A Measurement of BR($t \to \tau\nu q$)

by

Sarah Marie Demers Konezny

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Kevin McFarland

Department of Physics and Astronomy

The College

Arts and Sciences

University of Rochester

Rochester, New York

2004

Curriculum Vitæ

The author graduated with a Bachelor of Arts degree in Physics from Harvard University. She continued her physics education at the University of Rochester, receiving her Master of Arts degree, and pursuing high energy physics research under the direction of Professor Kevin McFarland.

Acknowledgements

I would like to thank the CDF Collaboration for building and maintaining an awesome detector. This analysis builds on so many hours of the work of others that the only pieces of it that I can claim as totally my own are the mistakes I have made along the way. JJ Schmidt and Kirsten Tollefson in particular among the CDF crew have been wonderful friends and sources of support but I am grateful to everyone I have worked with on the experiment. I have missed and will continue to miss all of you at Fermilab.

I want to thank my advisor, Kevin McFarland, for sharing so much physics knowledge and for his work on the front lines of the analysis. I can not begin to quantify how much physics I have learned from Kevin, or just as importantly, how much I have learned about how to DO physics from Kevin. He is the very best in the business. On the non-physics front, I am grateful for his support of all of my decisions, including the crazy ones. And, of course, I need to thank him for the gallons of espresso that he has shared with me over the years we have worked together.

Thanks to Melissa Franklin and John Huth for giving me the initial opportunity to become involved with particle physics.

I would like to thank the faculty and the staff at the University of Rochester for making the physics program everything they claimed it would be during recruiting weekend. Connie Jones and Barbara Warren in particular have helped me out of many binds and I will be forever grateful.

Special thanks go out to my in-laws, John and Ann Konezny. They have fed me, housed me, encouraged me, listened to me, and on top of it all provided me with a spouse! What more could I ask for? My extended family has been a wonderful source of support. I would like to thank my parents, Mark and Jan Demers, for their wisdom and their faith in me. I have always had a bottomless well of encouragement at most a phone call away and I could not ask for better friends.

I am grateful for the frequent reminders that life is not on pause while we attend school that have come from my husband, Steven. Thanks for understanding EXACTLY what I am talking about so much of the time.

Abstract

A search for top, anti-top events with a semi-leptonically decaying tau lepton in the final state is presented. The data sample corresponds to an integrated luminosity of 194 pb^{-1} and was collected with the upgraded Collider Detector at Fermilab using proton, anti-proton collisions at a center-of-mass energy of 1.96 TeV. We observe two events with an expected background of 1.3 $\pm$ 0.3 events. We set a 95$\%$ confidence level upper limit on an anomalous rate enhancement factor, $r_\tau <$ 5.0, consistent with the Standard Model prediction of $r_\tau = 1$.

Contents

1 Introduction **1**

2 Physics Motivation **5**

2.1 The Standard Model 5

2.1.1 Bosons 6

2.1.2 Fermions 7

2.1.3 Quantum Electrodynamics 9

2.1.4 The Weak Force 10

2.1.5 Quantum Chromodynamics 13

2.1.6 The Top Quark 13

2.2 $t \to \tau$ 18

3 Collider Detector at Fermilab **20**

3.1 Accelerator Chain 20

3.2 CDF 22

3.2.1 Luminosity 24

3.2.2 Tracking 25

3.2.3 Calorimetry 28

3.2.4 Muon Detectors 32

3.2.5 Trigger 35

4 Event Selection **38**

4.1 Triggers 40

4.2 Electron Identification 43

4.3 Muon Identification 46

4.4 Jets 49

4.4.1 b tagging jets 51

4.5 Tau Reconstruction and Identification 52

4.6 Kinematics 55

4.7 Z Mass Veto 57

4.7.1 Monte Carlo Samples Used 58

4.7.2 Angular Cuts 61

4.7.3 Mass Reconstruction 63

4.8 Optimization of H_T and Jet E_T 66

5 Acceptance **69**

5.1 Acceptance from signal Monte Carlo 71

5.2 Systematics 75

6 Tau Identification in $W \to \tau\nu$ **80**

6.1 Event Selection 81

6.2 Corrections to Monte Carlo 82

6.3 Backgrounds 85

6.4 Scale Factor 86

7 Backgrounds **94**

7.1 Fakes 95

7.1.1 $j \to \tau$ Fakes 95

7.1.2 $e \to \tau$ Fakes 112

7.1.3 $\mu \to \tau$ Fakes 118
7.2 $\gamma^+/Z \to \tau\tau + jets$ 128
7.3 WW 129
7.4 WZ 129
7.5 Background Summary 130

8 Results **131**
8.1 Control Sample: Low Jet Multiplicity 131
8.2 The Unblinded Data Samples 136
8.3 r_τ 136

9 Conclusions **142**

List of Figures

2.1 The three generations of quarks and leptons 8
2.2 Fundamental vertex of Quantum Electrodynamics 10
2.3 Fundamental vertices of Weak Interactions 11
2.4 Decay via the Weak Force 12
2.5 Fundamental vertices of Quantum Chromodynamics 14
2.6 Parton Distribution Functions 16
2.7 $t\bar{t}$ Production 17
2.8 Standard Model Constraint on the Higgs Mass 19

3.1 The Accelerator Chain at Fermilab 21
3.2 CDF Detector Schematic 23
3.3 Delivered and Recorded Luminosity 25
3.4 Silicon Layers in r-z. 26
3.5 A cell in the COT 29
3.6 End Plug Calorimeter 31
3.7 Muon Coverage (η - ϕ 33

4.1 Electron ID efficiency, plotted as a function of ϕ and detector η. 47
4.2 Example event for Z Mass Veto 59
4.3 Example event for Z Mass Veto 60

4.4 $t\bar{t}$ and $\gamma^*/Z \to \tau\tau$ angular distributions 62
4.5 Reconstructed mass for events passing angular cuts. 65
4.6 Integral event fractions surviving H_T and E_T cuts. 66
4.7 Optimization variables for H_T and leading jet E_T cuts. 68

6.1 Jets in $Z \to \mu\mu$ data and Monte Carlo 84
6.2 Predicted jet$\to \tau$ fake background in $W \to \tau\nu$ 87
6.3 $W \to \tau\nu$: tau track multiplicity 89
6.4 $W \to \tau\nu$: tau cal iso ratio 90
6.5 $W \to \tau\nu$: tau track + π^0 mass 91
6.6 $W \to \tau\nu$: tau track + π^0 P_t 92
6.7 $W \to \tau\nu$: $\Delta\phi$ between tau and $\not{E}_{\mathrm{T}}$ 93

7.1 Distributions for Level1, Level2 and Level3 trigger objects. 97
7.2 $\mathrm{E_T}$ and pseudo-rapidity distributions for biased and unbiased jets. . . . 98
7.3 $\Delta\mathrm{R}$ distance from each jet nearest tau candidate. 99
7.4 τ ID variables for fake rate denominator 101
7.5 τ ID variables for fake rate numerator 102
7.6 More τ ID variables for fake rate numerator 103
7.7 Comparison for fake rate from JET20, JET50, JET70 and SUMET . . . 106
7.8 Isolation variable after all tau ID cuts are applied 107
7.9 Relative fake rates in each isolation region. 108
7.10 Jet to tau fake background 111
7.11 Invariant mass in $Z \to ee$ events 114
7.12 Electron $E_{HAD}/\Sigma P$ 115
7.13 $E_{HAD}/\Sigma P$ for electron background 117
7.14 $M_{\mu\mu}$ in data and the Monte Carlo 120
7.15 $M_{\mu\mu}$ in data and the Monte Carlo below the Z pole 121

7.16 Hadronic energy deposition for muon candidate in the data 123
7.17 Hadronic energy for muon candidates in background subtracted data and Monte Carlo 124
7.18 Hadronic energy for muon candidates in background subtracted data and Monte Carlo (10–100 GeV) 124
7.19 E_T/p_T (muon veto) for muons which are tau candidates in background subtracted data and in Monte Carlo 126

8.1 Jet multipicity *a priori* test of the data probability 135
8.2 Event display for run 167299, event 2376337. 138
8.3 The unnormalized probability distribution $P(N_{\rm obs} = 2 \| r_\tau)$ 140

List of Tables

2.1 Relative strength of forces averaged at low energy. 6
2.2 Properties of the Quarks. 8
2.3 Properties of the Leptons. 9
2.4 Branching Ratios of W decays 11

3.1 Properties of Muon Detectors 34

4.1 Final states of $t\bar{t}$ events with their branching ratios 39
4.2 Tight Electron Identification Cuts 44
4.3 Tight CMUP and CMX Identification Cuts 48
4.4 Tau Identification Cuts 54
4.5 $\gamma^*/Z \rightarrow \tau\tau$ Monte Carlo samples 58
4.6 Effect of Z mass removal 64

5.1 Acceptance Scale Factors 71
5.2 Top Monte Carlo after cuts 74
5.3 Top Monte Carlo after cuts (reordered) 74
5.4 Acceptance of tau identification cuts 76
5.5 Systematic uncertainty due to choice of generator. 76
5.6 Systematic uncertainty due to ISR. 77
5.7 Systematic uncertainty due to FSR 78

5.8 Summary of Systematics for Acceptance 79

7.1 Predictions of events passing all tau cuts from each jet sample 109
7.2 Electron Fake Rate as a function of veto variable 113
7.3 $e \to \tau$ fake dependence on $\not{E}_{\mathrm{T}}$ 116
7.4 Effect of cumulative cuts on the data and the Monte Carlo 119
7.5 $\mu \to \tau$ Fake rates calculated from data and Monte Carlo $Z \to \mu\mu$ Samples 123
7.6 Backgrounds from $\gamma^*/Z \to \tau\tau + jets$ 129
7.7 e, τ channel signal and background predictions 130
7.8 μ, τ channel signal and background predictions 130

8.1 Jet multiplicity table for e-tau channel, opposite sign events. 133
8.2 Jet multiplicity table for mu-tau channel, opposite sign events. 133
8.3 Jet multiplicity table for e-tau channel, same sign events. 134
8.4 Jet multiplicity table for mu-tau channel, same sign events. 134
8.5 Consistency tests applied to the low jet multiplicity data 137
8.6 Events passing final jet multiplicity cut 137
8.7 Events predicted and found in the signal region 137
8.8 Events predicted and found in the signal region in Run1 137
8.9 Run 167299 event 2376337 details. 140
8.10 Run 151434 event 158200 details. 140

Chapter 1

Introduction

In order to gauge our understanding and progress in physics we make a hypothesis about the laws governing an interaction and we then develop an experiment that allows us to compare our hypothesis with what we measure. In the case of Particle Physics we have organized all interactions into categories governed by four forces and we have developed a well defined model, known as the standard model, that encompasses three of the four forces and all particles that have been discovered. The standard model gives a prediction for how these particles should interact and decay and it has held up under scrutiny to amazing levels of accuracy. A more technical discussion of the standard model can be found in the following chapter but for now I will focus on the pieces that are particularly interesting, in terms of this analysis, in a more general way.

An example of the organizational structure in the standard model is the categories into which we separate particles. Each particle is either a fermion or a boson. The fermions can be broken into categories of quarks and leptons. Combinations of quarks make up hadrons, which can further be broken into categories of mesons, with two quarks, and baryons, with three quarks. This language, while it may seem thick, allows us to acknowledge symmetries, or patterns of behavior, within the model. A further

piece of organization in the model is the separation of quarks and leptons into three generations, and it is these generations that I would like to examine further.

All of the matter that we are familiar with in our normal, non-accelerated lives consists of first generation particles. The electron is a first generation lepton and the proton and neutron are both made up of first generation quarks. First generation particles are lighter than their second and third generation partners, and unlike particles in the higher generations they are stable, meaning they do not decay. If this was not the case we would not exist for very long in our present forms because our atoms, made up of first generation particles, would find their building blocks of electrons, protons and neutrons decaying into other particles! In order to create the more massive particles of the higher generations we need the energies and densities that are available when accelerated particles collide at facilities like the TeVatron at Fermilab.

The research described in this thesis is the search for a third generation quark, the top, that decays to third generation leptons, the tau and tau neutrino and other particles. The top quark is produced mostly with its antimatter partner, the anti-top. These pairs are rarely created, making up only about 10 of every trillion collisions in the detector. The top (anti-top) quarks decay nearly always to a W boson and a b quark (b anti-quark), which is a member of the third generation. The W boson decays about 11% of the time to a tau. The tau, as a third generation particle, does not exist for very long before it decays to lighter, stable particles. Its fleeting presence is difficult to detect in part because the particles it decays into are common decay products from other particles.

Why do we go through the trouble of creating and looking for these rare and tricky events with both tops and taus? First of all, the only way we know if a model has truth in it is to understand what it predicts and then to test the predictions with experiments. The standard model makes predictions about how often the collisions in our detector will result in top anti-top pairs with the tau leptons as decay products. We refer to a collision as an “event.” When we measure the frequency of the top and tau events, or

the rate, we therefore provide a check of the model.

We also have reasons for being interested in third generation particles in general. The premise is that in the early universe enough energy existed with enough density for third generation particles to exist "naturally." These particles have since then decayed away as the universe has cooled and expanded and they are now only found in high energy cosmic rays or when we create them with accelerators. This means that when we create and examine third generation particles we are in a way looking back in time to the early universe. The special case of one third generation (top or anti-top) particle decaying to another third generation particle (tau) before further decays makes this study that much more intriguing.

One unanswered question in the model is, "Why are there three generations?" What fundamental reason is there for more than one generation and for three in particular? Also, though there is evidence that there are only three generations, are there higher generations that we have not yet discovered? By studying the third generation we may make progress on the answers to these questions.

Aside from the issue of generations, these particles are interesting merely because they have high masses. Note that the standard model encompasses three of the four forces. The force that is not a part of the model is the force of gravity. This is because our understanding of gravity, while it has a long history and is impressive on the macroscopic scale of solar systems, planets, and people, is almost non-existent on the microscopic particle scale. We can hypothesize that gravity has a mediating particle just as the other four forces and we can even name the particle the graviton and realize that it must be massless but we have no experimental evidence for its existence. One governor of gravity's strength, the mass of an object or particle, is also not understood on a microscopic scale. We have no method in the model for predicting, a priori, the mass of a particle. Studying the most massive particles is therefore interesting in its own right as these particles may hold clues to the formation of mass.

Finally, specific extensions to the standard model have predicted rates for the process we study so any anomalous rate we find could point us in a new physics direction. Some of the potential processes that would give us an excess of our final state do not even originate with top quarks.

We predict 2.3 events, including the process itself and background processes that we misidentify as the process we are looking for, and we measure two events. Our result allows us to say within a confidence level of 95%* that the rate of this process is less than five times that which is predicted by the standard model, meaning it is consistent with the standard model. Improved detection methods coupled with more data will allow for improvements on this limit in the future.

*Our 95% confidence level means that there are only 5 chances out of 100 that our data is consistent with a rate that is five times higher than the standard model rate.

Chapter 2

Physics Motivation

2.1 The Standard Model

A basic premise of the standard model is that all particle interactions and decays can be described by exchanges of force carrying particles. The four forces of gravity, the strong force, the weak force and the electromagnetic force account for all known interactions. The strengths of the forces are compared in Table 2.1. The standard model in particle physics encompasses the latter three forces and all particles that have been discovered to date. The model's limited scope ensures that the current standard model is not the entire story of the laws of the universe; in addition to leaving out gravity it has nothing to say about dark energy and it is unlikely that a solution to the dark matter puzzle can be found in the model as it stands. However, the successes of the model, particularly the agreement between predictions and precise electroweak measurements[1], give many physicists confidence that the mod el is a worthwhile organization of our current understanding and a solid foundation upon which to add future discoveries. In this chapter the fundamental particles in the standard model are described and the processes that are critical to this analysis are discussed.

Force	Typical Strength	Mediator
Strong	10	g
Electromagnetic	10^{-2}	γ
Weak	10^{-13}	$W^{\pm}$, Z
Gravity	10^{-42}	Graviton

Table 2.1: Relative strength of the forces at low energy. This needs to be taken with a grain of salt because the strength of the force is a function of many things, including the energy of the interaction in question. But this chart gives one an idea of average values at low (non-accelerator) energies.

2.1.1 Bosons

The force carrying particles in the standard model are known as the gauge bosons. These particles have integer spin.

The photon (γ) is the neutral, massless boson that is the mediator of the force of electromagnetism. Because the force is massless it is able to travel long distances and electromagnetism is therefore a long range force.

Gluons (g), of which there are eight types, are the mediators of the strong force. Gluons have no electromagnetic or weak charge but they do have color, the charge of the strong force. Color charge comes in three varieties usually denoted as red, green and blue. Gluons always carry both a color and an anti-color charge. The color charge of the mediating particle of the strong force complicates the strong theory (Quantum Chromodynamics, or QCD) dramatically over the theory of electrodynamics (Quantum Electrodynamics, or QED) where the mediator (γ) is not itself charged.

The weak force has two bosons associated with it: the charged W boson (W^+ and W^-) and the neutral Z boson (Z^0). Both bosons carry the weak charge with the $W^{\pm}$ bosons carrying the electromagnetic charge as well. Unlike the other mediating bosons, the massless γ and g, the W and the Z have masses at 80.425 $\pm$ 0.038 GeV and 91.1876 $\pm$ 0.0021 GeV respectively[19]. When the weak force and the electromagnetic force are

combined in electroweak theory the masses of the W and Z bosons break the symmetry between the two forces.

The particle responsible for regulating mass is the as yet undiscovered Higgs Boson. The strength of a particle's coupling to the Higgs determines the mass of the particle.

2.1.2 Fermions

The fermions are often referred to as the matter particles, or the building block particles. When we think of particles we tend to think of fermions, particles with half-integer spin, that account for the matter that surrounds us (electrons, protons and neutrons.) The fermions can be conveniently arranged in three generations with the highest mass particles in the third generation.

The first generation consists of electron (e) and electron neutrino (ν_e) leptons and up (u) and down (d) quarks. Most of the matter we interact with, with the exception of when we are using particle accelerators, is comprised of first generation particles.

The second generation consists of muon (μ) and muon neutrino (ν_μ) leptons and charm (c) and strange (s) quarks.

The third generation consists of tau (τ) and tau neutrino (ν_τ) leptons and top (t) and bottom (b) quarks.

Figure 2.1 shows the particles organized by generation.

Each of these fermions has an anti-matter partner particle with the same mass but opposite charge, color and weak charge. Anti-matter particles of quarks and neutrinos are denoted with a bar above the particle symbol, so the anti-top quark is represented by $\bar{t}$. Anti-matter partners of the negatively charged leptons can be recognized by their positive charge. For example, the antimatter particle of the electron, e^-, is the positron, e^+.

A summary of the properties of the quarks is given in Table 2.2 and a summary of the

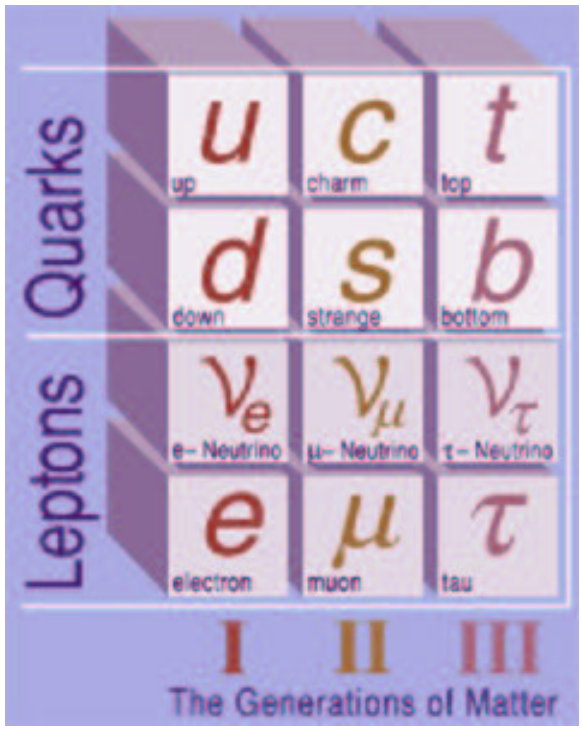

Figure 2.1: The flavors of Quarks and Leptons, organized into the three generations[3].

Quark	Electric Charge	Mass
up	+2/3e	1.5 to 4 MeV
down	-1/3e	4 to 8 MeV
strange	-1/3e	80 to 130 MeV
charm	+2/3e	1.15 to 1.35 GeV
bottom	-1/3e	4.1 to 4.9 GeV
top	+2/3e	174.3 $\pm$ 5.1 GeV

Table 2.2: Properties of the quarks. All data is from the 2004 Particle Data Group (PDG)[2]. The mass of the top quark quoted is from direct observation at CDF and D0. A standard model electroweak fit gives a mass of $178^{+10.4}_{-8.3}$ GeV. Charges are shown as multiples of the electron charge.

Lepton	Electric Charge	Mass	Lifetime
e	-1e	0.51099892 $\pm$ 0.00000004 MeV	$>$ 4.6 x 10^{26} years
ν_e	0	$<$ 3 eV	
μ	-1e	105.658369 $\pm$ 0.000009 MeV	2.19703 $\pm$ 0.00004 X 10^{-6} s
ν_μ	0	$<$ 0.19 MeV at 90% CL	
τ	-1e	$1776.99^{+0.29}_{-0.26}$ MeV	2.906 $\pm$ 1.1 X 10^{-13} s
ν_τ	0	$<$ 18.2 MeV at 95% CL	

Table 2.3: Properties of the Leptons from the 2004 PDG[2]. The mass eigenstates of the neutrinos are related to their flavor eigenstates through the MNS matrix, as described in section 2.1.4.The mixing between generations is large enough so that the mass eigenstates cannot be approximated as the flavor eigenstates as is shown in this table. This mixing does not have an impact on the analysis described in this thesis.

properties of the leptons is given in Table 2.3. The large difference between the lightest quark and the heaviest quark is worth noting as an important puzzle. Also, the fact that the quarks have fractional charge is worth pointing out. While this may seem to threaten the idea that the charge of the electron is the fundamental unit of charge, free quarks, or single quarks, have not been observed in nature. When a single quark is produced it hadronizes, combining with one or more additional quarks to form a hadron, a bound state of quarks with integer charge, on the order of 10^{-23} seconds. So while there are particles with 1/3 and 2/3 the charge of the electron in the standard model, there still does seem to be something fundamental about the integer unit of charge.

Another puzzle we have with the fermions is the existence of three generations. Why do we not have only the first generation? And why are there exactly three generations?

2.1.3 Quantum Electrodynamics

The fundamental vertex of Quantum Electrodynamics is shown in Figure 2.2. This vertex, where a photon is emitted or absorbed by a charged particle, can be used to build all interactions in electromagnetism. The strength of the force is a function of the

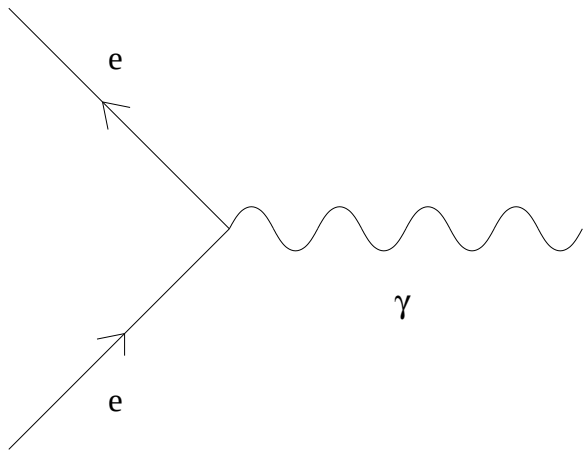

Figure 2.2: Fundamental vertex of Quantum Electrodynamics. Shown is the diagram for an electron but any quark or charged lepton could replace the electron in this diagram.

charge.

2.1.4 The Weak Force

The fundamental vertices of Weak interactions are shown in Figure 2.3. The vertex where a W boson is emitted or absorbed accounts for flavor changes. The vertex where a Z boson is emitted or absorbed accounts for the neutral interactions. The strength of the force is the same order of magnitude as the strength of the electromagnetic force. This may seem to be in conflict with the relative strength of the forces show in Table 2.1. The reason the weak force is, in practice, so much weaker than the electromagnetic force is because the very large masses of the W and Z bosons reduce the intrinsic strength of the force in the interaction.*

When a lepton decays via a W boson the decay vertex connects members of the same generation[8]. For example, if a tau decays to a tau neutrino, emitting a W^- boson, the

*The mass of the boson is in the denominator of the propagator of the weak force.

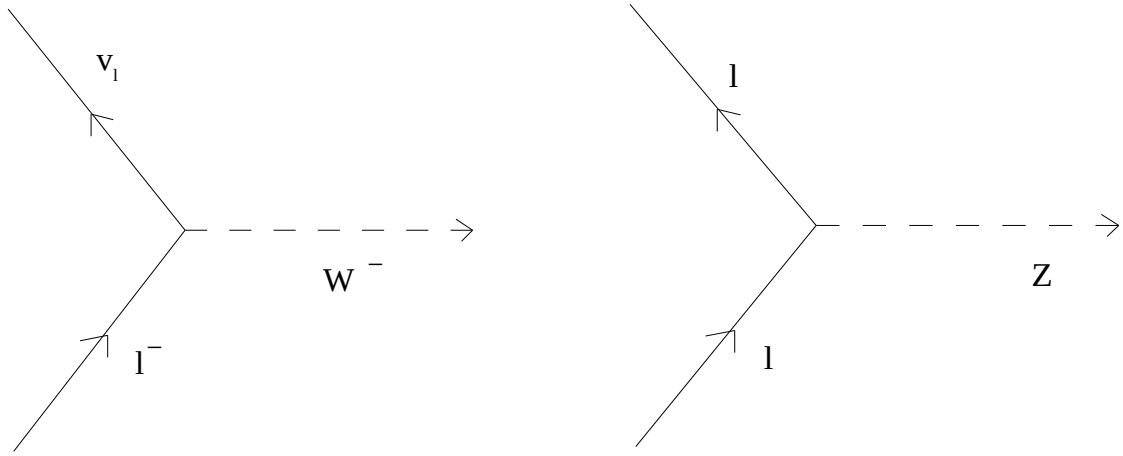

Figure 2.3: Fundamental vertices of Weak Interactions. Quarks can replace the leptons in the left diagram as long as charge is conserved at the vertex. All leptons and quarks can interact with the Z boson.

W boson may then go on to decay to μ and $\bar{\nu_\mu}$, e and $\bar{\nu_e}$, or to quarks, but at each vertex the leptons would be grouped together by generation as shown in Figure 2.4. A summary of the branching ratios of W boson decays is given in Table 2.4.

$W^\pm$ decay	Branching Ratio
$e^\pm \nu_e$	10.72 ± 0.16
$\mu^\pm \nu_\mu$	10.57 ± 0.22
$\tau^\pm \nu_\tau$	10.74 ± 0.27
hadrons	67.96 ± 0.35

Table 2.4: Branching Ratios of W Decays[19].

You might assume that the same is true for quarks, meaning that weak decays involving the W boson must stay within the same generation at each vertex. However, quark generations are more complicated than the charged lepton generations . If we use $V_{q_1 q_2}$ to denote the coupling between two quarks, q_1 and q_2, we can show the CKM Matrix (for Cabibbo, Kobayashi and Maskawa) as[7]:

V_{ud} = 0.975 to 0.9770 $\quad V_{us}$ = 0.21 to 0.24 $\quad V_{ub}$ = 0. to 0.014

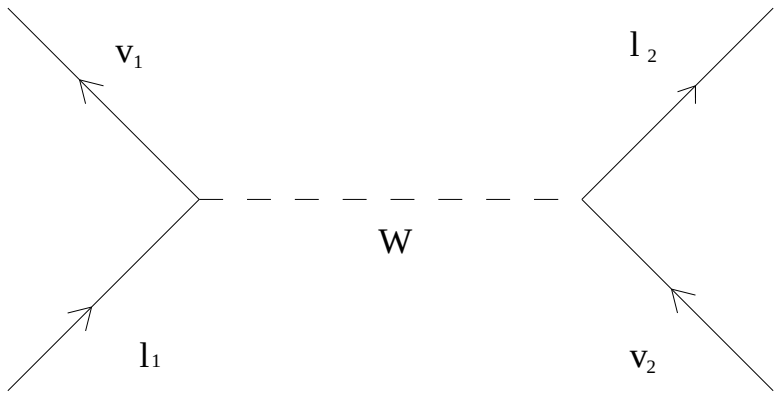

Figure 2.4: In this process a charged lepton, l_1, decays to a neutrino, v_1, and emits a W boson. The W goes on to decay to a charged lepton , l_2, and a neutrino, v_2. The subscript numbers represent the generations. At each of the two vertices involving the W only one generation is allowed.

V_{cd} = 0.21 to 0.24	V_{cs} = 0.971 to 0.973	V_{cb} = 0.036 to 0.070
V_{td} = 0. to 0.024	V_{ts} = 0.036 to 0.069	V_{tb} = 0.997 to 0.999

The non-zero off-diagonal terms show us that quarks can decay, via the weak force, outside their generations. Note that

$$V_{tb} >> V_{ts} > V_{td}$$

This tells us that the top quark decays to a W boson and a bottom quark almost 100% of the time. The decays of top to a W and s are supressed by the factor $\frac{|V_{ts}|^2}{|V_{tb}|^2}$. The top decay to a W and d is suppressed by the factor $\frac{|V_{td}|^2}{|V_{tb}|^2}$.

Recent experiments involving neutrino oscillations have shown us that our description of the neutrinos in terms of their weak flavor is not a safe approximation of their mass eigenstates[4][5][6]. Unlike in the case of quarks, where small mixing angles mean that mass eigenstates can be approximated as flavor eigenstates, there are large mixing angles in the neutrino sector. The relationship between weak eigenstates ν_ℓ, where ℓ is an e, μ, or τ, and mass eigenstates, ν_i, with i equaling 1, 2, or 3 is:

$$|\nu_\ell> \;=\; \Sigma U_{\ell i}\,|\nu_i>$$

with mixing matrix $U_{\ell i}$, known as the MNS matrix for Maki, Nakagawa, and Sakata. We refer to neutrinos as ν_e, ν_μ and ν_τ in this analysis according to the charged leptons with which they are produced, enforcing conservation of electron number, muon number, and tau number in weak decays.

2.1.5 Quantum Chromodynamics

The fundamental vertices of quantum chromodynamics are shown in Figure 2.5. The gluon-gluon interactions are a result of the color charge of the gluon. The strength of the coupling constant for the strong force, α_s, is greater than 1 but the force has a very short range.

Because gluons interact with themselves the strength of the strong force increases with distance. The result is that quarks that are close together behave as if they do not feel the strong force. This quality of the strong force is known as asymptotic freedom. As the quarks begin to separate the force of attraction between them increases. This quality of the strong force is known as confinement. As a result, the strong force is often described as the “rubber band” force. If enough energy is given to the system to separate the two quarks then additional quarks are pulled from the vacuum so that no quark exists alone but must be in the presence of at least one other quark.

2.1.6 The Top Quark

The top quark was discovered via $t\bar{t}$ events by both the CDF experiment[9] and the D0 experiment[10] in 1995 at Fermilab. The high mass of the top quark, shown in Table 2.2, was unknown when the bottom quark was found in 1977 via the discovery of the upsilon.[11]. After the discovery of the bottom quark the top was predicted as the isospin

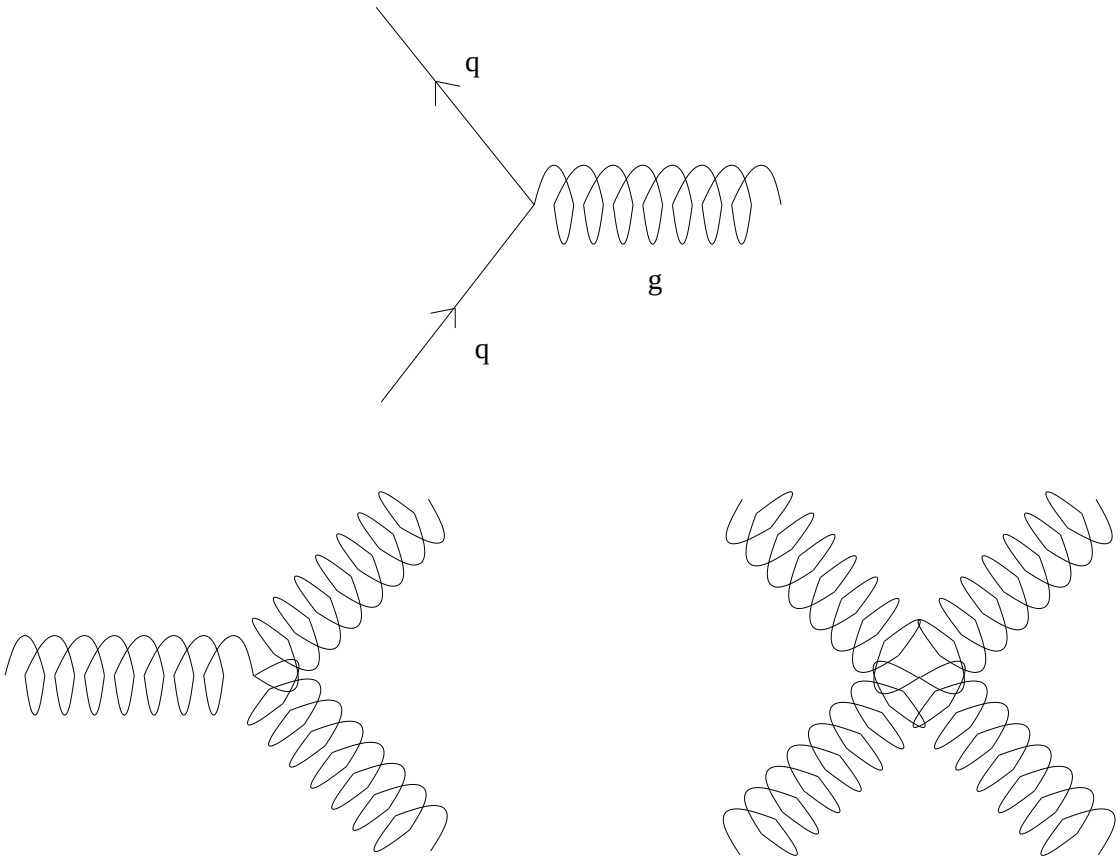

Figure 2.5: Fundamental vertices of Quantum Chromodynamics.

partner of the bottom but its surprisingly high mass, at 35 times the mass of the bottom quark, allowed it to elude detection for almost twenty years. The large amount of energy required to create top quarks continues to make studying them technically challenging. As a result, precision tests of top quark physics are difficult and there is unexplored phase space in the standard model in the top sector where new physics may be found.

The top quark was discovered during a period of data-taking at the Fermilab Tevatron from 1992 to 1996 that is identified as Run1. During Run1 protons and anti-protons were collided with a center of mass energy, $\sqrt{s}$, of 1.8 TeV. The current data-taking period, with physics quality data beginning in 2002, is referred to as Run2. The Run2 $\sqrt{s}$ is 1.96 TeV. To a good approximation, when protons and anti-protons collide it is actually a quark or gluon inside each proton or anti-proton that participates in the collision. The quarks and gluons are referred to as partons. Each parton carries a fraction, x, of the momentum of the proton or anti-proton. The probabilities corresponding to various values of x for partons in the proton are shown in Figure 2.6. In order to create a $t\bar{t}$ pair the following requirement on the colliding partons must be met:

$$x_p \cdot x_{\bar{p}} \geq \frac{4 \cdot m_t^2}{s}$$

At the Tevatron $t\bar{t}$ pairs can be created via $q\bar{q}$ annihilation or gluon-gluon fusion. Because the u and d have larger distribution functions at the x required for $t\bar{t}$ production the $q\bar{q}$ interaction accounts for 85% of the $t\bar{t}$ events. The lowest order Feynman diagrams for these processes are shown in Figure 2.7.

In addition to being the least explored known particle in the standard model the top is interesting because its high mass means it couples strongly to the Higgs boson, an important missing piece of the standard model that will answer questions about the generation of mass. The mass is also high enough for the top decay to a real W boson, a condition that none of the other quarks or leptons meet. The top is also unique among the quarks in that its lifetime is so short, at 10^{-25} seconds, that it decays before it hadronizes.

