe Foucher, Tim Dubbs, Mary Roach, Lee Robbins, and many others. Carol Hawk and Mike Traynor have been and remain true friends, and no, Carol, I did not make you eat those "two big cookies."

Perhaps I could have made it without meeting Seska, Heinrich, Helmut, James Logan, Klaus (the one and only, thank God), Kylenne, Stephan, Luiz, Gargon, Engendro, Flint, and others ... perhaps. I am glad I do not have to find out. Thank you to the universe creators, Todd Huffman, Steve Pappas, Randal Hans, Ken Kelley, and to my fellow adventurers, Noreen Huffman, Joe Steele, Jim and Maureen Hoff, Nancy Chase, Troy Daniels, and Zigs (although Puller never really got to meet Serendipity). I will miss those nights.

Although the "cholesterol of CDF" (you all know who you are) pulled me away from physics analysis to the necessary task of blocking hallways, the welcome distraction helped me to be more focussed when sitting in front of WHCDFX, my faithful VAXStation. My days would have been considerably more boring without the wit and company of Tom LeCompte, Carol Anway-Weise, Leslie Groer, and Leigh Ann Markosky. I also enjoyed many late evenings with the CDF potluck crowd including Leila Belkora, Petar Maksimovic, Wasiq Bokhari, and Kara Hoffman.

I would be remiss if I did not mention the Drug Sniffing Dogs, who twice a year gave me an opportunity to remember why my "day job" is as a physicist and not as a singer. If I hadn't thrown caution to the wind, I would have missed out on jam sessions, performances, and the chance to say "I'm a singer in a garage band." I have also greatly enjoyed practicing with the Fermilab Choir.

My many friends on FemRel—too numerous to even attempt to name—provided a much needed opportunity to talk intelligently (and sometimes less than intelligently) about something other than physics. Thank you. I even got to meet and get to know many of you: Sherri, Lili, Sue, Jon, Jay, Laura, and others. Robin, it means a lot that you believe in me.

I was happy to discover that the city of Chicago has culture! (Yes, indeed.) My summers would have been emptier without the annual Blues Fest, Taste of Chicago, fireworks, free concerts, and the excuse to spend a weekend with friends. Neil and Jeff introduced me to the Blues Fest and Ed Lopez introduced me to a more expensive event: WorldCon. As long as I am discussing distractions, "Babylon 5" is the best television, not to mention the best science fiction series, I have ever seen. Each Thursday night

for the past four years, I have had a reason to rush home, slow down, and actually cook dinner before returning to my office.

What would I have done without Nathan, Nancy, and Marion, who urged and cajoled me to "write, write, write"? You might think they have been there before or something. I appreciate their relentless encouragement; sometimes I needed a kick in the rear. I would never have made it this far without the love and support of family, my cats, the "Tuesday night" crowd, Bob, and the many friends and colleagues listed and not listed here. Thank you all.

This work was supported in part by United States National Science Foundation Grant PHY-94-17820.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BBC	Beam-Beam Counter
BFM	BuFfer Manager
BMX	Buffer MultipleXor
CDF	Collider Detector at Fermilab
CEM	Central ElectroMagnetic calorimeter
CES	Central Electron Strip chamber
CHA	Central HAdronic calorimeter
CFT	Central Fast Tracker
CTC	Central Tracking Chamber
DAQ	Data AcQuisition system
DST	Data Summary Tape
EWK	ElectroWeaK
EVB	EVent Builder
FEM	Forward ElectroMagnetic calorimeter
FHA	Forward HAdron calorimeter
FRED	Front-end REadout and Decision
ISO	ISOlated
NISO	Not ISOlated
PAD	Physics Analysis Dataset
PEM	Plug ElectroMagnetic calorimeter
PHA	Plug HAdron calorimeter
SVX	Silicon VerteX detector
TAUO	TAU Object
TS	Trigger Supervisor
VTX	VerTeX time projection chamber
WHA	Wall HAdronic calorimeter

Chapter 1

Introduction

> The Tao that can be told is not the eternal Tao.
> The name that can be named is not the eternal name.
> The nameless is the beginning of heaven and earth.
> The named is the mother of the universe.
> — Lao Tsu [1]

Throughout recorded history, people have used a variety of explicit and implicit models to describe the natural world, some attributing events to the actions of a God or gods, some attributing events to a "natural law" of cause and effect. In the centuries since philosophers first thought some component of matter might be "atomic," many wondered: Is there any component of the world which is indivisible and (thus) fundamental? Of what is the natural world composed?

An early division of nature into components divided all we see into four elements: earth, air, fire, and water. Later, atoms, and thus the chemical elements, were thought to be fundamental. Discovering the periodic table brought order to the increasing number of elements; this order led to the discovery of protons and electrons. During the late 20th century, the discovery of an increasing number of "elementary" particles confounded the belief that these leptons, mesons and baryons are all fundamental particles. Physicists are drawn to models which are simple and elegant, and a large "zoo" of fundamental particles is neither.

In addition to seeking elegance, physicists seek models which are precise, accurate and testable, that is, models which provide testable predictions confirmed by measurement. The division of nature into four elements, for example, a model which is both simple and elegant, fails this test. What to do? In answer to this quandary, the "Standard Model" evolved into its present form, seeking to explain and predict, yet with as much simplicity and elegance as can be identified. The Standard Model identifies

quarks and leptons as elementary particles and specifies the interactions between these particles.

In addition to the belief that the number of elementary particles is small, the desire for an elegant theory leads physicists to believe that these elementary particles (and thus, all matter and energy) interact via only a few fundamental forces. These fundamental forces may exhibit themselves in many ways, just as electromagnetism exhibits electrostatic forces, paramagnetism, visible light, the quantum Hall effect, the van der Waals force, and many other seemingly unrelated forces and effects. The Standard Model identifies four fundamental forces; no compelling evidence yet exists to support the existence of a fifth force.

This dissertation investigates one aspect of the Standard Model, the charge asymmetry of $W^{\pm}$ decays, with data collected at the Collider Detector at Fermilab (CDF) using the Fermilab Tevatron, the highest energy proton-antiproton collider in the world. Although this dissertation focuses on a narrow segment of the Standard Model, such focussed measurements test the Standard Model and provide constraints on the possibilities of particle physics beyond the Standard Model.

I shall briefly introduce the Standard Model in the next section, followed by a few comments describing $\overline{\mathrm{p}}\mathrm{p}$ collisions at the Tevatron. Finally, I shall discuss charge asymmetry and how I measure it. The details of the Standard Model which are important for this dissertation shall receive more attention in the next chapter.

1.1 The Standard Model

The Standard Model is the culmination of many decades of experiments and theory, and is an attempt to bring everything known about particle physics into one theory. (See, for example, Ref [2] and the many references therein.) In the Standard Model, two forces, electromagnetism and the weak nuclear force, are unified in the electroweak theory. The strong nuclear force is described by quantum chromodynamics (QCD). Only one of the four known forces is unaccounted for—gravity.

Generation	Quarks	charge	mass	Leptons	charge	mass
1st	up	$+2/3$	$\sim$ 5 MeV	ν_e	0	0
	down	$-1/3$	$\sim$ 10 MeV	e	-1	0.511 MeV
2nd	charm	$+2/3$	$\sim$ 1.3 GeV	ν_μ	0	0
	strange	$-1/3$	$\sim$ 0.2 GeV	μ	-1	106 MeV
3rd	top	$+2/3$	$\sim$ 175 GeV	ν_τ	0	0
	bottom	$-1/3$	$\sim$ 4.3 GeV	τ	-1	1.78 GeV

Table 1.1: The fundamental particles; units of electric charge are $|e|$. ν_τ is the only particle which has not been directly observed. Each particle has an antiparticle with the same mass and opposite quantum numbers. All neutrinos have zero mass in the Standard Model. Quark masses are not precisely measured. Note that the total electric charge of each generation, with each quark counted three times to account for color degrees of freedom, is zero. Quarks are listed here in their QCD eigenstates; see Section 2.3.

In the Standard Model, all matter is composed of fermions, particles with 1/2-integral spin. Forces are mediated by "gauge bosons," particles of integral spin which are the physical representation of a "gauge field." "Gauge" is a historical misnomer for "phase," referring to the mathematical phase symmetry (resulting in a conserved charge) present in gauge field theories, which I discuss further in Section 2.1.

Fermions interact only through the exchange of a gauge boson (the Standard Model includes no "action at a distance") and fermions do not interact directly with one another. Each force is mediated by specific gauge bosons. Gluons mediate the strong force, $W^\pm$ and Z^0 mediate the weak force, the photon (γ) mediates the electromagnetic force, and the postulated graviton mediates gravity although no successful quantum theory of gravity exists.

All particles respond to the electroweak force. The fundamental fermions are divided into quarks and leptons; leptons do not respond to the strong force; quarks do. Both quarks and leptons are assumed to be point particles with no structure.

The fermions are organized into three generations, each a doublet, as shown in Table 1.1. Since quarks respond to the strong force, they carry "color charge" with one of three possible values. Thus, eighteen fundamental quarks exist: six "flavors" each in three colors. Leptons do not carry color charge, so there are only six fundamental leptons. Note the careful way in which electric charge sums to zero. The total electric

charge of each generation is equal to zero:

$$3 \times \left(\frac{2}{3}\right) + 3 \times \left(-\frac{1}{3}\right) + 0 - 1 = 0,$$

where the factor of three comes from the three colors of each quark flavor. Since the electric charge enters linearly into electromagnetic interaction amplitudes, it enters quadratically into a measurement of an observable quantity. In each generation, the difference between the charge-squared of the more positive fermion and the more negative fermion *also* sums to zero:

$$3 \times \left(\frac{2}{3}\right)^2 - 3 \times \left(-\frac{1}{3}\right)^2 + 0^2 - (-1)^2 = 0.$$

Due to "quark confinement," free quarks do not exist, so quark masses must be measured indirectly and are not precisely measured. Lepton masses, however, can be measured directly and are well known. The electron and muon mass are measured to a precision better than 1 part in 10^6; the tau mass is measured to a precision of nearly 1 part in 10,000. In the Standard Model, neutrinos have zero mass; experiments have not yet ruled out small masses for the electron or muon neutrinos. The tau neutrino is the only fundamental fermion in the Standard Model which has not been observed.

Quark confinement means that quarks are bound into colorless objects, hadrons, via the strong force. Thus, hadrons are composite particles which interact strongly. Hadrons and leptons are observable particles, although only one hadron is stable (in the Standard Model) when free, the proton (uud). All other hadrons, when free, decay through a cascade of particles into (eventually) electrons, photons, neutrinos, and occasionally protons. The time scale for hadron decay ranges from 10^{-24} sec for resonances (e.g., ρ, η, ϕ) to 10^{-8} sec for the "almost stable" pions and kaons to about 10^3 sec for the neutron.

The quarks (and gluons exchanged between the quarks) constituting a hadron are called "partons." "Valence" quarks are those which contribute to the quantum numbers of the hadron, but "sea" partons—gluons and quark-antiquark pairs which the valence quarks radiate and absorb, temporarily "borrowing" energy from the vacuum—are important constituents of a hadron.

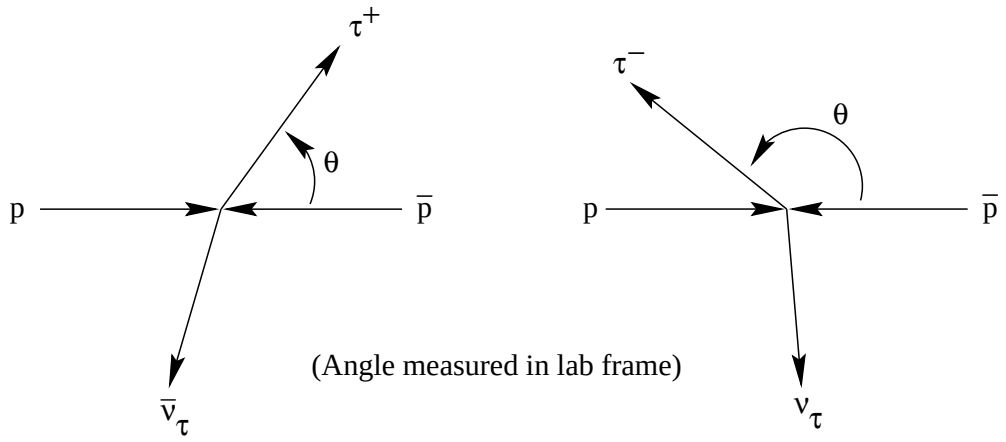


Figure 1.1: Tau charge asymmetry in $\overline{\mathrm{p}}\mathrm{p}$ collisions. This figure is illustrative only, showing the features of $\overline{\mathrm{p}}\mathrm{p} \rightarrow W \rightarrow \tau\nu$ as seen in the lab. Not all final collision products are shown; the W mediating this interaction is not seen in the lab. If positive and negative taus have identical distributions versus polar angle θ, then no charge asymmetry exists.

The proton contains three valence quarks (u, u, d), each a different color, thus keeping the proton colorless. In a high momentum proton, the valence quarks carry less than half of the proton's momentum, thus illustrating the significance of the parton sea.

In the Fermilab Tevatron, protons and antiprotons, each with an energy of 900 GeV, collide with a center-of-mass energy of 1800 GeV. The actual energy (or momentum transfer) involved in the interaction is usually small, but some interactions involve a large momentum transfer between the proton and antiproton—usually between a single parton in each particle. When the momentum transfer is large enough, a u and $\overline{d}$ quark from the proton and antiproton can annihilate, forming a W^+ particle. Such interactions and the decay of the W form the basis for this analysis.

1.2 Measuring Charge Asymmetry

The fundamental concept behind charge asymmetry is illustrated in Figure 1.1. In a proton-antiproton collision, some interaction takes place resulting in the creation of a tau and a tau neutrino. The neutrino is unobserved by any particle detector, but the tau can be detected and measured through its decay daughters, usually one or three charged tracks plus neutral electromagnetic energy. Defining polar angle θ from

the direction the proton is travelling as shown in Figure 1.1, one can measure the angular distributions of positive and negative taus. If positive and negative taus have identical distributions in θ, then no charge asymmetry exists for taus. In fact, positive and negative taus have different angular distributions. Although the total number of positive and negative taus should be the same, the number in a given range of polar angle differs. Assuming CP symmetry (see Section 2.3), the positive and negative tau angular distributions can be shown to be mirror images of each other, flipped about $\theta = \pi/2$. The asymmetry in positive and negative tau distributions probes the details of the $\overline{\mathrm{p}}\mathrm{p} \rightarrow W + X$ collision and subsequent W decay into a tau + neutrino.

More than one physics process contributes to tau charge asymmetry. Production of the W^+ and W^- is asymmetric ("W production charge asymmetry") with polar angle θ. In addition, W's do not decay to leptons isotropically in the W rest frame ("W decay charge asymmetry"). The tau distribution versus polar angle θ observed in the lab ("tau charge asymmetry") is the net result of these two effects, all of which are described in more detail in the next chapter.

A tau's pseudorapidity may be calculated (at least approximately, since neutrinos from its decay are unobserved) from the polar angle of its decay products:

$$\eta = -\ln \tan \frac{\theta}{2}.$$

Rapidity y is approximately equal to the "pseudorapidity" η as defined here. Pseudorapidity and rapidity are equal for a massless particle, and nearly equal for a highly relativistic massive particle.

I wish to measure the tau charge asymmetry as a function of the η of its decay products. To measure the tau charge asymmetry, I first measure the number of positive and negative taus in bins of η. Since some other physics processes ("backgrounds") leave an event signature similar to that left by $W \rightarrow \tau\nu$, event selection includes "cuts" to reject backgrounds, entailing a balance between efficiency in selecting taus and a high rejection of backgrounds.

When the efficiency for measuring positive and negative taus is the same, the charge

asymmetry in a bin of η is parameterized as

$$A(\eta) = \frac{N_\tau^+(\eta) - N_\tau^-(\eta)}{N_\tau^+(\eta) + N_\tau^-(\eta)},$$

where $N_\tau^+(\eta)$ and $N_\tau^-(\eta)$ are the number of positive and negative taus in that bin. Event selection can not remove all background events without unacceptably low efficiency for taus, so some background processes are necessarily present in the final event sample. If the background processes possess no charge asymmetry, their presence dilutes the measured asymmetry. Thus, even if backgrounds do not possess charge asymmetry, the final asymmetry measurement must account for backgrounds properly and carefully.

Since the backgrounds can be modeled, a binned log likelihood fit (which by design accounts for the Poisson statistics present when the number of events per bin is small) is an appropriate way to account for the backgrounds left in the event sample after all selection cuts. The log likelihood fit requires, as input, a set of $W \to \tau\nu$ candidate events selected from the data, a model of the $W \to \tau\nu$ signal, and models of the background processes. In this manner, background processes are properly accounted for and the number of positive and negative taus in each bin of η is carefully measured.

The end result of event selection and the log likelihood fit is a measurement of the tau charge asymmetry as a function of η.

Chapter 2

Theory

> In theory there is nothing to hinder our following what we are taught; but in life there are many things to draw us aside.
> — Epictetus, *Discourses* [3]

The three strongest forces in the Standard Model are well described by quantum electrodynamics (QED) for electromagnetism, quantum chromodynamics (QCD) for the strong force, and the unified electroweak theory for both electromagnetism and the weak force. Although electromagnetism is described well by the unified electroweak theory, QED still enjoys a robust life in particle theory; this is comparable to using non-relativistic mechanics to describe the motion of non-relativistic objects. The Standard Model has a place-holder gauge boson for gravity, the graviton, but no satisfactory quantum theory of gravity exists.

The strong, electromagnetic, and weak forces involve a conserved "charge," electric charge for the electromagnetic force, color charge for the strong force, weak isospin and weak hypercharge for the electroweak force. The net value of these charges is strictly conserved globally (the total electric charge of the universe is constant, for example) and locally in any particle interaction. Such strict charge conservation has important consequences, as described in the next section.

2.1 Gauge Fields and Gauge Bosons

The gauge bosons are listed in Table 2.1 along with the force mediated by each boson. In a gauge field theory, arbitrary gauge transformations (global or local) acting on matter and gauge fields must leave the Lagrangian, and thus all observable quantities, invariant. Since gauge transformations are continuous transformations (i.e., allowing infinitesimal transformations), gauge field theories are categorized into Lie

Gauge Boson	Force	Range (m)	Theory
gluon (g)	strong	10^{-15}	QCD; SU(3)
$W^{\pm}$, Z^0	weak	10^{-18}	Electroweak; SU(2) $\times$ U(1)
photon (γ)	electromagnetic	∞	Electroweak; SU(2) $\times$ U(1)
graviton (G)	gravitation	∞	General Relativity

Table 2.1: The gauge bosons and the force they mediate, listed in order of the strength of the interaction. The graviton, the predicted carrier of gravitation, is massless, as are gluons and the photon. The $W^{\pm}$ and Z^0, on the other hand, are massive. The Higgs boson, not listed in this table, provides the mechanism by which particles, notably the $W^{\pm}$ and Z^0, gain mass. Many models, including the Standard Model, predict the existence of one or more Higgs bosons, but none have yet been observed.

groups, mathematical groups which describe quantities that exhibit a symmetry (or measurable quantity which remains invariant) under some continuous transformation. Each such symmetry directly relates to a quantity which is conserved.

For example, classically, one requires equations of motion to be invariant under translations in space and time as well as spatial rotations. That is, the laws of physics should not depend upon any privileged position or orientation in space. As a direct result of these symmetries, momentum, energy, and angular momentum, respectively, are conserved quantities.

2.2 QED and QCD

Electromagnetism, described by QED, is a U(1) theory, that is, described by the U(1) Lie group; QED is unitary and U(1) has one gauge phase. Local gauge invariance in QED requires the introduction of a photon "field," the gauge field. The gauge field adjusts for the ability to impose arbitrary gauge transformations without changing any observables. The photon is massless. QED's single conserved quantity is electric charge. Further, QED is Abelian, meaning two consecutive gauge transformations commute, just as two consecutive arbitrary translations in space commute. The coupling constant for QED is α, which is a "running" coupling constant; its value depends on the momentum transfer, or Q^2, of an interaction. At $Q^2 = 0$, $\alpha = 1/137$; at $Q^2 = M_W^2$, $\alpha = 1/128$.

On the other hand, the strong force, described by QCD, is described by the SU(3)

Lie group, composed of eight 3×3 unitary traceless matrices. QCD thus requires eight gauge fields (corresponding to gluons) to maintain local gauge invariance. As with QED, gluons are massless. QCD is a non-Abelian theory; two consecutive gauge transformations do not necessarily commute, just as two consecutive rotations in 3-dimensional space do not necessarily commute. Due to this effect, gluons carry color charge; carrying the charge of the gauge field is a general feature of the bosons of non-Abelian theories. Although gluons are massless, QCD does not actually act at infinite range due to "quark confinement," described next. The coupling constant for QCD is α_s, also a running coupling constant. For $Q^2 < (300 \text{ MeV})^2$, α_s approaches unity, while $\alpha_s = 0.12$ at $Q^2 = M_Z^2$.

It is because gluons carry color charge (in contrast to photons, which carry no electric charge) that the strong force behaves so differently from the otherwise similar electromagnetic force. Gluons interact with each other directly and as strongly as with quarks. Due to this gluon-gluon interaction, the strong force grows linearly with distance: As two quarks move apart from each other, gluons exchanged between the two quarks interact with each other as well as with the quarks. The increasing force either binds the quarks together, or the "color string" breaks when the energy of the color field between the quarks is great enough to create a quark-antiquark pair. This results in "quark confinement": Quarks can appear only in colorless objects, SU(3) singlets, not as free particles. Quarks can be bound into a colorless object in one of two ways: as a meson ($q\bar{q}$) or a baryon (qqq), collectively referred to as hadrons. Thus, "hadronization" is the process by which a quark, knocked out of a hadron in a high Q^2 interaction, and the rest of the hadron reform into colorless objects. The confinement radius is approximately 1 fm.

Just as the strong force grows linearly with distance, it becomes weaker as two colored objects approach each other, although at very small distances $r \ll 1$ fm the strong force is believed to behave like a Coulomb field. (See, for example, Ref [4] pp 13–14.) Therefore, at quark separations much smaller than the quark confinement radius (or equivalently, for interactions with high momentum transfer $Q^2 \gg (300 \text{ MeV})^2$), the quarks in a hadron behave as free particles. That is, a single parton may participate

in an interaction; the remaining partons are "spectators" and only interact with the struck parton during hadronization.

This property of QCD, "asymptotic freedom," is a second consequence of the strong force being non-Abelian, that gluons carry color charge. Another way of putting this is that high Q^2 interactions can be calculated perturbatively in QCD and hadronization occurs well after the high Q^2 interaction and at a much lower energy scale, i.e., $\approx$ 300 MeV. Hadronization is non-perturbative and not directly calculable, but due to the lower energy scale of hadronization (as compared with the high Q^2 interaction) each outgoing parton from the high Q^2 interaction hadronizes into a "jet" of hadrons travelling roughly collinearly with each other and with the original direction of the outgoing parton. Thus, information about the outgoing partons from the high Q^2 interaction is largely preserved in the resulting jets.

2.3 The Electroweak Force

QED and QCD each comprise a single Lie group, but the electroweak force comprises two disjoint sets, SU(2)×U(1). That is, the electroweak force can be divided into two potentially independent forces with different coupling constants; the two forces in the electroweak theory are imperfectly unified. The SU(2) Lie algebra describes the force involving weak isospin, which is non-Abelian, while the Abelian weak hypercharge is described by the U(1) Lie algebra. The electroweak force is usually described as $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$; the subscripts refer to weak isospin, which interacts only with left-handed fermions (and right handed antifermions) for $\mathrm{SU}(2)_L$ and weak hypercharge for $\mathrm{U}(1)_Y$. In contrast, QED contains exactly one type of charge, electric charge, and QCD contains exactly one type of charge: color. The electroweak coupling constants are conventionally taken to be g for weak isospin and $g'/2$ for weak hypercharge.

Weak hypercharge Y is a linear combination of electric charge Q and the third component of weak isospin, T_3; Y is defined by the following equation:

$$Q = T_3 + \frac{Y}{2}.$$

The values of Y, T, and T_3 for all fermions are given in Table 2.2. Since U(1) is Abelian,

Quarks	T	T_3	Q	Y	Leptons	T	T_3	Q	Y
u_L, c_L, t_L	1/2	+1/2	+2/3	+1/3	ν_e, ν_μ, ν_τ	1/2	+1/2	0	−1
d_L, s_L, b_L	1/2	−1/2	−1/3	+1/3	e_L^-, μ_L^-, τ_L^-	1/2	−1/2	−1	−1
u_R, c_R, t_R	0	0	+2/3	+4/3					
d_R, s_R, b_R	0	0	−1/3	−2/3	e_R^-, μ_R^-, τ_R^-	0	0	−1	−2
$\overline{u}_R, \overline{c}_R, \overline{t}_R$	1/2	−1/2	−2/3	−1/3	$\overline{\nu}_e$, $\overline{\nu}_\mu$, $\overline{\nu}_\tau$	1/2	−1/2	0	+1
$\overline{d}_R, \overline{s}_R, \overline{b}_R$	1/2	+1/2	+1/3	−1/3	e_R^+, μ_R^+, τ_R^+	1/2	+1/2	+1	+1
$\overline{u}_L, \overline{c}_L, \overline{t}_L$	0	0	−2/3	−4/3					
$\overline{d}_L, \overline{s}_L, \overline{b}_L$	0	0	+1/3	+2/3	e_L^+, μ_L^+, τ_L^+	0	0	+1	+2
Charged EWK Gauge Bosons					Neutral EWK Gauge Bosons				
W^+	1	+1	+1	0	Z^0	1	0	0	0
W^-	1	−1	−1	0	γ	0	0	0	0

Table 2.2: Weak hypercharge for all quarks, leptons, and the gauge bosons for the electroweak (EWK) force. Notice that the Z^0 and photon have the same quantum numbers (T_3 and Y); thus, they can mix. Gluons have $T = Q = Y = 0$, but also carry color charge, thus differing from the electroweak gauge bosons which do not. Massless neutrinos appear only as left-handed particles and right-handed antiparticles.

none of the gauge bosons carry weak hypercharge. The gauge bosons *do* carry weak isospin (and thus, electric charge), however, since SU(2) is non-Abelian. SU(2) requires weak isospin = 1, thus, with three z projections: +1, 0, −1; the three SU(2) gauge bosons carry charge $Q = T_3$. Weak isospin, therefore, involves a charged current and a neutral current. The charged current is mediated by the $W^\pm$, the neutral current by the Z^0.

Global gauge symmetry for $\mathrm{SU}(2)_L$ and $\mathrm{SU}(1)_Y$ requires weak isospin and weak hypercharge, respectively, to be globally conserved. As with QED, local gauge symmetry requires the introduction of gauge fields. For SU(2) gauge symmetry, three gauge fields, $W_\mu^{(i)}$, $i = 1, 2, 3$, adjust for the ability to arbitrarily change the local gauge without changing observables. U(1) gauge symmetry, as with QED, requires a vector field B_μ. Since each electroweak charge is conserved locally and globally, the electroweak force necessarily conserves electric charge.

As mentioned in Table 2.2, the Z^0 and γ gauge bosons have the same quantum numbers, and thus may mix. The Weinberg angle, θ_W, parameterizes the amount of

this mixing. Thus, the electroweak gauge fields may be written as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^{(1)} \mp iW_\mu^{(2)}\right)$$

$$Z_\mu = W_\mu^{(3)} \cos\theta_W - B_\mu \sin\theta_W$$

$$A_\mu = W_\mu^{(3)} \sin\theta_W + B_\mu \cos\theta_W$$

with $W_\mu^\pm$, Z_μ, and A_μ identified with the $W^\pm$, Z^0, and γ, respectively. The $W^\pm$ gauge bosons mediate the weak "charged current" and the Z^0 gauge boson mediates the weak "neutral current." The most recent world-averaged measurement of the Weinberg angle is $\sin^2\theta_W = 0.2315 \pm 0.0004$ [5]. Note that if $\sin^2\theta_W = 0$, then the neutral weak and electromagnetic forces are unmixed.

In the Standard Model, the coupling constants are related to $|e|$, the electron electric charge, by:

$$g \sin\theta_W = g' \cos\theta_W = |e|.$$

The relative strength ρ of the charged and neutral current is defined as

$$\rho \equiv \left(\frac{g_Z^2}{M_Z^2}\right) / \left(\frac{g^2}{M_W^2}\right),$$

where the Z^0 coupling g_Z^2 can be shown to be equal to $g/\cos\theta_W$. (Remember that the physical Z^0 boson is a mixture of the weak isospin and weak hypercharge gauge fields.) Thus,

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2\theta_W}.$$

In the Standard Model, $\rho \equiv 1.0$, leading to $M_W/M_Z = \cos\theta_W$.

Since weak isospin couples only to left-handed particles, this force violates parity (P) and charge conjugation (C) symmetries maximally. Right-handed fermions (and left-handed antifermions) have zero weak isospin. Weak isospin *almost* conserves the combined CP symmetry. Parity violation is critically important in determining the W charge asymmetry (see Section 2.5), but in the Standard Model no CP violation occurs in $\overline{\mathrm{p}}\mathrm{p} \to W \to \tau\nu$ interactions.

Weak isospin divides the leptons into left-handed doublets and right-handed singlets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, e_R^-,\ \mu_R^-,\ \tau_R^-.$$

Unlike leptons, quarks *do* mix (change flavor) in weak interactions. This mixing is parameterized by a 3×3 unitary matrix which transforms between the strong flavor eigenstates and the weak isospin eigenstates. Such a matrix has three independent real parameters (describing the mixing) and one complex phase which, if non-zero, generates CP-violation. The up-type and down-type quarks may separately mix; however, an arbitrary choice of phase allows the up-type quarks to have the same strong and electroweak eigenstates. Thus, mixing is parameterized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}.$$

The diagonal V_{ij} entries are close to one; the off-diagonal entries are close to zero. The electroweak down-type quark eigenstates, which are mixtures of the strong force (and mass) eigenstates, are indicated by d', s', and b'. To a good enough approximation, the elements of the CKM matrix may be identified (using numbers from Ref [5]) as follows: $V_{ud} \approx V_{cs} \approx 0.97$, $V_{tb} \approx 1.0$, $V_{cd} \approx V_{us}$, $V_{ts} \approx V_{cb}$, and $V_{td} \approx V_{ub}$. Mixing between the generations, thus, can be measured by $V_{ud}^2 \approx 20 \times V_{us}^2$, $V_{cs}^2 \approx 600 \times V_{cb}^2$, and $V_{tb}^2 \approx 25000 \times V_{ub}^2$.

2.4 $\overline{\mathrm{p}}\mathrm{p}$ Collisions

In the Fermilab Tevatron, protons and antiprotons collide at very high energy. While some of these collisions are diffractive or elastic, many are inelastic with a high momentum transfer (or Q^2). In a collision, it is the internal constituents of the proton and antiproton that interact. Thus, a high Q^2 collision can probe deeply the structure of the proton and antiproton. As described in the previous section, the forces binding quarks into a proton are non-perturbative, so the structure of the proton cannot yet be calculated from first principles.

One way to describe the internal structure of the proton is through parton distribution functions (pdfs), $f_i(x, Q^2)$, distributions for each parton i which describe the

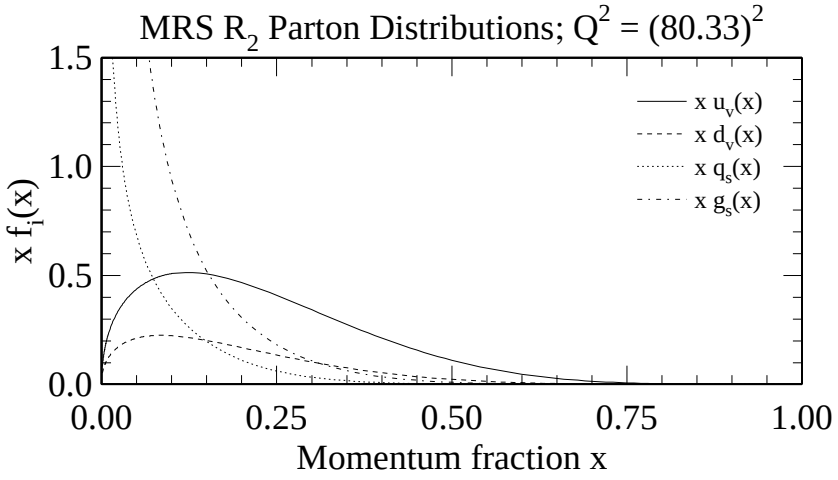


Figure 2.1: MRS-R_2 parton distributions at $Q^2 = M_W^2$. This figure presents one parameterization of the parton distribution functions, MRS-R_2, a recent fit by the MRS group (Martin, Roberts, and Stirling). x is the fraction of the proton momentum carried by the parton. As is the convention, the functions $x \times f_i(x)$ are plotted. The sea quark distributions are summed together in this plot as q_s—note the parton sea diverges for low x and the gluon sea dominates for $x \lesssim 0.1$.

probability of finding that parton with a given momentum fraction x of a moving proton. (Pdfs do not describe a static proton.) The value of the integral $\int_0^1 f_i(x, Q^2)\, dx$ is therefore the total number of parton i found in the proton, and the value of $\int_0^1 x\ f_i(x, Q^2)\, dx$ gives the total momentum fraction of the proton carried by this parton. Therefore,

$$\sum_{i=u,d,\ldots} \int_0^1 x\ f_i(x, Q^2)\, dx = 1.0.$$

At large x, i.e., $x \gtrsim 0.2$, the valence quarks u and d dominate, but as the momentum fraction drops, the parton sea takes over. At low x, gluons dominate. Gluons carry about $1/2$ of the proton's momentum; the valence quarks carry most of the rest. The parton distribution functions are independent of the proton's momentum; they depend only on x and Q^2. A pdf for a given parton is often written as, for example, $u \equiv u(x, Q^2) \equiv f_u(x, Q^2)$. Given an interaction at a particular Q^2, the momentum scale, Q^2, is often implied rather than specified, thus listing pdfs as $u(x), d(x), \overline{c}(x), g(x)$, etc. I will follow this convention.

Several sum rules constrain the values of distribution functions. For example, the

number of c and $\overline{c}$ quarks must be identical—as sea partons, these quarks can appear in a proton only when created in pairs from the vacuum. Also the sum $u(x) - \overline{u}(x)$ over all x must equal 2, the number of valence u quarks in a proton. ($u(x)$ contains both valence and sea contributions, but $\overline{u}(x)$ can appear in a proton only as a sea parton.) Thus,

$$u(x) - \overline{u}(x) = [u_v(x) + u_s(x)] - \overline{u}_s(x) = u_v(x) + [u_s(x) - \overline{u}_s(x)] = u_v(x),$$

where the subscripts v and s indicate valence and sea quark distributions, respectively.

The pdf $\overline{u}(x)$ for an antiproton is the same as the pdf $u(x)$ for a proton. Parton distribution functions are measured (parameterized) in fits to deep inelastic scattering (DIS) data and other experimental data which constrain the parton distributions. Two families of pdfs, MRS and CTEQ, are named after the groups who fit a parameterization to the data. Recent parameterizations by these groups include MRS-R_2 [6] and CTEQ-4M [7].

Figure 2.1 presents a current MRS parameterization of the parton distribution functions. The distributions are presented on a linear scale to show that the valence quarks dominate for $x \gtrsim 0.2$ and the parton sea, especially g_s, dominates for $x \lesssim 0.2$. Notice that the distribution functions for the u and d valence quarks are different, not only in amplitude (as the proton contains twice as many u valence quarks), but also in shape. The u quark carries more of the proton's momentum, on average, than a d quark.

In Figure 2.2, the same parton distributions are plotted on a log scale. On a log scale, the sea distributions can be more easily compared. Since the $s(x)$, $c(x)$, and $b(x)$ distributions are less important for this analysis, they are summed in this figure.

As yet, no theoretical basis exists to calculate the parton distributions, so the several groups (e.g., MRS and CTEQ) who parameterize parton distributions use different, though similar, parameterizations. Differences between parton distributions are clearly visible in Figure 2.3, presenting the u and d valence quark distributions for different pdfs. The magnitude and the shape of the pdfs differ.

When a proton and antiproton collide at high energy (i.e., with large momentum transfer), usually only one parton from the proton interacts with one parton from the

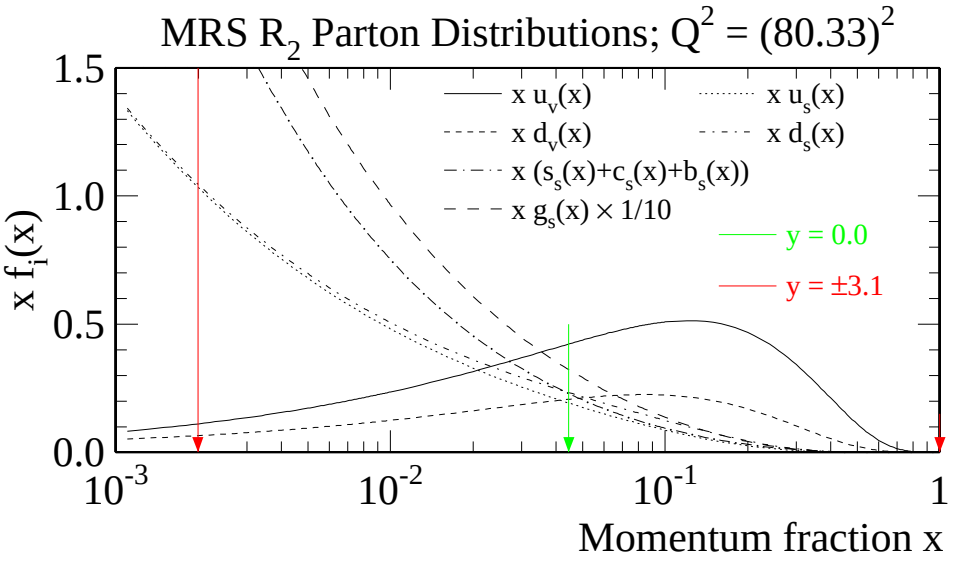


Figure 2.2: MRS-R_2 parton distributions at $Q^2 = M_W^2$, presented on a log scale. The log scale allows the sea distributions to be compared. Note that $d_s(x) > u_s(x)$, if not by a large factor. Separately, each of $s_s(x)$ and $c_s(x)$ and $b_s(x)$ is smaller than $u_s(x)$ and $d_s(x)$. The gluon distribution $g_s(x)$ is scaled by 1/10 for clarity. The arrows indicate the x carried by partons which create a W with a given boost, indicated by y.

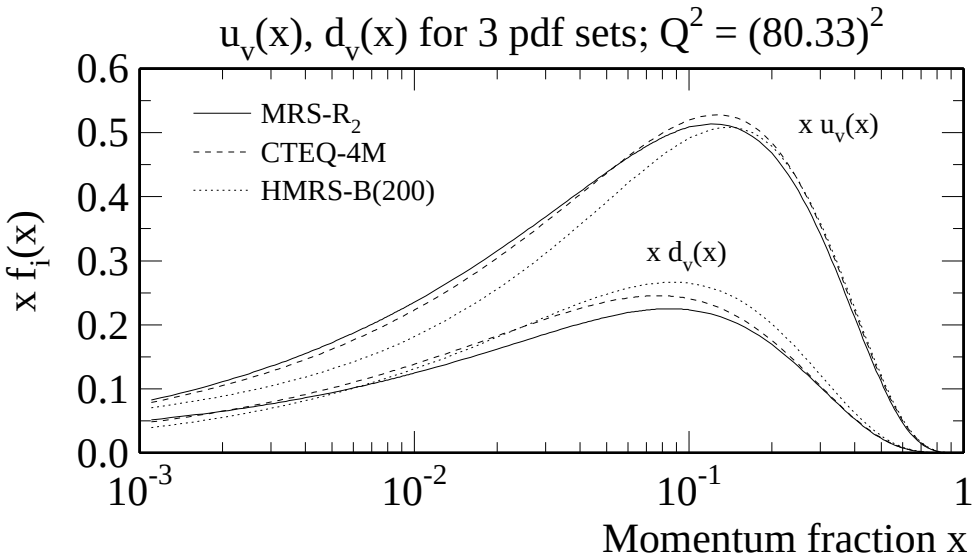


Figure 2.3: Comparing MRS-R_2 and CTEQ-4M parton distributions, both at $Q^2 = M_W^2$. Only the u and d valence distributions are shown here, for illustration. Notice that the *shapes* of the distributions differ as a function of x.

antiproton. Remember that partons bound into a hadron are asymptotically free; thus, *while still bound inside the proton*, a parton can interact with another particle, at high Q^2, as if free. Only after the parton scatters does it interact with the rest of the proton's partons. The QCD force increases with strength linearly as the parton separates from the proton. For a high Q^2 reaction, this force will become large enough to create particle-antiparticle pairs from the vacuum. In this manner, the struck parton and the rest of the proton hadronize to remain colorless objects, and the outgoing parton hadronizes into a jet of particles travelling roughly collinearly to the original direction of the outgoing parton.

The remaining partons from the baryon are "spectators" to the interaction, and interact with the struck parton only at low Q^2 during hadronization. These low Q^2 non-perturbative interactions add low energy hadrons throughout the event in η and ϕ, collectively referred to as the "underlying event," and the spectator quarks hadronize into forward and backward jets. Underlying event hadrons generally have an energy around 300 MeV, the energy scale of hadronization. The underlying event includes all particles not directly associated with the high Q^2 interaction.

2.5 $\overline{\mathrm{p}}\mathrm{p} \rightarrow W^{\pm} + X$ and W Production Charge Asymmetry

W^+ (W^-) bosons are created at CDF when a u (d) quark from a proton annihilates with a $\bar{d}$ ($\bar{u}$) quark from an antiproton. Figure 2.4 contains two Feynman diagrams showing the creation of a W^+: $\overline{\mathrm{p}}\mathrm{p} \rightarrow W + X$ is necessarily a weak interaction. Figure 2.6(a) shows the first-order process including the presence of the other partons in the proton and antiproton. The spectator partons, not shown in Figure 2.4, hadronize into forward and backward jets.

At high center-of-mass energy, a $\overline{\mathrm{p}}\mathrm{p} \rightarrow W$ interaction probes a limited range of momentum fraction, x. Pretending the W width is zero (the width , of a particle is inversely related to the lifetime of the particle: $, _W = 1/\tau_W$), a first-order process (i.e., Figure 2.4(a)) can create a W only when the center-of-mass energy of the *parton* interaction equals M_W. In fact, the parton interaction center-of-mass energy for a first-order

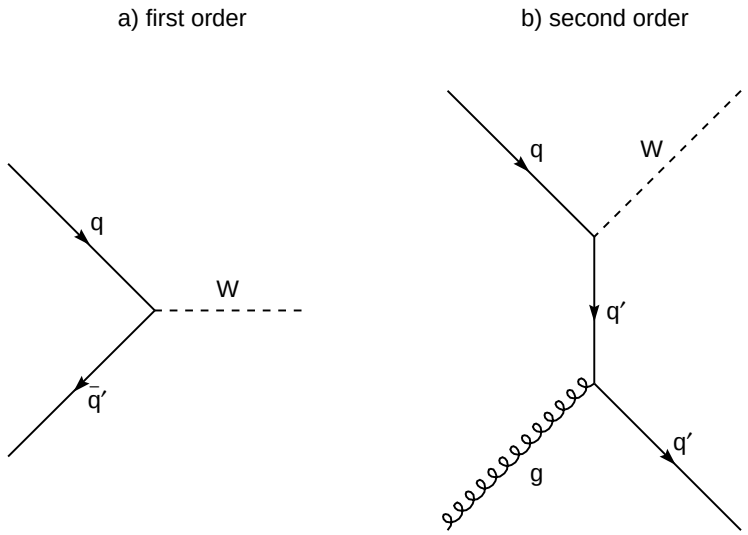


Figure 2.4: Feynman diagram showing W creation. (a) shows the first-order Feynman diagram involved in W creation in a $\overline{\mathrm{p}}\mathrm{p}$ collider. To create a W^+, q = u and $\overline{\mathrm{q}}'$ = $\overline{d}'$. (The c-$\overline{d}'$ and c-$\overline{s}'$ contributions are negligible.) (b) shows the second-order Feynman diagram involved in W creation, reduced over the first diagram by a factor of $\alpha_s(M_W) \approx 0.1$. To create a W^+, q = u as before; however, q′ = d'.

process may equal $M_W \pm$, $_W$, and the W width is large (, $_W$ = 2.07 GeV/$\hbar$) reflecting its very short lifetime. In addition, second and higher order processes allow the parton center-of-mass energy to be different from M_W. These details do not substantially change the following discussion of kinematics.

To create a W exactly at rest in a first-order process, each parton must have a momentum fraction of

$$x_0 = \frac{M_W}{\sqrt{s}} = \frac{80.33 \text{ GeV}}{1800 \text{ GeV}} = 0.0446.$$

Again pretending the W width is zero, a W^+ with a longitudinal Lorentz boost y is created when the u and $\overline{d}$ have momentum fractions x_1 and x_2, respectively:

$$x_1 = x_0 e^y; \qquad x_2 = x_0 e^{-y}.$$

For a first-order process, the W Lorentz boost is entirely longitudinal; however, multiple soft gluon emission (i.e., from higher-order processes such as shown in Figure 2.4(b) when the gluon is soft) gives the W a small (less than 20 GeV/c) but significant transverse momentum. Multiple soft gluon emission does not change the calculation for x_1 or x_2. See Ref [8] for a thorough discussion.

The maximum possible parton momentum fraction x_1 is 1.0, so the maximum W^+ rapidity is equal to $e^{-y_{max}} = x_0$. Thus, $y_{max} = 3.1$. The minimum momentum fraction which can create a W^+, given $\sqrt{s} = 1800\ GeV$, is $x_1(\text{min}) = (x_0)^2 = 0.002$, corresponding to a minimum W^+ rapidity of $y = -3.1$. The x fractions corresponding to these limits (i.e., $|y| = 3.1$ and $y = 0$) are marked with arrows in Figure 2.2. For the momentum fractions x which can result in $\overline{\text{p}}\text{p} \rightarrow W^{\pm}$, as indicated by the arrows on this figure, the parton sea contributes significantly. At a $\overline{\text{p}}\text{p}$ center-of-mass energy of 1.8 TeV, however, non u and d quarks do not make a large contribution to the cross section to create a $W^{\pm}$. (The non-u-d contribution is of order 3%.) As $\sqrt{s}$ increases, lower momentum fractions are probed, and thus, the sea distributions (especially non u and d distributions) become increasingly significant in W creation.

As shown in the previous section, u and d quarks carry different fractions of the proton's momentum. Thus, in general, W's are generated with substantial longitudinal

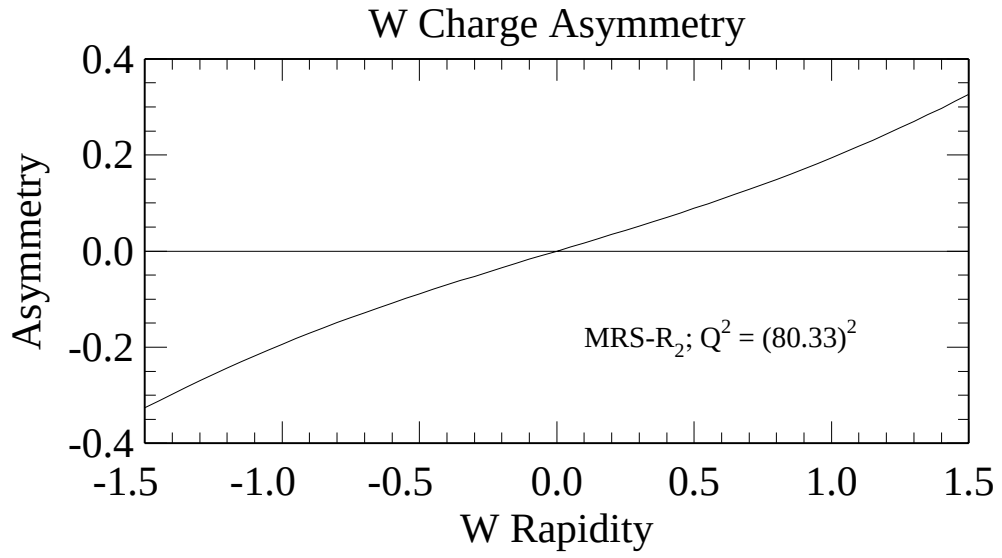


Figure 2.5: Charge asymmetry of W using the MRS-R_2 pdfs, ignoring next-to-leading order processes but including all combinations of quarks which can produce a W at leading order. The W asymmetry is a monotonically increasing function across the full kinematically realizable region, about half of which is shown in this figure.

momentum; the distribution of W boost depends on the slope of $u(x)/d(x)$ around $x_0 = M_W/\sqrt{s}$; as shown below. (See, for example, Ref [9, Section IV.B.].) The net W^+ boost is equal in magnitude and opposite in direction to the net W^- boost.

The W charge asymmetry is defined as

$$A_W(y) = \frac{d\sigma^+/dy - d\sigma^-/dy}{d\sigma^+/dy + d\sigma^-/dy}$$

where $d\sigma^\pm/dy$ is the differential cross section to produce a $W^\pm$. Neglecting non $u - \overline{d}$ contributions (which are small at $\sqrt{s} = 1.8$ TeV) and next-to-leading order processes, the W charge asymmetry is approximately

$$A_W(y) \approx \frac{u(x_1)d(x_2) - u(x_2)d(x_1)}{u(x_1)d(x_2) + u(x_2)d(x_1)}$$

with x_1 and x_2 functions of y as above. That is, $d\sigma^+/dy$ is proportional to $u(x_1)d(x_2)$ and $d\sigma^-/dy$ is proportional to $u(x_2)d(x_1)$; all remaining factors in the differential cross sections for W^+ and W^- cancel in the ratio. Figure 2.5 shows the result of such a calculation using the MRS-R_2 pdfs. Remember that a u quark (from the proton) with

momentum fraction x_1 and a $\overline{d}$ quark (from the antiproton) with momentum fraction x_2 interact to create a W^+ with rapidity y.

Introducing the ratio $R_{du}(x) = d(x)/u(x)$, the asymmetry can be rewritten:

$$A_W(y) \approx \frac{R_{du}(x_2) - R_{du}(x_1)}{R_{du}(x_2) + R_{du}(x_1)}$$

and for small y, $R_{du}(x_2) - R_{du}(x_1) \approx -2y(dR/dy) = -2yR'$, so

$$A_W(y) \approx -x_0\ y\ \frac{R'_{du}(x_0)}{R_{du}(x_0)}.$$

Therefore, the W charge asymmetry, at least for central rapidities, probes the *slope* of the ratio $u(x)/d(x)$ at $x = x_0$. Since the W itself cannot be detected—it decays in $1/,_W$ sec, or approximately 10^{-25} sec—the W charge asymmetry cannot be measured directly. Thus, the W charge asymmetry must be measured by identifying its decay daughters. Since jet-jet processes from W decays (e.g., $W \rightarrow c\overline{s}$) are largely indistinguishable from jet-jet processes from purely QCD interactions, W's are normally identified through their leptonic decays.

The effect of W decay on the observed charge asymmetry of the lepton is described in Section 2.7.

2.6 Lepton Universality

To first order, when a W decays, nine equally likely possibilities exist:

$$W^+ \rightarrow e^+\nu_e, \quad W^+ \rightarrow \mu^+\nu_\mu, \quad W^+ \rightarrow \tau^+\nu_\tau,$$

$$3 \times (W^+ \rightarrow u\overline{d}'), \qquad 3 \times (W^+ \rightarrow c\overline{s}'),$$

where the factors of 3 come from the three color degrees of freedom available to quarks. In the Standard Model, "lepton universality" expresses the assumption that the W and Z couple to all leptons with universal strength. In fact, experimentally, $B(W \rightarrow l\nu) = (10.8 \pm 0.4)\%$ for $l = e, \mu, \tau$, where $B(W \rightarrow X)$ indicates the branching fraction (or decay fraction) of the W to X. The W decays to hadrons slightly more frequently than one would naively expect: $B(W \rightarrow \text{hadrons})\ /\ 6 = (11.3 \pm 0.25)\%$.

Using the assumption of lepton universality, thus,

$$\sigma(\overline{\mathrm{p}}\mathrm{p} \to W) \cdot B(W \to e\nu) = \sigma(\overline{\mathrm{p}}\mathrm{p} \to W) \cdot B(W \to \mu\nu) = \sigma(\overline{\mathrm{p}}\mathrm{p} \to W) \cdot B(W \to \tau\nu)$$

and a precision measurement of $\sigma \cdot B(W \to e\nu)$ can be taken as a measurement of $\sigma \cdot B(W \to l\nu)$ for all charged leptons l. This assumption, used in Chapter 5, allows me to properly scale a $W \to \tau\nu$ monte carlo simulation to the measured cross section for $W \to e\nu$.

2.7 Lepton Charge Asymmetry

As mentioned in the previous section, about 11% of the time a W decays into a tau plus a neutrino or anti-neutrino. Given that positive and negative W's have different rapidity distributions, positive and negative taus also have different rapidity distributions. Tau rapidity is correlated with W rapidity; the lepton rapidity distribution convolutes the W rapidity distribution with the W angular decay distribution, where the polar angle distribution is the one of interest. Given lepton universality, all leptons have the same angular distribution in the W rest frame, and thus, in the lab frame, except to the degree that lepton mass differences change the relativistic factor γ for each lepton. This discussion focuses on taus.

Even if W's decay into taus isotropically in the W rest frame, the net $W^{\pm}$ boost causes positive and negative taus to have different event rapidity distributions. Due to the W boost, positive (negative) taus tend to be boosted in the proton (antiproton) direction, the same direction the $W^{\pm}$ is travelling.

In fact, the W does *not* decay into taus isotropically in its rest frame, since the W's created in high energy $\overline{\mathrm{p}}\mathrm{p}$ interactions are fully polarized. In a $\overline{\mathrm{p}}\mathrm{p}$ interaction with $Q^2 = M_W^2$, the u and $\overline{d}$ quarks are *highly* relativistic: relativistic $\gamma > 1000$ for W creation across all central rapidity, $|y| < 1.2$. In fact, at Tevatron energies, the lowest possible γ occurs when $x_u = 1.0$ and $x_d = 0.002$; the d quark's relativistic γ is about 180, still highly relativistic, i.e., $\beta = 0.999985$. Thus, at Tevatron energies, W's are practically 100% polarized across the entire kinematic range for W creation. This polarization is shown in Figure 2.6(a).

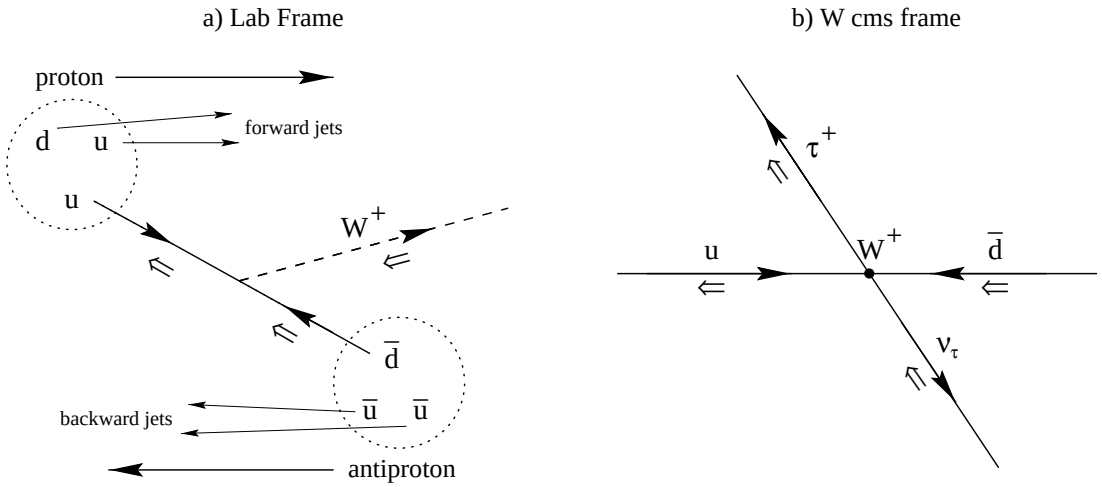


Figure 2.6: Helicity of W and tau in $\bar{\mathrm{p}}\mathrm{p} \to W + X; W \to \tau\nu$. In (a), the full proton-antiproton "collision" is shown, with one parton from each hadron annihilating to form a W. The remaining partons in the proton and antiproton then hadronize into forward and backward jets. In (b), the same interaction is shown, simplified, in the W center-of-mass frame. Also shown in this frame is the subsequent decay to tau + neutrino. The helicity of each particle is shown; at high Q^2, the W overwhelmingly interacts only with negative helicity particles and positive helicity antiparticles, corresponding to right-handed antiparticles and left-handed particles.

Particles of spin 1/2 and non-zero mass may exist in either positive (right-handed) or negative (left-handed) helicity states, where the helicity sign indicates whether the particle spin is aligned $(+1)$ or anti-aligned (-1) with the direction of particle motion. Since a Lorentz transform can "overtake" a massive particle, causing it to change its direction of travel (relative to one's reference frame) and thus its helicity, helicity is not a good quantum number for massive particles. For highly relativistic particles, however, helicity is a good quantum number to a very good approximation.

A massless spin 1/2 particle may exist in either a positive or negative helicity state, but due to the $V - A$ (that is, $1 - \gamma_5$) nature of W interactions, W's interact only with negative helicity neutrinos and positive helicity anti-neutrinos. That is, the weak interaction violates parity maximally. If right-handed massless neutrinos exist, no interaction in the Standard Model can detect them; in the Standard Model, neutrinos are massless and exclusively left-handed.

