Precision Measurement of the Top Quark Mass in the Lepton + Jets Channel Using a Matrix Element Method with Quasi–Monte Carlo Integration

by

Paul Joseph Lujan

B.A. (Harvard University) 2000

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:
Professor James Siegrist, Co-Chair
Dr. Angela Galtieri, Co-Chair
Professor Robert Jacobsen
Professor Eugene Haller

Fall 2009

The dissertation of Paul Joseph Lujan is approved:

Co-Chair	Date
Co-Chair	Date
	Date
	Date

University of California, Berkeley
Fall 2009

Precision Measurement of the Top Quark Mass in the Lepton + Jets Channel
Using a Matrix Element Method with Quasi–Monte Carlo Integration

© 2009

by Paul Joseph Lujan

Abstract

Precision Measurement of the Top Quark Mass in the Lepton + Jets Channel
Using a Matrix Element Method with Quasi–Monte Carlo Integration

by

Paul Joseph Lujan

Doctor of Philosophy in Physics

University of California, Berkeley

Professor James Siegrist, Co-Chair

Dr. Angela Galtieri, Co-Chair

This thesis presents a measurement of the top quark mass obtained from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron using the CDF II detector. The measurement uses a matrix element integration method to calculate a $t\bar{t}$ likelihood, employing a Quasi–Monte Carlo integration, which enables us to take into account effects due to finite detector angular resolution and quark mass effects. We calculate a $t\bar{t}$ likelihood as a 2-D function of the top pole mass m_t and Δ_{JES}, where Δ_{JES} parameterizes the uncertainty in our knowledge of the jet energy scale; it is a shift applied to all jet energies in units of the jet-dependent systematic error. By introducing Δ_{JES} into the likelihood, we can use the information contained in W boson decays to constrain Δ_{JES} and reduce error due to this uncertainty. We use a neural network discriminant to identify events likely to be background, and apply a cut on the peak value of individual event likelihoods to reduce the effect of badly reconstructed events. This measurement uses a total of 4.3 fb^{-1} of integrated luminosity, requiring events with a lepton, large $\not\!\!E_T$, and exactly four high-energy jets in the pseudorapidity range $|\eta| < 2.0$, of which at least one must be tagged as coming from a b quark. In total, we observe 738 events before and 630 events after

applying the likelihood cut, and measure $m_t = 172.6 \pm 0.9$ (stat.) $\pm$ 0.7 (JES) $\pm$ 1.1 (syst.) GeV/c^2, or $m_t = 172.6 \pm 1.6$ (tot.) GeV/c^2.

Co-Chair	Date
Co-Chair	Date

To my parents

Contents

List of Figures

List of Tables

Acknowledgements

There are many, many people who have helped me in the work that has led to this thesis, and I could not possibly hope to cover everyone, but the following people deserve special mention:

First, of course, is my advisor, Lina Galtieri. Lina is not only an excellent physicist, but an excellent mentor; she has always treated her students like family, and helped this analysis along with infinite patience through countless hurdles and difficulties. It has truly been an honor to work for her, and I am immensely grateful for the opportunity.

The rest of the LBL top mass group also deserves recognition, as they have all contributed quite a bit of work towards the results presented here, so thanks to Igor Volobouev, Jeremy Lys, and Jason Nielsen, as well as John Freeman, who, while no longer in the group, was my partner in crime through many of the trickiest parts of the analysis (and also deserves thanks for providing me with the LaTeX template for this thesis, which he in turn got from Henri Bachacou).

Many people at CDF have helped out this analysis with helpful suggestions and contributions; special thanks go to the conveners of the CDF Top Mass subgroup, Mousumi Datta, Fabrizio Margaroli, and George Velev; the conveners of the CDF Top Group, Florencia Canelli and Veronica Sorin; and the internal reviewers for our analysis, Jahred Adelman, Petteri Mehtälä, and Nathan Goldschmidt, as well as past conveners and reviewers. Our analysis is much better for their sharp eyes.