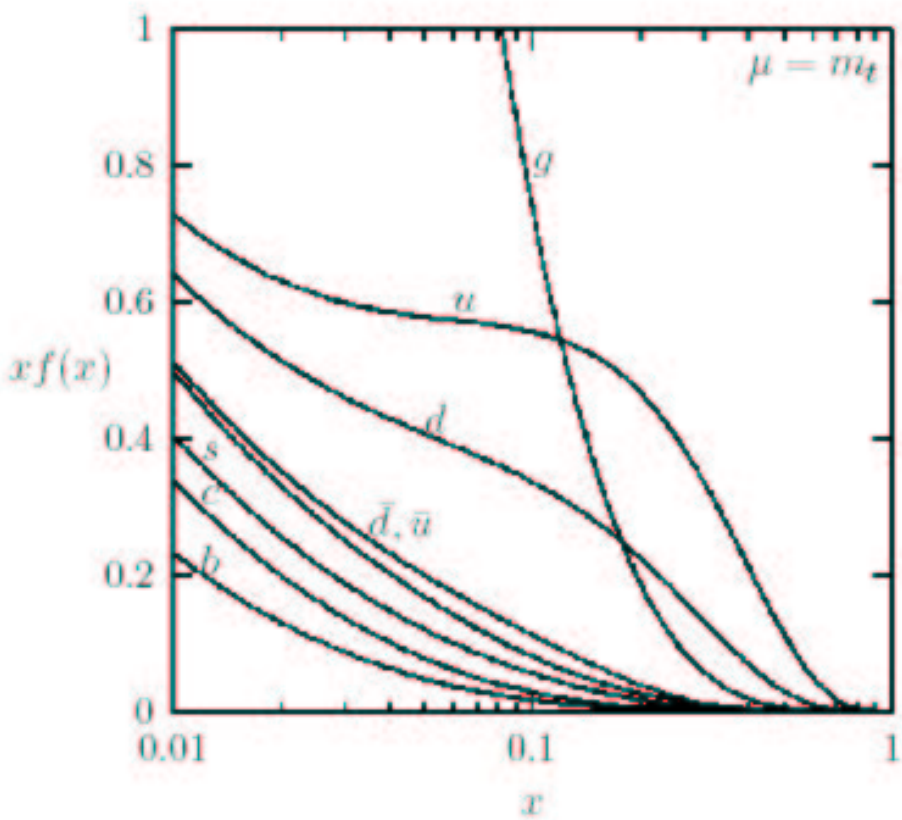

Figure 2.6: Parton Distribution Functions for partons in the proton at the scale, μ, of the mass of the top quark. Above, x·f(x) is plotted as a function of x with x equaling the fraction of the proton momentum carried by the parton and f(x) equaling the probability of finding a given parton with momentum x. The typical value of x at the Tevatron is 0.18 at this energy scale, so quarks dominate over gluons in carrying the momentum of the proton and they therefore have a higher probability of contributing to $t\bar{t}$ production[12].

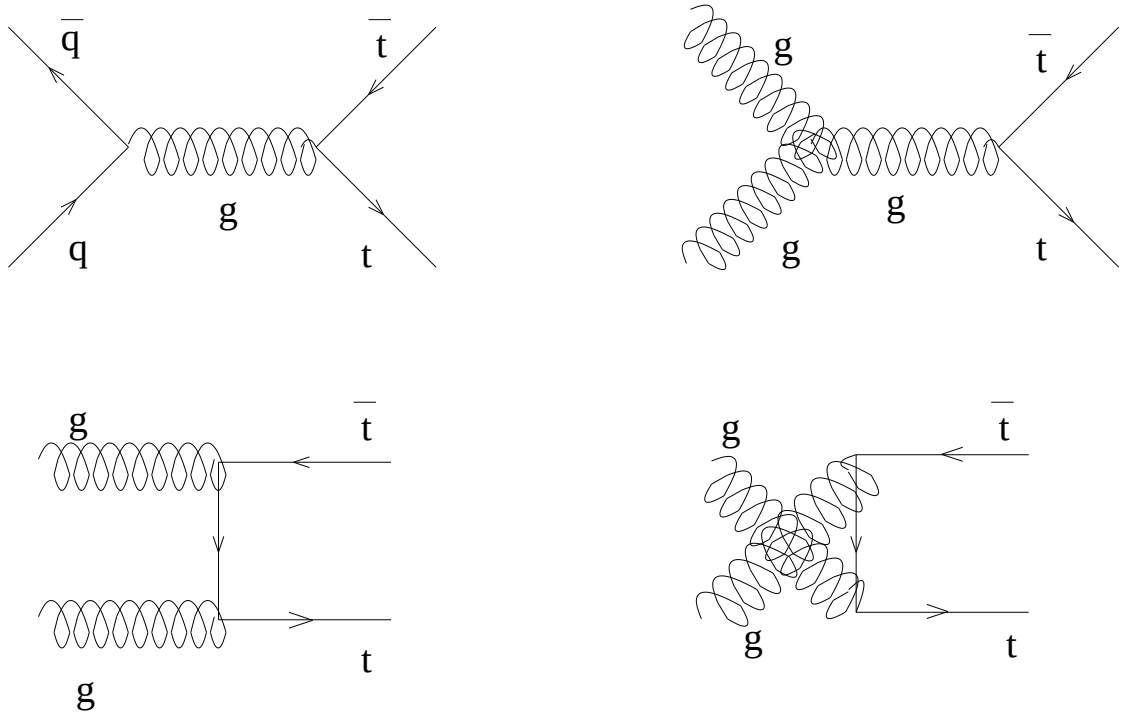

Figure 2.7: Lowest order Feynman diagrams for $t\bar{t}$ Production: Upper left shows the $q\bar{q}$ interaction and the other three diagrams show gluon-gluon fusion production. The dominant production at the Tevatron is the $q\bar{q}$ process.

Studying the top therefore gives us the opportunity to learn something about bare quarks that might not be available by studying the other quarks.

2.2 $t \to \tau$

This analysis searches for $t\bar{t}$ events with a semi-leptonic tau decay in their final state. The decay chain is interesting because it involves only third generation particles. If there is something special about the heavy third generation this is a promising place to look for it. Current measurements of the interactions and decays of these particles (top and tau) leave room for new physics[13][14][15].

As shown in Figure 2.8, the most likely mass of the standard model Higgs boson is nearly excluded by experiment. While it is possible that the standard model Higgs is just above our energy reach it is also possible than an extension to the model is needed in order to explain the generation of mass. One such extension, the minimal supersymmetric standard model, includes a charged Higgs with a lower mass than the top quark. The coupling of the Higgs, which is proportional to the mass of the particle it couples to, would be stronger to taus than to the other leptons because of the high mass of the tau. This means that the following decay chain, with a final state identical to our standard model search final state, could account for a significant fraction of charged Higgs decays:

$$t\bar{t} \to H^{\pm}Wb\bar{b},\ H^{\pm} \to \tau^{\pm}\nu_{\tau} \text{ [16] [17].}$$

By testing the standard model predicted rate of our decay chain we are searching at the same time for any such anomalous process that could show up in the final state as an enhanced (or suppressed) rate for tau leptons in top decay.

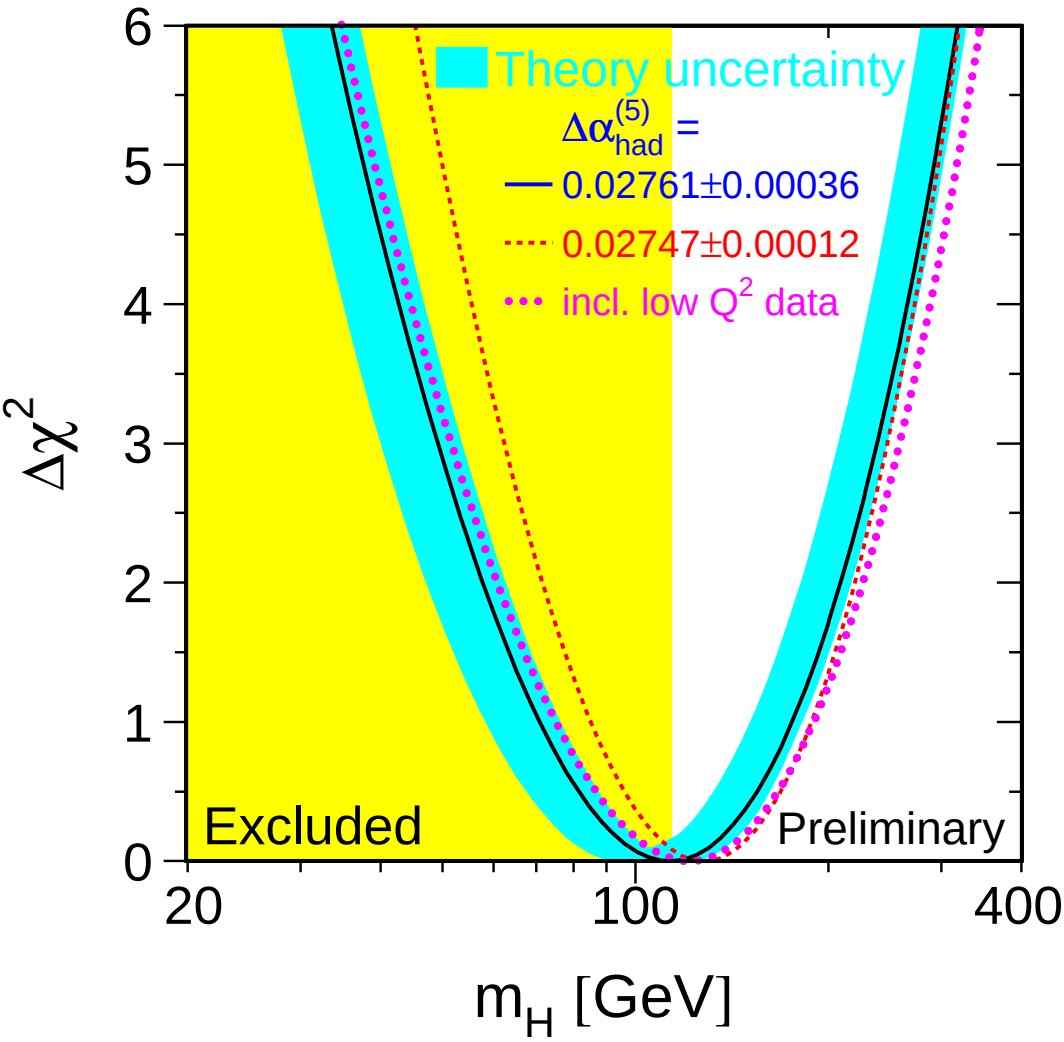

Figure 2.8: Standard Model Constraint on the Higgs Mass from `http://lepewwg.web.cern.ch/LEPEWWG`.

Chapter 3

Collider Detector at Fermilab

The energy necessary for top quark production is only available at one accelerator in the world, the TeVatron, which is located at Fermilab in Batavia, Illinois. Fermilab, originally the National Accelerator Laboratory, was established under a bill signed by President Lyndon B. Johnson in 1967. Both the bottom quark (May-June 1977) and the top quark (1995) were discovered at the lab. Also, the first direct observation of the tau neutrino was made at the lab in July, 2000.

A chain of accelerators is required in order to accelerate particles to high enough energies to produce $t\bar{t}$ events. A very sophisticated detector and data analysis network are required in order to find and analyze the top events that occur only 10 times in every trillion events. This chapter describes the accelerator chain used to create and collide high energy protons and anti-protons. The CDF detector, one of the two detectors along the ring of the TeVatron and the detector used for this analysis, is also described here.

3.1 Accelerator Chain

The Fermilab accelerator chain[18] is shown schematically in Figure 3.1.

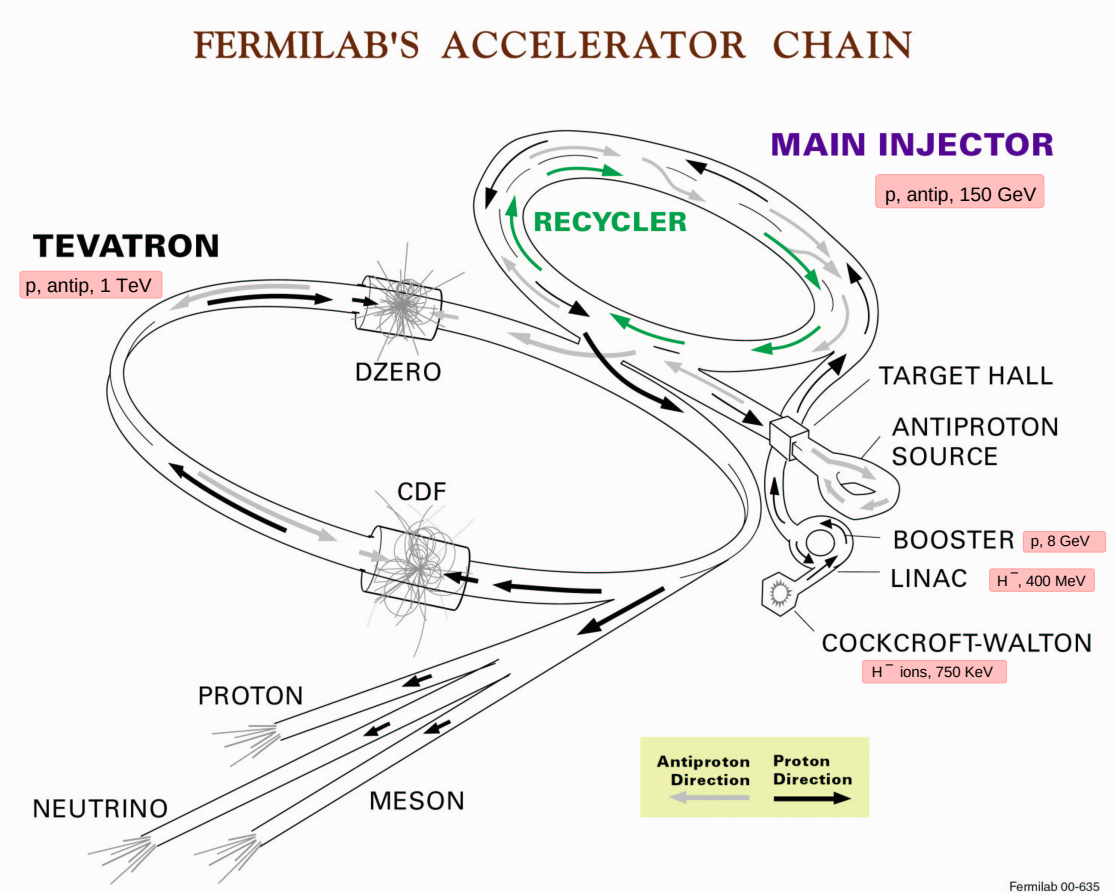

Figure 3.1: The Accelerator Chain at Fermilab.

The acceleration process begins with hydrogen gas. The gas is negatively ionized and accelerated to a kinetic energy of 750 KeV by the Cockcroft-Walton accelerator. The ions then enter a 500 foot linear accelerator that accelerates them to a kinetic energy of 400 MeV. At this point the ions are passed through a carbon foil that strips away the electrons leaving nearly relativistic protons.

The protons then enter the Booster, a circular accelerator that uses magnets to focus and bend the proton beam around the ring and radio frequency electric fields to accelerate the protons. The circular structure of the Booster enables the protons to be accelerated by electric fields many times, traversing the same path over and over again.

The field strength can be increased as the protons get to higher energies on successive rotations. The protons travel about 20,000 laps in the Booster and have been accelerated to 8 GeV when they move on to the next stage.

After the Booster the protons enter the Main Injector, a much larger circular accelerator, where the protons are accelerated from 8 GeV to 150 GeV. The 150 GeV protons are then injected into the TeVatron. The Main Injector also accelerates some of the 8 GeV protons to 120 GeV and sends them to the Anti-Proton Source where they collide with a nickel target to produce a large number of particles. A tiny fraction of the particles resulting from the collision will be anti-protons. The anti-protons are collected and then focused using a lithium lens and sent to the Accumulator ring where the method of stochastic cooling is employed to make the spray of particles into a beam[20]. Once enough anti-protons have accumulated they are sent back to the Main Injector where they are accelerated to 150 GeV and injected into the TeVatron, traveling in the opposite direction of the protons.

The TeVatron is the final accelerator in the chain. It is a circular accelerator with a radius of 1 kilometer. In the TeVatron the beams of protons and anti-protons are both accelerated to 980 GeV. The beams are focused with superconducting magnets and are not filled evenly with protons and anti-protons but are structured in bunches to increase the density of the particles and therefore increase the probability of a collision. These bunches collide in the center of the CDF and D0 detectors.

3.2 CDF

The Collider Detector at Fermilab (CDF) was first commissioned for physics data taking in 1988. Since that time the detector has been upgraded for Run1 (1992 - 1996) and again for Run2 (2002 - 2008). This section describes the upgraded detector used for data taking in Run2. A schematic drawing of CDF[19] is shown in Figure 3.2.

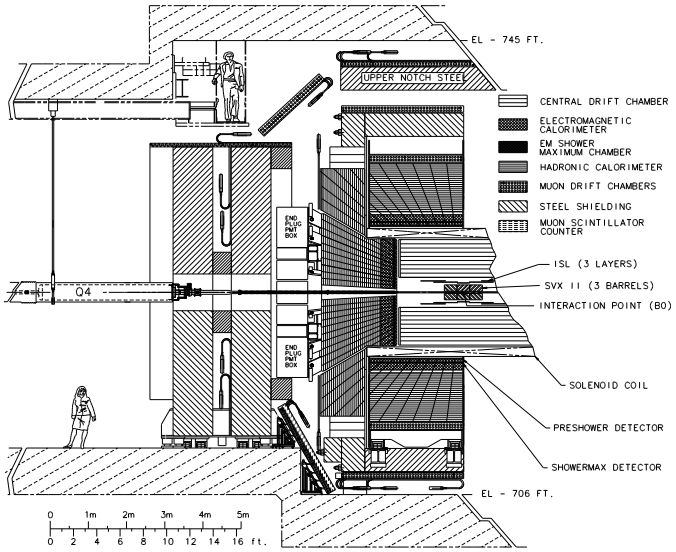

Figure 3.2: CDF Detector Schematic.

CDF is roughly the size of a house and weighs about 5,000 tons. It has over 1,000,000 electronic channels and miles of cable. While CDF is taking data there are a minimum of four people in the detector control room on shift at all times with dozens of experts on the various detector systems on call.

The beamline defines the z axis at CDF, with the detector a cylinder around the beam. Protons travel in the positive z direction (West to East through the detector.) The partons that collide have negligible momentum in x and y compared to their z momentum and we use this in our analyses by enforcing conservation of transverse momentum in each collision. Neutrinos which do not interact in our detector and therefore carry away missing momentum can be inferred in events where the visible transverse momentum balance is badly broken.

The azimuthal angle, ϕ, measures the angle from the plane defined by the TeVatron. The pseudo-rapidity, η, is defined by

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

where θ measures the angle from the positive proton direction, the polar angle. The angle η is useful because the quantity $\Delta\eta$, or $\eta_1 - \eta_2$, is invariant under boosts.

Collisions occur in the center of CDF. The tremendous amount of energy that goes into the collision in the form of the kinetic energy of the colliding partons may be conserved via the formation of high energy particles that can have momentum in the x and y, or transverse, directions. The particles resulting from the collision first travel through the tracking chambers which are inside a 1.4 Tesla solenoid. The particles, or their decay products, then enter the calorimeters. The particles that exit the calorimeter are primarily muons and neutrinos. These particles then pass through the muon chambers where muons are observed.

3.2.1 Luminosity

The Cherenkov Luminosity Counter Detectors (CLC) are located on both the East and West sides of CDF along the beamline, covering $3.75 <| \eta |< 4.75$. These detectors measure the instantaneous luminosity:

$$\mathcal{L} = fn\frac{N_1 N_2}{A}$$

with f the revolution frequency, n the number of bunches in each beam, N_1 and N_2 the numbers of particles in the two bunches that collide, and A the cross-sectional area of the bunches in the beams[21]. This formula assumes that the bunches completely overlap as they collide. Figure 3.3 shows the amount of time-integrated luminosity delivered by the TeVatron and the amount recorded by CDF during Run2 to date.

The CLC is also a critical tool for real time beam monitoring. TeVatron operators use CLC beam information to tune the accelerator and the CDF shift crew has access to real time luminosity information for tuning detector triggers.

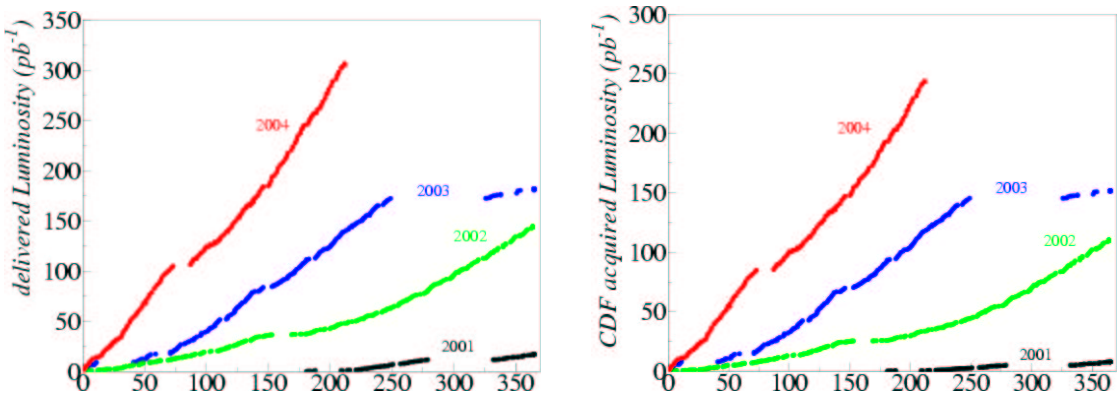

Figure 3.3: Delivered and recorded Luminosity by year.

3.2.2 Tracking

The goal of the Tracking Chambers at CDF is to be transparent, having as little effect as is possible on the particle as it travels through, while at the same time measuring the momentum and charge of charged particles. There are two types of tracking chambers at CDF: silicon detectors and a gas drift chamber. The silicon detectors consist of the Silicon Vertex Detector (SVX), the Intermediate Silicon Layer (ISL), and a layer placed on the beamline known as Layer00. The drift chamber is known as the Central Outer Tracker (COT).

Silicon

The silicon detectors are solid state detectors that are “radiation hard” which means they can withstand large doses of radiation without being harmed. They can be made with high segmentation to give excellent resolution. The are put very close to, or on, the beam pipe. They give high resolution information about the primary vertex of the collision, where the original interaction took place, as well as information about secondary vertices from particles that live long enough to travel a few millimeters in the detector before they decay, such as b mesons. While the silicon detectors are very useful they

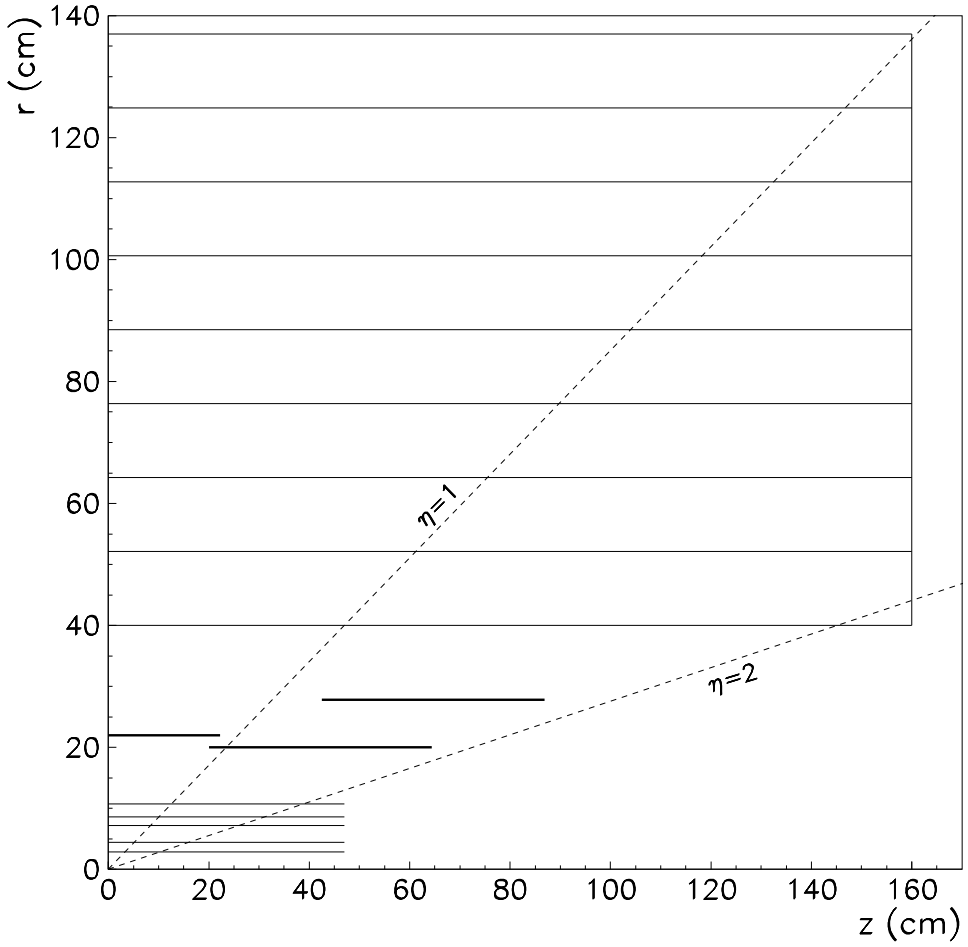

Figure 3.4: A cross-sectional view of a corner of the Silicon Detectors with the COT. The x axis is the z coordinate and the y axis is the radius from the center of the detector. The lower five layers are the SVX and the three layers from 20 cm to 28 cm are the ISL.

are also extremely expensive. They also represent a significant amount of material for particles to interact with. They are therefore used sparingly, occupying only the critical section of CDF tracking that is located close to the beamline. A diagram of the SVX and ISL with the COT is shown in Figure 3.4.

At the center of CDF the beryllium pipe beamline is only 1 cm in diameter. Layer00 is a placed directly on the beamline and is not used in this analysis. The SVX inner radius is at 2.4 cm and it extends to 10.7 cm, covering the region with $\mid \eta \mid < 2.0$. The SVX is 90 cm long and has 5 double-sided layers. A resolution of 24.8 μm was found in $t\bar{t}$ event impact parameter measurements with the SVX[22]. The ISL begins at a radius of 20 cm and its outer layer is at a radius of 28 cm. The ISL improves tracking in the forward region of the detector for tracks that do not pass completely through the COT, which is true for all tracks with $\mid \eta \mid > 1.0$. It also improves tracking efficiency in the central region. The ISL consists of 3 double-sided layers.

COT

The COT extends from a radius of 44 cm to 132 cm. It is a drift chamber filled with 50:35:15 Ar-C_2H_6-CF_4 with a drift velocity of about 100 μm per nanosecond over a maximum drift distance of about 1 cm. The COT is filled with cells, shown in Figure 3.5, containing potential wires, sense wires, shaper wires, gold-mylar field planes and bare mylar shaper panels. The electric field in each cell is approximately uniform at 2.5 kV/cm.[19]

When a charged particle passes through the COT it ionizes the gas creating free electrons. The electric field causes the electrons to drift toward the sense wires. Near the sense wire the 1/r field accelerates the electron to create an “avalanche” of secondary ionization. The time and amplitude of the electrons deposited on the wire are recorded by the electronics. There are eight layers of cells, referred to as “superlayers”, with 12

possible measurements in each layer in the radial direction for a total of 96 possible hits for a particle with high enough $Psin\theta$ (P_T) to travel through the COT. The COT contains a total of 63,000 wires and has 30,240 readout channels.

The magnetic field produced by the solenoid is in a direction parallel to the z axis. A charged particle passing through the magnetic field in the COT will move in the path of a helix. The curvature of the track is used to determine the particle's momentum and the direction of the curvature reveals the sign of the charge of the track. The following is the relationship between the momentum, p, of the particle and the radius of curvature, R:

$$p = \frac{0.3\ z\ B\ R}{cos\lambda} \text{ [19]}$$

with ze being the charge of the particle, B the magnetic field, and λ the pitch angle.

3.2.3 Calorimetry

The CDF Calorimeters are sampling detectors with alternating layers of shielding (lead or iron) and scintillator. Their purpose is to stop charged and neutral particles completely, with the exception of neutrinos and muons that are minimum ionizing and therefore do not interact enough to be stopped, and to measure their energy deposited as they interact in the detectors. Both E and $Esin\theta$ (E_T) are measured in the calorimeters.

The central calorimeters include the Central Electromagnetic calorimeter (CEM) and the Central Hadronic calorimeter (CHA). Both central detectors were assembled in wedges of roughly 15° in ϕ and varying units of η, with an span of 0.11 in η at the center, defined by η = 0. The central calorimeters extend to an $\mid \eta \mid$ of 1.1. Leakage of energy into the cracks between wedges has been extensively modeled with low energy, isolated particles. Both the East and West sides of the detectors have Plug calorimeters that aid in the detection of particles that are high in $\mid \eta \mid$, from about 1.1 to 3.6.

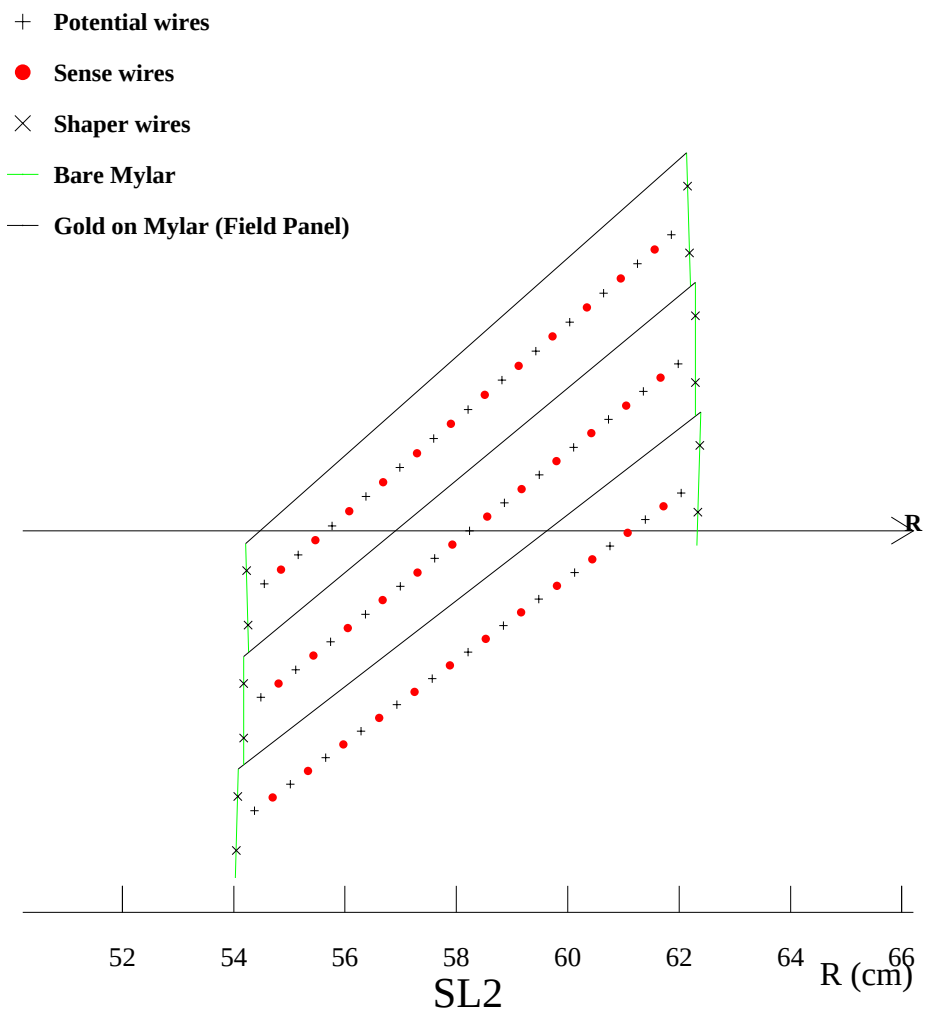

Figure 3.5: A few cells in superlayer 2 of the COT.

CEM

The CEM consists of 23 layers of scintillator alternating with lead. This detector is designed to detect electrons and photons. When electrons travel through the calorimeter they decelerate via bremsstrahlung, the electromagnetic radiation associated with changes in velocity resulting from atomic collisions. The pre-showering detector, the Central PreRadiator (CPR), is located in the first layer of the CEM. The CPR is a wire detector that uses the solenoid and the tracking chamber as a radiator. This detector adds a factor of two to rejection of electron backgrounds when identifying low energy electrons as coming from b quarks (soft lepton b-tagging) and allows for a reduction by a factor of 3 on the systematic uncertainties associated with direct photon measurements. A shower maximum detector, the Central Electromagnetic Shower detector (CES), is located at a depth of 6 radiation lengths, where electron showers are at their peak. The CES contributes to the identification of electrons and photons and allows for the separation of photons that are a result of π^0 decays into two photons.

The entire CEM detector is 19 radiation lengths thick with an E_T resolution of $16\%/E_T$[19]. The CEM is calibrated using the momentum of electrons from W decays.

CHA

The CHA consist of 23 layers of alternating scintillator and iron. It has a thickness of 4.5 nuclear absorption lengths. This detector is designed to give energy information for hadrons.

Plug

A diagram of the plug calorimeter is shown in Figure 3.6. The plug calorimeter's EM section contains 23 layers of alternating scintillator and lead and a shower maximum detector, referred to as the "position detector." This detector is 21 radiation lengths

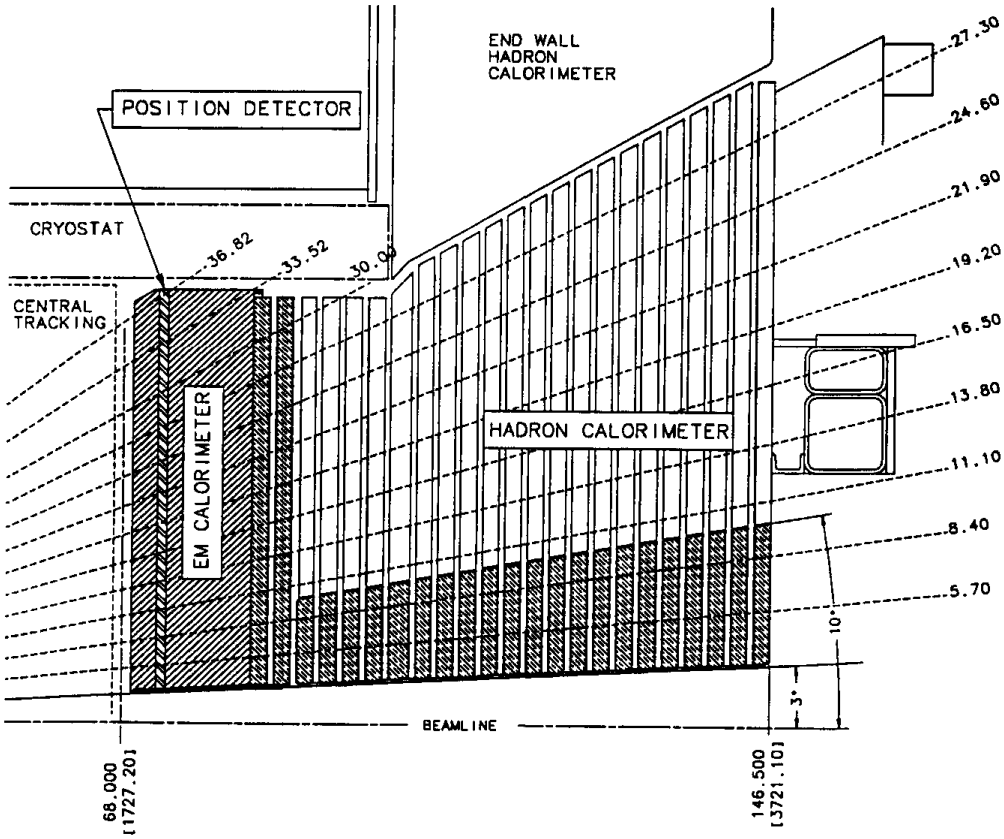

Figure 3.6: End Plug Calorimeter.

thick. The HAD section is 22 layers of alternating scintillator and iron and is 7 nuclear absorption lengths thick..

We do not use the plug calorimeters to detect electrons or taus in this analysis, requiring an $\mid \eta \mid$ maximum of 1.0 for both, but we do use the plug to identify jets, which we accept out to an $\mid \eta \mid$ maximum of 2.0.

3.2.4 Muon Detectors

Muons tend to be minimum ionizing particles (mips), meaning they deposit very little energy in the calorimeters. They therefore can travel through the entire CDF detector without interacting enough to deposit much of their energy in the detector. Because the other common charged particles that we detect do tend to stop and decay inside the detector we can safely assume that the charged particles that exit the detector are muons.

Muon chambers surround much of the CDF detector, as can be seen in the η - ϕ view in Figure 3.7. These detectors consist of scintillators and proportional chambers. We rely on the assumption that if a particle has made it to the muon detector and interacts in the muon chamber it is likely a muon.. There are four types of muon chambers at CDF: Central Muon Chamber (CMU), Central Muon Upgrade (CMP), Central Muon Extension (CMX), and the Intermediate Muon Detector (IMU). The characteristics of each are summarized in Table 3.1 and each system is described in the following sections.