When a W decays into a tau and a neutrino, the constrained neutrino helicity forces the tau to be polarized as well, as demonstrated in Figure 2.6(b). In the $W^{\pm}$ rest frame,

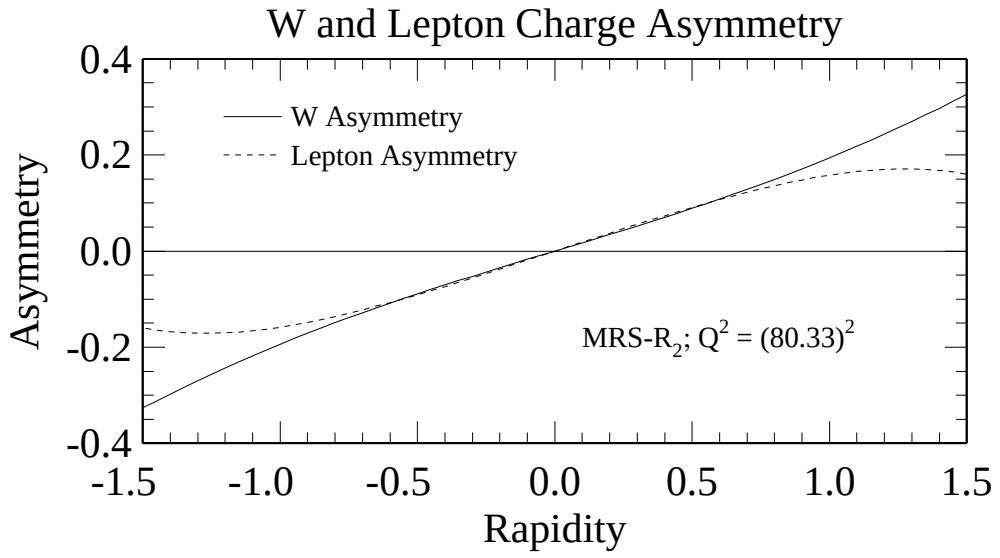


Figure 2.7: W and tau charge asymmetry using the MRS-R_2 pdfs. The W asymmetry is the same as that in Figure 2.5. The lepton charge asymmetry is calculated using a monte carlo integration at next-to-leading order [10].

taus are distributed as $(1 \mp \cos\theta)^2$, where θ is measured with respect to the proton beam direction. Note that this angular distribution is opposite in direction to the W boost, so the two effects compete. For W's with a small boost, the W polarization is a larger effect in the lab frame; for W's with a large boost, the W polarization is a small effect in the lab frame.

As mentioned in Chapter 1, "pseudorapidity" η is a useful coordinate in $\overline{\mathrm{p}}\mathrm{p}$ collisions; for a massless particle, the Lorentz boost y is equal to the pseudorapidity η. Pseudorapidity η is measured geometrically, and is in fact defined relative to the polar angle θ from the proton beam direction:

$$\eta = -\ln\tan\frac{\theta}{2}.$$

Thus, although y for a particle produced in $\overline{\mathrm{p}}\mathrm{p}$ collisions can be measured only if the mass of the particle is known or can be measured, the pseudorapidity η can be measured for all detected particles from the collision. Although, in general, $y \neq \eta$ for the W (due to its large mass relative to its momentum), this is not the case for the much lighter tau. Since the η is what is actually measured using the detector, the following discussion,

when referring to the rapidity of the tau, uses η.

Tau charge asymmetry is parameterized in the same manner as W asymmetry:

$$A(\eta_\tau) = \frac{d\sigma^+/d\eta_\tau - d\sigma^-/d\eta_\tau}{d\sigma^+/d\eta_\tau + d\sigma^-/d\eta_\tau},$$

where $d\sigma^+/d\eta_\tau$ ($d\sigma^-/d\eta_\tau$) is the cross section for producing W^+ (W^-) times the branching ratio for W^+ (W^-) decays to taus as a function of tau rapidity (η_τ). Positive rapidity is defined as being in the proton beam direction. Using the results of a monte carlo integration at next-to-leading order [10] in W production, in Figure 2.7, I compare the charge asymmetry of the tau to the charge asymmetry of the W. The effect of W polarization is clear, although minimal for central η.

When the efficiency for measuring positive and negative taus is the same, the charge asymmetry in a bin of rapidity may be expressed as

$$A(\eta) = \frac{N_\tau^+(\eta) - N_\tau^-(\eta)}{N_\tau^+(\eta) + N_\tau^-(\eta)} \tag{2.1}$$

where $N_\tau^+(\eta)$ and $N_\tau^-(\eta)$ are the number of positive and negative taus, respectively, in a bin. In this analysis, I measure only those taus which decay hadronically. A hadronically decaying tau appears in the detector as a narrow hadronic jet with only a few charged particles (low charged multiplicity), as discussed in more detail in the next section.

Assuming CP conservation, which holds for leptons in the Standard Model,

$$\frac{d\sigma^+}{d\eta_\tau} = \frac{d\sigma^-}{d(-\eta)_\tau}.$$

That is, the cross section times branching ratio vs. rapidity to produce a τ^+ is the same as the cross section times branching ratio vs. $-$rapidity to produce a τ^-. Given this, $A(\eta_\tau) = -A(-\eta_\tau)$ and the data can be folded about $\eta = 0$ to improve statistics.

2.8 τ Lepton

The tau bears the distinction of being the only lepton heavy enough to decay into hadrons. Precision measurements of the tau during the last several years [11, and references therein] are summarized in Table 2.3 and Table 2.4. Note that the n-prong

Property	Value
Mass	1777.00 $\pm$ 0.3 MeV
Mean lifetime	(290.7 $\pm$ 1.3) $\times 10^{-15}$ s
$c\tau$	87.2 $\pm$ 0.5 μm
$\gamma c\tau$	2.0 $\pm$ 0.01 mm (for 40.165 GeV/c)
B(1-prong)	84.96 $\pm$ 0.14% (total)
B(1-prong)	35.18 $\pm$ 0.14% (leptonic)
B(1-prong)	49.76 $\pm$ 0.20% (hadronic)
B(3-prong)	14.91 $\pm$ 0.14%
B(5-prong)	0.10 $\pm$ 0.01%

Table 2.3: Properties of the tau lepton. $c\tau$ is a measure of the distance a relativistic tau travels before decaying. For a tau of a given γ, this measurement can be given in units measured in the lab. B(n-prong) lists the branching fraction for a tau to decay to n charged particles. Due to rounding, these branching fractions do not add exactly to 100%.

decays with $n > 1$ are essentially purely hadronic. The only leptonic $n > 1$ prong tau decay discovered so far is $\tau^- \rightarrow e^-e^-e^+\overline{\nu}_e\nu_\tau$ with a branching fraction of $(2.8 \pm 1.5) \times 10^{-5}$ [11]. It is thus safe to consider tau decays with $n > 1$ to be purely hadronic.

The most recent world-average measurement of the tau decay modes is summarized in Table 2.4. All tau decays are used in this analysis except those which are purely leptonic, but only the 1-prong and 3-prong decays contribute significantly.

A typical hadronic $W \rightarrow \tau\nu$ decay contains a narrow hadronic jet with one or three tracks. The invariant mass of all the tau decay products (including neutrino(s)) is equal to the tau mass, 1.78 GeV. Since neutrinos are undetected, the invariant mass of all *detected* tau decay daughters is less than this. In addition, W events are "clean," with little activity in the event beyond the tau decay daughters and underlying event. This is especially true when the W is not recoiling against a gluon or quark jet (the second and higher order processes). Remember that the underlying event is generated by color strings, e.g., a quark being separated from a hadron, as discussed in Section 2.1. In a first-order W process such as the one shown in Figure 2.6(a), color strings exist only between the spectator quarks. The W is colorless, so W events are relatively "cleaner" (in the central region) than QCD 2-jet events. The net result is that leptons from W decay are often isolated in the event, with little activity from the underlying event nearby.

τ^- Decay Mode	Type	Branching Fraction (%)
$\mu^-\bar{\nu}_\mu\nu_\tau$	leptonic	17.4 ± 0.1
$e^-\bar{\nu}_e\nu_\tau$	leptonic	17.8 ± 0.1
$\pi^-\nu_\tau$	1-prong	11.3 ± 0.2
$\rho\nu_\tau;\ \rho \to \pi^-\pi^0$	1-prong	25.2 ± 0.2
$a_1^-\nu_\tau;\ a_1^- \to \rho^-\pi^0;\ \rho^- \to \pi^-\pi^0$	1-prong	9.3 ± 0.1
$\pi^- \geq 3\pi^0$	1-prong	1.3 ± 0.2
$\pi^-\bar{K}^0 \geq 0$ neutrals	1-prong	1.3 ± 0.1
$K^- \geq 0\pi^0$	1-prong	1.4 ± 0.1
$K^-K^0 \geq 0\pi^0$	1-prong	0.3 ± 0.04
$a_1^-\nu_\tau;\ a_1^- \to \rho^0\pi^-;\ \rho^0 \to \pi^+\pi^-$	3-prong	9.4 ± 0.1
$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	3-prong	2.6 ± 0.1
$\pi^-\pi^+\pi^- \geq 2\pi^0\nu_\tau$	3-prong	0.2 ± 0.07
$3h^-2h^+\nu_\tau$	5-prong	0.1 ± 0.01
$\pi^-\eta^0\pi^0\nu_\tau$	misc	0.2 ± 0.03
$h^-\omega^0 \geq 0\pi^0\nu_\tau$	misc	2.3 ± 0.1

Table 2.4: Tau branching fractions, rounded to the tenth of a percent, from the PDG fit basis modes [11]. "Neutrals" refers to π^0's, K^0's and $\overline{K^0}$'s; "$h^\pm$" refers to either $\pi^\pm$ or $K^\pm$. Due to rounding, the branching fractions do not add to exactly 100%. The two resonances a_1 and ρ contribute a significant fraction of the hadronic decay modes.

Leptonically decaying taus from $W \to \tau\nu$ are, in general, impossible to separate from direct production of $W \to \mu\nu$ and $W \to e\nu$. Therefore, events containing identifiable electrons and muons are rejected using selection cuts which are efficient for keeping hadronic tau decays. The details of these selection cuts are discussed in Chapter 4. Rejecting electron and muon events increases the purity of the tau sample. This is especially important because $W \to \mu\nu$ and $W \to e\nu$ also produce leptons with charge asymmetry, and I wish to measure the charge asymmetry of taus, not the charge asymmetry of electrons or muons.

An important feature of taus from $W \to \tau\nu$ is their polarization, which merits discussion. Since the W bosons are produced fully polarized, the tau from W decay is also fully polarized—in the rest frame of the W. However, remember W's are normally produced with a substantial longitudinal Lorentz boost relative to the lab frame. Thus, in the lab frame, taus are not 100% polarized [12].

The tau polarization is important in hadronic tau decays since—due again to the

left-handedness of the electroweak force—the tau neutrino constrains the angular distribution of the hadron or hadrons in the decay. If taus *were* fully polarized from the W, the more common hadronic decays would be constrained as follows. For $\tau \to \pi\nu$, since the pion is a scalar (spin 0) particle, the neutrino must carry all the spin. Checking Figure 2.6(b), a positive tau from W decay is right-handed. Due to parity violation, the tau anti-neutrino from the τ^+ decay must be right-handed. Thus, in $\tau \to \pi\nu$, the neutrino travels in the tau direction and the pion travels in the opposite direction (with $1 \pm \cos\theta$ distributions where θ is measured with respect to the tau direction and the "+" case is for the neutrino). Thus, the neutrino from tau decay is emitted in the opposite direction from the neutrino involved in the tau creation, and the pion energy in the lab is reduced.

The $\rho^\pm$ and $a_1^\pm$ are both vector (spin 1) particles; each has three possible z projections for spin. However, since the total spin must add to 1/2, the z component of spin must be conserved, and the tau neutrino is left-handed, only two decay configurations are possible. In the first, the $\rho^\pm$ or $a_1^\pm$ is preferentially emitted in the tau direction with the same handedness as the tau; in the second, the $\rho^\pm$ or $a_1^\pm$ is preferentially emitted (in the tau rest frame) in the direction opposite the tau, with its z projection of spin equal to zero. However, the $\rho^\pm$ and $a_1^\pm$ decays to pions are more involved, especially for the $a_1^\pm$ since it decays to $\rho\pi$—that is, $\rho^\pm\pi^0$ and $\rho^0\pi^\pm$.

Since, in the lab frame, taus are not 100% polarized, these polarization effects are reduced but still important. All monte carlo simulations involving tau decays carefully account for all polarization effects, as described in Chapter 5.

2.9 Other Measurements of Asymmetry

CDF has published measurements of the W decay asymmetry using electrons and muons [13]. These results are somewhat more precise than any possible measurement using taus, simply because it is easier to get a relatively pure and high-statistics sample of $W \to e\nu$ and $W \to \mu\nu$ events. Figures 1 and 2 from Ref [13] are reproduced here as Figure 2.8(a) and (b). In Figure 2 from Ref [13], the data are folded about $y = 0$ to

increase statistics, as mentioned in Section 2.7 above.

The tau charge asymmetry has not been measured before. Assuming lepton universality, it may seem superfluous to measure the charge asymmetry of taus, especially given the reduced statistics of a measurement using taus, but lepton universality is not certain. Lepton universality is an assumption of the Standard Model. Some models posit a force whose strength depends on the mass of the object the force couples to. Since the tau is the heaviest lepton, such processes couple to the tau at a much higher rate than the other leptons. A charged Higgs boson, as well as certain SUSY models, contains such a coupling.

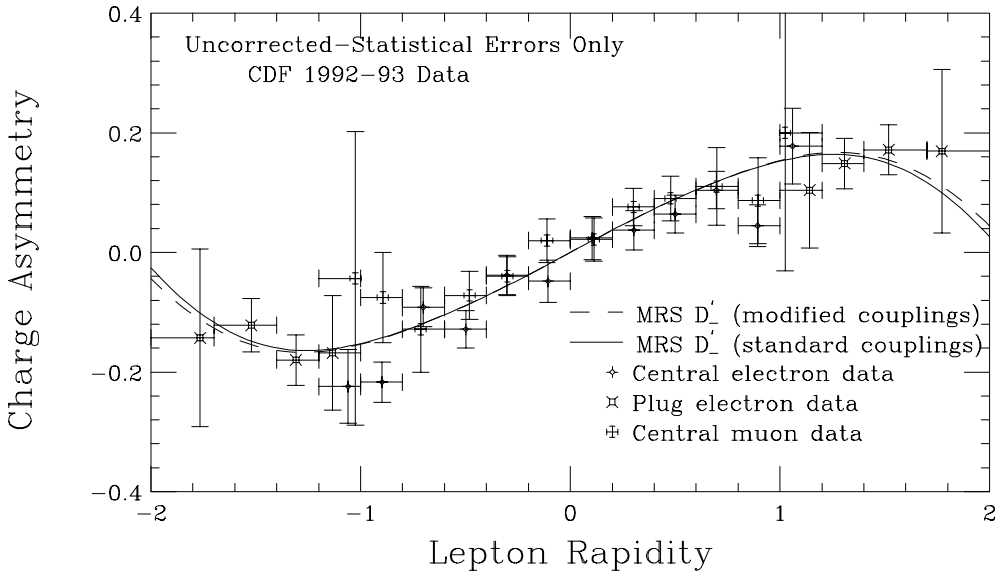
In these models, the electron and muon asymmetry measurements primarily contain W decays; the tau asymmetry measurement contains a mixture of W decays and these other processes. Thus, the tau asymmetry, if consistent with electron and muon asymmetry, provides limits on physics beyond the Standard Model.

2.10 Background Processes

An event selection designed to select $W \to \tau\nu$ events with a high efficiency also selects several background processes with high or fair efficiency. For example, $W \to e\nu$ events closely resemble $W \to \tau\nu$ events; events containing electrons need special attention in the event selection cuts. That is, events containing loosely-identified electrons are rejected. These electron rejection cuts are not 100% efficient at removing electrons from the event sample; the tau efficiency would be unacceptably low if they were. Thus, some $W \to e\nu$ events remain in the final event sample. Similar backgrounds are present from other Standard Model electroweak processes, namely $Z \to ee$. Event selection cuts which reject electrons from $W \to e\nu$ are effective at removing $Z \to ee$ as well.

Although $W \to e\nu$ and other electroweak processes are easily modeled using monte carlo simulations, other backgrounds are more difficult to model well. The QCD process $\overline{\mathrm{p}}\mathrm{p} \to$ two or more jets (e.g., Figure 2.9(a)–(c)) has a *much* higher cross section than $W \to \tau\nu$, and such QCD processes depend strongly on non-perturbative physics. Thus,

(a) e, μ charge asymmetry; all rapidity



(b) e, μ charge asymmetry folded about $y = 0$

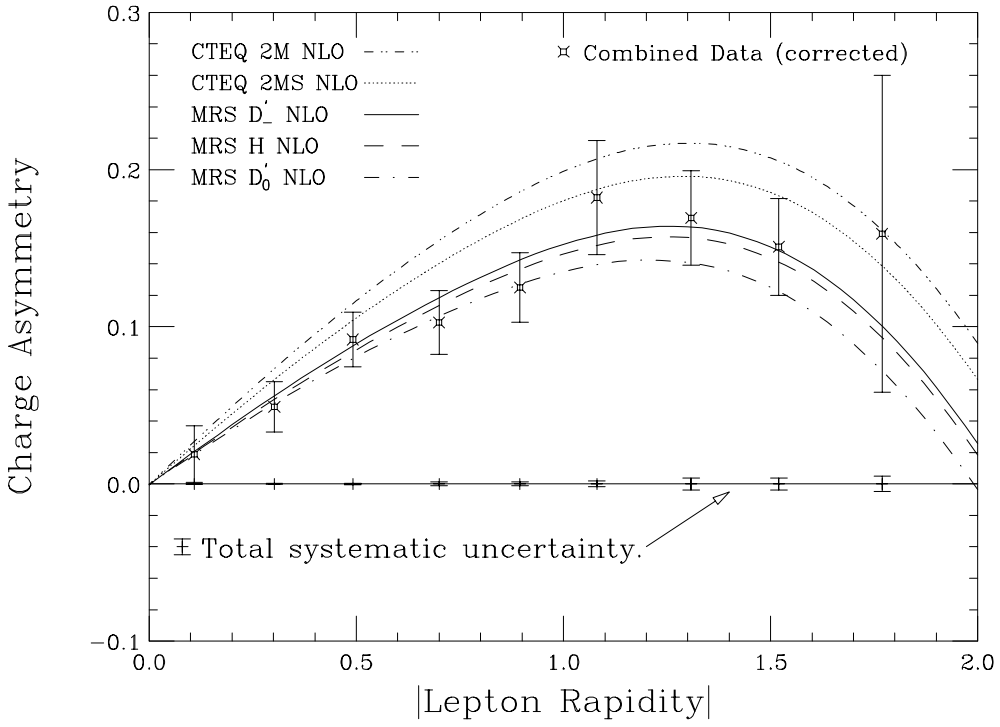


Figure 2.8: Electron and muon asymmetry measurement at CDF from Ref [13]. Figures 1 and 2 from this publication are reproduced above as (a) and (b), respectively.

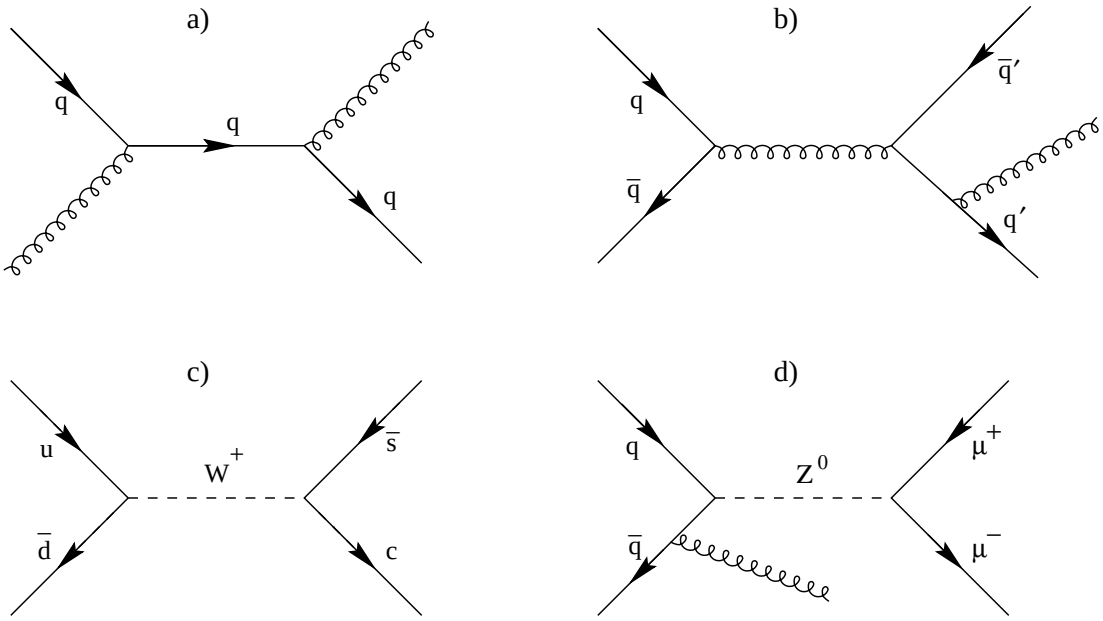


Figure 2.9: Feynman diagrams for background processes. a) and b) demonstrate QCD processes leading to 2- and 3-jet processes. In c) a W^+ decays hadronically. Since the result is two hadronic jets, each jet coming from an initial-state parton, this processes is classified as QCD background even though it is an electroweak process. d) presents an electroweak background process.

these QCD processes can be modeled only phenomenologically. Current phenomenological monte carlo simulations model the underlying event—especially the forward and backward jets—poorly. Thus, a separate event sample is used here to model the QCD background.

Although only a small percentage of prompt hadronic jets (i.e., jets whose parent particle is a quark or gluon from the $\overline{\text{p}}\text{p}$ interaction) "fragment" (hadronize) in a way that resembles a hadronic jet from hadronic tau decay, the much higher cross section for purely QCD processes means this small number is still a substantial fraction of the tau event sample.

The processes $W \to \mu\nu$ + jets and $Z \to \mu\mu$ + jets—both next-to-leading order processes—can enter the tau event sample if the QCD jet recoiling against the W fragments in a way which resembles the hadronic decay of a tau. To remove such background events, muons are loosely identified as a track which deposits little energy in the calorimeter. Events with such loosely identified muon are rejected. Not only does this background rejection remove a large fraction of $W \to \mu\nu$ and $Z \to \mu\mu$ backgrounds, but

it also removes much of the $Z \rightarrow \tau\tau$ background.

Thus, electroweak backgrounds are efficiently, if not fully, removed from the final event sample. The remaining small background from electroweak processes may be modeled using a monte carlo event generator.

The most troublesome background, therefore, is the prompt hadronic jet, or QCD background, mainly because the cross section for QCD $\overline{\mathrm{p}}\mathrm{p}$ interactions resulting in 2-jet processes is so much larger than the cross section for $W \rightarrow \tau\nu$. (As compared with the electroweak processes, which have a cross section equal to or smaller than the $W \rightarrow \tau\nu$ cross section.) The different processes—electroweak for $\overline{\mathrm{p}}\mathrm{p} \rightarrow W + X$, strong for $\overline{\mathrm{p}}\mathrm{p} \rightarrow$ jets—involved in the creation of taus and prompt hadronic jets provide some handles to differentiate between a prompt hadronic jet (hereafter called a "QCD" jet for brevity) and a tau jet. A QCD jet is composed solely of strongly-interacting particles; the partons from the $\overline{\mathrm{p}}\mathrm{p}$ collision interact with one another while hadronizing. These partons are connected by color strings to the spectator quarks. The ensuing strong interactions spread underlying event particles throughout the event, some near the QCD jets.

In contrast, since the W is not a strongly-interacting particle, it is not involved in the strong interactions which form the underlying event. The spectator quarks still interact strongly with one another—W events are not completely free of underlying event. Still, W events are generally much cleaner than prompt hadronic jet processes. Therefore, when the W decays to a tau, that tau and its decay daughters are usually isolated in the event. This difference provides one handle for separating prompt hadronic jets from tau hadronic jets.

Additionally, as a tau becomes more energetic in the lab frame, its decay daughters subtend an increasingly smaller solid angle due to the Lorentz boost given the particles by the tau momentum. On the other hand, a high momentum parton knocked out of a proton, in general, hadronizes into a wide jet with a large number of particles when compared with a lower momentum parton: As these partons increase in momentum, more energy is put into color strings and the color strings thus break more often. Also, since these partons interact with spectator quarks during hadronization, prompt

hadronic jet hadronization is "messier" than that of a hadronic jet from a tau decay.

No matter what a tau's momentum, a strictly limited amount of energy is available to the tau decay daughters in the tau rest frame. When a tau decays through a virtual W into a quark pair, the color string between these quarks has only so much energy, limiting both the total number of hadrons and the invariant mass of the hadrons. Unlike prompt hadronic jet hadronization which varies with the momentum of the initial parton, the hadronization of tau decay daughters is independent of the tau momentum.

I use these differing characteristics of QCD jets and hadronic jets from tau decay to reduce the QCD background as much as possible without compromising efficiency for identifying taus.

Chapter 3

Apparatus

Look, it cannot be seen—it is beyond form.
Listen, it cannot be heard—it is beyond sound.
Grasp, it cannot be held—it is intangible.
— Lao Tsu [1]

This chapter describes the Fermilab accelerators and the Collider Detector at Fermilab (CDF), including the data acquisition and trigger systems.

3.1 The Tevatron

The Tevatron is a superconducting accelerator which collides beams of protons and antiprotons at a center of mass energy of 1800 GeV, currently the highest energy collider in the world.

During Run 1a of the Tevatron (August 1992–May 1993), six proton bunches and six antiproton bunches circulated in opposite directions. The energy of each particle in the bunches is 900 GeV; thus, the center-of-mass energy of collisions is 1800 GeV. The maximum luminosity during Run 1a was $9.2 \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$, and the average was about $3.5 \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$. A typical store was injected with a luminosity of $5\text{–}8 \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$ and lasted about 12–18 hours with the luminosity dropping to $(1\text{–}3) \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$ at the end of the store.

The event rate R, given a cross section σ and luminosity $\mathcal{L}$, can be calculated as

$$R = \sigma \times \mathcal{L}.$$

From this, the total number of events, N, generated or observed given a cross section and luminosity is

$$N = \sigma \times \int \mathcal{L}\, dt.$$

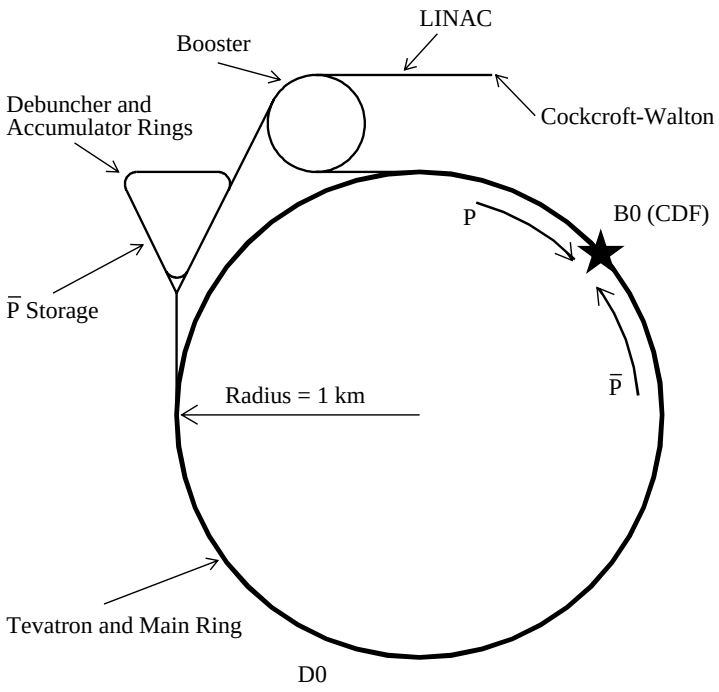


Figure 3.1: Schematic of the Fermilab $\bar{p}p$ accelerators.

The integral $\int \mathcal{L}\, dt$ is called "integrated luminosity," and its units are often given as events/picobarn, abbreviated pb^{-1}.

Given a small, fixed cross section of an interesting physics process, the only way to increase the rate at which this physics process occurs is to increase the luminosity. Given a process with a cross section of 1 nb, or 10^{-33} cm^2, and an accelerator luminosity of $1\times10^{30}\text{cm}^{-2}\text{s}^{-1}$, this process occurs with a rate of 1/1000 Hz, or roughly four events per hour. The Tevatron's high luminosity (compared with other hadronic colliders) is thus instrumental in extending the physics reach of experiments at Fermilab, including CDF.

The luminosity attained in the Tevatron can be calculated from accelerator properties as

$$\mathcal{L} = \frac{N_p N_{\bar{p}} B f_0}{4\pi\lambda^2},$$

where N_p ($N_{\bar{p}}$) is the number of protons (antiprotons) per bunch, B is the number of bunches of each type, f_0 is the revolution frequency, and λ is the cross sectional area of

the bunches. Since the Tevatron radius is 1 km and the beams travel at approximately the speed of light, $f_0 = c/2\pi$ km = 47.7 kHz. There are six bunches of each type, thus, the bunches are separated by just under 3.5 μs. Since the beam spot is about 60 μm across, $\lambda \approx 3.6 \times 10^{-5}$ cm^2. During Run 1a, typical values for the bunch occupancy were about $N_p = 15 \times 10^{10}$ and $N_{\bar{p}} = 5 \times 10^{10}$. Using these values in the above equation yields a typical luminosity of about 5×10^{30} cm^{-2}s^{-1}.

With increasing luminosity, the probability of a proton and antiproton interacting during each bunch "crossing" approaches unity. In fact, at luminosities much above 5×10^{30}cm^{-2}s^{-1}, more than one $\overline{\mathrm{p}}\mathrm{p}$ interaction occurs, *on average*, during each crossing. These "multiple interactions" are discussed in more detail in Section 3.6.

The Tevatron is described in more detail in Appendix C.

3.2 CDF Detector Overview

The CDF detector is described in detail elsewhere [14]. Here, I briefly discuss those components of the CDF detector which are important in this analysis.

The CDF detector is a general-purpose detector designed to measure electrons, photons, muons, hadronic jets, and neutrinos coming from high energy $\overline{\mathrm{p}}\mathrm{p}$ interactions in the Tevatron. Many detector subsystems are used in these measurements; these subsystems are divided by polar angle in the detector (central, plug, and forward) and by function (tracking, electromagnetic calorimetry (EM), electromagnetic strip chambers for precise position resolution in the EM calorimeters, hadronic calorimetry (HAD), and muon detection).

The CDF detector is approximately 27 meters long, 10 meters high, and weighs about 5000 tons. Figures 3.2 and 3.3 schematically show the detector. From the interaction region out, a particle produced transversally in a collision encounters, in this order, the silicon vertex chamber (SVX), vertex time projection chamber (VTX), central tracking chamber (CTC), the superconducting solenoid, electromagnetic (EM) calorimeters with embedded strip chambers, hadronic (HAD) calorimeters, and muon chambers.

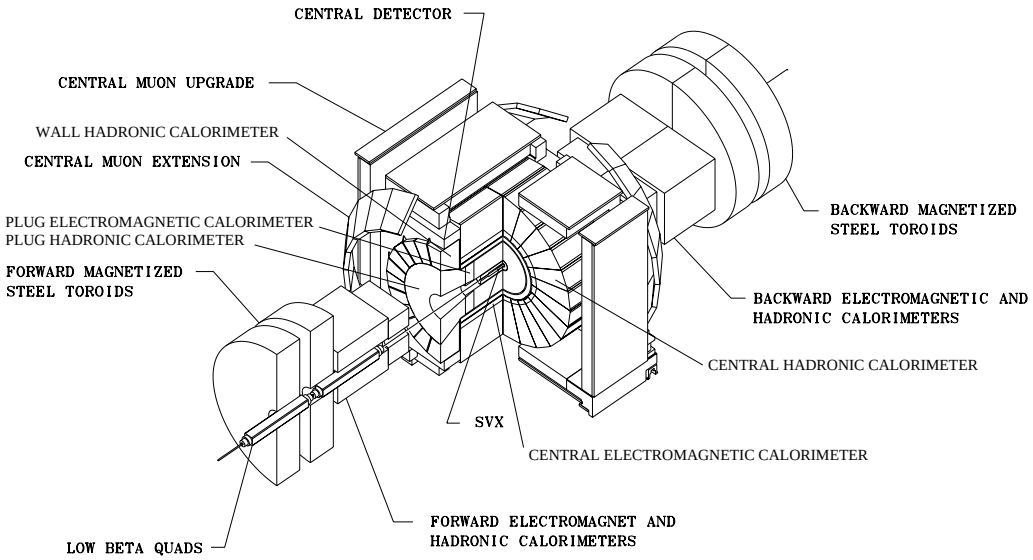


Figure 3.2: 3d cut-away view of the CDF detector.

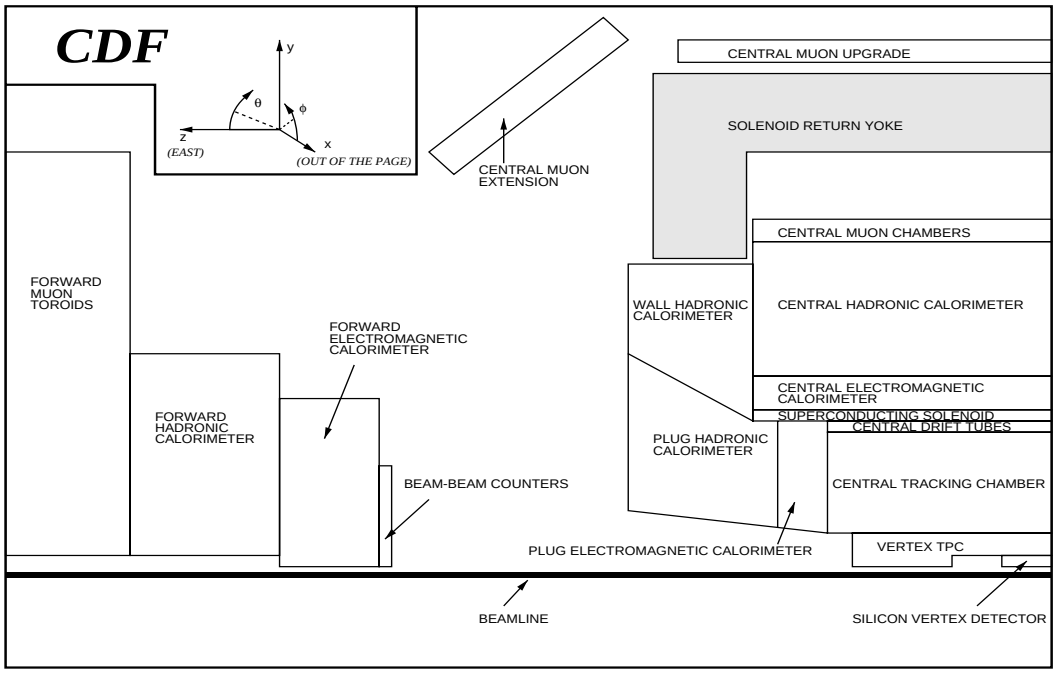


Figure 3.3: Side view of 1/4 of the CDF detector. The interaction region is on the right edge of the figure at the beamline. The detector is forward-backward symmetric and most systems are axially symmetric about the beamline.

At small angles with respect to the beam, in front of the forward calorimeters, the beam-beam counters (BBC) monitor instantaneous luminosity. Separate from the detector are the trigger and data acquisition (DAQ) systems. The trigger selects which events are to be saved to tape and the DAQ electronically reads data from detector components and moves this information to the proper electronics subsystems, ultimately to magnetic tape for those events accepted by the trigger.

Thus, as an example of how these subsystems are used together to identify particles, electrons and photons are measured using EM calorimeters (including embedded strip chambers in the central and plug EM calorimeters) and tracking chambers (VTX, CTC, and sometimes SVX). Experimentally, a photon appears as an electron without a track. In another example, energetic muons penetrate the calorimeters without depositing much energy, so muons are identified and measured using the tracking and muon chambers. (A muon chamber is essentially an abbreviated tracking chamber mounted behind layers of steel thick enough to absorb most other particles.)

Jets, e.g., from the hadronization of a gluon or quark or from tau decay, are collimated groupings of particles associated with the same parent particle, and so are measured using the tracking chambers and both EM and HAD calorimeters to group together closely associated neutral and charged particles. Since the b-quark has a long lifetime, and thus, a B meson travels a measurable distance before decaying (one to several mm, even many cm for very energetic B mesons), b-quark jets can be separated from other jets by using the SVX to measure a displaced vertex, that is, the charged tracks from the B meson's decay daughters meet at a point which is displaced in a direction transverse to the beamline. Finally, the presence of neutrinos must be inferred, since neutrinos do not interact in any CDF detector. Thus, using all of the calorimeters to look for an imbalance of E_T, one can infer the presence or absence of neutrinos, as described in Section 3.3.2.

3.3 CDF Coordinate Systems

Two primary coordinate systems are used at CDF, a cylindrical coordinate system and a cartesian coordinate system, sharing the same origin and z axis. The origin is in the center of the detector where collisions occur. The z axis is defined as the proton direction and θ is the polar angle from the z axis. The y axis points straight up out of the center of the detector and the x axis points radially outward away from the center of the Tevatron. The azimuthal angle ϕ is measured from the x axis. We usually describe events using ϕ and pseudorapidity, η, defined from the polar angle θ as

$$\eta = -\ln \tan \frac{\theta}{2}.$$

Two forms of η are used at CDF. Detector η (η_D) is measured using the center of the detector ($z = 0$) as the origin. Event η is measured using the event z-vertex as the origin.

"Forward" ("backward") is used to refer to the direction the proton (antiproton) travels; positive η is forward, negative η is backward. Forward and backward are used to refer to detector systems at large $|\eta|$. Sometimes all the detector systems at large $|\eta|$ are collectively referred to as forward. The CDF detector is also divided into an "east" and "west" half; the forward direction is east, backward is west.

One of two variables is used, most often, to measure the distance (the angle, really) between two objects in the detector. The flow of particles from a $\overline{\mathrm{p}}\mathrm{p}$ collision is more naturally described using η than using θ; hadronic collisions distribute particles more uniformly in η. Thus, the distance is more commonly measured in $\eta - \phi$ space using $\triangle R$, defined as

$$\triangle R = \sqrt{(\phi_1 - \phi_2)^2 + (\eta_1 - \eta_2)^2},$$

where ϕ is measured in radians. It is occasionally useful to measure the distance (angle) between two particles in $\theta - \phi$ space, with both angles measured in radians.

The calorimeter is divided into many "towers," with each tower measuring the energy in a rectangular region of the calorimeter as seen from the center of the detector. CDF's calorimeters are designed with a projective tower geometry, meaning towers do

not overlap in η or ϕ when looking from the center of the CDF detector, at $z = 0$. To put it another way, a straight line starting from the center of the CDF detector (say, a very high momentum particle) encounters exactly one tower. This is not strictly true when the event z-vertex is not zero; for the range of z-vertices encountered at CDF, however, the overlap between towers is small.

When referring to position local to one calorimeter tower, it is useful to use the coordinates x-local and z-local. These coordinates are measured in cm from the center of a calorimeter tower, most frequently in an EM calorimeter tower at the depth of the central electron strip detector. Although x-local is measured in the ϕ direction, it is measured as a distance, not an angle, so this coordinate is given a name to suggest a distance, not an angle; z-local is measured along the z axis.

Two common slices or projections through the CDF detector merit discussion here. A slice parallel to and through the beamline such as that seen in Figure 3.3 is referred to as an $r - z$ view, where r measures distance from the event vertex. An $r - z$ view is usually about 45° "thick" in ϕ. For example, as described in Section 3.4.1, the VTX detector measures tracks in the $r - z$ view. The ϕ of most VTX tracks is known only to within the 45° segmentation of the VTX. An axial projection through the CDF detector, i.e., down the z axis, is referred to as an $r - \phi$ view. The CFT, as described in Section 3.7.5, measures the projection of tracks in the $r - \phi$ plane only.

3.3.1 Clustering and Jet Clustering

Since a single particle, not to mention a jet of particles from a common parent particle, is usually measured by many detector subsystems, the trigger and offline processing correlate related detector information when possible. A hadronic shower from a one or more pions, for example, often shares its energy among many calorimeter towers. The *physics* quantity of interest is the energy of the pion(s), so the energy in calorimeter towers must be "clustered." Jet clustering looks for groups of calorimeter towers with correlated energy deposits; collimated jets of particles are assumed to be from the same parent particle, so jet clusters may contain many towers.

Several means of jet clustering exist, but two are used most often at CDF: cone

clustering and nearest-neighbor clustering. Both clustering methods start with a "seed" tower, where the seed tower is either the first tower found over an E_T threshold or the highest E_T tower in the cluster, depending on the details of the clustering algorithm. Both clustering algorithms also calculate a "cluster centroid," the E_T-weighted center of the cluster in η and ϕ, as well as the total cluster E_T and the invariant mass of the shower.

In cone clustering, all towers in a certain cone (i.e., range of $\triangle R$) from the cluster centroid are added into the cluster. The cone size is a fixed quantity; CDF uses cone clustering with cones of 0.4, 0.7, and 1.0 radian in $\triangle R$. I use only the 0.4 radian $\triangle R$ clusters. Cone clustering is an iterative process: For each iteration, the cluster centroid location is recalculated; this new centroid is used for the next iteration. Iterating stops when the cluster centroid position is stable or after a maximum number of loops.

In nearest-neighbor clustering, all neighboring towers to the seed tower are checked. Each tower over an E_T threshold is added to the cluster, and then the neighbors of each new tower are checked. The Level 2 trigger cluster finder uses this algorithm. Some forms of nearest-neighbor clustering limit how wide (in $\triangle R$) the cluster can get; other forms do not, depending on the needs of the clustering algorithm.

Additional information can be correlated with calorimeter clusters—charged tracks for example. A charged track can be extrapolated to the radius of the calorimeter to see if it is associated (by closeness in $\triangle R$, for example) with a cluster. Sometimes the initial track direction is compared in η and ϕ (i.e., $\triangle R$) to the calorimeter cluster centroid.

Different initial states can often be distinguished by correlating information from many detector subsystems.

3.3.2 E_T, P_T, $\not\!\!E_T$, and Neutrinos

Most physics studied at CDF involve high-momentum interactions; this includes W and Z physics, e.g., the W physics discussed in Chapter 2. These interactions generate particles with significant transverse momentum, that is, momentum transverse to the beam line. By comparison, most elastic or diffractive $\overline{\mathrm{p}}\mathrm{p}$ collisions result in events

with little transverse momentum. In addition, most inelastic $\overline{\mathrm{p}}\mathrm{p}$ collisions are "soft," generating little transverse momentum.

Since the proton and antiproton have no transverse momentum and equal and opposite longitudinal momentum, the total momentum, transverse and longitudinal, of the products of the collision sums to zero in a perfect calorimeter (i.e., full 4π solid angle coverage). The detector must leave a space for the beampipe, however, so some particles from the collision (especially hadrons from the spectator quarks' hadronization, as mentioned in Section 2.4) travel down the beampipe or the uninstrumented space around the beampipe. These particles completely elude the calorimeters and all other detectors.

To so elude the calorimeters, a particle must travel at a very small polar angle, two degrees or less, so its transverse momentum is less than 3.5% of its total momentum. Such unmeasured particles do not carry away significant amounts of transverse momentum, but they may carry away a significant amount of longitudinal momentum. Therefore, longitudinal momentum usually does not balance in the detector. Within the detector's accuracy, however, transverse momentum *does* balance, save for presumably small losses down the 2° forward and backward holes. A particle travelling at a narrow angle to the beamline from the center of the detector, if it is not measured by the forward EM or forward HAD calorimeters, travels down the 2° hole, as shown in Figure 3.3. See also Figure 3.8.

For these reasons, rather than using E and P to describe a particle, we use transverse energy, E_T, and transverse momentum, P_T, which are defined as follows:

$$E_T = E \times \sin\theta,$$

$$P_T = P \times \sin\theta.$$

In the large energies found at CDF, E_T is nearly equal to $P_T$for most particles. We use these variables in a specific way, however: We use E_T to describe transverse energy deposited in a calorimeter and P_T to describe transverse momentum measured in a tracking chamber. I will follow this convention.

In order to account for transverse momentum balance in a collision, we also find

useful the variable "missing transverse energy" ($\not\!\!E_T$), usually measured only in calorimeters.

In general, E_T does balance in an event—a high-E_T jet on one side of the detector is balanced by a high-E_T jet on the other side of the detector, but several effects can cause an imbalance of E_T. These effects include calorimeter energy resolution, non-interacting particles (such as neutrinos) which deposit no energy in the calorimeters, minimum-ionizing particles (such as muons) which deposit little of their energy in the calorimeters, and particles which travel down a non-instrumented region ("crack," e.g., the gap between two calorimeter towers) of the detector. Minimum-ionizing particles and charged particles travelling down a crack are detected by the track they leave in any tracking chambers they traverse, but particles which don't interact with any part of the detector, (e.g., neutrinos) can be measured only by finding a significant imbalance of E_T when summed over all calorimeter towers, treating the E_T as measured in each tower as a vector quantity, essentially, momentum.

Thus, we define $\not\!\!E_T$ as the magnitude of the vector sum over towers of E_T:

$$\not\!\!E_T = \sqrt{(\sum_{towers} E_T \cdot \sin\phi_{tower})^2 + (\sum_{towers} E_T \cdot \cos\phi_{tower})^2}.$$

If all particles in an event deposit all of their energy in the calorimeters, only calorimeter energy measurement error will cause $\not\!\!E_T$ to be non-zero. The calorimeter energy resolution scales with the square root of E_T for all calorimeter subsystems. While the electromagnetic calorimeters measure energy with relatively small uncertainty, the energy uncertainty in each hadronic calorimeter is approximately $(100/\sqrt{E_T\ (GeV)})\%$. (The uncertainty is somewhat smaller in the central region, somewhat larger everywhere else.) With this simplification, the E_T measured in a tower can be considered to be $(E_T \pm \sqrt{E_T})$ GeV. Adding in quadrature the uncertainty in the E_T measurement for each tower results in $\sqrt{\sum_{towers} E_T\ (GeV)}$ as the uncertainty of any measurement involving the E_T in all towers.

With this in mind, "$\not\!\!E_T$ significance," defined as $\not\!\!E_T/\sqrt{\sum_{towers} E_T\ (GeV)}$, is a measure of how significant the $\not\!\!E_T$ in an event is relative to the calorimetry energy measurement uncertainty. Values of $\not\!\!E_T$ significance well above 1.0 indicate some particle (such

as a neutrino) escaped detection. When the $\not\!\!E_T$ significance is less than about 1.0, the $\not\!\!E_T$ is more likely to be a result of measurement uncertainty.

One source of "fake" $\not\!\!E_T$ significance is minimum-ionizing particles such as muons which deposit little of their energy in the calorimeter. Although it is possible to correct the measured $\not\!\!E_T$ and $\not\!\!E_T$ significance for the presence of such particles, such a correction is unnecessary here since events containing minimum ionizing particles are rejected in this analysis.

Putting it all together, the presence of neutrinos is inferred by measuring a large amount of $\not\!\!E_T$ (above some threshold) which is significant ($\not\!\!E_T$ significance much greater than 1.0) when there are no signs of a particle travelling down a crack.

3.4 Tracking

The CDF detector contains three tracking chambers—in order from the beampipe out, the Silicon Vertex Chamber (SVX), the Vertex Time Projection Chamber (VTX), and the Central Tracking Chamber (CTC). The SVX is not used in the analysis, so I do not discuss it further.

3.4.1 Vertex Time Projection Chamber (VTX)

The VTX [15, 16] surrounds the SVX detector with inner radii of 20 cm (for the central-most modules surrounding the SVX) and 10 cm (for the modules at each end). All VTX modules have an outer radius of 25 cm. The VTX provides track information for $|\eta| \lesssim 3.25$. For most tracks, the VTX provides 2-dimensional information in the $r-z$ plane only (as defined in Section 3.3). Unless a track passes through two neighboring VTX modules, the ϕ of the track is known only to within the 45° of one octant of one VTX module.

The VTX is a gas chamber containing 50/50 Argon/Ethane bubbled through isopropyl alcohol at -7° C, and comprises 28 drift modules. Each drift module surrounds the beam pipe and covers about 10 cm in z. The layout of the VTX is shown schematically in Figure 3.4 from an event display. Each of the 28 drift modules is divided into

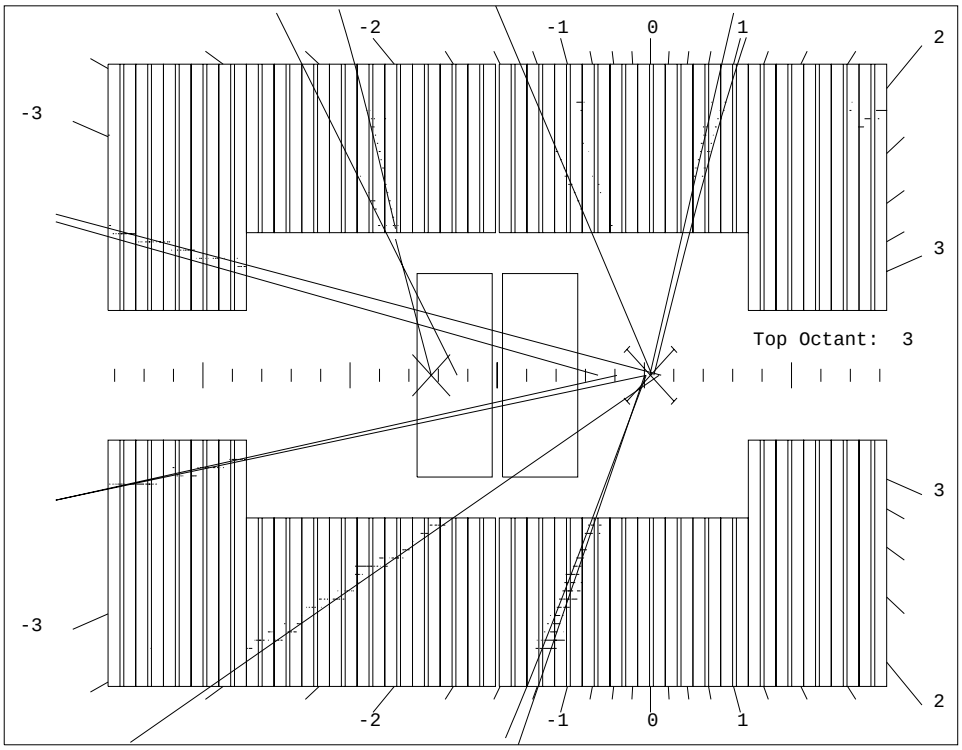


Figure 3.4: Event display showing one view, of four, of the VTX. The beam travels horizontally through the center with protons travelling to the right. The straight lines are VTX track segments found by reconstruction software. Hits are visible along the track segments. The X's along the beamline are vertices found by the VTX reconstruction software. The X with bars at the tips is a high quality vertex. Event eta (using the high quality vertex as $z = 0$) is indicated by the scale, -3 to $+3$.

two drift regions (in z), each about 5 cm long, for a total of 56 drift regions. The drift regions in each module are further divided into eight octants, which together surround the beam pipe and provide full coverage in ϕ. Each octant covers 45° in ϕ.

The two large boxes in the middle of Figure 3.4 represent the SVX. The division of the VTX into modules each with two drift regions is shown in this figure.

The ten outer drift modules, five on each side of $\eta = 0$, have 24 sense wires in each drift region. The nine internal modules on each side of $\eta = 0$ have 16 sense wires in each drift region. The 18 inner modules have a larger inner radius (thus, fewer sense wires), allowing the SVX to fit inside the VTX. Sense wires are strung azimuthally with eight straight sections, one in each octant. A sense wire plane is strung in each drift region of each module. Charged particles drift parallel to the beam toward the center of each drift module to the sense wires. The drift time provides z information.

Drift modules alternate between one of two ϕ orientations; each drift module is tilted 15° in ϕ relative to neighboring modules. Tracks which pass through more than one module have rudimentary ϕ information, but most tracks are measured only in the $r - z$ plane.

The VTX measures the z vertex of tracks with a 2 mm resolution. CTC track reconstruction software uses the z vertices found in the VTX as endpoints for CTC tracks. The VTX z-vertex resolution is much more accurate than the CTC z-vertex resolution. The radial position of the wire "hit" plus the drift time together provide $r - z$ tracking information.

3.4.2 Central Tracking Chamber (CTC)

The Central Tracking Chamber [17] is CDF's primary tracking chamber and the only tracking chamber at CDF inside the solenoid which measures a full 3-dimensional track trajectory. The CTC is a cylindrical drift chamber 3.2 m long covering the range $|\eta_D| < 1.0$; it surrounds the VTX and SVX. The inner and outer radii of the chamber are 277 mm and 1380 mm; the radii of the inner and outermost sense wires are 309 mm and 1320 mm. The axis of the CTC is the z axis of the CDF detector. The CTC is contained within the superconducting solenoid which generates an axial 1.41 T magnetic

field. Thus, charged particles follow a helical trajectory in the CTC. A charged particle's momentum is measured from the curvature of the helix with a resolution of $\delta P_T / {P_T}^2 \leq 0.002\ (GeV/c)^{-1}$.

Charged particles ionize the argon-ethane-ethanol gas (49.6/49.6/0.8 %) filling the CTC as they pass through it. The ionized electrons then drift to sense wires in a drift field of $\approx$ 1350 V/cm, with a drift time of 800 ns (which is shorter than the 3.5 μs crossing time of the Tevatron). The maximum drift distance is 40 mm.

As shown in Figure 3.5, the sense wires are grouped in 84 layers, which are further grouped into nine superlayers numbered 0 (innermost) to 8 (outermost). Each superlayer contains many drift cells, each with multiple sense wires. The limited size of each drift cell reduces the maximum drift time, and multiple sense wires per cell allow ambiguous or corrupted information to be dealt with efficiently by comparing to neighboring sense wires.

Five superlayers (layers 0, 2, 4, 6, and 8) are axial superlayers; these superlayers each contain drift cells with twelve sense wires mounted axially, i.e., parallel to the z axis. Four superlayers (layers 1, 3, 5, and 7) are stereo superlayers. Stereo superlayers each contain six sense wires and are tilted $\pm 3^\circ$ with respect to the z axis. The cells in superlayers 1 and 5 are tilted $+3^\circ$; the cells in superlayers 3 and 7 are tilted -3°. The axial superlayers provide $r - \phi$ (or $x - y$) information for a track; the stereo layers add information on z. Together, the axial and stereo layers measure the full helical path taken by a charged particle.

The position of a charged particle at the radius of a given sense wire is determined from the position of the sense wire and the distance corresponding to the drift time measured for that hit. In the plane of the wires, sense wires are separated by 10 mm. The two-track resolution is about 5 mm. The individual wire resolution is about 200 μm, and the efficiency per wire is greater than 98%.

Drift cells are tilted 45° with respect to the radial direction to compensate for the Lorentz angle of an ionization electron drifting in a crossed 1.41 T magnetic field and 1350 V/cm electric drift field. Thus, the net force on drift electrons is approximately azimuthal. Note in Figure 3.5 that drift cells overlap. Thus, a radial high-P_T track will

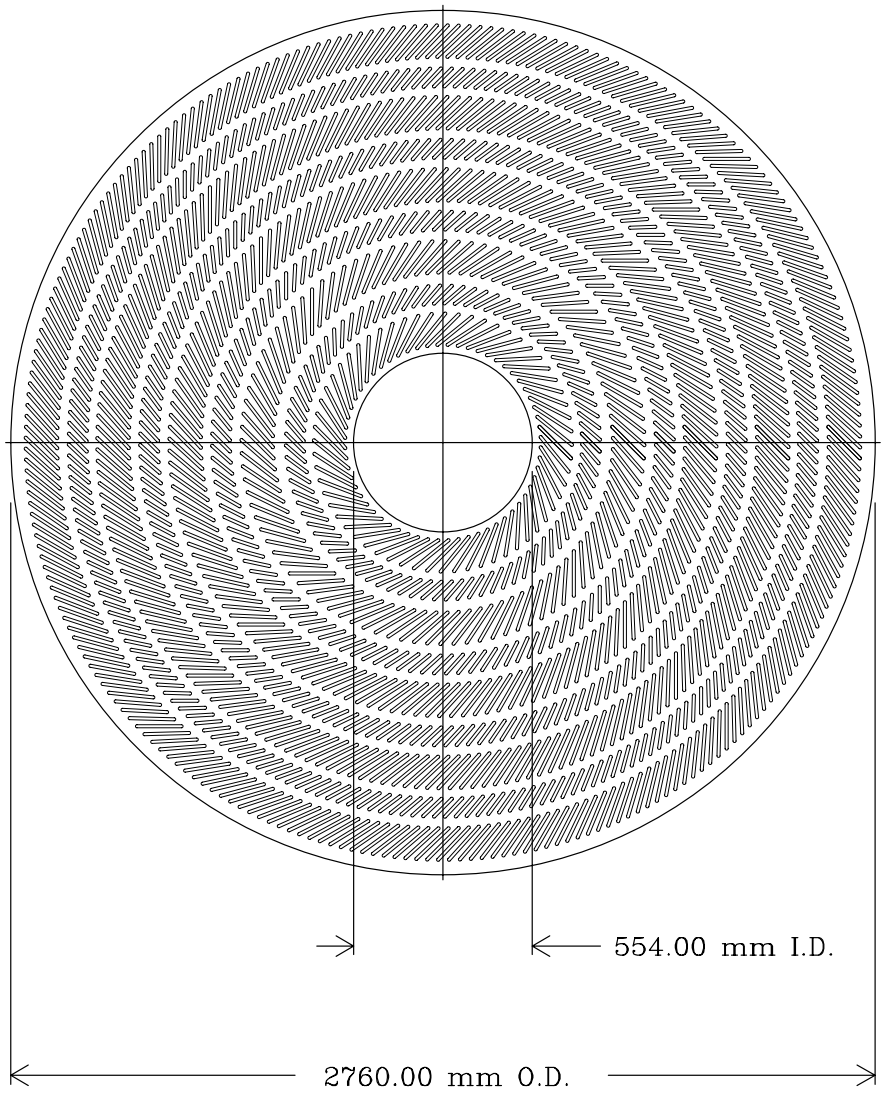


Figure 3.5: End view of the central tracking chamber. This schematic shows the grouping of sense wires into superlayers, as well as the tilt and overlap of the sense wire cells; every second slot contains sense wires. This view looks in the proton beam direction, down the positive z axis. In this view, positively charged particles curl clockwise with increasing ϕ; negative particles curl counter-clockwise.