Naturally, big thanks go to the rest of my thesis committee, Jim Siegrist, Bob Jacobsen, and Eugene Haller, for reading this thesis despite an extremely tight schedule; I greatly appreciate their contribution of time and effort.

At the beginning of my CDF career, I had the opportunity to work on the CDF Run IIb silicon upgrade project with Marc Weber, who was an excellent supervisor, and I enjoyed my time there immensely. I also had the privilege of working on the b-tag scale factor with Weiming Yao, which gave me a solid start in doing physics

analyses.

Thanks also go to my fellow CDF students and officemates, Amanda Deisher and Hung-Chung Fang, who were always a source of sanity and camaraderie; Anne Takizawa, without whom the Berkeley physics graduate department simply could not function; and Marjorie Shapiro, who, with Lina, has been such a key part of the LBL CDF group.

Finally, very special thanks go to my parents Judy and Larry Lujan, to whom I owe so much, and to my beloved wife Juliana Froggatt, who is simply wonderful in every way.

Chapter 1

Introduction

Since the '70s, the Standard Model of particle physics has proven to be an extremely successful description of nearly all known particle physics measurements. The Standard Model, which contains the two families of fundmental particles, the quarks and leptons, and three fundamental forces and the gauge bosons which carry those forces, has seen a series of experimental successes, most notably the prediction of a third generation of quarks in the bottom quark (discovered in 1977 [1]) and its heavier partner, the top quark (discovered in 1995 [2]), as well as the prediction of the W and Z gauge bosons, discovered in 1983 [3].

However, there are still several questions that the Standard Model has not yet answered; one of the most important of these is the origin of mass. The Standard Model predicts that the mass of particles arises from interactions with a scalar field, the Higgs field, and the corresponding gauge boson, the Higgs boson. Unfortunately, despite being the subject of intense search, the Higgs boson has not yet been experimentally observed. Discovery of the Higgs boson would represent another success for the Standard Model and answer one of the most important remaining questions, while failure to discover a Standard Model Higgs would open up a new area of potential physics beyond the Standard Model. Consequently, the Higgs search is of great importance to the field of particle physics at present.

Measurement of the top quark mass provides a very useful tool to help us in this Higgs search. Aside from being intrinsically important — as the Standard Model does not contain any *a priori* predictions of particle masses, the only way to determine them is through direct measurement — measurement of the top quark mass also allows us to constrain the Higgs boson mass, giving us a better idea of how to best conduct this search.

The Fermilab Tevatron is currently the highest-energy particle accelerator in the world, and will remain so until LHC, the Large Hadron Collider, begins physics operations (currently expected to be in late 2009). As the Tevatron is the only collider with enough energy to produce top quarks and (potentially) Higgs bosons,

it is naturally home to a large amount of activity in both of these areas. While it was initially expected that the Higgs boson would not be discovered until the LHC began operations, improved performance of the Tevatron coupled with advances in analysis techniques have increased the possibility that the Higgs boson could be observed at the Tevatron before the LHC.

This thesis presents a measurement of the top quark mass performed at CDF, one of the two large, general-purpose detectors (along with D0) located at the Tevatron. We first employed the technique presented here to measure the top quark mass in 1.9 fb^{-1} of data, which we later updated to 2.7 fb^{-1} and then to 3.2 fb^{-1} [4]. This paper further updates the measurement to include 4.3 fb^{-1} of integrated luminosity, totalling more than seven years' worth of data from the Tevatron, and includes a total of 738 events selected as $t\bar{t}$ candidates from the data.

The particular technique used in this analysis is known as a matrix element integration, in which the matrix element $|M|^2$ for $t\bar{t}$ production and decay is integrated over unmeasured quantities to obtain a likelihood of observing the events seen in the detector as a function of the top pole mass m_t. The matrix element method has become increasingly popular as a technique to measure the top mass, as the computational power available to perform this integration has increased; first used to measure the top quark mass in a D0 Run I analysis [5], it has now seen widespread adoption at both CDF and D0 [6] for top mass analyses in a variety of channels. Matrix element techniques have generally proven to be the most precise method available to measure the top mass; some advantages of the matrix element method are that events are naturally weighted by the amount of information that they contain, as the likelihood will tend to have sharper peaks for better-measured events, and the ability of the matrix element integration, through the inclusion of transfer functions, to model non-Gaussian detector responses.