CMU

The CMU detector is located in the outer layer of the central hadronic calorimeter. This proportional drift chamber detector covers the region $\mid \eta \mid < 0.6$ with the exception of a gap at η = 0 of 18 cm. Drift time information in the radial direction is available for rough calculations of the stub P_T for trigger decisions.

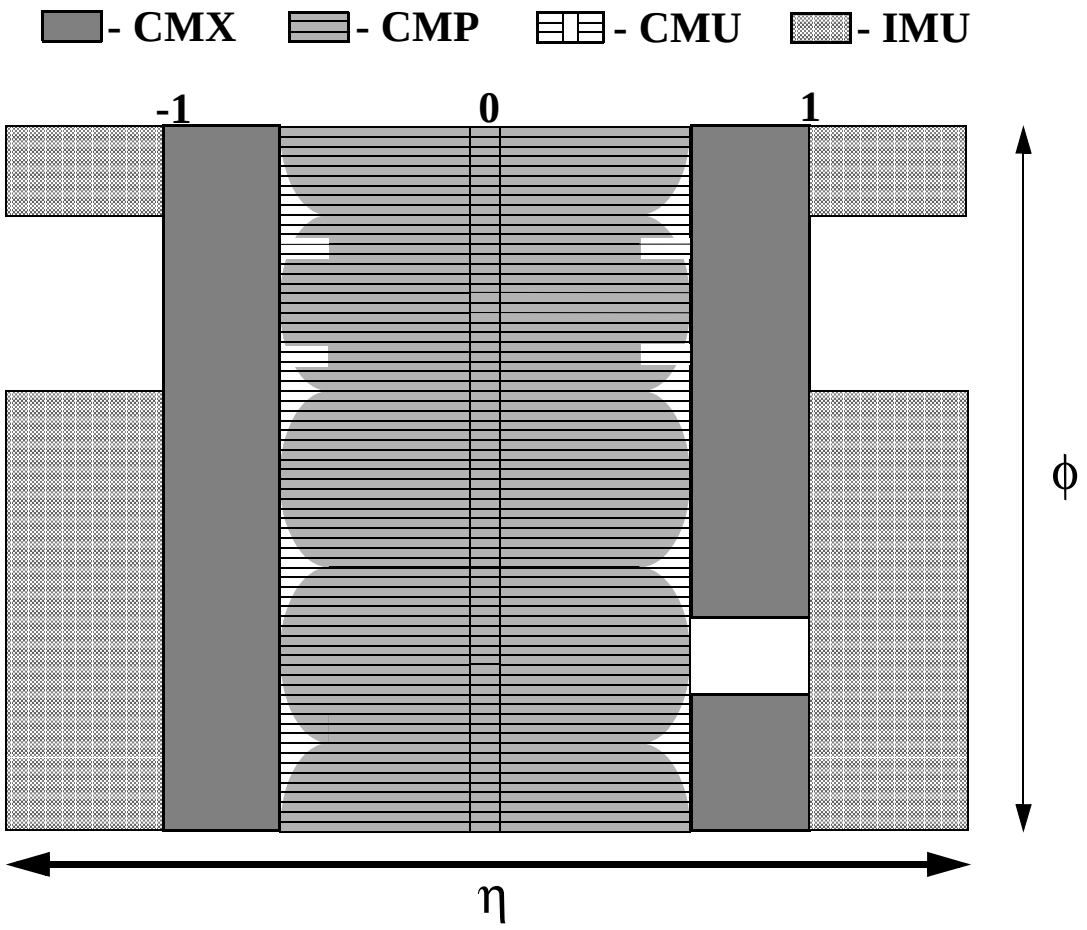

Figure 3.7: Muon Coverage in terms of η and ϕ.

	CMU	CMP/CSP	CMX/CSX	IMU
η coverage	$\mid\eta\mid\leq$0.6	$\mid\eta\mid\leq$0.6	0.6$\leq\mid\eta\mid\leq$1.0	1.0$\leq\mid\eta\mid\leq$1.5
Drift tube cross-section	2.68x6.35 cm	2.5x15 cm	2.5x15 cm	2.5x8.4 cm
Drift tube length	226 cm	640 cm	180 cm	363 cm
Max drift time	800 ns	1.4 μs	1.4 μs	800 ns
Total drift tubes	2304	1076	2208	1728
Scint. counter thickness		2.5 cm	1.5 cm	2.5 cm
Scint. counter width		30 cm	30-40 cm	17 cm
Scint. counter length		320 cm	180 cm	180 cm
Total counters		269	324	864
Pion interaction len.	5.5	7.8	6.2	6.2-20
Min. detectable muon P_T	1.4 GeV/c	2.2 GeV/c	1.4 GeV/c	1.4-2.0GeV/c

Table 3.1: Properties of Muon Detectors[19].

CMP/CSP

The CMP detector is a proportional drift chamber that covers the area defined by $\mid\eta\mid<$ 0.6. This detector is on the outside of the CMU detector and is separated from the CMU by two feet of steel shielding. The CMU and CMP detectors are used together in this analysis to define CMUP muons. Requiring both a CMU and a CMP stub removes the background from hadrons that punch through the hadronic detector, making a muon stub in the CMU, but are stopped in the steel shielding between the CMU and CMP detectors.

The CSP is a layer of scintillation counters installed on the outside of the CMP drift chambers.

CMX/CSX

The CMX detector consists of conical sections of drift tubes that cover a range of 0.6 - 1.0 in $\mid\eta\mid$. Rough P_T information is recorded for use in Level 1 of the trigger. The CSX detector consists of scintillation counters installed on the outside and the inside layers of the CMX detector chambers.

The gap in coverage on the East side is a result of the cryogenics instrumentation for the solenoid.

IMU

The IMU chambers extend the reach of the muon system to higher η, roughly covering $1.0 <| \eta |< 1.5$. These chambers are new with the Run2 upgrade. The IMU chambers were not used in this analysis.

3.2.5 Trigger

In Run2 the CDF detector sees a collision every 396 nanoseconds. There are far too many collisions for us to keep information about every one, and not every collision is even interesting. In order to make decisions about which events to keep we have a three level trigger system. The first level has the highest volume of events to deal with and the least amount of time. This means that fast decisions need to be made with the most basic information. At the second level of the trigger the volume of events has decreased so there is a bit more time for decision making. Some correlations between quantities of interest are available. At the third level of the trigger the amount of time we have per event is limited by our CPU power. The events are almost fully reconstructed in a PC farm and algorithms searching for specific physics processes are run.

An event is stored for complete reconstruction and later analysis if it satisfies all of the requirements of one of the 100 or so defined trigger paths during physics running. Trigger paths consist of a Level1/Level2/Level3 chain of triggers. This means that the system is smart; it requires a specific prerequisite to be passed before an algorithm will be run on an event. As a result of having a smart trigger lots of time is saved by not running an advanced stage of an algorithm on an event that failed a more basic requirement. Monitoring Streams are maintained with a prescaled number of events automati-

cally passed through a trigger level on to the next level regardless of the outcome of the trigger decision. These events are then run through the algorithms offline and checked to make sure that the events did pass a trigger if they should have passed. These types of tests give us confidence that we are not losing events that we should be keeping.

The trigger system is critical. If an event is not selected it is impossible to get it back. The challenge with such a system is that we must know what we are looking for in order to find it and we must understand how we bias our results by the way we choose the events we keep.

The information that is available at each level of the trigger[23] is described in the following sections.

Level1

Level1 has the task of reducing the trigger rate from 2.5 MHz to under 20 kHz.

There are 24x24x2 trigger towers in the calorimeter. These towers cover 0.2 in η and 15° in phi, which does not correspond to the physical segmentation of the towers. The calorimetry available at Level1 includes

- ΣE_T of all trigger towers with a minimum of 1 GeV
- $\not{E}_{\mathrm{T}}$: All trigger towers with a minimum of 1 GeV are included in the calculation
- E_T, P_T, E_{EM}, E_{HAD}

In addition to calorimetry there is information about muons and tracking:

- CMU, CMP, CMUP, and CMX stub information with rough P_T calculations from the CMU and CMX
- limited tracking, in the form of extremely fast tracks, or XFT tracks, which are created with a reduced number of superlayers and simplified tracking reconstruction.

A very limited number of thresholds for requirements on each quantity are allowed to be programmed at Level1. A total of 64 Level1 triggers may be used in each physics run.

Level2

Level2 receives the trigger information from Level1 and must reduce the rate from 20 kHz to 300 Hz. Information from single trigger towers is not available at Level2, but basic clustering algorithms can be performed. Also, the energies are measured with a finer resolution than is available at Level1.

- Tracks can be matched to clusters in the shower max detector.
- Basic isolation in the calorimeter by trigger tower is available.
- Basic track isolation can be required.
- Muon stub matching is done with improved resolution.
- Linking XFT tracks to SVX hits is possible, including information about track impact parameter.

Level3

Level3 must reduce the rate from 300 Hz to a maximum of 75 Hz which is written out to tape and analyzed offline.

The events coming into Level3 are sent to a farm of about 300 CPUs. Each event takes on average 0.5 seconds to 0.8 seconds to analyze. Most of the offline reconstruction code is run in the Level3 trigger in addition to algorithms that are searching for particular physics processes that depend on the trigger path. The triggers used in this analysis are described in Chapter 4.

Chapter 4

Event Selection

We are looking for $t\bar{t}$ events in the $e\tau_{had}$ channel or the $\mu\tau_{had}$ channel with a tau lepton in the decay chain. The "$e\tau_{had}$" channel refers to events in the following decay chain:

$$t\bar{t} \rightarrow WWbb \rightarrow e\tau_{had}\nu_e\nu_\tau bb.$$

The "$\mu\tau_{had}$" channel refers to events in the following decay chain:

$$t\bar{t} \rightarrow WWbb \rightarrow \mu\tau_{had}\nu_\mu\nu_\tau bb,$$

where in each case τ_{had} refers to a τ decay with a tau neutrino and one or more hadrons in the final state. These taus decays are referred to as semi-leptonic.

Top quarks decay to a W boson and a b quark about 100% of the time, as described in Section 2.1.4. Electrons, muons, and taus each account for about 11% of W decays[2]. The two decay chains combined account for 4 out every 81 $t\bar{t}$ events.[2] Table 4.1 shows the possible final states from $t\bar{t}$ events. Taus decay semi-leptonically about 64% of the time. These decays can be grouped into a category with one charged hadron final states and a category with with three charged hadron final states. The branching ratios of taus

to five or more charged particles is less than 1%[2] and we do not concern ourselves with these decays in this analysis.

description	final state	approximate BR
standard dilepton	$ee/e\mu/\mu\mu$ + b jets	4/81
tau dilepton	$e\tau/\mu\tau$ + b jets	4/81
lepton + jets	$e/\mu/\tau$ + jets + b jets	36/81
all-hadronic	jets + b jets	36/81

Table 4.1: Final States of $ttbar$ events with their branching ratios. This analysis focuses on the tau dilepton categories which account for 4 out of every 81 $t\bar{t}$ event.

Our data comes from central electron and central muon triggers which will be described in this chapter. In each case we then require the primary lepton (the electron or muon) to be in the central rapidity region of the detector and pass strict identification requirements in order to purify our sample. We also require the semi-leptonically decaying tau lepton to be in the central detector. In order to account for the presence of b quarks in our events we require two central jets, where a "jet" is a spray of particles, both charged and neutral, that are associated together via a clustering algorithm that groups neighboring energy deposits in the calorimeter. Jets can form as a result of hadronized gluons or quarks. The clustering algorithm is described in detail in Section 4.4. We do not require the jets to be identified as b jets in our event selection. We do, however, check to see if any of our jets were tagged as b jets after we have selected our events. The process of b tagging is described in Section 4.4.1.

The presence of neutrinos in the event allows us to require missing transverse energy because neutrinos pass through our detectors without leaving a trace. We are able to remove a significant amount of the background due to the process $\gamma^*/Z \rightarrow \tau\tau$ by rejecting events with a reconstructed electron (muon) and tau mass that is near the mass of the Z boson. Note that this rejection, referred to as our "Z mass veto cut" is only

applied to events that are reconstructible with our method. The Z mass veto cut is described in detail in Section 4.7. Finally, other $t\bar{t}$ event kinematics are taken advantage of as we make requirements that reduce backgrounds. All cuts are described in detail in the following sections.

4.1 Triggers

This analysis uses data from the three trigger paths, `ELECTRON_CENTRAL_18`, `MUON_CMUP_18`, and `MUON_CMX_18`.

Central electrons will deposit the majority of their energy in the CEM so we make requirements on the maximum allowed value of HAD/EM energy. Electrons are charged particles so they will therefore be tracked by the COT. We expect to find a track that matches to the cluster of energy in the calorimeter with a P_T that is a significant fraction of the E_T deposited. We have the added advantage of knowing what electrons' energy deposit profiles look like in our calorimeters because of test beam electrons that were fired into the calorimeters. We can therefore compare the shape of the candidate electron lateral shower profile, L_{shr}, to the test beam electron L_{shr}.

The `ELECTRON_CENTRAL_18` has the following triggers at each level:

- Level1 `L1_CEM8_PT8` trigger:
  - $E_T > 8$ GeV in a single CEM tower
  - XFT $P_T > 8.34$ GeV/c where the track extrapolates to the CEM tower
  - HAD/EM energy < 0.125
- Level2 `L2_CEM16_PT8`:
  - $E_T > 16$ GeV

  - XFT $P_T > 8.34$ GeV/c
  - matching in ϕ between track and energy cluster
  - HAD/EM energy < 0.125

- `L3_ELECTRON_CENTRAL_18`
  - $E_T > 18$ GeV
  - $P_T > 9.0$ GeV/c
  - Lshr < 0.4
  - Δz between the electron vertex and the primary vertex < 8 cm

The muon CMUP trigger path, `MUON_CMUP18`, has P_T requirements at each trigger level. Most muons deposit only very small amounts of energy in the calorimeter so there are no requirements on the minimum amount of energy. In fact when we identify muons offline, during analysis of the event after it has been stored on tape, we have requirements on the maximum amount of energy allowed to have been deposited in the calorimeter.

- `L1_CMUP6_PT4` :
  - CMU Stub $P_T > 6$ GeV
  - CMU XFT $P_T > 4.09$ GeV/c

- `L2_CMUP6_PT8` :
  - Minimum of 4 XFT layers
  - XFT $P_T > 8.34$ GeV/c

- `L3_MUON_CMUP_18` :

  - $P_T > 18$ GeV/c
  - CMU and CMUP stubs both required
  - Δ x between CMP stub and extrapolated track < 20 cm
  - Δ x between CMU stub and extrapolated track < 10 cm

The muon CMX trigger path, `MUON_CMX18`, has momentum requirements and stub requirements. As with the CMUP trigger path, there are no requirements on the E_T at the trigger level :

- `L1_CMX6_PT8_CSX` :
  - CMX Stub $P_T > 6$ GeV/c
  - CMX XFT $P_T > 8.34$ GeV/c
- The Level2 decision is an automatic accept, meaning there are no requirements. More recent versions of this trigger require 10 GeV P_T at Level2, but this analysis had no Level2 requirement
- `L3_MUON_CMX18` :
  - P_T minimum of 18 GeV/c
  - Δ x between CMx stub and extrapolated track < 10 cm

The low rate of $t\bar{t}$ events combined with the fact that we only include 4 of every 81 $t\bar{t}$ events means that before any analysis cuts we have a challengingly low acceptance. We therefore depend on our triggers having high efficiency. The `ELECTRON_CENTRAL_18` trigger path is 96.6% efficient at triggering on high E_T electrons. This was measured in $Z \rightarrow ee$ data with the denominator all central electrons that pass the E_T, P_T and HAD/EM requirements and the numerator all events that pass the trigger. The `MUON_CMUP_18`

trigger path is 89.0% efficient and the `MUON_CMX_18` trigger path is 96.6% efficient at triggering on high P_T CMUP and CMU muons, respectively. The muon trigger efficiencies were measured in $Z \rightarrow \mu\mu$ data where the denominators include the P_T threshold cuts and the numerators are events that pass the respective triggers.

Our data is now separated into the two categories of electron trigger data and data from the muon triggers. Our next step is to make our offline electron and muon identification cuts on the corresponding datasets.

4.2 Electron Identification

Our electron cuts are summarized in Table 4.2. These cuts are the standard tight cuts for a CDF top quark analysis that selects high E_T electrons with high purity.

In addition to the cuts made by the online trigger, some of which are listed above, tight selection cuts are made on electron candidates. We apply cuts on the E_T and P_T at 20 GeV and 10 GeV/c respectively, reflecting the high energy available in $t\bar{t}$ decays. We require the electrons to be isolated, as defined below, in a region of 0.4 in ΔR, where ΔR is determined using η and ϕ coordinates:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

We take advantage of the isolation of the electrons in top events by placing a cut on the amount of energy allowed to be deposited in the calorimeter near the electron candidate. The total energy deposited in a region of 0.4 in ΔR surrounding the electron candidate is summed. We subtract the contribution of the energy associated with the electron candidate from this sum. The ratio of the energy in the region after the subtraction of the electron energy to the energy of the electron is our isolation quantity. We require this to be less than 0.1 to reflect the tendency of our signal electrons, unlike a track in a generic jet, to be very isolated.

Variable	Cut
E_T	$>$ 20 GeV
P_T	$>$ 10 GeV/c
Isolation Ratio (ΔR 0.4)	$<$ 0.1
E_{had}/E_{em}	$<$ 0.055 + 0.00045*E
E/P	$<$ 2.0 (or $E_T >$ 50 GeV)
L_{shr}	$<$ 0.2
Q * Δx	-3.0 cm, 1.5 cm
$\mid \Delta z \mid$	$<$ 3.0 cm
χ^2_{strip}	$<$ 10
$\mid z_0 \mid$	$<$ 60.0 cm
track quality	$\geq$ 3 stereo SL $\geq$ 7 hits
	$\geq$ 3 axial SL $\geq$ 7 hits
Fiducial	$\mid$extp x$\mid$ $\leq$ 21 cm
	9 cm $\leq$ $\mid$extp z$\mid$ $\leq$ 230 cm

Table 4.2: Electron identification cuts.

Our calorimeters were designed such that electrons deposit most of their energy in the electromagnetic section of the calorimeter. We take advantage of this by imposing a requirement on the maximum allowed ratio of energy deposited in the hadronic calorimeters to energy deposited in the electromagnetic calorimeters. This quantity is allowed to vary with the E_T of the electron in order to not lower our efficiency for high E_T electrons which occasionally punch into the hadronic calorimeter.

There is a requirement on the ratio of E/P (the energy over the momentum) of the electron candidate. In the case of electrons we expect roughly the same amount of energy and momentum. This would not be the case for neutral particles where energy would be detected in the calorimeter but no track would be found so no momentum would be measured. This cut is therefore reducing our background from jets that get a significant fraction of their energy from photons that were either directly produced or are a result of π^0 decays.

A cut on the quantity describing the lateral profile of the electromagnetic shower (L_{shr}) compares the electron candidate's profile to the known profile shape of a standard electron. In order to determine a standard electron shape a test beam of electrons was sent into the detector wedges and the profile was measured before the detector was fully assembled.

The electron candidate's track is extrapolated from the COT to the shower max detector, the CES. The distance in x from the extrapolated track to the nearest hit in the CES is identified as the electron candidate's Δx. A cut on the charge (Q) of the track times Δx ensures that the candidate track is associated with an electron-like object. A similar requirement is made on the Δz between the extrapolated track and the CES hit. Once again test beam electrons come in useful as we can compare the shape of the CES cluster associated with the electron candidate to the known shape of an electron cluster from test beam electrons. The chi squared associated with this comparison, χ^2_{strip}, is required to be less than 10.

The z_0 is the point along the z-axis where the track originates. There is a cut requiring that the track originates close enough to the center of the detector so that we can measure the track well with our centrally positioned tracking chambers. Further restrictions are put on the track to ensure that enough hits in the tracking chamber were recorded. At least three stereo super-layers must have a minimum of seven hits. The same is required of at least three axial super-layers

Finally, our electron candidate must pass a fiducial requirement that ensures that it passed through a responsive and well understood section of the calorimeter.

The electron identification efficiency for these cuts was measured in the data and in the Monte Carlo using $Z \rightarrow ee$ events. The efficiency is plotted as a function of both ϕ and detector η for electrons with a minimum E_T of 20 GeV in Figure 4.1[24].

4.3 Muon Identification

Our muon cuts are summarized in Table 4.3. These cuts are the standard tight cuts for a CDF top quark analysis that selects high P_T muons with high purity. Our muons come in two categories from the two triggers. The CMUP muons and the CMX muons have identical identification cuts with the exception of the cuts associated with muon stubs.

The muon candidate track is required to have a minimum P_T of 20 GeV/c, the same requirement we put on the electron E_T. The isolation requirement is calculated the same way as described above for the electron with the same isolation requirement imposed.

Muons are minimum ionizing particles; They travel through the calorimeters and then the shielding between the calorimeters and muon detectors without depositing much energy. The upper limit of energy allowed deposited in the hadronic calorimeter is 6 GeV for muons with momentum below 100 GeV/c. The upper limit allowed deposited in the electromagnetic calorimeter is 2 GeV for muons with momentum below 100 GeV/c. The muons above 100 GeV/c have maximum allowed deposited energies on

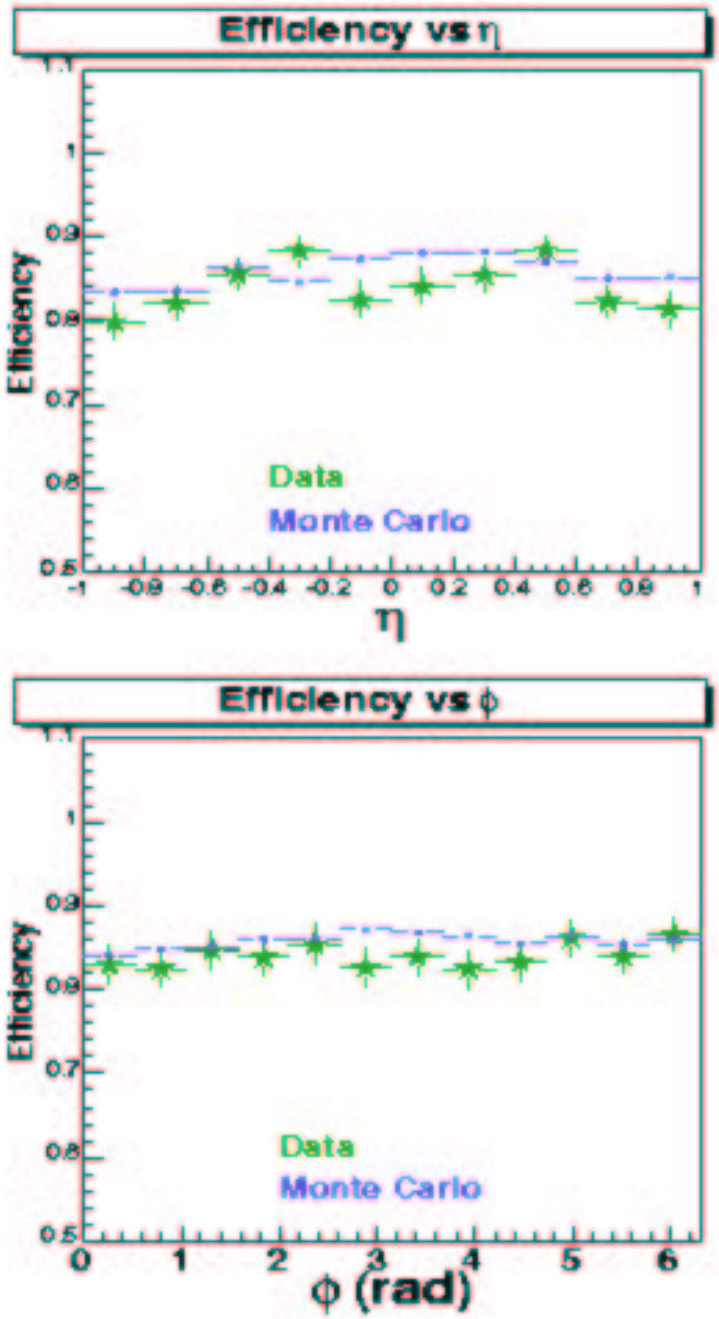
Efficiency vs η
Efficiency
Data
Monte Carlo
η
Efficiency vs ϕ
Efficiency
Data
Monte Carlo
ϕ (rad)

Figure 4.1: [24].

Variable	Cut
Track P_T	$>$ 20 GeV/c
Isolation Ratio (ΔR 0.4)	$<$ 0.1
E_{had}	$<$ 6.0 GeV (and sliding for P above 100 GeV)
E_{em}	$<$ 2.0 GeV (and sliding for P above 100 GeV)
$\mid z_0 \mid$	$<$ 60.0 cm
$\mid d_0 \mid$	$<$ 0.02 cm (with silicon)
	$<$ 0.2 cm (no silicon)
track quality	$\geq$ 3 stereo SL $\geq$ 7 hits
	$\geq$ 3 axial SL $\geq$ 7 hits
muon stub	has CMU and CMP stubs (for CMUP)
	has CMX stub (for CMX)
$\mid \Delta X_{CMU} \mid$	$<$ 3.0 cm (for CMUP muon)
$\mid \Delta X_{CMP} \mid$	$<$ 5.0 cm (for CMUP muon)
$\mid \Delta X_{CMX} \mid$	$<$ 6.0 cm (for CMX muon)

Table 4.3: Muon Identification Cuts.

a sliding scale that is a function of their momentum. This is to account for the fact that the radiative cross section for bremsstrahlung increases as a function of the logarithm of a particle's energy[25].

The standard z_0, axial hits and stereo hits quality cuts are made.

A challenge we have with muons that we did not face with electrons is the presence of a background due to cosmic rays. This means that not every muon we detect in our detector came from a collision at the center of our detector. The frequency of cosmic ray muons in the CDF detector results in a background that must be actively reduced because many surface cosmic rays are high momentum muons. Luckily, we can largely eliminate these events by requiring all muons to pass very close to within the center of our detector. The point of closest approach to the center, known as the impact parameter of the track (d_0) must be less than 0.02 cm in events where we have silicon information and less than 0.2 cm in events where we do not have silicon information.

The remaining cuts on muon candidates relate to stubs being found in the CMU, CMP or CMX detectors. The CMUP muons must have a CMU stub associated with them within 3 cm of the extrapolated track in Δx. These muons must also have a CMP stub associated with them within 5 cm of the extrapolated track in Δx. The CMX muons need only a CMX stub associated with them within 6 cm of the extrapolated track in Δx.

4.4 Jets

Jets, as defined above, are sprays of particles in the detector that result from the hadronization of quarks or gluons. We form jets in our events with a clustering algorithm, JetClu, that groups together neighboring energy deposits. In JetClu the energy in each tower is calculated. The tower with the largest deposit above 3 GeV is considered for the "seed" tower of the jet. Towers within a ΔR of 0.4 are added to the jet cluster. After all towers that qualify have been added to the cluster the center of the cluster is calculated by

weighting the energy contributions of the included towers. The jet is reclustered using a cone of 0.4 around the new jet center. This clustering and re-centering repeats until the center of the jet does not change with further iterations.

Our jet energy requirements were determined in an optimization study described in the Section 4.8. Our leading jet, which must not match to our electron (muon) or tau within a ΔR of 0.4 must have a minimum E_T of 25 GeV. An additional jet with a minimum E_T of 15 GeV must also be found in the event, also not matching to the electron(muon), the tau, or the other jet. Both of the jets must be within an absolute value of η less than 2. Beyond an η of 2 the calorimeter segmentation changes, the energy resolution decreases, and the backgrounds increase substantially.

Our main backgrounds fall off dramatically as we increase the E_T requirement on our jets so we are extremely sensitive to the measurement of jet energies in this analysis. Unfortunately, jet energies are difficult quantities to measure as there are many effects (both physics and detector related) that bias the energy measurement away from the true energy of the quark in the top decay. We therefore correct our jet energies for the following, which could result in incorrect jet energy measurements:

- Relative Jet Corrections: The best understood regions of the CDF calorimeters are within absolute values of η between 0.2 and 0.6. These regions are within the CEM and CHA detectors and they are away from the cracks. The relative corrections correct for known deficiencies in the less understood regions of the calorimeters. They are applied as a function of the size of the clustering cone, which is 0.4 in ΔR for this analysis, and the eta and the P_T of the jet. The result of these corrections is a more uniform distribution of jet energies as a function of η.

- Time Dependence: This correction accounts for fluctuations in the calorimeter photomultiplier tubes that occur as a function of time either as a result of changes

due to radiation damage or due to the running conditions, like changes in the high voltage settings for balancing gains. Calibration studies between data taking allow these fluctuations to be tracked and then later compensated for as time dependent corrections to the jet energies.

- Raw Scale: These corrections come in two parts. The first part is a correction that is a result of a comparison between Run1 and Run2. The second part of the Raw Scale correction is a correction applied to Monte Carlo after a comparison with the data. Photon and jet balancing is used where the photon is measured well, allowing comparisons between the photon and jet P_Ts and energies.

- Multiple Interactions: This correction accounts for the fact that there can be events coinciding with the $t\bar{t}$ events that we try to measure from the same bunch crossing. The level of correction is determined using data taken without requiring triggers to be satisfied, also referred to as minimum bias data.

- Absolute: This corrects for calorimeter non-linearity and energy that is lost in regions of the calorimeter that are not instrumented, otherwise known as the cracks[26].

4.4.1 b tagging jets

When a b quark is created in a collision it hadronizes within about 10^{-23} seconds, forming a b meson. These mesons are not stable and they travel only a few centimeters before they decay. The decay creates what is referred to as a secondary vertex, a point that is the origin for the tracks of two or more charged particles that is a distance away from the primary vertex where the parton-parton collision occurred.

We therefore have an algorithm that we run on jets in candidate top quark events, where we know that actual $t\bar{t}$ events should have two jets that originated from b mesons,

to see if the jets come from a displaced vertex.

$t\bar{t}$ events have a 50% probability of having at least one jet that is identified by our b tagging algorithm as coming from a b quark. Our largest background, W bosons produced in association with jets, has only a 5% chance of containing a b tagged jet. Because the efficiency for finding b tags is low we do not require an event to have a tag in order to make it into our signal sample. We do, however, check the events that we select to see if any of them contain a jet that can be tagged because this is evidence that the event is $t\bar{t}$ - like.

4.5 Tau Reconstruction and Identification

We reconstruct taus at CDF using both calorimeter and tracking information. We first find a calorimeter energy tower with E_T greater than 6 GeV. All adjacent towers with E_T above 1 GeV are added to the tau cluster. The mass of the tau is 1.8 GeV/c^2 yet we require our taus to have a minimum of 15 GeV in E_T so we expect tau decays to be very narrow, or collimated, as they travel through the calorimeter. Because we expect taus to be narrow the tau candidate must have no more than 6 adjacent towers above the adjacent tower minimum threshold E_T.

Once we have the seed cluster we look for a track pointing to the cluster. This track, known as the tau candidate seed track, must have a minimum P_T of 4 GeV. If more than one track is found pointing to the seed cluster the highest P_T track is chosen for the seed track. We form a cone around the seed track that we call the tau cone. The outer edge of the cone is defined by an angle with the seed track that is a function of the energy of the tau:

$$\theta_{cone} = \min(0.17\text{rad}, \frac{5.0\ rad/GeV}{E^{tau\ cluster}})$$

All tracks inside the cone that have P_Ts above 1 GeV and that intercept the z-axis

within 10 cm of the seed track z-axis intercept are associated with the tau candidate.

The tau candidate also has an isolation annulus associated with it. The inner radius of the annulus is θ_{cone} as defined above and the outer radius is 0.52 radians. All tracks within the annulus are identified as isolation tracks. After reconstruction when we make our tau identification cuts we will require the number of isolation tracks to be zero.

We associate π^0's with the tau candidate in a procedure similar to that used to associate tau tracks. The π^0s decay about 98% of the time to two photons[2]. The pairs of photons are reconstructed as π^0s in the CES detector, and the clusters must be within the extrapolated tau cone in order to be associated with the tau. The π^0's positions are reconstructed using the two photon clusters found in the finely segmented CES calorimeter. Their energies are measured with the CEM calorimeter.

Once the tau candidates have been reconstructed we can use the information we have gathered about Identifying semi-leptonically decaying taus at CDF is a tricky business. The decay products which can include both charged and neutral particles, can look like a typical quark or gluon jet in our detector. The P_T and E_T spectra of the tau are softer than for its electron (muon) counterpart from the other W boson because some of the energy and momentum is carried away by the tau neutrino in the decay. This means that we must take advantage of every characteristic the semi-leptonic tau decay has to offer us. Our tau selection is summarized in Table 4.4.

The most accurate measure of the energy/momentum of the tau was determined to come from combining the P_T of the tracks of the tau with the energy deposited by identified π^0s in the calorimeter. As noted above, the tau neutrino carries away some of the energy of the tau means that the energy (momentum) spectrum of the tau is softer than its electron (muon) counterpart. Our track P_T + $\pi^0 E_T$ cut, at 15 GeV, is therefore lower than the corresponding E_T (P_T) cut for electrons (muons.)

The impact parameter (d_0) of the tau must be less than 1 cm.

We know that the mass of the tau is 1.8 GeV/c^2 so we require our tau candidate to

Tau Variable	Cut
π^0 + Track P_T	$>$ 15 GeV
$\mid z_0 \mid$	$<$ 60 cm
$\mid d_0 \mid$	$<$ 1 cm
(Track + π^0) Mass	$<$ 1.8 GeV/c^2
Cal Iso: ΔR=0.4/$E_T^{cluster}$	$<$ 0.06
# tracks in Iso Annulus	0
# π^0's in Iso Annulus	0
# tracks in Tau Cone	$<$ 4
$\mid \Sigma$(track charge) in Tau Cone $\mid$	1
μ veto: E_T/seed track P_T	$>$ 0.5
e veto: E_{HAD}/SUM (P)	$>$ 0.15
seed track quality	$\geq$ 3 stereo SL $\geq$ 7 hits $\geq$ 3 axial SL $\geq$ 7 hits
seed track $\mid$ Z CES $\mid$	$>$ 9 cm $<$ 216 cm

Table 4.4: Tau identification cuts. Note that the tau cone and isolation annulus depend on the energy of the tau candidate.

have a track + π^0 mass less than 1.8 GeV/c^2.

A calorimeter isolation variable is calculated in a similar manner as with electron and muon isolation. In the case of the tau the jet fake rate is much higher than it is for electrons or muons so our isolation requirement is tighter as well, at a maximum of 0.6. Further isolation requirements allow no tracks or π^0s to be found inside the tau isolation annulus formed from the θ_{tau} of the tau cone to 0.52 radians.

The tau cone is required to contain less than four tracks. Also, the sum of the charges of the tracks must have an absolute value of one. The result is that only one track or three tracks can be in the tau cone. This reflects the one-prong and three-prong tau decays.

An explicit veto on muons is formed by requiring the E_T divided by the seed track P_T to be greater than 0.5. Muons, which will only have one track associated with them and will deposit very little energy in the calorimeter, will likely be removed by this cut except in a case where the muon was badly mis-reconstructed. Occasionally a muon will come along that does interact enough in the calorimeter to pass this requirement, which partly accounts for the muon to tau fake rate discussed in a later chapter.

Our electron veto requires the hadronic energy divided by the sum of the momentum of the tracks to be greater than 0.15. Electrons tend to fail this cut because they deposit most of their energy in the electromagnetic section of the calorimeter.

The standard z_0, axial hits, and stereo hits track quality cuts are made. The fiducial cuts on the z location at the CES are also the same as the fiducial requirements on the electron.

4.6 Kinematics

- $\not\!\!E_{\mathrm{T}}$: We have many sources of missing energy in $t\bar{t}$ events. Both of the Ws in our event decay leptonically, so there is a tau neutrino as well as an electron(muon) neutrino in the event. Our tau decays semi-leptonically to hadrons plus a tau neu-

trino. The neutrinos are not detected in the calorimeters so they carry energy away that is not included in the transverse energy balance in the calorimeters. We take advantage of this by requiring a minimum of 20 GeV in missing transverse energy($\not\!\!E_{\mathrm{T}}$.) We use $\not\!\!E_{\mathrm{T}}$ that has been corrected for muons passing tight identification cuts because muons tend to escape from the calorimeters without depositing their energy so they could account for a mis-balance in the measured E_T. We also adjust the $\not\!\!E_{\mathrm{T}}$ for the jet corrections of all jets above E_T of 15 GeV, within an absolute value of η less than 2. The $\not\!\!E_{\mathrm{T}}$ calculation is also sensitive to our determination of the origin of the collision, known as the primary vertex. Our primary vertex must therefore be consistent with our high P_T leptons and the high P_T tracks in the event. Our $\not\!\!E_{\mathrm{T}}$ cut is placed at 20 GeV.