Calorimeter	$\lvert\eta\rvert$ coverage	Tower Width η	Tower Width ϕ [°]	Energy Resolution σ [$\%/\sqrt{E_T}$]	Energy Resolution constant [%]
Central EM	0–1.1	0.11	15	13.5	1.7
Central HAD	0–0.9	0.11	15	75	3
Wall HAD	0.7–1.3	0.11	15	75	3
Plug EM	1.1–2.4	0.09	5	28	2
Plug HAD	1.3–2.4	0.09	5	130	4
Forward EM	2.2–4.2	0.1	5	25	2
Forward HAD	2.3–4.2	0.1	5	130	4

Table 3.1: CDF calorimeter information. The energy resolution is $\frac{\sigma(E)}{E} = (\frac{\sigma}{\sqrt{E_T}} \oplus C)\%$ where σ is the E_T dependent resolution and C is the constant resolution uncertainty. The E_T is measured in GeV and the E_T-dependent uncertainty is added in quadrature with the constant resolution term, as indicated by $\oplus$.

pass close to at least one wire in each superlayer. Another benefit from this 45° tilt is that it reduces the right-left ambiguity—an ionization electron can approach a sense wire from the right or left, and there is no instrumentation to distinguish between these alternatives. With the 45° tilt to the drift cells, only one of these two choices points toward the event vertex for a high-P_T track from the event vertex.

3.5 Calorimetry

Neutral particles, and charged particles with a transverse momentum greater than about 350 MeV/c, escape the solenoid's magnetic field and may be detected by the calorimeters. All CDF calorimeters are sampling calorimeters. That is, they are constructed from alternating layers of an absorbing material and an active material. Electrons lose energy in the absorber through bremsstrahlung, photons through pair production. All charged particles slowly lose energy through ionization, and pions and other hadrons lose energy through interactions with nuclei in the absorbing material.

The active layers measure the energy at that depth of the shower. At CDF, the active layers are scintillating plastic, which measures energy via the light deposited in a layer of plastic, and proportional gas tubes, which measure ionization produced when the shower passes through the tube. The signal from the active layer is proportional to the energy contained in the particle shower at that depth. In successive layers of

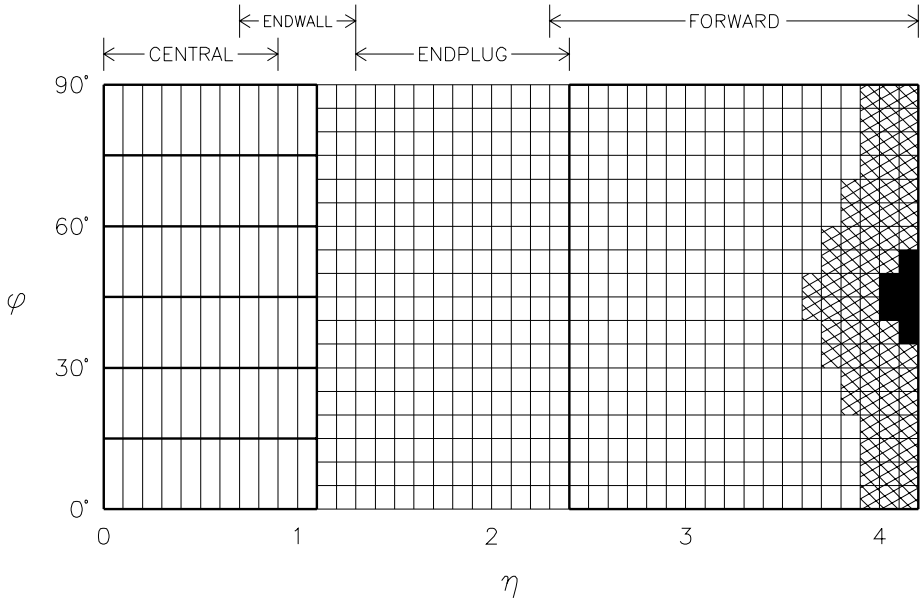


Figure 3.6: Tower segmentation of the CDF calorimeters, showing one of eight identical $\eta - \phi$ quadrants. The hatched area has only partial hadronic depth coverage, and the black area has no hadronic depth coverage, due to a cut for accelerator focussing magnets. The EM calorimeters have complete ϕ coverage out to $\eta = 4.2$. Heavy lines indicate module or chamber boundaries.

Figure 3.7: Central quadrant closeup. The relative position of the Wall Hadron and Plug calorimeters is clear in this side view. The Forward calorimeters are artificially moved toward the central detector, but everything else is to scale. The CES embedded in the CEM is also visible in this view.

absorber, the particles in a shower lose energy and shower, themselves, until all particles in a shower have low energy, well less than 1 GeV. The layers of the calorimeters are arranged so that particle showers are mostly perpendicular to the layers.

The thickness of an EM calorimeter is measured in "radiation lengths." An average high energy electron or photon loses all but $1/e$ ($e = 2.71828$) of its energy, through bremsstrahlung and pair production, in one radiation length. HAD calorimeter thickness is measured in nuclear interaction lengths. Of N neutral hadrons passing through one nuclear interaction length of absorber, all but N/e interact with a nucleus in the material. Muons are "minimum-ionizing" particles, losing little energy through bremsstrahlung and not interacting strongly with nuclei. Muon energy is not normally

measured with calorimeters, and is not measured using calorimeters at CDF.

Information from several separate calorimeter systems is combined to measure the electromagnetic and hadronic energy of such particles and jets in the CDF detector. Different calorimeters are separated by detector region. The central calorimeters include the Central EM [18], Central HAD [19], and at some pseudorapidities, Wall HAD [19] calorimeters. (See Figure 3.7.) The CEM contains strip chambers for improved position resolution. A plug fits into the central detector as an end cap and covers the η range $1.1 \leq |\eta| \leq 2.4$. The plug region includes the Plug EM [20] and Plug HAD [21] calorimeters. The forward detectors are located away from the central detector, as shown in Figure 3.3. The forward region includes the Forward EM [22] and Forward HAD [21] detectors.

The calorimeters at CDF are divided into 480 towers in a projective tower geometry as defined in Section 3.3. Towers are a constant width in η and ϕ for each subsystem. Each tower includes a hadronic calorimeter placed behind an electromagnetic calorimeter, allowing the HAD/EM ratio to be measured for any individual tower. The tower segmentation is shown in Figure 3.6. All together, the calorimeters cover the full 2π in azimuth and pseudorapidity out to $|\eta| = 4.2$. The η and ϕ widths of towers for each subsystem are listed in Table 3.1.

The CDF hadronic calorimeters are not compensating; their response to π^0's differs from their response to hadronic shower components of equal energy. This, in part, accounts for the worse performance of hadronic calorimeters compared with electromagnetic calorimeters at CDF. The energy measured for a jet depends on how the jet fluctuates and what fraction of the jet fluctuates into π^0's. In addition, the CDF hadronic calorimeters are designed for 95% containment for 50 GeV pions, i.e., a 50 GeV pion deposits, on average, 95% of its energy in the hadronic calorimeters with 5% of the shower energy not measured in any calorimeter. (This phenomenon is referred to as "punch-through.") The energy resolution of high energy particles degrades when hadrons are not contained within the calorimeter.

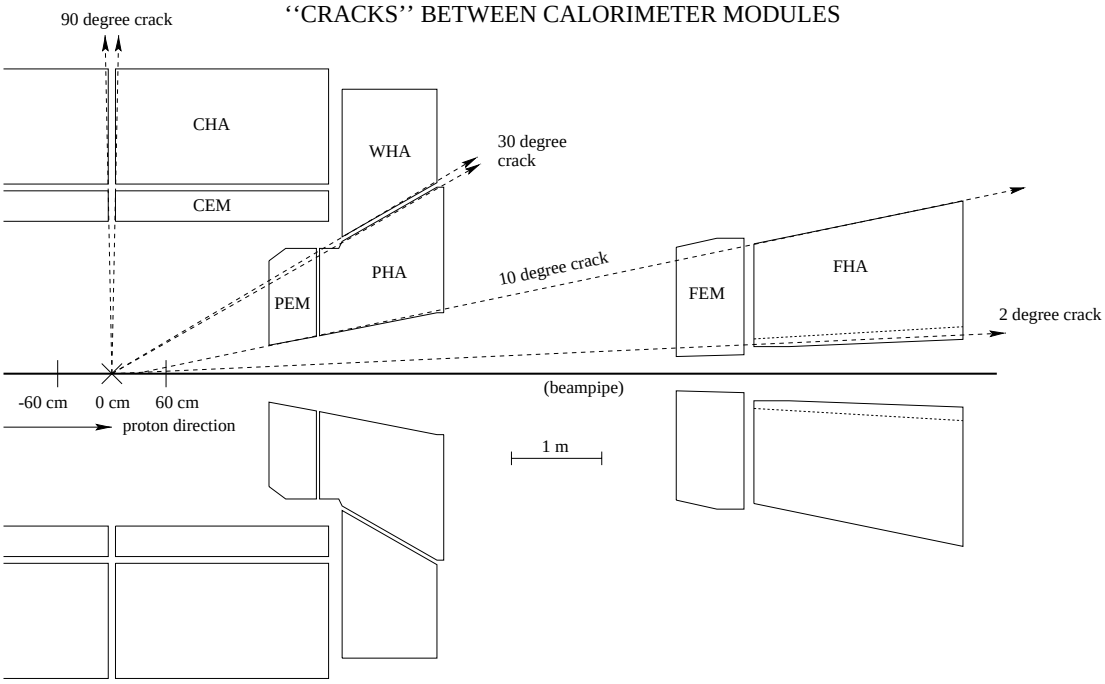


Figure 3.8: Calorimeter schematic showing "cracks." Only the *instrumented* regions of calorimeters are shown in this diagram; no other detector is shown. The four "η cracks" are plainly visible. Most collisions occur at $|z| < 60$ cm. The inner radius of the FHA depends on ϕ since the low beta quadrupoles surround the beampipe inside the FHA. In contrast, only the beampipe passes through the FEM. The FHA has a square opening for the low beta quads.

3.5.1 Calorimeter "Cracks"

Uninstrumented regions, or cracks, occur between calorimeter towers, where cabling or structural support is provided, and between the plug and forward calorimeters in η. Cracks appear in both η and ϕ, but the η cracks are substantially larger and so receive special attention here. The energy of a particle travelling through a crack is poorly measured; only that part of the particle shower entering an instrumented region of a calorimeter is measured. The rest of the particle shower, and thus the rest of the energy, is unmeasured.

The η cracks are illustrated in Figure 3.8. These cracks appear at a polar angle θ = 90°, 30°, and 10°. (Of course, corresponding cracks exist in the backwards region at 150° and 170°.) In addition, this Figure illustrates the 2° hole about the beampipe.

Electromagnetic calorimeter coverage is better than hadronic coverage, especially in the forward region, as illustrated in Figure 3.8. The hadronic calorimeter η cracks are larger. In addition, particles which enter a calorimeter near a crack, travelling toward the crack, are poorly measured. Any part of the shower which enters the gap between calorimeter modules is unmeasured.

3.5.2 Central Calorimeters

The central calorimeters are divided azimuthally into 24 wedges, and into an east and a west half in η. Thus, the central region is composed of a total of 48 wedges, each covering 15° in ϕ and a range of η: $0 < |\eta| <$ about 1.1. One wedge is shown in Figure 3.9. The wedges combine into four "arches"; each arch covers 180° in ϕ and either the east or west half of central calorimetry. Cracks appear every 15° in ϕ, between wedges, as well as at $\eta = 0$ (the "90° crack") between east and west arches. The ϕ cracks, together, account for 4.8% of central azimuthal angle coverage.

The Central EM calorimeter is constructed from 31 layers of lead absorber, 3.2 mm thick, alternating with layers of polystyrene scintillator, 5 mm thick, used as the sampling medium. The high Z of lead nuclei promotes electromagnetic showering of photons and electrons. The CEM is about 18 radiation lengths thick.

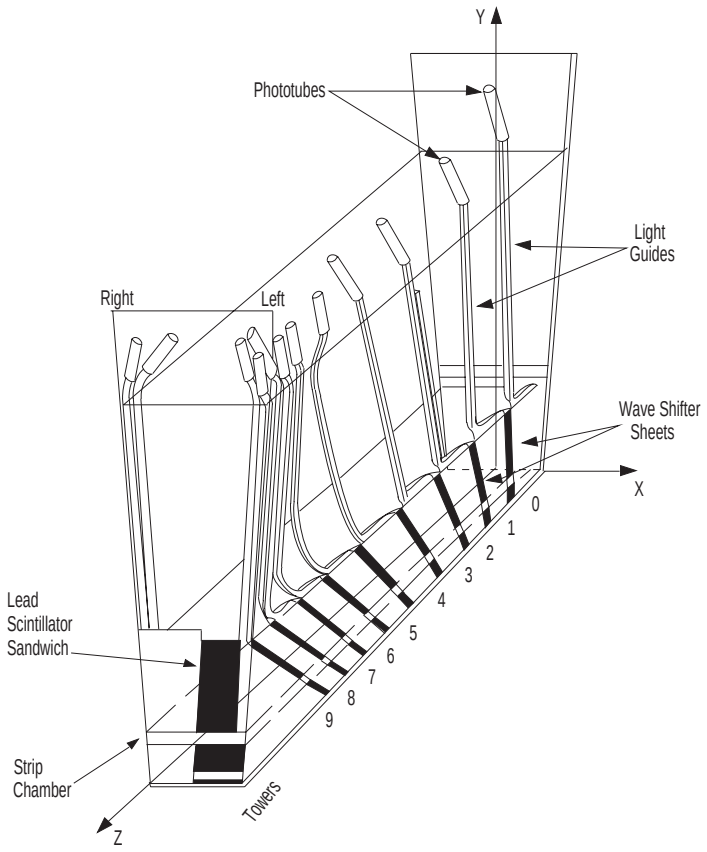


Figure 3.9: One central calorimeter wedge, made up of ten towers numbered 0–9. A central wedge contains ten EM towers, each with an embedded strip chamber (CES), and eight HAD towers (not shown in detail in this schematic) The HAD towers sit atop the lead-scintillator sandwich comprising the CEM. HAD towers 6–8 continue in the WHA; HAD towers 9–11 are contained entirely in the WHA.

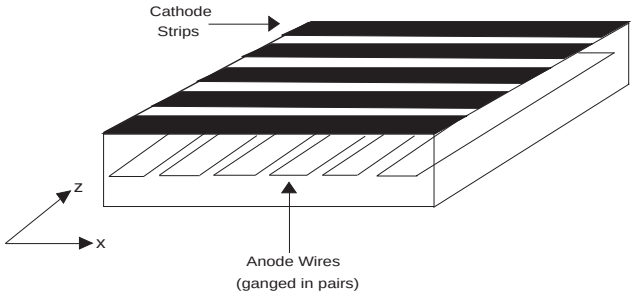


Figure 3.10: The central electron strip chamber. This diagram schematically illustrates one tower's length of the CEM in a wedge. The CES position resolution is about 2 mm in ϕ and η.

Embedded within the CEM is a proportional strip chamber (CES), shown in Figure 3.10. The CES is inserted in each CEM wedge at a depth corresponding to the maximum average transverse (to the shower) electromagnetic shower development, or the distance at which the greatest amount of initial photon or electron energy is deposited in the shower: about 5.9 radiation lengths deep. Orthogonal strips (perpendicular to the beam direction) and wires (parallel to the beam direction) measure the shower profile in the z and ϕ directions, respectively. The CES position resolution is about ± 2 mm in each view.

The Central HAD calorimeter is constructed from 32 layers of steel absorber, 2.5 cm thick, alternating with plastic scintillator 1.0 cm thick. The CHA is about 4.5 interaction lengths deep. Both the WHA and CHA measure part of the depth for towers in the region $0.7 < |\eta_D| < 0.9$; a particle in this region first passes through the CHA, then the WHA. Only the WHA measures particles in the region $0.9 < |\eta_D| < 1.3$.

The Wall HAD calorimeter is constructed from 15 layers of steel absorber, 5.1 cm thick, alternating with plastic scintillator 1 cm thick. The steel absorber layers in the WHA are twice as thick as those in the CHA because particles of the same E_T going into these calorimeters have $\sqrt{2}$ times more energy in the WHA. The WHA is about 4.5 interactions lengths deep.

3.5.3 Plug Calorimeters

Unlike the central calorimeters, the plug calorimeters are divided into 72 wedges in ϕ, each 5° wide. The plug EM and HAD calorimeters use gas proportional tubes (filled with argon-ethane) with cathode pad readout to measure energy. The gas proportional tubes are constructed of conductive plastic with a square cross section, and are strung with gold-plated tungsten wire. The tubes are arranged in layers perpendicular to the beam line.

The central hadron calorimeters leave uncovered the region with polar angle less than 30°. The plug calorimeters fill this hole, but there is a crack, named the "30° crack," at the interface.

The Plug EM calorimeter is built from four azimuthal quadrants which circle the

beam pipe; ϕ cracks appear between these quadrants. Thirty-four proportional tube arrays are interleaved with lead absorber 2.7 mm thick. The PEM is about 19 radiation lengths deep.

The Plug Hadronic Calorimeter (PHA) is constructed from 20 layers of steel, 5.1 cm thick, interleaved with drift tubes as in the PEM. The PHA is 5.7 interactions lengths deep.

3.5.4 Forward Calorimeters

The PEM and PHA leave a 10° hole in polar angle. The forward calorimeters fill this hole, but there is a crack (the "10° crack") in η at the interface. The forward calorimeters are constructed from gas proportional tubes like the plug calorimeters.

The Forward EM calorimeter is located about 6.5 m from the interaction point. It is constructed from 30 layers of lead (containing about 6% antimony for strength and flatness, and 4.8 mm thick) interleaved with gas proportional tubes. The FEM is about 25.5 radiation lengths deep.

The Forward HAD calorimeter is constructed from 27 layers of steel, 5.1 cm thick, interleaved with gas proportional tubes. It is about 7.7 interaction lengths deep.

3.6 Beam-Beam Counters

Several planes of scintillation counters mounted on the front of the forward and backward electromagnetic calorimeters, the beam-beam counters (BBCs) [14, Section 3.5], provide a minimum-bias trigger indicating that some $\overline{\mathrm{p}}\mathrm{p}$ interaction took place during a given $\overline{\mathrm{p}}\mathrm{p}$ bunch crossing. The minimum-bias trigger is so named because it triggers on almost any inelastic $\overline{\mathrm{p}}\mathrm{p}$ interaction without otherwise requiring any specific physics signals to be present.

The BBCs are used as CDF's primary luminosity monitor. The BBCs cover the angular region 0.32° to 4.47° ($3.24 \leq |\eta| \leq 5.90$). The luminosity monitor and minimum-bias trigger require at least one hit in each plane for both sets of BBCs, requiring a 15 ns coincidence around the beam crossing time, thereby ensuring a $\overline{\mathrm{p}}\mathrm{p}$ scattering occurred.

The BBC timing resolution is less than 200 ps.

The BBCs monitor luminosity by "counting" $\overline{\mathrm{p}}\mathrm{p}$ bunch crossings in which an inelastic interaction occurs. (Strictly speaking, the BBCs count nondiffractive $\overline{\mathrm{p}}\mathrm{p}$ interactions.) The total rate of crossings $f_0 \times B$ (as defined in Section 3.1) is known: 286 kHz. One important detail is that the BBC does not count *multiple* interactions. A crossing in which more than one $\overline{\mathrm{p}}\mathrm{p}$ interaction occurs is counted the same as a crossing in which exactly one $\overline{\mathrm{p}}\mathrm{p}$ interaction occurs. Thus, at high luminosities, the BBC rate underestimates the luminosity unless the probability of multiple interactions is accounted for. Thus, the luminosity measured at CDF is calculated as [23]:

$$\mathcal{L} = -\frac{f_0 \times B}{\sigma_{\mathrm{BBC}}} \ln\left(1 - \frac{R_{\mathrm{BBC}}}{f_0 \times B}\right),$$

where R_{BBC} is the rate (in Hz) of BBC coincidence and σ_{BBC} is the $\overline{\mathrm{p}}\mathrm{p}$ inelastic cross section, as measured by the BBC, with a value of 51.2 ± 1.7 mb. The details of this calculation are beyond the scope of this dissertation, but the $\overline{\mathrm{p}}\mathrm{p}$ interaction cross section is measured in [24, 25, 26], and these measurements result in the calculation of σ_{BBC} [27].

Looking at minimum-bias data samples (i.e., with all events passing the minimum bias trigger written to tape), only a small fraction of minimum-bias events deposit large amounts of transverse energy in the calorimeters. Since multiple $\overline{\mathrm{p}}\mathrm{p}$ interactions during a crossing are independent, it follows that most high-E_T events which contain additional interactions add only a small contribution to that E_T, on average, from the additional interactions.

As $R_{\mathrm{BBC}}/(f_0 \times B)$ approaches unity, small uncertainties in the BBC rate cause large uncertainties in the luminosity measurement. Fortunately, during Run 1a, the BBC did not approach saturation. For the range of luminosities encountered during Run 1a (see section 3.1), the ratio $R_{\mathrm{BBC}}/(f_0 \times B)$ ranged from 0.2–0.8 [23, See Table 1].

Since at lower luminosities several sequential crossings may occur with no $\overline{\mathrm{p}}\mathrm{p}$ interaction, requiring a BBC coincidence allows CDF to process only those crossings which include at least one $\overline{\mathrm{p}}\mathrm{p}$ interaction.

3.7 The Trigger and DAQ

With six proton bunches and six antiproton bunches in the Tevatron, $\overline{\text{p}}\text{p}$ bunch crossings occur at CDF at the rate of about 285 kHz, with 3.5 μs between crossings. Since, depending on the luminosity, $\overline{\text{p}}\text{p}$ interactions occur about 1–4 out of every 5 crossings (see previous section), potentially 230000 events per second can be written to tape to be processed offline. Given the constraint of limited computing resources offline, however, events were written to tape at a rate of 5–7 Hz. Thus, CDF needs a trigger system with a rejection factor of about 40000 to 1 so that only those events deemed most interesting are saved to be processed offline. Trigger criteria are based on patterns of detector response and are discussed in Section 3.7.2.

Digitizing and reading out detector information takes a very long time ($\approx$ 30 ms) compared with the time between beam crossings. Therefore, is is *highly* desirable to reject as many events as possible before fully reading all detector information. In fact, it is desirable to reject as many events as possible during the time between $\overline{\text{p}}\text{p}$ crossings in order to keep the detector "live" as much as possible. The trigger system is considered live (dead) when it is able (unable) to trigger on a beam-crossing. Livetime (deadtime) is the fraction of the time that beam is delivered to CDF during which CDF's trigger is able (unable) to process a $\overline{\text{p}}\text{p}$ crossing. Deadtime wastes delivered luminosity.

To reduce deadtime, the CDF trigger system is divided into three trigger levels. The first trigger level is designed to make a decision between $\overline{\text{p}}\text{p}$ crossings, i.e., in less than 3.5 μs. The Level 1 trigger thereby incurs no deadtime, but is only able to make trigger decisions based on simple detector quantities with no time to group detector information in order to make a more informed decision. Only some events can be rejected in such a simple way. If the Level 1 trigger does not accept one crossing, it is immediately ready to process the next crossing and no livetime is lost. The Level 1 trigger is able to process a crossing quickly due to using fast analog electronics. If the Level 1 trigger cannot reject a crossing, the fast analog information is passed to the Level 2 trigger for processing. The Level 1 trigger rejects more than 99 out of 100 crossings, passing events at a rate of a few kHz to the Level 2 trigger.

The Level 2 trigger takes about 25–35 μs to make a decision; thus, the next 7–10 crossings are ignored while the Level 2 trigger is processing one crossing, and the experiment incurs deadtime. When the Level 2 trigger rejects a crossing, the Level 1 trigger can process the next available crossing. When the Level 2 trigger accepts a crossing, the full detector information is digitized, read out, and passed to the third trigger level for processing. The DAQ takes about 3 ms to digitize and read out the detector information; almost one thousand crossings are ignored while this happens. The Level 3 trigger processes up to 48 events in parallel, taking about 1–2 seconds for each, and adds no deadtime at a bandwidth up to about 20 Hz. Events passing the Level 3 trigger are written to 8 mm tape for later offline analysis. Unlike the Level 1 and Level 2 trigger systems, the Level 3 trigger is implemented in FORTRAN software running on commercially available computer processors, not custom hardware.

The trigger system successfully maintained an average of about 85% livetime during Run 1a.

Event flow is controlled by the data acquisition system (DAQ) which communicates with each trigger level. The trigger levels are discussed in more detail in the following sections; appropriate parts of the DAQ are discussed along with each trigger level.

3.7.1 Data Acquisition

The data acquisition system (DAQ) controls two parallel data paths and is shown schematically in Figure 3.11. The Level 1 and Level 2 triggers use fast analog outputs from a subset of the CDF detector while the Level 3 trigger uses fully digitized information from each detector system. The Buffer Manager (BFM) coordinates the actions of the entire DAQ; the hardware trigger system (Level 1 and Level 2 plus FRED, the front end readout and decision board, described below) is supervised (for the BFM) by the Trigger Supervisor (TS).

Timing in the DAQ is controlled by the Master Clock, which picks up a timing marker from the Tevatron RF system. The Master Clock then forwards this beam-crossing signal through the TS to the hardware trigger system and the front-end electronics. This is how the front-end electronics sample the detector at an appropriate

time with respect to the beam crossing. The front-end electronics encompass about 120000 signals from the detector hardware. These signals pass through "scanners," a number of custom-designed boards which read digital information from the front-end electronics and hold the information in a buffer until the scanner is told (by the TS) to reuse that buffer. The scanners which pull information from the front-end electronics have four buffers each, so fully digitized events can be processed in parallel. The event data are read out of the scanner buffers by the next board in the chain to process the event data.

I describe the Level 1 and Level 2 triggers separately. The final Level 1 decision is made by and communicated to the TS from the Final Decision Logic boards (FRED). FRED collects one bit for each Level 1 trigger component and uses a lookup table to make the decision. The Level 1 and Level 2 triggers communicate with the TS through FRED, and FRED also generates the Live/Dead gates for luminosity determination. When the Level 1 trigger accepts an event, FRED sends this information to the TS, which then inhibits the front-end electronics from clearing the sample-and-hold circuits. When the Level 1 trigger rejects an event, the TS sends a signal to the front-end electronics to clear the sample-and-hold circuits.

When the Level 2 decision is sent to the TS (through FRED), the TS signals the front-end scanners to digitize and read the event out into one of the scanner buffers (if a Level 2 accept) or to clear the sample-and-hold circuits (if a Level 2 reject). The scan requires about 3 ms, and when complete, the trigger hardware is ready to consider the next event (as long as the scanners have a spare buffer). The detector is now live again.

Once an event passes Level 1 and Level 2 and the scanners have digitized the detector information, the BFM instructs one of two event builders (EVB) to pull the event information from the many scanner buffers (one buffer for each scanner), reformat the event information into a standard format recognized by the event reconstruction FORTRAN code, and push the event into a Level 3 event buffer. Each event builder is a custom hardware board; the experiment had one additional spare board. Only three boards were built. The event builders require about 30 ms to completely process an event. The two event builders processed events in parallel, except they can not pull

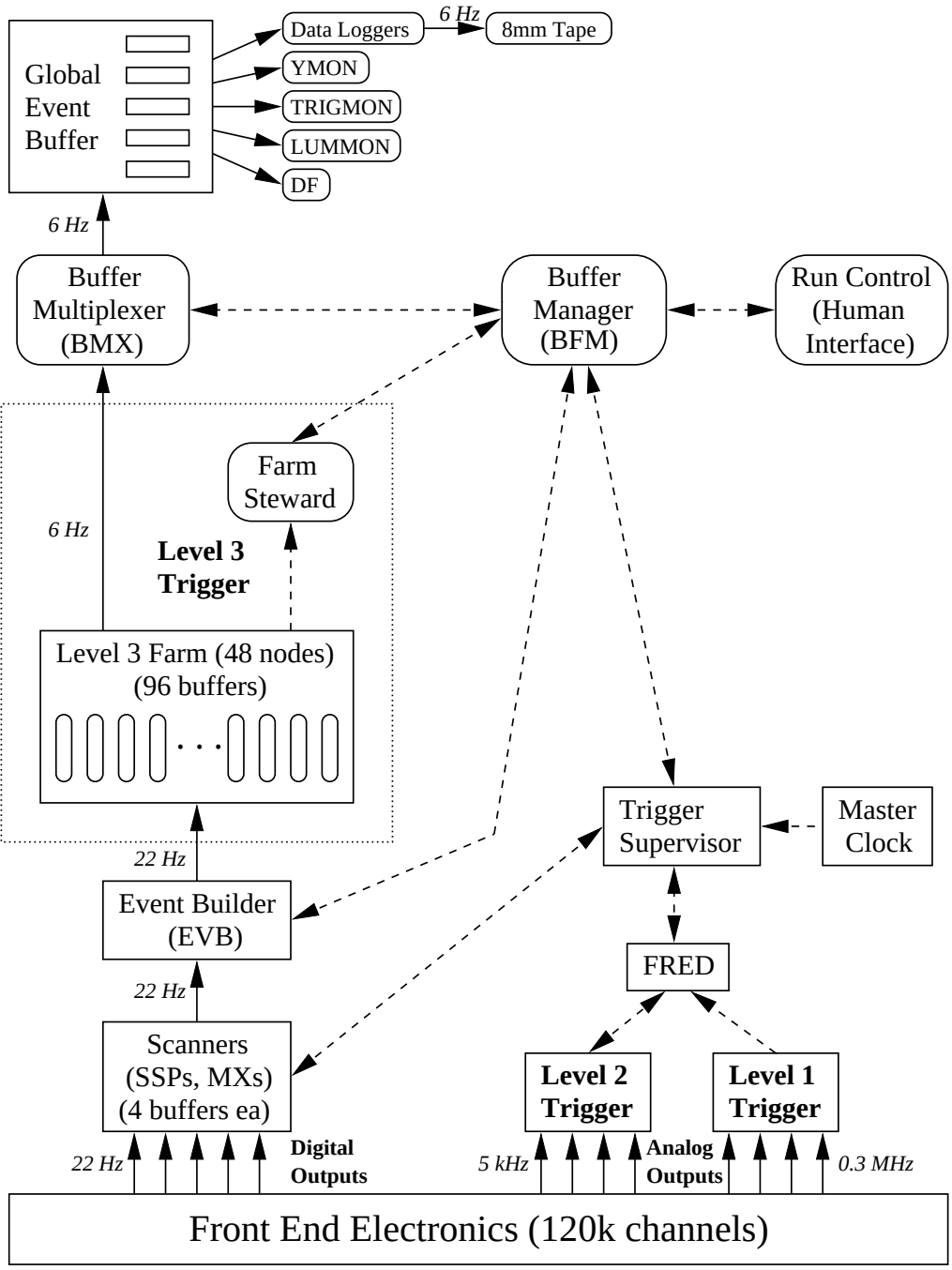


Figure 3.11: Schematic of the data acquisition system. Data paths are shown by solid arrows and communication paths are shown by dotted arrows. This schematic shows five data consumers: the data loggers which write events to 8 mm tape, three data quality monitor processes, and the event display process, DF.

event information from the scanner buffers simultaneously; they share a data path to the scanners. The long time required for an event builder to process an event proved to be a bottleneck during Run 1a, limiting the total event rate through the event builders to about 20 Hz.

When a Level 3 event is accepted, the BFM instructs the Buffer Multiplexer (BMX) to pull the event into a global event buffer. The various data consumers (including the data loggers which write events to 8 mm tape) read events out of the global event buffers.

3.7.2 Trigger Tables

A trigger is a collection of physics requirements designed to look for a specific kind of physics event. The sophistication of the physics requirements depends on the trigger level, as explained earlier. While taking data, many triggers are tested at once (in each trigger level) to support diverse physics analyses. With such a large number of individual triggers (about 10 triggers at Level 1, 40 at Level 2, and 70 at Level 3), some order is necessary to keep track of what parameters control data taking. Thus, all necessary information on the requirements for the triggers is collected in one text file, a "trigger table," which defines the kinds of collisions recorded to tape during a run.

A run is a collection of events taken with the same trigger table—runs are ended when a store ends, when the run gets very long (close to a million events), or when a part of the DAQ or detector hardware fails. An exception to this general rule is "special runs," which are usually used to study some aspect of the detector or trigger. Events from special runs are not included in data samples for physics analysis and are not described here.

Many parameters of the hardware triggers and the Level 3 triggers are related to physics, and adjustable between runs. For example, many thresholds can be changed—if a trigger's rate is too high when requiring 20 GeV jet clusters, this cut can be changed to require 25 GeV jet clusters. In addition, you may want to change which cuts are applied to the event data, adding or removing a track requirement, for example, or changing the manner in which tracking is done. The list of cuts and thresholds for the

entire trigger system is stored in the trigger table.

At the beginning of a run, the trigger table is parsed; appropriate constants from the trigger table and from databases referenced in the trigger table are downloaded to the scanners and triggers. Additionally, calibration constants (described in the next section) are downloaded to the scanners and triggers.

During a run, all downloaded parameters, cuts, and calibration constants are held constant. The sole exception to this rule is that dynamic prescales, described in Section 3.7.6, may change during a run. Between runs, the trigger table can change and calibration constants are updated. Thus, each run uses known and documented cuts and filters. Although which trigger table is used can change between sequential runs, most triggers did not change during Run 1a.

3.7.3 Calibration Constants

The conditions under which the CDF detector took data changed considerably over the course of Run 1a. For example, the temperature and air pressure changed, affecting the gas calorimeters and some other systems. CDF's timing (in digital electronics) relative to the Tevatron varied by a few ns over months, affecting tracking efficiency. Additionally, detector components and electronics aged, broke, and were fixed. Detector aging most commonly caused certain detectors to slowly become less efficient.

For all of these reasons, the entire CDF detector is calibrated periodically, usually daily. The results of calibration are stored in a database, and the constants and tables (of dead calorimeter towers, for example) are downloaded to the trigger system before each run. In this manner, the trigger properly takes into account timing drift, dead channels, temporarily disabled detector components, and calibration shifts.

Many of the various trigger components use lookup tables to make a decision. These lookup tables are also stored in the database, allowing a lookup table to change (e.g., be corrected or improved) between runs.

3.7.4 Level 1 Trigger

The Level 1 trigger is required to make a decision in less than one crossing (3.5 μs) in order to be without deadtime. Thus, it can only select events based on simple detector quantities. The Level 1 trigger is implemented in custom hardware and its decision is based on calorimetry and muon information. The Level 1 trigger contains several separate components working in parallel; each component sends its decision to FRED which then computes the Level 1 trigger decision. The Level 1 trigger reduces the event rate from about 230 kHz to a few kHz.

The Level 1 and Level 2 triggers, when looking at calorimetry, use "trigger towers," which are one physical tower wide in ϕ (in the central) but two physical towers wide in η. In the plug and forward calorimeters, a trigger tower is three physical towers wide in ϕ and two in η; thus, all trigger towers are 15° wide in ϕ and approximately $\Delta\eta = 0.2$ wide in η. No tracking information is available for the Level 1 trigger.

The Level 1 general calorimetry trigger (used for all physics requiring jets) simply requires any single trigger tower over an E_T threshold. This threshold can be set independently for each separate calorimeter component. For Run 1a, this trigger requires a CEM trigger tower over 6 GeV, a CHA or WHA tower over 8 GeV, or a PEM or FEM tower over 8 GeV. The Level 1 calorimeter trigger threshold for the PHA and FHA varied between runs during Run 1a, as described in Section 4.4.1.

Level 1 also contains a prescaled calorimeter trigger with lower E_T thresholds; effectively, a prescaled trigger is one which "sees" only a fraction of the luminosity seen by an unprescaled trigger. The Level 1 and Level 2 triggers apply prescales differently, but the end result is similar. At Level 1, every 20th $\overline{\mathrm{p}}\mathrm{p}$ crossing, the trigger hardware applies the lower thresholds *instead of* the thresholds listed above. Thus, the two Level 1 calorimetry triggers are mutually exclusive.

The prescaled calorimetry trigger is designed for QCD and trigger studies; the higher tower thresholds of the standard calorimetry trigger bias lower energy QCD dijets by accepting only those events where a jet deposited its energy in a narrow cone. By taking some events with lower thresholds, this bias and the efficiency of the higher thresholds

can be studied.

Events passing the Level 1 calorimetry triggers are events with high transverse E_T and thus are events with a $\overline{\mathrm{p}}\mathrm{p}$ hard scattering. This trigger does not accept diffractive $\overline{\mathrm{p}}\mathrm{p}$ interactions.

During part of Run 1a, the Level 1 trigger also required a BBC coincidence. This requirement was removed later in Run 1a after the average luminosity rose and the small inefficiency of the BBC requirement became more important than the rejection it provided.

The Level 1 trigger generated no deadtime during Run 1a.

3.7.5 Level 2 Trigger

The Level 2 trigger uses custom designed hardware. Level 2 takes about 25–35 μs to process an event, thus, the next 7–10 crossings are ignored. During Run 1a, Level 2 incurred a total deadtime of about a few percent. Rejection in Level 2 is set to maintain a maximum Level 2 accept rate of about 20 Hz, the maximum rate the event builders (described in Section 3.7.1) can handle.

Since the Level 2 trigger uses more programmable hardware, it can cut against physics objects rather than just detector responses. The Level 2 trigger hardware identifies jet clusters, calculates the total transverse energy and $\not\!\!E_T$ of the event, finds two-dimensional tracks (r-ϕ) from the Central Fast Tracker (CFT), and matches CFT tracks to clusters.

Level 2 calorimeter clusters are identified by the cluster finder, custom-built hardware, which looks for trigger towers over the seed threshold (3 GeV) and then adds adjacent trigger towers over the shoulder threshold (1 GeV) to the cluster. At this stage, the tower E_T information is still analog. Only after cluster finding is complete is the cluster information (E_T, average η and ϕ, and other information) digitized.

The CFT [28] finds and measures high-P_T tracks with high efficiency. The transverse momentum (P_T, defined in Section 3.3.2) resolution is also good: $\delta P_T/{P_T}^2 = 3.5\%$. The CFT is a hardware track finder which uses the axial CTC layers to find two-dimensional tracks (r-ϕ). CFT tracks are binned by P_T into eight bins each for positively and

negatively charged tracks (16 P_T bins total), and contain no z information.

The CFT normally requires about 8 μs to complete its list of CTC tracks. It counts the number of wires with a prompt hit (a hit with a small drift time) in a "road" (an expected pattern of prompt hits) based on a prompt hit in the outer CTC superlayer. Each sense wire in superlayer 8 is associated with 32 roads—8 P_T bins plus two azimuthal bins for each curvature direction (i.e., positive or negative charge). A road, essentially, is a list of wires in each axial superlayer which should possess a prompt hit for a given signed transverse momentum and prompt hit in superlayer eight. That is, it is a lookup table.

When the CFT follows a road from superlayer eight inward, it looks for track hits only in the road. Therefore, the CFT only finds a track efficiently in an appropriate P_T bin—otherwise the track's curvature will take it outside the road in which the CFT is looking for hits. The CFT track-finding efficiency for a given P_T bin is therefore dependent on track P_T. The P_T bin boundaries are defined by where the CFT is 90% efficient at finding tracks; the bin boundaries are at P_T = 3.0, 3.7, 4.8, 6.0, 9.2, 13.0, 16.7, and 25.0 GeV/c. The sixteen P_T bins are numbered 0–7 separately for positively and negatively charged tracks.

When the cluster finder and CFT have completed processing, a processor board in the Level 2 trigger uses a lookup table to extrapolate CFT tracks to the calorimeter, in ϕ. The CFT track lookup table uses two indices: the wire number of the track's prompt hit in superlayer 8, and the P_T bin (including charge) of the track. The output of the lookup table is a ϕ index 0–23 which is compared to the ϕ index of Level 2 clusters to see if the track is associated with that calorimeter cluster.

The manner in which CFT tracks are matched to Level 2 clusters is relevant to a bias in the data which is discussed later.

3.7.6 Trigger Prescaling

The Level 2 trigger hardware allows individual Level 2 triggers to be prescaled, but remember that Level 2 prescales are applied in a different manner than Level 1 prescales. If a given trigger, at Level 2, is prescaled by 20, then only every 20th event passing

that trigger's selection requirements is actually accepted by the Level 2 trigger. The previous 19 passing events are considered to fail the Level 2 trigger unless some other Level 2 trigger accepts the event. Sometimes it is preferable to prescale a trigger to reduce its trigger rate, as opposed to making the trigger cuts more stringent. This is equivalent to reducing the luminosity seen by that trigger.

Since both the event builders and Level 3 trigger can not handle an event rate much above 20 Hz without unacceptably high deadtime, many triggers are prescaled to reduce the total Level 2 accept rate to a maximum of about 20 Hz.

During the middle of Run 1a, dynamic Level 2 prescales were introduced. A dynamic prescale is one which may change *during a run.* Thus, when taking data at high luminosity, the trigger rate for many triggers is reduced, then restored during the run as the luminosity drops with time. Dynamic prescales allow CDF to keep the detector live for rare interactions at high instantaneous luminosity while still supporting a diverse physics program.

For dynamic prescales, the current total number of events passing the trigger and the number of events passing the trigger and prescale are saved in each event. Thus, even though the prescales are subject to change during a run, the integrated luminosity seen by each trigger is calculable.

3.7.7 Level 3 Trigger

The Level 3 trigger is implemented using commercial computer processors and standard CDF offline event analysis code, written in FORTRAN, with only a few changes. Most of the event reconstruction and filtering code used in Level 3 is the same as that used offline, but certain code modules differ to allow Level 3 to process events more quickly (if less completely) than the offline system does. For offline processing, one desires the best possible reconstruction of the raw data into physics quantities, but the Level 3 trigger must process events quickly enough to prevent deadtime. For this reason, Level 3 does not process all detector data—for example, no SVX track reconstruction is performed online. The Level 3 trigger incurred little deadtime. Each event requires about 1–2 seconds to process. Events in Level 3 are processed in parallel and are buffered.

The FORTRAN analysis code in Level 3 differs from the offline code for two additional reasons. To ensure maximum stability, the Level 3 FORTRAN analysis code was frozen early in the run and updated only when a critical bug was fixed. Additionally, since only a small amount of the Level 3 trigger internal processing can be saved, certain algorithms are modified to be predictable from the information available offline. For instance, at Level 3, the event vertex is always taken as z=0. Additional minimum bias events contribute additional z vertices in the event; choosing z=0 allows the Level 3 trigger decision to be free of bias from these additional vertices.

Some FORTRAN code modules in Level 3 are abbreviated compared with the code used offline. Tracking, for example, is simplified in Level 3. Full 3-dimensional track reconstruction is done in Level 3, but the reconstruction algorithm includes a bias toward tracks coming from the beamline and is quicker (and less efficient) than the full offline tracking is. Level 3 uses calibration constants available at the time of the run; offline processing, run days later, uses the best available calibration constants which are not available to Level 3.

The Level 3 farm consists of six Silicon Graphics Power Servers, each with eight RS4000 CPUs. The six Power Servers (1 SGI 280, 1 SGI 380, 4 SGI 480's) contain a total CPU power of 1100 MIPS (as measured using CDF analysis code), where 1 MIPS is defined as the processing power of a VAX 11/780. This can be compared with the roughly 37 MIPS of the '88-'89 run Level 3 farm.

Each CPU on each Power Server has one Level 3 farm process running on it. Each farm process has two buffers so that one buffer can be filled or emptied by the DAQ while it processes the other.

Level 3 event processing consists of two parts: event reconstruction and trigger path processing. Event reconstruction processes the raw data into physics quantities such as tracks and jets; a trigger path contains a list of analysis and filtering code modules to be executed as well as parameters for those code modules. Filtering modules are FORTRAN code modules which return a pass/fail after applying cuts to an event. Event reconstruction and tau filters are described in greater detail in Section 4.4.3. The Level 3 trigger, using separate but often overlapping trigger paths, searches for

electrons, muons, taus, photons, jets, heavy flavor (QCD jets from c and b quarks) and exotic physics, or searches for physics beyond the Standard Model.

Each trigger in Level 3 is implemented as one path which is independent of every other trigger path. Thus, to understand one Level 3 trigger, you need to understand event reconstruction and that one trigger path only; Level 3 trigger paths do not interact. Once event reconstruction is complete, each trigger path is executed. An event which passes *any* Level 3 trigger path is accepted by the Level 3 trigger and written to 8 mm tape.

The Farm Steward is a process which controls all of the other Level 3 processes and communicates with the BFM as shown in Figure 3.11.

Chapter 4

Event Selection

This was the most unkindest cut of all.
— Julius Caesar, Act 3

In this and the next chapter, I present the path events follow from the data acquisition system (or monte carlo simulator for the simulated electroweak event samples) into the event samples used to measure the tau charge asymmetry. First I describe what $W \to \tau\nu$ and background events look like in the CDF detector, motivating the discussion which follows where I present the event selection (trigger cuts and analysis cuts) which result in the final event samples.

To measure the tau charge asymmetry as a function of η, I must first measure the relative differential cross section for the processes $\overline{\mathrm{p}}\mathrm{p} \to W^{\pm} + X; W^{\pm} \to \tau^{\pm}\nu$. I do not need to know the absolute cross sections, only the relative cross sections for positive and negative taus. Assuming the same efficiency for identifying positive and negative taus (from $\overline{\mathrm{p}}\mathrm{p} \to W + X$), I only need to count the number of positive and negative taus in each η bin. If the efficiency for positive and negative taus is *not* the same, but the efficiency difference is measurable, the measured counts of positive and negative taus can be corrected to the assumption of equal efficiency.

Balancing the competing needs to have a pure sample of taus (few background events in the sample) and to have as large a sample of taus as possible for statistical accuracy, the final event selection contains more than a negligible number of background events. These background events must be subtracted statistically from each η bin in order to measure the tau charge asymmetry. To subtract the backgrounds statistically, I must first model the background processes.

After event selection, I have one "real" event sample for tau events (including some

background contamination) and one simulated event sample for each significant background process. The distribution of positive and negative taus versus η is determined once the background samples are properly normalized, as described in Chapter 6.

Event selection comprises several steps. First, the events must have been accepted by the data acquisition system (or simulated) and written to tape or disk. Since real events depend upon calibration constants, these events are reprocessed before continuing event selection; the most precise calibration possible was determined only after Run 1a ended. Event selection in the triggers is "loose"; efficiency for accepting desired events is valued above the purity of events passing the triggers. Therefore, following event reprocessing, the data stream which passed the triggers is filtered to a smaller (purer) and more manageable sample. After preselection, I apply a series of analysis event selection cuts to increase the purity of the tau sample as much as possible without sacrificing too much efficiency.

Event selection for the tau event sample is described in detail in Sections 4.4 through 4.7. The simulated event samples pass through a similar process and are described in the next chapter.

4.1 The 1992 Tevatron 1a Run

This analysis uses data only from the 1992–1993 Run 1a of the Tevatron, operating at a center-of-mass energy of 1.8 TeV. During Run 1a, CDF collected a total of 20.6 pb^{-1} of data (19.6 pb^{-1} considered to be good data) with an uncertainty on the integrated luminosity of 3.5% [27]. The Tevatron delivered about 30 pb^{-1} to CDF during Run 1a, so the total experimental livetime was about 70%, including both DAQ deadtime (when the electronics reads out an event) and time when beam was delivered to CDF and the detector was not working. A total of about 16 million events were written to tape. Of these events, about 2 million were written to the XOX1_3P stream used in this analysis.

4.2 $W \rightarrow \tau\nu$ Events at CDF

Events containing hadronically decaying taus from $W \rightarrow \tau\nu$ at CDF present several distinguishing characteristics, some of which are discussed in Sections 2.8 and 2.10. To summarize, a jet from a high-E_T tau deposits energy in a narrow cone in the calorimeter (thus, in only a few calorimeter towers), and only a few charged tracks are associated with this calorimeter cluster. The invariant mass of the charged tracks plus neutral energy identified in the CES is less than the tau mass since the tau neutrino is not included in the measurement. Also, the opening angle of the tau decay daughters decreases as the tau momentum increases.

Taus from W decay are generally isolated in the detector with little activity from the underlying event. The only additional high-E_T hadronic jets present in W events are those recoiling against the W; the leading order process to create a W in a $\overline{\mathrm{p}}\mathrm{p}$ collider only contains one high-E_T jet—from the lepton. Only next-to-leading order and more rare interactions generate $W \rightarrow \tau\nu$ events containing more than one high-E_T hadronic cluster. An event selection that removes events with more than one high-E_T cluster, therefore, efficiently selects taus and removes many background events.

A tau always decays to an odd number of charged particles, but the underlying event can sometimes add a track "beneath" the tau cluster. Also, some charged tracks are not identified: Charged tracks very close to one another can be identified as just one track due to finite position resolution in the CTC. That is, the track reconstruction code is not perfectly efficient at identifying tracks, especially for pairs of tracks which share a large number of hits in the CTC. Still, a jet from a hadronically decaying tau contains only a few identified charged tracks.

The presence of neutrinos is inferred by looking for $\not\!\!E_T$, as discussed in Section 3.3.2. With $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ processes, only one neutrino is produced and the $\not\!\!E_T$ signal is clear and unmistakable. Two neutrinos are produced for $W \rightarrow \tau\nu$ when the tau decays hadronically. Due to the W polarization and corresponding tau polarization, these neutrinos tend to travel in opposite directions. Thus, the neutrino from tau decay usually partially cancels, in transverse momentum, the energetic neutrino involved in

the tau production from W decay. Although $\not\!\!E_T$ is generally smaller for $W \to \tau\nu$ events, it is still a significant signal.

Figure 4.1 shows a typical 3-prong tau candidate. The upper plot is a "lego" plot showing the calorimeter E_T deposits vs. η and ϕ. Dark hatches show electromagnetic E_T and lighter hatching shows hadronic E_T. The energy of the tau candidate cluster is deposited in a narrow cone, predominantly in one tower. As mentioned in Section 2.8, the rest of the event in a W decay should be clean; it is.

The lower plot presents 2-dimensional tracking information (r-ϕ, see Section 3.3) as well as calorimeter transverse energy deposits versus ϕ (summed over η). The calorimeter energy is deposited transverse to the $\overline{\mathrm{p}}\mathrm{p}$ beam. The innermost two dark circles are the inner and outer radii of the CTC. The box around the calorimetry deposits is the cursor showing which calorimeter wedge is displayed in the sideview. Track segments which end at the outer radius of the CTC are central tracks which pass through all nine superlayers. Short track segments represent tracks which enter the plug or forward parts of the detector, and thus do not traverse all superlayers of the CTC.

The side view of the CTC display requires some description. The side view shows a 2-dimensional $r - z$ view of a ϕ slice one tower wide and passing through the beamline. Only the central calorimeters, including the Wall HAD and the strip chambers embedded in the CEM, appear in this view; see Figure 3.7 for comparison. The wedge displayed in this figure is wedge number 5, as indicated at the bottom of the side view. Wedges are numbered 0–23 starting at $\phi = 0$, which is just above the horizontal on the right side of the CTC display. Comparison with the lego plot shows positive η is at the top of the side view, negative η at the bottom. The tau candidate charged tracks' polar angle is about 51°.

In the side view, tracks are (somewhat incorrectly) shown as straight lines, having passed through the solenoid and thus being outside the magnetic field. (In the side view, charged tracks follow a sinusoidal path, a consequence of curvature in ϕ, although tracks with $P_T \gtrsim 5$ GeV do not change direction very much.) The first box the tracks pass through, with a wavy line down the middle, represents the CES (see Section 3.5.2); CES energy deposits are represented by the amplitude of the wavy line. Hadronic particles,

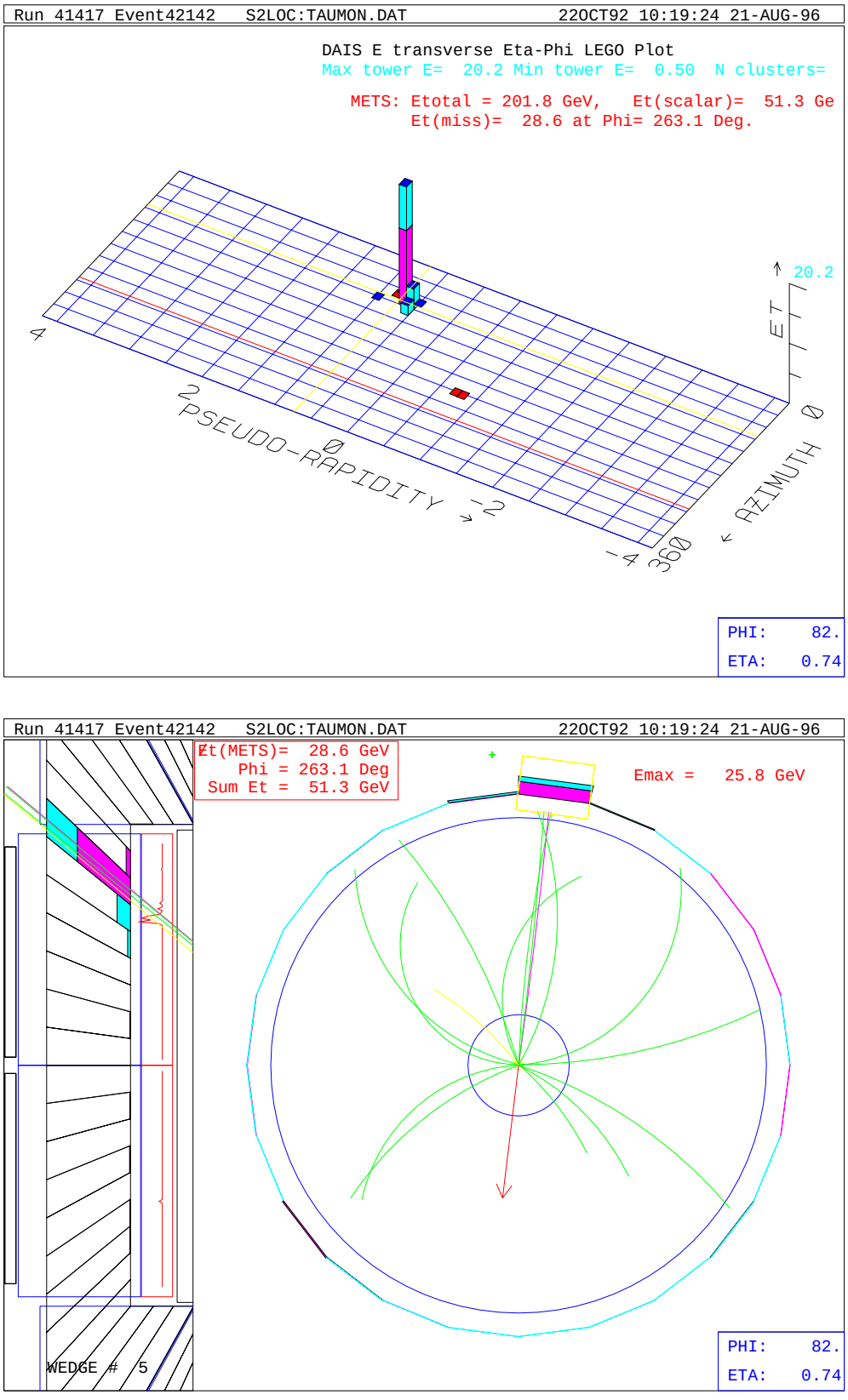


Figure 4.1: A 3-prong tau candidate

photons, and electrons deposit energy in the CES, although generally photons and electrons leave a much larger energy deposit than hadronic particles. Following the CES, the EM and HAD E_T deposits in the selected wedge are shown for each central tower in that wedge. As with the lego plot, darker hatching represents EM E_T and lighter hatching represents HAD E_T.

In the CTC display, three tracks in a narrow cone project to the central tower (in ϕ) of the tau candidate cluster. Looking carefully at the curvature of the three tracks, the total charge is $+1$. (In this display, positively charged tracks travel counter-clockwise as they leave the event vertex.) The $\not\!\!E_T$ (arrow) points almost directly opposite the tau candidate cluster and there is no other activity—either tracks or calorimetry—opposite the tau candidate cluster. In the side view, the three tracks project (in η) to the correct tower. The large two-peaked CES deposit suggests a π^0 (or other neutral hadron) decayed to two photons. Thus, this event is likely a $\tau^+ \to \pi^+\pi^+\pi^-\pi^0\overline{\nu}_\tau$ event.

The double-handful of low momentum tracks visible in the CTC display represent the underlying event. All of these tracks have less than about 2 GeV of transverse momentum, several of the tracks much less.