Because a $t\bar{t}$ event includes a large number of unknowns, integrating over the

full phase space of these unknown variables is not computationally practical. One can reduce the number of unknowns, and hence the computation time required to evaluate the likelihood, by making a variety of kinematic assumptions; for instance, an earlier version of this analysis [7] integrated over a total of seven unknown variables by making assumptions about the masses and angles of the quarks in the event. While this makes the calculation more computationally tractable, the disadvantage of this procedure is that these assumptions are, of course, not always satisfied, resulting in an imperfect modeling of the actual physics and hence a worse result. In this analysis, we employ, for the first time in a top mass analysis, an improved integration technique, Quasi–Monte Carlo integration, which allows us to discard many of the assumptions we used previously and integrate over 19 dimensions without significantly increasing the computation time required to evaluate the likelihood, resulting in an improved mass resolution.

The remainder of the thesis proceeds as follows: Chapter 2 gives a brief overview of the relevant physics in the Standard Model for top quark production and decay. In Chapter 3, the Fermilab Tevatron and the CDF detector are described. Chapter 4 describes the procedure by which we select $t\bar{t}$ candidate events from the millions of events per second produced at the Tevatron. Chapter 5 describes the matrix element procedure by which the likelihood is calculated for signal events, while Chapter 6 describes how our method handles background events and other events not well-modeled by our analysis. Chapter 7 shows how we test and calibrate our method using Monte Carlo simulated events. The actual measured value in data is presented in Chapter 8, while Chapter 9 discusses the systematic uncertainties in our measurement. Finally is a brief conclusion in Chapter 10.

Chapter 2

Top Physics

In this chapter, I discuss the physics underlying the measurement of the top quark mass described in this thesis. First is a brief review of the Standard Model, followed by a discussion of the top quark properties in the Standard Model; finally, the relation of the top quark measurement to the Higgs is presented.

2.1 The Standard Model

The Standard Model contains our current understanding of nearly all of particle physics; it relates all observed particles and interactions (with the exception of gravity). Included are 12 spin-1/2 fermions, which are the constituents of matter, and four spin-1 bosons, which carry three fundamental forces: the electromagnetic force, the weak force, and the strong force, with the first two unified into the electroweak force. In the Standard Model, particle interactions are described by the gauge group $SU(3) \times SU(2) \times U(1)$; in addition, the Standard Model posits the existence of the Higgs boson, whose interactions with other particles is responsible for the origin of their mass. While the Standard Model has proven enormously successful, there are several issues that it does not address; for instance, gravity is not included, and the problem of how to create a unified theory containing gravity is the subject of much active theoretical work.

The basic particles making up the Standard Model are shown in Figure 2.1. Fermions are further classified into two types, the quarks and leptons. The quarks carry color charge and thus interact via the strong force (as well as the other two forces); there are a total of six quarks (and six corresponding antiquarks) divided into three generations of increasing mass. Each generation contains one quark with charge $(+2/3)e$ and one quark with charge $(-1/3)e$, which are often referred to as "up-type" and "down-type", respectively. The leptons do not interact via the strong force, and there are also six leptons divided into three generations (with their corresponding antileptons); each generation contains one charged lepton with charge $-e$ and a corresponding uncharged neutrino.