- H_T: The large mass of top quarks results in lots of energy being available for $t\bar{t}$ decay products. This means that the sums of the energies of the final state particles in top quark events tend to be larger than for many background processes that we try to suppress. We therefore cut on the quantity H_T, defined as

$$H_T = E_T^e(P_T^\mu) + P_T^\tau + (\Sigma E_T \text{ of jets within } |\eta| < 2) + \not\!\!E_{\mathrm{T}}$$

  Jets must have a minimum of 15 GeV E_T in order to be included in the sum. The tau P_T also includes energy from the associated π^0s. We require events to have a minimum H_T of 205 GeV. This cut was optimized in a 2-dimensional optimization with the cut on the E_T of the first jet, as described in a later section on optimization.

- Opposite Charge: We take advantage of the conservation of charge in the decay chain by requiring our electron(muon) to have opposite charge from the tau.

4.7 Z Mass Veto

Drell-Yan production of τ pairs, $p\bar{p} \rightarrow \gamma^*/Z \rightarrow \tau\tau$, in association with extra jets can be a significant background to both the $e\tau_{had}$ and $\mu\tau_{had}$ channels. The Z boson can decay to two taus with one of the taus decaying to an electron(muon) and the other tau decaying semi-leptonically. In the Run1 tau dilepton analysis [27] looking for the same final states listed above, the signal to background counting only γ^*/Z background was 2:1. *

Armed with the knowledge that this would be a formidable background we developed a cut specifically to reduce this background without significantly harming our signal acceptance. If we were to try to eliminate a background from $Z \rightarrow ee$ we would be able to reconstruct the mass of the Z boson using a the two electrons and cut around the Z mass window, targeting the background. The same would work for the case of $Z \rightarrow \mu\mu$. For an example of the clean reconstruction of the Z mass where the Z boson has decayed to muons see Figure 7.14. The complication with $Z \rightarrow \tau\tau$ is that the taus have decayed semi-leptonically so some of their energy has been carried away by the tau neutrinos and therefore the Z mass cannot be reconstructed as simply as with the prior cases. We follow the idea demonstrated in other tau analyses of distributing the event $\not\!\!E_{\mathrm{T}}$ to the tau decay products in the event, gaining back the energy lost in the neutrinos [28]. We assign the $\not\!\!E_{\mathrm{T}}$ to the tau decay products based on the ϕ of the $\not\!\!E_{\mathrm{T}}$ by assuming that the τ decays are the only source of $\not\!\!E_{\mathrm{T}}$ in the event.

This cut was tuned using signal and background Monte Carlo samples. We isolated events that are $\gamma^*/Z \rightarrow \tau\tau$ like with an angular requirement. If an event passes the angular requirement we then reconstruct the mass and reject events that fall inside our

*In 109 pb^{-1} of Run I data the expected number of events from this background source was 1.48 ± 0.38, compared with an expectation of 0.7 signal events. An expected number of signal events of 1.1 is quoted, but the standard model predicted top cross-section was not used. Our analysis uses the standard model predicted top cross-section so that is what we use in our comparison.

Z mass window of 65 GeV/c^2 to 115 GeV/c^2. The details are given in Sections 4.7.1 - 4.7.3

4.7.1 Monte Carlo Samples Used

We use exclusively Monte Carlo in this study. Our signal events come from an inclusive $t\bar{t}$ Monte Carlo sample generated by Pythia Monte Carlo with TAUOLA for correct treatment of tau polarization.[29][30] We begin with 386049 events from this sample and we filter the events for an electron or a muon from one W and a semi-leptonically decaying tau from the other W at the generator level.

MC sample	Generated Z Mass	CS * BR (pb)
Sample A	10-75 GeV	50.2±0.1
Sample B	75-105 GeV	23.3±0.1
Sample C	105-800 GeV	0.631±0.001

Table 4.5: Summary of Herwig + Alpgen $\gamma^*/Z \rightarrow \tau\tau$ + 2parton Monte Carlo used

We need three Monte Carlo samples to determine the contribution due to the $\gamma^*/Z \rightarrow \tau\tau$ background because the samples were generated in three mass regions, below, at, and above the Z pole which we refer to as Sample A, Sample B, and Sample C respectively. We use Herwig + Alpgen Monte Carlo created with two partons in order to have a realistic estimate of events with two or more jets. The samples and their (Cross Sections)*(Branching Ratios) are described in Table 4.5. The largest contribution to our $\gamma^*/Z \rightarrow \tau\tau$ background comes from events within the 75-105 GeV range, or Sample B. We focus on this sample for the remainder of the study, unless otherwise noted.

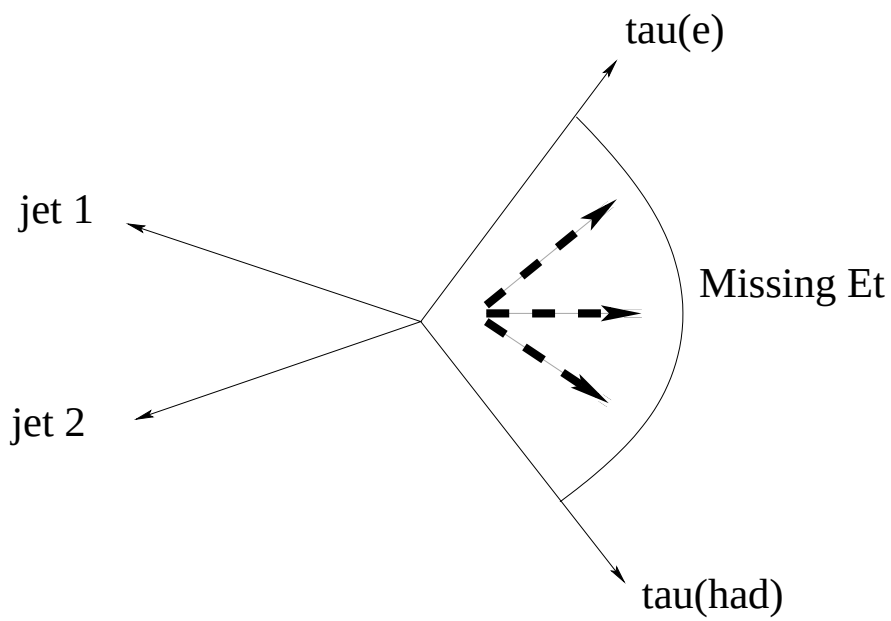

Figure 4.2: An example of an event in which the ditau mass could be reconstructed. The tau with an electron in its final state, tau(e), and the tau that decayed semi-leptonically to a tau neutrino and to hadrons, tau(h), are not back-to-back, and the $\not{E}_{\mathrm{T}}$ falls between the tau(e) and the tau(h).

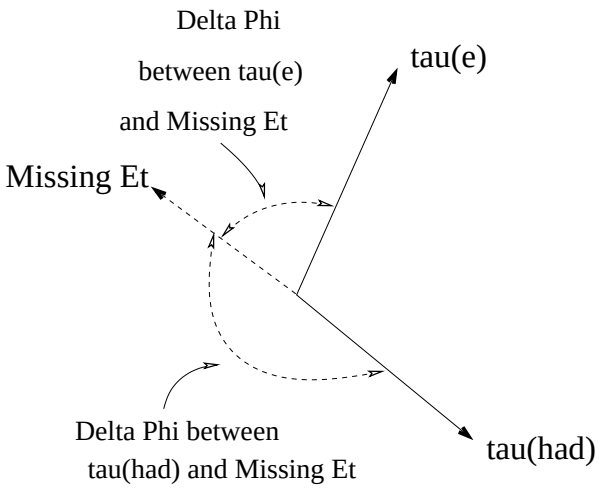

Figure 4.3: An example of an event in which the ditau mass could not be reconstructed. The $\not{E}_{\mathrm{T}}$ does not fall between the two tau decays.

4.7.2 Angular Cuts

As noted in the introduction, the reconstruction of the ditau mass can only proceed if the $\not{E}_{\mathrm{T}}$ in the event can be apportioned among the two taus by comparing the ϕ of the lepton candidates to the ϕ of the $\not{E}_{\mathrm{T}}$. This technique fails if *either* the taus are back-to-back or if the direction of the $\not{E}_{\mathrm{T}}$ does not lie between the two taus as illustrated in Figures 4.2 and 4.3.

If the mass reconstruction fails for one of these two reasons the event is not a candidate for this mass veto cut. For the $t\bar{t}$ signal events a significant fraction of the events are not reconstructible in this way because the correlation between the direction of the $\not{E}_{\mathrm{T}}$ and the lepton directions that is always present in $Z \rightarrow \tau\tau$ is absent when the $\not{E}_{\mathrm{T}}$ comes from the lepton decays of two W bosons.

The angular distributions in the $r - \phi$ plane are shown in Figure 4.4. The difference between the angular distributions between the γ/Z events and the $t\bar{t}$ events can be clearly seen. $\delta\phi_{\ell\tau}$ must not be too close to π or under the $Z \rightarrow \tau\tau$ hypothesis, the two τ candidates are nearly back-to-back. In the below definitions, let "τ" denote the semi-leptonic decaying τ and let ℓ denote either an electron or muon from W decay or from τ decay. $\delta\phi_{\tau\not{E}_{\mathrm{T}}} + \delta\phi_{\ell\not{E}_{\mathrm{T}}}$ must also be very nearly equal to $\delta\phi_{\ell\tau}$. If the sum is greater than the $\ell\tau$ separation, this means that the $\not{E}_{\mathrm{T}}$ is not between the τ and the ℓ, indicating that the $\not{E}_{\mathrm{T}}$ cannot be due solely to $\not{E}_{\mathrm{T}}$ along the directions of the τ and the ℓ, a frequent though not necessary signature of the $t\bar{t}$ events.

We use these distributions to set angular requirements for the mass reconstruction. From these distributions we have chosen,

$$\Delta\phi_{\tau_{\ell},\tau} < (\pi - 0.5)$$

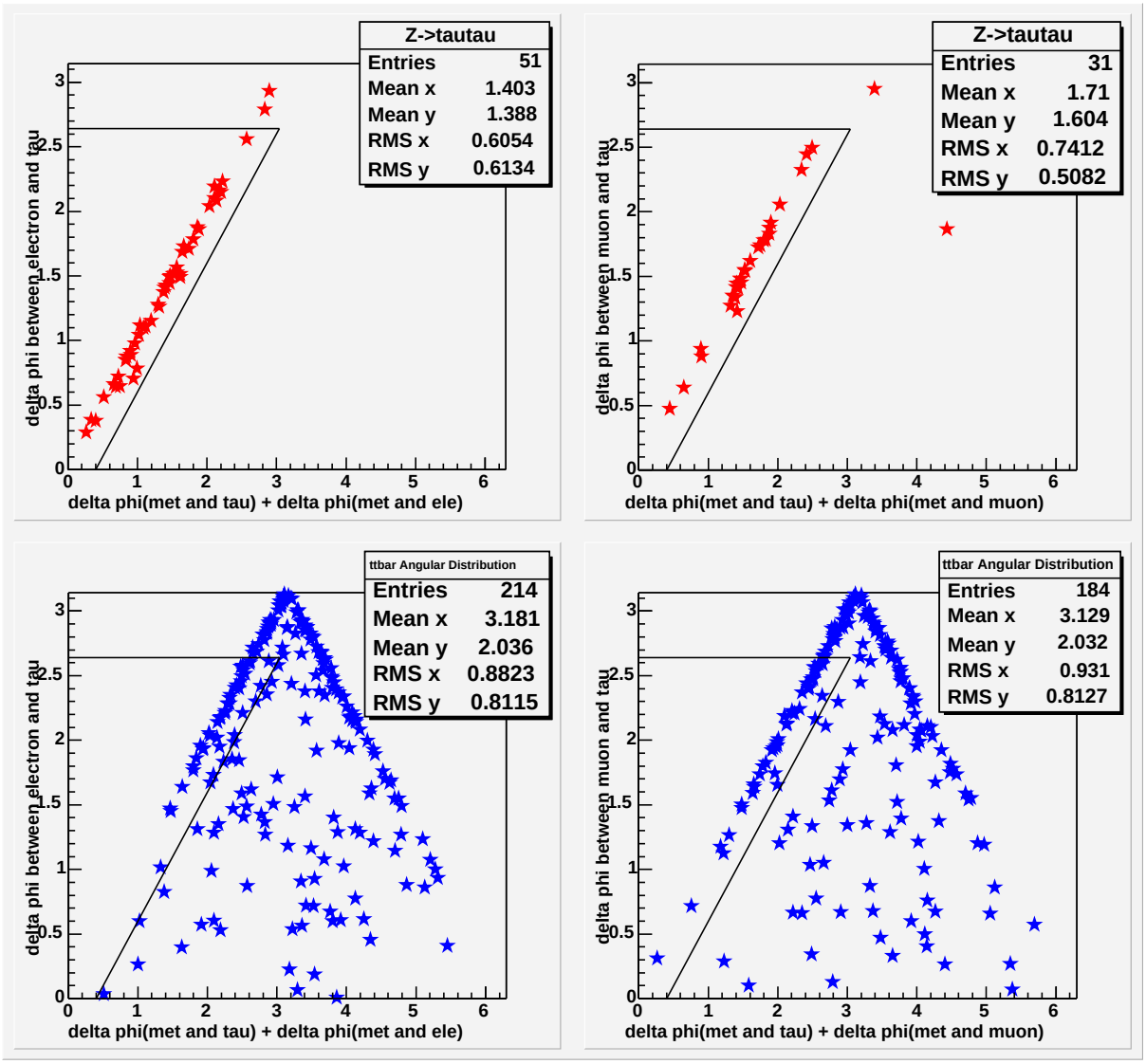

Figure 4.4: $r - phi$ plane angular distributions of e/μ, τ and $\not{E}_{\mathrm{T}}$ from $Z \to \tau\tau$ (Sample B) and $t\bar{t}$ sample. The variable on the x-axis is the sum of $\delta\phi_{\tau\not{E}_{\mathrm{T}}}$ and $\delta\phi_{\ell\not{E}_{\mathrm{T}}}$ which will the equal to $\delta\phi_{\ell\tau}$ when the mass is "reconstructible". **top left:** $\gamma^*/Z \to \tau\tau$ with $\tau \to e$, **top right**: $\gamma^*/Z \to \tau\tau$ with $\tau \to \mu$, **bottom left**: $t\bar{t}$ with $e\tau$, **bottom right**: $t\bar{t}$ with $\mu\tau$.

and

$$(\Delta\phi_{\tau_\ell,\not{E}_{\mathrm{T}}} + \Delta\phi_{\tau,\not{E}_{\mathrm{T}}} - \Delta\phi_{\tau_\ell,\tau}) < 0.4$$

4.7.3 Mass Reconstruction

In the $\tau^+\tau^-$ final state with a decay into $\ell^\pm\nu\bar{\nu}_\ell\nu_\tau\bar{\nu}_\tau j$, only charged lepton and the final state hadrons (denoted j above for the narrow "jet" due to the boosted τ) are visible. However, for the same reason the jet is narrow, two neutrinos will be approximately collinear with the charged lepton and one with the jet. Therefore, neglecting the τ lepton mass, we can define ratio of true to visible energy for the τ,

$$f \equiv \frac{E^{(\tau)}}{E^{(visible)}}.$$

Label the two τ in the event with the index $i \in \{1, 2\}$. Again neglecting terms of order m_τ/E_τ, we can write $\vec{p}_i^{\,\tau} = f_i\vec{p}_i^{\,visible}$. The $\tau^+\tau^-$ invariant mass,

$$m_{\tau\tau} = (E_1^\tau + E_2^\tau)^2 - (\vec{p}_1^{\,\tau} + \vec{p}_2^{\,\tau})^2,$$

can therefore be written as

$$\begin{aligned} m_{\tau\tau} &= (E_1^\tau)^2 - (\vec{p}_1^{\,\tau})^2 + (E_2^\tau)^2 - (\vec{p}_1^{\,\tau})^2 + 2(E_1^\tau E_2^\tau - \vec{p}_1^{\,\tau}\cdot\vec{p}_2^{\,\tau}) \\ &= 2m_\tau^2 + 2f_1f_2(E_1^{visible}E_2^{visible} - \vec{p}_1^{\,visible}\cdot\vec{p}_2^{\,visible}), \end{aligned} \tag{4.1}$$

which only involves visible quantities and the f_i.

The f_i can then be calculated from the observed visible energy, the directions of the visible τ daughter candidates, and the missing transverse energy in the calorimeter. Let the x and y components of the energy and missing energy transverse be indicated by

subscripts. Then

$$f_1 = 1 + \frac{\not{E}_{\mathrm{T}x} E_{2,y}^{visible} - \not{E}_{\mathrm{T}y} E_{2,x}^{visible}}{E_{1,x}^{visible} E_{2,y}^{visible} - E_{2,x}^{visible} E_{1,y}^{visible}},$$

and

$$f_2 = 1 + \frac{\not{E}_{\mathrm{T}x} E_{1,y}^{visible} - \not{E}_{\mathrm{T}y} E_{1,x}^{visible}}{E_{1,x}^{visible} E_{2,y}^{visible} - E_{2,x}^{visible} E_{1,y}^{visible}}.$$

Sample	% passing angular requirement	% outside Z Mass window	% failing angular requirement	% passing Z Mass removal
Sample B (Z pole)	86.6%	2.9%	13.4%	16.3%
Sample C (high mass DY)	77.4%	51.3%	22.6%	73.9%
ttopei ($t\bar{t}$)	15.1%	11.4%	84.9%	96.3%

Table 4.6: Effect of Z mass removal. The second column shows the percentage of events that pass the angular requirements for $\tau\tau$ mass reconstruction and the third column shows the fraction of events with a reconstructed mass outside the Z mass window after all other analysis cuts. The third and fourth column sum to give the total fraction of events surviving the Z mass removal. The low mass Drell-Yan (Sample A) is not shown because no events survive other analysis cuts.

Plots of the reconstructed mass derived from this relationship are shown in Figure 4.5. Note that the peak of the reconstructed mass is somewhat below the Z mass. Because of this, we adopt a wide cut, removing events with a reconstructed mass centered on the Z mass in a window between 65 and 115 GeV. Note also that the majority of $t\bar{t}$ events have a reconstructed $\tau\tau$ mass well above the Z pole. The efficiency of the cut for the Drell-Yan and $t\bar{t}$ samples is shown in Table 4.6. A significant fraction of the $\gamma^*/Z \rightarrow \tau\tau$ background is eliminated with this cut (83.7% of Sample B and 26.1and the majority of the background from this process comes from events near the Z pole, in Sample B) while we lose only 4.7% of our signal acceptance.

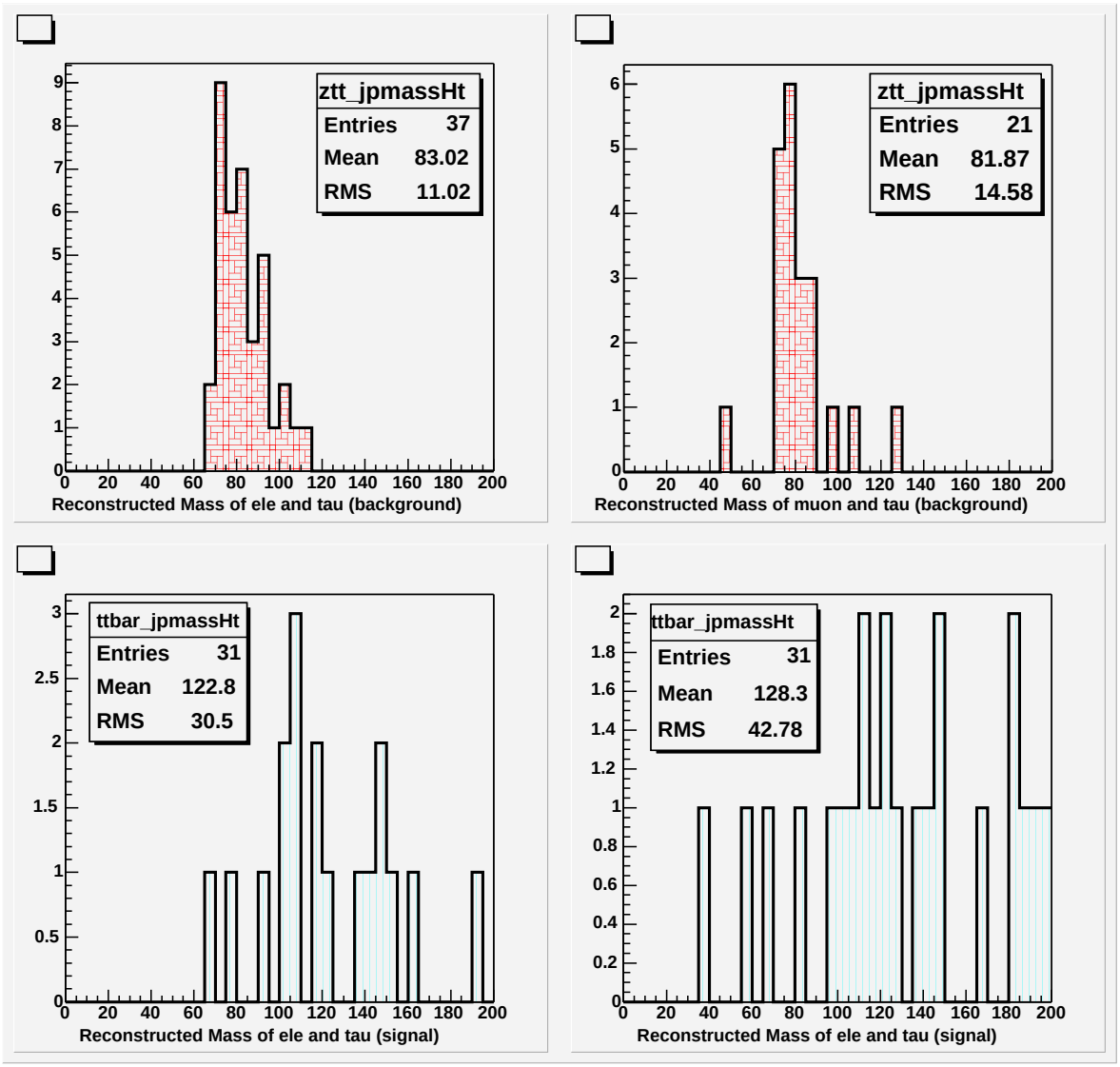

Figure 4.5: Reconstructed mass for events passing angular cuts. Note that because most of the $t\bar{t}$ events do not meet the angular requirements and are therefore not candidates for the mass reconstruction the number of $t\bar{t}$ events has decreased dramatically from what was shown in Figure 4.4. **top left:** $\gamma^*/Z \rightarrow \tau\tau$ with $\tau \rightarrow e$, **top right**: $\gamma^*/Z \rightarrow \tau\tau$ with $\tau \rightarrow \mu$, **bottom left**: $t\bar{t}$ with $e\tau$, **bottom right**: $t\bar{t}$ with $\mu\tau$.

4.8 Optimization of H_T and Jet E_T

We chose our cuts on H_T and the E_T of the leading jet ($E_T^{(1)}$) with the results of an optimization study. Both of these variables are good discriminators between the signal and the background. Figure 4.6 shows the percentage of remaining events as a function of the cuts on H_T and $E_T^{(1)}$. The signal survives to dramatically higher values of H_T and significantly higher values of $E_T^{(1)}$ than the background. The signal is taken from our Pythia $t\bar{t}$ Monte Carlo used in the acceptance calculation in Chapter 5. The background is taken from both data and Monte Carlo sources, as is done in our calculation of background levels in Chapter 7.

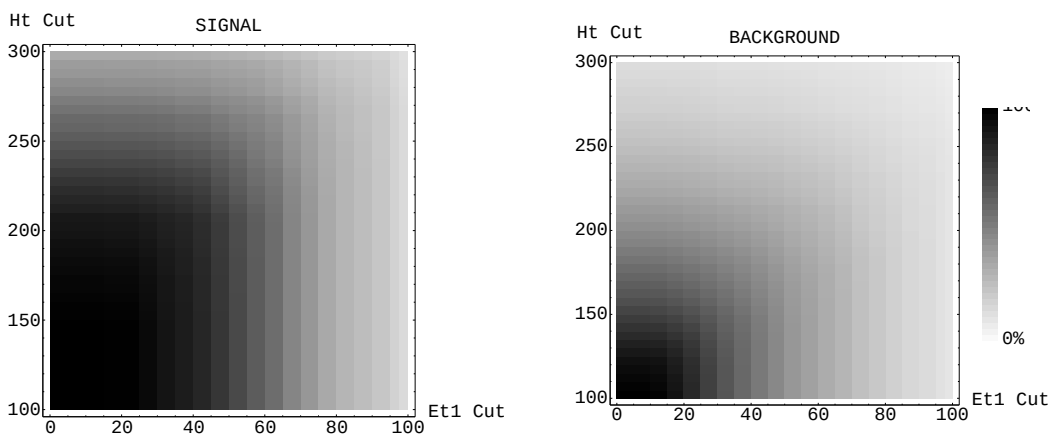

Figure 4.6: Integral event fractions surviving cuts in a space of H_T cut vs. $E_T^{(1)}$ cut. As shown by the key, darker regions have a higher event fraction surviving the cuts.

We are unable to include more variables in our optimization due to our limited statistics in both the Monte Carlo and the data. With increased statistics one could include more than two variables in an optimization study, perhaps adding the E_T of the second jet, event missing E_T and even thresholds of various tau cuts.

In order to quantify our optimization we look at two variables. The standard signal

over the square root of background optimization was performed:

$$\mathcal{O}_{\mathcal{X}}^{(1)} \equiv \frac{S_{\mathcal{X}}}{\sqrt{B_{\mathcal{X}}}},$$

with the number of signal events for a given set of cuts, $\mathcal{X}$, denoted as $S_{\mathcal{X}}$, and the background similarly denoted as $B_{\mathcal{X}}$.

This variable, $\mathcal{O}^{(1)}$, is correct when there is no observed significant signal in the limit of Gaussian statistics. A more correct method in our situation, where we have very low numbers of events expected and therefore need Poisson statistics as opposed to Gaussian statistics, is the likelihood ratio variable:

$$\mathcal{O}_{\mathcal{X}}^{(2)} \equiv \sum_{N} \frac{P(S_{\mathcal{X}} + B_{\mathcal{X}}, N)}{P(B_{\mathcal{X}}, N)},$$

where $P(\mu, N)$ denotes the Poisson probability of observing N events with an expected mean of μ events. This variable sums over all numbers of possible observed events, and therefore correctly takes into account the probability of observing any arbitrary number of events in the final analysis.

The optimization variables $\mathcal{O}^{(1)}$ and $\mathcal{O}^{(2)}$ are shown as a function of H_T and $E_T^{(1)}$ cuts in Figure 4.7. The light-colored regions of each plot correspond to the most optimal regions to cut. There are well-defined, if very broad, maxima in both optimization variables in roughly the same region (in the "eyes and mouth" of Figure 4.7).

We choose the region in this broad optimum which has the *highest* signal and background acceptance. In this case, this is the region in the left-most corner of the "mouth" of Figure 4.7. For both optimization variables, the selected cut by this procedure is $H_T > 205$ GeV and $E_T^{(1)} > 25$ GeV, which we therefore use in our analysis.

We note that a decision to make no H_T cut in this analysis would increase background by a factor of 2.3 and gain only 10% in signal acceptance.

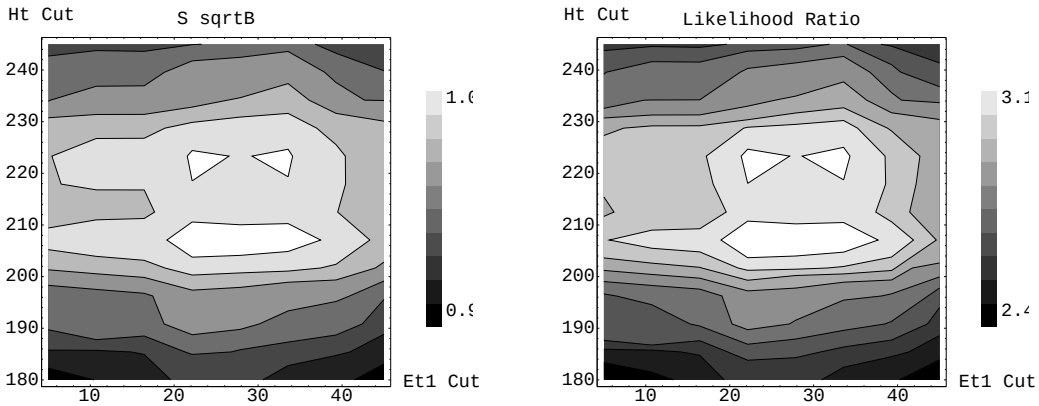

Figure 4.7: Optimization variables **(left)** $\mathcal{O}^{(1)}$ $(S/\sqrt{B})$ and **(right)** $\mathcal{O}^{(2)}$ (a likelihood ratio) shown as a function of the H_T and $E_T^{(1)}$ cuts. Happy Halloween!

Chapter 5

Acceptance

We use inclusive $t\bar{t}$ Monte Carlo, or Monte Carlo that includes all standard model final states with appropriate branching ratios, to calculate the fraction of $t\bar{t}$ events our analysis cuts select, which is known as our signal acceptance. Our sample was generated with PYTHIA, using TAUOLA to correctly treat the polarization of taus[29][30].

Our signal acceptance, $\epsilon_{t\bar{t}_{e/\mu,\tau}}$, can be written as:

$$\epsilon_{t\bar{t}_{e/\mu,\tau}} = \epsilon_{geom-pt}\ \epsilon_{ID}^{e/\mu}\ \epsilon_{reco}^{e/\mu}\ \epsilon_{ID}^{\tau}\ \epsilon_{vertex}\ \epsilon_{\not{E}_T}\ \epsilon_{H_T}\ \epsilon_{2jet}\ \epsilon_{OS}\ \epsilon_{Zveto}\ \epsilon_{trigger}$$

with the following definitions for the pieces of the acceptance:

- $\epsilon_{geom-pt}$: refers to the geometric and P_T acceptance. In order to ensure that we are measuring quantities well and to reduce backgrounds we require our physics objects (electrons, muons, taus and jets) to be central. This means that we lose events with particles at small angles to the beamline. The geometric acceptance accounts for these losses. The P_T threshold cut is also included in this efficiency.

- $\epsilon_{ID}^{e/\mu}$: the efficiency associated with the identification cuts we use for electrons and muons.

- $\epsilon_{reco}^{e/\mu}$: the efficiency associated with electron and muon reconstruction. The value for muons was measured in the data and the value for electrons was taken from the Monte Carlo and verified with the data.

- ϵ_{ID}^{τ}: the efficiency associated with the identification of taus is taken initially from the Monte Carlo. However, the number is corrected by the result of a study using $W \rightarrow \tau\nu$ events. In this sample we measure a "scale factor", the ratio of $\epsilon_{ID}^{W\rightarrow\tau\nu}$ in data to that in Monte Carlo. This study is described in detail in Chapter 6.

- ϵ_{vertex}: the efficiency associated with our cut on the primary vertex, measured with the data.

- $\epsilon_{\not{E}_{\mathrm{T}}}$: the efficiency of our $\not{E}_{\mathrm{T}}$ cut.

- ϵ_{H_T}: the efficiency of our H_T cut.

- ϵ_{2jet}: the combined efficiency of our cuts on the two jets.

- ϵ_{OS}: the efficiency of our opposite sign cut where we require the tau to have the opposite charge of the electron (muon.)

- ϵ_{Zveto}: the efficiency of our Z mass veto cut.

- $\epsilon_{trigger}$: takes into account the trigger efficiencies. Our three triggers (CEM, CMUP, CMX) all have an efficiency associated with them that was measured from the data. The three values are weighted as determined by the fraction of the signal from each of the three categories.

$\epsilon_{geom-pt}$, ϵ_{reco}^{e}, $\epsilon_{\not{E}_{\mathrm{T}}}$, ϵ_{H_T}, ϵ_{2jet}, ϵ_{OS}, ϵ_{Zveto} are calculated with Monte Carlo only. The other contributions are shown in Table 5.1

When using the acceptance to calculate our expected number of events we take the $p\bar{p} \rightarrow t\bar{t}$ cross section to be 6.7 pb from the NLO calculations[31].

	scale factor
$\epsilon_{trigger}$: CEM	0.966±0.001
$\epsilon_{trigger}$: CMUP	0.890±0.009
$\epsilon_{trigger}$: CMX	0.966±0.007
ϵ^{e}_{ID}: CEM	0.965±0.006
ϵ^{μ}_{ID}: CMUP	0.94±0.01
ϵ^{μ}_{ID}: CMX	1.015±0.007
ϵ^{μ}_{reco}: CMUP	0.927±0.010
ϵ^{μ}_{reco}: CMX	0.992±0.011
ϵ^{τ}_{ID}	0.95±0.06
ϵ_{vertex}	0.948±0.003

Table 5.1: Scale factors that must be applied to Monte Carlo acceptance (and Monte Carlo backgrounds.)

5.1 Acceptance from signal Monte Carlo

We begin with 389,067 inclusive $t\bar{t}$ events and apply all of our event selection cuts. Our first step is to apply our z vertex requirement to the event primary vertex. The scale factor of 0.948 $\pm$ 0.003 in table 5.1 is applied after all analysis cuts to account for the difference between data and Monte Carlo.

The effect of the remainder of our analysis cuts is shown in Table 5.2. The cuts are defined as follows:

- N(sig): Our events are separated into four categories using exact information about the simulated interaction. We require the presence of a semi-leptonic tau decay from one of the W's in each event. The $e\tau_h$ category requires the other W in the event to decay to an electron. The $\mu\tau_h$ category requires the other W to decay to a muon. We also have two categories where both Ws decayed to taus. Since one of those taus decayed semi-leptonically the other tau is required to decay into an electron or a muon. The number of events in each category is referred to as N(sig). We begin with 5545 $e\tau_h$ events, 5689 $\mu\tau_h$ events, 1000 $\tau_e\tau_h$ events

and 1031 $\tau_\mu\tau_h$ events in the simulation. This corresponds to about 400 times our number of events in data.

- N($e\mu$geoP_T): Requires a reconstructed central electron (muon) over 20 GeV (20 GeV/c) that passed the fiducial cuts. The large drop in number of events for the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ categories is to a large extent a result of the softer E_T and P_T spectra when the electron or muon is not the W daughter but comes from $W \rightarrow \tau\nu$, $\tau \rightarrow \ell\nu\nu$.

- N($e\mu$ ID): Requires the electron or muon to pass our standard, tight identification cuts described in the previous chapter with the exception of the isolation cuts that are treated separately below.

- N($e\mu$ iso): Requires the electron or muon to pass the isolation cut.

- N(τ cand): Requires a reconstructed tau candidate which does not match, within a ΔR of 0.4, to the electron (muon).

- N(τ P_T,ID,iso): Requires the tau candidate to pass all identification cuts. Note that there is a very large drop in the number of events in each category after the tau identification requirements are made. This is because many of our tau candidates in the Monte Carlo came from the b jets in the event, not from the $t \rightarrow bW$, $W \rightarrow \tau\nu$. A much higher percentage of jets is removed by the tau identification cuts than what we expect from real taus. In order to more clearly see the effect of the identification cuts on real taus we have an alternate table, Table 5.3 that is described later in this section. We also look at the effect of each of the tau identification cuts separately later in this chapter.

- N(opp sign): Requires the identified tau to have the opposite charge of the electron (muon.) This cut is highly efficient for our signal events as the two charges are

absolutely always opposite. This cut will be effective in terms of reducing backgrounds where there is less correlation between the charges of the two objects which is often the case when the "tau" is actually a jet that faked a tau.

- N(1jet): Requires a minimum of one extra jet in the event that does not match the electron (muon) or the tau in ΔR of 0.4, and has $|\eta| < 2.0$ and $E_T > 25$ GeV
- N(2jet): Requires a second jet that does not match the electron (muon) or the tau in ΔR of 0.4 and has $|\eta| < 2.0$ and $E_T > 15$ GeV.
- N($\not{E}_{\mathrm{T}}$): Requires event $\not{E}_{\mathrm{T}}$ to be greater than 20 GeV.
- N($\mathrm{H_T}$): Requires event H_T to be above 205 GeV
- N(Z mass veto): Requires the event to survive our mass reconstruction veto cut.

If we reorder our cuts and require our tau candidate to be matched to the generated semi-leptonically decayed tau from $t \rightarrow bW$, $W \rightarrow \tau\nu$ we will be able to show the effect of some of the cuts more clearly since we will not suffer from the confusion of having jets fake our taus in the acceptance table. The results can be seen in Table 5.3. When we do this we also add a new category:

- N(τ geoP_T): Requires the tau candidate, that matches in ΔR to the generated tau, to pass the central fiducial and P_T cuts.

We also have separated the tau isolation cuts from the rest of the tau identification requirements. The events that pass all cuts are the same in both acceptance tables, in spite of our requirement that the tau candidate match to the generated tau in Table 5.3. This confirms for us that we are not including jets faking taus in our calculation of acceptance. It can be determined from Table 5.3 that our tau identification efficiency, with the isolation included, is about 35%.