Figure 4.2 shows a 1-prong tau candidate. Again, the tau candidate cluster deposits energy in just a few towers in the lego plot. In the CTC display, the $\not\!\!E_T$ points opposite the tau cluster and the track extrapolates to the cluster in both η and ϕ. There is no significant activity opposite the tau candidate. The CES display suggests a π^0 deposited the EM E_T while the charged track deposited the HAD E_T. When the EM E_T is deposited by an electron, a charged track extrapolates to the CES energy deposit; instead, the charged track extrapolates to the tower containing nearly pure HAD E_T. The one peak in the CES (with no track extrapolating to it) suggests either a very energetic π^0 (thus decaying into nearly collinear photons) or $\pi^0 \to \gamma\gamma$ where the photons are separated in ϕ but not in η. The few additional tracks in this event are all low momentum. Thus, this event is likely to be $\tau^- \to \nu_\tau\rho^-$ with $\rho^- \to \pi^-\pi^0$.

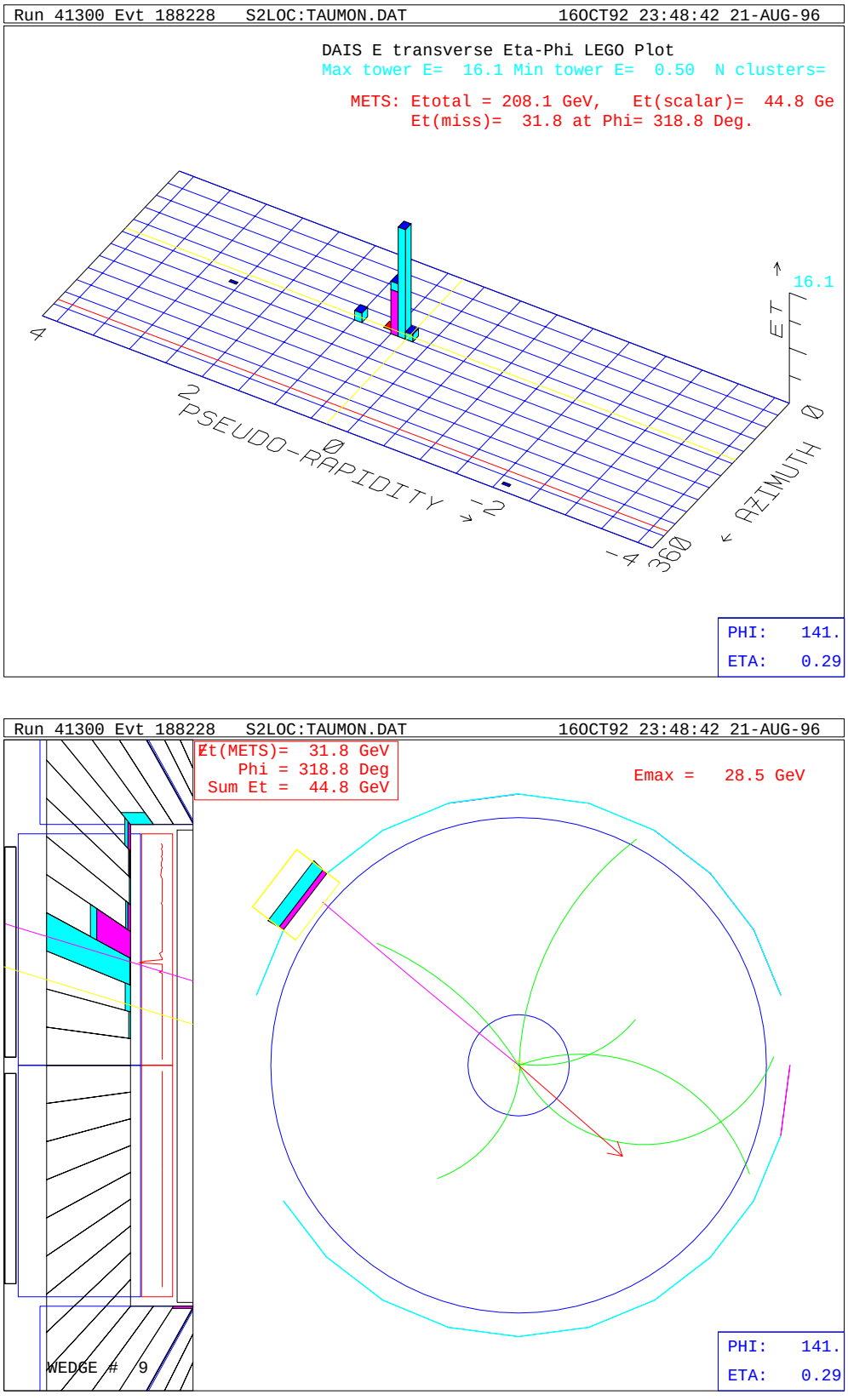


Figure 4.2: A 1-prong tau candidate

4.3 Background Events in $W \rightarrow \tau\nu$ Data Sample

The backgrounds to a $W \rightarrow \tau\nu$ analysis are summarized in Section 2.10. In short, both electroweak and QCD jet processes contribute a significant number of background events. Electroweak backgrounds are reduced to a small level through event selection cuts; these event selection cuts are discussed in Section 4.7.4. QCD background processes cannot be reduced to a similarly small level without greatly reducing the efficiency for identifying taus, so the jet background is the most significant background process in this analysis.

Electroweak processes contributing to backgrounds include all leptonic W and Z processes other than $W \rightarrow \tau\nu$. The most prominent electroweak background processes are $W \rightarrow e\nu$, and $Z \rightarrow \mu\mu$ + jets. An electron leaves a "tau-like" cluster, fitting closely all the characteristics of a tau cluster as described in the previous section. In addition, the one neutrino present in $W \rightarrow e\nu$ has high E_T, 40 GeV if the W is produced with no transverse momentum. Since electron clusters are distinguishable from hadronic clusters, however, electrons are efficiently removed from the hadronic decay tau event sample.

Normally, Z processes do not generate significant $\not\!\!E_T$. When a Z decays into two charged leptons, no neutrinos are present unless the Z decays into two taus. In this case, the tau helicities are correlated, so the neutrinos from the tau decays are correlated and tend to cancel one another. Again, there is little $\not\!\!E_T$. The process $Z \rightarrow ee$ is rejected similarly to $W \rightarrow e\nu$.

The process $Z \rightarrow \mu\mu$ plus jets, on the other hand, appears similar to a $W \rightarrow \tau\nu$ event when the Z recoils against a gluon or quark jet and when the jet from that parton hadronizes in a manner similar to a tau jet. Muons do not deposit a significant amount of energy in the calorimeters, so the hadronic jet will be back-to-back with apparent $\not\!\!E_T$. Fortunately, events containing muons are easily rejected by looking for the high-P_T track muons leave in the CTC. Such muon rejection can fail, for example, when the Z is produced with a large longitudinal boost: one of the muons is forward (that is, it does not pass through all of the CTC superlayers) and the other muon is soft

enough in P_T to fail high-P_T track selection. If the muon rejection cut is too tight, the efficiency for finding taus is unacceptably low. Still, rejecting events containing muons by looking for a high-P_T track successfully reduces this background to a small level.

In summary, even the most significant electroweak backgrounds are successfully reduced to a small background contribution, easily modeled by a monte carlo simulation. The monte carlo simulation and event selection cuts for these simulated event samples are described in the next chapter.

The jet background (that is, the background from prompt hadronic jets) is not so easily reduced to a negligible level. The cross section for $\overline{\mathrm{p}}\mathrm{p} \rightarrow$ jets where at least one jet has E_T greater than 20 GeV is about 8 μb, but the cross section for $\overline{\mathrm{p}}\mathrm{p} \rightarrow W \rightarrow \tau\nu$ where the tau jet has E_T greater than 20 GeV is about 0.4 nb. Although, in principle, many event selection cuts distinguish background jets from tau jets, significant background contamination of the final data sample is impossible to avoid due to the disparity in production cross sections.

The characteristics of tau jets as described in the previous section can be compared with characteristics of QCD processes that create jets from quarks and gluons. High energy prompt hadronic jets generally fragment to a larger number of tracks and spread calorimeter energy in a larger $\triangle R$ cone in η and ϕ than lower energy QCD jets do. The transverse momentum of tracks in such events can be low, since the total transverse energy and momentum of the initial parton is spread across many hadrons. Indeed, some prompt hadronic jets do not contain any charged energy at all.

Purely QCD processes rarely generate $\not\!\!E_T$; when they do, usually the $\not\!\!E_T$ arises from fluctuation in measurement of the event. The invariant mass of the particles in a prompt hadronic jet is frequently larger than that of the particles in a tau jet since much more energy is available to go into fragmentation. Requiring the tau candidate jet cluster to be isolated removes many background jets for two reasons: Prompt hadronic jets tend to fragment in a wider cone at higher energy, and QCD processes generate more underlying event particles spread throughout the event. Finally, since the transverse energy must balance, high transverse energy QCD processes usually result in two high-E_T jets nearly back-to-back in ϕ, although higher-order (and thus more rare) processes

can generate multiple jets spread throughout the detector in ϕ. Note that QCD jets are not commonly back-to-back in η except in the rest frame of the parton interaction.

To remove background jet events from the tau sample, I select events that contain large $\not\!\!E_T$ and do not contain two high-E_T clusters, especially when the clusters are back-to-back in ϕ. The tau candidate cluster must be narrow in opening angle (only a few calorimeter towers), have a small invariant mass, and be isolated. Of these event selection cuts, the one which removes the greatest percentage of jet background events is the $\not\!\!E_T$ cut.

About 5% of hadronic jets fluctuate to be narrow and pass loose tau identification cuts, but only about 1% of these jets pass tighter tau identification cuts, as shown by Cal Loomis [29, Section 5]. (Note that these tau identification cuts do not include cuts on global event quantities.) To enter the tau event sample as background, such events must also contain significant $\not\!\!E_T$; $\not\!\!E_T$ in jet background events is usually fake—not from a neutrino. For example, when one jet in an event is undetected, as by entering the calorimeter in an uninstrumented region, the E_T belonging to that jet is undetected and the event appears to contain $\not\!\!E_T$. Additionally, some hadronic jets begin showering late (deep) in the hadronic calorimeter. The energy of such jets is poorly measured and the event appear to contain $\not\!\!E_T$.

These background jet events which so closely mimic the $W \to \tau\nu$ signature are difficult to remove without losing all efficiency for identifying taus. Still, since the two processes required for a QCD process to mimic $W \to \tau\nu$ are independent, the tau event sample contains about 20% jet background (in 1- and 3-prong isolated clusters) after all event selection cuts (described in the rest of this chapter). This background is modeled using a separate sample of events selected to resemble the prompt hadronic jet background but with some event selection cuts relaxed to increase the presence of QCD jets over tau jets. This event sample and its selection are described in Section 5.2.

```
TRIGGER L1_CALORIMETER_BBC
    SELECT/VETO  LEVEL1_CALORIMETRY_PRESCALE ! Must be here!
        PARAMETER   L1_PRESCALE        = 1
    SELECT  LEVEL1_CALORIMETRY_SINGLE_TOWER
        PARAMETER WFEM_TOWER_THRESHOLD  >  8. (GeV)
        PARAMETER WPEM_TOWER_THRESHOLD  >  8. (GeV)
        PARAMETER CEM_TOWER_THRESHOLD   >  6. (GeV)
        PARAMETER EPEM_TOWER_THRESHOLD  >  8. (GeV)
        PARAMETER EFEM_TOWER_THRESHOLD  >  8. (GeV)
        PARAMETER WFHAD_TOWER_THRESHOLD >  12. (GeV)
        PARAMETER WPHAD_TOWER_THRESHOLD >  12. (GeV)
        PARAMETER CHAD_TOWER_THRESHOLD  >  8.  (GeV)
        PARAMETER EPHAD_TOWER_THRESHOLD >  12. (GeV)
        PARAMETER EFHAD_TOWER_THRESHOLD >  12. (GeV)
    SELECT BBC
```

Figure 4.3: Level 1 calorimeter trigger. See text for details.

4.4 The Trigger

The first cuts applied on the data are those cuts applied in the triggers. The data acquisition and trigger system are described in general in Section 3.7. Here, I describe specifically those triggers used for the tau event sample.

4.4.1 Level 1

The level one trigger used in this analysis is the standard "calorimeter" trigger used in all jet analyses. This trigger selects any event with a trigger tower over a threshold. Figure 4.3 presents the basic Level 1 trigger used for this analysis [30]. The trigger parameter names give the thresholds independently for each trigger tower position (central, plug, forward), location (east or west), and type (EM, HAD). Thus, the parameter `WFEM_TOWER_THRESHOLD` in Figure 4.3 refers to the West, Forward, EM calorimeter towers. Central towers are not separated into east and west. If any trigger tower anywhere in the detector is over the appropriate threshold, then the trigger accepts the event.

This trigger is efficient for $W \rightarrow \tau\nu$ events when the tau jet cluster E_T is at least 20 GeV (see Figure 4.4); the elements of this trigger are discussed below.

Two Level 1 calorimetry triggers were used during Run 1a. The primary trigger

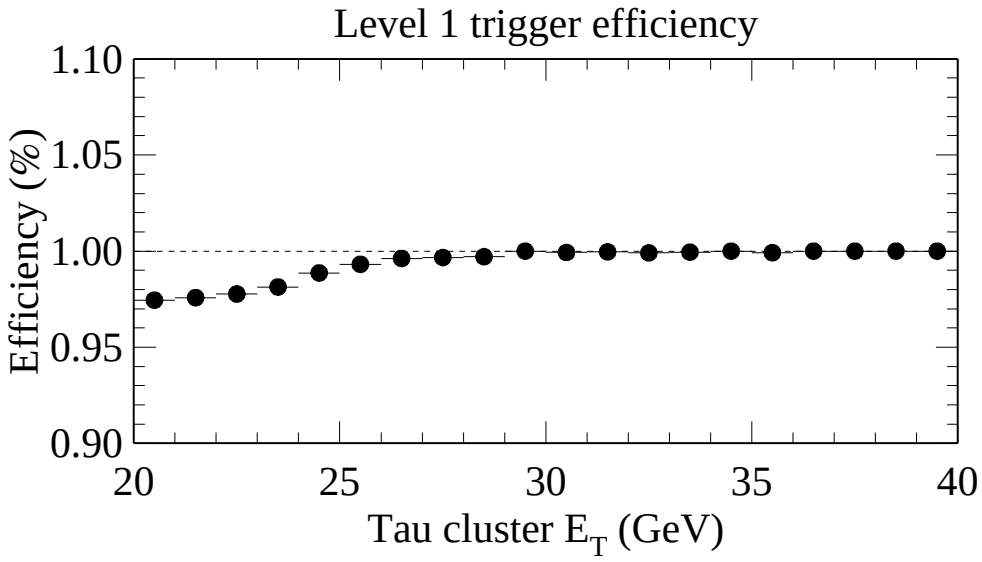


Figure 4.4: Level 1 trigger efficiency, shown for events which pass the Level 2 trigger described in the next section. The data in this plot come from a $W \rightarrow \tau\nu$ monte carlo event sample, and the Level 1 trigger efficiency, for events passing the `TAU_20_MET_20` trigger, is 99% summed over all E_T.

used is the one presented in Figure 4.3, and this trigger changed little during Run 1a. The changes to this trigger during Run 1a do not affect tau physics, as discussed below.

Since the forward and plug hadronic calorimeters present a noise problem (from particle showers of particles not coming from the collision vertex), this trigger changed several times from the version shown in Figure 4.3. Thus, the four thresholds `WFHAD`, `WPHAD`, `EFHAD`, and `EPHAD` changed from 12 to 51 to 25 GeV during the course of Run 1a. These changes do not affect tau analyses for two reasons: Tracking requirements, discussed later, force tau candidates to be central objects, and monojet selection, also discussed later, removes events with a second high-E_T jet. That is, an event with two clusters, a central cluster passing the tau identification cuts and a plug or forward cluster with one tower over 12 or 25 GeV, is rejected by requiring only one high-E_T jet cluster—the tau. A central Level 1 tau-identifying inefficiency cannot be recovered by the presence of a second jet in the plug or forward HAD calorimeters, so this change in the Level 1 triggers is of no consequence for this analysis.

At the beginning of Run 1a, the BBC coincidence shown in Figure 4.3 was required, but this requirement was removed later in Run 1a. Early in Run 1a, when the typical luminosity was small, many $\overline{\mathrm{p}}\mathrm{p}$ crossings did not contain a hard (high momentum transfer) interaction. Later, when the typical Tevatron luminosity increased, BBC coincidence occurred for most crossings. Thus, the BBC requirement was removed. The BBC is discussed in greater detail in Section 3.6.

The second Level 1 calorimeter trigger applies lower E_T thresholds to the central calorimeters, and is prescaled. (see Section 3.7.4) Since the two Level 1 calorimeter triggers are mutually exclusive by design, rejecting events which pass this second trigger needlessly discards 5% of my event sample. Thus, I simulate the higher threshold calorimeter trigger (using offline calorimeter energies) for events which pass the prescaled trigger.

The Level 1 calorimeter trigger cross section varied during Run 1a when the plug and forward HAD thresholds changed. When the BBC coincidence was required, the standard Level 1 trigger cross section changed from 300 μb to 55 μb to 65 μb when the thresholds changed from 12 GeV to 51 GeV to 25 GeV. When the BBC requirement was removed, the cross section rose slightly to 70 μb. The prescaled Level 1 calorimeter trigger, on the other hand, accepts events at a rate equivalent to a cross section of 35 μb. This equivalent cross section takes the prescale into account—the actual cross section is 20 times greater.

4.4.2 Level 2

The Level 2 tau trigger is the most complicated Level 2 trigger used during Run 1a, and so required careful study. Due to the difficulty in implementing the trigger, the first 5 pb^{-1} of data from Run 1a—about 1/4 of the full data set—cannot be used for tau physics. The text of this trigger is listed in Figure 4.5; this trigger contains six components.

The first two parameters control prescaling, which is discussed in Section 3.7.5. For much of Run 1a, this trigger was not prescaled. The `LEVEL2_PRESCALING` step is present for technical reasons. The dynamic prescale was set to a minimum of 1 and

```
TRIGGER TAU_20_MET_20_V1
    EXECUTE LEVEL2_PRESCALING
        PARAMETER   PRESCALE_FACTOR = 1
    EXECUTE DYNAMIC_PRESCALING
        PARAMETER   MINIMUM_PRESCALE_FACTOR = 1
     PARAMETER   MAXIMUM_PRESCALE_FACTOR = 100
    SELECT TAU
        PARAMETER   EM_ET_OVER_TOTAL_ET <= 1        ! Fixed cuts in
        PARAMETER   CENTRAL_TOWERS <= 3             ! Mercury-TAU PROM
        PARAMETER   ETA_MIN = 0.0                   !     "        "
        PARAMETER   ETA_MAX = 1.19                  !     "        "
        PARAMETER   ET_THRESHOLD >= 20.0 (GeV)      ! Selectable cuts
        PARAMETER   TRACK_PT >= 4.8 (GeV/c)
                  ! 90% eff at 3.0,3.7,4.8,6.0,9.2,13.0,16.7,25.0(GeV/c)
        CUT         NUMBER_OF_TAUS >= 1
    SELECT  MISSING_ET
        CUT         MISSING_ET_SQUARED >= 400. (GeV**2)
    SELECT TRL_TAU_CUT
    SELECT TWO_CLUSTER_CUT
        PARAMETER   CLUSTER1_TYPE=1            ! Highest Et Cluster
        PARAMETER   CLUSTER2_TYPE=2            ! 2nd Highest Et Cluster
        PARAMETER   CLUSTER_QUANTITY=18        ! Phiseed
        PARAMETER   COMPARE_TYPE=6             ! Less than
        PARAMETER   OPERATION=3                ! Subtract with Abs
        CUT         THRESHOLD=11    ! Trigger is Delta phi less than 11
```

Figure 4.5: Level 2 tau trigger. See text for details

a maximum of 100. Thus, at high instantaneous luminosity, this trigger is prescaled by a factor of 100. As the luminosity dropped over the course of a run, the dynamic prescale is adjusted in steps down to 1. During a run using dynamic prescales, the average prescale across the run was typically of order 3–6. That is, the `TAU_20_MET_20` trigger processed 1/3–1/6 of the integrated luminosity of that run. Enough information was stored from each run (and saved in a database) to reconstruct the total luminosity seen by each trigger.

The `TAU` requirement is the central requirement of this trigger. This set of cuts, except the track P_T threshold cut, was implemented in a special hardware board called the tau Mercury module using Level 2 cluster finder information to select tau-like clusters. This part of the trigger requires a candidate cluster to have fewer than four trigger towers, $|\eta_D| < 1.19$ for the cluster centroid, and cluster $E_T > 20$ GeV. The first threshold, `EM_ET_OVER_TOTAL_ET`, is not used for taus; it looks at the fraction of E_T for a candidate jet which is electromagnetic and is disabled by allowing any value in the range 0–1. Also, the track P_T threshold is set here, but applied separately, later. The `CUT` requirement asks that at least one cluster in the event passes these requirements. If no cluster in the event passes these requirements, the event does not pass the tau trigger.

The highest E_T Level 2 jet cluster passing the `TAU` requirement is stored in the tau Mercury module. The E_T, seed tower index for η and ϕ, the η and ϕ centroid, and some other information is stored about this cluster. (See the discussion of clustering in Section 3.3.1.) The seed tower indices store the tower number of the cluster seed tower using Level 2's tower numbering: $0 \leq \eta \leq 41$, $0 \leq \phi \leq 23$. The Level 2 cluster which passes the `TAU` requirement is called a Level 2 tau cluster.

Another trigger module simultaneously calculates the $\not\!\!E_T{}^2$ of the event. To save time at Level 2, all cuts on $\not\!\!E_T$ are applied instead to $\not\!\!E_T{}^2$. Thus, this trigger requires at least 20 GeV of $\not\!\!E_T$.

Next, the track list cut is applied, using a small amount of code on the Level 2 executive board. (Continuing the planetary motif, the Level 2 trigger executive board is called the "Jupiter" module.) This code loops over the list of tracks identified by

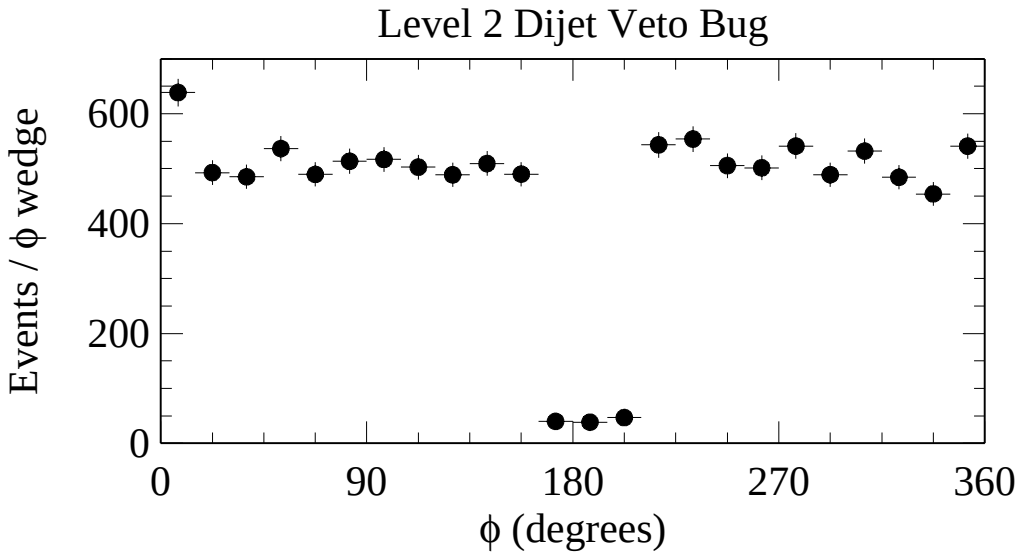


Figure 4.6: The dijet veto bug, shown for a $W \rightarrow \tau\nu$ monte carlo simulated event sample after all event selection cuts. The 24 ϕ bins (one for each ϕ wedge) are numbered 0–23, clusters are indexed by the ϕ index of their seed tower, as described in the text. This bug removes most events in bins 11-13. The spike in ϕ bin 0 is a side effect of the Level 2 clustering algorithm but does not affect the tau efficiency.

the CFT (see Section 3.7.5). Each track passing the track P_T threshold stored in the `TAU` cut (4.8 GeV, or P_T bin 2) is extrapolated to the calorimeter radius using a lookup table. This lookup table returns a list of calorimeter ϕ wedges to which this track may extrapolate, allowing overlap when a track extrapolates to a region between two ϕ wedges. The `TRL_TAU_CUT` returns true when at least one track extrapolates to the same ϕ as the tau cluster ϕ seed tower index.

The final cut in the Level 2 tau trigger is the dijet veto. The two highest-E_T jets are selected. If the two highest-E_T jets at Level 2 are opposite each other in ϕ within 30°, then the event is rejected. Remembering that at Level 2, the calorimetry segmentation is 24 towers in ϕ everywhere, this cut says that the two highest-E_T jet clusters' seed towers must be fewer than 11 towers in ϕ from each other. This cut rejects background jet events where one jet fluctuated to be narrow and tau-like.

The dijet veto processing contains a bug, causing an efficiency loss of 12.5%. In events containing only one Level 2 jet cluster, the dijet veto takes the second jet's seed

ϕ index to be ϕ index = 0. Thus, events with only one Level 2 cluster are rejected when that cluster is at $11 \leq \phi$ index ≤ 13. Figure 4.6 shows this, using the $W \rightarrow \tau\nu$ monte carlo simulation event sample discussed in the next chapter. After all event selection cuts, the spectacular loss of efficiency for 1/8 of the detector (in ϕ) is clear. Also noticeable is a 25% enhancement for ϕ bin 0 relative to the other ϕ bins. This small enhancement does not affect the efficiency for identifying taus.

Note that the CFT track match and the dijet veto are both applied against the *seed tower* ϕ index 0–23. Due to details of Level 2 clustering, discussed with other systematic effects in Section 7.2.4, the efficiency for identifying tau clusters is not charge symmetric.

This trigger did not change during Run 1a in any significant manner. Once the trigger was verified to be working, the only change was the addition of dynamic prescaling.

In addition to the Level 2 trigger used for the analysis data, a second Level 2 trigger selects events for background studies. This trigger is named `TAU_20`, and it includes all of the cuts of the `TAU_20_MET_20` trigger except the $\not\!\!E_T$ cut. Removing this cut raises the trigger rate substantially, so this trigger is heavily prescaled. Therefore, about 30000 events were written to tape from this trigger, equivalent to about 1/42 of the integrated luminosity of the analysis stream.

Still, the `TAU_20` trigger provides a sample of tau-like background jets with a small "contamination" of $W \rightarrow \tau\nu$ which for the purposes of background studies can be removed using a monte carlo simulation. Events which pass this trigger (and the corresponding Level 3 trigger, described in the next section) receive special attention in Section 5.2.

Without taking dynamic prescales into account, the cross section of the main Level 2 tau trigger is 75 nb (30 times greater than the physics cross section $W \rightarrow \tau\nu$) and the cross section of the `TAU_20` trigger is 400 nb. Since the `TAU_20` trigger, by default, is prescaled by a factor of 40, the actual *rate* of the `TAU_20` trigger is equivalent to a cross section closer to 10 nb.

4.4.3 Level 3

The Level 3 trigger is discussed in general in Section 3.7.7. Briefly, the Level 3 trigger processes events on a commercial computing platform using almost the same FORTRAN event reconstruction software as used offline. After performing event reconstruction, described briefly here, the Level 3 trigger checks each event to see if it passes any trigger. The tau triggers are described in this section.

The event reconstruction code first processes the raw digitized information from the detector components into "physical" quantities: track hits on CTC wires, energy deposits in calorimeter towers, and so on. Next, reconstruction calculates physics quantities such as the $\overline{\mathrm{p}}\mathrm{p}$ collision point along the z axis (called the z vertex), calorimeter clusters (see Section 3.3.1), and event $\not\!\!E_T$. Note that charged tracks are not automatically reconstructed for every event, saving a considerable amount of processing time for most events.

Calorimetry reconstruction includes the standard calorimetry "Cleanup" code module which accounts for known detector effects that cause a spurious measurement of energy. Single phototube spikes—noise amplified by a phototube into an apparent measurement of energy—are removed; each calorimeter tower feeds two phototubes. Cleanup also removes "Texas towers" in the gas calorimeters where a neutron appears to deposit an enormous amount of energy. Sometimes trouble with a calorimeter chamber or its electronics causes the tower to always appear to have energy deposits (a "hot tower"). The standard calibration processing locates these sorts of problems, allowing Cleanup to account for them.

For Level 3 processing, the event z vertex is set to $z = 0$ for all events. Since $\sin\theta$ depends upon the z of the vertex, the choice of this z affects both E_T and $\not\!\!E_T$ calculations. The Collaboration made the decision to choose a *consistent* z vertex rather than one which, depending on calibration constants and multiple $\overline{\mathrm{p}}\mathrm{p}$ interactions, might fluctuate in a non-predictable manner. Calibration constants are discussed in Section 3.7.3.

The text of the Level 3 tau trigger is presented in Figure 4.7. Tau reconstruction is

```
TRIGGER EX01_TAU_10_MET_20_V4
    SELECT PREREQ
        PARAMETER TALK_TO SETNAME TAU_PREREQ
            PREREQUISITE TAU_20_MET_20_V1/LEVEL=2
            PREREQUISITE TAU_20/LEVEL=2
            PREREQUISITE TAU_10_2JETS_SUMET_65_V2/LEVEL=2
            PREREQUISITE CLF_ANALOG_NOISE_MONITOR/LEVEL=2
            RETURN
        END_TALK
    EXECUTE JETCLU                          !Jet clustering
        PARAMETER TALK_TO SETNAME TAU_JET
            ALPHA_MAX .4                !0.4 cone (narrow)
            RETURN
        END_TALK
    SELECT METFLT                           !Filter on Missing Et
        PARAMETER TALK_TO SETNAME TAUMET_V1
            LOW_MISS_ET_CUT 20
            DI_JET YES                      !Dijet cut ON
                PHI_WINDOW          30.0
                MET_FRACTION        0.5
                MIN_ET              10.0
                MAX_MET             40.0
                RETURN
            USE_MET Z_EQ_0         !Use MET from Z=0 vertex
            RETURN
        END_TALK
    EXECUTE TRCONTROL
        PARAMETER SETNAME GLOBTRK1
    EXECUTE FINDTAU                          !Tau reconstruction
        PARAMETER TALK_TO SETNAME TAU1
            L3TO_MAKE ON
            RETURN
        END_TALK
```

Figure 4.7: Level 3 tau trigger. See text for details.

performed while processing the tau trigger, (the `EXECUTE FINDTAU` step) and not during Level 3 event reconstruction, for technical reasons.

The first requirement of the Level 3 tau trigger is a Level 2 trigger prerequisite, reducing the number of spurious Level 3 accepts. The prerequisite returns true if any of the listed Level 2 triggers accepted the event. Some of the Level 2 triggers used as prerequisites are not used in this analysis and aren't described here. If none of the Level 2 prerequisite triggers accept the event, the tau trigger returns false and does not execute further.

Following the prerequisite check, the tau trigger ensures that jet clusters of $\triangle R < 0.4$ radians are available. Although many analyses use jets clustered using the default cone size ($\triangle R < 0.7$ radians) in the standard reconstruction paths, tau analyses use the narrower cone.

Once jet clustering is finished, the event $\not\!\!E_T$ is checked; if it is less than 20 GeV, the Level 3 tau trigger stops processing at this point and returns false. Since the Level 3 trigger performs a much more complete calorimetry cleanup than the Level 2 trigger, spurious calorimetry signals causing a Level 2 tau trigger accept (through apparent E_T or $\not\!\!E_T$) are weeded out at this step.

The `DI_JET YES` parameters increase the efficiency of the Level 3 $\not\!\!E_T$ cut by relaxing the cut under certain circumstances. The only parameter important for this analysis is `PHI_WINDOW 30.0`: If the two leading jet clusters in the event are back-to-back in ϕ within 30°, the event is rejected. This cut removes a large fraction of background jet events.

If an event passes the $\not\!\!E_T$ cut, then the tau reconstruction module is executed. Tau reconstruction requires that charged tracks be reconstructed, so global tracking is executed first. Following track reconstruction, the standard tau reconstruction module FINDTAU (described in detail in Section 4.5.1) executes using its default cuts. Briefly, FINDTAU accepts only events with at least one jet cluster passing the following cuts:

- the cluster has fewer than 6 calorimeter towers with $E_T > 1$ GeV,
- the cluster has a total E_T of at least 20 GeV,

- the cluster has a seed track (with $P_T > 5$ GeV) pointing toward the calorimeter cluster.

The Level 3 tau trigger passes any event accepted by FINDTAU. In addition to this Level 3 tau trigger, a second Level 3 trigger processes events for background studies. This second trigger, `EX01_TAU_10_MET_10`, differs from `EX01_TAU_10_MET_20` in two respects: It allows, as a prerequisite, only the Level 2 `TAU_20` trigger, and the $\not\!\!E_T$ threshold is 10 GeV instead of 20 GeV. Remember that the `TAU_20` trigger does not apply any $\not\!\!E_T$ threshold at all, so applying the cut here results in a substantial rate reduction.

The cross section for the standard Level 3 tau trigger is 10 nb; the cross section for the background study trigger is 2 nb.

4.5 Offline Filtering

After being written to tape by the DAQ, all events pass through a processing step common to all physics data streams, called "production" or offline processing. Before the data for a run is processed through production, a number of calibration jobs measure the most up-to-date constants for that run. For instance, while collecting data, the beam position calculated from the *previous* run is used as the beam position in Level 3. After a run ends, calibration processing finds the beam position for that run, which is then stored in a calibration database for use in offline processing. (See Section 3.7.3.) Thus, production uses the final calibration constants.

Also, production uses the full data analysis package which requires significantly more CPU time than the abbreviated package used online for Level 3. Following production, events are filtered into streams of related triggers. The tau trigger data is written to the XOX1_3P stream, the "exotic data" stream. After filtering, this stream contains about 2 million of the total 16 million events written to tape during Run 1a.

Note that these 2 million events may have passed any trigger in the exotic stream, not necessarily just a tau trigger. The exotic stream encompasses many triggers designed for doing many different kinds of physics, and the tau triggers of interest to this dissertation are a small fraction of the exotic stream.

Offline processing creates two versions of each dataset. The first version (a DST, for Data Summary Tape) not only contains the results of offline processing, but also contains all of the raw data. As a result, a DST with many events requires a large amount of storage space. To allow easy access to large datasets, production also creates a PAD (Physics Analysis Dataset) version of the data, which contains the results of offline processing but none of the raw data. In addition, calorimeter and track data are compressed with some loss of precision and also some loss of information. Very low momentum tracks ($P_T \ll 1$ GeV) and low energy calorimeter towers ($E_T < 0.1$ GeV) are discarded. This analysis uses the PAD dataset; the small loss of precision has a negligible effect on this analysis.

4.5.1 Tau Reconstruction

The FINDTAU [31, 32] analysis module distinguishes hadronic tau jets from background jets. This module implements two clustering methods – one used in the online system, the second used in offline analysis. The online clustering method first looks at 0.4 cone JETCLU clusters, counting the number of towers in each cluster with E_T over 1 GeV. Only a cluster with six or fewer towers over 1 GeV is accepted as a tau candidate. This is potentially inefficient when compared with the nearest-neighbor clustering (see Section 3.3.1) used offline in FINDTAU, but a monte carlo simulation study showed that this inefficiency—after other event selection cuts are applied—is negligible.

If the cluster contains six or fewer calorimeter towers with $E_T > 1$ GeV, and if the sum of the E_T in these towers is greater than 10 GeV, FINDTAU looks for a "seed track," a track with $P_T > 5$ GeV associated with the calorimeter cluster, that is, within $\triangle R < 0.5236$ radians (30°) from the calorimeter cluster centroid. Online, the seed track is the highest-P_T track within a 0.4 radian cone in $\triangle R$ from the calorimeter cluster centroid. Offline, the seed track is the highest P_T track within a 10° cone (in θ and ϕ, not η and ϕ) from the calorimeter cluster centroid. Thus, Level 3 seed track identification is overefficient: It applies a looser event selection cut than is used offline.

The z vertex of the tau cluster—the z position along the beamline of the $\overline{\mathrm{p}}\mathrm{p}$ collision generating the tau—is defined to be the z vertex of the seed track, or the point in z

where the seed track intercepts the z axis. Remember that the z axis is defined to be along the beamline. All tracks (including the seed track) which satisfy the following requirements are added to the tau cluster: $P_T > 1$ GeV, $\triangle R < 0.5236$ radians, and z vertex within 5 cm of the tau z vertex. Tracks which do *not* intercept the z axis near the tau z vertex are assumed to come from a separate $\overline{\mathrm{p}}\mathrm{p}$ interaction.

FINDTAU creates one TAUO, or Tau Object, for each calorimeter cluster which passes these cuts. Tau cluster information is collected so that CDF analysis code can access this information using the name "TAUO"; the abbreviation "TAUO" is often used in place of the words "tau cluster as identified by FINDTAU," and is used as such in this dissertation.

Offline clustering proceeds similarly. The primary difference with the offline FINDTAU clustering algorithm is that it provides more adjustable parameters and options and that it uses a neighbor-based clustering algorithm rather than a purely cone-based algorithm. (See Section 3.3.1.) As before, FINDTAU uses $\triangle R < 0.4$ radian JETCLU clusters to seed its cluster finding. FINDTAU chooses the highest E_T tower over 3 GeV in each JETCLU cluster and uses that as the *seed tower* of the candidate tau cluster. FINDTAU checks all calorimeter towers neighboring the seed tower; any towers with $E_T > 0.1$ GeV and within $\triangle R < 0.4$ radians from the seed tower are added to the cluster. This algorithm is repeated for each tower added to the cluster until no more towers are added.

As before, the total E_T in the cluster must be greater than 10 GeV, and the cluster is rejected if more than six towers have E_T greater than 1 GeV. The six tower limit is a narrowness cut. Tracks are associated with the cluster in the same manner as online.

Online, FINDTAU calculates each tower's E_T from its energy assuming the event z vertex is located at $z = 0$. Offline, FINDTAU uses the seed track z vertex, defined above, as the tau cluster z vertex, calculating the cluster E_T using that z vertex (or assumed z position of the $\overline{\mathrm{p}}\mathrm{p}$ interaction.) Although the calculated E_T of a cluster varies somewhat depending on the z vertex used, the Level 3 trigger is not inefficient; the Level 3 trigger requires $E_T > 10$ GeV and the analysis uses a higher threshold of $E_T > 20$ GeV.

As with the online version of FINDTAU, the offline version creates one TAUO for each tau candidate cluster.

4.6 Preselection

The 2 million events from the XOX1_3P data stream were further filtered and reprocessed to a more useful data sample by John Conway and Leslie Groer [33]. This data sample is referred to as the TAUMON (tau monojet) data sample, and it contains 66185 events from XOX1_3P. The reprocessing was done because improved calibration of many CDF detector components became available after Run 1a. For example, a new calibration of the CTC found a small non-uniformity in the solenoidal magnetic field. Any adjustment to charged track identification potentially affects tau reconstruction, so tau clusters were recalculated. Limited computing resources ruled out the option of reprocessing all of CDF's data through production, so during preselection, the XOX1_3P data stream was processed through the corrected version of production.

Preselection does not require any specific triggers; at the time of preselection, it had not been decided which triggers were most suitable for this analysis. The preselection saves events which are identifiable as a monojet event. To be selected as a monojet event, the event must have

- a jet cluster with $E_T > 15$ GeV, $|\eta| < 1$, and at least one track,
- no other jet with $E_T > 10$ GeV,
- no jet opposite the leading jet (a 3 dimensional angle with the monojet greater than 160°) with $E_T > 5$ GeV,
- $\not\!\!E_T > 20$ GeV and $\not\!\!E_T$ significance > 2.4.

This collection of event selection cuts constitutes the "loose dijet" cut; after applying this cut, the presence of background jet events in the event sample is greatly reduced. Creating the TAUMON event sample from the XOX1_3P data stream is a time consuming process which requires many weeks of computer processing—the XOX1_3P data

stream is stored on about 40 8 mm data tapes. To avoid having to repeat this step, the background jet rejection is purposefully loose. A "tight dijet" cut is defined later.

4.7 Analysis Cuts

The same event selection cuts, listed in Table 4.1, are applied to all data samples, data and simulation, although certain cuts are relaxed for background studies as listed in the table and discussed in Section 5.2.

4.7.1 Monojet Selection

Data taken from some runs is suspect because of detector, DAQ, trigger, or other problems. 57016 of the 66185 events occur in runs determined to be "good" runs [33]. Many of the runs marked as bad occur before the Level 2 tau trigger worked, as described in Section 4.4.2. The good run requirement is not applied during initial data selection or preselection; the good run list was still changing at that time. Thus, the good run requirement is imposed during analysis in the global cuts. Of course, all monte carlo simulation events come from a "good run."

These 57016 events still include a great many background events, so several cuts are applied to purify the tau signal. These cuts are:

- the run number is in the good run list,
- (total) out-of-time E_T (described below) < 5 GeV
- the Level 1 trigger passed,
- the Level 2 trigger passed,
- the Level 3 trigger passed,
- reapply the loose monojet cut, defined in Section 4.6,
- apply the tight monojet cut: No other jet opposite the monojet ($\delta\phi < 150^\circ$) with $E_T > 5$ GeV.

Monojet Selection:

0. Good Run
1. Out-of-time $E_T < 5$ GeV
2. L1 trigger passes
3. L2 trigger passes †
4. L3 trigger passes †
5. Loose Monojet cut †
6. Tight Monojet cut †

Tau ID:

7. At least one TAUO exists
8. Seed track $P_T > 6$ GeV
9. $|$ Tau z vertex $| < 60$ cm
10. Tau $E_T > 20$ GeV

QCD Rejection:

11. Level 2 cluster not 2x1 (ϕ, η)
12. Event $\not\!\!E_T > 25$ GeV †
13. Track+CES mass < 1.8 GeV
14. Reject single π's without neutral energy

EWK Rejection

15. electron rejection
16. muon rejection

Sample Division

17. separate isolated, nonisolated events
18. reject taus from nonisolated sample

Table 4.1: The analysis cuts are separated into five groups; each group is described separately in the text. Cuts marked with † are relaxed for TAU-20 background studies. Abbreviations are explained in the text.

Reapplying the loose monojet cut is a double-check on the preselection, and removes no events. The tight monojet cut removes QCD background jets.

The hadron calorimeters associate timing information with the energy measured in each tower. Energy not coincident in time with the $\overline{\mathrm{p}}\mathrm{p}$ interaction is labelled "out of time." Thus, the out-of-time E_T cut efficiently removes cosmic rays, which have only a small probability of being "in-time." In addition, the out-of-time E_T cut removes events which contain another $\overline{\mathrm{p}}\mathrm{p}$ interaction where the additional interaction involves a "satellite" bunch, see Appendix C.

4.7.2 Tau ID

The tau identification cuts are applied to tau objects (TAUO) as identified by FINDTAU (see Section 4.5.1).

- require at least one TAUO identified by FINDTAU,
- seed track $P_T > 6$ GeV,
- $|$ Tau z vertex$| < 60$ cm (described below),
- tau $E_T > 20$ GeV.

The first tau identification cut requires that FINDTAU has identified at least one TAUO. Since the monojet cuts require that there is only one jet over 10 GeV and the minimum tau cluster E_T is 10 GeV—and a later cut increases this threshold to 20 GeV—the tau cluster is guaranteed to be the lone high-E_T jet identified by the monojet cuts.

In the trigger, the seed track requirement is loosely applied. This step, however, uses the more stringent offline FINDTAU definition of a seed track, as discussed in Section 4.5.1.

Track "quality" cuts are applied to the seed track. The seed track must be associated with a minimum number of CTC wire hits—requiring particularly that the track's position is well measured in the CTC in at least two stereo layers and two axial layers. In addition the track must pass near the beam line; a track not passing close to the

beam position is most likely a cosmic ray, a decay-in-flight of a muon or kaon, or a decay product of a long-lived meson such as a high-momentum B meson (which can travel centimeters before decaying). The track quality cuts remove a small number of events, but these removed events are certainly not taus.

The CFT becomes efficient at finding tracks (in bin 2) at about a transverse momentum of 5 GeV. Thus, requiring at least 6 GeV transverse momentum avoids having to simulate the turn-on of the CFT efficiency.

The accuracy of the $\not\!\!E_T$ measurement degrades for events where the $\overline{\mathrm{p}}\mathrm{p}$ interaction is far from the center of the detector, so I only analyze events with a central z vertex. To an event farther than 60 cm from $z = 0$, the 2° crack in the forward calorimeters becomes larger on one side and a more significant amount of transverse energy can be lost in that direction—about a 5% difference 60 cm from $z = 0$, a 10% difference 1 m from $z = 0$, nearly a 20% difference 1.5 m from $z = 0$. This effect can cause a significant mismeasurement of $\not\!\!E_T$, so these events are removed from the event sample. See Figure 3.8 for an illustration of the size of the 2° hole.

Removing events whose $\overline{\mathrm{p}}\mathrm{p}$ interaction occurs far from $z = 0$ also selects events with a z vertex inside the volume of the VTX. Thus, the z vertex of the $\overline{\mathrm{p}}\mathrm{p}$ interaction is well-measured for events selected by this cut.

4.7.3 QCD Background Rejection

The monojet selection and tau identification steps remove many QCD background jet events, but a significant number still remain in the event sample. A set of event selection cuts targeting differences between hadronic tau jets and background jets further reduces the presence of this background. In addition, clusters with certain characteristics are sufficiently biased as to affect the charge asymmetry measurement unless removed from the event sample. These background jet events and events containing badly biased clusters are removed by the following cuts:

- level 2 cluster shape is not 2x1 (ϕ, η) in trigger towers
- event $\not\!\!E_T > 25$ GeV,

- charged track + (neutral) CES invariant mass < 1.8 GeV/c,
- reject single π's without neutral energy (i.e., associated π^0's).

The first cut removes a class of events known to have a charge bias due to trigger effects. This bias is further described in Section 7.2.4, but briefly, the trigger efficiency for a certain class of calorimetry clusters is strongly charge biased—these clusters, when associated with a positively charged track, are half as efficient as when associated with a negative track. That is, the Level 2 trigger has an egregious inefficiency for positive taus when the Level 2 calorimeter cluster is two towers wide in ϕ and one (trigger) tower wide in η. To avoid the need to correct for this inefficiency, these events are removed from the event sample.

The more stringent $\not\!\!E_T$ cut of $\not\!\!E_T > 25$ GeV removes a large percentage of the remaining dijet events.

Since electrons and muons are specifically rejected from the tau event sample (as described in the next section), the possibility that an electron or muon is the particle leaving a charged track can be neglected. Thus, under the assumption that charged tracks in tau clusters are purely hadronic (e.g., charged pions and kaons) and energy deposited in the CES is purely neutral and electromagnetic (e.g., photons from neutral pions), one can calculate the total invariant mass of all particles (except the neutrino) in the tau cluster. Remember that the CES is designed to accurately identify the position of electromagnetic particles and that charged tracks have a well-measured trajectory.

Daughter particles from a tau have an invariant mass less than the mass of the tau, 1.78 GeV, due to the undetected neutrino. Allowing for CES and tracking resolution, requiring the cluster invariant mass to be below 1.8 GeV efficiently accepts taus while removing many background jet events.

The last cut listed above removes jets that contain a lone pion unaccompanied by π^0's. Background QCD jets which hadronize into a single pion pass most of the QCD rejection cuts without any problem since such jets are as narrow as tau jets. Thus, no information is present in these jets to help one distinguish between a single-pion quark or gluon jet from a single pion tau decay. Such clusters are identified as those which

contain a single charged track in the 10° cone that is associated with a calorimeter cluster whose energy is more than 90% hadronic. This cut removes many quark and gluon jets, but few taus.

4.7.4 Electroweak Background Rejection

Electroweak rejection encompasses removing electroweak background events by loosely identifying electrons and muons. Ideally, one chooses cuts that efficiently identify electrons and muons without also efficiently selecting taus.

The event selection cuts listed so far efficiently select electrons from the process $W \rightarrow e\nu$; electrons generally deposit energy in one or two calorimeter towers with one associated charged track, and electrons from W decays are accompanied by substantial $\not\!\!E_T$ from a neutrino. In fact, the cuts listed so far select $W \rightarrow e\nu$ events with a higher efficiency than $W \rightarrow \tau\nu$ events: $W \rightarrow e\nu$ events contain only one neutrino, so the electron E_T and event $\not\!\!E_T$ are larger (in general) than the E_T and $\not\!\!E_T$ of a cluster from the decay of an identical tau. Thus, I explicitly remove electrons from the tau event sample.

Electrons, ideally, deposit all of their energy in the EM calorimeters, so the ratio of EM energy to the total energy (or "EM fraction") for the cluster is close to 1. Also, since the electron track momentum and EM calorimeter energy are well measured, the ratio of calorimeter energy to track momentum (E/P) is close to 1. Thus, in a plot of EM fraction versus E/P (where P is understood to be the momentum sum of all tracks in the cluster), electrons clump near 1.0 in each variable.

Taus that decay to hadronic jets, on the other hand, are unlikely to have an EM fraction close to 1.0. Additionally, since π^0's often accompany charged pions from tau decay, tau jets usually have $E/P > 1.0$.

Figure 4.8 shows a plot of EM Fraction vs. E/P for electrons and hadronic tau clusters in an ideal detector. In addition to fluctuations in the EM calorimeter energy measurement, electrons can be measured with an EM Fraction or E/P different from 1.0 for two independent reasons: Electrons sometimes bremsstrahlung photons, causing E/P to be greater than one, and high energy electron showers leak out the back of the

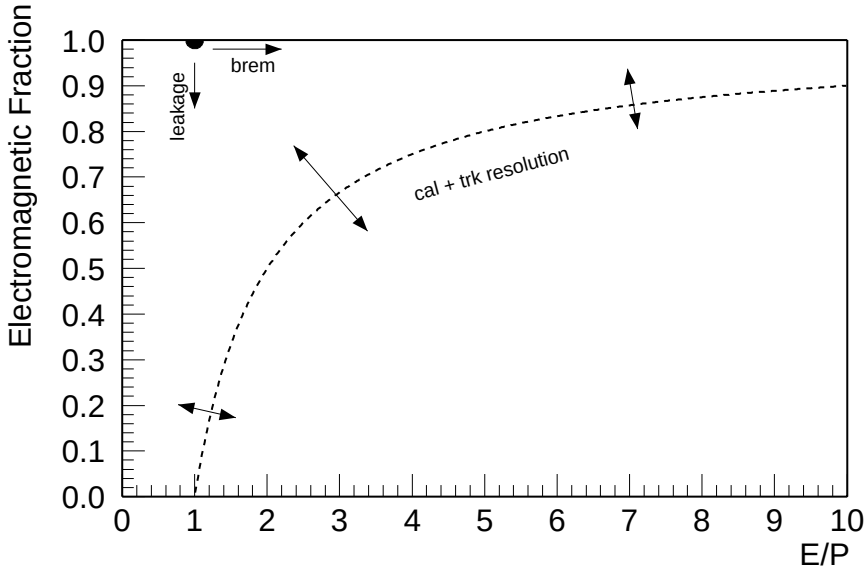


Figure 4.8: EM fraction vs. E/P for an ideal detector. Electrons are represented by the dot at EM Fraction = E/P = 1.0. Hadronic clusters from tau decay populate the area shown by the dotted line. Arrows show the effect of calorimeter resolution and electron bremsstrahlung.

EM calorimeter, causing the EM fraction to be less than 1.0.

For a hadronic tau cluster in an ideal detector, both the EM Fraction and E/P depend directly on the charged and neutral energy. A tau cluster containing only charged pions, for example, has an EM Fraction of 0 and E/P = 1.0. On the other hand, a tau cluster with an equal amount of charged and neutral energy has an EM Fraction = 0.5 and E/P = 2.0. In other words, in an ideal detector, a tau cluster has

$$\text{EM Fraction} = 1.0 - \frac{\mathit{1.0}}{(E/P)}.$$

Fluctuations in the measured calorimeter energy and track momenta, as well as EM showers leaking into the hadronic calorimeter and hadronic showers sometimes beginning in the EM calorimeters, cause this relationship to be inexact. The italicized "1.0" in the numerator represents the assumption that charged tracks deposit all their energy in the hadronic calorimeter *only*. Although inexact, this relationship inspires the cut to remove electrons. Solving for the italicized "1.0" for reasons that will shortly become clear:

$$\mathit{1.0} = \frac{E}{P}\,(1.0 - \text{EM Fraction}).$$

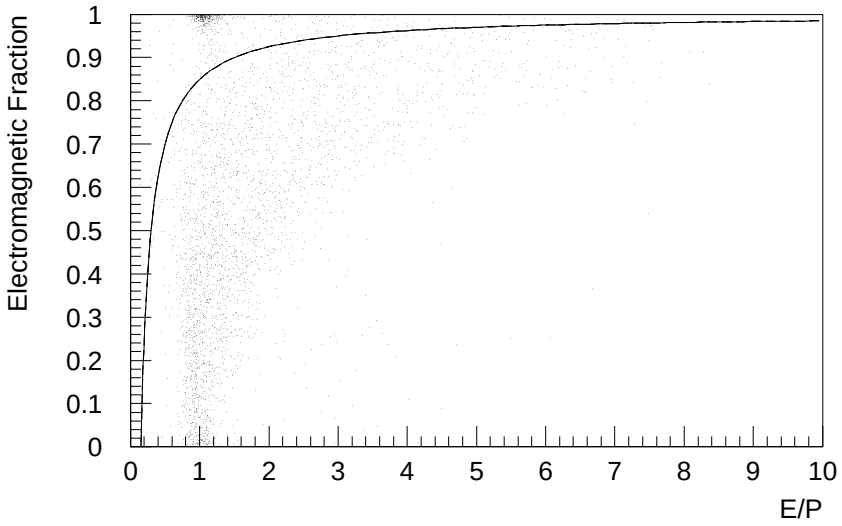


Figure 4.9: EM fraction vs. E/P, demonstrating electron rejection. Events plotted are from a tau monojet event sample without an electron rejection cut.

Given the separation between taus and electrons in EM fraction and E/P, a cut in this plane rejects electrons efficiently. With this in mind, the electron rejection cut variable, ξ, is defined as follows:

$$\xi = \frac{E}{\sum P}\left(1.0 - \frac{E_{\mathrm{EM}}}{E}\right) = \frac{E_{\mathrm{HAD}}}{\sum P};$$

making explicit that the momentum sum is over all tracks in the 10° cone about the tau cluster. In an ideal calorimeter, all electron clusters have $\xi = 0$ and hadronic tau clusters have $\xi = 1.0$. Allowing for finite calorimeter resolution, events are accepted only if $\xi > 0.15$. This cut rejects nearly all events containing electrons. Figure 4.9 demonstrates the electron-tau separation the variable ξ offers; the curve on the plot is $\xi = 0.15$ and only events *below* the curve are accepted by this cut.

The effectiveness of this cut at rejecting electrons and keeping hadrons is shown in Figure 4.10. Note the y axis uses a log scale. Nearly all events where the tau decays into an electron (plus neutrinos) fall below the cut on ξ; only a small fraction of the hadronic decays do.

The main electroweak background remaining after the electron rejection cut is W and Z processes that produce a muon plus jet(s), or $W \to \mu\nu$ + jet and $Z \to \mu\mu$ + jet. The muon or muons, of course, deposit little energy in the calorimeters, thus producing

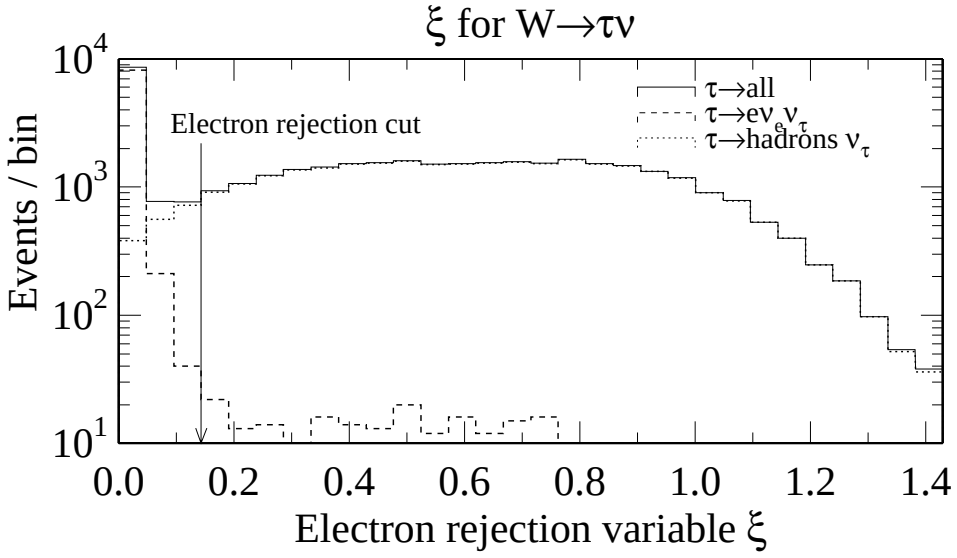


Figure 4.10: Electron rejection variable ξ for a $W \to \tau\nu$ monte carlo event sample. The solid line shows all $W \to \tau\nu$ events which pass most TAUMON event selection cuts. The dashed and dotted lines show the subsamples where the tau decays into an electron vs. where is does not; the arrow shows the cut of 0.15.

a fake $\not\!\!E_T$ signal and passing the monojet cuts. The QCD jet in these events sometimes fluctuate to appear like a tau cluster. These events are identified by the energetic track the muon(s) leave in the CTC. Thus, for the event to remain in the tau event sample, the highest P_T track outside a 30° cone from the tau cluster must have a P_T less than 10 GeV.