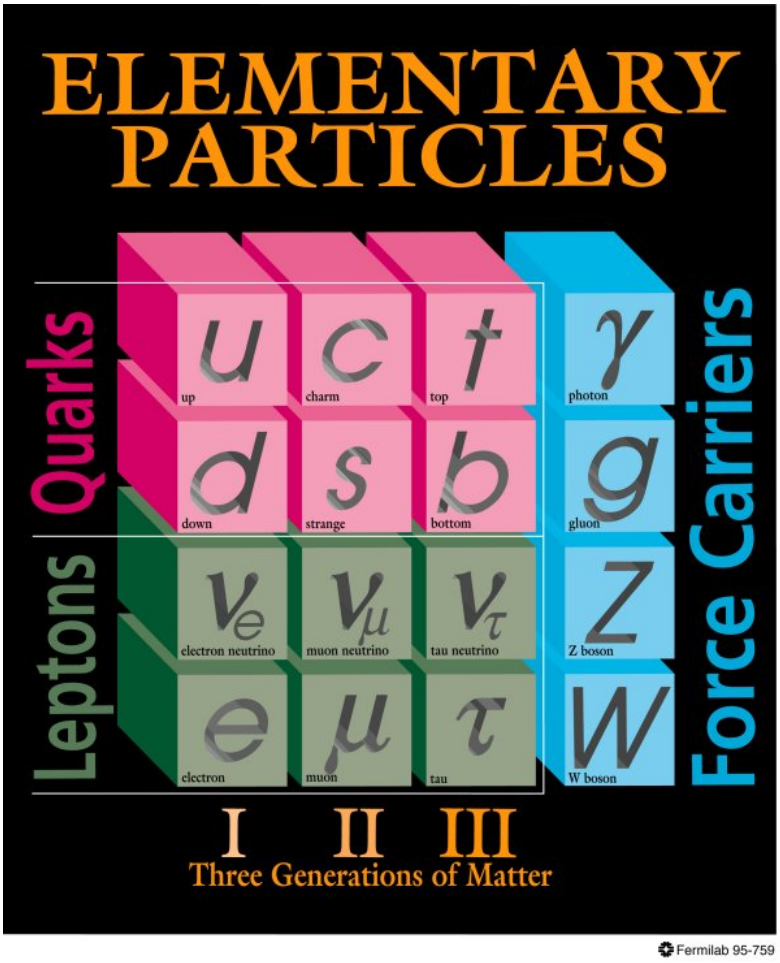

Figure 2.1: The fundamental particles included in the Standard Model: the six quarks, six leptons, and four bosons.

The masses of the twelve leptons are summarized in Table 2.1 [8]. Note that since the Standard Model does not predict any of the particle masses, the only way to determine them is by direct measurement. It was long believed that the neutrinos had zero mass, but in 1998 evidence from the Super-Kamiokande detector conclusively indicated the non-zero mass of the neutrino [9]. (As this measurement only provides information on the difference in neutrino masses, however, the values listed in Table 2.1 are constraints obtained from direct measurements; the actual values are generally assumed to be much smaller.) Note the dramatic increase in quark mass, especially in the top quark, which is substantially more massive than the next-lightest quark; this was much greater than initially expected and resulted in the discovery of the top quark taking place much later than the other quarks. The mass of individual quarks is dependent on the scheme used to define the mass; Table 2.1 quotes the values obtained in the modified minimal subtraction ($\overline{\text{MS}}$) scheme [10], with the exception of the top quark, for which the pole mass is quoted. How the measured mass can be corrected to obtain the value in the $\overline{\text{MS}}$ scheme is currently the subject of active debate [11].

2.1.1 The Electroweak Force

The electroweak force is described by a $SU(2) \times U(1)$ symmetry group which is spontaneously broken, resulting in the separate electromagnetic and weak forces. The electromagnetic force is carried by the photon (γ), a massless spin-1 particle, which couples to particles carrying electric charge. Because of the masslessness of the photon, the electromagnetic force can act over long ranges. The weak force is carried by the $W^{\pm}$ and Z bosons, which are both quite massive (80.40 ± 0.03 GeV/c^2 and 91.188 ± 0.002 GeV/c^2, respectively), which results in the weak force only acting over very short ranges. However, the weak force is unique among the forces in the Standard Model in that it is the only interaction which can change the flavor of quarks. Thus, decays of heavier quarks into lighter quarks can only

Table 2.1: Measured masses for the six quarks and six leptons in the Standard Model. Uncertainties are not included. For quarks other than the top, the masses are determined in the $\overline{\text{MS}}$ scheme.