	# of events			
Cut	eτ_h	$\mu\tau_h$	$\tau_e\tau_h$	$\tau_\mu\tau_h$
N(sig)	5545	5689	1000	1031
N($e\mu$ geoP_T)	3373	2331	371	232
N($e\mu$ ID)	2198	1983	178	195
N($e\mu$ iso)	2034	1828	156	172
N(τ cand)	1525	1419	117	134
N(τ P_T,ID,iso)	284	246	25	24
N(opp sign)	282	245	25	24
N(1jet)	274	240	22	24
N(2jet)	236	202	18	19
N($\not{E}_{\mathrm{T}}$)	219	185	18	18
N($\mathrm{H_T}$)	200	171	17	15
N(Z mass veto)	193	164	16	14

Table 5.2: Number of events in Pythia signal Monte Carlo passing each stage of the analysis cuts out of 386,037 events from this dataset.

	# of events			
Cut	eτ_h	$\mu\tau_h$	$\tau_e\tau_h$	$\tau_\mu\tau_h$
N(sig)	5545	5689	1000	1031
N($e\mu$ geoP_T)	3373	2331	371	232
N(τ geoP_T)	1323	957	160	84
N($e\mu$ ID)	841	809	71	66
N(τ ID)	469	463	39	40
N($e\mu$ iso)	440	425	35	34
N(τ iso)	284	246	25	24
N(opp sign)	282	245	25	24
N(1jet)	274	240	22	24
N(2jet)	236	202	18	19
N($\not{E}_{\mathrm{T}}$)	219	185	18	18
N(H_T)	200	171	17	15
N(Z mass veto)	193	164	16	14

Table 5.3: Number of events in Pythia signal Monte Carlo dataset passing a reordered version of the analysis cuts. Here, the N(τ geoP_T) cut requires the reconstructed tau candidate to match the generated tau.

We measure the effect of each of the tau identification cuts by first requiring an electron or muon to pass its tight identification cuts and also requiring the presence of a tau in the event. We make all tau identification cuts except for the one we study and we see the effect on the remaining tau candidates in Table 5.4.

Our $t\bar{t}$ efficiency can be re-written as

$$\epsilon_{t\bar{t}}^{MC} = (\text{Total } N_{Zmass\ veto} \text{ / Total Sample Size}) * \text{scale factors},$$

where the scale factors, listed in Table 5.1 are weighted over the different efficiencies associated with the finding of the electron (CEM) or muon (CMUP or CMX) in the event. We therefore calculate an acceptance of

$$\epsilon_{t\bar{t}} = 0.00080 \pm 0.00005(\text{stat}) \pm 0.00014\ (\text{sys})$$

The systematic errors are described in the Section 5.2.

5.2 Systematics

We have a number of systematic errors that are associated with the acceptance calculation. We summarize the systematic uncertainty contributions in Table 5.8 and explain how each was computed in this section. The systematics due to initial state radiation (ISR), final state radiation (FSR), parton distribution functions (PDFs), and Monte Carlo generator dependence were calculated by Dr. Anthony Vaiciulis, a postdoctoral researcher with the University of Rochester, following procedures set by the CDF Top Physics group.

- Jet Energy Corrections

  Our jet energy corrections have uncertainties associated with them that contribute

Cut	Percentage of remaining τs
π^0 + Track $P_T > 15$ GeV	90.7%
$\|z_0\| < 60$ cm	99.2%
$\|D0\| < 1$ cm	100%
(Track + π^0) Mass < 1.8 GeV/c^2	95.6%
Cal Iso: ΔR=0.4/$E_T^{cluster} < 0.06$	77.9%
No tracks in Iso Annulus	84.1%
No π^0's in Iso Annulus	97.5%
# tracks in Tau Cone $< 4 = 1$	99.5%
$\|\Sigma$(track charge) in Tau Cone$\|$	92.0%
μ veto: E_T/seed track $P_T > 0.5$	99.7%
e veto: E_{HAD}/SUM (P) > 0.15	71.3%
seed track quality	98.8%
seed track $\|$Z CES$\|$ (fiducial)	90.8%

Table 5.4: Effect of tau cuts on inclusive $t\bar{t}$ Pythia Monte Carlo after a tight lepton (electron or muon) has been selected. An event must have a generated tau to make it to this point of the analysis but we have not required the tau to be a semi-leptonically decaying tau. Electrons, unlike muons, are very efficiently reconstructed as taus in part because the energy they typically deposit in the calorimeter easily passes the threshold for tau reconstruction. Because of the efficient reconstruction of electrons as taus and because of the presence of electrons from tau decays in this table, we are not surprised by the reduction in the electron veto cut.

Category	Pythia	Herwig
$e\tau_h + \tau_e\tau_h$	0.76 ± 0.08	0.67 ± 0.06
$\mu\tau_h + \tau_\mu\tau_h$	0.53 ± 0.07	0.61 ± 0.06
total	1.29 ± 0.10	1.28 ± 0.08

Table 5.5: Number of expected events in 193.5 pb^{-1} for Pythia with QED FSR turned off and Herwig re-weighted for differing $W \to \tau\nu$ branching ratios.

Category	Pythia no ISR	Pythia with default ISR
$e\tau_h + \tau_e\tau_h$	0.55 ± 0.06	0.62 ± 0.05
$\mu\tau_h + \tau_\mu\tau_h$	0.38 ± 0.06	0.46 ± 0.04
total	0.93 ± 0.08	1.08 ± 0.06

Table 5.6: Number of expected events in 193.5 pb^{-1} for a Pythia sample with no QED/QCD ISR and a Pythia sample with ISR turned on.

to the systematic error. We estimate this systematic error by comparing the acceptance after shifting our jet energy corrections by $\pm 1\sigma$ of their combined uncertainty and taking half the difference as our error[32]. This procedure results in a 5.8% systematic error assigned to jet corrections.

- MC generator Dependence

  We determine our systematic uncertainty due to choice of Monte Carlo generator by comparing the difference between our acceptance calculated with a pythia sample and our acceptance calculated with a herwig sample. In order to make a fair comparison we account for the fact that Herwig does not include QED FSR by using a pythia sample that has QED FSR turned off. We also re-weight the Herwig results by a factor of $\frac{(0.108)^2}{(0.111)^2}$ to account for the fact that Herwig uses the theoretical value of 0.111 for the branching ratio of $W \to \tau\nu$ while Pythia uses the measured value of 0.108. After taking this factors into account we find that the difference in acceptance from the two generators is smaller than our 7% statistical uncertainty on the comparison. The result is shown in Table 5.5 in terms of expected number of events in 193.5 pb^{-1}. We take 7% as the systematic uncertainty.

- Initial State Radiation (ISR)

  We calculate half the difference in the acceptance between a $t\bar{t}$ Pythia Monte Carlo sample with ISR turned on and a Pythia $t\bar{t}$ Monte Carlo sample with ISR turned

Category	Pythia less FSR	Pythia more FSR
$e\tau_h + \tau_e\tau_h$	0.52 ± 0.04	0.52 ± 0.04
$\mu\tau_h + \tau_\mu\tau_h$	0.41 ± 0.04	0.37 ± 0.04
total	0.93 ± 0.06	0.89 ± 0.06

Table 5.7: Number of expected events in 193.5 pb^{-1} for an inclusive $t\bar{t}$ Pythia sample with less FSR and an inclusive $t\bar{t}$ Pythia sample with more FSR than our default.

off. As shown in Table 5.6, the result is 7%.

- Final State Radiation (FSR)

  The systematic uncertainty associated with FSR is calculated by changing the FSR evolution Kfactor to increase or decrease the amount of FSR in the sample. With the Kfactor changed from 1.0 to 2.0 we have less FSR in the sample than in the default inclusive Pythia $t\bar{t}$ Monte Carlo. With the Kfactor changed to 0.5 we have more FSR in the sample. The difference in the acceptance between the two modified samples is less than the statistical uncertainty of 7% so we take 7% to be our systematic uncertainty. The study details are shown in Table 5.7.

- PDFs

  Our default inclusive Pythia $t\bar{t}$ Monte Carlo sample is based on CTEQ5L[33] with α_s = 0.118. We calculate the systematic uncertainty due to PDFs by varying the internal parameters of the PDF, varying the the choice of PDF group and varying α_s[34]. With the new PDF set CTEQ6M, the CTEQ group made available 40 complementary PDF sets CTEQ6M.01...CTEQ6M.40 each of which represents an up or down variation along one of the twenty eigenvectors (corresponding to the ~ 20 free parameters) which collectively form an orthonormal basis set spanning the PDF parameter space [35]. Each up and down variation pair represents the range of PDF behavior that is consistent with the current global data. Each event

Source	Systematic Uncertainty
Jet Corr/Energy Scale	$\pm 5.8\%$
Electron/Muon ID	$\pm 5\%$
Tau ID	$\pm$ 6%
MC Generator	$\pm 7\%$
ISR	$\pm 7\%$
FSR	$\pm 7\%$
PDF	$\pm 1\%$
Total	16%

Table 5.8: Summary of Systematics for Acceptance

in the Pythia ttopei inclusive $t\bar{t}$ sample is reweighted according to the ratio of the CTEQ6M PDF values and the CTEQ6M.xx PDF values. Then all normal selection cuts are applied using full simulation and reconstruction. Our total systematic uncertainty is 1%.

- Electron/Muon ID

  The dependency of the efficiency of e, μ identification on the number of jets in the event gives us a systematic uncertainty of 5% on our primary lepton identification[40].

- Tau ID We measure this tau ID scale factor and calculate its systematic uncertainty by comparing $W \rightarrow \tau\nu$ data and Monte Carlo, as described in the following chapter. We calculate a scale factor with an uncertainty of 6%.

Our total systematic uncertainty on the acceptance is 16%.

Chapter 6

Tau Identification in $W \to \tau\nu$

In our acceptance calculation in Chapter 5 we take the efficiency from the Monte Carlo. We need to understand how well the Monte Carlo models our tau identification efficiency. We multiply our result from the Monte Carlo by a normalization, or scale factor, that accounts for the differences between our Monte Carlo predicted tau identification efficiency and our true tau identification efficiency as is measured a control sample, $W \to \tau\nu$, in the data. We use the same procedure when we calculate our background contributions from physics processes involving real taus ($Z \to \tau\tau$, WW, WZ) described in Chapter 7 where we take the efficiency from the Monte Carlo and apply a scale factor. This chapter describes our calculation of the scale factor.

In order to study tau identification in the data we would like to have a clean sample of taus. The physics process $W \to \tau\nu$ provides us with this sample. The signature of this process is a large amount of $\not{E}_{\mathrm{T}}$ and a high E_T, isolated tau. We apply cuts to a data sample that is independent of our main analysis data sample to isolate a fairly pure sample of taus and then we apply the same cuts to $W \to \tau\nu$ Monte Carlo. We take into account and correct for our trigger efficiencies and backgrounds. We also account for control process differences that do not depend on our tau ID but would result in the data

giving different results than the Monte Carlo. We then use our Monte Carlo, assuming the well measured theoretical $W \rightarrow \tau\nu$ cross section and branching ratio, to predict the number of events we should see in the data. The difference between the number of taus that we predict with our Monte Carlo and the number that we measure in the data gives us our scale factor.

6.1 Event Selection

We use 57.6 $\pm$ 3.4 pb^{-1} of Run2 data for this study. This sample is used, as opposed to the full 194 pb^{-1}, because this data was taken with a stable tau trigger that has been studied and is well understood. Our events come from the `TAU_MET` trigger path which, with its requirements on high $\not\!\!E_{\mathrm{T}}$ and an isolated tau candidate, includes taus from W decays. The following requirements are in the trigger path:

- Level1: $\not\!\!E_{\mathrm{T}} > 25$ GeV
- Level3: $\not\!\!E_{\mathrm{T}} > 20$ GeV, $\geq 1\ \tau$ candidate

The cuts at Level3 on the tau candidate are as follows:

- $E_t > 20$ GeV
- $|$detector $\eta|\ < 1$
- seed track $P_t > 4.5$ GeV
- seed tower $E_t > 6$ GeV
- no tracks in the 10-30 degree annulus around the seed track

We use a Pythia $W \rightarrow \tau\nu$ Monte Carlo sample and we apply all cuts listed above to this sample. We enrich our data sample with $W \rightarrow \tau\nu$ events with the following selection cuts that are applied to both Monte Carlo and data unless noted. The event must have passed the `TAU_MET` path (data only.) No tracks in 10 to 30 degree isolation annulus around the seed track are allowed. There is a large overlap between this cut and our isolation annulus tau ID cut. We require the corrected $\not\!\!E_{\mathrm{T}}$ in the event to be greater than 30 GeV. The tau candidate must pass all ID cuts shown in Table 4.4, with Track + π^0 P_t greater than 25 GeV. We veto events with $\geq$ one 5 GeV jet, $|\eta| < 2.0$. This requirement significantly reduces backgrounds but adds the complication of requiring an understanding of the Monte Carlo modeling of jet multiplicity. This cut is also referred to as the "monojet" cut because it leaves us with a sample of events with only one jet, the tau candidate itself.

6.2 Corrections to Monte Carlo

We need to correct our Monte Carlo for trigger efficiencies and the modeling of jet multiplicity. We also must scale the Monte Carlo using the data luminosity.

Trigger Efficiency Corrections

We correct our Monte Carlo for the trigger efficiency of the Level1 $\not\!\!E_{\mathrm{T}}$ trigger. The Level1 trigger $\not\!\!E_{\mathrm{T}}$ requirement is 20 GeV. For $\not\!\!E_{\mathrm{T}}$ values of 30 GeV and above, as we require in this study, the largest correction we apply due to the trigger is 5%.[36] Our 30 GeV $\not\!\!E_{\mathrm{T}}$ cut, at 10 GeV above the Level3 trigger threshold, ensures that we never need to correct by a large factor because of the high trigger efficiencies at high $\not\!\!E_{\mathrm{T}}$.

Monojet Cut Correction

In our selection cuts we have a requirement that the number of jets that do not match in ΔR to the tau candidate must be zero. This cut significantly reduces the non-$W \rightarrow \tau\nu$ backgrounds in the data. However, the number of jets in an event, or jet multiplicity, is a difficult quantity to model in the Monte Carlo. Even small discrepancies between jet multiplicity in the data and Monte Carlo will have an effect on our scale factor measurement so we need to correct for the difference. We therefore compare event jet multiplicity in a very clean sample of $Z \rightarrow \mu\mu$ data to jet multiplicity in Pythia $Z \rightarrow \mu\mu$ Monte Carlo. The jet multiplicity distribution in Z boson events should be similar to the jet multiplicity distribution in W boson events so we expect corrections we find in Z boson events to be applicable to our W events.

We isolate a clean sample of $Z \rightarrow \mu\mu$ in both the data's inclusive muon sample (193.5 pb^{-1}) and a Monte Carlo Pythia sample. We make the following cuts:

- Tight CMX or CMUP muon with $P_T > 20$ GeV
- Track with $P_t > 20$ GeV
- Muon and muon object (track) are back-to-back ($\Delta\phi > \pi - 0.5$)
- Reconstructed mass of the two objects is between 75 GeV and 105 GeV

We estimate our background in the data using the same-sign events. After background subtraction we have 72023 Monte Carlo events and 7091 data events. We plot the number of 5 GeV jets in the data and Monte Carlo events in Figure 6.1 with both distributions normalized to unity. We see that we have a higher fraction of Monte Carlo events in the 0 jet bin than data events. We therefore scale our Monte Carlo by the ratio of the fraction of events in the 0 jet bin in data divided by the fraction of the events in the 0 jet bin in the Monte Carlo before we can compare data and Monte Carlo distributions.

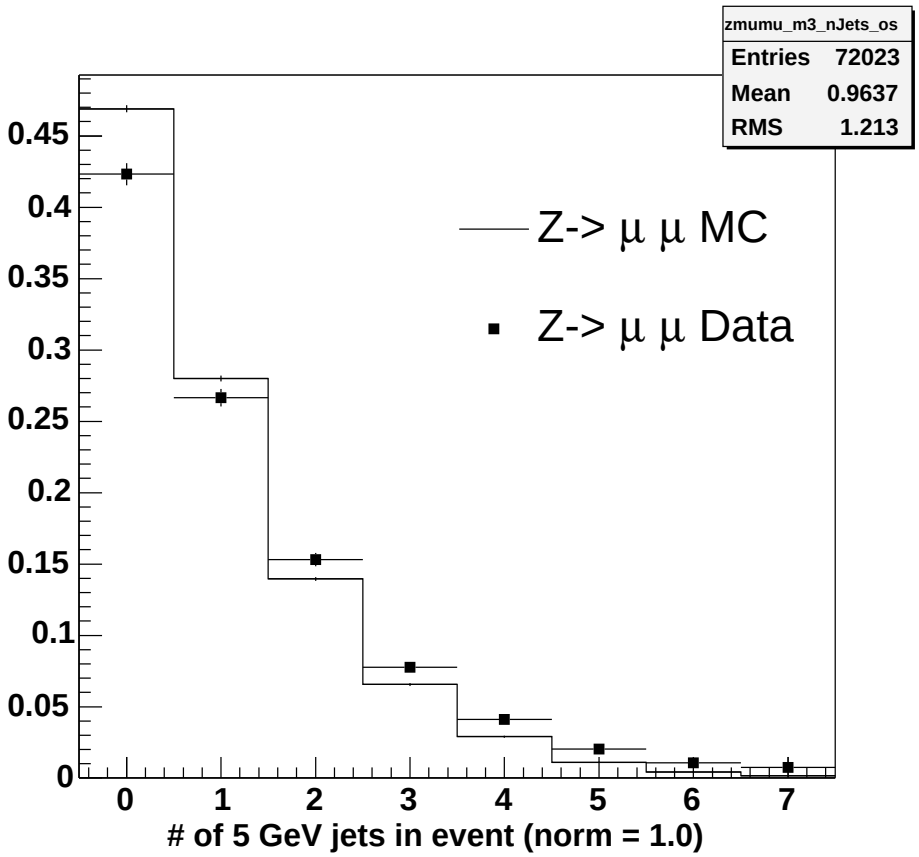

Figure 6.1: The distribution of the number of 5 GeV jets in $Z \to \mu\mu$ events in data and Monte Carlo. Both distributions have been normalized to unity.

All plots of the $W \to \tau\nu$ Monte Carlo in this note have this monojet scale factor applied to them.

We must also scale for the luminosity of the samples involved. Assuming a $W \to \tau\nu$ σX BR of 2690$\pm$100 pb, the NNLO theoretical prediction[37], we have the equivalent of (171 $\pm$ 6) pb^{-1} of Monte Carlo. In addition to the correction for the monojet cut we apply the correction for luminosity (57.6/171) to the Monte Carlo sample before we compare the Monte Carlo with the data.

6.3 Backgrounds

To measure the QCD background we loosen our tau ID cuts to:

- track + π^0 $P_t > 25.0$ GeV
- electron veto: HadE/ΣP > 0.15
- |charge of tau| = 1
- $<$ 4 tracks in tau cone

This sample includes our $W \rightarrow \tau\nu$ events and an increased number of jet to tau fakes in the data. When we create the same sample in the Monte Carlo we simply have loose taus, since there is a tau available in each of the $W \rightarrow \tau\nu$ events.

We apply a relative jet to tau fake rate to the data sample to understand our amount of QCD background. The relative fake rate is calculated with the same procedure described in Section 7.1.1 with one additional denominator cut, where we have required the tau candidate to be isolated in an annulus from 10 degrees to 30 degrees around the seed track. One problem with this procedure is that we assume that all of the events we apply our fake rate to are jets, but we have reason to believe that a significant number of these events contain real taus due to our selection cuts.

One way for us to understand how much of an overestimate the above method is on the QCD background is to apply the exact same procedure to the Monte Carlo where we are assured that we do have a tau in each event. While applying a jet to tau fake rate on a sample of Monte Carlo taus is not a meaningful thing to do in and of itself, it will give us an understanding of what the contribution is to our QCD background calculated above that is present because of taus in our events and not jets.

When we apply our jet to tau fake rate to our data sample, which is presumably a mixture of jets and taus, and to our Monte Carlo sample, which is a sample of pure taus,

we have similar numbers “QCD background” events. The jet and tau mixture from the data give us 56.8 $\pm$ 14.8 events. The taus from the Monte Carlo give us 53.7 $\pm$ 14.0 events. This exercise allows us to put a limit on the total number of QCD background events that our data sample can support.

The comparison can be seen between the Monte Carlo and data predictions in Figure 6.2, where we have dropped the cut on the tau charge and the cut on the number of tracks in the tau to show the track multiplicity distribution of the two samples, after the fake rate has been applied. We see that the data distribution is consistent with the distribution of pure taus in the Monte Carlo. This shape can be contrasted with the shape of the track multiplicity for a sample of actual jets. The first and third bins in track multiplicty are plotted for jets in Figure 7.4. Note that the number of jets with three tracks is greater than the number of jets with one track in the jet distribution. We conclude that our monojet data cannot support a significant amount of QCD background.

We reach a different conclusion with our electron background where we believe we are able to detect a contribution coming from $W \rightarrow e\nu$. We determine this background by measuring the number of events that has a tau candidate that passes all cuts except for the electron veto cut. Following a procedure outlined in Section 7.1.2 we apply our relative e to tau fake rate to this sample of (1.2$\pm$0.3)$\%$ as measured in Section 7.1.2. Our background due to electrons faking taus is therefore predicted to be 105.8$\pm$26.5 events.

6.4 Scale Factor

We divide the number of events we see in the data with the number of of events we predict for the given amount of luminosity with the Monte Carlo. This number is our scale factor than can later be applied to Monte Carlo samples to correct the tau identification efficiency so that it agrees with the tau identification efficiency in the data. Our resulting scale factor is 0.95 $\pm$ 0.06. The largest contributions to the uncertainty, in decreasing

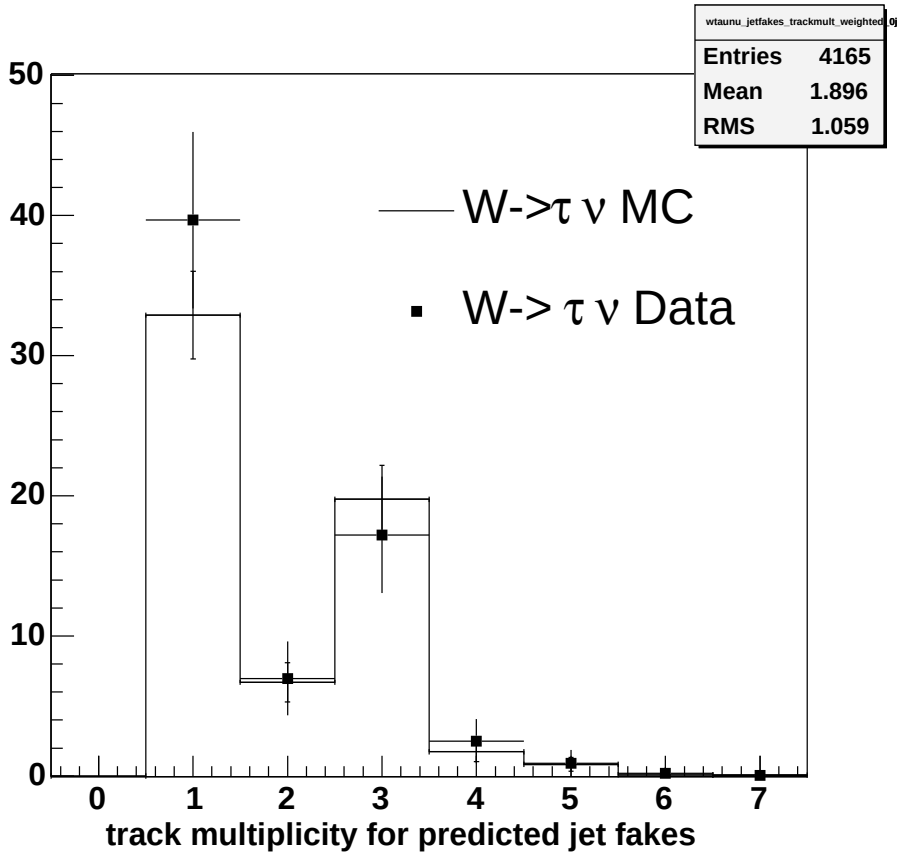

Figure 6.2: We show the predicted background from jet fakes in data and Monte Carlo. The agreement between the two where, in the Monte Carlo case we expect no background, gives us confidence to neglect this background in our scale factor calculation.

order, are the uncertainty on the 57.6 pb^{-1} luminosity, the 2690 $\pm$ 100 pb cross section, and the 25% uncertainty on the electron background subtraction.

We show the track multiplicity distribution of our $W \rightarrow \tau\nu$ events, where we have removed the charge and number of tracks in tau cone cuts to see the shapes in Figure 6.3. We show the Monte Carlo stacked on top of the electron background contribution, allowing the total to be compared to the data points before the electron subtraction. Other comparisons, plotted in the same style but after all tau ID cuts, are shown in Figures 6.4 - 6.7. Note that the calorimeter isolation variable is not well modeled in the Monte Carlo (Figure 6.4.) We believe this discrepancy accounts for the majority of the measured scale factor.

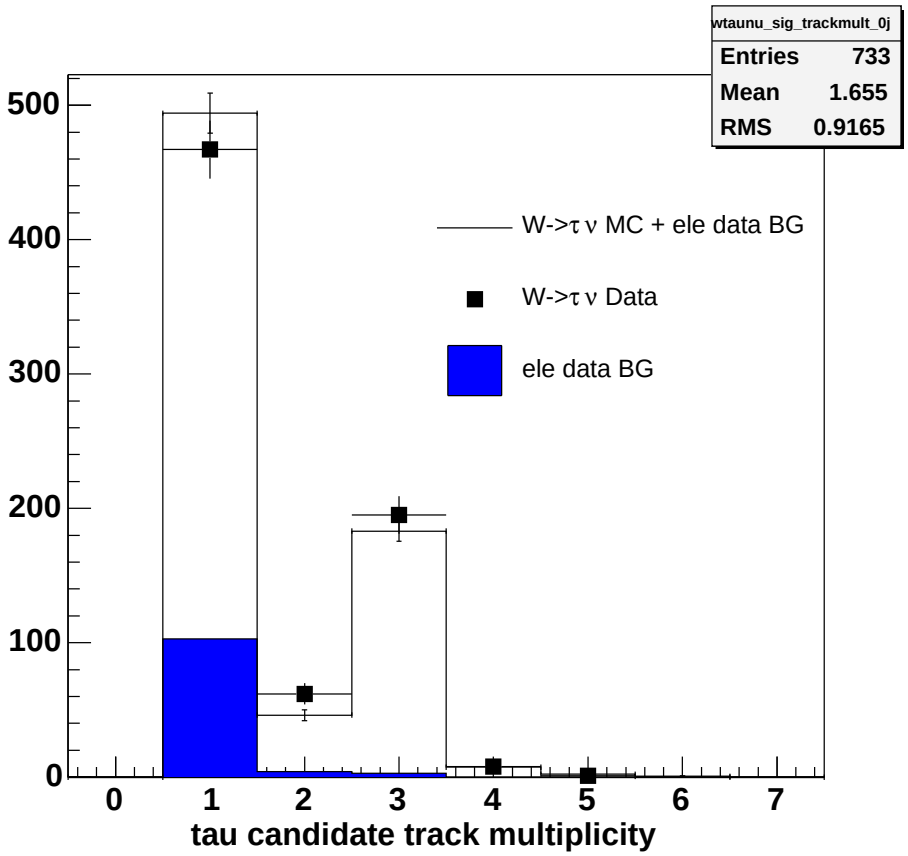

Figure 6.3: Final track multiplicity prediction in the data and the Monte Carlo. We show absolute predicted number of events from the Monte Carlo and measured events in the data before the electron subtraction. The electron background and Monte Carlo contributions have been stacked on top of each other so that the total can be compared to the data.

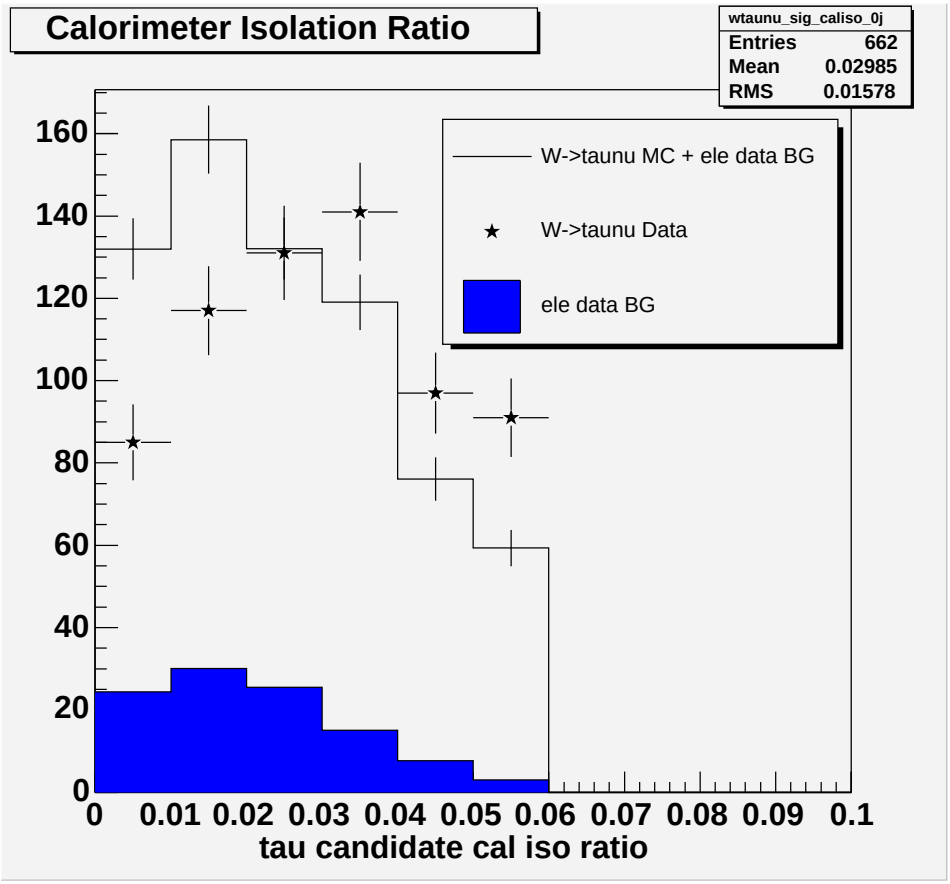

Figure 6.4: The data (points) and Monte Carlo plus electron background (histogram) calorimeter isolation ratio distributions are shown. Calorimeter isolation ratio is poorly modeled in the Monte Carlo which contributes to the overall scale factor that we must apply to our acceptance and to our real tau Monte Carlo backgrounds.

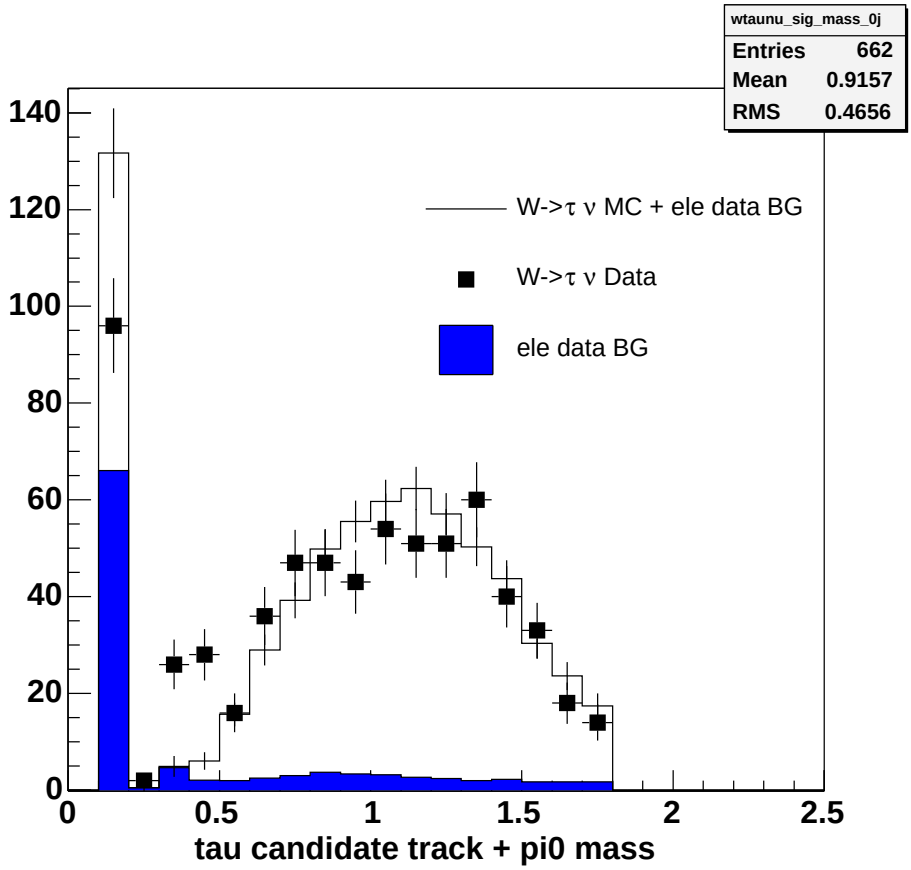

Figure 6.5: The data (points) and Monte Carlo plus electron background (histogram) tau candidate track + π^0 mass distributions are shown.

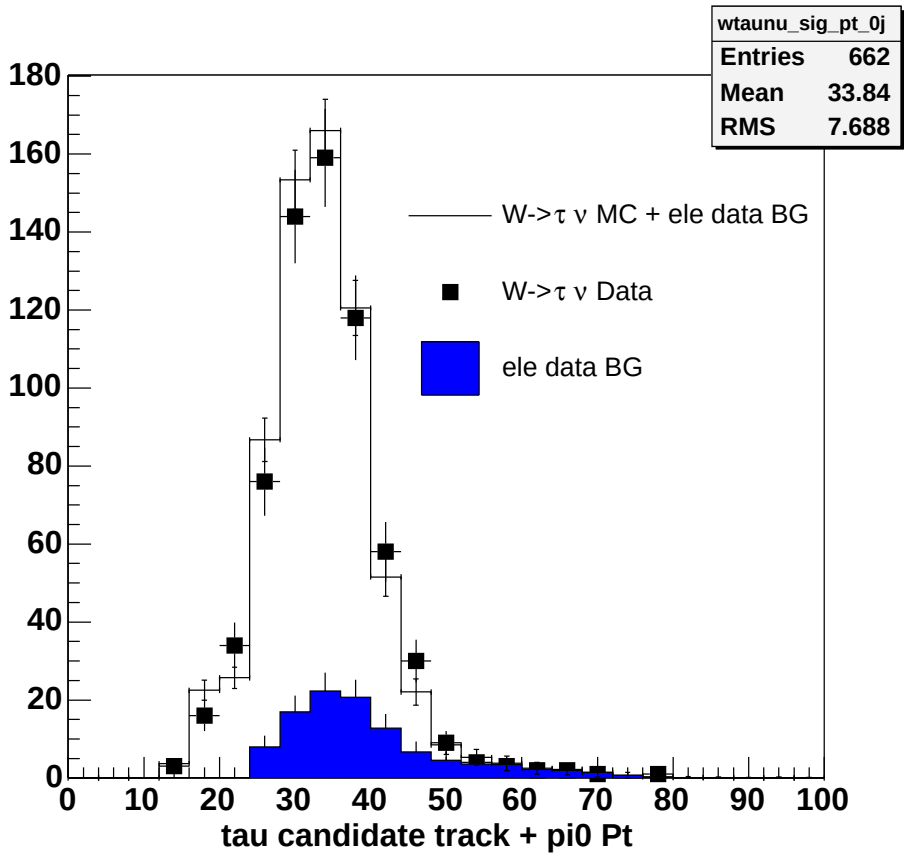

Figure 6.6: The data (points) and Monte Carlo plus electron background (histogram) track + π^0 P_t distributions are shown.