In addition to removing electroweak muon backgrounds, this cut reduces the QCD jet background contamination: When a jet is not measured in the calorimeters because it travels into an uninstrumented region (for example), the jet can still be identified if it contains a high momentum charged track.

4.7.5 Sample Division

The final set of cuts separates the analysis sample into two subsamples, isolated (ISO) and nonisolated (NISO). The purpose of the track isolation cut is to separate the data sample thus far into a tau-rich sample (ISO) and a QCD-background rich sample

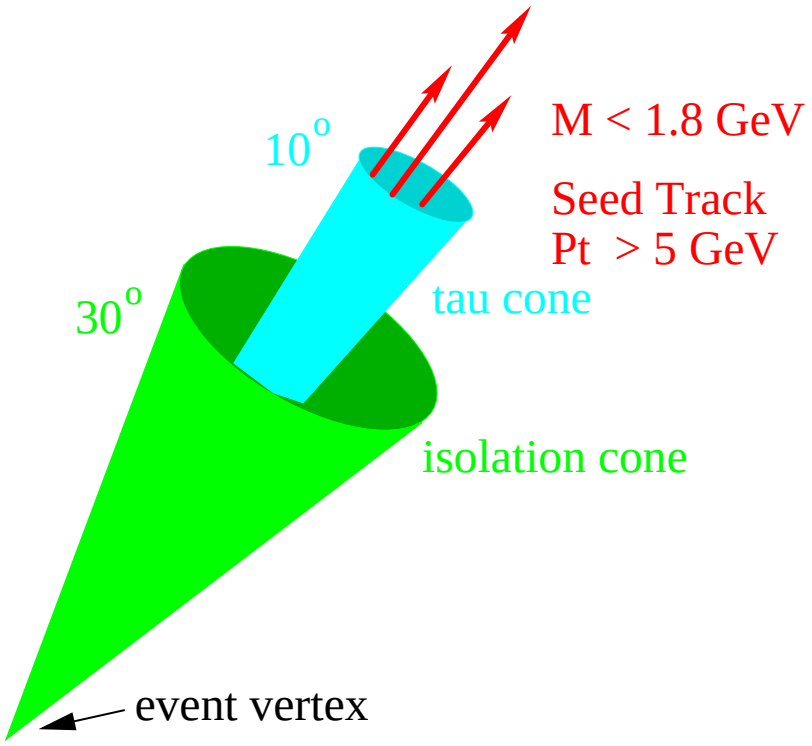


Figure 4.11: Isolation annulus

(NISO). The motivation for the track isolation is discussed in Section 4.2. Briefly, a calorimeter cluster with associated tracks in an annulus outside the 10° cone is more likely to a background jet event than it is to be a tau event.

Since the underlying event and possible additional minimum bias interactions can contribute tracks to a jet which are not part of that jet, one would like to keep the isolation annulus reasonably narrow to avoid unnecessary inefficiency for tau jets. The isolation annulus is defined to be 10–30°, schematically shown in Figure 4.11.

An event with at least one track with momentum greater than 1 GeV in the isolation annulus is tagged as a QCD background event. Otherwise, it is tagged as a tau event. To further improve the separation between the two data samples, one addition cut is imposed on the QCD data sample: The closest track in the isolation annulus to the tau candidate must have a momentum greater than 2 GeV. This cut removes the bulk of the $W \rightarrow \tau\nu$ contribution to the NISO data sample and is shown graphically in Figure 4.12.

When a tau cluster is not isolated, the track in the isolation annulus is usually of low momentum, from underlying event or an additional $\overline{\mathrm{p}}\mathrm{p}$ interaction. In contrast,

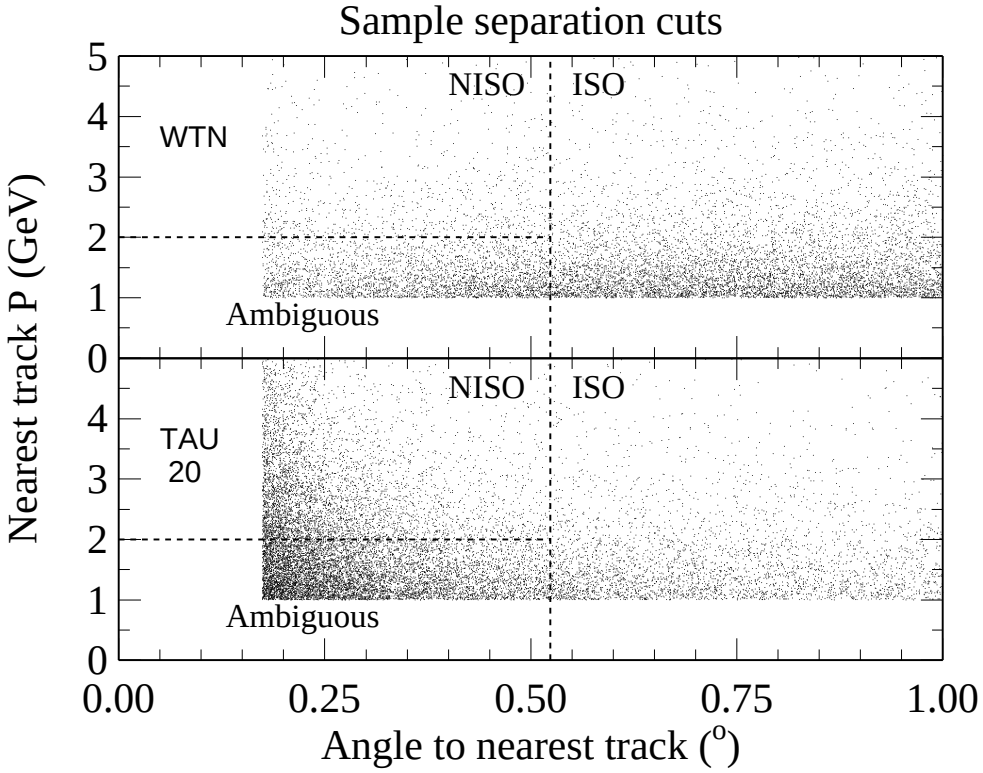


Figure 4.12: Sample separation cut for WTN, TAU-20. The division of the event sample thus far into three subsamples is shown in this plot for WTN (top) and TAU-20 (bottom). Events with no track of momentum over 1 GeV in a 10–30° annulus are isolated. Events which contain a 1 GeV track in the isolation annulus are further divided into non-isolated events (where that track's momentum is greater than 2 GeV) and events which are ambiguous, as described in the text.

Monojet Selection		
Good Run	161018	
Out-of-time $E_T < 5$ GeV	146290	
L1 trigger passes	144522	
L2 trigger passes	32799	
L3 trigger passes	32644	
Loose Monojet cut	20488	
Tight Monojet cut	14738	
Tau ID		
At least one TAUO exists	14383	
Seed track $P_T > 6$ GeV	13590	
$\mid$ Tau z vertex $\mid < 60$ cm	12849	
Tau $E_T > 20$ GeV	12791	
QCD Rejection		
L2 cluster not 2x1 (ϕ, η)	12021	
Event $\not\!\!E_T > 25$ GeV	9336	
Track+CES mass < 1.8 GeV	8228	
Reject $\pi^{\pm}$ w/ no π^0	8069	
EWK Rejection		
electron rejection	1308	
muon rejection	1208	
Sample Division		
separate ISO, NONISO	644	564
reject taus from NONISO	644	187

Table 4.2: Number of TAUMON events remaining after each cut. See Table 4.1 and following discussion in the text for the definitions of the above cuts and abbreviations.

background jets are more likely to contain a higher-momentum track in the isolation annulus. Thus, this cut on the isolation track reduces the number of taus in the non-isolated sample, increasing the contrast between the signal and background data samples. The contrast is clearly visible in Figure 4.12. The $W \to \tau\nu$ monte carlo event sample and TAU-20 event sample show markedly different shapes in these variables; removing events where the nearest track (within 30°) has a momentum less than 2 GeV clearly removes a greater fraction of the $W \to \tau\nu$ sample than of TAU-20.

4.8 Final Tau Sample

The number of events in the event sample which pass each cut is shown in Table 4.2. Of the 644 isolated events passing all cuts, 517 are 1- or 3-prong.

A 1- and 3-prong excess, showing a presence of taus, is clear in the TAUMON event sample. Also evident is the presence of a significant number of background events, especially for events with 2 tracks in the 10^{o} cone.

Chapter 5

Monte Carlo and Background Event Samples

> Do not quench your inspiration and your imagination; do not become the slave of your model.
>
> — Vincent van Gogh [3]

5.1 Monte Carlo Simulated Datasets

No large, pure sample of tau events exists at CDF with which one can study event selection and trigger cut efficiencies and biases. Electron and muon analyses, on the other hand, use data sets taken with many overlapping triggers and with different physics processes. For example, $Z \to ee$ events can be loosely identified by cutting tightly on one "leg" of the Z and loosely on the other. The efficiency and bias of event selection cuts can be studied by applying these cuts to the second "leg" (that is, lepton) of the Z.

Such samples are not readily available for taus in samples large and pure enough to be useful, so I must rely on simulations of the process $W \to \tau\nu$. In addition to studying event selection cut efficiency and biases, since jet backgrounds (especially) are a significant fraction of the tau event sample, the $W \to \tau\nu$ process must be simulated in order to provide information about distributions of positive and negative taus from $W \to \tau\nu$ to the log likelihood fit which I use to measure charge asymmetry.

I could measure the charge asymmetry without a fit, and thus, without needing to generate a $W \to \tau\nu$ simulated event sample except as needed to study biases and efficiencies. A log likelihood fit, however, ensures that Poisson statistics are handled properly and that I make the best possible measurement of the *uncertainty* of the asymmetry measurement. A log likelihood fit, as described in the following chapter,

naturally returns an accurate measurement of the statistical uncertainty of each parameter measured by the fit. Correlations and event sample size are automatically taken into account in this calculation.

Thus, for these reasons, I use a monte carlo simulated $\overline{\mathrm{p}}\mathrm{p} \to W \to \tau$ event sample both to search for charge bias in the event selection and trigger cuts and to provide information about positive and negative tau distributions to the log likelihood fit. This latter purpose of the $W \to \tau\nu$ simulation is described in more detail in the next chapter.

Of course, no monte carlo simulation can be perfect—especially for a physics process such as charged lepton asymmetry which depends on a model which is phenomenological only, e.g., the parton distribution functions. Thus, this "imperfect" modeling of $W \to \tau\nu$ introduces a measurable uncertainty into the charge asymmetry measurement. This uncertainty is discussed in Chapter 7.

The electroweak backgrounds are studied similarly to $W \to \tau\nu$, with a monte carlo simulated event sample. The four monte carlo event samples are named as follows and are described in the sections that follow.

- WTN - $W \to \tau\nu$ monte carlo,
- WEN - $W \to e\nu$ monte carlo,
- WMN - $W \to \mu\nu$ monte carlo,
- ZLL - $Z \to ll$ monte carlo, including $Z \to ee$, $Z \to \mu\mu$, and $Z \to \tau\tau$.

When referring to the electroweak backgrounds, the WEN, WMN, and ZLL monte carlo event samples are collectively referred to as the EWK event sample.

5.1.1 Generating the Monte Carlo Simulation

The Herwig monte carlo generator [34] is used with the CDF detector simulation package QFL [35]. Herwig generates a $\overline{\mathrm{p}}\mathrm{p}$ interaction resulting in a given process, here $\overline{\mathrm{p}}\mathrm{p} \to W$ or $Z + X$ where the electroweak boson decays into a lepton. Herwig follows the spectator quarks' hadronization and fills in the underlying event. (See Section 2.4.) QFL takes the particles as Herwig generated them and propagates these

particles through the CDF detector; QFL turns the generated particles into detector quantities including track fits and calorimeter energy deposits.

Although Herwig decays the tau into daughter particles, it does not correctly handle resonances in the tau decay (such as the a_1 and the ρ) or tau polarization. In addition, this version of Herwig uses out-of-date tau decay branching fractions. Thus, I use TAUOLA [36] to resimulate the tau decay before executing QFL. TAUOLA is a tau monte carlo package written at CERN; by design, it correctly handles the tau polarization and tau decay resonances and it also uses more recent branching fractions. Cal Loomis integrated TAUOLA into the CDF simulation package [37].

During Run 1a, the luminosity was often high enough that many events contain detector information from more than one $\overline{\mathrm{p}}\mathrm{p}$ collision. The additional collisions are usually minimum bias events, as discussed in Section 3.6. Although these additional minimum bias events rarely add high-P_T jets or leptons to the event, they still add additional charged tracks and calorimeter energy which can affect the efficiency of tau identification cuts. Thus, an addition step in the simulation used the simulation module MIXEVT to mix minimum bias events into the W or Z events.

We attempt to include all known detector and trigger effects in the simulation. For example, when simulating the Level 3 trigger, the event z vertex (or presumed position in z of the $\overline{\mathrm{p}}\mathrm{p}$ collision) is set to $z = 0$, and the Level 2 dijet veto bug is simulated. Some effects are difficult to simulate, however, and are imperfectly simulated in Herwig or QFL.

For example, although the simulation mixes minimum bias events into the Herwig-generated events, the simulation distribution of $\sum E_T$ shows a serious discrepancy from the data. Although $\sum E_T$ is not cut upon in this analysis, the $\not\!\!E_T$ significance, which depends on the value of $\sum E_T$, *is*. Monte carlo generators (e.g., Herwig) underestimate the amount of energy deposited in the forward calorimeters by the underlying event. The charged energy, or the $\sum P_T$ of all charged particles detected in the CTC is well simulated, suggesting that the distribution of particles in the central parts of the detector is well simulated. Therefore, at worst, the $W \to \tau\nu$ monte carlo overestimates the $\not\!\!E_T$ significance and thus overestimates the overall efficiency for selecting taus. This

effect should be charge-independent, however, and thus, should not change the ratio of positive to negative taus passing the selection cuts.

In addition, the simulation of the hardware trigger does not take into account the differences between the (analog) energy scale of the hardware trigger and the (digitized) energy scale of the calorimeter towers. Since the difference in energy scale does not depend on particle charge, failing to simulate this effect does change the *relative* efficiency of positive and negative particles, only the overall efficiency.

For these reasons, any measurement of the trigger efficiency contains substantial systematic uncertainties. Fortunately, this analysis does not need to measure the trigger efficiency, only the *relative* efficiency for positive and negative taus. For this purpose, the simulation is adequate, but the "imperfections" in the monte carlo event samples require study to understand what effect they cause, if any, on the tau charge asymmetry measurement.

An important detector effect which is imperfectly modeled is the shower shape of pion showers in the calorimeters. A pion often showers into more than one calorimeter tower; thus, the energy of the pion is shared between towers. QFL-simulated pions, in general, have narrower showers than real pions do. In addition, they have too few single-tower showers [38]. This topic receives deeper scrutiny, with other systematic effects, in Chapter 7.

5.1.2 Monte Carlo Sample Normalization to TAUMON

The luminosity for WTN and the EWK monte carlo event samples can be normalized absolutely to the data; this normalization needs some discussion. CDF has published a measurement of the cross section for $\overline{\mathrm{p}}\mathrm{p} \to W + X$ followed by $W \to e\nu$ [39], often abbreviated as $\sigma \cdot B(W \to e\nu)$, that is, cross section times branching ratio. CDF has also measured the corresponding $Z \to ee$ cross section times branching ratio.

$$\sigma \cdot B(W \to e\nu) = 2.49 \pm 0.02(\mathrm{stat}) \pm 0.08(\mathrm{syst}) \pm 0.09(\mathrm{lum})\ \mathrm{nb}$$

$$\sigma \cdot B(Z \to ee) = 0.231 \pm 0.006(\mathrm{stat}) \pm 0.007(\mathrm{syst}) \pm 0.008(\mathrm{lum})\ \mathrm{nb}$$

Process	Cross Section (nb)	No. Events expected
$N(W \to e\nu)$	2.49 ± 0.12	38595 ± 1240
$N(W \to \mu\nu)$	2.49 ± 0.12	38595 ± 1240
$N(W \to \tau\nu)$	2.49 ± 0.12	38595 ± 1240
$N(Z \to ee)$	0.231 ± 0.012	3581 ± 143
$N(Z \to \mu\mu)$	0.231 ± 0.012	3581 ± 143
$N(Z \to \tau\tau)$	0.231 ± 0.012	3581 ± 143

Table 5.1: Number of events expected for electroweak processes. These estimates use the cross sections in Equations 5.1 and 5.2 as well as the integrated luminosity of the tau sample. See discussion in the text. $N(B \to yz)$ is shorthand for $\overline{p}p \to B + X; B \to yz$, e.g., where B is a W or Z and y and z are leptons.

Adding the uncertainties in quadrature, the cross sections are

$$\sigma \cdot B(W \to e\nu) = 2.49 \pm 0.12 \text{ nb} \tag{5.1}$$

$$\sigma \cdot B(Z \to ee) = 0.231 \pm 0.012 \text{ nb} \tag{5.2}$$

These cross section measurements are used to normalize WTN and the other electroweak monte carlo event samples. For the electroweak backgrounds, these cross section uncertainties result in a small systematic uncertainty discussed in Chapter 7. For WTN normalization, however, the uncertainty in the cross section times branching ratio cancels, as discussed in Chapter 6.

Assuming lepton universality, I apply these measured cross sections to muons and taus. The integrated luminosity of the tau data sample is 15.5 pb^{-1} [40]. Given a cross section and integrated luminosity, I can calculate the number of events one would observe for the various electroweak processes, see Table 5.1.

The monte carlo event samples can now be properly normalized. For example, assume a 100000 event $W \to e\nu$ monte carlo event sample is generated. With the known integrated luminosity and cross section, one expects 38595 $W \to e\nu$ events, so the integrated luminosity in the monte carlo sample is $100000/38595 = 2.6$ times the luminosity in the data sample. Therefore, the monte carlo sample needs to be scaled by $38595/100000 = 0.386$ to have the same integrated luminosity as the data. Table 5.2 summarizes the electroweak monte carlo samples, the number of events generated for each, and the scale factor for each.

Monte Carlo name	Process	No. Events generated	Scale Factor
WTN	$W \to \tau\nu$	1000000	0.0386 ± 0.0012
WEN	$W \to e\nu$	500000	0.0772 ± 0.0025
WMN	$W \to \mu\nu$	500000	0.0772 ± 0.0025
ZLL	$Z \to ll$	500000	0.0215 ± 0.00086

Table 5.2: Electroweak monte carlo samples and scale factors.

The one W/Z process not included is $\overline{\mathrm{p}}\mathrm{p} \to Z + X; Z \to \nu\overline{\nu}$. This process has about twice the rate of $Z \to ll$ where l is any one charged lepton e, μ, τ, not the sum over all three. Thus one expects about 21500 ± 900 $Z \to \nu\overline{\nu}$ events in 15.5 pb^{-1}. These events can only enter the tau analysis event sample as background when the Z recoils off a jet and this jet fakes a tau. In about one out of five $\overline{\mathrm{p}}\mathrm{p} \to W$ or Z events, the boson recoils against a jet, and the rate for a QCD jet (from an initial state parton) to hadronize similar to a tau jet is about 5% [29].

Thus, if one out of five of the $Z \to \nu\overline{\nu}$ events is $Z \to \nu\overline{\nu}$ + jets and one in twenty of these jets fakes a tau, this still leaves about 215 background events which can make it partway through the tau sample cuts. A naive monte carlo simulation of this process revealed that this process contributes—after all cuts—approximately 19 events to the tau sample, and that these events are about equally split between the isolated (signal) and non-isolated (QCD) data samples.

All of these events enter as background through a QCD jet faking a tau, so the QCD background should correctly match the shape of these jets in charge pattern and isolation, which are presumably independent of the physics source of the jet. Thus, I lump $Z \to \nu\overline{\nu}$ in with QCD backgrounds and account for the rest of the electroweak background processes by directly normalizing the monte carlo data samples, applying the tau selection cuts, and subtracting the resulting distributions.

5.1.3 Final Monte Carlo Samples

The number of monte carlo events passing the event selection is listed in Tables 5.3 and 5.4. Also listed for the electroweak monte carlo samples is the absolute number of electroweak background events once the samples are scaled to an integrated luminosity

	WTN	
Monojet Selection		
Number of events simulated	1000000	
Out-of-time $E_T < 5$ GeV	973171	
L1 trigger passes	355647	
L2 trigger passes	52372	
L3 trigger passes	49340	
Loose Monojet cut	39843	
Tight Monojet cut	38708	
Tau ID		
At least one TAUO exists	38010	
Seed track $P_T > 5$ GeV	36509	
|Tau z vertex | < 60 cm	36409	
Tau $E_T > 20$ GeV	36084	
QCD Rejection		
L2 cluster not 2x1 (ϕ, η)	31910	
Event $\not\!\!E_T > 25$ GeV	22496	
Track+CES mass < 1.8 GeV	21071	
Reject $\pi^{\pm}$ w/ no π^0	20502	
EWK Rejection		
electron rejection	14363	
muon rejection	14267	
Sample Division		
separate ISO, NONISO	11629	2638
reject taus from NONISO	11629	540

Table 5.3: Number of $W \rightarrow \tau\nu$ monte carlo events passing cuts. The details of the cuts are described in detail in the text.

	WEN		WMN		ZLL	
Monojet Selection						
Number of events simulated	500000		500000		500000	
Event passes preselection	115983		1296		5372	
Out-of-time $E_T < 5$ GeV	115763		1058		4983	
L1 trigger passes	110021		968		4705	
L3 trigger passes	109759		967		4676	
Tau ID						
Seed track $P_T > 5$ GeV	109322		850		4488	
|Tau z vertex | < 60 cm	109303		850		4485	
Tau $E_T > 20$ GeV	108993		842		4459	
QCD Rejection						
Event $\not\!\!E_T > 25$ GeV	100038		719		3438	
Track+CES mass < 1.8 GeV	93103		373		3086	
Reject $\pi^{\pm}$ w/ no π^0	93103		367		2815	
EWK Rejection						
electron rejection	157		352		1599	
muon rejection	143		132		905	
Sample Division						
separate ISO, NONISO	78	65	33	99	731	174
reject taus from NONISO	78	30	33	39	731	36
Scaled to 15.5 pb^{-1}	8.3		5.6		16.5	

Table 5.4: Number of electroweak monte carlo events passing cuts. The preselection includes three event selection cuts: the tight monojet cut, the Level 2 trigger cuts, and at least one TAUO exist. In order to save disk and tape space, events not passing the preselection were not saved.

of 15.5 pb^{-1}.

Of the 12169 WTN events passing the event selection, 10882 are 1- or 3-prong isolated clusters, representing 420 events once scaled to 15.5 pb^{-1}. In comparison, 770 electroweak monte carlo events are 1- or 3-prong isolated clusters, or 22 events when scaled to 15.5 pb^{-1}. The electroweak backgrounds, thus, add a small background to events passing the event selection cuts—about 5% of the number of $W \rightarrow \tau\nu$ events passing the event selection.

5.2 TAU-20 Cuts and Data Sample

The log likelihood fit requires, as input, an event sample which simulates the jet background, that is, an event sample containing jets that pass the event selection cuts for the tau event sample, but for which the parent particle is a parton, not a lepton. Since monte carlo simulations still poorly model many important effects, notably the underlying event, and since structure functions are imperfectly known, a monte carlo simulation is deemed inadequate for this purpose. Thus, a sample of real events—the TAU-20 event sample—is used as the jet background simulation.

When using real events to simulate a background, there is a tradeoff between accuracy of simulation and the size of the background event sample. To create a background event sample with an increased presence of background events, one must relax, remove, or otherwise change event selection cuts. The different event selection may cause systematic differences between the characteristics of the actual background events and the simulated background events.

In this analysis, the event selection cuts remove jet backgrounds efficiently, leaving only a very small fraction of QCD jets as a background; this small fraction of background jets is not representative of QCD events as a whole. Event selection cuts designed to accept taus from W decay accept only those QCD jets which fluctuate to be narrow, and only when some process in the event or mismeasurement in the calorimeters generates $\not\!\!E_T$.

Jets which fluctuate to be narrow, in general, contain fewer charged tracks than

```
TRIGGER TAU_20
    EXECUTE LEVEL2_PRESCALING
        PARAMETER   PRESCALE_FACTOR = 40
    EXECUTE DYNAMIC_PRESCALING
        PARAMETER   MINIMUM_PRESCALE_FACTOR = 40
     PARAMETER   MAXIMUM_PRESCALE_FACTOR = 100
    SELECT TAU
        PARAMETER   EM_ET_OVER_TOTAL_ET <= 1        ! Fixed cuts in
        PARAMETER   CENTRAL_TOWERS <= 3             ! Mercury-TAU PROM
        PARAMETER   ETA_MIN = 0.0                   !     "        "
        PARAMETER   ETA_MAX = 1.19                  !     "        "
        PARAMETER   ET_THRESHOLD >= 20.0 (GeV)      ! Selectable cuts
        PARAMETER   TRACK_PT >= 4.8 (GeV/c)
                  ! 90% eff at 3.0,3.7,4.8,6.0,9.2,13.0,16.7,25.0(GeV/c)
        CUT         NUMBER_OF_TAUS >= 1
    SELECT TRL_TAU_CUT
    SELECT TWO_CLUSTER_CUT
        PARAMETER   CLUSTER1_TYPE=1             ! Highest Et Cluster
        PARAMETER   CLUSTER2_TYPE=2             ! 2nd Highest Et Cluster
        PARAMETER   CLUSTER_QUANTITY=18         ! Phiseed
        PARAMETER   COMPARE_TYPE=6              ! Less than
        PARAMETER   OPERATION=3                 ! Subtract with Abs
        CUT         THRESHOLD=11    ! Trigger is Delta phi less than 11
```

Figure 5.1: Level 2 TAU-20 trigger. See text for details

QCD jets in general. Due to this and similar biases, the event selection for background studies is as similar as possible to the tau event selection. A few cuts are relaxed, mainly in the triggers and event preselection, thus increasing the statistics of the background simulation event sample. More specifically, as shown in Table 4.1, the Level 2 and Level 3 triggers, the monojet cuts, and the $\not\!\!E_T > 25$ GeV event selection cut are relaxed. The resulting event sample is named after its Level 2 trigger: The TAU-20 event sample. This event sample is used as the background simulation in the log likelihood fit, and is described in the rest of this section.

5.2.1 TAU-20 Triggers

The Level 2 trigger used for the TAU-20 event sample is shown in Figure 5.1. This trigger differs in two steps from the standard Level 2 tau trigger shown in Figure 4.5. Most critically, no Level 2 $\not\!\!E_T$ cut is applied in this trigger. Without this event selection

cut, the TAU-20 trigger passes events at a rate equivalent to a 400 nb cross section. At an instantaneous luminosity of 5×10^{30} cm^{2}s^{-1}, this trigger *alone* will consume the entire Level 2 event rate budget of 20 Hz. Therefore, the Level 2 `TAU_20` trigger is prescaled by a factor of 40, accepting events at a rate equivalent to a 10 nb cross section or about 1/8 the rate of the Level 2 `TAU_20_MET_20` trigger.

The dynamic prescale is present in the `TAU_20` trigger for the same reason a dynamic prescale is present in the Level 2 `TAU_20_MET_20` trigger; see Sections 4.4.2 and 3.7.6. During a run using dynamic prescales, the average dynamic prescale (after taking into account the usual prescale of 40) was typically of order 3–5; the `TAU_20` trigger "saw" about 1/4 the cross section it otherwise would have seen during such runs.

The `TAU_20` trigger is otherwise identical to the `TAU_20_MET_20` trigger.

Since the `TAU_20` trigger was added to the trigger table somewhat later than the `TAU_20_MET_20` trigger was, the total integrated luminosity seen by the `TAU_20` trigger is slightly less than the 1/40 one would otherwise expect. In fact, the `TAU_20` trigger saw about 1/42 the integrated luminosity exposed to the `TAU_20_MET_20` trigger. All-in-all, almost 30000 events passing the `TAU_20` trigger (and the following Level 3 trigger) were written to 8 mm tape.

The Level 3 trigger used for the TAU-20 event sample, `EX01_TAU_10_MET_10`, is shown in Figure 5.2. As with the TAU-20 Level 2 trigger, the Level 3 trigger differs only slightly from the Level 3 trigger used for the tau event sample, shown in Figure 4.7. The two differences are that the Level 3 $\not\!\!E_T$ cut is set at 10 GeV rather than 20 GeV, and this trigger processes only those events which are accepted by the Level 2 `TAU_20` trigger (prescale and all).

The `EX01_TAU_10_MET_10` trigger in Figure 5.2 is abbreviated (when compared with the `EX01_TAU_10_MET_20` trigger) for technical reasons: Once a set of "cuts" is defined for a given module, this set of cuts can be referred to by name. Thus, the discussion of the Level 3 trigger components in Section 4.4.3 applies to the trigger modules used in Figure 5.2.

This trigger is present mainly as a "cleanup" step, making certain the Level 2 cluster information is verified by the full digital event readout. The event rate out of this Level 3

```
TRIGGER EX01_TAU_10_MET_10_V4
    SELECT PREREQ
        PREREQUISITE TAU_20
    EXECUTE JETCLU                          !Jet clustering
        PARAMETER SETNAME TAU_JET
    SELECT METFLT                           !Filter on Missing Et
        PARAMETER TALK_TO SETNAME TAUMET_V2
            LOW_MISS_ET_CUT 10
            DI_JET YES                      !Dijet cut ON
                PHI_WINDOW          30.0
                MET_FRACTION        0.5
                MIN_ET              10.0
                MAX_MET             40.0
                RETURN
            USE_MET Z_EQ_0         !Use MET from Z=0 vertex
            RETURN
        END_TALK
    EXECUTE TRCONTROL
        PARAMETER SETNAME GLOBTRK1
    EXECUTE FINDTAU                          !Tau reconstruction
        PARAMETER SETNAME TAU1
```

Figure 5.2: Level 3 TAU-20 trigger. See text for details.

trigger is equivalent to a cross section of about 2 nb.

In summary, the triggers used for the TAU-20 event sample and the tau event sample, differ only in the $\not\!\!E_T$ cut and prescale applied. The event selection cuts to identify the tau cluster candidate are the same in both event samples.

5.2.2 TAU-20 Event Selection

As described in the introduction to this section, the TAU-20 event sample selection is nearly the same as the event selection for the TAUMON event sample; five event selection cuts are relaxed or otherwise different, the rest are the same. The trigger differences are described in the previous section.

The monojet cuts (loose and tight), described in Sections 4.6 and 4.7.1, are not applied to the TAU-20 events. The tight monojet cut, which is the loose monojet cut plus one additional cut, comprises selecting events which

- contain one cluster with $E_T > 15$ GeV, $|\eta| < 1$, and at least one charged track,
- do not contain another cluster with $E_T > 10$ GeV,
- do not contain another cluster, back-to-back in ϕ within 30° with the first cluster, with $E_T > 5$ GeV,
- have $\not\!\!E_T > 20$ GeV,
- have $\not\!\!E_T$ Significance > 2.4.

The first of these event selection cuts is applied, essentially, by requiring at least one TAUO to be identified; however, the remaining four monojet event selection cuts are not applied. In addition, the $\not\!\!E_T > 25$ GeV cut, later in event selection, is not applied. Instead, the Level 3 $\not\!\!E_T$ cut ($\not\!\!E_T > 10$ GeV) is reapplied, removing events where the Level 3 $\not\!\!E_T$ is fake, e.g., due to a "hot tower" identified by later calibrations but not identified when the event was written to tape. No $\not\!\!E_T$ significance cut is applied to the TAU-20 event sample.

Since the TAU-20 triggers differ from the TAUMON triggers only in the $\not\!\!E_T$ cut and technical details which do not affect the physics (that is, prescales and prerequisites),

	TAU-20	
Monojet Selection		
Number of events simulated	26269	
Out-of-time $E_T < 5$ GeV	26247	
L1 trigger passes	25888	
L2 trigger passes	25757	
L3 trigger passes	22109	
Tau ID		
At least one TAUO exists	21732	
Seed track $P_T > 5$ GeV	18022	
|Tau z vertex | < 60 cm	16769	
Tau $E_T > 20$ GeV	16103	
QCD Rejection		
L2 cluster not 2x1 (ϕ, η)	13102	
Event $\not\!\!E_T > 10$ GeV	11922	
Track+CES mass < 1.8 GeV	7573	
Reject $\pi^{\pm}$ w/ no π^0	7188	
EWK Rejection		
electron rejection	6824	
muon rejection	6361	
Sample Division		
separate ISO, NONISO	2161	4200
reject taus from NONISO	2161	1691

Table 5.5: Number of TAU-20 events passing cuts. The cuts applied are those described in this section. That is, the Level 2 and Level 3 trigger cuts are those for the TAU-20 event sample, the $\not\!\!E_T$ cut selects events with $\not\!\!E_T > 10$ GeV, and no monojet selection is applied.

Monte Carlo name	Process	No. Events generated	Scale Factor
WTN	$W \to \tau\nu$	1000000	$(9.2 \pm 0.4) \times 10^{-4}$
WEN	$W \to e\nu$	500000	$(18.4 \pm 0.9) \times 10^{-4}$
WMN	$W \to \mu\nu$	500000	$(18.4 \pm 0.9) \times 10^{-4}$
ZLL	$Z \to ll$	500000	$(5.1 \pm 0.3) \times 10^{-4}$

Table 5.6: Electroweak monte carlo samples and scale factors, scaling the electroweak monte carlo samples to the TAU-20 event sample.

the TAU-20 event selection differs from the TAUMON event selection in two ways. TAU-20 uses a looser $\not\!\!E_T$ identification and accepts events containing multiple jets or back-to-back jets. Loosening these event selection cuts increases the size of the TAU-20 sample from about 10 events to about 4000 events while preserving *most* of the tau identification cuts used in TAUMON. The number of TAU-20 events remaining after each cut is list in Table 5.5.

5.2.3 Monte Carlo Normalization

Since the TAU-20 event sample does not include any cuts not applied to the TAUMON sample, it includes all the same physics processes found in TAUMON, that is, QCD processes, $W \to \tau\nu$, and other electroweak processes. In fact, the TAU-20 event sample, in principle (although not in practice) is a superset of the TAUMON event sample. If no prescale were applied to the `TAU_20` trigger used in selecting the TAU-20 event sample, TAU-20 would contain all events in the TAUMON event sample plus additional events. The primary difference between the event samples is the *fraction* of events in the sample from $W \to \tau\nu$; TAU-20, by design, contains a much higher fraction of QCD events.

The TAU-20 event sample is intended to model only the QCD processes, so additional physics processes must be removed to avoid background or signal double-counting in the fit. The electroweak backgrounds ($W \to e\nu$, $W \to \mu\nu$, $Z \to ll$) contribute a small fraction of events to TAU-20, but the $W \to \tau\nu$ process contributes a significant fraction.

Since all four of these electroweak processes are already simulated for use in the fit, the electroweak processes (including $W \to \tau\nu$) can be subtracted from TAU-20 statistically, that is, by normalizing the number of events in each sample to the number

	WTN	
Monojet Selection		
Number of events simulated	1000000	
Out-of-time $E_T < 5$ GeV	973171	
L1 trigger passes	355647	
L2 trigger passes	62949	
L3 trigger passes	62375	
Tau ID		
At least one TAUO exists	61266	
Seed track $P_T > 5$ GeV	57929	
$\|$Tau z vertex $\| < 60$ cm	57898	
Tau $E_T > 20$ GeV	55852	
QCD Rejection		
L2 cluster not 2x1 (ϕ, η)	49076	
Event $\not\!\!E_T > 10$ GeV	49029	
Track+CES mass < 1.8 GeV	46126	
Reject $\pi^{\pm}$ w/ no π^0	44731	
EWK Rejection		
electron rejection	30629	
muon rejection	29926	
Sample Division		
separate ISO, NONISO	24488	5438
reject taus from NONISO	24488	1116
Scaled to 0.37 pb^{-1}	23.1	

Table 5.7: Number of $W \to \tau\nu$ monte carlo events passing TAU-20 event selection cuts, as described in this section.

expected due to the total integrated luminosity "seen" by the Level 2 `TAU_20` trigger, or 0.37 pb^{-1} (measured with the program LUM_CONTROL [41]). As mentioned in Section 5.2.1, the TAU-20 event sample contains approximately 1/42 of the integrated luminosity of TAUMON.

The electroweak monte carlo samples are normalized to the TAU-20 integrated luminosity in the same manner that they are normalized to the TAUMON integrated luminosity, as described in Section 5.1.2. The scale factors are listed in Table 5.6.

The number of monte carlo events passing the TAU-20 event selection is listed in Tables 5.7 and 5.8. After the TAU-20 event selection cuts are applied, and after scaling each monte carlo sample to 0.37 pb^{-1}, only 23 $W \to \tau\nu$ and about 1 electroweak background events remain in TAU-20, compared with 3852 data events. Thus, these

	WEN		WMN		ZLL	
Monojet Selection						
Number of events simulated	500000		500000		500000	
Event passes preselection	115983		1296		5372	
Out-of-time $E_T < 5$ GeV	115763		1058		4983	
L1 trigger passes	110021		968		4705	
L3 trigger passes	109759		968		4705	
Tau ID						
Seed track $P_T > 5$ GeV	109582		851		4515	
|Tau z vertex | < 60 cm	109563		851		4512	
Tau $E_T > 20$ GeV	109251		843		4486	
QCD Rejection						
Event $\not\!\!E_T > 10$ GeV	109251		843		4486	
Track+CES mass < 1.8 GeV	102032		452		4073	
Reject $\pi^{\pm}$ w/ no π^0	102032		445		3738	
EWK Rejection						
electron rejection	169		423		2155	
muon rejection	149		166		1279	
Sample Division						
separate ISO, NONISO	82	67	47	119	1044	235
reject taus from NONISO	82	31	47	49	1044	44
Scaled to 0.37 pb^{-1}	0.2		0.2		0.6	

Table 5.8: Number of electroweak monte carlo events passing cuts. The preselection includes three event selection cuts: the tight monojet cut, the Level 2 trigger cuts, and at least one TAUO exist. In order to save disk and tape space, events not passing the preselection were not saved.

electroweak processes account for less than 1% of the TAU-20 event sample. The rest of the TAU-20 event sample corresponds to the QCD jet backgrounds to this analysis.

In the isolated 1- and 3-prong region, the electroweak processes (mostly $W \rightarrow \tau\nu$) contribute 21 events out of TAU-20's 1014. Thus, $W \rightarrow \tau\nu$ accounts for about 2% of the signal region (1- and 3-prong isolated clusters) and a negligible fraction of the rest.

Chapter 6

The Log Likelihood Fit

KING CLAUDIUS: Let's further think of this;
Weigh what convenience both of time and means
May fit us to our shape
— Hamlet, Prince of Denmark, Act 4

Measuring the charge asymmetry using the TAUMON data sample, backgrounds must be removed or accounted for. The presence of QCD background jets dilutes any $W \to \tau\nu$ asymmetry signal. Remember that the charge asymmetry is measured in bins of η using

$$A(\eta) = \frac{N_\tau^+(\eta) - N_\tau^-(\eta)}{N_\tau^+(\eta) + N_\tau^-(\eta)}.$$

Thus, a charge symmetric QCD jet background mixed in with $W \to \tau\nu$ events increases N_τ^+ and N_τ^- equally, reducing the measured value of A. Even worse, a charge asymmetric QCD background can distort the *shape* as well as the magnitude of the asymmetry measurement, depending on the relative shape and magnitude of asymmetry of QCD background jets.

In addition to QCD background effects, electroweak backgrounds contribute biases to the charge asymmetry; $W \to e\nu$ and $W \to \mu\nu$ processes contain a charge asymmetry for the same reason $W \to \tau\nu$ contains a charge asymmetry (barring physics beyond the standard model); Z processes contain an entirely different form of asymmetry (a forward-backward asymmetry). Any of these asymmetries can bias the measurement of $W \to \tau\nu$ charge asymmetry. Therefore, these backgrounds, although small, must be properly accounted for.

Since, using cuts alone, these backgrounds cannot be reduced to a negligible level, they will be "removed" statistically. In order to best account for statistics as well as measure the statistical significance of the measurement, a log likelihood fit is used to

measure the relative contributions to the TAUMON data sample from $W \to \tau\nu$, QCD background jets, and electroweak backgrounds. Of course, the fit must measure N_τ^+ and N_τ^- separately, which motivates the choice of variables used in the fit.

6.1 Binning Variables

Ideally, the input to a log likelihood fit is given as distributions of events in bins of variables which provide clearly different shapes for signal and backgrounds. If the signal and backgrounds look the same in the variables provided to the fit, the fit has no information about the relative contribution of the signal and backgrounds. Using a binned log likelihood fit, one wants to bin the data and monte carlo samples in variables which provide the most distinct contrast possible.

Charge multiplicity, the count of the number of charged tracks in a 10° cone, provides a clear contrast between taus and the backgrounds of this analysis. An event sample rich in taus displays an excess of 1- and 3-prong clusters; any cut increasing the purity of taus increases this excess. QCD background jets, on the other hand, show no excess of this kind in charge multiplicity; for these jets, charge multiplicity is a smoothly varying variable peaking at some small number of tracks 2–4, depending the event selection used.

Charge multiplicity is less useful for discriminating against electroweak backgrounds. $W \to e\nu$ and $Z \to ee$ backgrounds overwhelmingly contribute tau candidate clusters with one track; $W \to \mu\nu$ and $Z \to \mu\mu$ events appear in the event selection only when the W or Z recoils against a QCD jet and this jet fragments in a tau-like manner. That is, the μ itself *very* rarely deposits enough energy in the calorimeters to be identified as a tau candidate. Finally, $Z \to \tau\tau$ background events mostly contribute events containing real taus, i.e., one or three tracks, where the second tau in the event is undetected.

Fortunately, the electroweak backgrounds are small enough and well enough simulated that the fit does not need to adjust the relative level of these backgrounds. They are normalized absolutely to the luminosity of the TAUMON event sample, as discussed in Section 5.1.2.

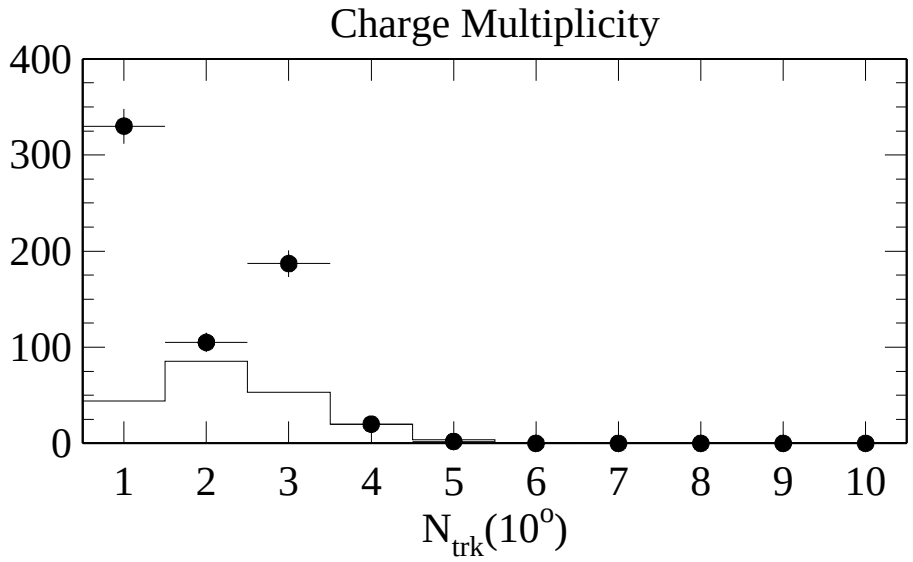


Figure 6.1: Track multiplicity. The points show the TAUMON isolated data ($W \rightarrow \tau\nu$ events plus backgrounds); the histogram shows the TAU-20 isolated data (predominantly background events). The signal sample shows an excess of 1- and 3-prong tracks relative to the background sample.

Figure 6.1 presents the charge multiplicity plot of the TAUMON isolated (tau-rich) and TAU-20 isolated (background-rich) event samples after all event selection cuts are applied. No background is yet subtracted from the TAUMON data sample. For this demonstration of relative shape, the normalization applied to the TAU-20 histogram is simple: It is normalized so that the TAUMON and TAU-20 histograms have the same number of 4-prong events. Note that some taus are identified as 2-prong clusters due to tracking inefficiency or underlying event tracks. The background sample varies smoothly as the number of tracks increases, but the signal sample shows an excess of events with 1 and 3 tracks in addition to a smoothly-varying component. This figure clearly shows the relative contribution of tau jets and background jets to the TAUMON event sample.

Although charge multiplicity is clearly a useful variable to use as input to a log likelihood fit—the $W \rightarrow \tau\nu$ signal and jet background present clearly distinguishable

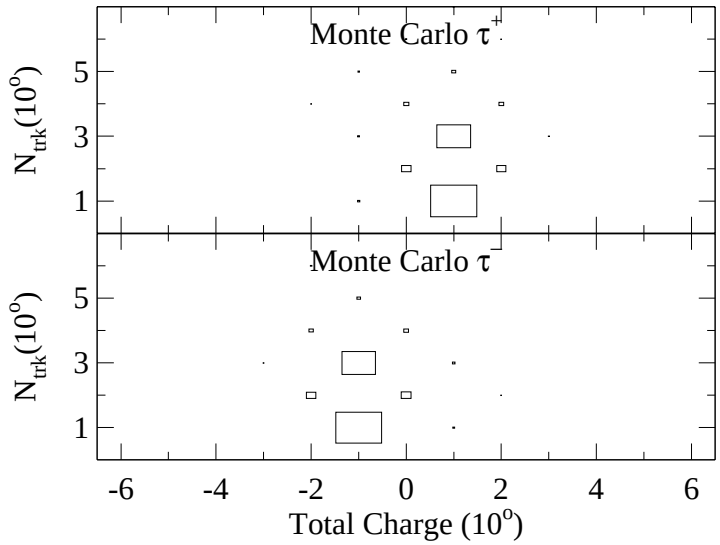


Figure 6.2: Charge multiplicity versus total charge in WTN; the area of the data squares is proportional to the count of events. The upper plot shows generated τ^+ events, and the lower plot shows generated τ^- events. The tau charge is misidentified for a small number of clusters.

shapes in this variable—charge multiplicity does not contain enough charge information. For example, a 3-track cluster can have a total charge of $-3, -1$, $+1$, or $+3$. Without discarding the signal-to-background contrast available in charge multiplicity, this variable will be much more useful with additional charge information.

The simplest way to add charge information is to add a second dimension to the plot, that is, to plot charge multiplicity of the cluster versus the total charge in the 10^o cone. This information is presented in Figure 6.2 for the WTN monte carlo event sample, where the charge of the generated tau is known. It is clear from Figure 6.2 that most positive (negative) taus enter the WTN monte carlo sample with a charge of $+1$ (-1), and with 1 or 3 total charged particles.

The total charge and the total number of charged tracks are misidentified for a small fraction of the WTN monte carlo events—as the total charge and total number of tracks are occasionally misidentified for real taus. That is, the underlying event and additional $\overline{\mathrm{p}}\mathrm{p}$ interactions sometimes contribute tracks which appear to be associated with the tau jet. In addition, as mentioned in Section 4.2, a cluster is sometimes misidentified as

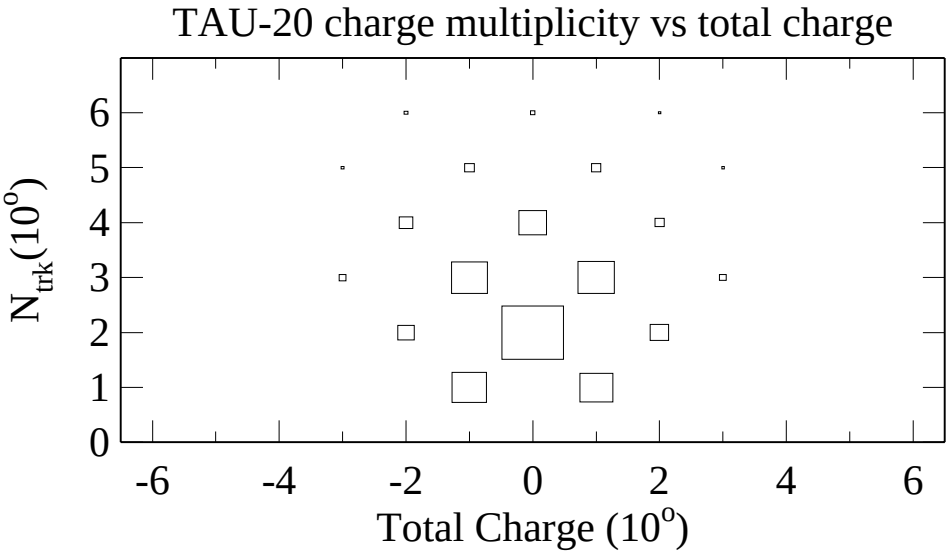


Figure 6.3: Charge multiplicity versus total charge in TAU-20. The charge multiplicity shape difference between taus and background is still evident when comparing TAU-20 to the $W \rightarrow \tau\nu$ monte carlo. Additionally, total charge information is present, allowing QCD charge asymmetries to be naturally accounted for.

a tau, e.g., the QCD jet of a W + jets event is identified as a tau, when the tau is undetected. Events which contain jets misidentified in this manner also populate the plot in Figure 6.2.

To check that the charge information in the 2-dimensional plot distinguishes signal from background, Figure 6.3 presents the TAU-20 event sample, representing the QCD jet background. The charge multiplicity contrast between signal (i.e., Figure 6.2) and background is still clear. It is also clear that the QCD background in Figure 6.3 presents a markedly different shape, when compared with the WTN monte carlo sample, not only in charge multiplicity but also in total cluster charge.

Given this binning, enough information is now present to allow a measurement of the charge asymmetry for $W \rightarrow \tau\nu$ events; with the total charge of the cluster, the fit has information to separate τ^+ from τ^-. Providing total charge information has the additional benefit of ensuring that any charge asymmetry in the QCD jet backgrounds is properly accounted for.

Using both the charge multiplicity and total charge variables also ensures that the

	Total Charge												
N_{trk}	-6	-5	-4	-3	-2	-1	0	$+1$	$+2$	$+3$	$+4$	$+5$	$+6$
1						1		2					
2					3		4		5				
3				11		6		7		11			
4			11		8		9		10		11		
5		11		11		11		11		11		11	
6	11		11		11		11		11		11		11

Table 6.1: This table shows the definition of the charge pattern index as it is derived from the charge multiplicity and total charge. About 1/2 of the entries added into charge pattern index 11 are empty.

fit properly handles tau charge misidentification, e.g., when an underlying event track happens to extrapolate to the tau cluster. Since the WTN monte carlo simulation does model these effects, one can see from Figure 6.2 that most events are identified with the correct charge.

Binning the data samples in all possibilities of charge and charge multiplicity divides the data unnecessarily. Given N tracks, there are $N + 1$ possible total charges. For example, with multiplicity=4, there can be 0 (4), 1 (3), 2 (2), 3 (1), 4 (0) positive (negative) tracks (for a total charge of -4, -2, 0, 2, 4). Therefore, for charge multiplicities 1–6, there are a total of 27 possible "charge patterns."

Of the 27 possible charge patterns, many have low statistics or no events at all. These low statistics bins present no information separating signal and background; overwhelmingly, these bins contain only background. For example, a tau jet uncommonly contains n tracks all of the same sign when $n > 2$. Thus, in the interests of simplifying the fit where possible, 17 bins (about 1/2 of which are empty) are combined into one. Thus, the fit uses 11 charge pattern indices.

These charge patterns form an important part of this analysis. The correspondence of each charge pattern to total charge and the track multiplicity may be seen in Table 6.1. To show what the monte carlo $W \to \tau\nu$ data sample looks like, separately for positive and negative generated taus, Figure 6.4 presents the same data as that plotted in Figure 6.2. As expected, positive and negative taus have quite different shapes in charge pattern.

Charge pattern indices alone are not enough information for the log likelihood fit.

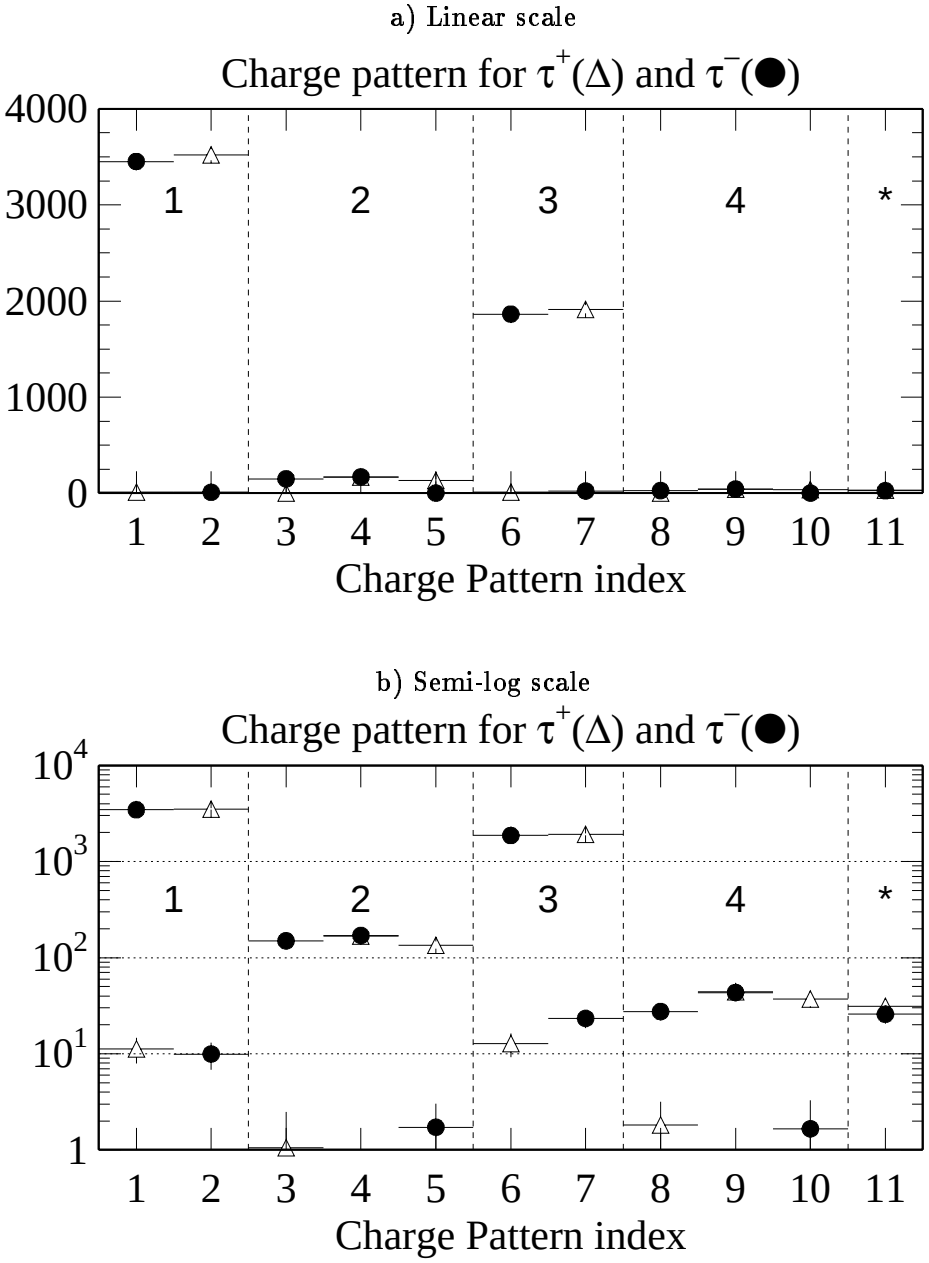


Figure 6.4: Charge patterns for + and − taus in WTN. Negative taus are represented by points, positive taus by triangles. The same information is plotted on a linear scale in a), a semi-log scale in b). As expected, the patterns are different and nearly complementary. The number of tracks for each charge pattern is indicated by the number 1–4 present at the top of the plot. The "*" bin is bin 11 which, as described in the text, combines the low statistics bins.

First, events must be binned in η in order to measure the charge asymmetry as a function of η. Second, one more factor is used to distinguish between the tau signal and the QCD jet background, i.e., the data is binned in isolated and nonisolated samples using the sample division cut described in Section 5.1.2. As described in that section, isolated and nonisolated data samples present markedly different shapes in charge multiplicity, and thus, in charge pattern.

To sum up the variables in which event samples are binned, all event samples used as input to the log likelihood fit are binned in three variables, i, j, and k, defined as:

- i = 1–11; track charge pattern index;
- j = 1,2; j=1 selects isolated clusters (rich in tau signal); j=2 selects non-isolated clusters (rich in QCD jet background); Electroweak backgrounds are split about equally among j=1, j=2;
- k = 1–8; η bins from -1.2 to 1.2.

The first two variables provide information which, statistically, separates the data samples; the last variable allows a measurement of charge asymmetry for bins in η.

6.2 Binned Data Sets and Fit Parameters

Several data sets are binned in the three variables listed above. Some of the binned data sets are data or monte carlo event samples; some are generated by the fit itself. Binned samples from data and monte carlo simulation are

- TAUMON data, D_{ijk};
- WTN monte carlo simulation, $W^{\pm}_{ijk}$;
- TAU-20 data, J_{ijk}, defined in Section 5.2.3;
- a small electroweak background (other W and Z contributions), E_{ijk} determined from monte carlo simulation.