	1st generation	2nd generation	3rd generation
Quarks	u	c	t
	1.5 – 3.3 MeV/c^2	1.27 GeV/c^2	173.1 GeV/c^2
	d	s	b
	3.5 – 6.0 MeV/c^2	105 MeV/c^2	4.20 GeV/c^2
Leptons	e	μ	τ
	511 keV/c^2	105.7 MeV/c^2	1.78 GeV/c^2
	ν_e	ν_μ	ν_τ
	< 2 eV/c^2	< 2 eV/c^2	< 2 eV/c^2

take place via the weak interaction; this fact often results in clear experimental signatures for this type of event.

The amplitude of the mixing between quark flavors mediated by the W boson is described by the CKM matrix, named after its inventors Cabibbo, Kobayashi, and Maskawa:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.1}$$

where V_{ij} is the factor attached to a vertex containing a W boson, an up-type quark i, and a down-type quark j. The terms in the CKM matrix are complex, and the matrix is required to be unitary in the Standard Model; traditionally, the matrix is parameterized by three mixing angles and a CP-violating phase. Using fits which apply the unitarity constraint to direct measurements, the current values for the magnitudes of the entries [8] are:

$$\begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & (3.59 \pm 0.16) \times 10^{-3} \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & (4.15 \pm 0.01) \times 10^{-2} \\ (8.74 \pm 0.32) \times 10^{-3} & (4.07 \pm 0.01) \times 10^{-2} & 0.999133 \pm 0.000044 \end{pmatrix} \tag{2.2}$$

2.1.2 The Strong Force

In the Standard Model, the strong force is described by a $SU(3)$ non-Abelian gauge theory, known as Quantum Chromodynamics (QCD). The strong force acts on particles which carry color charge (quarks, but not leptons); an individual quark can carry one of three different colors (red, green, or blue). The strong force is carried by massless spin-1 gluons, which carry a color and an anticolor charge[1];

[1]In total, one would thus expect there to be a total of nine gluons. However, the gluon corresponding to the color singlet of the $SU(3)$ group does not exist, so only eight gluons in total

because of this, the gluons can interact with each other as well as with quarks.

The strong force exhibits two unique properties which make it behave unlike any other fundamental force. The first is asymptotic freedom, in which interactions become negligibly weak at large momenta or short distances. An important consequence of this fact is that it is possible to analyze QCD processes (e.g., scattering cross sections) by treating partons (i.e., quarks and gluons) as free particles independent from others in the event.

The second is color confinement: quarks cannot exist singly, but must always be bound into color-neutral hadrons, either as three quarks in a baryon, or one quark and one antiquark in a meson. Attempting to separate the quarks in a hadron will simply result in more quarks being created from the vacuum, resulting in the production of more hadrons. This has a very significant experimental implication: while we can speak of a given process producing a quark, that quark will not be directly observed in our detector; rather, it will hadronize into a number of collimated hadrons, known as a "jet". Measuring the momentum or energy of these jets, as we will see in Section 3.4.2, is a very difficult task.

2.1.3 The Higgs Mechanism

One might expect that the masses of fermions and the W and Z bosons would simply appear as mass terms in the Standard Model Lagrangian. However, there is no way to add the mass terms directly in a gauge-invariant way, and breaking the gauge invariance would cause non-renormalizable divergences in the theory. Consequently, a more complicated mechanism to introduce masses into the Standard Model is required; this mechanism is known as the Higgs mechanism [12].

To illustrate the basic principle behind the Higgs mechanism, we consider a somewhat simplified example [13]. Consider a complex scalar field ϕ with a quartic self-coupling term in the potential and a negative "mass" term: $\mathscr{L} = |\partial_\mu \phi|^2 +$

exist.