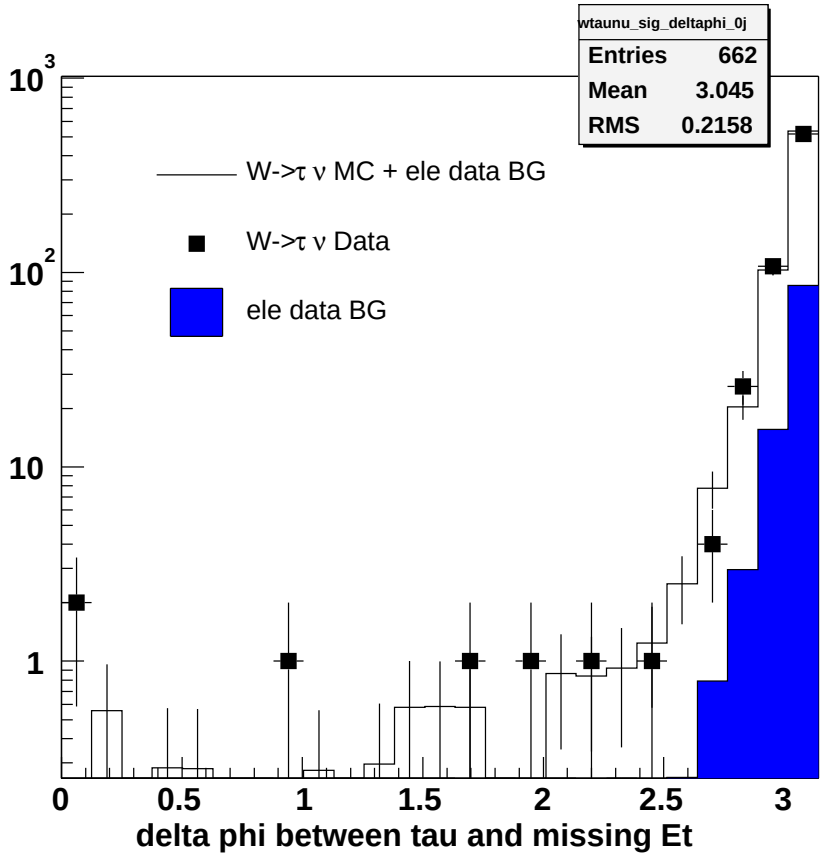

Figure 6.7: The data (points) and Monte Carlo plus electron background (histogram) $\delta\phi$ between the tau and the event $\not{E}_{\mathrm{T}}$ distributions are shown. We expect to see a peak at π for $W \to \tau\nu$ events.

Chapter 7

Backgrounds

Our background events can be separated into two categories; those with real taus and those without real taus. This distinction is useful because the probability of fake taus is higher by more than an order of magnitude than the probability of a fake electron or muon.

The highest contributions to the background consist of events that pass our signal cuts with an electron, muon, or jet incorrectly identified as a tau. When a non-tau object passes all tau identification cuts it is said to "fake" a tau. Each type of tau fake (electron, muon, and jet) has its own contribution to the overall background. The methods used to estimate these background contributions are described in this chapter.

A small number of background events comes from physics processes with hadronic tau decays in their final states, $p\bar{p} \rightarrow \gamma^*/Z, WW$ and WZ with at least one τ in the final state. The methods used to estimate these background contributions are also described in this chapter.

7.1 Fakes

We first consider the cases where jets, electrons, and muons pass tau identification cuts. The events that have a tau fake and also pass all other event selection cuts will enter our final sample as background events. The jet fakes are the largest contribution to the background because they are the most numerous. Electrons and muons, while they are more rare in events than jets, are more likely than jets to fake taus.

7.1.1 $j \rightarrow \tau$ Fakes

Jets tend to fake taus when they have low track multiplicity and are narrow, or collimated. We do not trust the simulation's modeling of the jet fake rate so we measure the rate from data.

We have measured the rate at which a jet fakes a tau using the four datasets Jet20, Jet50, Jet70, and SUMET. The Jet20 dataset is fed by a trigger that requires a central, 20 GeV jet in the event. Similarly, the Jet50 and Jet70 datasets are based on triggers requiring a central, 50 GeV jet and a central, 70 GeV jet respectively. The SUMET dataset comes from a trigger that requires a minimum of 4 central jets, each with $E_T >$ 15 GeV and the total energy deposited in the event to be at least 125 GeV.

We use only non-trigger jets in this study so that our fake rate is not biased by the trigger. A reconstructed jet is considered a trigger jet if it passes the requirements for a Level1 and Level2 and Level3 trigger jet. Figure 7.1, plot f, shows the number of trigger jets in each event.

Note that several events do not have a trigger jet defined. There are several reasons for this: a) the jet trigger requirement at each level was not satisfied by the same calorimeter cluster, b) the matching of offline jets to trigger objects is not exact, and c) calibration constants used in the Level3 trigger may differ from those used offline so a jet that passed the online trigger requirements may not pass the same requirements

offline when it is reconstructed.

The following algorithm was used to define the “biased” and “unbiased” jet samples. Events are divided into three categories based on the number of identified trigger jets. If the event does not have any trigger jets it is not used in this study. If the event has one trigger jet, that jet is considered biased and is not used to determine the jet to tau fake rate but all other jets in the event are used. If the event has two or more trigger jets then all jets in the event are considered unbiased and they are all used in the calculation of jet to tau fake rate. Because four jets are required in the SUMET sample the definition of a trigger jet is very different there than it is in the Jet20, Jet50 and Jet70 datasets. We take all jets as unbiased in the SUMET dataset for the purposes of this study.

The E_T and η distributions for unbiased and biased jets are shown in Figure 7.2. Note that the biased Jet E_T plot has entries below the E_T thresholds. This is possible because the clustering cone size is different online, where $R = 0.7$, and offline, where $R = 0.4$. Also, the online and offline jets have different algorithms to determine their vertex position.

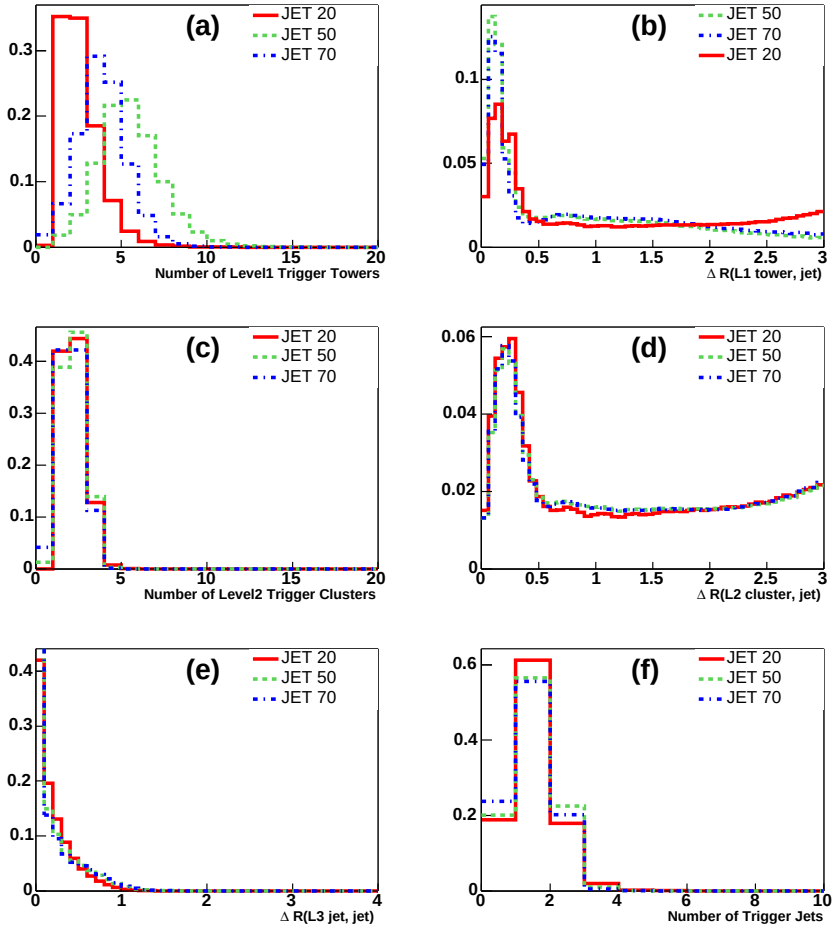

Figure 7.1: Distributions for Level1, Level2 and Level3 trigger objects. The ΔRs shown are the distances between the online (trigger) jets and offline jets.

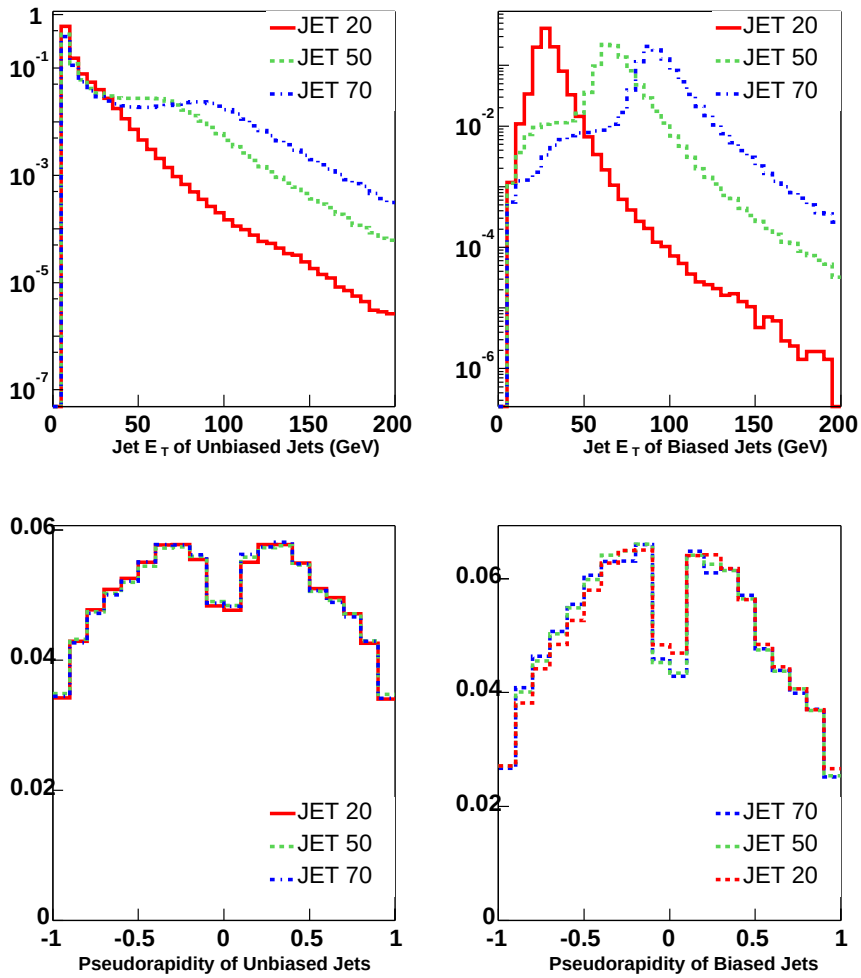

Figure 7.2: E_T and η distributions for biased and unbiased jets.

We look for tau candidates, reconstructed as described in section 4.5, that are near the jets. Figure 7.3 shows the distance in ΔR from each unbiased jet to the nearest tau candidate. Each reconstructed jet is required to have $E_T > 10$ GeV and to be within 0.2 in ΔR of a tau candidate. Only tau candidates matched to a jet are used to determine the jet to tau fake rate.

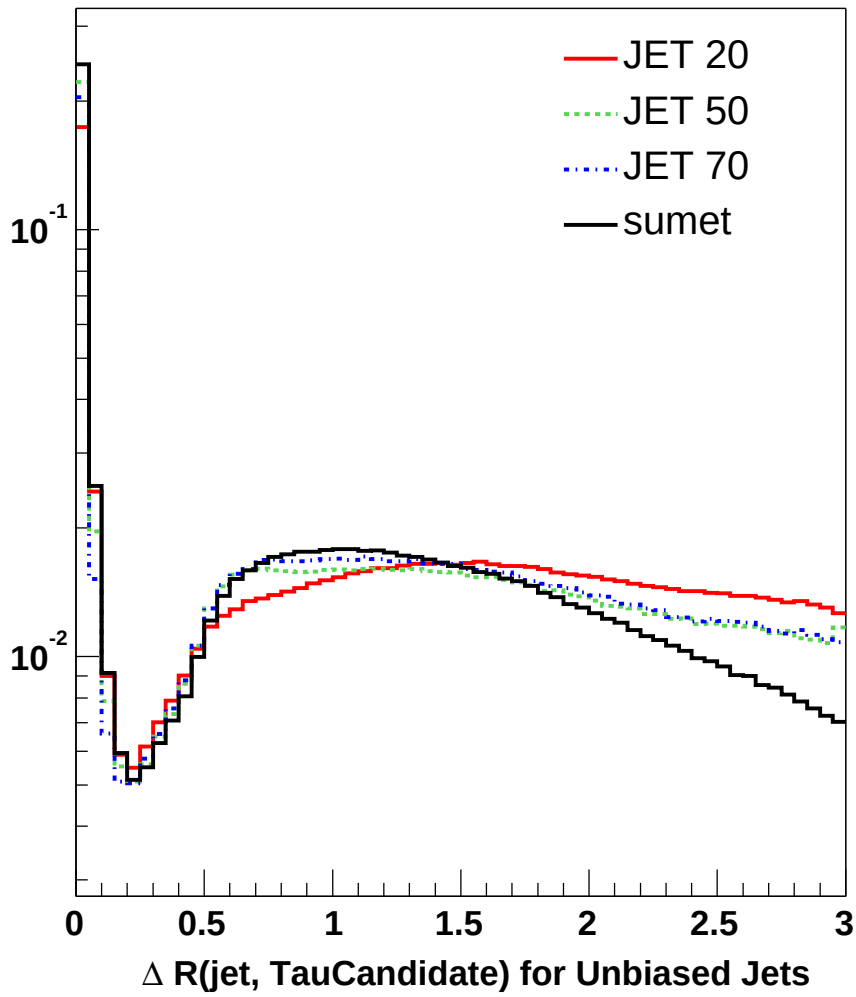

Figure 7.3: ΔR distance from each unbiased jet to the nearest tau candidate.

Relative Fake Rate

In the analysis of the data we will create a sample of loosely identified taus which do not pass all tau cuts. This sample, dominated by jets, can be used to statistically predict the number of jets which do pass all tau cuts by applying a "relative fake rate" to this sample.

The denominator objects are required to pass the basic tau candidate cuts described in Section 4.5 as well as the following tau identification cuts: (π^o + Track) $P_T > 15$ GeV, $|\Sigma$(track charge) in tau cone$|$ = 1, # tracks in Tau Cone < 4, and $E_{HAD}/\Sigma P > 0.15$. The track multiplicity cuts, $|\Sigma$(track charge) in tau cone$|$ = 1 and # tracks in Tau Cone < 4, are necessary so we are able to have access to a meaningful tau candidate charge.

We will be looking at the jet to tau fake rate in both same sign events, where the tau candidate has the same sign as our primary electron or muon, and opposite sign events, where our tau candidate has the opposite sign of our primary electron or muon. The electron veto cut, $E_{HAD}/\Sigma P > 0.15$, is in the denominator in order to ensure that our jet to tau fake rate sample is orthogonal to our electron to tau fake rate sample. The numerator objects pass all denominator cuts as well as the remaining tau identification cuts.

Figure 7.4 shows the four tau identification variables in the denominator after all denominator cuts are applied. Figure 7.5 and 7.6 show the remaining tau identification variables after all numerator cuts are applied except for the cut on the variable being plotted.

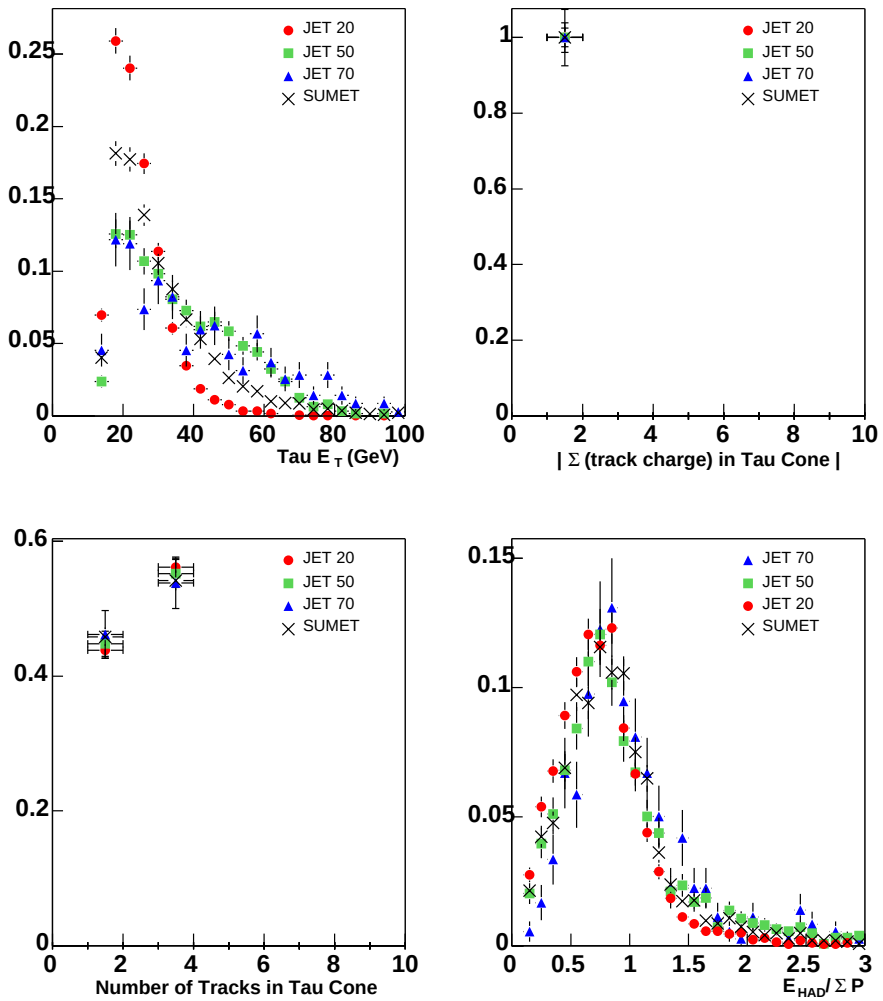

Figure 7.4: Comparison of tau identification variables for the denominator of the fake rate. All denominator cuts have been applied.

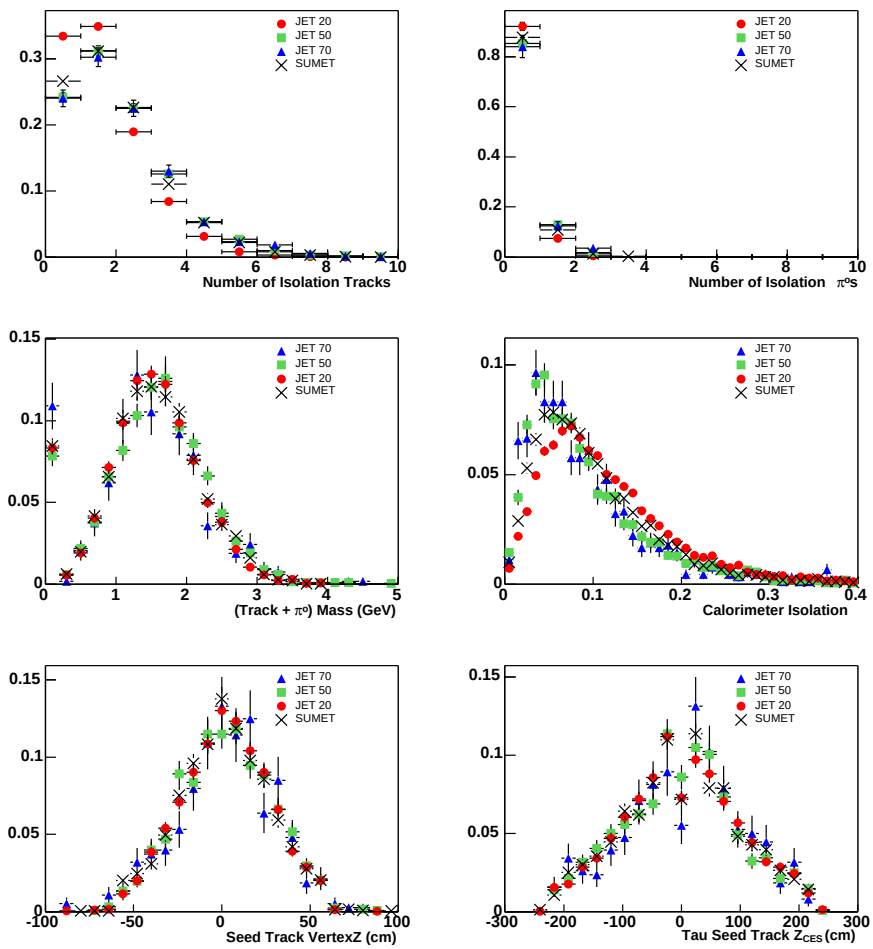

Figure 7.5: Comparison of tau identification variables for the numerator of the fake rate. All tau ID cuts have been applied except for the cut on the variable being plotted.

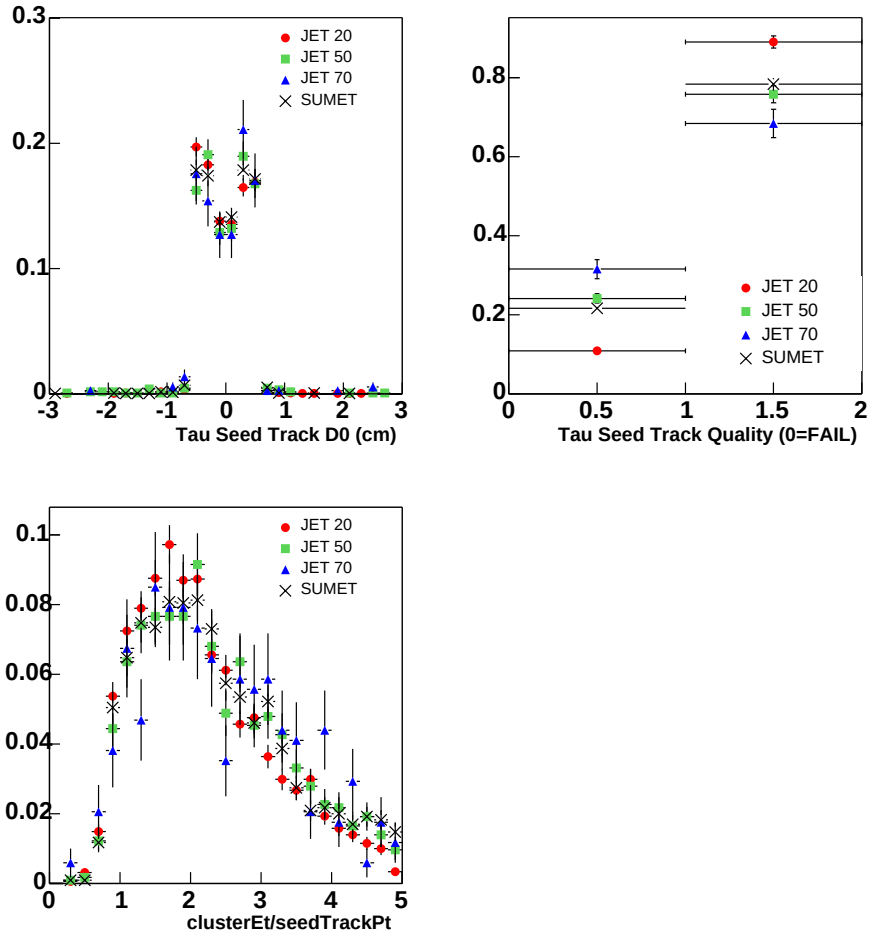

Figure 7.6: Comparison of tau identification variables for the numerator of the fake rate. All tau ID cuts have been applied except for the cut on the variable being plotted.

The relative fake rate for unbiased jets as a function of (π^o + Track) P_T is shown in Figure 7.7a. The relative fake rates for biased jets are also shown in Figure 7.7. There is a trend for the fake rate to be higher for the biased jets, presumably due to the characteristics of the trigger selections. As mentioned above, only unbiased jets are used to determine the fake rate that is used in this analysis.

Figure 7.7a shows that the relative fake rate measured from the Jet20 sample is considerably larger than than the relative fake rate measured from the other samples. It is reasonable to take this difference as a measure of the uncertainty in our fake rate prediction procedure. However, this difference is so large that it would significantly degrade our attempt to measure the background due to jet fakes.

In order to reduce this systematic uncertainty we need to choose another variable, in addition to (π^o + Track) P_T, and parametrize the fake rate as a function of the two variables.

Studies show that the fake rate is strongly correlated with the isolation of the jet, or the amount of energy deposited in regions of the calorimeter near the jet. An isolation variable is defined for each tau candidate as the sum of transverse energy over calorimeter towers in the range $0.4 < \Delta R < 1.0$. This sum is normalized by dividing by the (π^o + Track) P_T. The normalized isolation variable is shown in Figure 7.8 after all tau ID cuts are applied.

Six regions of isolation are defined: 0.0-0.05 (ISO1), 0.05-0.1 (ISO2), 0.1-0.2 (ISO3), 0.2-0.3 (ISO4), 0.3-0.5 (ISO5), > 0.5 (ISO6) where ISO1 corresponds to the most isolated tau candidates and ISO6 corresponds to the least isolated tau candidates. The relative fake rates within each isolation region are shown in Figure 7.9. The plots show that the less isolated tau candidates correspond to lower fake rates, because the tau identification algorithm requires some isolation. With the two variable parametrization of the jet to tau fake rate the four samples, Jet20, Jet50, Jet70, and SUMET, measure rates that are much more consistent with each other than they were when only one variable

was used for the parametrization.

As a test of the consistency of the fake rates determined from the jet samples, the fake rate measured in each sample is used to predict the number of events passing all cuts in the other samples. Table 7.1 shows that the ability of each jet sample to predict the number of events faking a tau in the other samples usually improves when the samples are first divided into the six isolation regions. For example, multiplying the fake rate determined from JET50 by the number of events in JET20 passing the denominator cuts (only as a function of E_T) gives a prediction of 1814 $\pm$ 66 events compared to the measured value 3317 $\pm$ 57 events. When the same calculation is done using the fake rate in each isolation region, the prediction is 2924 $\pm$ 123 events.

If the JET50 fake rate is used to predict the number of events expected in the other three samples, the difference ((predicted - measured)/predicted) is 13%, 1% and 26% for the JET20, JET70 and SUMET samples respectively. 26% is the largest such difference in Table 7.1 and can be taken as a measure of the systematic uncertainty in the jet to tau fake rate estimate. Note that without taking isolation into account, differences as large as 96% occur.

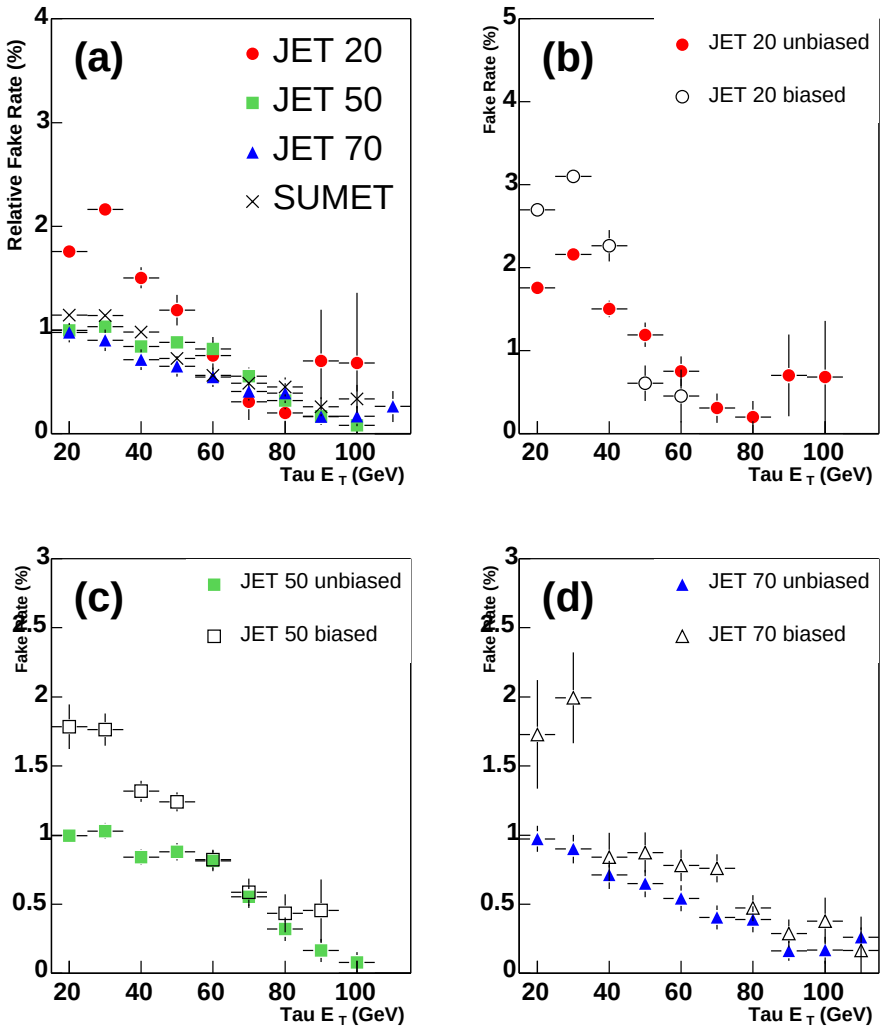

Figure 7.7: (a) Comparison of relative fake rate for unbiased jets in JET20, JET50, and JET70 samples and all jets in the SUMET sample. (b)-(d) Comparison of relative fake rate for unbiased and biased jets.

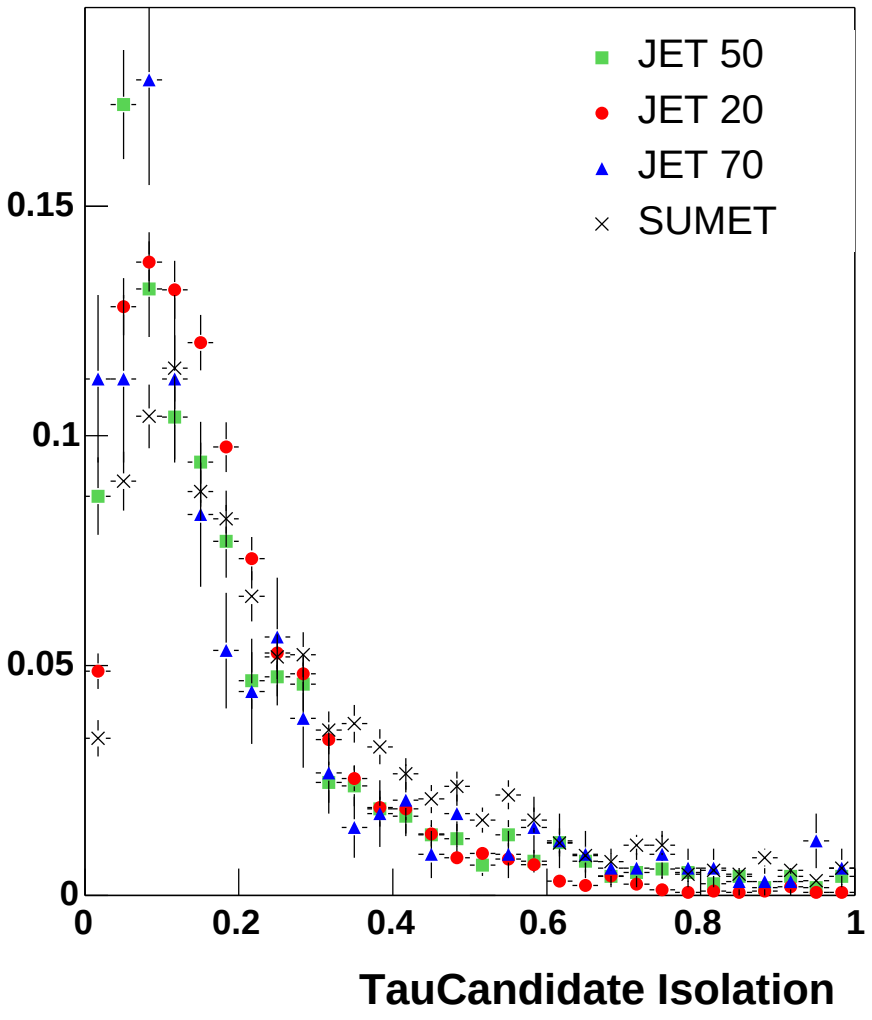

Figure 7.8: Isolation variable after all tau ID cuts are applied.

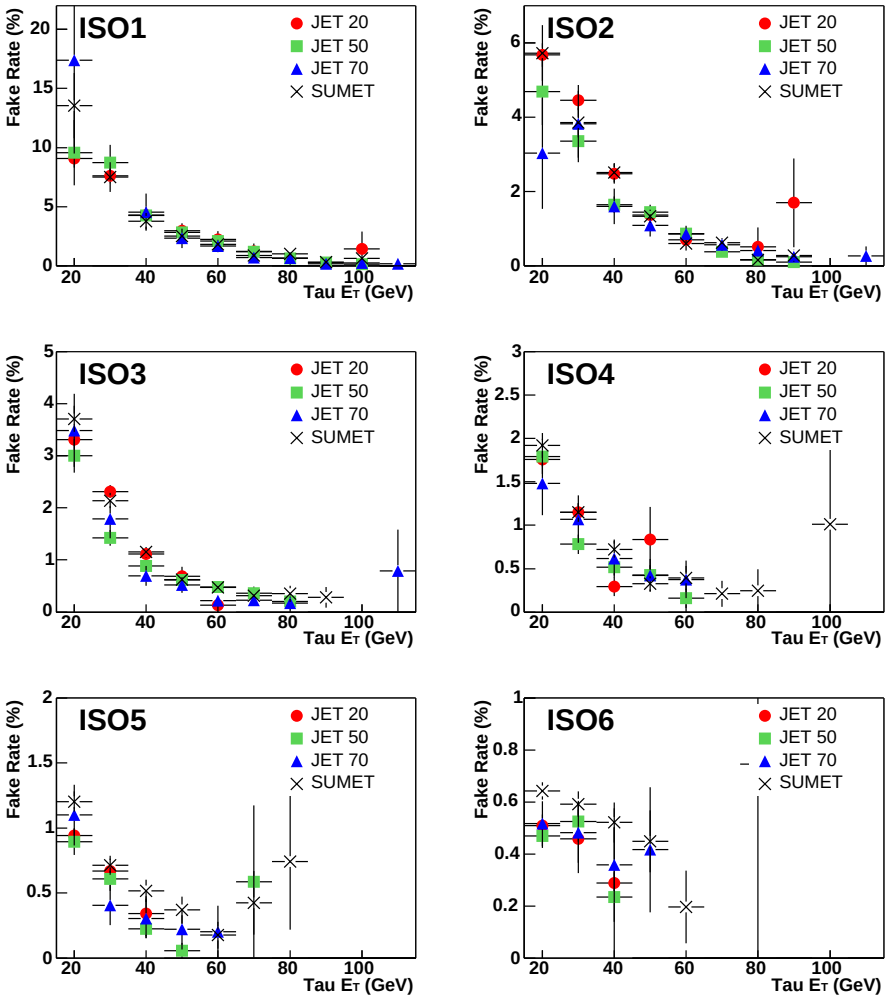

Figure 7.9: Relative fake rates in each isolation region.

		Predictions	
data sample	events passing all cuts	no isolation bins	with isolation bins
JET20	3317 ± 57	1814 ± 66 (JET50)	2924 ± 123 (JET50)
		1694 ± 120 (JET70)	2970 ± 255 (JET70)
		2043 ± 45 (SUMET)	3581 ± 100 (SUMET)
JET50	1265 ± 35	2112 ± 59 (JET20)	1408 ± 50 (JET20)
		1091 ± 59 (JET70)	1239 ± 69 (JET70)
		1324 ± 28 (SUMET)	1517 ± 33 (SUMET)
JET70	358 ± 18	661 ± 29 (JET20)	392 ± 25 (JET20)
		412 ± 12 (JET50)	356 ± 11 (JET50)
		431 ± 10 (SUMET)	439 ± 11 (SUMET)
SUMET	2466 ± 49	3984 ± 85 (JET20)	2176 ± 68 (JET20)
		2277 ± 66 (JET50)	1951 ± 60 (JET50)
		2028 ± 118 (JET70)	1999 ± 117 (JET70)

Table 7.1: Comparison of number of events passing all tau ID cuts in each jet sample with predictions using the other jet samples.

Calculating Background

In order to determine our background due to jet fakes we select a sample of tau candidates from both the tight electron and tight muon datasets that pass the fake rate denominator cuts as well as the following event selection cuts:

- One central tight electron, CMUP muon or CMX muon above 20 GeV
- One central jet ($|\eta| < 1.1$) that matches to a tau candidate passing the denominator cuts shown above
- event corrected $\not{E}_{\mathrm{T}} > 20GeV$
- $\geq$ 2 jets with $|\eta| < 2$
  - E_T of first jet $>$ 25 GeV

  - E_T of second jet $>$ 15 GeV

- $H_T > 205$ GeV

- opposite charge of tau candidate and electron/muon

Note that if an event has more than one tau candidate that passes the denominator fake rate cuts it may be entered into our denominator background sample more than once.

This fake rate sample, which we believe to be dominated by W+jet events, does contain both our signal and other sources of real taus ($Z \rightarrow \tau\tau$, WW, WZ). In order to account for this we apply the fake rate only to events where the tau candidate passes the above cuts but fails the final tau ID cuts. To avoid under-counting our fakes this way, we replace our measured relative fake rate, f, with the fake rate corrected for the missing fake taus, $f/(1-f)$. This procedure minimizes counting our signal events as jet fake background events.