The fit adjusts the relative scale of $W^{\pm}_{ijk}$ and J_{ijk}, as described below, but D_{ijk} is fixed and E_{ijk} is normalized as described in Section 5.1.2. Note that, unlike all other data samples, WTN is divided into positive and negative samples based on the charge of the generated tau.

Compared against TAUMON data, D_{ijk} (containing both $W \to \tau\nu$ and background events), is X_{ijk}, defined as the expected contents from fitting to the histogram bins D_{ijk}. That is, D_{ijk} is the experimental measurement and X_{ijk} is the log likelihood fit prediction of signal plus background. Contributions to the bins of X_{ijk} come from

- a positive tau signal, T^{+}_{ijk}, defined in Section 6.4;
- a negative tau signal, T^{-}_{ijk}, defined in Section 6.4;
- a QCD background Q_{ijk}, defined in Section 6.3;
- an electroweak background, E_{ijk}, as defined above.

Summing all terms for X_{ijk},

$$\boxed{X_{ijk} = T^{+}_{ijk} + T^{-}_{ijk} + E_{ijk} + Q_{ijk}}$$

Some additional quantities used by the fit merit brief discussion here. These quantities are variables whose value is calculated by the log likelihood fit. The parameters used by the fit are

- asymmetry in a given η bin, A_k;
- total number of taus in a given η bin, t_k;
- the number of taus in an η bin, as above, divided into $+$ and $-$ subsamples, $t^{\pm}_{k}$;
- the number of QCD events in the isolated and nonisolated regions, f_j.

The charge asymmetry (A_k) and the total number of taus in each η bin (t_k) and the uncertainty on these values are the ultimate output from the fit. These quantities, as well as $T^{\pm}_{ijk}$ and Q_{ijk}, are defined in the two following sections.

6.3 Modeling the QCD Jet Background

The QCD jet background (Q_{ijk}) must be modeled somehow. Since this is the primary background to $W \to \tau\nu$ physics, it must be modeled carefully. I cannot use an unbiased sample of QCD jets; the cuts applied to generate the tau sample strongly bias those QCD background jets which pass the cuts. Thus, as promised in Section 5.2, I use the TAU-20 event sample to model the QCD background. The choice of TAU-20 as the background model is motivated in Section 5.2.

The TAU-20 event sample contains electroweak background processes just as the TAUMON event sample does. These non-jet-background events (including $W \to \tau\nu$) are statistically removed from the TAU-20 sample as described in Section 5.2.3.

Some cuts applied in the TAUMON event sample are relaxed for the TAU-20 background sample, as discussed in Section 5.2. The relative proportion of events in the isolated and nonisolated samples varies with these cuts. Therefore, the relative normalization of the isolated and nonisolated subsamples from TAU-20 is biased. One way to remove this bias is to tighten those cuts which were relaxed—but the result is a vanishingly small number of events in the TAU-20 event sample. Therefore, the isolated and nonisolated subsamples of TAU-20 are normalized independently in the fit.

TAU-20 is normalized as follows: First, the number of events in the TAU-20 event sample, J_{ijk}, is counted independently for isolated and nonisolated samples. Dividing each bin of J_{ijk} by the appropriate sum (i.e., the sum with the same value for j) scales the total number of isolated TAU-20 events to 1.0 and the total number of nonisolated TAU-20 events to 1.0. To allow the fit to scale each TAU-20 sample, the parameter f_j is defined so that $f(1)$ equals the total number, from the fit, of jet background events in the isolated sample and $f(2)$ equals the total number, again from the fit, of jet background events in the nonisolated sample:

$$Q_{ijk} = f_j \times \frac{J_{ijk}}{\sum_{ik} J_{ijk}}.$$

Thus, J_{ijk} is used to describe the *shape* of the QCD jet background in charge pattern and η, and f_j is used to measure the normalization or amount of the QCD jet background in the fit.

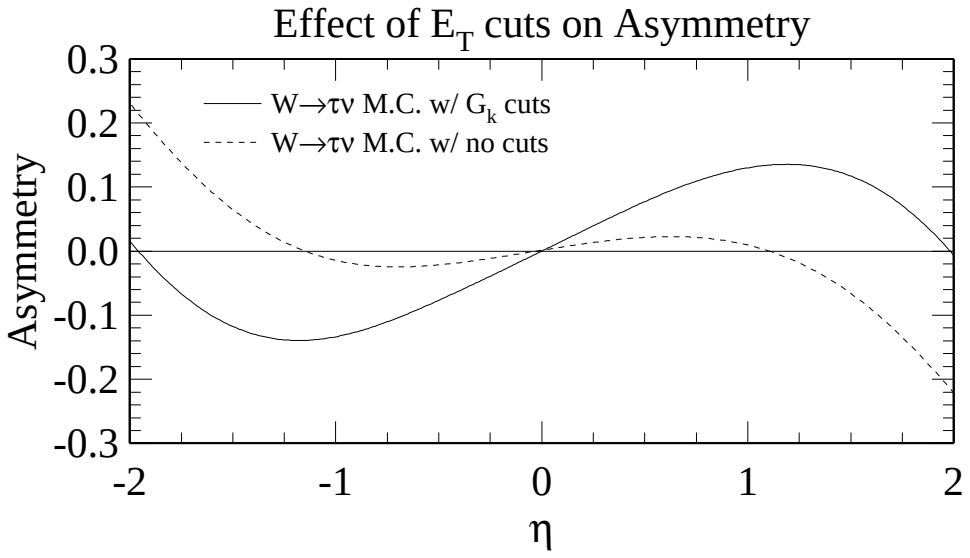


Figure 6.5: Tau asymmetry for different thresholds, from the $W \rightarrow \tau\nu$ monte carlo event sample. The detector and trigger simulation are not used here. The solid line represents the tau asymmetry for monte carlo events passing the G_k cuts listed in the text; the dashed line represents the tau asymmetry for *all* hadronic tau events in the monte carlo sample.

6.4 Modeling the $W \rightarrow \tau\nu$ Signal

To determine the amount of positive and negative tau signal present, I need the expected shapes of the charge pattern distribution for each charge. I can measure this shape using the WTN monte carlo simulation. WTN necessarily includes the physics effects which generate charge asymmetry—this asymmetry should not be included in the fit—but the event selection cuts introduce charge biases (asymmetries) which *should* be included in the fit.

At the monte carlo generator level, before the detector and trigger simulation have propagated all particles through the detector, no charge biases are present except the charge asymmetry I am measuring. Thus, normalizing each eta bin, positive and negative taus separately, by the generator level event counts, $G_k^{\pm}$, solves this problem. The measured charge asymmetry, however, depends strongly on whatever E_T cuts are applied. (See Figure 6.5 for a demonstration of this effect.) Thus, I apply several cuts at

the generator level to measure $G_k^{\pm}$:

- hadronic tau decay,
- generated tau $E_T > 25$ GeV,
- $W \to \tau\nu$ neutrino $E_T > 25$ GeV.

These are the same cuts as those applied to the CDF electron and muon charge asymmetry measurements [13]; thus, these analyses may be compared directly. These cuts assume perfect detector resolution, for simplicity.

Since many more WTN events pass the generator-level cuts than the full event simulation cuts, due to detector and trigger inefficiencies, each bin must be normalized by the ratio $R_{\rm WTN}$:

$$R_{\rm WTN} = \frac{\sum_k (G_k^+ + G_k^-)}{\sum_{ijk} (W_{ijk}^+ + W_{ijk}^-)}.$$

Using this normalization, the number of taus calculated by the fitting program is scaled properly:

$$T_{ijk}^{\pm} = t_k^{\pm} \times \frac{W_{ijk}^{\pm}}{G_k^{\pm}} \times R_{\rm WTN},$$

where, remember, $t_k^{\pm}$ is the count of positive and negative taus in each η bin from the fit.

The quantities $t_k^{\pm}$ can be calculated in two different ways. One choice uses the twelve $t_k^{\pm}$ as parameters in the fit and then calculates A_k from these values. The other choice instead uses six $t_k = t_k^+ + t_k^-$ and six A_k, thus calculating A_k and the uncertainty on A_k directly from the fit. With this choice, the twelve $t_k^{\pm}$ used above are calculated via

$$t_k^+ = t_k \times \frac{1 + A_k}{2},$$

$$t_k^- = t_k \times \frac{1 - A_k}{2}.$$

This fit uses the latter alternative, as the uncertainties and correlations of A_k are then calculated directly by the fit. To summarize, the number of taus, positive and negative, is calculated by the fit from:

$$T_{ijk}^+ = t_k \times \frac{1 + A_k}{2} \times \left(\frac{W_{ijk}^+}{G_k^+} \times R_{\rm WTN} \right),$$

$$T^{-}_{ijk} = t_k \times \frac{1 - A_k}{2} \times \left(\frac{W^{-}_{ijk}}{G^{-}_{k}} \times R_{\mathrm{WTN}} \right).$$

6.5 Degrees of Freedom

Counting data bins, there are 2×8 histograms each with 11 bins, for a total of 176 bins. In addition to these degrees of freedom are 18 unknown parameters:

- A_k; asymmetry vs. η; 8 total;
- t_k; tau signal vs. η; 8 total;
- f_j; qcd scale for isolated and non-isolated histograms; 2 total.

The fit therefore has $(176 - 18) = 158$ degrees of freedom.

As described in Appendix A, a maximum log likelihood fit for histogram bins with Poisson statistics minimizes

$$\chi^2 = -2 \ln \frac{\mathcal{L}}{\mathcal{L}_0} = -2 \sum_{i,j,k} (D_{ijk} \log X_{ijk} - D_{ijk} \log D_{ijk} + D_{ijk} - X_{ijk}).$$

Note that since this is a *normalized* log likelihood fit, the value of the fit can be considered equivalent to a χ^2 [42, 43], and is expected to be distributed as χ^2 defined in the usual manner. Given 158 degrees of freedom, therefore, one expects the normalized log likelihood to be $158 \pm \sqrt{2 \times 158} = 158 \pm 18$.

Calculating the "folded" asymmetry is equivalent to adding four constraints to the fit:

- $A_1 = A_8$,
- $A_2 = A_7$,
- $A_3 = A_6$,
- $A_4 = A_5$.

Thus, one expects the normalized log likelihood for the "folded" asymmetry measurement to be 162 ± 18.

6.6 Testing the Fit

To ensure the log likelihood fit is a stable fit against Poisson fluctuations, given the level of statistics present in this analysis, a "toy" monte carlo simulation is used to fit 1000 iterations of Poisson fluctuated toy data. The electroweak backgrounds are simulated in the toy using E_{ijk}. For the $W \to \tau\nu$ signal in the toy, arbitrary t_k and A_k shapes are chosen restricted by the desire to roughly imitate the expected signal and sample size. The nonisolated jet background is simulated in the toy using the TAUMON nonisolated sample. The isolated jet background is simulated using the TAUMON nonisolated sample, scaled up by a factor of 1.15 to roughly simulate the slightly increased statistics in that region.

For the purposes of the toy fit, an approximate resemblance to the expected data is sufficient. Each of the 1000 loops through the toy monte carlo simulation generates the same D_{ijk} as described above, then fluctuates the number of events in each bin using Poisson statistics. The normal log likelihood fit, as described in the rest of this Chapter, is then applied to this Poisson-fluctuated D_{ijk}.

After the 1000 loops, the asymmetry and number of taus used as input to the toy are compared with the distribution of these quantities coming from the fit. For all variables used as input to the toy D_{ijk}, the correspondence is excellent. The mean of each distribution is much less than one standard deviation from the value used as input and within the uncertainty of the mean. That is, for example, if the 1000 loops result in a distribution for the fit value for asymmetry in η bin number 3 $(-0.6 < \eta < -0.3)$, A_3, with a mean of m and a width of σ, then $m \pm \sigma/\sqrt{1000}$ well represents the value used as input for A_3 in the toy monte carlo simulation.

This indicates the fit is consistent and stable; that the measurement and uncertainties produced by the fit are reliable.

6.7 Results of the Fit

Applying the fit to the data samples, applying no corrections, the resulting log likelihood is 155, well within the 158 ± 18 expected from degrees of freedom. Calculating the folded

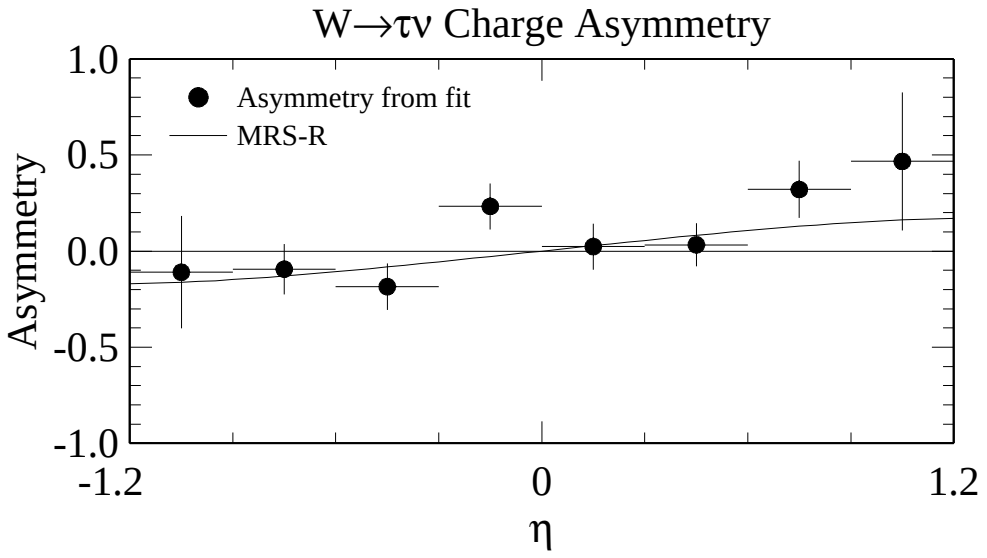


Figure 6.6: Uncorrected asymmetry from fit. The full physical range of asymmetry, -1 to $+1$, is shown.

fit, the log likelihood is 160, agreeing with the expected 162 ± 18.

Results of the fit with corrections and systematic uncertainties indicated are given in Chapter 8. The parameters and uncertainties measured by the fit, uncorrected, are shown in Table 6.2 (for the unfolded plot, $-1.2 < \eta < 1.2$) and Table 6.3 (for the folded plot, $0 < \eta < 1.2$). The asymmetry from the fit is shown in Figure 6.6, compared to a theoretical prediction; the folded asymmetry from the fit is presented in Figure 6.7. Only the statistical uncertainty is presented in these figures.

To compare the fit to data, Figure 6.8 presents the results of the unfolded asymmetry fit as a function of the charge pattern index. The points show the data; the three histograms show each contribution to X_{ijk}, QCD background, electroweak background, and the sum $T^+_{ijk} + T^-_{ijk}$. The data in this histogram are summed over all η bins. The same information for each η bin is present in Figure 6.9 for negative η and in Figure 6.10 for positive η.

Finally, Figure 6.11 shows the number of tracks in the 10° cone from the fit, with all contributions shown.

Minuit [44] (see also Appendix A), while calculating the fit, also calculates the

A_k	Asymmetry		E+	E−
A_1	$-0.10 \pm$	0.28	0.25	0.24
A_2	$-0.06 \pm$	0.13	0.14	0.14
A_3	$-0.17 \pm$	0.12	0.14	0.14
A_4	$0.25 \pm$	0.12	0.13	0.13
A_5	$0.04 \pm$	0.12	0.13	0.13
A_6	$0.05 \pm$	0.11	0.12	0.12
A_7	$0.35 \pm$	0.14	0.15	0.16
A_8	$0.50 \pm$	0.33	0.28	0.33
t_k	N_{tau}		E+	E−
t_1	$71.77 \pm$	16.22	18.64	16.29
t_2	$64.12 \pm$	10.27	9.33	8.67
t_3	$54.39 \pm$	7.87	8.03	7.51
t_4	$65.07 \pm$	8.82	8.81	8.28
t_5	$63.79 \pm$	9.46	8.93	8.36
t_6	$70.27 \pm$	9.88	8.69	8.15
t_7	$49.27 \pm$	8.43	8.51	7.83
t_8	$39.03 \pm$	13.17	14.49	12.14
f_j	QCD isolation shape		E+	E−
f_1	$170.25 \pm$	20.71	21.29	20.05
f_2	$155.28 \pm$	15.14	13.76	13.06

Table 6.2: Results of the fit, uncorrected. E+ and E− are the upper and lower MINOS uncertainties. The quoted uncertainty on each parameter is the "parabolic" uncertainty measured by the curvature at the fit value.

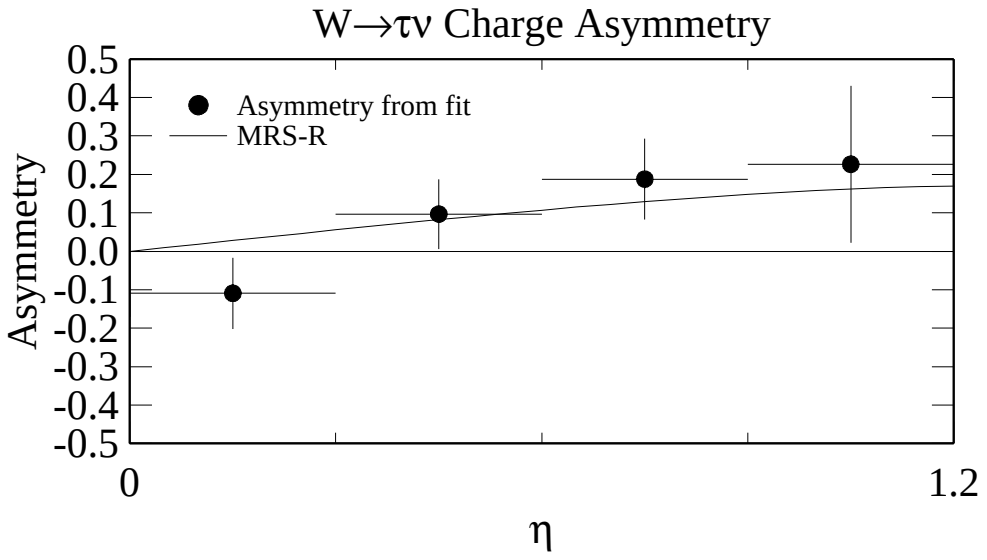


Figure 6.7: Uncorrected folded asymmetry from fit.

A_k	Asymmetry		E+	E−
A_1	$-0.23 \pm$	0.20	0.21	0.20
A_2	$-0.19 \pm$	0.11	0.11	0.10
A_3	$-0.10 \pm$	0.09	0.09	0.09
A_4	$0.11 \pm$	0.09	0.09	0.09
A_5	$-0.11 \pm$	0.09	0.09	0.09
A_6	$0.10 \pm$	0.09	0.09	0.09
A_7	$0.19 \pm$	0.11	0.11	0.10
A_8	$0.23 \pm$	0.20	0.21	0.20
t_k	N_{tau}		E+	E−
t_1	$71.63 \pm$	18.04	19.28	16.84
t_2	$65.19 \pm$	9.25	9.60	8.91
t_3	$54.66 \pm$	7.80	8.07	7.53
t_4	$62.85 \pm$	8.17	8.43	7.92
t_5	$61.60 \pm$	8.35	8.62	8.09
t_6	$70.17 \pm$	8.43	8.70	8.17
t_7	$49.97 \pm$	8.39	8.75	8.04
t_8	$37.21 \pm$	13.69	14.93	12.48
f_j	QCD isolation shape		E+	E−
f_1	$170.97 \pm$	20.71	21.36	20.07
f_2	$158.33 \pm$	13.54	13.88	13.21

Table 6.3: Results of the fit, uncorrected. E+ and E− are the upper and lower MINOS uncertainties. The quoted uncertainty on each parameter is the "parabolic" uncertainty measured by the curvature at the fit value.

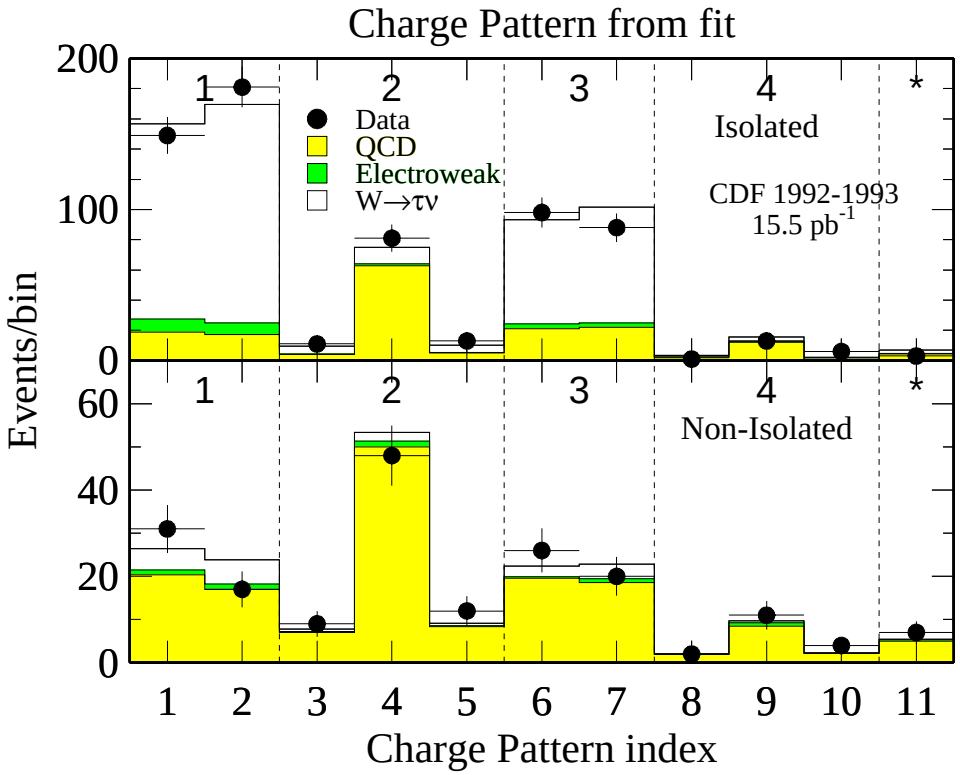


Figure 6.8: Fit Charge Pattern distributions, summed over all η bins. Each contribution to X_{ijk} is shown separately.

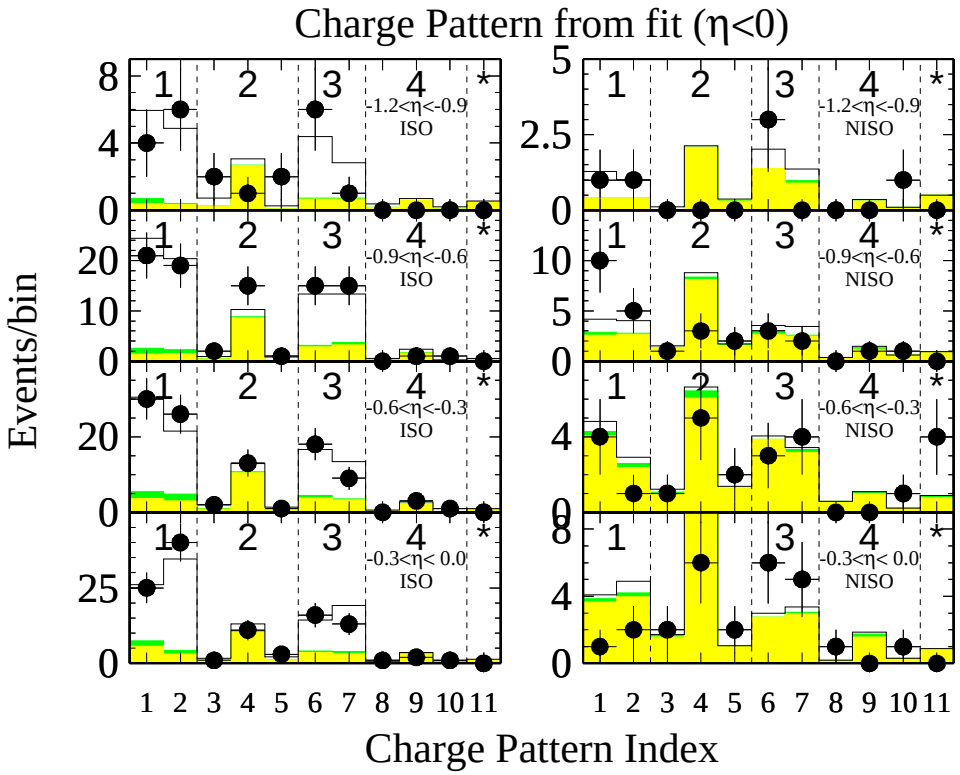


Figure 6.9: Fit Charge Pattern distributions for each η bin at negative η. All information available to the fit is shown in this plot, except T^{+}_{ijk} and T^{-}_{ijk} are added in each charge pattern and η bin. The plots on the left are isolated; the plots on the right are non-isolated.

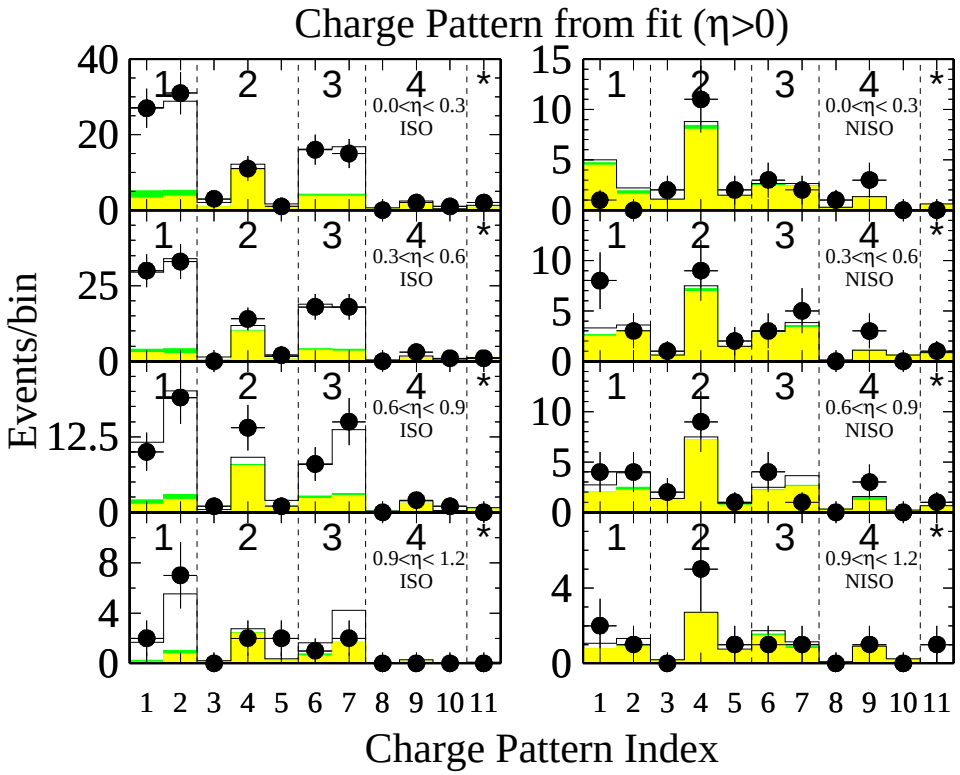


Figure 6.10: Fit Charge Pattern distributions for each η bin at positive η. All information available to the fit is shown in this plot, except T^{+}_{ijk} and T^{-}_{ijk} are added in each charge pattern and η bin. The plots on the left are isolated; the plots on the right are non-isolated.

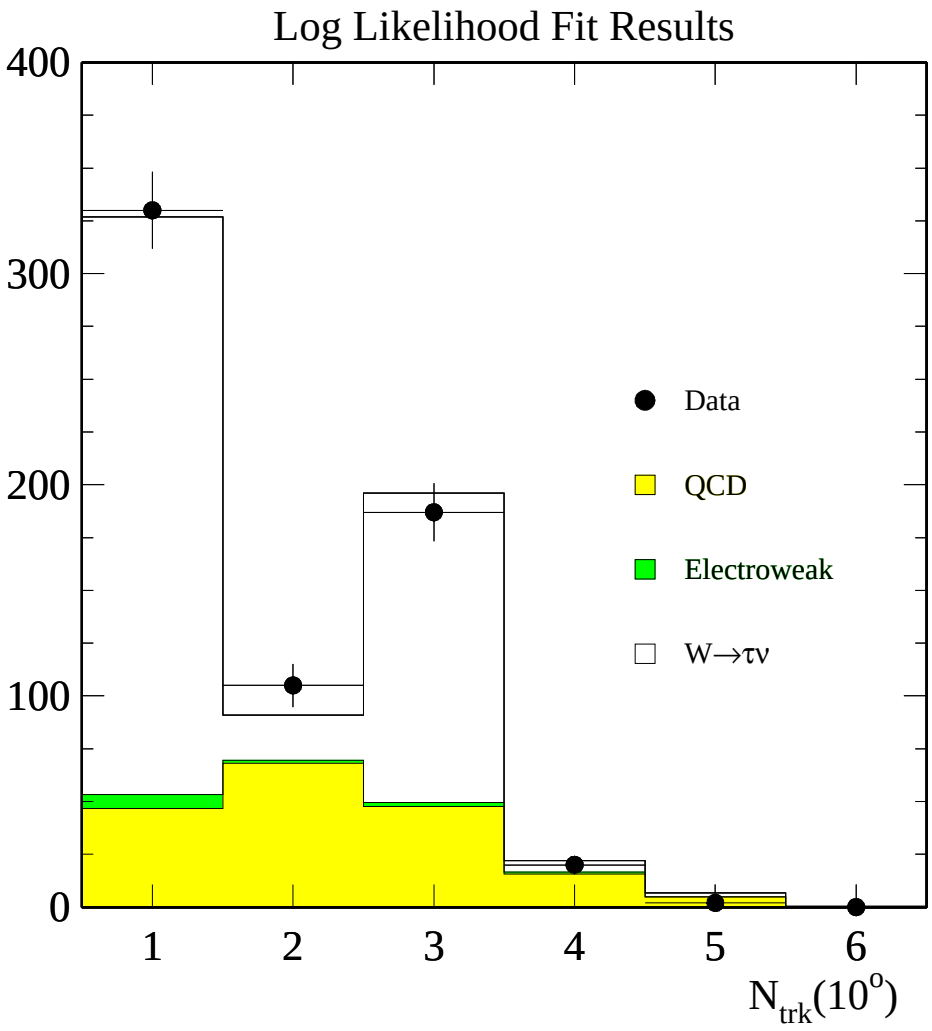


Figure 6.11: Number of tracks in 10° cone for uncorrected fit, showing all contributions.

covariance matrix for the uncertainties of the asymmetry in each η bin. The correlation matrix can be calculated from the covariance matrix. The correlation matrix is given below with η increasing left to right, top to bottom:

$$\begin{pmatrix}
1.000 & & & & & & & \\
0.001 & 1.000 & & & & & & \\
- & 0.001 & 1.000 & & & & & \\
-0.001 & -0.002 & -0.025 & 1.000 & & & & \\
- & - & 0.002 & -0.010 & 1.000 & & & \\
- & -0.001 & -0.001 & 0.006 & -0.001 & 1.000 & & \\
0.001 & 0.001 & -0.005 & 0.022 & -0.001 & - & 1.000 & \\
0.002 & -0.002 & -0.002 & 0.001 & -0.004 & - & -0.006 & 1.000
\end{pmatrix}$$

Elements marked by "—" have a negligible correlation, less than 0.1%. All correlations are small; only three off-diagonal bins show a correlation of 1% or larger.

Taking the square root of the diagonal elements of the covariance matrix, the uncertainty of the asymmetry of each η bin is given below with η increasing left to right:

$$\begin{pmatrix} 0.277 & 0.129 & 0.121 & 0.122 & 0.120 & 0.111 & 0.144 & 0.325 \end{pmatrix}$$

Chapter 7

Corrections and Systematic Uncertainties

We will now discuss in a little more detail the Struggle for Existence.
— Charles Darwin [45]

Physics analyses in high energy experimental physics and many other fields of science are complicated. To quote just statistical uncertainty as the total uncertainty of a measurement is to understate the uncertainty. For example, the CDF detector energy scale is known to a finite accuracy. When measuring the mass of a particle, uncertainty in the energy scale adds uncertainty to the mass measurement. This additional uncertainty—called systematic uncertainty—must be estimated.

The systematic uncertainty of a measurement reflects its stability with respect to the analysis criteria and includes all uncertainties other than the pure statistical uncertainty of the measurement. All corrections applied or studied in this analysis are, due to the way the corrections are identified and measured, limited in the accuracy with which the correction is known. Thus, I study how changing the correction changes the measured asymmetry. For example, since the $W \to \tau\nu$ charge asymmetry depends on the E_T scale, which is known to a finite accuracy, I study how the asymmetry varies as the E_T scale is changed.

Finally, some event selection cuts, assumptions embodied in the analysis or monte carlo programs, or differences between event samples introduce a bias to the measurement. Such biases, and thus, their effect on the asymmetry measurement, must be estimated.

7.1 Overview

The important systematic effects in this analysis can be neatly categorized into three groups: accuracy of the $W \rightarrow \tau\nu$ monte carlo simulation (described in Chapter 5) used in the log likelihood fit, differences between the QCD jet background model (i.e., TAU-20, also described in Chapter 5) and the actual QCD jet background, and assumptions in the event selection cuts and method of the analysis.

First, I study systematic effects related to the $W \rightarrow \tau\nu$ monte carlo simulation. The simulation, for example, assumes that the $\overline{\mathrm{p}}\mathrm{p}$ interactions are distributed along the z axis (along the beamline) as a Gaussian centered at the center of the CDF detector. Due to changing accelerator conditions during Run 1a, both parts of this assumption are wrong. Although the distribution of the z position of $\overline{\mathrm{p}}\mathrm{p}$ interactions, in itself, should not introduce a charge bias, the possible interaction of this bias and others merits study.

Since jets from hadronic tau decay are narrow objects in the calorimeter, especially when compared with QCD jets, it is important that QFL (i.e., CDF's detector simulation) well models the development of pion showers in the calorimeter. In fact, known deficiencies in QFL pion shower modeling (which are not trivial to correct, either in QFL or after its processing) influence the asymmetry measurement.

The Level 1 and Level 2 triggers are both inefficient; some events which pass all of the other event selection cuts are not accepted by the hardware triggers. This inefficiency must be properly modeled; biases present in the hardware trigger should be present in the simulation of the trigger. Those which are not modeled well must be studied as a potential source of bias, and thus, systematic uncertainty.

Next, I study the QCD jet background, which is troublesome to model; the event selection cuts for $W \rightarrow \tau\nu$ remove nearly all QCD jet background events and strongly bias those which remain to be "tau-like." Relaxing these cuts to increase the statistics in the QCD jet background sample, thus, may change its distributions of events away from the distributions of the actual QCD jet background; not relaxing the cuts results in too small a QCD jet background sample to be statistically useful. I study the bias

introduced by relaxing event selection cuts for the TAU-20 event sample.

Finally, I study the several remaining systematic effects. The finite resolution of the calorimeter E_T scale influences the accuracy with which the asymmetry can be measured. The log likelihood fit relies on the electroweak backgrounds to be scaled to the proper number of events. Since this scale is known to a finite accuracy, it introduces a potential bias. In addition, the distributions of $W \to \tau\nu$ events are known only to the statistical accuracy of the generated event sample. This statistical uncertainty introduces a systematic uncertainty to the asymmetry measurement.

This analysis has many potential sources of systematic uncertainty and bias, but the systematic uncertainties of the asymmetry measurement are all smaller than the statistical uncertainty. Due to the low efficiency for identifying $W \to \tau\nu$ events once the QCD backgrounds are reduced to an acceptable level, the statistical uncertainty in each asymmetry bin is of order 12%, while all systematic uncertainties are of order 2%.

7.2 $W \to \tau\nu$ Simulation Corrections

First, I discuss how I measure each of the several $W \to \tau\nu$ simulation corrections. Once all the corrections are defined, I discuss the systematic uncertainty, in asymmetry, due to these corrections.

7.2.1 Z Vertex Shape and Offset

The $W \to \tau\nu$ monte carlo simulation does not simulate the z distribution of $\overline{\mathrm{p}}\mathrm{p}$ interactions (i.e., the z vertex distribution) seen in the data. The functional form for $d\mathcal{L}/dz$ is complicated [46], but for a given set of beam conditions, it is approximately Gaussian. For Run 1a, the Tevatron was set to one of two sets of beam conditions (see Ref [46]). Thus, I fit the data z vertex shape to the sum of two Gaussian functions. The $W \to \tau\nu$ simulation, however, assumes a single Gaussian shape to the z distribution.

In addition, the $W \to \tau\nu$ simulation assumes that $\overline{\mathrm{p}}\mathrm{p}$ interactions are centered (in z) in the CDF detector, which is not the case for Run 1a. The average collision position is actually at about $z = -2$ cm.

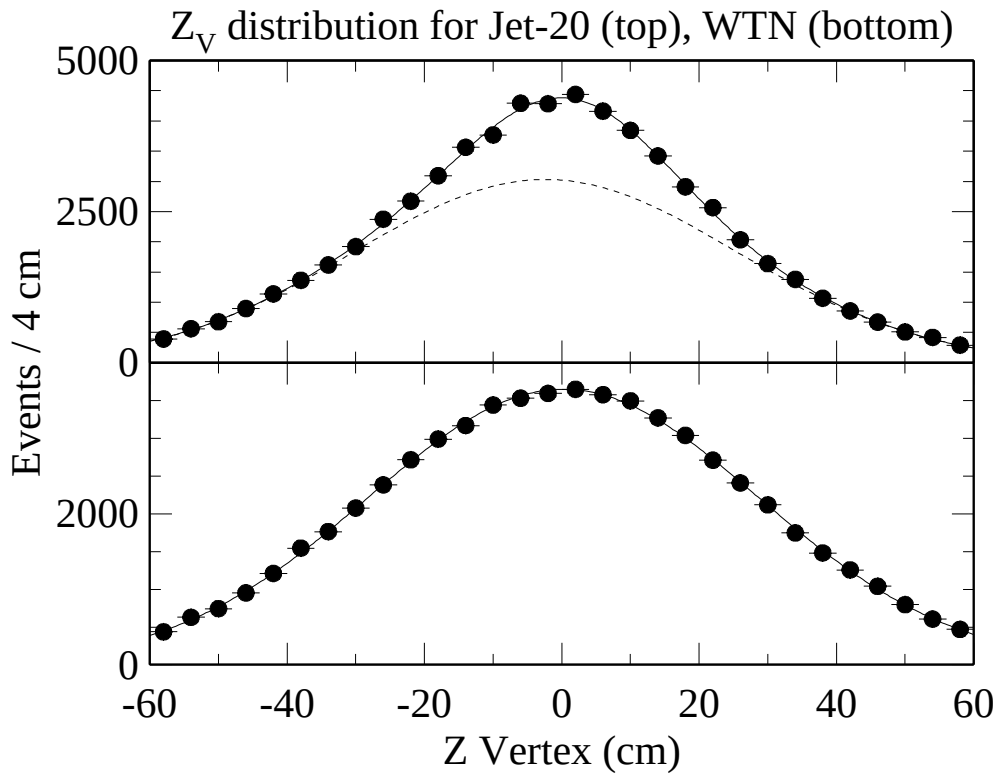


Figure 7.1: Z_V shape for data and simulation. The data well fits the sum of two Gaussians. The dotted line shows the wider of the two Gaussians for the data. The simulation z vertex distribution is Gaussian.

I use a form for the Gaussian distribution which is appropriate for a binned quantity:

$$G = \frac{A}{\sqrt{2\pi}}\frac{\delta}{\sigma}e^{-\frac{1}{2}(\frac{Z_v - m}{\sigma})^2}$$

where A is the total area (or number of events) under the Gaussian, σ is the width in z, δ is the bin width—4 cm here—and m is the Gaussian's mean.

I fit a single Gaussian to the $W \rightarrow \tau\nu$ simulation and the sum of two Gaussians to the data z vertex distribution, then normalize the curves to the same area. The fits are presented in Figure 7.1; the χ^2 of each fit is approximately one per degree of freedom, indicating a reasonable fit. The parameters for the two fits are listed in Table 7.1. Since there is no expected dependence of the z vertex distribution on the kind of collision, I used a large sample of QCD jets (the JET-20 sample) to have a high-statistics measurement of the z vertex distribution in the data.

Z vertex distribution parameters

	Area (events)	Mean (cm)	Sigma (cm)
Data	52846 ± 1710	-2.40 ± 0.20	27.8 ± 0.4
	11615 ± 1776	1.13 ± 0.43	13.6 ± 0.8
$W \rightarrow \tau\nu$ simulation	65080 ± 214	0.22 ± 0.10	28.5 ± 0.1

Table 7.1: z vertex distribution parameters

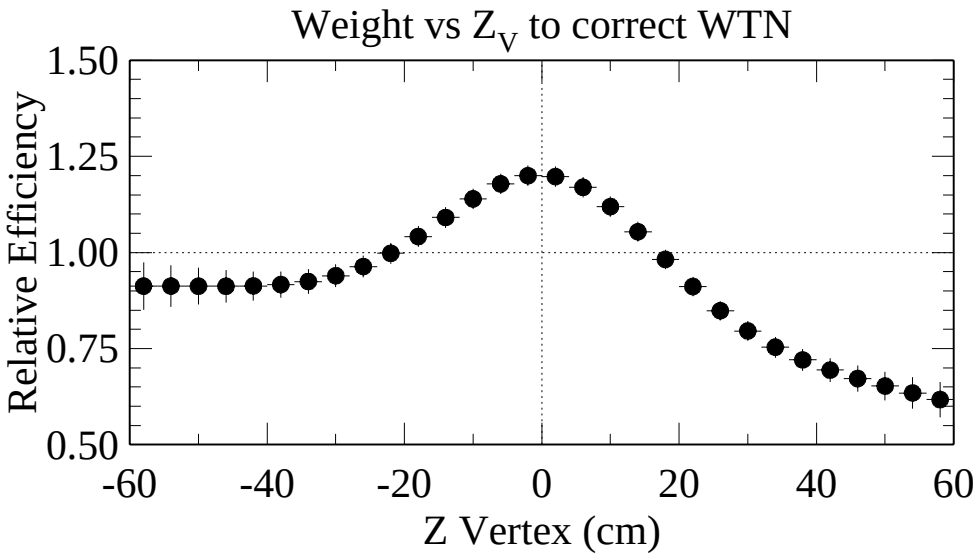


Figure 7.2: WTN weight vs. Z_V

Although the differences between simulation and data z vertex distribution, alone, do not change shower shapes or pion efficiencies, it is possible that this difference between simulation and data can interact with event selection or other simulation details to introduce a bias. Thus, while testing systematic effects, I use the ratio of the z vertex distributions, as a function of z, as a "weight." That is, each event is given a statistical weight due to its z vertex position. If the weight is equal to 2.0, then it is as if *two* events occurred at this z vertex. Applying this weight to all $W \to \tau\nu$ simulation events restores the correct z vertex distribution. The weight, as a function of z, is shown in Figure 7.2. The statistical uncertainty of this weight is easily calculated, assuming Poisson statistics.

Representing the two data Gaussians as $G_{D1}(z)$ and $G_{D2}(z)$ and the $W \to \tau\nu$ Gaussian as $G_W(z)$, an event's weight, given z, is

$$w = \frac{G_{D1}(z) + G_{D2}(z)}{G_W(z)}$$

and the uncertainty on that weight is given by

$$s = w\sqrt{\frac{1}{G_{D1}(z) + G_{D2}(z)} + \frac{1}{G_W(z)}}.$$

I discuss the systematic uncertainty due to this correction in Section 7.5.

7.2.2 Single Pion Response

QFL is well-tuned to simulate the averaged response of the calorimeter to hadronic jets, but the simulation does not model well hadronic showers from *single*, isolated pions. Few CDF analyses select such events, so no such tuning of QFL has been necessary. The details of pion showers matter for tau physics. That is, if these details are not properly simulated, the monte carlo simulation may not accurately measure any possible charge bias in the trigger and analysis cuts.

What deficiencies exist in single pion simulation? Most critical is that the shower width, measured in towers, is poorly simulated. QFL underestimates the number of single-tower clusters and underestimates the shower width of multiple-tower clusters [38]. There is no simple way to correct the simulation of pion shower shapes and

widths. Thus, rather than correct the pion shower shapes in a detailed manner, I studied the efficiency of single pions in variables of direct interest to this analysis and applied corrections in those variables.

7.2.3 Correction for Pion Efficiency Near Cracks

QFL poorly simulates the way pion showers share energy among towers; its pion shower simulation overestimates the energy measured for a pion which showers into or through a crack between calorimeter towers. This effect is especially significant for narrow hadronic jets which traverse a gap in η between two trigger towers; remember that the Level 1 and Level 2 triggers measure calorimeter energy in pairs (in η) of physical calorimeter towers. Since taus are central objects due to tracking requirements, only the trigger towers in the central region need study; there are 10 trigger towers in the central region ($-1.0 < \eta < 1.0$).

Narrow jets which traverse a trigger tower gap present a clear loss of efficiency due to the Level 1 trigger; jets traversing a gap between two physical towers in the middle of a trigger tower do not show this loss of efficiency. The Level 1 trigger applies a stringent requirement to the tau candidate jet: It must have at least one trigger tower with EM E_T greater than 6 GeV or HAD E_T greater than 8 GeV. This means, for example, a 24 GeV tau cluster two *physical* towers wide where the cluster's energy is divided between the two towers as 5 GeV EM and 7 GeV HAD E_T each, might pass or fail the Level 1 trigger, depending on it's η position. If the two physical towers are in separate trigger towers, the event will fail the Level 1 trigger.

To model the Level 1 trigger efficiency accurately, QFL must correctly partition the cluster energy between towers and cracks (where it is "lost" energy). The Level 1 trigger efficiency depends on the x-local and z-local (as defined in Section 3.3) at which each particle enters the trigger tower, and on the direction the particle is travelling when it enters the trigger tower. In the data, a particle entering the calorimeter near a trigger tower crack, but travelling away from the crack, has a much higher Level 1 trigger efficiency than a particle entering the calorimeter at the same position but travelling toward the crack. QFL does not simulate *this* effect at all.

Given the statistics-dominated uncertainty of this analysis, a qualitative correction allows me to measure this systematic uncertainty with enough accuracy. Thus, rather than attempt to resimulate the pion energy sharing between towers, I "correct" the overall efficiency for clusters near a crack.

I compare the $W \rightarrow \tau\nu$ simulation with TAU-20, although I see the same behavior when I instead compare with JET-20. Since the CDF detector has 10 central trigger towers, 5 on each side of $z = 0$, I fold the event samples around $z = 0$ to double the statistics of the parameterization. Then, for each tau cluster candidate, I identify the z vertex (Z_V) of the seed track (the highest momentum charged track associated with the cluster) and the z ($Z_{\rm cal}$) at which the track enters the EM calorimeter. In the $W \rightarrow \tau\nu$ simulation, the Level 1 efficiency depends only on $Z_{\rm cal}$, but in the data, the efficiency clearly depends on both variables. In fact, the efficiency in the data is approximately constant along lines of $1/2Z_V - Z_{\rm cal}$.

Plotting both TAU-20 and WTN as a function of $1/2Z_V - Z_{\rm cal}$, it is evident that multiplying QFL's efficiency by a quadratic function of about 1.1 in the middle of a trigger tower (for $Z_V = 0$), dropping to about 0.6 at the edge of each trigger tower, corrects QFL's overestimate of the efficiency near trigger tower cracks without changing the total number of events. The χ^2 for the quadratic fits is approximately 1 per degree of freedom for all central trigger towers, indicating that this parameterization is an adequate model of this effect.

I discuss the systematic uncertainty due to this correction in Section 7.5.

7.2.4 Level 2 Cluster Shape Bias

Due to a geometric bias in the Level 2 trigger calorimeter cluster finding, the efficiency of matching a CFT track to a Level 2 cluster is charge dependent. This is not a failure of the CFT to properly find a track; it is an effect of the way that a CFT track is matched to a Level 2 calorimeter cluster. As explained in Section 3.7.5, CFT tracks are matched (extrapolated) to calorimeter clusters using a lookup table; the two indices into the lookup table are the CFT track P_T bin and track's ϕ as it exits the CTC. The lookup table returns a 24-bit bitmask with one bit set for each tower 0–23 to which this

track may extrapolate. As each P_T bin includes a range of momenta, the lookup table allows tracks that extrapolate to ϕ boundaries between towers to match both towers.

Since the bitmask returned from the lookup table is a discrete quantity, it must be compared to a discrete quantity calculated for a Level 2 cluster. Although the ϕ centroid of each Level 2 cluster is digitized, and to this extent is a discrete quantity, it also is a quantity ranging from 0–2π, not 0–23, and is thus simply not useful. Since the ϕ *seed tower* index, 0–23, is stored for each cluster, the track extrapolated ϕ is compared to this quantity instead.

The seed tower for a Level 2 cluster is the first tower with E_T greater than the seed tower threshold (3 GeV) during clustering. For a ring of towers at a given η, the cluster finder checks each tower in turn, starting at ϕ index = 0 and continuing through increasing ϕ to ϕ index = 23. Thus, if a Level 2 cluster is two towers wide in ϕ, both towers with E_T greater than the seed threshold, then the seed tower is always the tower at lower ϕ. The one exception is that clusters straddling $\phi = 0$ have a ϕ seed index of 0. On average, therefore, the seed tower ϕ is biased toward lower ϕ in Level 2 clusters.

Positively charged tracks curve with increasing ϕ; negatively charged tracks curve with decreasing ϕ. Thus, as Figure 7.3 shows, track matching to the ϕ seed tower introduces an inefficiency for positive taus. Indeed, the inefficiency is present for other single-track hadronic jets as well. Jets including more than one charged track provide more opportunities to match the seed ϕ index, so this inefficiency is diluted for clusters with no negative track and more than one positive track.

I reduce the trigger charge bias by rejecting events where the Level 2 tau cluster contains more than one tower in ϕ but only one tower in η above the clustering seed threshold. Rejecting these clusters removes the most egregiously biased clusters from the sample. The efficiency difference between positive and negative clusters which remain is smaller than a few percent, and given the statistics is consistent with no efficiency difference.

Since charge biases potentially remain, if much smaller, I take 1/4 of the amount that the asymmetry changes in the fit as the uncertainty due to this effect, once the most biased cluster are rejected. Given the relative scale of the systematic and statistical

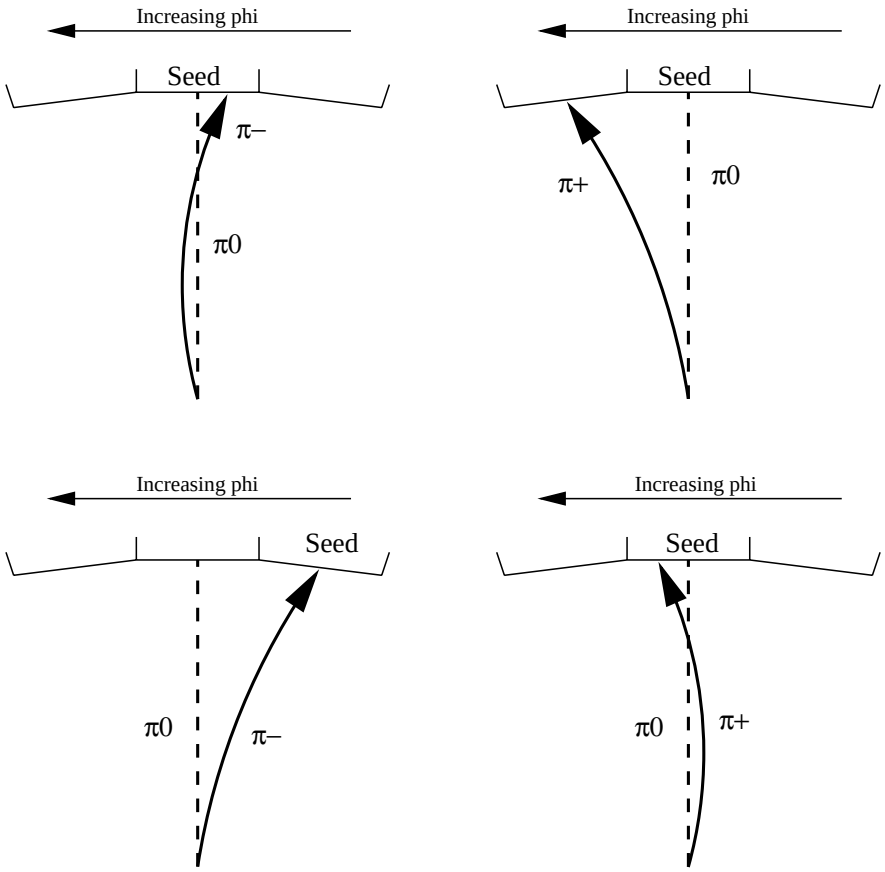


Figure 7.3: CFT track matching. Assuming the π^0 and $\pi^{\pm}$ each deposit at least 3 GeV E_T in the calorimeter, the seed tower is always the tower at lowest ϕ. For π^-'s, the pion always extrapolates to the tower at lowest ϕ, but π^+'s sometimes extrapolate to a tower other than the seed tower in ϕ. This effect is responsible for an inefficiency for positive τ's relative to negative τ's.

uncertainties, this is an acceptable estimate of the uncertainty. For two η bins, the asymmetry changes negligibly. The systematic uncertainty for the asymmetry in each η bin is

$$\left(\begin{array}{cccccccc} 0.028 & 0.014 & - & 0.020 & 0.005 & 0.004 & - & 0.016 \end{array} \right)$$

where η increases from left to right and the two bins marked "—" have negligible uncertainty due to this effect.

7.2.5 CFT Simulation Overview

The CFT (described in Section 3.7.5) is difficult to simulate well since the CFT track-finding efficiency depends on so many variables. One such variable is the efficiency of measuring prompt hits; this efficiency depends on the total distance the track travels in a given CTC cell. Remember that the CFT allows a finite and limited (compared with the offline) time for a "prompt hit," and a slightly longer time for "delayed hits." A well-measured track leaves a minimum of one prompt hit and two delayed hits in each superlayer, but not all tracks are well measured.

To most properly simulate the CFT, one must simulate hits in the CTC, including noise hits. The QFL detector simulation does not simulate tracks by simulating hits in the CTC; instead, it assumes 100% track-finding efficiency and directly generates track information from charged particle simulation information, although QFL does "smear" the track parameters by the known track parameter resolutions. In order to simulate track hits in the CTC, an alternate (and CPU-intensive) detector simulation must be used, at least to simulate tracking information. For these simulated hits, the Level 2 data path must then be simulated with its strict timing limits for prompt and delayed hits; *then* the full CFT simulation can process those prompt and delayed hits.

For this analysis, the CFT was simulated more naively in a parameterized manner, described in Appendix B. The following two sections present effects not accounted for in that simulation.

7.2.6 CFT vs. $\cot\theta$ Correction

As mentioned above, the CFT efficiency for finding a track depends, in part, upon the track's path length through each CTC cell. Thus, as the absolute value of $\cot\theta$ increases, and a track's path length through the CTC increases, the CFT should become more efficient. (Remember that θ is the polar angle from the z axis.) In addition, the strength of this effect may depend on the track's charge, especially for lower-momentum tracks.

Since positive tracks curve with increasing ϕ, they follow a longer path past each CTC sense wire than negative tracks do. For high-momentum tracks, the difference is negligible. For tracks of low momentum, especially as the momentum drops toward 3 GeV (the lowest momentum tracks the CFT is designed to detect), positive tracks follow a longer path through the outer superlayers than negative tracks do.

The relative efficiency for positive and negative tracks still ought to be close, as the path length difference is not striking except for *very* low momentum tracks. More striking is the efficiency variation versus $\cot\theta$: A track at $\eta = 1$, or about $\cot\theta = 1.2$, has a path length in the CTC over half-again as long in each CTC cell as a track at $\eta = 0$.

Figures 7.4 and 7.5 clearly show, respectively for negative and positive tracks with P_T greater than 5 GeV, that the efficiency varies strongly with $\cot\theta$, and also that the difference between the efficiency of positive tracks and negative tracks is relatively small. Curiously, the efficiency difference between positive and negative tracks is not symmetric with $\cot\theta$; negative tracks are slightly more efficient than positive tracks. Still, the total efficiency difference between positive and negative tracks is smaller than 2%.

I discuss the systematic uncertainty due to this correction in Section 7.5.