We have 67 background candidate events in the e, τ channel and 37 background candidate events in the μ, τ channel to which we apply the fake rate. The total background in 193.5 pb^{-1} is 0.45 $\pm$0.10 (stat) $\pm$0.12 (sys) events for the e, τ channel and 0.30 $\pm$0.06 (stat) $\pm$0.08 (sys) events in the μ, τ channel. The first error given is the statistical error due to the number of events in our loose tau sample in the data. The second error is the systematic error due to the uncertainty on our jet to tau fake rate. The contribution of the events to the background as a function of tau candidate track + π^0 Pt can be seen in Figure 7.10.

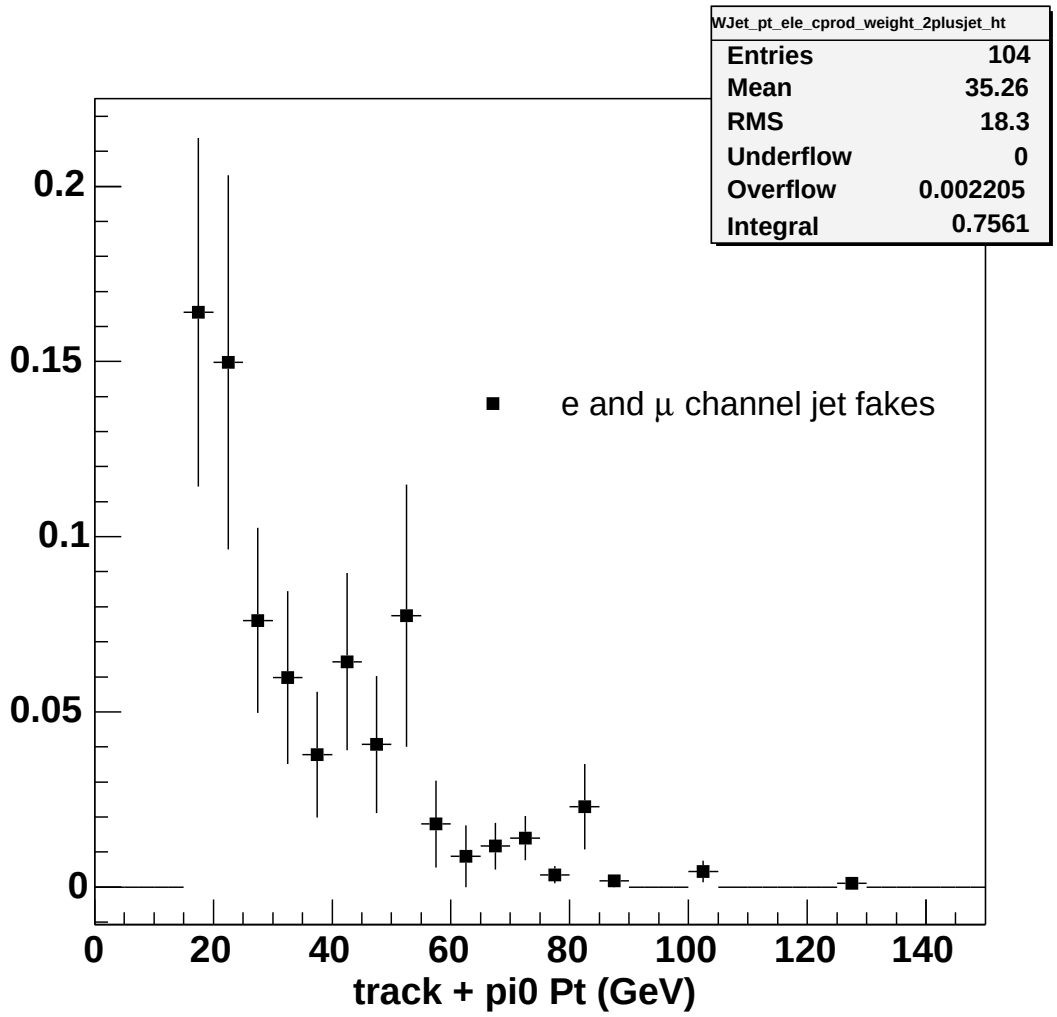

Figure 7.10: The background contribution due to jets faking taus is shown as a function of tau candidate track + π^0 P_t for the combined e, τ and μ, τ channels.

7.1.2 $e \rightarrow \tau$ Fakes

Electrons, which deposit energy in the calorimeter and tend to be isolated, fairly efficiently fake taus that have decayed hadronically with one charged track. We therefore include an explicit veto on electrons in our tau identification cuts. We determine the rate of electrons faking taus by measuring the rate of electrons passing the tau identification cuts in the data.

In order to make this measurement we require a clean and unbiased sample of electrons. We use the physics process $Z \rightarrow ee$ where tight cuts are placed on one of the electrons and very loose cuts that will not bias our results are placed on the second electron. We can then use the second electron to probe the efficiency of our electron veto cut. The details of the study are described in this section, along with the application of the electron to tau fake rate to our top to tau analysis.

Calculating $e \rightarrow \tau$ Fake Rate

We use the inclusive high P_T electron dataset for this study. This is the same dataset as is used for the e, τ channel in the main analysis. The full 195 $\pm$ 12 pb^{-1} for the $t \rightarrow \tau$ analysis was not available at the time of this study, so only the available 75 pb^{-1} was used. We require an electron with $E_T >$ 30 GeV that passes all standard inclusive high E_T cuts for central electrons, shown in Table 4.2.

The second electron must be central, with detector $\mid \eta \mid <$ 0.9. Also, $\Delta\phi$ between the first (tight) electron and the second (loose) electron must be greater than π - 0.5 because $Z \rightarrow ee$ events tend to have their electrons produced back-to-back in ϕ. We do not include events with jets of $E_T >$ 8 GeV and $\mid \eta \mid <$ 2.2 that do not match to the electrons within ΔR of 0.4 in order to reduce the non-$Z \rightarrow ee$ events in the sample. Our second electron must match, within ΔR of 0.4, to a reconstructed tau candidate in the event, with standard tau candidate requirements and tau identification cuts as described

in section 4.5. We do not yet make the electron veto cut as the point of this study is to determine a reasonable choice for the value of the cut given the rate of electrons faking taus, which is a function of the cut.

In order to increase the purity of the process $Z \rightarrow ee$ in our sample we only include events with the invariant mass of the two electrons greater than 75 GeV. Note that the mass of the Z is 91.1876 $\pm$ 0.0021 GeV[19]. A significant source of our background due to this physics process may be due to events where an electron has been mis-reconstructed. In order to not eliminate these events with our invariant mass cut we add the event $\not{E}_{\mathrm{T}}$ to our invariant mass calculation if the $\Delta\phi$ between the $\not{E}_{\mathrm{T}}$ and the second (loose) electron is less than 0.7. The resulting invariant mass is plotted for our remaining events in Figure 7.11 before the cut at 75 GeV is imposed. The strong peak at 90 GeV and low level of events outside the Z peak region gives us some confidence that our sample is dominated by $Z \rightarrow ee$ events.

There are 991 events that pass the invariant mass cut. We plot the electron veto value, $E_{HAD}/\Sigma P$, for the loose electrons (tau candidates) in these events in Figure 7.12. We calculate our fake rate as a function of $E_{HAD}/\Sigma P$ with 991 events as our denominator, and the number of events above a given $E_{HAD}/\Sigma P$ threshold as the numerator. The results are shown in Table 7.2.

We choose to place the electron veto variable cut, on $E_{HAD}/\Sigma P$, at 0.15, for a $e \rightarrow \tau$

$E_{HAD}/\Sigma P >$ (value)	Events Passing Veto	Electron Fake Rate
0.05	98	0.09 $\pm$ 0.01
0.075	54	0.055 $\pm$ 0.007
0.10	36	0.36 $\pm$ 0.006
0.125	21	0.021 $\pm$ 0.005
0.15	12	0.012 $\pm$ 0.003

Table 7.2: Electron fake rate as a function of the veto variable,$E_{HAD}/\Sigma P$.

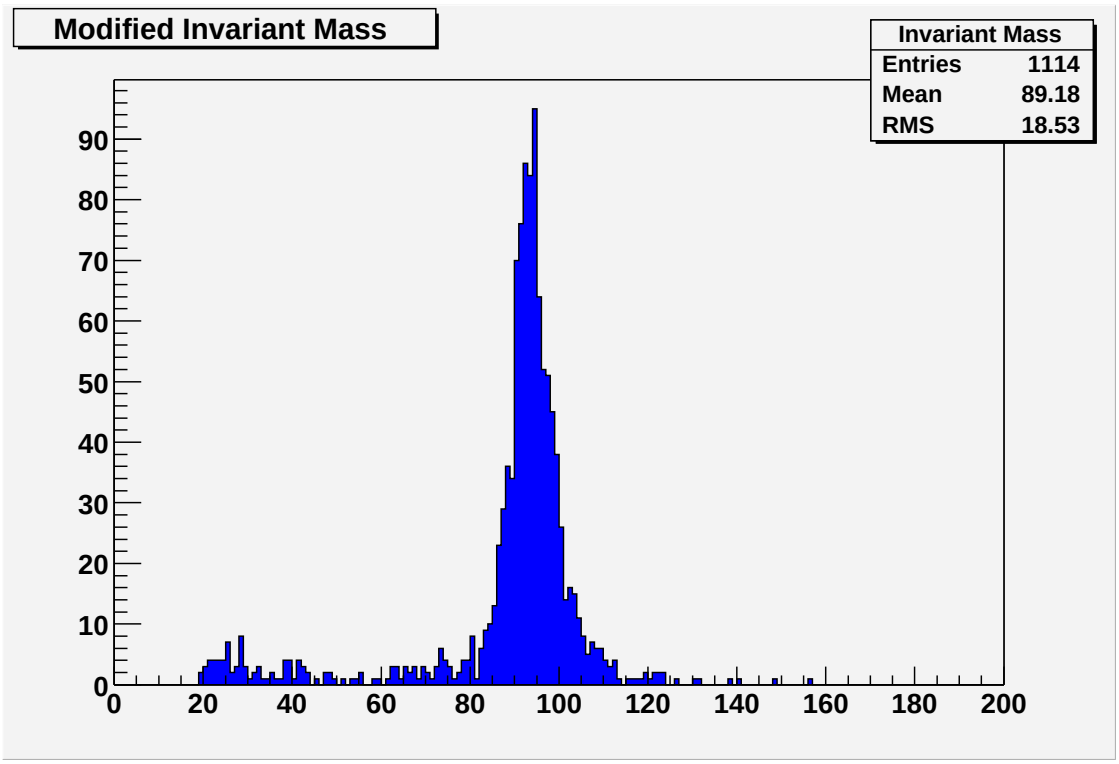

Figure 7.11: The invariant mass for electrons in our $Z \rightarrow ee$ sample after all cuts have been made on the electrons but before the 75 GeV cut on invariant mass. Note that the $\not{E}'_{\mathrm{T}}$ has been added to the loose electron if it is within a $\Delta\phi$ of 0.7.

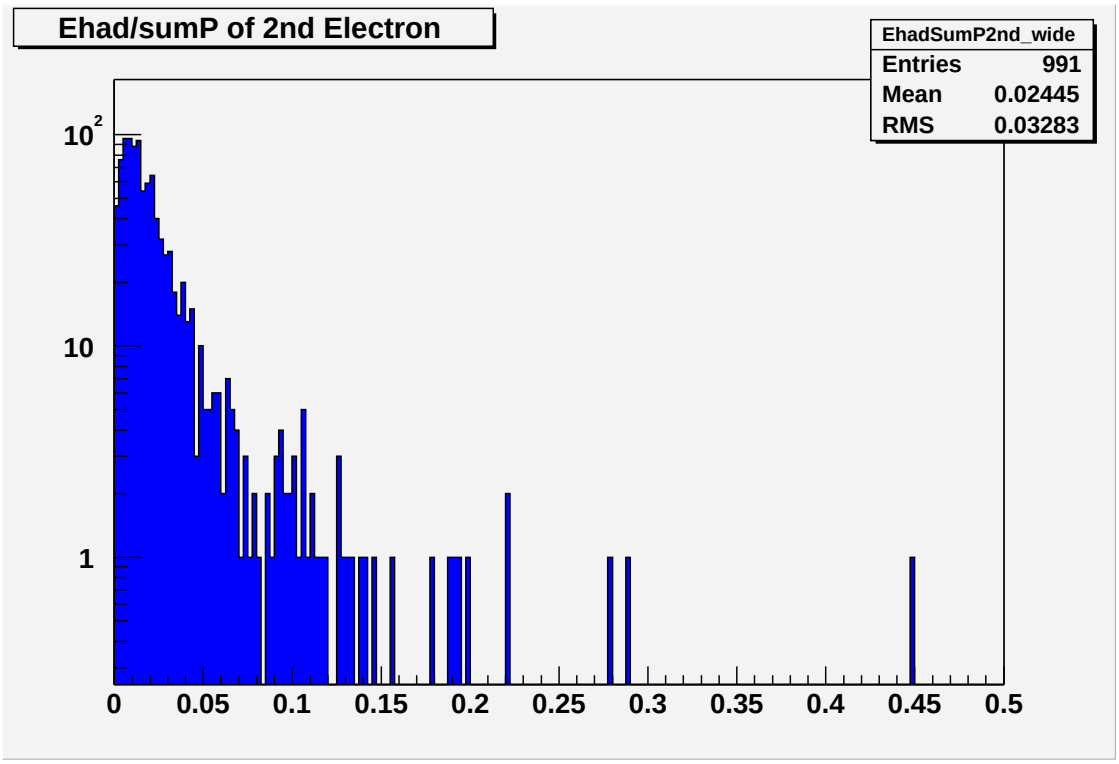

Figure 7.12: $E_{HAD}/\Sigma P$ of electron after invariant mass cut.

fake rate of (1.2 $\pm$ 0.3) $\%$.

One distinction between the sample where we measure the fake rate ($Z \rightarrow ee$) and the sample to which we apply the fake rate in the analysis is the amount of $\not\!\!E_{\rm T}$ in the events. Our $Z \rightarrow ee$ events, where there are no extra jets in the events, have a much softer $\not\!\!E_{\rm T}$ distribution than our analysis events, where we have made a minimum requirement of 20 GeV in $\not\!\!E_{\rm T}$. We therefore look at our measured fake rate as a function of event $\not\!\!E_{\rm T}$ in order to see if there is a dependence on $\not\!\!E_{\rm T}$. The results are shown in Table 7.1.2. Within statistics we see no $\not\!\!E_{\rm T}$ dependence.

$\not\!\!E_{\rm T}$Bin (GeV)	E_{HAD}/ΣP<.075	E_{HAD}/ΣP<.10	E_{HAD}/ΣP<.125	E_{HAD}/ΣP<.15
0-5	.05±.01	.039±.009	.024±.007	.014±.005
5-10	.05±.01	.033±.009	.018±.007	.013±.006
10-15	.06±.02	.03±.02	.02±.02	0.0±0.0
< 15	.054±.007	.036±.006	.021±.005	.012±.003
15-20	.05±.05	.05±.05	0.0±0.0	0.0±0.0

Table 7.3: Ratios of Events passing E_{HAD}/ΣP cuts in different $\not\!\!E_{\rm T}$ bins

Background due to $e \rightarrow \tau$ fakes

We count the number of events in the data that pass all of our event selection cuts where the tau candidate passes all tau ID cuts except it fails our electron veto cut. There are 8 events before the H_T cut, and we plot the $E_{HAD}/\Sigma(P)$ of these events in Figure 7.13 to assure ourselves that they are electron-like. The peak at zero gives us evidence of this. After the H_T cut we have seven events that pass all cuts except for the electron veto. Applying our fake rate to this sample gives us a background of 0.08±0.03(stat)±0.02(sys) events. The systematic error is due to the error on the $e \rightarrow \tau$ fake rate. Note that this background includes the physics background due to $Z \rightarrow ee$ events.

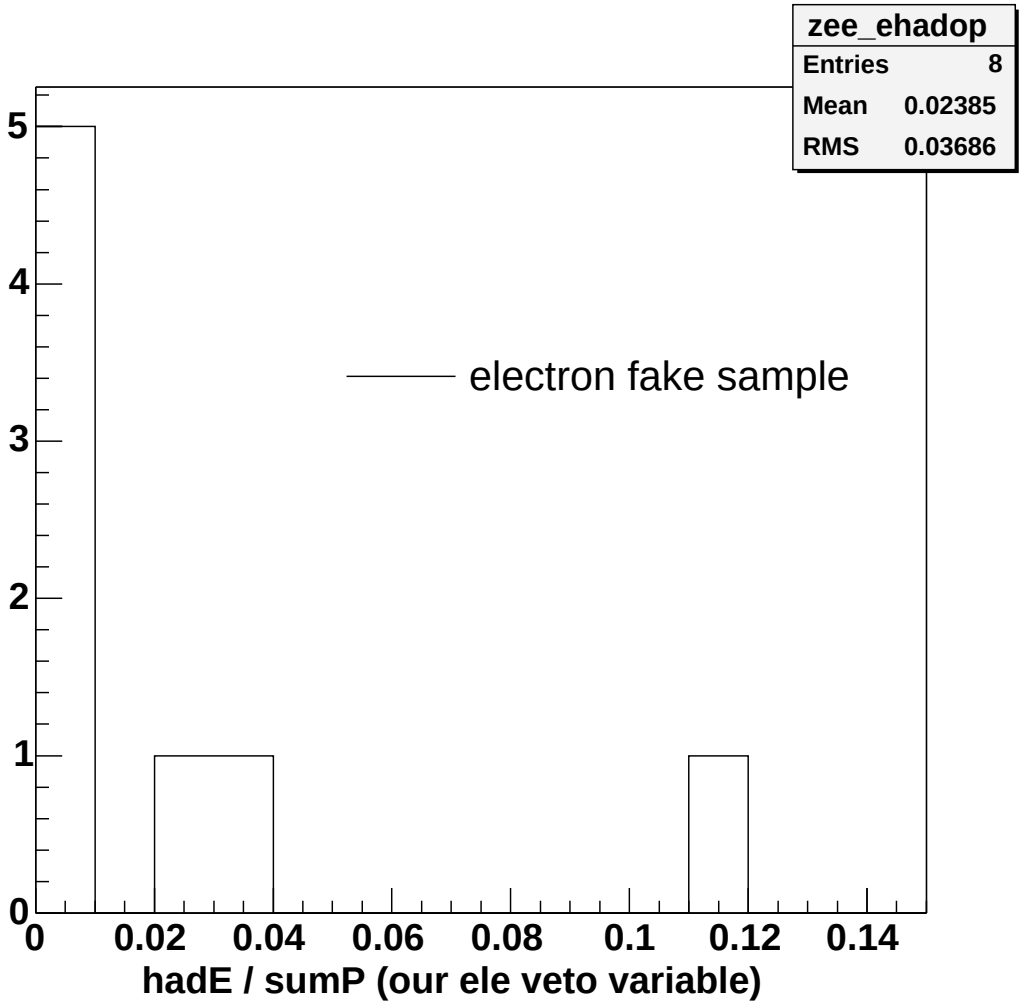

Figure 7.13: The electron veto variable, $E_{HAD}/\Sigma P$, for our electron candidate events before the H_T cut. The peak at zero gives us confidence that our sample is dominated by real electrons. Applying the H_t cut brings the sample to 7 events.

7.1.3 $\mu \rightarrow \tau$ Fakes

A muon is less likely than an electron to form a tau candidate because muons usually deposit very little energy into the calorimeter and a minimum of 6 GeV of calorimeter energy is required in order to form a tau candidate. However, the occasional muon that forms a tau candidate will share many characteristics with hadronically decaying taus. For example, the muon is likely to be isolated and will have a tendency to deposit its energy into the hadronic section of the calorimeter as opposed to the electromagnetic section. We therefore do have need of an explicit muon veto to reduce the background due to muons faking taus. The physics source for this background is $Z \rightarrow \mu\mu + jets$ events. This is the physics process that we examine, without the extra jets, both in data and Monte Carlo, to determine the $\mu \rightarrow \tau$ fake rate.

Calculating $\mu \rightarrow \tau$ Fake Rate

To measure this fake rate, we examine $Z \rightarrow \mu\mu$ events where one of the muons is identified as our primary tight lepton and the other muon fakes a tau. As mentioned above and as described in section 4.5, we have an explicit veto to reduce our background due to muons. We cut on the ratio of cluster E_T to τ seed track P_T, requiring it to be greater than 0.5.

In order to study the efficiency of the cut in the data, we need to isolate a very pure sample of muons, and the only obvious sample is $Z \rightarrow \mu\mu$. This is consistent with the approach taken to study the $e \rightarrow \tau$ fake rate in section 7.1.2. We use the inclusive high P_T muon dataset for this study. This is the same dataset as is used for the μ, τ channel in the main analysis.

In order to isolate a clean sample of muons in the data we look for $Z \rightarrow \mu\mu$ events by requiring one 20 GeV muon to pass our standard tight muon cuts, shown in Table 4.3 and another track with $P_T >$ 20 GeV is required in the event. The event must also

pass the cosmic removal filter. The muon and the track must be back-to-back in ϕ, with ($\Delta\phi > \pi$-0.5) and the mass of the two tracks must be between 75 GeV and 105 GeV under the $\mu\mu$ final state hypothesis. Although the 20 GeV track requirement is much higher in P_T than the requirement on taus in the data, we see empirically in our analysis that events failing the muon veto have very high track P_T, typically well over 20 GeV. Therefore, the 20 GeV track cut still leaves events representative of the $\mu \to \tau$ background.

To compare to Monte Carlo, we use a PYTHIA sample. The only difference in the analysis here between our data study and this Monte Carlo study is that we use generator level information to ensure that the two muon candidates are indeed the Z daughter muons.

cut	# data opp sign ev	# data same sign ev	# mc opp sign ev
μ and track $P_t > 20$ GeV	9799	591	86054
back-to-back	8791	403	81773
75< mass < 105	7265	116	72023
match to tau cand	316	44	3511
pass tau ID	10	0	79

Table 7.4: Effect of cumulative cuts on the data and the Monte Carlo

To understand the number of non $Z \to \mu\mu$ events in our data sample, we separate events into opposite and same sign categories depending on whether the primary muon and the tau candidate have opposite or same sign. The result of the cut sequence in the analysis for events with a muon and a track in the opposite sign data, same sign data and Monte Carlo is shown in Table 7.4.

The same sign events can be used to estimate background, or non- $Z \to \mu\mu$, levels in the opposite sign category in the data. However, some background sources will exhibit a charge preference between a leading track and a muon. Therefore, we check the agreement between the data and Monte Carlo off the Z pole on the low side tail. In

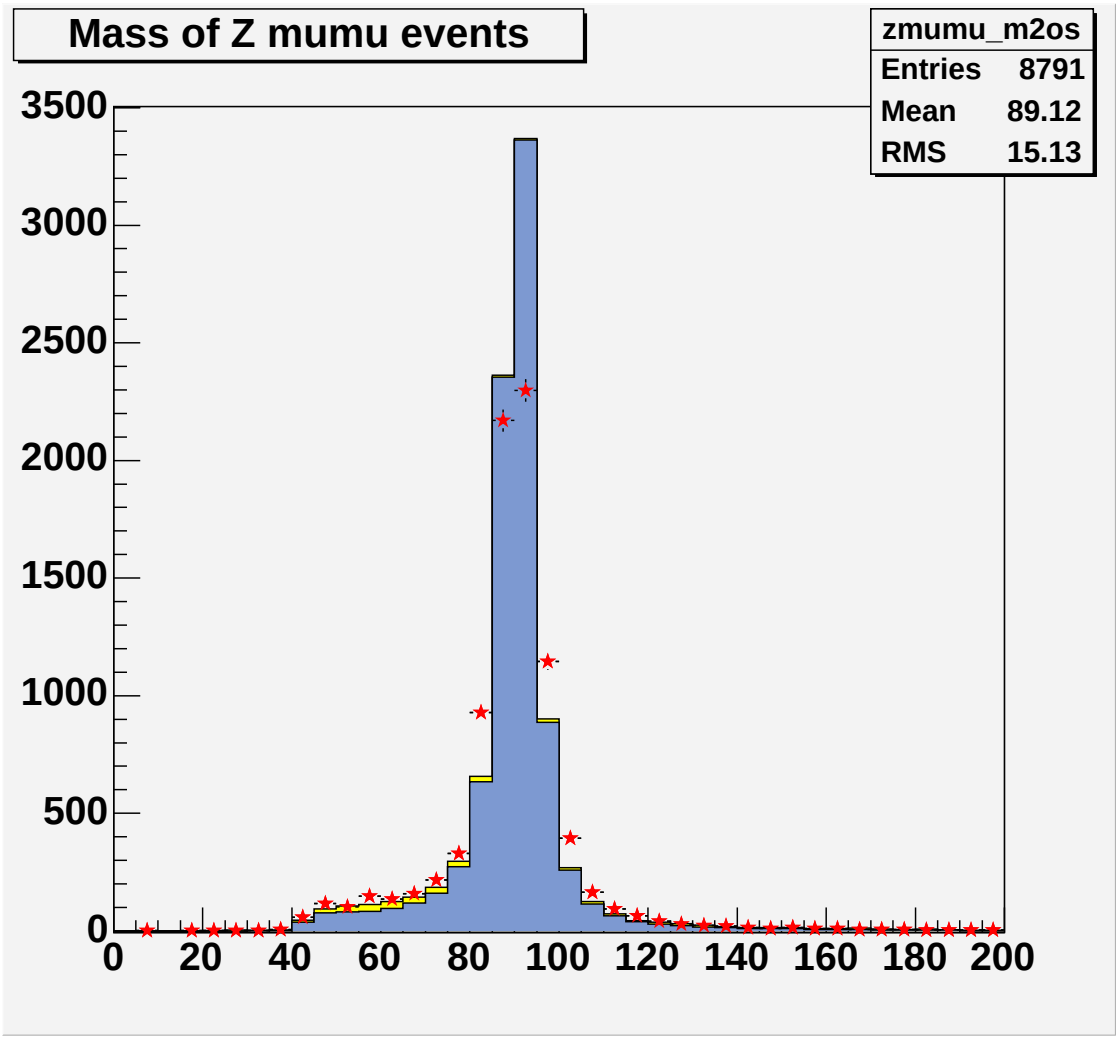

Figure 7.14: $M_{\mu\mu}$ (GeV/c^2) in data and the Monte Carlo. The opposite sign data events are shown as the plotted points while the Monte Carlo mass distribution is the blue histogram, normalized to the number of data events. The same sign data events are stacked on top of the Monte Carlo distribution in yellow, allowing us to check our background method in the low mass tail.

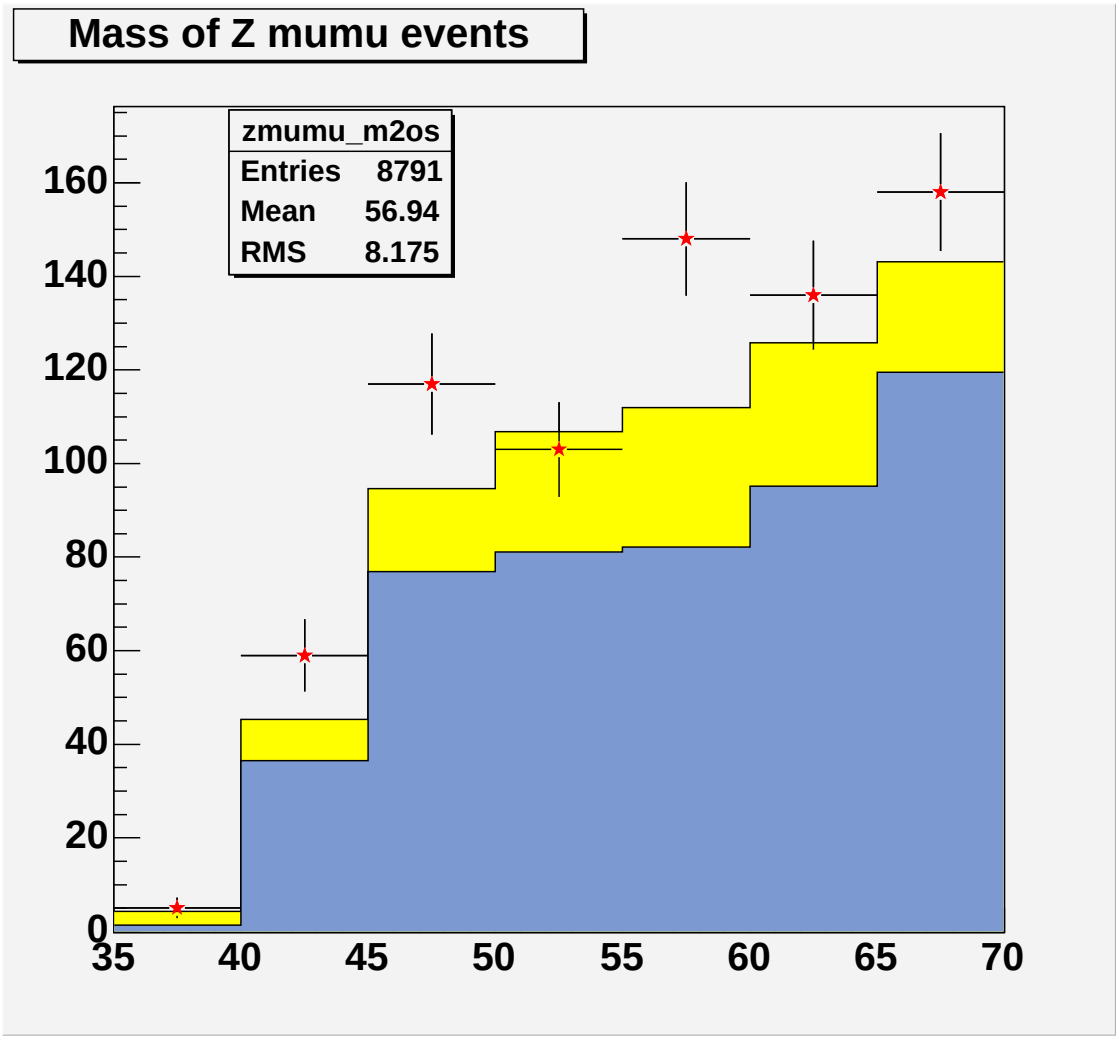

Figure 7.15: $M_{\mu\mu}$ (GeV/c^2) in data and the Monte Carlo below the Z pole. The opposite sign data events are the points while the Monte Carlo is the blue histogram, normalized to the number of opposite sign data events. The same sign data events have been stacked on top of the Monte Carlo events.

Figure 7.14 we plot the mass distribution of our 8791 opposite sign events in the data. On the same plot is our Monte Carlo mass distribution, normalized to our number of data events. Stacked on top of the Monte Carlo distribution is our same sign contribution from the data. If the same sign events are a measure of the background then we have a measure of how well we are determining our background by the comparison of the low mass tails between the opposite sign data and Monte Carlo + same sign data. Figure 7.15 zooms in on this region. Assuming we wish to determine a "scale factor" by which to multiply the same sign events to obtain estimates of opposite sign background, we see a slight tendency for the Monte Carlo plus same sign to under-predict the data. A conservative scale factor based on this plot would be 1.5$\pm$0.5.

After the mass cut we have 7265 opposite sign data events and 116 same sign data events, so the background contribution at that stage of the analysis, including the scale factor, is estimated to be in the range (2.4$\pm$0.8)%.

We compare the hadronic energy distributions for our opposite sign and same sign events in Figure 7.16. The average value of hadronic energy deposited by $Z \rightarrow \mu\mu$ signal event muons is 2.42 GeV while the average value for background events is 18.52 GeV. In order to compare the data with the Monte Carlo (where the fraction of same sign events is 0.1$\%$) we subtract the data same sign hadronic energy distribution from the data opposite sign hadronic energy distribution. The resulting distribution is compared with the Monte Carlo in Figure 7.17. The average amount of hadronic energy deposited in the background subtracted data is 2.2 GeV while the average in the Monte Carlo is 2.7 GeV, although much of the difference is in the very low energy range. These bins are not relevant for this analysis, because they are below the calorimeter threshold for tau candidates.

The high side tail in coarser bins is shown in Figure 7.18. We conclude that these distributions are consistent.

We can measure the $\mu \rightarrow \tau$ fake rate in the data and Monte Carlo. Numbers of events

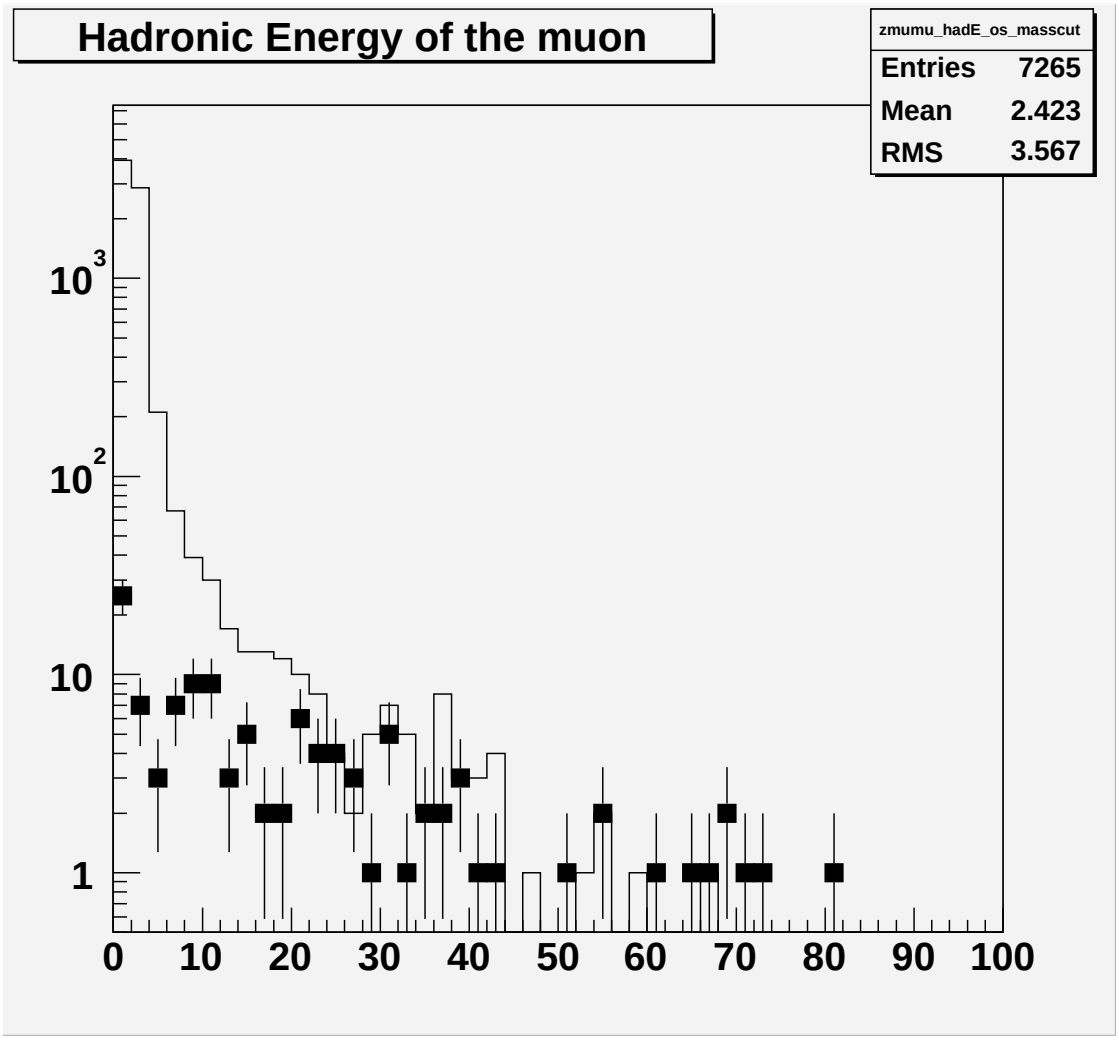

Figure 7.16: Hadronic energy deposition (GeV) for muon candidates in the data. The 7265 opposite sign events are shown as the histogram and the 116 same sign events are the black points.

	Fake rate	
	Data (Background subtracted)	Monte Carlo
$\mu \to \tau$ candidate	0.035±0.002(stat)±0.003(scale factor)	0.049±0.001(stat)
$\mu \to \tau$ (after ID cuts)	0.0014±0.0004(stat)±0.0001(scale factor)	0.0011±0.0001

Table 7.5: $\mu \to \tau$ Fake rates calculated from data and Monte Carlo $Z \to \mu\mu$ Samples. This data measurement is limited by low statistics and background uncertainty.

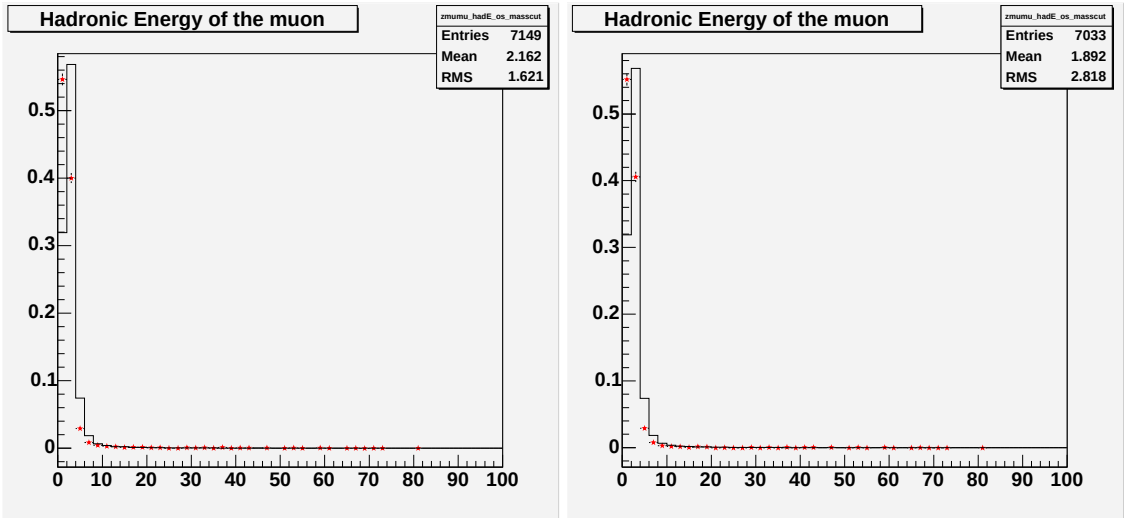

Figure 7.17: Hadronic energy for muon candidates in background subtracted data (points) and Monte Carlo (histogram), normalized to the data. The left plot shows the comparison with a background scale factor of 1, and the right plot with a scale factor of 2. A comment about the disagreement at low E_{had} is in the text.

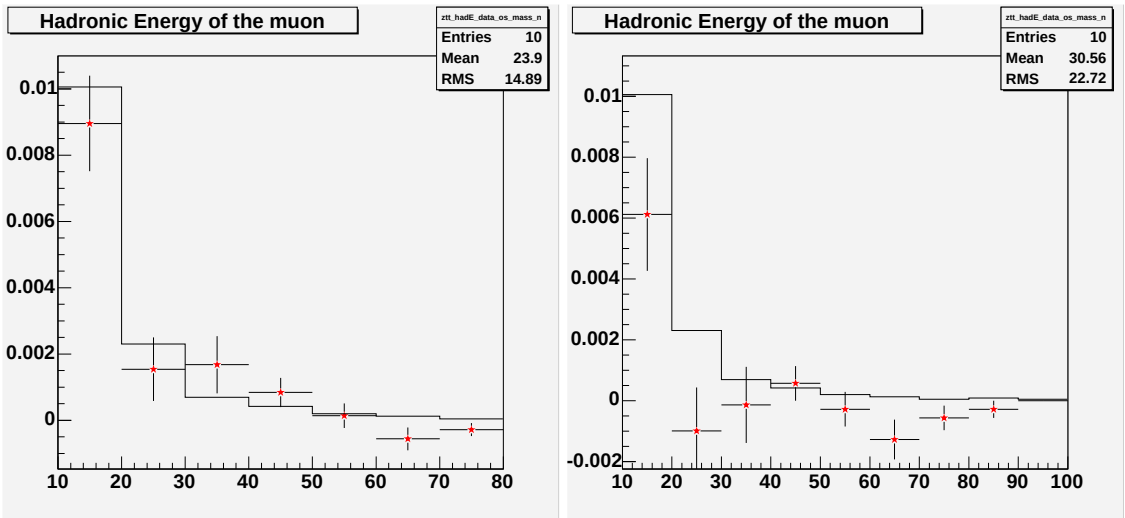

Figure 7.18: Hadronic energy (10–100 GeV) for muon candidates in background subtracted data (points) and Monte Carlo (histogram), normalized to the data. The left plot shows the comparison with a background scale factor of 1, and the right plot with a scale factor of 2.

in each category passing candidate and final τ event selections are shown in Table 7.4. After subtracting the background estimated from the same signs, we measure a muon to tau candidate fake rate of

$$0.035 \pm 0.002(\text{stat}) \pm 0.003(\text{scale factor})$$

in the data, to be compared with 0.049±0.001(stat) in the Monte Carlo. The fake rate after all cuts is estimated as

$$0.0014 \pm 0.0004(\text{stat}) \pm 0.0001(\text{scale factor})$$

from the limited data statistics, and is 0.0011±0.0001 in the Monte Carlo. These results are summarized in Table 7.5.

While there is a ∼30±10% reduction in the tau *candidate* fake rate from data relative to the Monte Carlo prediction, we see no evidence for an inconsistency between data and Monte Carlo after all tau identification cuts. This consistency test is far from conclusive because of the 30% statistical uncertainty from the data determination of the latter fake rate.

It is possible to isolate the muon veto component of the τ selection and to use this study to examine any possible difference between the background rejection of the muon veto in the data and Monte Carlo. Shown in Figure 7.19 is the muon veto variable E_T/P_T (cluster energy over seed track momentum) for data and background subtracted Monte Carlo. The efficiency of this cut for removing real muons that are tau candidates is measured to be $0.82 \pm 0.03(\text{stat}) \pm 0.05(\text{scale factor})$ in the data, and it is observed to be 0.907 ± 0.002 efficient in $Z \to \mu\mu$ Monte Carlo, so these fake rates are consistent. We can also form a fake rate in data and Monte Carlo by using the tau candidate fake rates given above. We conclude that the fraction of muons which are tau candidates passing this fake removal cut is $0.006 \pm 0.001(\text{stat}) \pm 0.002(\text{scale})$ in the data and 0.0046 ± 0.00015 in the simulation, which are consistent within the large uncertainties.

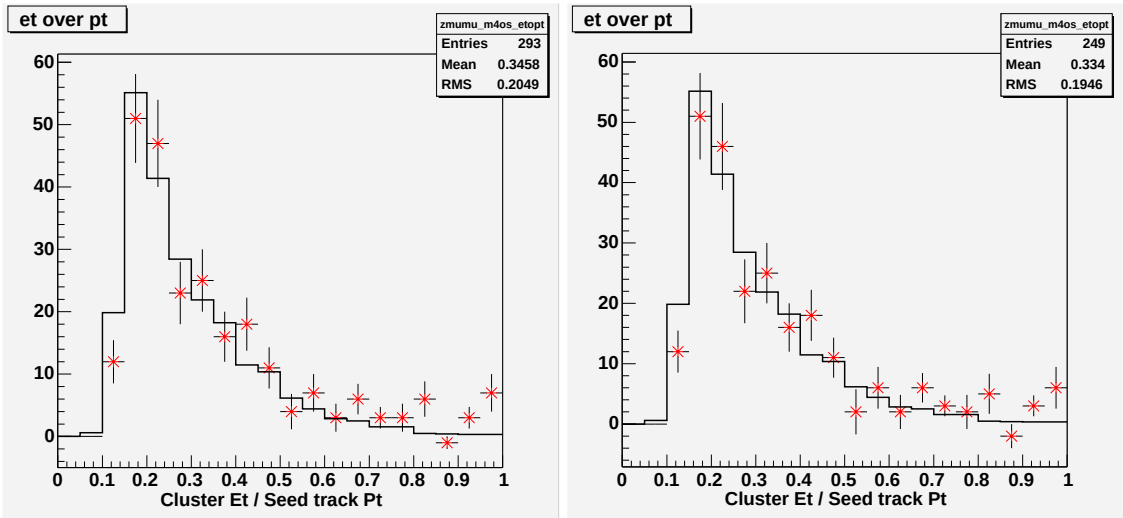

Figure 7.19: E_T/p_T (muon veto) for muons which are tau candidates in background subtracted data (points) and in Monte Carlo (histogram), normalized to the data. The left plot shows the comparison with a background scale factor of 1, and the right plot with a scale factor of 2.

Note that this study has a high overlap with the previous study of the muon to tau fake rate since it uses the same events in the data and since the muon veto is such a significant part of the background rejection.

Since we have some confidence in the Monte Carlo, and because it is difficult to isolate the few muons faking tau candidates in the data (and even if we could we would be plagued by low statistics and therefore high uncertainties) we calculate this background using $Z \to \mu\mu$ Monte Carlo.

Background due to $Z \to \mu\mu$ Events

To estimate the background for events with zero or one jet in the final state, we use 1000000 events from a PYTHIA Monte Carlo dataset. However there are not enough statistics in this dataset for us to determine the $\geq$2 jet bin contribution, after all cuts, to the total background. For this, we use 284946 events from the HERWIG + ALPGEN

(+ 2 parton) $Z \rightarrow \mu\mu$ Monte Carlo. The latter sample has a σX BR (cross section times branching ratio) of approximately 23 pb. After all cuts, we predict a background contribution of 0.05$\pm$0.03 events for 193.5 pb^{-1}.

7.2 $\gamma^+/Z \rightarrow \tau\tau + jets$

We use 464433 events from the $\gamma^*/Z \rightarrow \tau\tau$ PYTHIA Monte Carlo sample to determine the contribution from this background as a function of jets for the jet multiplicity study described in Chapter 8. The σX BR of this sample is 283.5 pb. Only one event from this sample passes all of our analysis cuts. Therefore, to estimate the background contribution after all cuts we use HERWIG + APLGEN (+2 parton) Monte Carlo.

There are three datasets covering three mass regions that need to be taken into account for the HERWIG + ALPGEN (+2 parton) Monte Carlo. Each has a different σX BR and its own contribution to the background from $\gamma^*/Z \rightarrow \tau\tau$:

- We used a sample of 286823 events from a dataset with events generated in the Z mass window, with masses between 75 GeV and 105 GeV. The σX BR is 23.3 pb.
- We used a sample of 274295 events from a dataset with generated masses between 10 GeV and 75 GeV. The σX BR is 50.2 pb.
- We used a sample of 166942 events from a dataset with generated masses between 105 GeV and 800 GeV. The σX BR is 0.631 pb.

The contribution due to each Monte Carlo sample is shown for both the e, τ channel and the μ, τ channel in Table 7.6, after all relevant Monte Carlo corrections have been applied as described in Table 5.8. The PYTHIA background based on the one event is listed for completeness but it is not used in the analysis. Note that our expected number of background events from $\gamma^+/Z \rightarrow \tau\tau$ is one fourth the size of our expected number of signal events. This is very different from this analysis in Run 1 where the background from this physics process was larger than the expected signal[27]. The mass veto cut described in Section 4.7 is responsible for the dramatic reduction of this background.

Generated Z Mass range	# of e, τ BG events	# of μ, τ BG events
75-105 GeV	0.09±0.04	0.06±0.03
10-75 GeV	0.00+0.03	0.00+0.03
105-800 GeV	0.06±0.01	0.05±0.01
Total (10-800 GeV)	0.15±0.05	0.11±0.04
PYTHIA ≥30 GeV	0.11±0.11	0.10±0.10

Table 7.6: Summary of backgrounds from $\gamma^*/Z \rightarrow \tau\tau + jets$. The PYTHIA prediction, which is not used, is listed for comparison.

7.3 WW

In WW events if one W decays to an electron or muon and the other decays to a tau the event can fake our signal region. This background is small due to the small σX BR of this process, particularly when two jets are required to be in the event.

We calculate our jet multiplicity contributions from WW with 195897 events from the HERWIG + ALPGEN (+0 parton) Monte Carlo sample. This sample has a σX BR of 8.282 pb. We calculate our background contribution with 828061 WW events from the HERWIG + ALPGEN (+1 parton) sample. This sample has a σX BR of 4.363 pb. We apply all relevant Monte Carlo corrections listed in Table 5.8. Our resulting background for the e, τ channel is 0.08 $\pm$ 0.01 events and for the μ, τ channel we see a background of 0.06 $\pm$ 0.01 events.

7.4 WZ

WZ events can be in our signal region if an electron or muon and a tau are decay products of the bosons. The σX BR of this process is four times smaller than the σX BR of WW, so we expect this background contribution to be very small.

We calculate our WZ background and jet multiplicity study contributions with 152500

events from a HERWIG + ALPGEN (+0 parton) Monte Carlo sample. This sample has a σX BR of 2.367 pb. After all analysis cuts we have eight events in both the e, τ and μ, τ channels. Our resulting backgrounds, after all relevant Monte Carlo scale factors have been applied, are 0.01 $\pm$ 0.01 events for both channels.

7.5 Background Summary

	number of events expected
$\gamma^*/Z \to \tau\tau + jets$	0.15 $\pm$0.05(stat) $\pm$0.03(sys)
$j \to \tau$ fakes	0.45 $\pm$0.10(stat) $\pm$0.12(sys)
$e \to \tau$ fakes	0.08 $\pm$0.03(stat) $\pm$0.02(sys)
WW	0.08 $\pm$0.01(stat) $\pm$0.02(sys)
WZ	0.01 $\pm$0.01(stat)
Total expected background events	0.77 $\pm$0.12(stat) $\pm$ 0.13(sys)
Signal exp from MC	0.59 $\pm$0.05(stat) $\pm$0.10(sys)

Table 7.7: e, τ channel signal and background predictions

	number of events expected
$\gamma^*/Z \to \tau\tau + jets$	0.11 $\pm$0.04(stat) $\pm$0.02(sys)
$j \to \tau$ fakes	0.30 $\pm$0.06(stat) $\pm$0.08(sys)
$Z \to \mu\mu$	0.05 $\pm$0.03(stat)
WW	0.06 $\pm$0.01(stat) $\pm$0.01(sys)
WZ	0.01 $\pm$0.01(stat)
Total expected background events	0.53 $\pm$0.08(stat) $\pm$0.08(sys)
Signal exp from MC	0.44 $\pm$0.04(stat) $\pm$0.07(sys)

Table 7.8: μ, τ channel signal and background predictions

We show a summary of our backgrounds in Tables 7.7 and 7.8.

Chapter 8

Results

We use a control sample, the low jet multiplicity sample, to test our method of calculating background contributions. We make predictions on the number of events we we expect to see in our control sample and compare our predictions with what we measure. When we are assured that our predictions are statistically consistent with our measurements in the control sample we have the confidence in our analysis methods to look at the signal region after all event cuts. With our measured number of events we are able to set a limit on the ratio of the rate of top decays to tau leptons that we see to the number that we expect based on the standard model. This ratio, r_τ, is our final result. The above procedure is described in detail in this chapter.

8.1 Control Sample: Low Jet Multiplicity

Before we measure the number of events after all analysis cuts we check our background calculations by comparing the number of events we expect to see in a non-signal region, with the number we measure from the data. Any large discrepancies we see are either statistical fluctuations or mistakes in our background or signal predictions. This check

allows us to change the analysis if we find problems without being biased by the final result. In order to keep our analysis blind in this way we do not look in the jet multiplicity region of interest where two or more jets are required. The resulting tables are referred to as jet multiplicity tables because we use the jet multiplicity to form our regions for comparison.

We separate the events into same-sign and opposite-sign categories and then further into the two channels, e, τ and μ, τ. Inside each of these four categories we make a prediction of our total expected background plus total expected signal as a function of jet multiplicity, after lepton identification cuts and the event $\not{E}_{\mathrm{T}}$ requirement. The predicted number of events is compared with the measured number in our 0 and 1 jet categories, allowing us to check our background and signal calculation methods without looking at the data in our true region of interest.

The results of these jet multiplicity comparisons for the zero and one jet bins are shown in Tables 8.1–8.4. Note that we have not made the the H_T cut or the Z mass veto cut before forming these tables.

sample	0 jets	1 jet	$\geq$2 jets
$jet \to \tau$ fakes	12.54±0.57±3.26	2.35±0.22±0.61	0.92±0.13 ±0.24
$e \to \tau$ fakes	0.92±0.10±0.23	0.22±0.05±0.05	0.10±0.03±0.03
$\gamma^*/Z \to \tau\tau$	8.15±1.07±0.86	1.79±0.50±0.19	0.89±0.18±0.09
WW	2.66±0.17±0.28	0.22±0.02±0.02	0.10±0.01±0.01
WZ	0.03±0.01	0.07±0.02	0.02±0.01
Signal ($t\bar{t}$)	0.03±0.01	0.12±0.02±0.01	0.70±0.05±0.07
Total Expected	24.3±1.2±3.4	4.8±0.6±0.6	2.7±0.2±0.3
Data	17	5	**blind**

Table 8.1: Jet multiplicity table for e-tau channel, opposite sign events.

sample	0 jets	1 jet	$\geq$2 jets
$jet \to \tau$ fakes	11.70±0.58±3.04	1.29±0.17±0.34	0.53±0.08±0.14
$\gamma^*/Z \to \mu\mu$	3.40±0.36	0.22±0.09	0.08±0.06
$\gamma^*/Z \to \tau\tau$	4.49±0.82±0.44	0.95±0.38±0.10	0.95±0.38±0.10
WW	2.03±0.15±0.21	0.19±0.02±0.02	0.10±0.01±0.01
WZ	0.05±0.01	0.03±0.01	0.02±0.01
Signal ($t\bar{t}$)	0.01±0.01	0.10±0.02±0.01	0.52±0.05±0.05
Total Expected	21.7±1.1±3.1	2.8±0.4±0.4	2.2±0.4±0.2
Data	11	4	**blind**

Table 8.2: Jet multiplicity table for mu-tau channel, opposite sign events.

sample	0 jets	1 jet	$\geq$2 jets
$jet \to \tau$ fakes	6.86±0.39±1.78	1.87±0.19±0.49	1.05±0.13±0.27
$e \to \tau$ fakes	0.02±0.01	$0^{+0.01}_{-0}$	$0^{+0.01}_{-0}$
$\gamma^*/Z \to \tau\tau$	0.30±0.20	0.0+0.11	0.02±0.02
WW	0.06±0.02	0.005±0.003	0.005±0.003
WZ	0.05±0.01	0.03±0.01	0.01±0.01
Signal ($t\bar{t}$)	$0^{+0.003}_{-0}$	0.003±0.003	$0^{+0.003}_{-0}$
Total Expected	7.3±0.4±1.8	1.9±0.2±0.6	1.1±0.1±0.3
Data	8	3	**blind**

Table 8.3: Jet multiplicity table for e-tau channel, same sign events.

sample	0 jets	1 jet	$\geq$2 jets
$jet \to \tau$ fakes	5.34± 0.36±1.39	0.78±0.15±0.20	0.51±0.14±0.13
$\gamma^*/Z \to \mu\mu$	0.08±0.08	$0^{+0.04}_{-0}$	0.10±0.04
$\gamma^*/Z \to \tau\tau$	0.10±0.10	0	0.01±0.01
WW	0.02±0.01	0.004±0.002	0.003±0.002
WZ	0.05±0.01	0.03±0.01	0.01±0.01
Signal ($t\bar{t}$)	0	0	0.002±0.002
Total Expected	5.6±0.4±1.4	0.8±0.2±0.2	0.6±0.2±0.1
Data	3	0	**blind**

Table 8.4: Jet multiplicity table for mu-tau channel, same sign events.

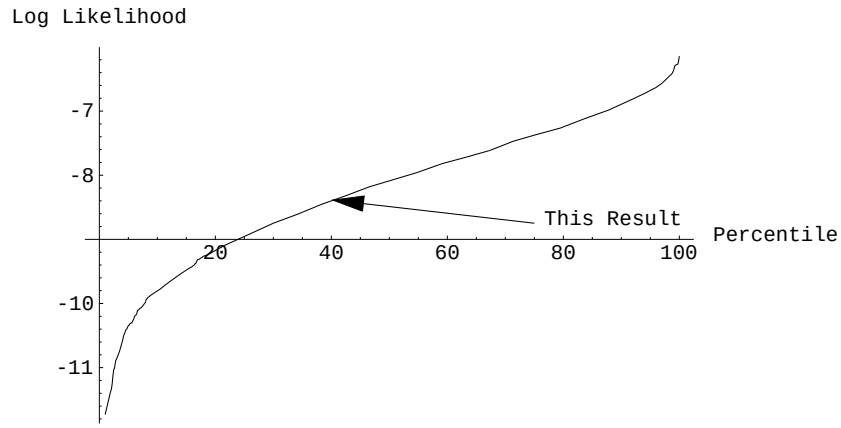

Figure 8.1: The *a priori* test of the data probability of the observed low jet multiplicities compared against the predictions shown with the distribution of the same quantity in pseudo experiments. The pseudo experiments have a lower probability 40% of the time.

To check consistency with the predictions, a joint probability of these low jet multiplicity observations, given the predicted rates, was formed. We account for the uncertainties in the predicted rates by comparing our result against generated pseudo experiments which chose "true" predictions from the measured predictions and their errors, and then generated Poisson fluctuations about those "true" means. The main *a priori* test that we chose was to look at all eight of the values for opposite and same sign, zero and one jet multiplicity, electron and muon data samples, and this resulted in a probability which was higher than in 40% of generated pseudo experiments. The distribution of these probabilities by percentile and the identified value in the data are shown in Figure 8.1. We expect this probability to be lower because it is performed *a posteriori* on the most unlikely bins.

We also performed a number of similar tests on subsets of the low jet multiplicity data in order to test for consistency with a variety of pathologies. Table 8.5 summarizes these tests and their results. The least probable agreement is in the zero jet, opposite sign muon case, where the probability is in the 5th percentile.

8.2 The Unblinded Data Samples

Having passed the *a priori* criterion for agreement, we then examine the signal region and same sign region for higher jet multiplicities. With no H_T or Z mass cut, the results are shown in Table 8.6. Electron same and opposite sign rates appear higher than predicted, whereas muon rates are consistent. However, upon applying the final event selection cuts, including the H_T and Z mass requirement, the results in Table 8.7 are obtained where no visible disagreement is evident. We show the Run1 predicted and measured events in Table 8.8. In total, we have two candidate signal events (both $e\tau$ events) and one same sign event which would otherwise pass signal criteria ($e\tau$). These events are:

- Run = 167299, Event = 2376337, Njets = 3, H_T = 286 GeV, MET = 59.4 GeV (opp sign eτ candidate). See Table 8.9 and Figure 8.2. One jet is b-tagged.
- Run = 151434, Event = 158200, Njets = 2, H_T = 239 GeV, MET = 71.7 GeV (opp sign eτ candidate). See Table 8.10. No jets are b-tagged.

The thirteen events in the $\geq$2 jet bins of the Njet tables were checked for b-tagged jets but none were found except in run 167299, event 2376337. This is consistent with the hypothesis that these bins are dominated by background processes that do not contain b quarks.

8.3 $\mathbf{r}_\tau$

This result provides no evidence for the the $e\tau$ or $\mu\tau$ plus jets final state in $t\bar{t}$-like events. Because of its relatively poor acceptance and high background, it is clear that the tau analysis can contribute little to a cross-section derived under the limit of lepton universality. Where this result can contribute, however, and what motivated this analysis in the

Subsamples included	Probability Percentile
All electron bins	62
All muon bins	20
All opposite sign bins	16
All same sign bins	78
Electron and muon zero jet opposite sign	9
Only muon zero jet opposite sign	5

Table 8.5: Consistency tests applied to the low jet multiplicity data

Sample	OS e	SS e	OS μ	SS μ
Background	2.03±0.2±0.3	1.1±0.1±0.3	1.7±0.4±0.2	0.6±0.2±0.2
$t\bar{t}$	0.70±0.05±0.07	<0.003	0.52±0.05±0.05	0 .002±0.002
Data	8	4	1	0

Table 8.6: The ≥ 2 jet multiplicity data where the $t\bar{t}$ signal is expected compared to predictions. No H_T or Z veto cut has yet been app lied.

Sample	OS e	OS μ
Background	0.77±0.12±0.13	0.53±0.0.08±0.08
$t\bar{t}$	0.59±0.05±0.10	0.47±0.04±0.07
Data	2	0

Table 8.7: The signal region, including the H_T, Z mass, opposite charge and ≥ 2 jet multiplicity requirements

Background	2.50±0.43
$t\bar{t}$	1.1±0.4
Data	4

Table 8.8: We show here the results of the Run1 tau dilepton analysis[27]. The expected number of events shown above was calculated with the Run1 CDF measured value of the $t\bar{t}$ cross section of $7.7^{+1.8}_{-1.5}$pb. In this analysis we use the theoretical value of the cross section. The corresponding value for Run1 conditions is 4.8 ±0.7pb [39].

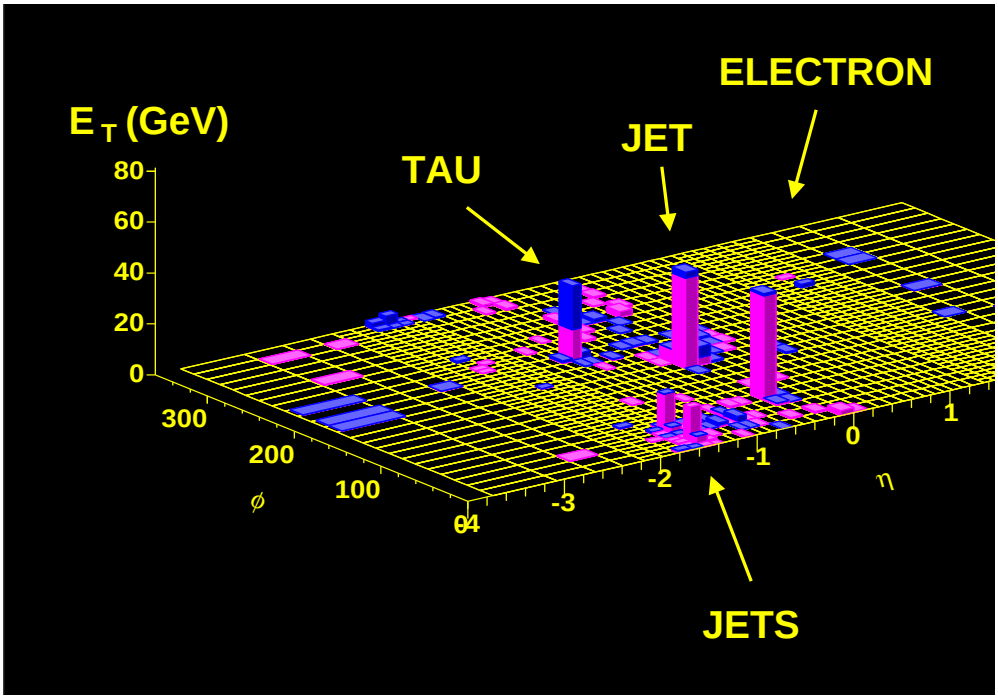

Figure 8.2: Calorimeter lego plot for run 167299, event 2376337 which is an eτ candidate.

first place, was the search for potentially anomalous contributions that could show up in the final state as an enhanced (or suppressed) rate for τ leptons in top decay.

To cast this analysis in this light, we choose to measure the parameter, r_τ where

$$r_\tau \equiv \frac{\mathrm{BR}(t \to b\tau\nu)}{\mathrm{BR}_{\mathrm{SM}}(t \to b\tau\nu)}.$$

One practical observable from which to derive r_τ is the ratio of measured rates for the $e\tau$+$\mu\tau$ dileptons to the ee, $e\mu$ and $\mu\mu$ dileptons. This observable ratio has several advantages experimentally, including largely common systematic uncertainties on acceptance. However, there is a problem with this technique in that a significant fraction (approximately 15% [40] under the assumption that r_τ is unity) of the acceptance in the $\{e, \mu\}$ dilepton acceptance comes from tau leptons. Therefore, in the limit of very large r_τ, the ratio of the two rates becomes insensitive to r_τ, and, in fact, because the likelihood as a function of r_τ approaches a small but non-zero constant, the integral probability over a flat prior r_τ distribution is infinite.

Therefore, we choose instead to determine this variable by comparison to the standard model predicted rate. The probability distribution for the observable r_τ, given this measurement, is calculated by numerically integrating over hidden true variables representing the true $\sigma_{t\bar{t}}$ constrained by the uncertainties of the NLO calculation [31], the standard model branching ratios, the number of $e\tau + \mu\tau$ dileptons predicted given $\sigma\times$BR, and the backgrounds to this analysis, all of which are constrained by experimental measurement or derivation.

The unnormalized probability distribution given this measurement as a function of r_τ is shown in Figure 8.3. The most probable value in this distribution is at $r_\tau \approx 0.8$.

This probability distribution can be used to set limits in the Bayesian approach by assuming a flat prior in r_τ. The resulting “one sigma” symmetric 68% confidence level

Object	E_T (GeV)	Pseudorapidity	Phi (rad/deg)
electron	39.8	-0.28	1.1(61)
tau	38.6	-0.95	3.8(215)
jet	73.3	-0.27	2.9(166)
jet	39.5	-1.48	0.2(12)
jet	35.4	-1.40	0.8(45)

Table 8.9: Details for run 167299, event 2376337.

Object	E_T (GeV)	Pseudorapidity	Phi (rad/deg)
electron	78.9	-0.99	1.7(98)
tau	20.0	-0.60	3.8(218)
jet	34.9	-0.92	1.2(71)
jet	33.6	0.30	5.0(286)

Table 8.10: Details for run 151434, event 158200.

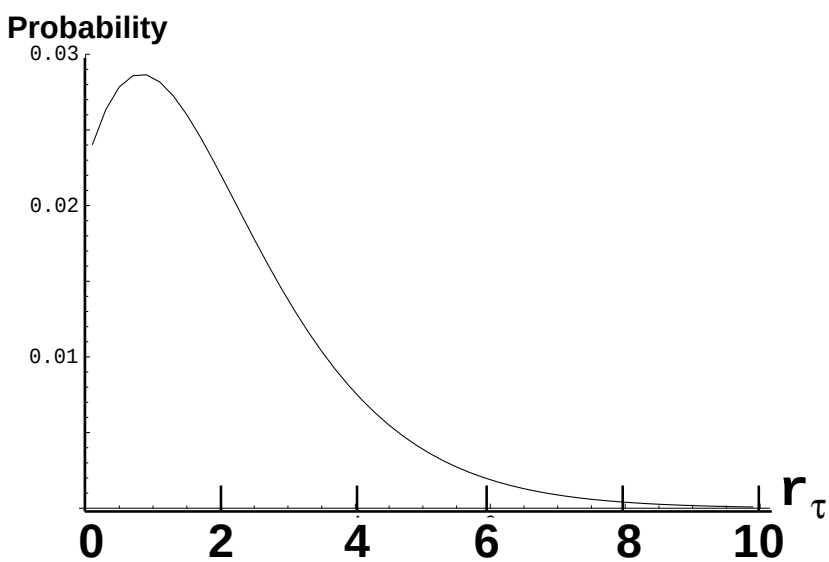

Figure 8.3: The unnormalized probability distribution $P(N_{\rm obs} = 2 \| r_\tau)$, where r_τ is the universality parameter

range is

$$0.6 < r_\tau < 3.6 \text{ at } 68\% \text{ confidence.}$$

The 95% lower limit on r_τ is driven more by the requirement that $r_\tau > 0$ and the assumption of a flat prior than by the measurement as Figure 8.3 illustrates. We set therefore an upper limit,

$$r_\tau < 5.0 \text{ at } 95\% \text{ confidence,}$$

which we consider the main result of this analysis. Clearly this measurement is consistent with the the lepton universality prediction of $r_\tau = 1$.

Chapter 9

Conclusions

Our value of r_τ is consistent with the standard model. However, the uncertainty on our measurement is large enough so that there is still room for non-standard model physics in this decay chain. In order to confirm that the standard model is correct in its prediction of the rate of top decays to tau leptons we need more data and improved analysis techniques.

This test of the model, combined with the hundreds of tests being performed at CDF and other high energy physics detectors around the world, allows us to develop a more complete picture of the fundamental forces and fundamental particles of nature. As we examine every prediction of the standard model and test every proposed testable extension we come closer to the new physics that we know must exist to explain the generation of mass, neutrino oscillations, a quantum theory of gravity, or even perhaps the puzzle of dark matter or dark energy.

This new physics could be in the form of an extension to the standard model, it could be a correction, or it could even replace the standard model. Regardless of its fate, the standard model is an excellent framework within which we have been able to probe fundamental properties. Our current favored model, with its stunning successes

paired with its gaping holes, captures for us what an exciting and uncertain time this is in particle physics.

Bibliography

[1] P.B. Renton, Rept. Prog. Phys. 65, 1271-1330 (2002).

[2] S. Eidelman et al., Phys. Lett, B 592, 1 (2004).

[3] `http://www.particleadventure.org`

[4] K2K Collaboration: M.H. Ahn, et al., Phys. Rev. Lett. 90, 041801 (2003).

[5] V. Barger et al., Phys. Lett. B537, 179-186 (2002).

[6] S. Bose, A. Raychaudhuri, J. Phys. G29 1069-1074 (2003).

[7] F.J. Gilman, Rev. Mod. Phys. 56, S296 (1984).

[8] D. Griffiths, “Introduction to Elementary Particles,” John Wiley and Sons, Inc., 1987.

[9] F. Abe et al., Phys. Rev. Lett. 74, 2626 (1995).

[10] S. Abachi et al., Phys. Rev. Lett. 74, 2632 (1995).

[11] S.W. Herb et al., Phys. Rev. Lett. 39, 252 (1977).

[12] S. Willenbrock, “The Standard Model and the Top Quark,” hep-ph/0211067 (2002).

[13] F. Abe et al., Phys. Rev. Lett. 79, 3585 (1997).

[14] C. Yue, H. Zong, L. Liu, Mod. Phys. Lett. A18, 2187-2194 (2003).

[15] T. Han, M.B. Magro, Phys. Lett. B 476, 79-86 (2000).

[16] V. Barger, R.J.N. Phillips, Phys. Rev. D 41, 884 (1990).

[17] T. Affolder et al., Phys. Rev. D 62, 012004 (2000).

[18] `http://www.fnal.gov/pub/inquiring/physics /accelerators/chainaccel.html`

[19] "The CDF II Dectecor Technical Design Report," FERMILAB-Pub-96/390-E, 1996.

[20] D. Broemmelsiek, R. Pasquinelli, "Fermilab Recycler Stochastic Cooling Commissioning and Performance," FERMILAB-Conf-03/122, (2003).

[21] D. Perkins, "Introduction to High Energy Physics, 4th Edition," Cambridge University Press, 2000.

[22] S. Rappoccio et al., "First Look at SecVtx Using Event Primary Vertex Finder PrimeVtx", `CDF/PUB/SEC_VTX/PUBLIC/6417`, (2003).

[23] `www-cdf.fnal.gov/internal/upgrades/daq_trig/ twg/tools/trigopts.ps`

[24] M. Coca, E. Halkiadakis, "Central Electron Identification Efficiencies for the 200 pb^{-1} Run 2 Dataset," `CDF/DOC/ELECTRON/CDFR/6580`, (2004).

[25] J.D. Jackson, "Classical Electrodynamics," John Wiley and Sons, Inc., (1998).

[26] J-F Arguin, et al., "Generic Jet Energy Corrections for RunII data used for the Winter Conferences", `CDF/ANAL/JET/CDFR/6280`, (2003).

[27] M. Hohlmann, Ph.D. thesis, University of Chicago, (1997).

[28] A. Connolly, Ph.D. thesis, University of California, Berkeley, (2003).

[29] T. Sjostrand, Comput. Phys. Commun. 82, 74 (1994).

[30] S. Jadach, $et\ al.$, TAUOLA 2.5, CERN-TH-6793 (1992).

[31] M. Cacciari et al., hep-ph/0303085. N. Kidonakis and R. Vogt, Phys. Rev. D68 114014 (2003).

[32] `http://hep.physics.utoronto.ca/JeanFrancoisArguin /JetCorrDoc/SystUncert.html`.

[33] `http://cteq.org`

[34] S. Demers, J. Insler, K. McFarland, A. Vaiculis, "Measurement of $t\bar{t}$ in the $e-\tau_{had}$ $\mu-\tau_{had}$ Dilepton Channels," CDF/ANAL/TOP/CDFR/6921, (2004).

[35] J. Pumplin et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", hep-ph/0201195 v3.

[36] P. Murat, "$W \rightarrow \tau\nu$ Signal in Run II Data and Preliminary Measurement of Sigma*BR($W \rightarrow \tau\nu$)/Sigma*BR($W \rightarrow e\nu$)," `CDF/PUB/ELECTROWEAK/PUBLIC/6010`, (2002).

[37] Van Neerven et al., Nucl. Phys. B359, 343, (1991).

[38] A. Anastassov et al., "Tau Reconstruction Efficiency and QCD Fake Rate for Run 2", CDF Note 6308.

[39] S. Catani et al., “The top cross section in hadronic collisions”, Phys. Lett. B 378, 329 (1996).

[40] CDF Collaboration, “Measurement of the ttbar Cross-Section in Dilepton Events”, CDF Note 6882.