7.2.7 CFT vs. Nearby Tracks Correction

Neighboring tracks can influence the track-finding efficiency for a given track. Usually, extra hits from nearby tracks are irrelevant; however, extra hits *can* compensate for

MINUIT χ^2 Fit to Plot 1157&1
- Eff vs CoTan(theta)
File: [EKUNS.TRKEFF]TRKEFF.PAW;4 6-DEC-97 21:49
Plot Area Total/Fit 30.616 / 30.616 Fit Status 3
Func Area Total/Fit 30.630 / 30.630 E.D.M. 2.013E-06
χ^2= 14.0 for 30 - 6 d.o.f., C.L.= 94.6%

Errors		Parabolic	Minos	
Function 1: Gaussian (sigma)				
AREA	-0.43078	± 5.1965E-02	- 0.0000E+00	+ 0.0000E+00
MEAN	-7.94579E-03	± 1.3305E-02	- 0.0000E+00	+ 0.0000E+00
SIGMA	0.26281	± 2.1528E-02	- 0.0000E+00	+ 0.0000E+00
Function 2: Polynomial of Order 2				
NORM	1.0205	± 3.9666E-03	- 0.0000E+00	+ 0.0000E+00
* POLY01	0.00000E+00	± 0.0000E+00	- 0.0000E+00	+ 0.0000E+00
POLY02	1.95663E-02	± 3.7402E-03	- 0.0000E+00	+ 0.0000E+00
OFFSET	-0.10811	± 4.2146E-02	- 0.0000E+00	+ 0.0000E+00

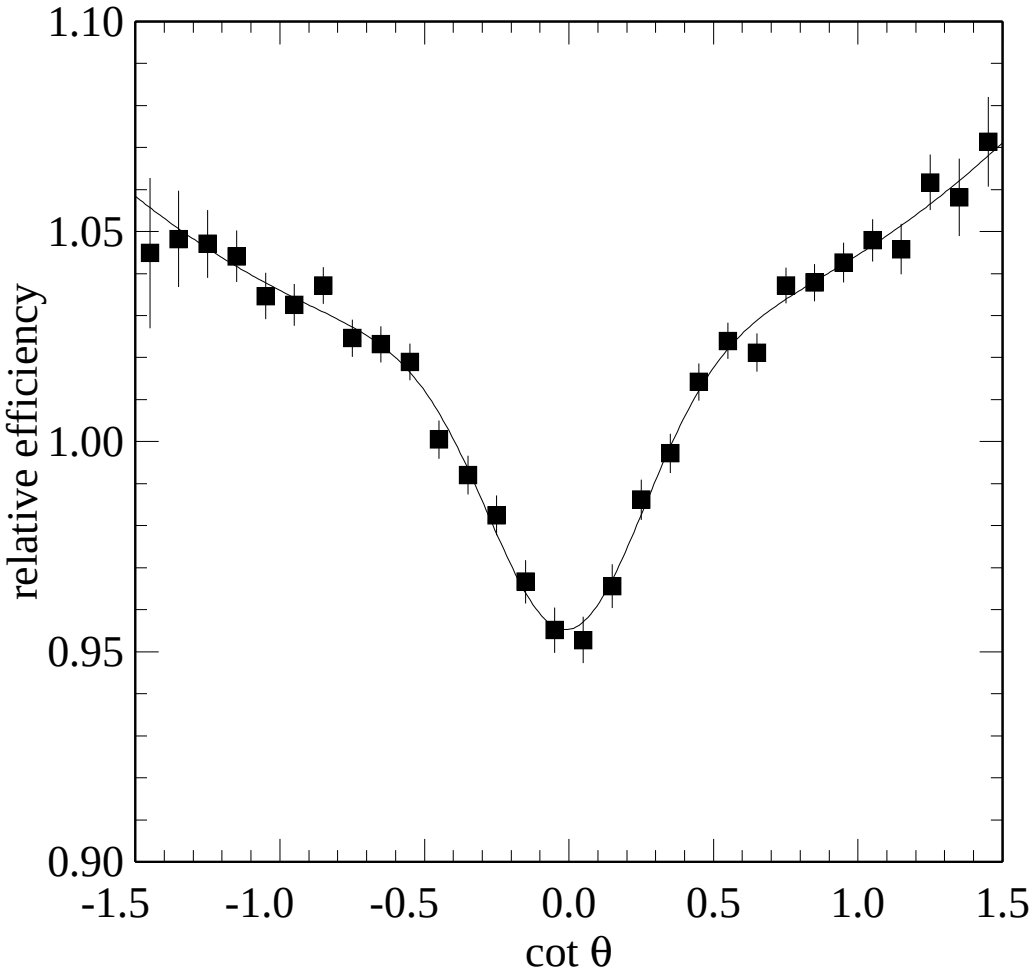


Figure 7.4: CFT negative track efficiency vs. $\cot\theta$ for tracks with $P_T > 5$ GeV. Since this is applied as a weight to correct CFTSIM, the overall scale is chosen to not change the total number of events.

MINUIT χ^2 Fit to Plot 1167&1
+ Eff vs CoTan(theta)
File: [EKUNS.TRKEFF]TRKEFF.PAW;4 6-DEC-97 21:49
Plot Area Total/Fit 30.237 / 30.237 Fit Status 3
Func Area Total/Fit 30.281 / 30.281 E.D.M. 1.143E-05
χ^2= 29.0 for 30 - 6 d.o.f., C.L.= 22.1%

Errors		Parabolic	Minos	
Function 1: Gaussian (sigma)				
AREA	-0.65278	± 0.1014	- 0.0000E+00	+ 0.0000E+00
MEAN	8.57814E-03	± 1.3848E-02	- 0.0000E+00	+ 0.0000E+00
SIGMA	0.33082	± 2.7798E-02	- 0.0000E+00	+ 0.0000E+00
Function 2: Polynomial of Order 2				
NORM	1.0242	± 7.0233E-03	- 0.0000E+00	+ 0.0000E+00
* POLY01	0.00000E+00	± 0.0000E+00	- 0.0000E+00	+ 0.0000E+00
POLY02	9.23359E-03	± 5.7792E-03	- 0.0000E+00	+ 0.0000E+00
OFFSET	6.38061E-02	± 9.3086E-02	- 0.0000E+00	+ 0.0000E+00

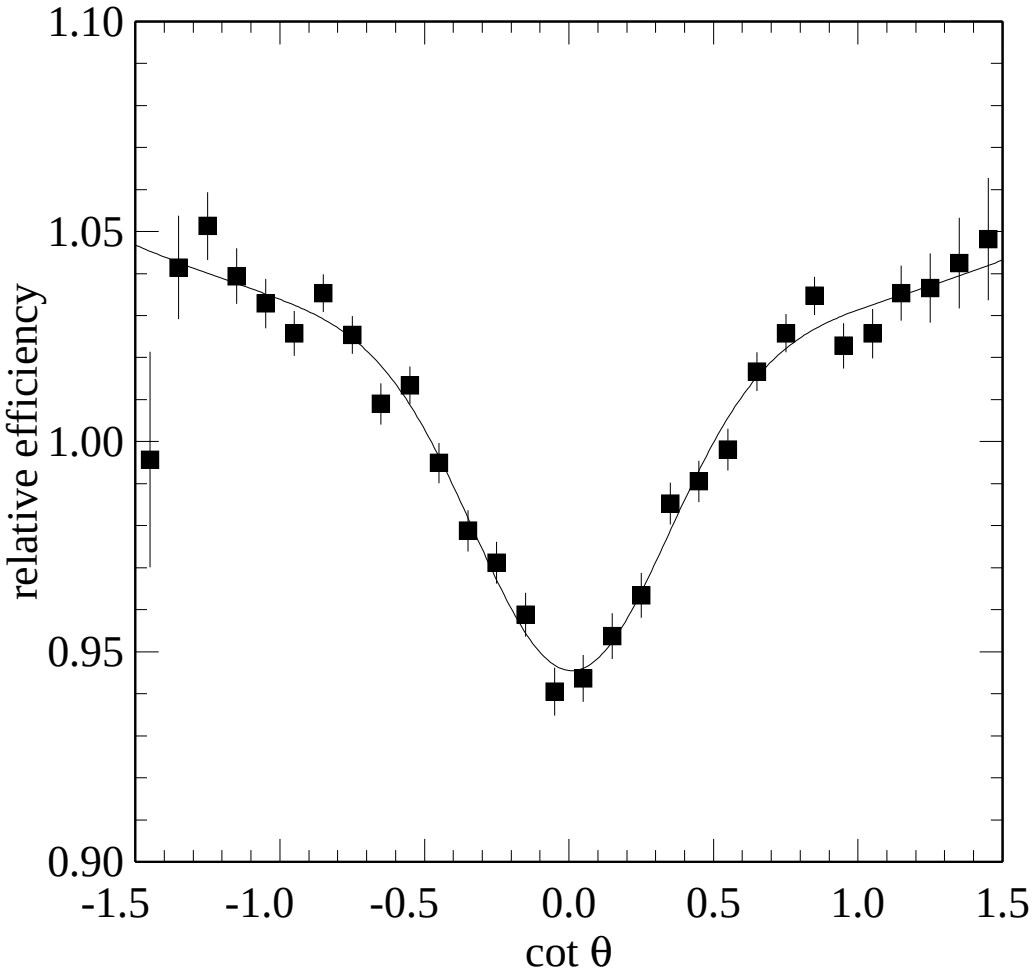


Figure 7.5: CFT positive track efficiency vs. $\cot\theta$ for tracks with $P_T > 5$ GeV. Since this is applied as a weight to correct CFTSIM, the overall scale is chosen to not change the total number of events.

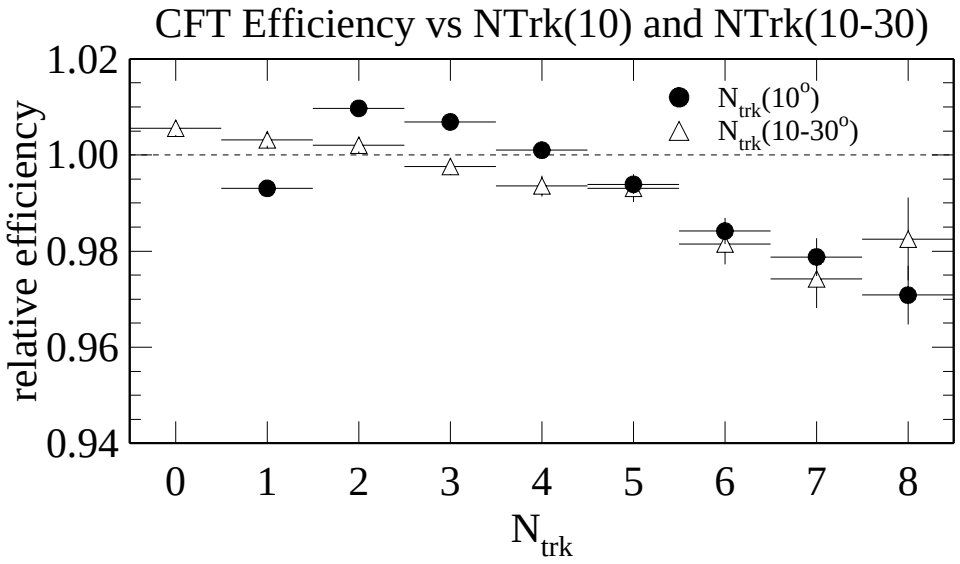


Figure 7.6: CFT efficiency with nearby tracks. The points show how the track-finding efficiency varies with the number of tracks in a cone 10° from the seed track; there is always at least one. Triangles show how the efficiency varies with the number of tracks in an annulus 10–30° from the seed track. Note the scale; the change in efficiency is small for all tracks.

inefficiencies in finding a prompt or delayed hit. Extra hits which lie in a CFT "road" can cause the CFT to mismeasure the momentum or charge of a track.

The total effect of nearby tracks on the efficiency for the CFT to identify a track is small. Looking at variables used in this analysis—the number of tracks in a 10° cone and the number of tracks in a 10–30° annulus—Figure 7.6 shows the size of each effect. Since what is important is the *relative* efficiency of each effect, the efficiencies are scaled to not change the total number of events if applied as a weight.

The overwhelming majority of events in this analysis contain five tracks or fewer in both the 10° cone and the 10–30° annulus, so the efficiency bias from tracks in the annulus is a very small effect—less than one percent. Although the efficiency bias from the number of tracks in the 10° cone seems to be more significant, it is still small—less than $\pm 1\%$ for 1–5 tracks in the 10° cone. The small drop in efficiency as tracks are added to the 10° cone is explained below.

Perhaps more significant is the $\approx 2\%$ efficiency difference between 1-prongs and,

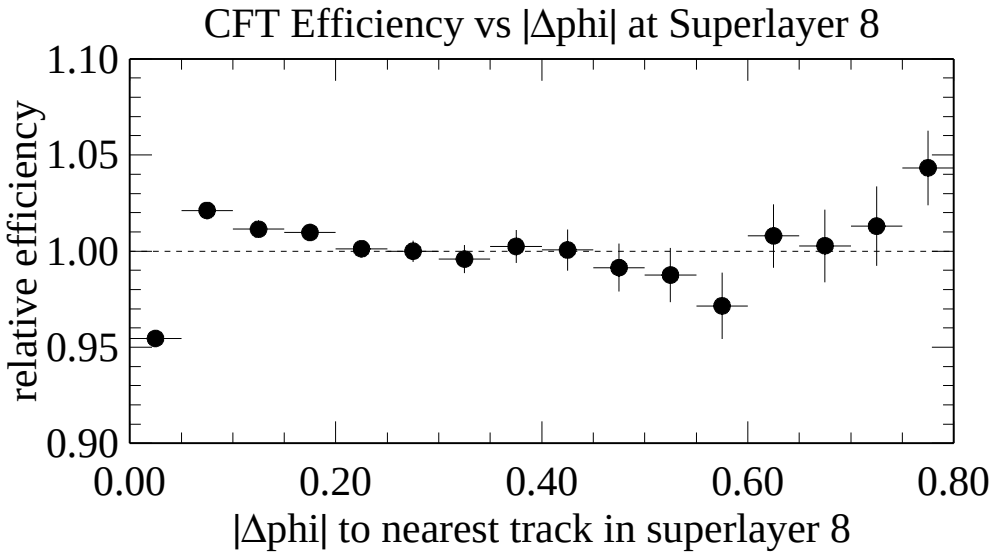


Figure 7.7: CFT efficiency vs. $\Delta\phi$(Superlayer 8) to nearest track. Note the drop in efficiency as $\Delta\phi$ is near zero and the slight enhancement in efficiency for tracks very near, but not coincident with the seed track.

as a group, 2- and 3-prongs. Most clusters with more than one track in the 10° cone contain tracks of both charges, thus, tracks which may cross in ϕ and contribute prompt or delayed hits to one another. The efficiency for a lone track in a 10° cone, thus, is closer to the CFT efficiency for an isolated track. Also, the "overefficiency" of tracks with another track nearby may be overstated for high momentum tracks with $P_T \gtrsim$ 15 GeV. Since the sign of such tracks is poorly measured by the CTC, the same CFT track may appear to match more than one CTC track. In any event, this is a small effect.

Nearby track prompt hits in superlayer 8 are especially difficult for the CFT since all roads start in that superlayer. To avoid duplication, the CFT records at most one track for a given keying wire and its immediate neighbors in superlayer 8. If two (or more) tracks extrapolate to within about 0.5° in ϕ at superlayer 8, the CFT records only a single CFT track matching the highest momentum track. For tracks in the same momentum bin, a negative track is arbitrarily chosen over a positive track.

Figure 7.7 shows, as a function of abs($\Delta\phi$) at superlayer eight from a seed track

to its nearest neighbor, how nearby tracks can affect the CFT track finding efficiency. Note that a nearby track in ϕ may be anywhere in η. As a result of this effect, only two of the three tracks of a three-prong tau are likely to be found by the CFT, especially when the two like-charge tracks have momenta close to one another, but this does not strongly affect the triggering efficiency.

I discuss the systematic uncertainty due to this correction in Section 7.5.

7.3 QCD Jet Background Model Uncertainty

A significant source of systematic uncertainty is the QCD jet background identification; I use the TAU-20 event sample to model the QCD jet background, as discussed in Section 5.2.

Two primary cuts are applied to the TAUMON event sample which are not applied to the TAU-20 event sample: Monojet selection cuts and $\not\!\!E_T$ cuts. Other than effects from these two cuts (and the difference in integrated luminosity), the TAU-20 event sample *is* the QCD jet background. I cannot study the effect of the missing cuts by applying them "full force"; only 13 TAU-20 events pass the full TAUMON event selection.

Thus, I apply the missing $\not\!\!E_T$ and monojet selection cuts one at a time, measuring the efficiency of these cuts as a function of charge pattern index, η, and isolation. Then, assuming the missing cuts are independent, I calculate the total TAU-20 efficiency (versus charge pattern, η, and isolation) for these cuts by multiplying their efficiencies. TAU-20 is used in the log likelihood fit to model the QCD jet background distributions versus η and charge pattern; the efficiency of the missing cuts as a function of these variables is presented in Figure 7.8. If the missing cuts impose no bias in these variables, a horizontal line will fit the isolated and nonisolated samples (not necessarily with the same efficiency) with an acceptable χ^2.

In fact, the only significant effect appearing in Figure 7.8 is a relative loss of efficiency for non-isolated 1-prong positive tracks which does not appear for isolated events.

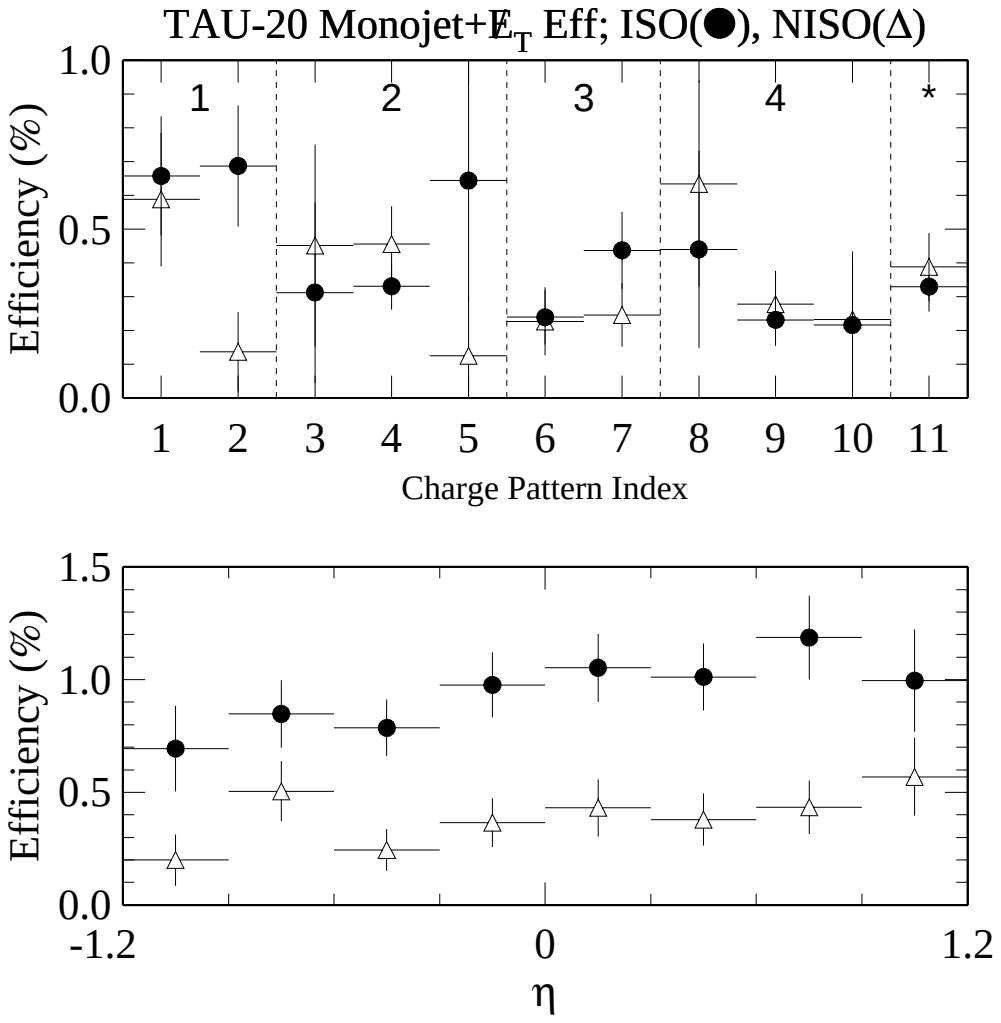


Figure 7.8: TAU-20 efficiency vs. charge pattern index and η. The only significant charge asymmetry is for the 1-prong positive events. There appears to be a systematic shift versus η in the efficiency, due to statistical uncertainty, but the distributions are consistent with no shift.

No other statistically significant systematic effect is visible. Since TAU-20 contains relatively more statistics than the actual QCD background does, an effect which is just visible in TAU-20 is insignificant in the QCD background given the statistics of this analysis. (The TAU-20 event sample contains about 4600 events; after fitting, the QCD jet background is identified as containing about 330 events.)

Therefore, to correct the one significant effect in the fit, I multiply the non-isolated 1-prong positive bin in TAU-20 by 1/3. I estimate the uncertainty of this correction as 1/2 the shift in asymmetry as a result of applying the correction. This uncertainty is negligible (much less than 0.5% and consistent with zero) in most of the central bins and about 1% in the outer η bins. The uncertainty for each η bin is listed below with η increasing left to right:

$$\left(\begin{array}{cccccccc} 0.010 & 0.013 & - & - & - & 0.005 & 0.009 & 0.010 \end{array} \right)$$

With an earlier set of cuts, a noticeable and significant loss of events at the outermost η bins was visible, so I assigned a correction factor of 0.80 to these outer bins in TAU-20. This effect is no longer visible (as seen in Figure 7.8), but the effect of this correction on the asymmetry is negligible, so I still apply it. One η bin—at the lowest η value—shows noticeable asymmetry shift (of 0.4%); all other η bins have a negligible asymmetry shift and are assigned zero systematic uncertainty for this correction.

7.4 Miscellaneous Uncertainties

Collected in this section are all uncertainties not directly related to the simulation details for the signal and background event samples.

7.4.1 E_T Scale

The charge asymmetry measured for a given electroweak sample depends on the E_T cut applied during analysis—a higher E_T cut results in a larger measured asymmetry. (See Figure 6.5.) When an asymmetry theory curve (based on a given parton distribution function) is compared to the asymmetry calculated by the fit, the two are comparable

only when the energy scale applied when calculating the theory asymmetry curve is the same as the energy scale in the real data.

Note that for this purpose, the energy scale used in WTN (i.e., in QFL) is irrelevant for the log likelihood fit. As described in Chapter 6, the log likelihood fit carefully removes the physics asymmetry present in WTN. Other analyses at CDF [47, 48] have carefully measured the jet energy scale uncertainty for Run 1a. In particular, one study [48] measures the systematic uncertainty of the jet energy scale as (3.2 ± 1.5(stat) $\pm$ 4.1(sys))% using 0.4 radian cone jets with an average P_T of 24 GeV—well suited to applying to tau jets.

Unfortunately, the jet energy scale uncertainty measured in Ref [48] is measured after jet energies are corrected for several detector effects. I do not apply these corrections to tau jets; tau jets are more narrow (in general) than the generic QCD jets used to measure the jet energy correction, so corrections for the amount of energy lost in a crack are incorrect. Also, tau jets contain relatively more EM energy than most generic hadronic jets, due to the frequency with which π^0's accompany $\pi^{\pm}$'s, for example, from tau decays through the $\rho^{\pm}$ resonance.

That jet energy corrections are not applied should not change the energy *scale* uncertainty, especially since such corrections derived for taus would be quite different from the ones for QCD jets. I thus assume a 4% uncertainty of the energy scale and a 6% uncertainty in the energy scale of the Level 1 and Level 2 triggers. (I allow a greater uncertainty in the trigger energy scale since the triggers use fast analog outputs with slightly different resolution than the digitized energy used offline.)

The uncertainty in the asymmetry measurement due to uncertainty in the energy scale is small, and artificially choosing a larger energy scale uncertainty (10%, for example) does not appreciably change the energy scale systematic uncertainty.

I discuss the systematic uncertainty due to this effect in Section 7.5.

7.4.2 Monte Carlo Sample Normalization

Since the electroweak monte carlo samples are scaled directly and are entered into the log likelihood fit as constants, the uncertainty on the electroweak monte carlo sample

scale factors, although small, must be accounted for. The electroweak backgrounds are a small fraction of the total background, so their scale uncertainty should be a small systematic effect. I study this effect as if it is an effect in the WTN monte carlo sample, i.e., as a "correction" to WTN.

If the electroweak monte carlo scale is overestimated, for example, the net effect in the fit is similar to subtracting extra electroweak background events from the $W \to \tau\nu$ simulated event sample. The "correction" used in this study, thus, is the electroweak monte carlo event sample, scaled by the scale *uncertainty* rather than by the scale.

I discuss the systematic uncertainty due to this correction in Section 7.5.

7.4.3 Background Double-counting

The fit separately accounts for electroweak backgrounds and jet backgrounds in the TAUMON data sample; however, the TAU-20 data sample contains some electroweak background events in addition to the jet events it is supposed to model. In addition, TAU-20 contains some $W \to \tau\nu$ events, which must be subtracted. The way these events are subtracted from TAU-20 is described in Section 5.2.3. Subtracting these contributions from TAU-20 before using it in the log likelihood fit prevents double-counting.

As discussed in Section 5.2.3, this correction is negligible: at most 2% for isolated 1- and 3-prong bins and negligible for the rest. The systematic effect of this correction on the asymmetry is negligible.

7.4.4 WTN Statistical Uncertainty

The last systematic effect studied is actually a statistical effect; although 1000000 $W \to \tau\nu$ events were simulated, only about 12000 survive all the event selection cuts. If the 12000 events are half positive and half negative, and if they are about evenly divided with η, then each η bins contains about 750 positive and 750 negative taus. (Of course, this is not the case!) The statistical uncertainty of the asymmetry measurement, given these numbers, is approximately 2%.

Using the actual WTN event sample distributions to calculate the statistical uncertainty on the asymmetry present, versus η, results in the uncertainties given below:

$$\left(\begin{array}{cccccccc} 0.050 & 0.025 & 0.022 & 0.022 & 0.023 & 0.022 & 0.025 & 0.049 \end{array} \right).$$

If a larger number of $W \to \tau\nu$ events is simulated, this uncertainty will decrease. Although the charge asymmetry present in the $W \to \tau\nu$ monte carlo is subtracted from the input to the log likelihood fit, I take this statistical uncertainty to reflect the systematic uncertainty introduced by the limited simulation statistics.

As mentioned before, all these uncertainties are small compared with the statistical uncertainty of the asymmetry from the log likelihood fit, which isn't surprising considering the $W \to \tau\nu$ monte carlo event sample is equivalent to about 26 times the integrated luminosity of the TAUMON event sample. Thus, the statistical uncertainty due to TAUMON statistics should be about five times larger than the statistical uncertainty of asymmetry measured directly from the $W \to \tau\nu$ event sample. In fact, the uncertainties are a little over five times larger, as expected, since the limited statistics of the QCD jet background sample also influence the result of the fit.

7.5 Systematic Uncertainty of Corrections

I study the systematic uncertainty of each correction using the same procedure, which I outline here. Specifically, I study the following uncertainties with a "toy" monte carlo simulation:

- E_T Scale,
- the shape and offset of the Z vertex,
- single pion efficiency near trigger tower cracks,
- electroweak monte carlo normalization,
- CFT track finding efficiency as a function of track isolation,
- CFT track efficiency as a function of $\cot\theta$.

I loop over the 1 million generated WTN events. For each loop, I generate a random number, N, between -3 and 3, the number of standard deviations by which I adjust the systematic effect under study. Each of the 1 million event is randomly assigned (independently for each loop) to either a test sample where I adjust the systematic effect, or to a control sample where I do not. After following all million events through one loop, I measure the asymmetry of the control and test samples and calculate the difference. For each systematic effect, I plot Asymmetry(test) – Asymmetry(control) vs. N. I loop through the WTN event sample 1000 times for each systematic effect.

The final result due to modifying the E_T scale (and reapplying all event selection cuts when the E_T scale is shifted up or down) is shown in Figure 7.9. Creating a profile histogram from the scatter plot and then fitting a line to the points results in Figure 7.10. Repeating this process for the electroweak normalization results in Figure 7.11; repeating on the correction for the CFT efficiency vs. $\cot\theta$ results in Figure 7.12.

Comparing the corrected Z vertex shape (systematically adjusted up and down by N sigma) against the uncorrected Z vertex shape, Figure 7.13 shows that there is no systematic shift in the asymmetry due to this effect or the uncertainty in this effect. I do not apply this correction, but I do measure the uncertainty due to not simulating the same Z vertex shape as in the data.

The last systematic effect and the second largest systematic uncertainty involves the efficiency of single pions in the Level 1 trigger when they pass near a trigger tower crack in η. Although this is among the largest systematic uncertainties, it is still much smaller than the statistical uncertainty of the asymmetry measurement. Figure 7.14 represents the result of the toy monte carlo simulation. Note that since I am comparing uncorrected data to corrected data (in which the correction is shifted by $N\sigma$), the lines fit to the profile histograms are not constrained to pass through the origin. Thus, the slope of each line (e.g., d(Asymmetry)/dσ) and the vertical offset each provide a measure of uncertainty.

The systematic uncertainties are summarized in Figures 7.15 and 7.16. The systematic uncertainties associated with E_T scale (in Figure 7.15) are possibly anti-symmetric

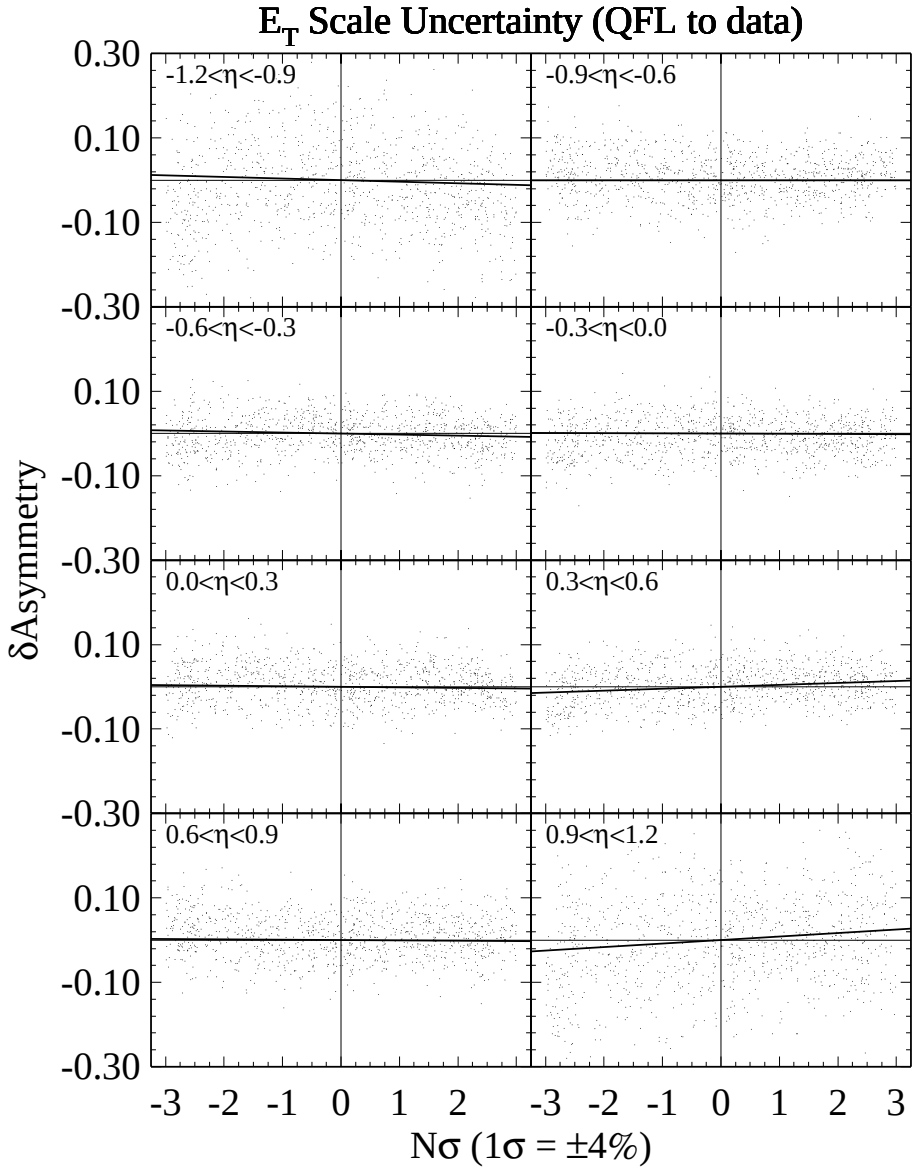


Figure 7.9: Scatterplot for the E_T scale of δAsymmetry vs. $N\sigma$, the amount by which the E_T scale is shifted. This effect is measured independently for each η bin. The line in each bin is the result of a linear fit, as described in the text.

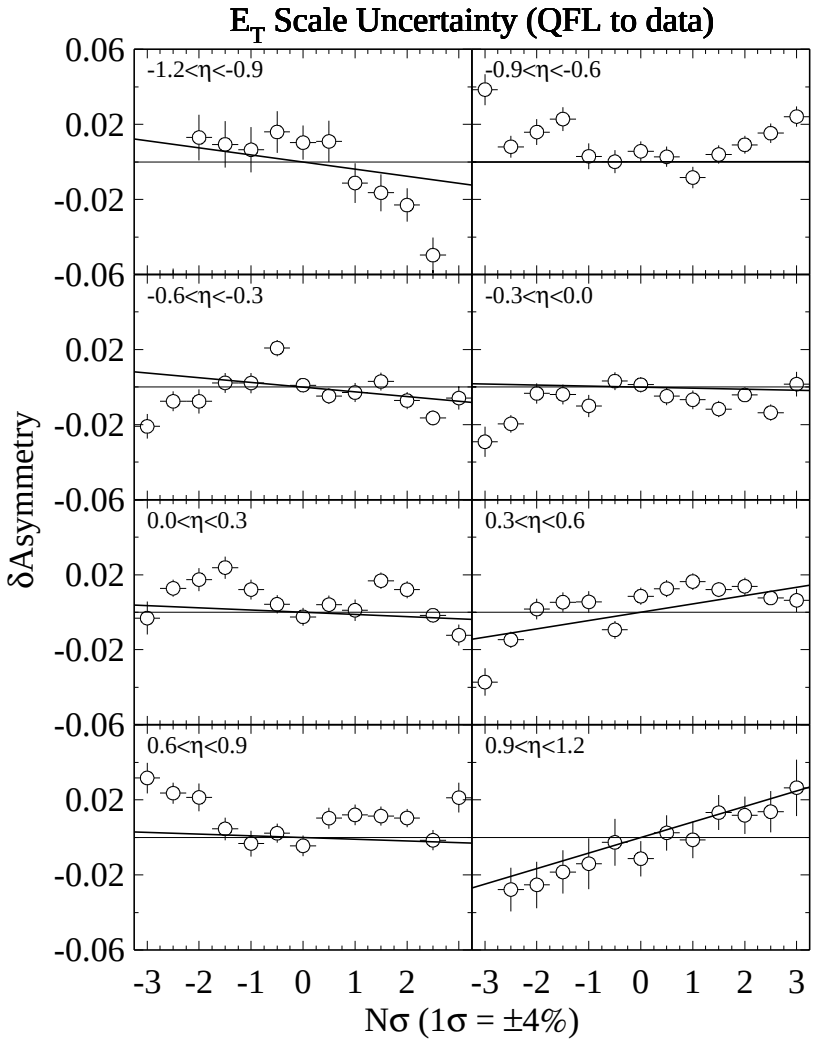


Figure 7.10: Profile histogram for E_T scale: δAsymmetry vs. $N\sigma$. This effect is measured independently for each η bin. The line in each bin is the result of a linear fit to the points. Note that the E_T scale is not necessarily linear in σ, as is clear from the scatter in this plot. The asymmetry depends on the E_T scale in a nonlinear manner. As expected, the outermost bins are anti-correlated with one another.

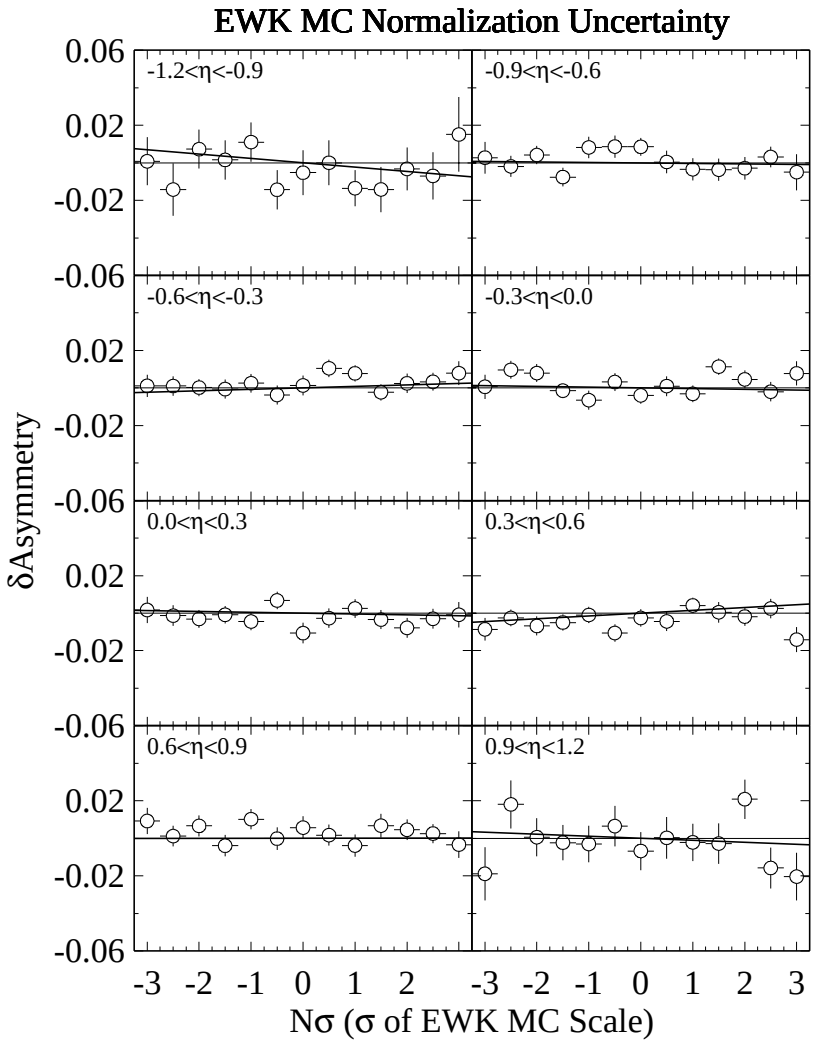


Figure 7.11: Profile histogram for EWK monte carlo scale: δAsymmetry vs $N\sigma$. There is no clear effect visible. The uncertainty due to the electroweak monte carlo normalization is negligible, not surprising since the electroweak backgrounds are such a small fraction of the final event sample.

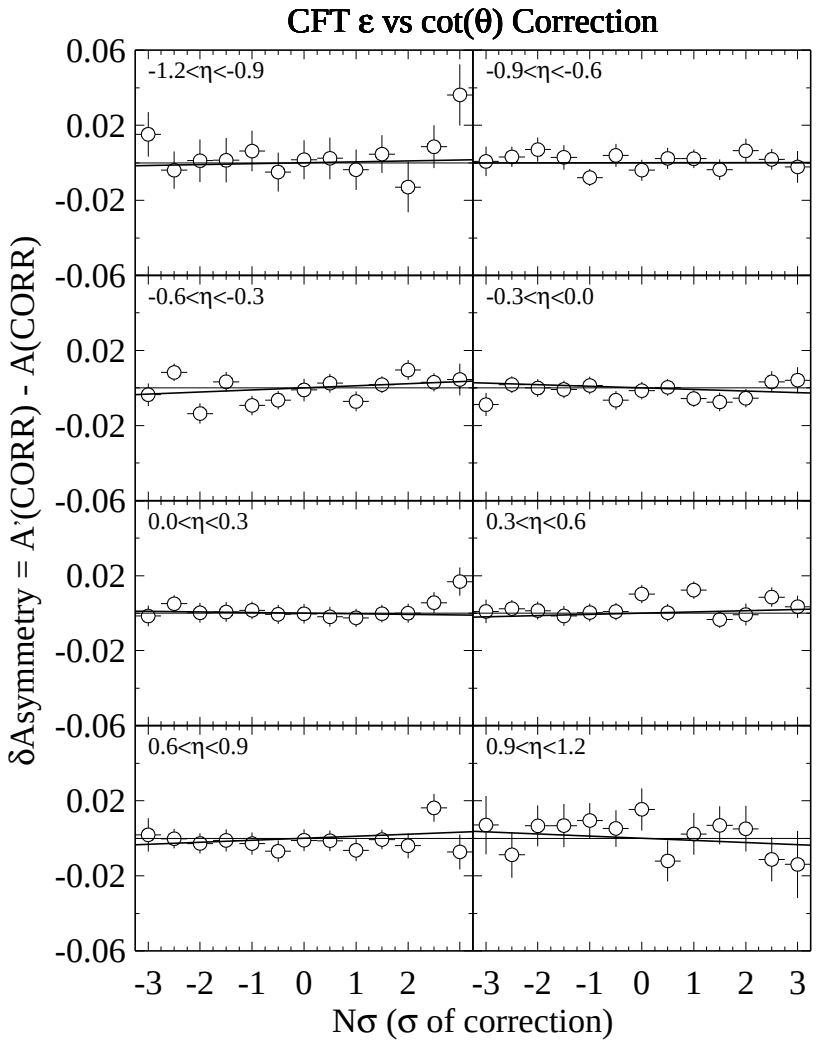


Figure 7.12: Profile histogram for uncertainty due to the CFT efficiency vs. $\cot\theta$ correction: δAsymmetry vs $N\sigma$. There is no clear effect visible. With the correction applied, varying the correction within the measured uncertainties does not appreciably change the asymmetry measurement.

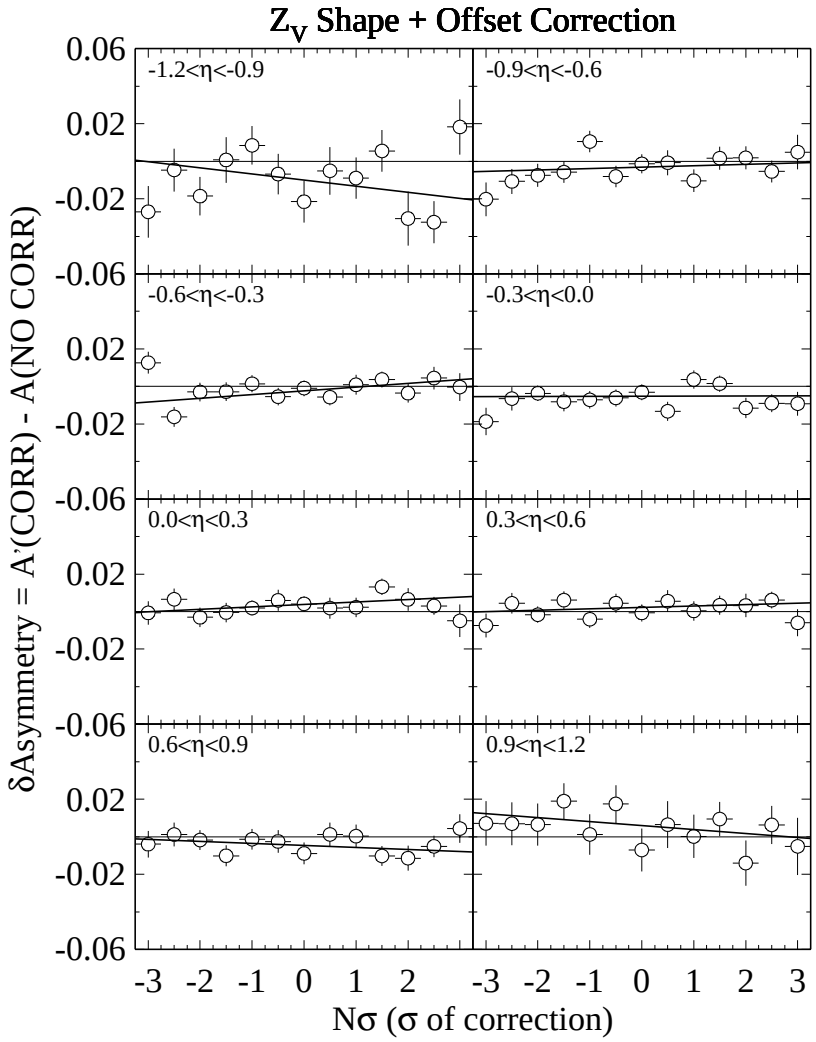


Figure 7.13: Profile histogram for uncertainty due to the Z vertex shape generated for WTN versus the actual Z vertex shape in the data. Since the corrected Z vertex shape is compared with the uncorrected Z vertex shape, a line fitted to this profile histogram is not constrained to pass through the origin. Thus, this effect adds uncertainty both from the slope (dA/dsigma) of the line and the vertical offset of the line.

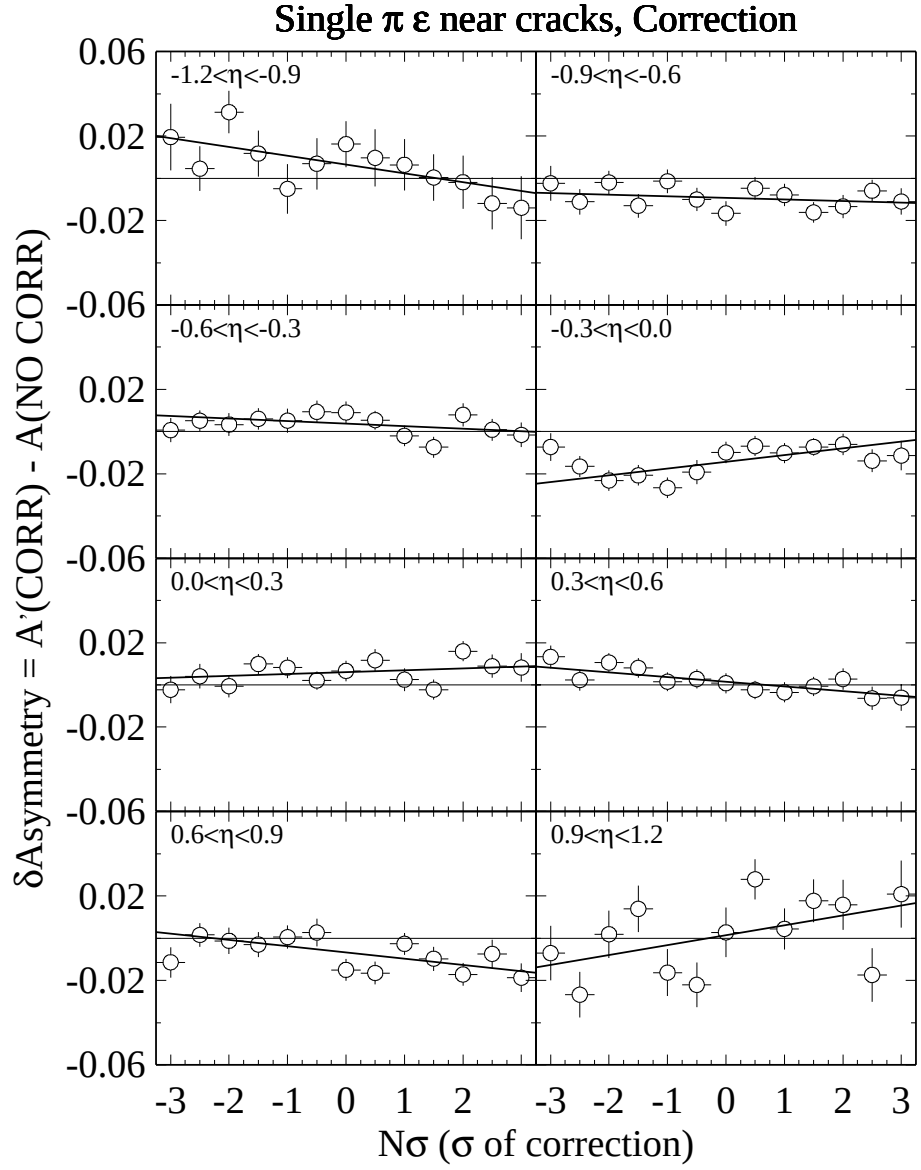


Figure 7.14: Profile histogram for uncertainty due to the single pion shower simulation near trigger tower calorimetry cracks. Since the corrected data are compared against the uncorrected data, a line fit to this profile histogram is not constrained to pass through the origin.

Systematic effect	Eta bins 1–8							
E_T Scale	0.010	0.005	0.005	0.005	0.005	0.005	0.005	0.010
Z_V Shape	0.010	0.005	0.005	0.005	0.005	0.005	0.005	0.010
QFL single π sim	0.010	0.010	0.010	0.015	0.010	0.010	0.010	0.010

Table 7.2: Systematic uncertainty of corrections vs. η

(as one might expect), but the effect is small (at most, about 1%), and especially small when compared with the statistical uncertainties in this analysis.

The Electroweak normalization uncertainty (also in Figure 7.15) is consistent with zero effect; the χ^2 per degree of freedom is less than one.

The rest of the systematic uncertainties are summarized in Figure 7.16. Note that the bottom two uncertainties—associated with the CFT—are negligible. The bottom left plot ($\cot\theta$ correction) shows the uncertainty *after* applying the correction. Applying the correction results in about a 1% systematic shift of the asymmetry, but this shift corrects the asymmetry to what would be measured if CFTSIM properly simulated the CFT efficiency. The bottom right plot shows the uncertainty in CFT track identification as a function of track isolation. This uncertainty is negligible.

The top two plots summarize the uncertainty of the Z vertex shape and the single pion efficiency near trigger tower cracks, as discussed in the previous section. Points show the slope (d(Asymmetry)/dsigma); triangles show the vertical offset. These uncertainties show no obvious correlations except for the two outermost η bins for the Z vertex shape correction.

The systematic uncertainties of the corrections are listed in Table 7.2. Effects not listed in that table (electroweak normalization and the two CFT corrections) have a negligible affect on the asymmetry measurement.

7.6 Systematic Uncertainty Correlations

To calculate the correlation matrix for the asymmetry versus η, I must first generate the correlation matrices for each systematic effect, calculate the covariance matrix from each correlation matrix, add the covariance matrices, and finally convert the summed

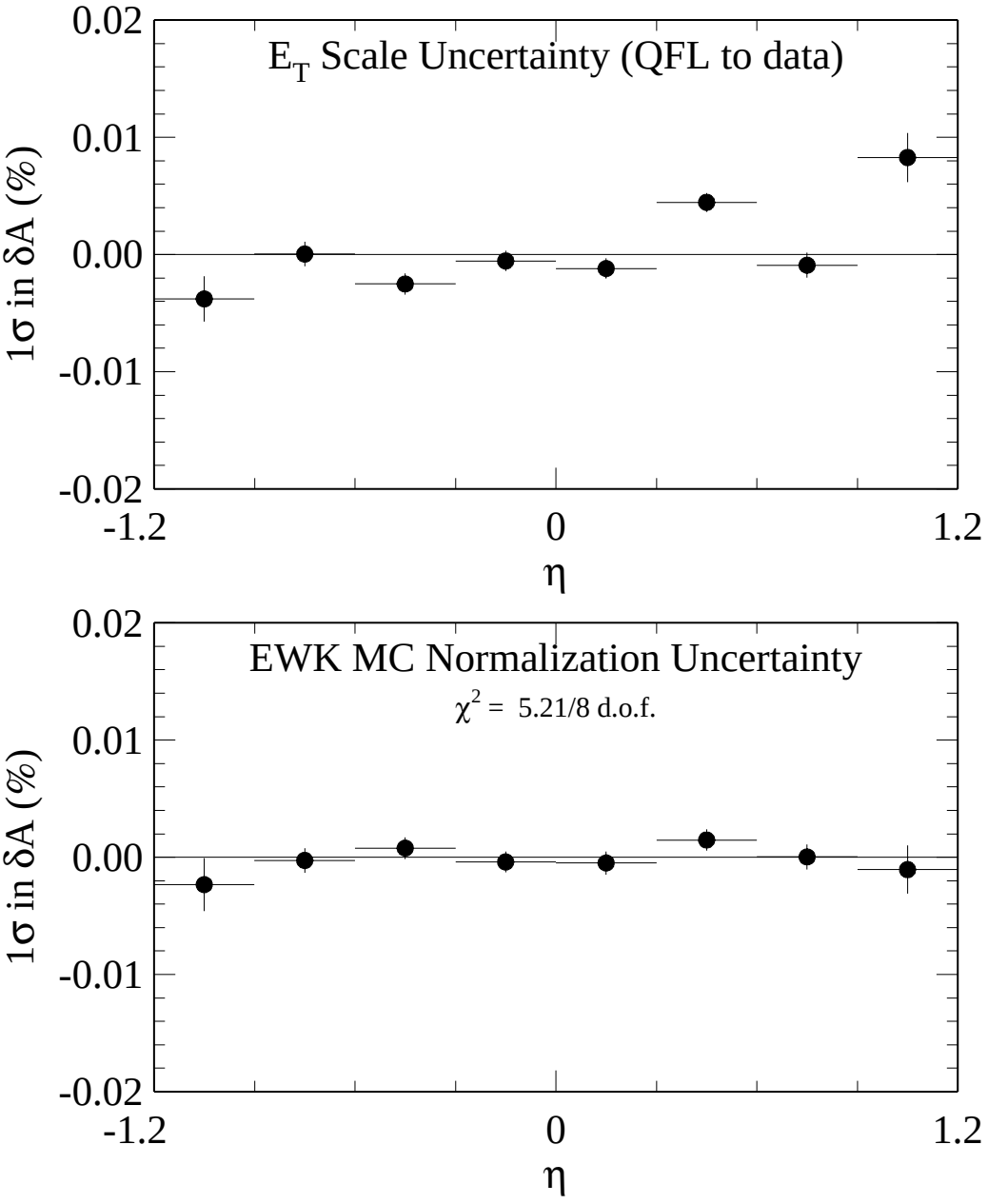


Figure 7.15: Systematic uncertainty as a function of η for the E_T scale (top) and the electroweak monte carlo sample normalization (bottom).

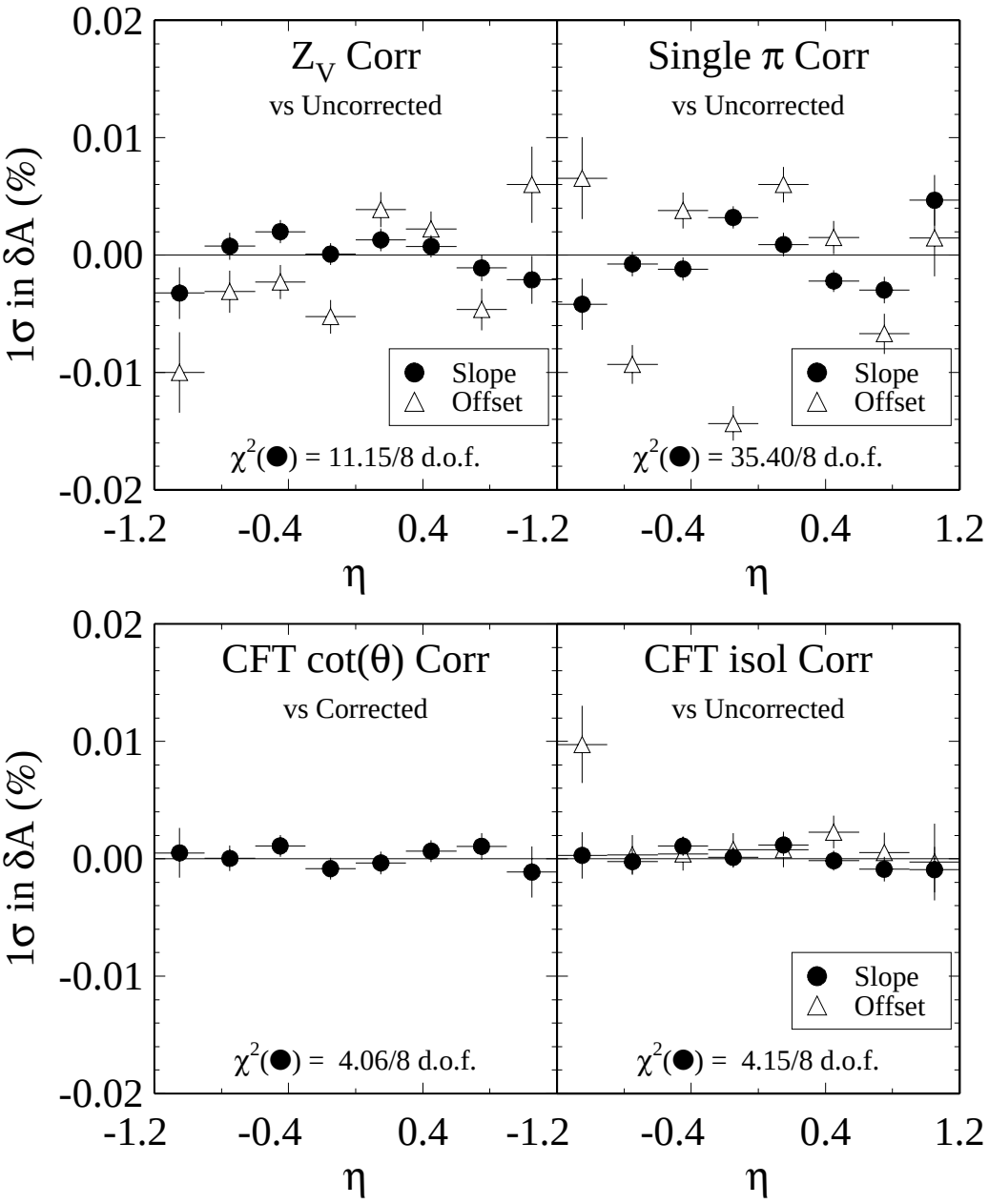


Figure 7.16: Graphical summary of systematic uncertainties. The uncertainties associated with the CFT are negligible. Both the Z_V correction and the single pion efficiency show noticeable uncertainty in both the slope and vertical offset in the profile histograms.

covariance matrix into a summed correlation matrix.

No clear method exists to calculate the correlations of the uncertainty in asymmetry for the systematic effects. Since the uncertainties are small and the correlations mostly appear small, I "eyeball" the uncertainty versus η for each correction.

The two TAU-20 corrections are simple. Looking at the TAU-20 versus charge pattern correction, I apply the same correction in each η bin, so the correlation matrix for this effect is uniformly 1.0. On the other hand, the uncertainty for the TAU-20 η correction is non-negligible in only one η bin, so correlations are irrelevant. The correlation matrix is thus 1.0 for the diagonal elements, 0.0 for the rest.

No clear correlation is visible in the asymmetry changes when accepting or rejection the Level 2 trigger cluster shape which is known to be severely charge biased. Thus, for this case I also assume a correlation matrix with 1.0 along the diagonal, 0.0 elsewhere. I also assume this form of correlation matrix for the $W \rightarrow \tau\nu$ statistical uncertainty, since each η bin is independent with respect to statistical fluctuations. I also assume this form of correlation matrix for the correction versus Z_V distribution and offset.

With the E_T Scale, four η bins seem not to change appreciably and the other four seem to change. Those bins that do change seem correlated/anti-correlated, so I assume the following form for the correlation matrix:

$$\begin{pmatrix}
1.0 & - & 1.0 & - & - & -1.0 & - & -1.0 \\
- & 1.0 & - & - & - & - & - & - \\
1.0 & - & 1.0 & - & - & -1.0 & - & -1.0 \\
- & - & - & 1.0 & - & - & - & - \\
- & - & - & - & 1.0 & - & - & - \\
-1.0 & - & -1.0 & - & - & 1.0 & - & 1.0 \\
- & - & - & - & - & - & 1.0 & - \\
-1.0 & - & -1.0 & - & - & 1.0 & - & 1.0
\end{pmatrix}$$

The last systematic correction studied is the QFL single pion efficiency near trigger tower cracks. The only bins which show any apparent (obvious) correlation are the extreme eta bins, which appear anti-correlated. I thus use the following correlation

matrix:

$$\begin{pmatrix} 1.0 & - & - & - & - & - & - & -1.0 \\ - & 1.0 & - & - & - & - & - & - \\ - & - & 1.0 & - & - & - & - & - \\ - & - & - & 1.0 & - & - & - & - \\ - & - & - & - & 1.0 & - & - & - \\ - & - & - & - & - & 1.0 & - & - \\ - & - & - & - & - & - & 1.0 & - \\ -1.0 & - & - & - & - & - & - & 1.0 \end{pmatrix}$$

Summing the covariance matrices for the systematic effects results in the following covariance matrix:

$$10^{-4} \times \begin{pmatrix} 37.0 & 1.3 & 0.5 & - & - & - & 0.9 & -1.0 \\ 1.3 & 11.6 & - & - & - & 0.7 & 1.2 & 1.3 \\ 0.5 & - & 6.4 & - & - & -0.3 & - & -0.5 \\ - & - & - & 10.4 & - & - & - & - \\ - & - & - & - & 7.2 & - & - & - \\ - & 0.7 & -0.3 & - & - & 6.9 & 0.5 & 1.0 \\ 0.9 & 1.2 & - & - & - & 0.5 & 8.7 & 0.9 \\ -1.0 & 1.3 & -0.5 & - & - & 1.0 & 0.9 & 30.6 \end{pmatrix}$$

which gives the following correlation matrix for the sum of all systematic uncertainties; elements above the diagonal are not shown:

$$\begin{pmatrix} 1.0 & & & & & & & \\ 0.063 & 1.0 & & & & & & \\ 0.033 & - & 1.0 & & & & & \\ - & - & - & 1.0 & & & & \\ - & - & - & - & 1.0 & & & \\ - & 0.079 & -0.046 & - & - & 1.0 & & \\ 0.049 & 0.118 & - & - & - & 0.064 & 1.0 & \\ -0.030 & 0.070 & -0.036 & - & - & 0.010 & 0.009 & 1.0 \end{pmatrix}$$

Systematic Effect	Uncertainty vs. η bin 1–8 (from $\eta = -1.2$ to $\eta = 1.2$)							
	1	2	3	4	5	6	7	8
TAU-20 vs. Ch.Ptrn	0.010	0.013	—	—	—	0.005	0.009	0.010
TAU-20 vs. η	0.004	—	—	—	—	—	—	—
L2 Cluster Shape	0.028	0.014	—	0.020	0.005	0.004	—	0.016
E_T Scale	0.010	0.005	0.005	0.005	0.005	0.005	0.005	0.010
Z_V Shape	0.010	0.005	0.005	0.005	0.005	0.005	0.005	0.010
QFL single π sim	0.010	0.010	0.010	0.015	0.010	0.010	0.010	0.010
WTN Statistics	0.050	0.025	0.022	0.022	0.023	0.022	0.025	0.049
Total Uncertainty	0.061	0.034	0.025	0.034	0.027	0.026	0.029	0.055

Table 7.3: Summary of systematic uncertainties. Uncertainties which are negligible for all η bins are not included in this table.

Clearly, few eta bins show a meaningfully correlated uncertainty in their asymmetry measurements.

7.7 Summary

The systematic uncertainties are summarized in Table 7.3, broken down by source and η bin. Since the smallest statistical uncertainty is over 11%, this measurement of tau charge asymmetry is dominated by statistical uncertainty. The results of the log likelihood fit, including corrections and systematic uncertainties, are given in the next chapter.

Systematic effects which are studied but not listed here are negligible.

Chapter 8

Results and Conclusions

> I do not know what I may appear to the world; but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.
>
> — Sir Isaac Newton [3]

Figure 8.1 presents the asymmetry measurement, including all corrections, for the full η range. This is the first direct measurement of tau charge asymmetry in W decays. The systematic uncertainties are indicated in the lower portion of the plot, the statistical uncertainties with the fit points. Clearly, the uncertainty in this analysis is statistics-dominated.

The quality of the fit may be gauged from Figure 8.2, showing the distributions versus charge pattern index of $W \to \tau\nu$, the electroweak and QCD jet backgrounds, and data, summed across all η bins. The $W \to \tau\nu$ signal is clearly evident for 1 and 3 track isolated clusters. This figure presents the direct output of the distributions in the fit, summing only across η.

Figures 8.3 and 8.4 present the asymmetry and fit distributions, respectively, with the data and fit folded about $\eta = 0$ to increase statistics and take advantage of the CP invariance of the underlying physics.

The tau asymmetry measurement is consistent with the Standard Model with a χ^2 per degree of freedom equal to 2.5 for 4 degrees of freedom when the asymmetry measurement is folded about $y = 0$ and 8.9 for 8 degrees of freedom when it is not.

The total uncertainty for the asymmetry measurement in each η bin, as shown in Figure 8.1, is

$$\left(\begin{array}{cccccccc} 0.28 & 0.13 & 0.12 & 0.13 & 0.12 & 0.11 & 0.15 & 0.33 \end{array} \right),$$

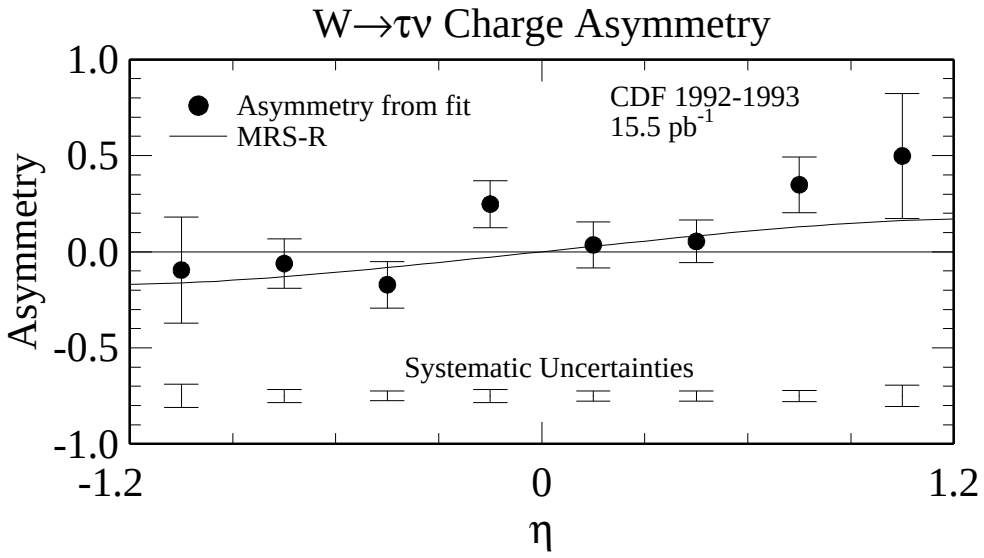


Figure 8.1: $W \rightarrow \tau\nu$ charge asymmetry for $1.2 < \eta < 1.2$. The systematic uncertainties are indicated in the lower half of each bin.

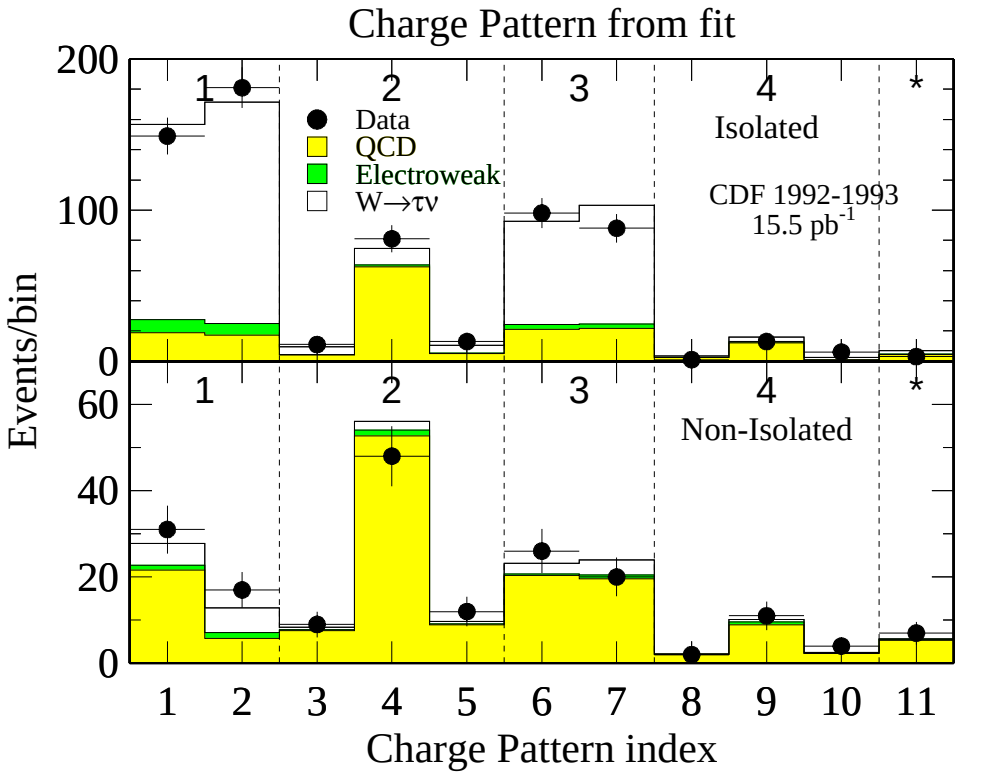


Figure 8.2: Charge pattern of fit, showing all contributions.

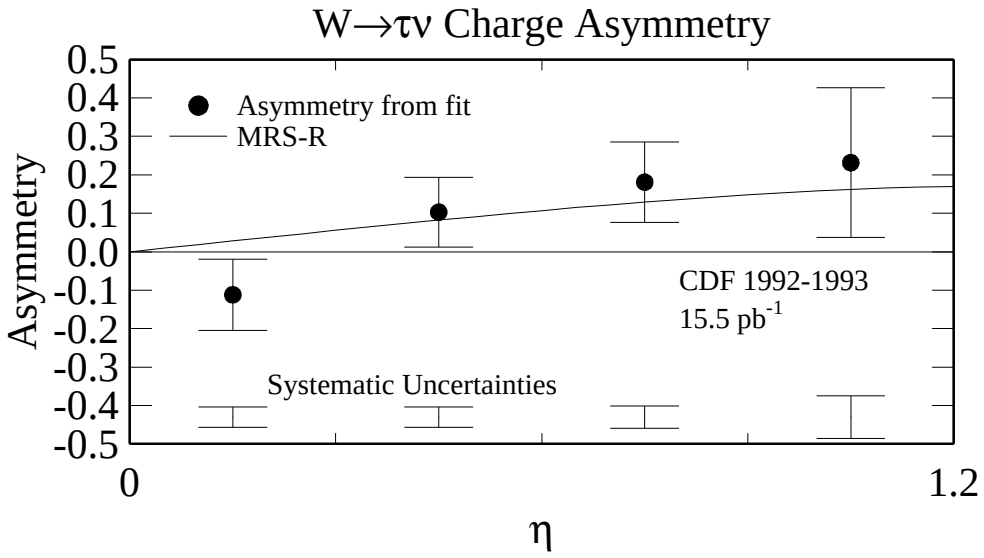


Figure 8.3: $W \rightarrow \tau\nu$ charge asymmetry for $0.0 < \eta < 1.2$. The systematic uncertainties are indicated in the lower half of each bin.

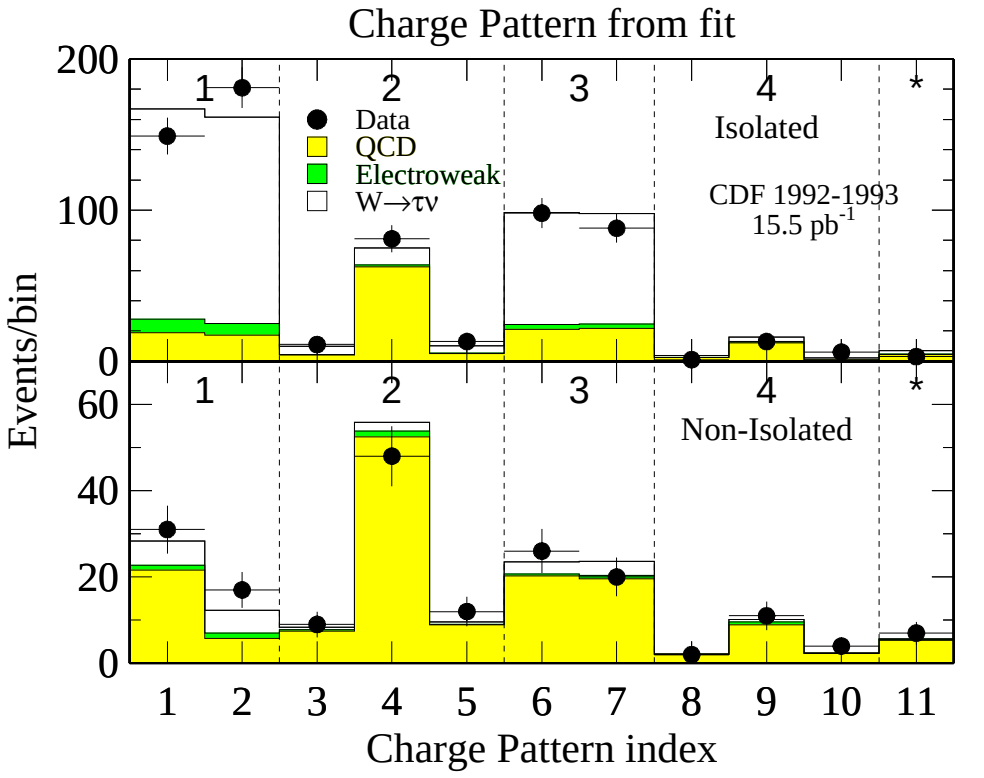


Figure 8.4: Charge pattern of fit, showing all contributions.

where η increases left to right.

The correlation matrix for the uncertainties of the asymmetry measurement for each η bin of the unfolded fit is given below. All uncertainties are included: statistical uncertainties from the log likelihood fit and the systematic uncertainties discussed in the last chapter. Eta increases from left to right and from top to bottom; elements above the diagonal are not listed and negligible correlations are indicated with "—".

$$\begin{pmatrix}
1.0 & & & & & & & \\
- & 1.0 & & & & & & \\
- & - & 1.0 & & & & & \\
- & - & -0.02 & 1.0 & & & & \\
- & - & - & -0.01 & 1.0 & & & \\
- & - & - & 0.01 & - & 1.0 & & \\
- & 0.01 & - & 0.02 & - & - & 1.0 & \\
- & - & - & - & - & - & - & 1.0
\end{pmatrix}$$

With the greater luminosity and much greater integrated luminosity approaching for CDF's Run II, the future of this measurement is bright. A much larger event sample and greater tracking reach in η will allow a measurement with greater precision. Together with the electron and muon charge asymmetry measurements, which are already providing constraints on parton distribution function parameterizations, these measurements will continue to improve our knowledge, and hopefully understanding, of the internal structure of the proton.

Appendix A

Normalized Log Likelihood Fitting

When fitting to the tau multiplicity distributions, Poisson statistics are more appropriate than Gaussian statistics; many bins contain few or no events. Poisson statistics motivate the choice to use a log likelihood function, not a χ^2, for the fit. While a χ^2 fit naturally provides quality-of-fit information, the functional value of a log likelihood function is meaningless. To recover this quality-of-fit information, I use a normalized log likelihood function [42, 43].

Using the customary λ and N to represent the expected number of events in a bin, and the data in that bin, respectively, the likelihood function may be written as

$$\mathcal{L} = \prod_{bins} \frac{\lambda^N e^{-\lambda}}{N!}$$

It is also necessary to define $\mathcal{L}_0$, the value of the likelihood function for a perfect fit:

$$\mathcal{L}_0 = \prod_{bins} \frac{N^N e^{-N}}{N!}$$

Thus, the normalized log likelihood function is

$$\begin{aligned} \ln \mathcal{L}' &= \ln \frac{\mathcal{L}}{\mathcal{L}_0} \\ &= \sum_{bins} \ln \frac{\lambda^N e^{-\lambda}}{N!} - \ln \frac{N^N e^{-N}}{N!} \\ &= \sum_{bins} \ln \lambda^N - \lambda - \ln N! - \ln N^N + N + \ln N! \\ &= \sum_{bins} N \ln \lambda - N \ln N + N - \lambda \end{aligned} \tag{A.1}$$

Finally, it can be shown (as in [42]) that

$$\chi^2 = -2 \ln \frac{\mathcal{L}}{\mathcal{L}_0} = -2 \ln \mathcal{L}'$$

Thus, minimizing $-2 \ln \mathcal{L}'$, one can treat the minimized value as a χ^2.

Rather than write a new fitting procedure to apply this log-likelihood fit, I use Minuit (release 93.08) [44]. Minuit calculates the uncertainties on the fit parameters automatically. In the fitting procedure, Minuit is instructed to calculate both the parabolic uncertainty calculated from the curvature at the fit and the MINOS uncertainty (see Ref [44], Chapter 7, for a description of MINOS) for which Minuit adjusts each parameter until the fit value changes by 1.0. Since the fit minimizes $-2 \ln \mathcal{L}'$, this is an appropriate choice. The MINOS uncertainties are used to check the parabolic uncertainty to ensure that Minuit calculates appropriate uncertainties.

Appendix B

CFT Simulation and CFTSIM

The Central Fast Tracker, CFT, is the hardware track-finding system used in the online Level 2 trigger and described in Section 3.7.5. The CFT identifies 2-dimensional $(r-\phi)$ charged tracks in the CTC. In this appendix, I describe the custom simulation package used to simulate the Level 2 CFT for this analysis.

Briefly, CFTSIM takes charged tracks as identified in the offline software and simulates the efficiency for the CFT to find the same tracks; CFTSIM does not simulate any CFT overefficiency. To most properly simulate the CFT, one should simulate track hits in the CTC, including noise hits, further simulating the CFT "prompt hit" and "delayed hit" resolution and efficiency. The QFL detector simulation, used to generate the WTN, WEN, WMN, and ZLL event samples, does not simulate charged tracks in such a detailed manner but rather in a parameterized manner: simulating only the measured resolution of track finding. That is, QFL generates tracks as would be identified by the track-finding code and not by a process of simulating track hits in the CTC and then processing those hits with the CDF track-finding code. The standard CFT simulation provided with CDF's Level 2 trigger simulation package relies on the existence of "track hits"—simulated or from data. Without these track hits, the standard CFT simulation does not have the information it needs to simulate CFT track-finding. This is the motivation for creating CFTSIM.

While an alternate (and CPU-intensive) detector simulation exists which can simulate CTC track hits, such a detailed simulation of the CFT is unnecessary for this analysis since the CFT track-finding efficiency is high.

Given a charged track identified by the offline track-finding code, the CFT efficiency

for identifying that track depends upon the *quality* of the offline track, where a high-quality track is very certain to correspond to a charged particle and a low-quality track may be a chance-identification of a track due to noise hits in the CTC, thus not corresponding to an actual charged particle.

The CFT is only efficient at identifying high-quality tracks with a few additional constraints not applied offline. For example, the CFT can not identify a track unless it passes through the full radius of the CTC; tracks which exit the CTC through an end plate, therefore, need not be simulated here. Additionally, very low momentum tracks which do not reach the outer radius of the CTC need not be simulated. Such tracks are not used in this analysis anyway, having a very small transverse momentum, less than 0.3 GeV.

In sum, several track quality cuts are applied before a track is simulated in the CFT. Tracks failing any of the quality cuts are "simulated" with zero efficiency in the CFT. The quality cuts are

- impact parameter $D_0 < 0.3$ cm,
- at least 2 stereo CTC layers contain 2 or more hits,
- at least 2 axial CTC layers contain 4 or more hits,
- the track passes through the full radius of the CTC.

A track's impact parameter, D_0, is its distance of closest approach to the beam position where collisions occur. Not only does the number of tracks fall markedly as one looks at larger impact parameters, but the CFT efficiency falls rapidly as well. The CFT track-finding "roads" (see Section 3.7.5) are designed expecting a track to come from the beam position, $x = y = 0$ and are thus much less efficient for tracks offset by a large distance from the beam position.

In general, with the tau's lifetime of 291 ns and $c\tau$ of approximately 90 μm, tracks from tau daughter particles are displaced only a few mm from the beam position, nearly always much less than 3 mm away from the beam position. A 40 GeV tau, with $\gamma \approx 23$, travels about 2 mm, on average, before decaying, thus leaving a track with an

impact parameter of *less than* 3 mm since the tau's mass is so small compared with its transverse momentum when it is a decay daughter of a W.

The next two quality cuts require the track position and direction to be well measured in at least two axial and two stereo layers. A track is considered well-measured in a given superlayer if it generates a hit in two out of the six wires in a stereo superlayer or four out of the twelve wires in an axial superlayer. A track which is not well measured in at least two axial layers is not likely to be found by the CFT. Although the CFT uses only the axial layers for track finding, a track which is not well measured in at least two stereo layers is not likely to be genuine. Reassuringly, tracks with zero or one well-measured stereo superlayers are quite inefficient in the CFT.

The final track quality cut requires the track to pass through the full radius of the CTC. This quality cut is applied by extrapolating the track from its origin (near the beam position) to the radius of CTC wire layer 78, or the 6th wire (of twelve) in the outer wire superlayer. If the track is outside the volume of the CTC at this radius, that is, $z_{extrap} > 150$ cm, then CFTSIM does not simulate this track in the CFT.

Using JET-20 data, the efficiency of CFT P_T bin 2 (the P_T bin used for tau triggers) is measured as a function of track momentum. The efficiency is parameterized as a function of $1/P_T$ (thus, essentially a function of track curvature) as follows:

$$\varepsilon(x') = \frac{\varepsilon_{\text{plateau}}}{2}\text{erfc}(x'); \quad x' = \frac{\frac{1}{P_T} - \frac{1}{P_{T_0}}}{\sqrt{2}\,\sigma}$$

where $\text{erfc}(x)$ is the complementary error function, defined as

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$$

The plateau efficiency for high momentum tracks, $\varepsilon_{\text{plateau}}$, is measured to be 96.3%. The value of P_T at which the trigger is measured to be 50% efficient, P_{T_0}, is $P_T = 3.8$ GeV. Note that this definition of P_{T_0} is different from the one employed by the usual bin identification of the CFT, where the bins are identified by the transverse momentum where they are 90% efficient, not 50% efficient. P_T bin 2 is 90% efficient for a transverse momentum of 4.8 GeV and 50% efficient for a transverse momentum of 3.8 GeV. Finally, the error function turn-on width, σ, is measured to be 0.05 GeV^{-1}.

To simulate the CFT efficiency for identifying a track, a random number 0–1 is thrown for each offline track that passes the quality cuts. If this number is less than $\varepsilon(x')$ then a CFT track of P_T bin 2 is simulated with appropriate parameters (i.e., translating the offline definition of a track to the Level 2 CFT definition of a track).

For very low momentum tracks with transverse momentum below 3.4 GeV, the CFT has a small linear efficiency equal to $P_T \times 0.048$ %. This extra efficiency above what is expected from the CFT resolution, is caused by the extra tracks (thus, extra hits) present in jet events. Tracks which are isolated in ϕ do not have this extra efficiency. Still, to be slightly over-efficient, this extra efficiency is modeled for all tracks. This is a small correction to the CFT efficiency.

Appendix C

The Tevatron

The proton beam in the Tevatron is created and initially accelerated in four other accelerators before being injected into the Tevatron. Thus, the five accelerators the proton beam traverses, in sequential order, are the Preaccelerator, the Linac, the Booster, the Main Ring, and the Tevatron [49]. In addition to these accelerators are two rings which together make up the antiproton source and storage ring: the Debuncher and the Accumulator. Figure 3.1 displays the layout of the accelerator complex.

C.1 Creating the Proton Beam

Proton bunches start in the preaccelerator, where H_2 gas is ionized to create H^- ions in the negative hydrogen ion source. These H^- ions are then accelerated to 750 keV in an electrostatic accelerating column powered by a commercial Cockcroft-Walton generator. This 750 keV H^- ion beam is injected into the Linac with a frequency of 15 Hz.

The Linac is a 150 meter long Alvarez drift-tube accelerator which produces a pulsed 200 MeV beam of H^- ions. The ions are accelerated through a series of nine electrically resonant radio frequency (RF) tanks; the electric field in each cylindrical RF tank resonates at 201.24 MHz. Inside and along the axis of each tank, a large number (23–59) of drift tubes shield the ions from the deaccelerating phase of the RF field. These drift tubes are separated by a small gap in which ions are accelerated by the RF field. Drift tubes increase in length along the beamline such that ions in each gap experience an accelerating field. Alternating focussing and defocussing quadrupole magnets in the drift tubes keep the ion bunches focussed. H^- ions out of phase with the RF field do not experience a net acceleration; thus, each RF cycle (or "bucket") contains one "bunch" of particles, bunches being separated by a gap.

After the Linac, the ions drift 46 meters down a transport line, during which the ion bunches spread longitudinally due the small momentum spread of the particles in each bunch. These bunches pass through an RF debuncher, an RF cavity placed so as to accelerate slower particles and deaccelerate faster particles, thus minimizing the momentum spread of ions in the beam before injection into the Booster.

The Booster is a 75.5 meter radius 8 GeV fast-cycling proton synchrotron. The H^- ions injected into the booster first travel through a dogleg—two dipoles of opposite polarity which displace the beam—then through a stripping foil, and then through another dogleg. The carbon stripping foil removes electrons from the H^- ions. A proton beam in the Booster follows a stable orbit in the Booster and also passes through both doglegs and the stripping foil. Thus, the second dogleg merges the new batch of protons (H^- ions from the Linac stripped of both electrons) with protons already in the Booster. H^- ions not stripped of both electrons are directed into a beam dump. The Booster fills in six turns with 3×10^{12} protons, at which point the doglegs are turned off to reduce losses from scattering in the foil.

Once the Booster is filled, the RF stations in the booster are turned on and brought into proper relative phase, thus capturing the proton beam into an RF bucket structure of 84 buckets. Increasing the RF frequency from 37.9 MHz to 52.813 MHz accelerates the proton beam to an energy of 8 GeV. This acceleration takes about 33 msec and, like the Preaccelerator and Linac, the Booster cycles at 15 Hz, or about 67 msec/cycle. The 8 GeV proton beam is now ready to be injected into the Main Ring. At least one of the 84 bunches is lost during injection due to the time the injection magnets require to turn on. When injecting proton bunches into the Main Ring which are to be injected into the Tevatron, 15 of the 83 possible bunches are injected, but all 83 possible bunches are injected into the Main Ring when the protons are to be injected into the Antiproton Source.

The Main Ring is a 400 GeV proton synchrotron with a 1 km radius which operates with 1113 RF buckets at a frequency of about 53 MHz. The Main Ring comprises 774 dipole magnets, 240 quadrupole magnets, and 18 RF cavities. The dipole magnets maintain a circular orbit in the Main Ring; the quadrupole magnets alternately focus

and defocus the beam. The Main Ring accelerates protons to 150 GeV to be injected into the Tevatron, or to 120 GeV to be injected into the Antiproton Source.

Before injecting protons into the Tevatron, the Main Ring "coalesces" the 15 proton bunches from the Booster into a single bunch at the position of the central bunch (which explains why an odd number of bunches are coalesced). The Main Ring RF is detuned so that bunches can cross bucket boundaries, and seven special coalescing RF cavities are turned on. These coalescing cavities provide a linear restoring force which causes all the bunches to reach the central bucket at the same time, at which point the normal Main Ring RF is restored and the coalescing RF cavities are turned off.

After coalescing, a typical bunch contains about 150×10^9 protons. Antiproton bunches are coalesced in a similar manner, although with only 11 bunches, resulting in a bunch of about 50×10^9 antiprotons. The coalescing efficiency drops rapidly for more than 11 bunches, and antiprotons are "expensive" to create, so 11 bunches are injected into the Main Ring from the Antiproton Source. Once coalesced, the proton or antiproton bunch is "cogged," that is, rotated in phase in the Main Ring to match the Tevatron bucket it is injected into. Each proton bunch is injected one at a time into the Tevatron, so this whole process is repeated six times when injecting a proton beam into the Tevatron.

The Main Ring is also used to inject protons into the Antiproton Source. The Main Ring can complete a pbar cycle every 2.4 seconds, thus, continuing to supply the Antiproton Source after the Tevatron is filled with six proton bunches.

C.2 Creating the Antiproton Beam

Antiproton stacking is the process of creating, collecting, accumulating, and storing antiprotons to be injected into the Main Ring, then the Tevatron. Since about 10^7 antiprotons are produced for every 10^{12} protons injected into the antiproton source, the Main Ring must go through many pbar cycles for the Antiproton Source to collect the desired number of antiprotons. After about 12–20 hours of "stacking," the antiproton stack contains enough antiprotons to inject about six times 50×10^9 antiprotons into

the Main Ring, i.e., to inject six bunches of that size, one at a time, into the Tevatron.

During antiproton stacking, the Main Ring injects 120 GeV protons into the Antiproton Source, which consists of nickel target, a Debuncher ring, an Accumulator ring, and transport lines. After protons from the Main Ring strike the nickel target, the particles produced in the target are focussed with a 15 cm long, 1 cm radius cylindrical lithium lens. A 0.5 MA pulsed current passed longitudinally through the lithium lens produces an azimuthal magnetic field which focuses these particles. Lithium is the least dense solid conductor, so it absorbs and scatters the least possible number of antiprotons. A pulsed dipole magnet selects 8 GeV negatively-charged particles into a transport line for the Debuncher.

The Debuncher is a triangle-shaped ring designed to reduce the antiproton's momentum spread through RF bunch rotation and adiabatic debunching. Initially, the Debuncher has 90 RF buckets—84 consecutive buckets occupied and the remaining 6 empty. Thus, when the beam is debunched, there is a gap in the otherwise continuous ribbon of antiprotons in the Debuncher. This gap allows the antiprotons to be injected into the Accumulator without losses; the injection magnet is turned on during the gap. Since a Main Ring pbar cycle is 2.4 seconds long, the Debuncher has more than 2 seconds to "cool" the antiproton beam before injecting it into the Accumulator. Betatron stochastic cooling reduces the transverse profile of the beam by a factor of two during this time.

On injection into the Accumulator, the antiproton beam energy is 8 GeV $\pm$ 18 MeV. The Accumulator is also triangular and resides in the same tunnel as the Debuncher, although the Debuncher is about 6.6% larger. Since the 8 GeV antiprotons are relativistic, different energy antiprotons occupy different size orbits in the Accumulator. Antiprotons from the Debuncher are injected into the Accumulator 80 mm from the central orbit. A 53 MHz RF system adiabatically captures and deaccelerates the antiprotons by about 60 MeV, and the beam is then debunched by adiabatically reducing the RF voltage. The antiprotons are now in the "tail" of the antiproton stack.

After about an hour of cooling, the antiprotons are cooled to an energy about 150 MeV lower than the injection energy of 8 GeV and to an orbit 63 mm inside

the central orbit. The antiprotons in the stack core are further cooled to reduce the momentum spread and the vertical and horizontal size of the beam. The size of the antiproton stack grows with time as more antiprotons are injected into the Accumulator.

When the antiproton stack is large enough, some fraction of the stack—usually totalling about 1/2—is rebunched and injected into the Main Ring. A fraction of the antiproton core is captured into an extraction orbit, which is the same as the injection orbit. A 53 MHz RF system bunches the antiprotons into 11 bunches for transfer to the Main Ring.

Upon injection into the Main Ring, the antiprotons are accelerated from 8 GeV to 150 GeV, then the 11 bunches are coalesced and cogged, as described in the previous section. The coalesced antiproton bunch is ready to be injected into the Tevatron; this process is repeated six times to inject six antiproton bunches, each with about 50×10^9 antiprotons, into the Tevatron. Since protons and antiprotons carry opposite charge, they can travel in the same beam pipe and be accelerated in opposite directions to prepare for proton-antiproton collisions.

C.3 The Tevatron

Thus, counter-rotating proton and antiproton beams of 150 GeV are injected, one at a time, into the Tevatron from the Main Ring. The Tevatron then accelerates these beams to 900 GeV. Not only is the Tevatron the highest energy collider in the world, with 1.8 TeV center-of-mass energy, but it is also the first large-scale superconducting synchrotron. The Tevatron resides in the same tunnel as the Main Ring and has the same radius of 1 km. It comprises 774 dipole magnets and 216 quadrupoles, similar to the Main Ring. The Tevatron also shares the 53 MHz, 1113 bucket structure of the Main Ring. A bunch occupies about 1 meter of a 5.6 m long bucket, and bunches are separated by 186 or 187 RF buckets. A bucket requires about 18.8 ns to pass a given spot in the accelerator, so bunches are separated by about 3.5 μs.

Electrostatic separators separate the proton and antiproton bunches transversely, except in those regions where the beams are designed to collide. In the two collision

regions (at CDF and her sister experiment DØ), special superconducting quadrupole magnets (called low-beta quads) squeeze the beam spot to about 40 μm across, increasing the luminosity. The longitudinal bunch length is about 60 cm.

Once the proton and antiproton bunches are injected and ramped up to 900 GeV each, the low beta quads are turned on and CDF and DØ can begin taking data. The period when the Tevatron beam energy is maintained at 900 GeV and the beams collide at CDF and DØ is called a "store." A store continues until the luminosity drops sufficiently, or until the store is lost due to something going wrong. Typically, during Run 1a, a store lasted about 12 hours. Antiproton stacking continues during Tevatron operation. When the antiproton stack is sufficiently large and the luminosity in the Tevatron decayed, the beam in the Tevatron is dumped so new bunches can be injected. In the two hours (minimum) between stores, CDF calibrates its detector components in order to have the best possible calibration for the next store.

Although most of the luminosity of the Tevatron is contained in the six proton and six antiproton bunches, each bunch is lead and trailed (by one RF bucket, or 18.8 ns and 5.6 m) by a "satellite" bunch of lower intensity. Since each proton and antiproton bunch is approximately 60 cm long longitudinally and the bunches collide in the center of the CDF detector, each bunch crossing lasts about 1.8 ns, and an entire bucket requires 18.8 ns to cross the center of the CDF detector. Thus, 9 ns before and after the proton-antiproton bunch collisions, a satellite (p or $\overline{\mathrm{p}}$) bunch collides with a main ($\overline{\mathrm{p}}$ or p) bunch, approximately 5.6 m from the center of the CDF detector. Such collisions can be clearly vetoed due to not creating an event vertex inside the central detector, but the particles from these collisions can still add energy to a main bunch crossing interaction. Also, the satellite bunches cross each other at the center of the CDF detector about 18 ns following the main bunch crossing. Timing information in the calorimeters and track chambers must be used to reject the energy and tracks from such additional $\overline{\mathrm{p}}$p interactions.

Appendix D

The CDF Collaboration

F. Abe,[13] M. G. Albrow,[7] S. R. Amendolia,[22] D. Amidei,[16] J. Antos,[28] C. Anway-Wiese,[4] G. Apollinari,[26] H. Areti,[7] M. Atac,[7] P. Auchincloss,[25] F. Azfar,[21] P. Azzi,[20] N. Bacchetta,[20] W. Badgett,[16] M. W. Bailey,[18] J. Bao,[35] P. de Barbaro,[25] A. Barbaro-Galtieri,[14] V. E. Barnes,[24] B. A. Barnett,[12] P. Bartalini,[22] G. Bauer,[15] T. Baumann,[9] F. Bedeschi,[22] S. Behrends,[3] S. Belforte,[22] G. Bellettini,[22] J. Bellinger,[34] D. Benjamin,[31] J. Benlloch,[15] J. Bensinger,[3] D. Benton,[21] A. Beretvas,[7] J. P. Berge,[7] S. Bertolucci,[8] A. Bhatti,[26] K. Biery,[11] M. Binkley,[7] F. Bird,[29] D. Bisello,[20] R. E. Blair,[1] C. Blocker,[3] A. Bodek,[25] W. Bokhari,[15] V. Bolognesi,[22] D. Bortoletto,[24] C. Boswell,[12] T. Boulos,[14] G. Brandenburg,[9] C. Bromberg,[17] E. Buckley-Geer,[7] H. S. Budd,[25] K. Burkett,[16] G. Busetto,[20] A. Byon-Wagner,[7] K. L. Byrum,[1] J. Cammerata,[12] C. Campagnari,[7] M. Campbell,[16] A. Caner,[7] W. Carithers,[14] D. Carlsmith,[34] A. Castro,[20] Y. Cen,[21] F. Cervelli,[22] H. Y. Chao,[28] J. Chapman,[16] M.-T. Cheng,[28] G. Chiarelli,[22] T. Chikamatsu,[32] C. N. Chiou,[28] L. Christofek,[10] S. Cihangir,[7] A. G. Clark,[22] M. Cobal,[22] M. Contreras,[5] J. Conway,[27] J. Cooper,[7] M. Cordelli,[8] C. Couyoumtzelis,[22] D. Crane,[1] J. D. Cunningham,[3] T. Daniels,[15] F. DeJongh,[7] S. Delchamps,[7] S. Dell'Agnello,[22] M. Dell'Orso,[22] L. Demortier,[26] B. Denby,[22] M. Deninno,[2] P. F. Derwent,[16] T. Devlin,[27] M. Dickson,[25] J. R. Dittmann,[6] S. Donati,[22] R. B. Drucker,[14] A. Dunn,[16] K. Einsweiler,[14] J. E. Elias,[7] R. Ely,[14] E. Engels, Jr.,[23] S. Eno,[5] D. Errede,[10] S. Errede,[10] Q. Fan,[25] B. Farhat,[15] I. Fiori,[2] B. Flaugher,[7] G. W. Foster,[7] M. Franklin,[9] M. Frautschi,[18] J. Freeman,[7] J. Friedman,[15] H. Frisch,[5] A. Fry,[29] T. A. Fuess,[1] Y. Fukui,[13] S. Funaki,[32] G. Gagliardi,[22] S. Galeotti,[22] M. Gallinaro,[20] A. F. Garfinkel,[24] S. Geer,[7] D. W. Gerdes,[16] P. Giannetti,[22] N. Giokaris,[26] P. Giromini,[8] L. Gladney,[21] D. Glenzinski,[12] M. Gold,[18] J. Gonzalez,[21] A. Gordon,[9] A. T. Goshaw,[6] K. Goulianos,[26] H. Grassmann,[6] A. Grewal,[21] L. Groer,[27] C. Grosso-Pilcher,[5] C. Haber,[14] S. R. Hahn,[7] R. Hamilton,[9] R. Handler,[34] R. M. Hans,[35] K. Hara,[32] B. Harral,[21] R. M. Harris,[7] S. A. Hauger,[6] J. Hauser,[4] C. Hawk,[27] J. Heinrich,[21] D. Cronin-Hennessy,[6] R. Hollebeek,[21] L. Holloway,[10] A. Hölscher,[11] S. Hong,[16] G. Houk,[21] P. Hu,[23] B. T. Huffman,[23] R. Hughes,[25] P. Hurst,[9] J. Huston,[17] J. Huth,[9] J. Hylen,[7] M. Incagli,[22] J. Incandela,[7] H. Iso,[32] H. Jensen,[7] C. P. Jessop,[9] U. Joshi,[7] R. W. Kadel,[14] E. Kajfasz,[7a] T. Kamon,[30] T. Kaneko,[32] D. A. Kardelis,[10] H. Kasha,[35] Y. Kato,[19] L. Keeble,[8] R. D. Kennedy,[27] R. Kephart,[7] P. Kesten,[14] D. Kestenbaum,[9] R. M. Keup,[10] H. Keutelian,[7] F. Keyvan,[4] D. H. Kim,[7] H. S. Kim,[11] S. B. Kim,[16] S. H. Kim,[32] Y. K. Kim,[14] L. Kirsch,[3] P. Koehn,[25] K. Kondo,[32] J. Konigsberg,[9] S. Kopp,[5] K. Kordas,[11] W. Koska,[7] E. Kovacs,[7a] W. Kowald,[6] M. Krasberg,[16] J. Kroll,[7] M. Kruse,[24] S. E. Kuhlmann,[1] E. Kuns,[27] A. T. Laasanen,[24] N. Labanca,[22] S. Lammel,[4] J. I. Lamoureux,[3] T. LeCompte,[10] S. Leone,[22] J. D. Lewis,[7] P. Limon,[7] M. Lindgren,[4] T. M. Liss,[10] N. Lockyer,[21] C. Loomis,[27] O. Long,[21] M. Loreti,[20] E. H. Low,[21] J. Lu,[30] D. Lucchesi,[22] C. B. Luchini,[10] P. Lukens,[7] J. Lys,[14] P. Maas,[34] K. Maeshima,[7] A. Maghakian,[26] P. Maksimovic,[15] M. Mangano,[22] J. Mansour,[17] M. Mariotti,[20] J. P. Marriner,[7] A. Martin,[10] J. A. J. Matthews,[18] R. Mattingly,[15] P. McIntyre,[30]

P. Melese,[26] A. Menzione,[22] E. Meschi,[22] G. Michail,[9] S. Mikamo,[13] M. Miller,[5] R. Miller,[17] T. Mimashi,[32] S. Miscetti,[8] M. Mishina,[13] H. Mitsushio,[32] S. Miyashita,[32] Y. Morita,[32] S. Moulding,[26] J. Mueller,[27] A. Mukherjee,[7] T. Muller,[4] P. Musgrave,[11] L. F. Nakae,[29] I. Nakano,[32] C. Nelson,[7] D. Neuberger,[4] C. Newman-Holmes,[7] L. Nodulman,[1] S. Ogawa,[32] S. H. Oh,[6] K. E. Ohl,[35] R. Oishi,[32] T. Okusawa,[19] C. Pagliarone,[22] R. Paoletti,[22] V. Papadimitriou,[31] S. P. Pappas,[35] S. Park,[7] J. Patrick,[7] G. Pauletta,[22] M. Paulini,[14] L. Pescara,[20] M. D. Peters,[14] T. J. Phillips,[6] G. Piacentino,[2] M. Pillai,[25] R. Plunkett,[7] L. Pondrom,[34] N. Produit,[14] J. Proudfoot,[1] F. Ptohos,[9] G. Punzi,[22] K. Ragan,[11] F. Rimondi,[2] L. Ristori,[22] M. Roach-Bellino,[33] W. J. Robertson,[6] T. Rodrigo,[7] J. Romano,[5] L. Rosenson,[15] W. K. Sakumoto,[25] D. Saltzberg,[5] A. Sansoni,[8] V. Scarpine,[30] A. Schindler,[14] P. Schlabach,[9] E. E. Schmidt,[7] M. P. Schmidt,[35] O. Schneider,[14] G. F. Sciacca,[22] A. Scribano,[22] S. Segler,[7] S. Seidel,[18] Y. Seiya,[32] G. Sganos,[11] A. Sgolacchia,[2] M. Shapiro,[14] N. M. Shaw,[24] Q. Shen,[24] P. F. Shepard,[23] M. Shimojima,[32] M. Shochet,[5] J. Siegrist,[29] A. Sill,[31] P. Sinervo,[11] P. Singh,[23] J. Skarha,[12] K. Sliwa,[33] D. A. Smith,[22] F. D. Snider,[12] L. Song,[7] T. Song,[16] J. Spalding,[7] L. Spiegel,[7] P. Sphicas,[15] L. Stanco,[20] J. Steele,[34] A. Stefanini,[22] K. Strahl,[11] J. Strait,[7] D. Stuart,[7] G. Sullivan,[5] K. Sumorok,[15] R. L. Swartz, Jr.,[10] T. Takahashi,[19] K. Takikawa,[32] F. Tartarelli,[22] W. Taylor,[11] P. K. Teng,[28] Y. Teramoto,[19] S. Tether,[15] D. Theriot,[7] J. Thomas,[29] T. L. Thomas,[18] R. Thun,[16] M. Timko,[33] P. Tipton,[25] A. Titov,[26] S. Tkaczyk,[7] K. Tollefson,[25] A. Tollestrup,[7] J. Tonnison,[24] J. F. de Troconiz,[9] J. Tseng,[12] M. Turcotte,[29] N. Turini,[22] N. Uemura,[32] F. Ukegawa,[21] G. Unal,[21] S. C. van den Brink,[23] S. Vejcik, III,[16] R. Vidal,[7] M. Vondracek,[10] D. Vucinic,[15] R. G. Wagner,[1] R. L. Wagner,[7] N. Wainer,[7] R. C. Walker,[25] C. Wang,[6] C. H. Wang,[28] G. Wang,[22] J. Wang,[5] M. J. Wang,[28] Q. F. Wang,[26] A. Warburton,[11] G. Watts,[25] T. Watts,[27] R. Webb,[30] C. Wei,[6] C. Wendt,[34] H. Wenzel,[14] W. C. Wester, III,[7] T. Westhusing,[10] A. B. Wicklund,[1] E. Wicklund,[7] R. Wilkinson,[21] H. H. Williams,[21] P. Wilson,[5] B. L. Winer,[25] J. Wolinski,[30] D. Y. Wu,[16] X. Wu,[22] J. Wyss,[20] A. Yagil,[7] W. Yao,[14] K. Yasuoka,[32] Y. Ye,[11] G. P. Yeh,[7] P. Yeh,[28] M. Yin,[6] J. Yoh,[7] C. Yosef,[17] T. Yoshida,[19] D. Yovanovitch,[7] I. Yu,[35] J. C. Yun,[7] A. Zanetti,[22] F. Zetti,[22] L. Zhang,[34] S. Zhang,[16] W. Zhang,[21] and S. Zucchelli[2]

(CDF Collaboration)

[1] *Argonne National Laboratory, Argonne, Illinois 60439*

[2] *Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40126 Bologna, Italy*

[3] *Brandeis University, Waltham, Massachusetts 02254*

[4] *University of California at Los Angeles, Los Angeles, California 90024*

[5] *University of Chicago, Chicago, Illinois 60637*

[6] *Duke University, Durham, North Carolina 27708*

[7] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510*

[8] *Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*

[9] *Harvard University, Cambridge, Massachusetts 02138*

[10] *University of Illinois, Urbana, Illinois 61801*

[11] *Institute of Particle Physics, McGill University, Montreal H3A 2T8, and University of Toronto, Toronto M5S 1A7, Canada*

[12] *The Johns Hopkins University, Baltimore, Maryland 21218*

[13] *National Laboratory for High Energy Physics (KEK), Tsukuba, Ibaraki 305, Japan*

[14] *Lawrence Berkeley Laboratory, Berkeley, California 94720*

[15] *Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

[16] *University of Michigan, Ann Arbor, Michigan 48109*

[17] *Michigan State University, East Lansing, Michigan 48824*

[18] *University of New Mexico, Albuquerque, New Mexico 87131*

[19] *Osaka City University, Osaka 588, Japan*
[20] *Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*
[21] *University of Pennsylvania, Philadelphia, Pennsylvania 19104*
[22] *Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*
[23] *University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[24] *Purdue University, West Lafayette, Indiana 47907*
[25] *University of Rochester, Rochester, New York 14627*
[26] *Rockefeller University, New York, New York 10021*
[27] *Rutgers University, Piscataway, New Jersey 08854*
[28] *Academia Sinica, Taiwan 11529, Republic of China*
[29] *Superconducting Super Collider Laboratory, Dallas, Texas 75237*
[30] *Texas A&M University, College Station, Texas 77843*
[31] *Texas Tech University, Lubbock, Texas 79409*
[32] *University of Tsukuba, Tsukuba, Ibaraki 305, Japan*
[33] *Tufts University, Medford, Massachusetts 02155*
[34] *University of Wisconsin, Madison, Wisconsin 53706*
[35] *Yale University, New Haven, Connecticut 06511*

References

[1] Lao Tsu. *Tao Te Ching.* Vintage, (1972). Translated by Gia-Fu Feng and Jane English, adapted by Edward W. Kuns.

[2] F. Halzen and A. Martin. *Quarks & Leptons: An Introductory Course in Modern Particle Physics.* John Wiley & Sons, Inc., (1984).

[3] John Bartlett. *Familiar Quotations.* Boston, Little, Brown, and company, 9th edition, (1901).

[4] U. Aglietti. Introduction to perturbative QCD. Electronic publication hep-ph/9705277, First School on Field Theory and Gravitation, URL: http://xxx.lanl.gov/abs/hep-ph/9705277, (May 1997).

[5] R.M. Barnett et al. Review of particle properties. *Phys. Rev.*, D54:1, (1996).

[6] A. D. Martin, R. G. Roberts, and W. J. Stirling. Parton distributions: A study of the new HERA data, α_s, the gluon, and $\overline{\mathrm{p}}\mathrm{p}$ jet production. *Phys. Lett.*, B387:419, (1996).

[7] H.L. Lai, J. Huston, S. Kuhlmann, F. Olness, J. Owens, D. Soper, W.K. Tung, and H. Weerts. Improved parton distributions from global analysis of recent deep inelastic scattering and inclusive jet data. *Phys. Rev.*, D55:1280, (1997).

[8] C. Balázs and C.-P. Yuan. Soft gluon effects on lepton pairs at hadron colliders. Electronic publication hep-ph/9704258, Dept. of Physics and Astronomy, Michigan State University, URL: http://xxx.lanl.gov/abs/hep-ph/9704258, (April 1997).

[9] A. D. Martin, R. G. Roberts, and W. J. Stirling. Parton distributions of the proton. *Phys. Rev.*, D50(11):6734, (1994).

[10] Fan Qun, (1997). Private communication.

[11] R.M. Barnett et al. Review of particle properties. *Phys. Rev.*, D54:1, (1996). and 1997 off-year partial update for the 1998 edition available on the PDG WWW pages (URL: http://pdg.lbl.gov/).

[12] J. Conway and C. Loomis. Kinematic dependence of lepton polarization in W decay. CDF Note 3146, CDF, (1995).

[13] F. Abe et al. Charge asymmetry in W-boson decays produced in $\overline{\mathrm{p}}\mathrm{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 74(6):850, (1995).

[14] F. Abe et al. The CDF detector: An overview. *Nucl. Instrum. and Methods Phys. Res.*, A271:387, (1988). Plus references in this article.

[15] B. Todd Huffman, (1996). Private communication.

[16] Bob Wagner, (1997). Private communication.

[17] F. Bedeschi et al. Design and construction of CDF central tracking chamber. *Nucl. Inst. Meth. Phys. Res.*, A268(50), (1988).

[18] L. Balka et al. The CDF central electromagnetic calorimeter. *Nucl. Inst. Meth. Phys. Res.*, A267:272, (1988).

[19] S. Bertolucci et al. The CDF central and endwall hadron calorimeter. *Nucl. Inst. Meth. Phys. Res.*, A267:301, (1988).

[20] Y. Fukui et al. The CDF end plug electromagnetic calorimeter using conductive plastic proportional tubes. *Nucl. Inst. Meth. Phys. Res.*, A267:280, (1988).

[21] S. Cihangir et al. The CDF forward/backward hadron calorimeter. *Nucl. Inst. Meth. Phys. Res.*, A267:249, (1988).

[22] G. Brandenburg et al. An electromagnetic calorimeter for the small angle regions of the Collider Detector at Fermilab. *Nucl. Inst. Meth. Phys. Res.*, A267:257, (1988).

[23] A. Beretvas and S. Segler. Multiple interactions as a function of instantaneous luminosity. CDF Note 3599, CDF, (1996).

[24] F. Abe et al. Measurement of the antiproton-proton total cross section at $\sqrt{s} =$ 546 and 1800 GeV. *Phys. Rev.*, D50(9):5550, (1994).

[25] F. Abe et al. Measurement of $\overline{\mathrm{p}}\mathrm{p}$ single diffraction dissociation at $\sqrt{s} = 546$ and 1800 GeV. *Phys. Rev.*, D50(9):5535, (1994).

[26] F. Abe et al. Measurement of small angle antiproton-proton elastic scattering at $\sqrt{s} = 546$ and 1800 GeV. *Phys. Rev.*, D50(9):5518, (1994).

[27] S. Belforte, P. Derwent, J. Marriner, and C. Grosso-Pilcher. σ_{bbc} Updates. CDF Note 2535, CDF, (1994).

[28] G. Foster et al. A fast hardware track-finder for the CDF central tracking chamber. *Nucl. Inst. Meth. Phys. Res.*, A269:93, (1988).

[29] J. Conway, L. Groer, E. Kuns, C. Loomis, and T. Watts. Tau identification. CDF Note 3545, CDF, (1996).

[30] P. Auchincloss, M. Campbell, P. Derwent, J. Hauser, E. Kuns, A. Roodman, E. Sexton, and T. Watts. The CDF trigger table. CDF Note 958, CDF, (1989).

[31] E. Kuns. The tau reconstruction module. CDF Note 1723, CDF, (1992).

[32] E. Kuns, T. Watts, and X. Wu. Reconstruction and filtering of taus: Modules FINDTAU, TAUFLT, and TJTFLT. CDF Note 2661, CDF, (1994).

[33] J. Conway and L. Groer. Reprocessing, preselection, and validation of Run Ia data for CDF tau analyses. CDF Note 2619, CDF, (1994).

[34] G. Marchesini et al. Herwig: A monte carlo event generator for simulating hadron emission reactions with interfering gluons. version 5.1. *Comput. Phys. Commun.*, 67:465, (1992).

[35] M. Shapiro, A. Bhatti, J. Benlloch, R. Harris, T. Rodrigo, P. Sphicas, and T. Westhusing. A user's guide to QFL. CDF Note 1910, CDF, (1992).

[36] S. Jadach, Z. Was, R. Decker, and J.H. Kuhn. The tau decay library TAUOLA: version 2.4. *Comput.Phys.Commun.*, 76:361, (1993).

[37] C. Loomis. Using TAUOLA, a decay library for polarized τ leptons, at CDF. CDF Note 2796, CDF, (1995).

[38] C. Loomis. Comparison of Run 1b isolated pions to QFL monte carlo. CDF Note 3554, CDF, (1996). and references herein.

[39] F. Abe et al. Measurement of sigma B($W \to e\nu$) and sigma B($Z^0 \to e^+e^-$) in $\overline{\mathrm{p}}\mathrm{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 76(17):3070, (1996).

[40] Leslie Groer, (1995). Private communication.

[41] D. Cronin-Hennessy, A. Beretvas, and S. Segler. Getting your luminosity. CDF Note 4317, CDF, (1997).

[42] T. Devlin. Correlations from systematic corrections to Poisson-distributed data in log-likelihood functions. CDF Note 3126, CDF, (1995).

[43] Particle Data Group. Review of particle properties. *Phys. Rev.*, D50:1276, (1994). Section 17.3.3.

[44] F. James. Minuit: Function minimization and error analysis. CERN Program Library Long Writeup D506, CERN, (1994).

[45] Charles Darwin. *The Origin of Species.* (1859). Chapter 3.

[46] W. Badgett and P. Derwent. Event z vertex cut efficiency as a luminosity correction for Run Ia. CDF Note 2703, CDF, (1994).

[47] L. Galtieri and J. Lys. Systematics on top mass from jet P_T uncertainties. CDF Note 3501, CDF, (1996).

[48] L. Galtieri and J. Lys. How well do we understand jets in Run I? Study of the jet energy scale for raw jet $E_T \geq 8$ GeV. CDF Note 3983, CDF, (1997).

[49] Joey Thompson. Introduction to colliding beams at Fermilab. Technical Report FERMILAB-TM-1909, Fermilab, (1994).

[redacted]

Edward W. Kuns

[redacted] [redacted]

1982	Teacher's Aid for summer school program teaching computers to gifted 6th–8th grade students.
1983	Graduated from Peabody Magnet High School, Alexandria, Louisiana.
1983-87	Attended Rensselaer Polytechnic Institute, Troy, New York. Majored in Physics.
1986	Summer job, programmer for RPI Chemistry Department.
1987	B.S., Rensselaer Polytechnic Institute.
1987	Summer job, programmer for RPI Chemistry Department.
1987-97	Graduate work in Physics, Rutgers, The State University of New Jersey, New Brunswick, New Jersey.
1987-88	Teaching Assistant, Department of Physics and Astronomy.
1989-97	Research Assistant, Department of Physics and Astronomy.
1989-97	List of publications available upon request.
1998	Ph.D. in Physics, Rutgers, the State University of New Jersey.