$\mu^2\phi^*\phi - \frac{1}{2}\lambda(\phi^*\phi)^2$. Now, the ground state of the field is no longer $\phi = 0$; rather, the ϕ field has a vacuum expectation value $|\phi_0| = \sqrt{\mu^2/\lambda}$, causing the symmetry to be spontaneously broken. If we now add an electromagnetic field to this Lagrangian and expand the Lagrangian around the new minimum, a positive mass term for the photon appears. Figure 2.2 illustrates the potential in this simplified case.

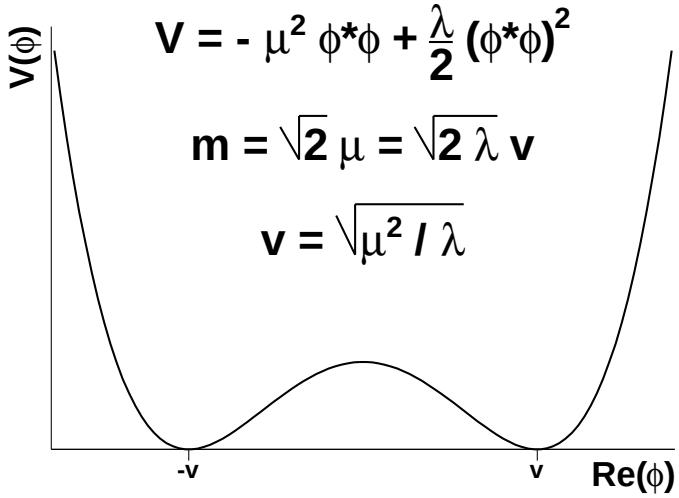

Figure 2.2: The potential for a complex scalar field ϕ, illustrating the basic principle of the Higgs mechanism.

Applying the Higgs mechanism to the full electroweak theory (since we don't actually want a massive photon) is a little more complex, as the Higgs field is now a $SU(2)$ doublet of complex fields, but the principle is still the same. Applying the Higgs mechanism results in the spontaneous symmetry breaking of the $SU(2) \times U(1)$ electroweak symmetry, causing the W and Z bosons to acquire mass (while the photon remains massless). A single spin-0 Higgs boson remains, with mass given by $m_H = \sqrt{2\lambda}v$, where λ is the quartic self-coupling of the Higgs field and v is the vacuum expectation value. This Higgs boson also couples to fermions in proportion to their mass, thus also accounting for fermion masses.

Because of its importance to the success of the Standard Model, the discovery of the Higgs boson is a very high priority in the particle physics community, but

so far no searches have successfully discovered the Higgs. Searches at LEP have ruled out $m_H < 114$ GeV/c^2, while recent searches at the Tevatron have ruled out $160 < m_H < 170$ GeV/c^2 [14]. Figure 2.3 shows the limits on the Higgs mass obtained from direct searches.

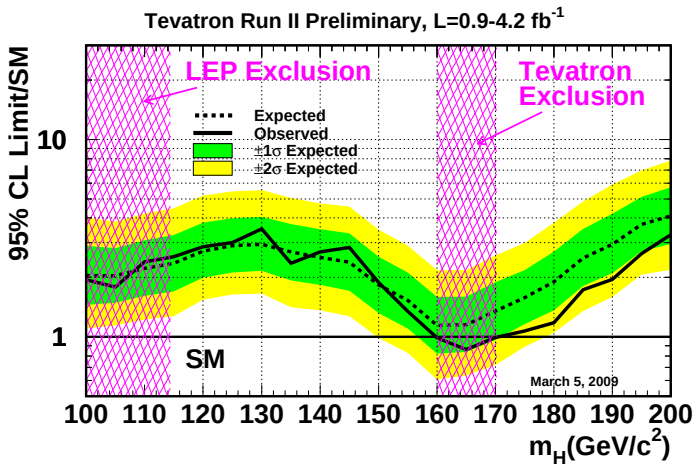

Figure 2.3: Current limits on the Higgs mass set by direct searches at LEP and the Tevatron.

2.2 Top Quark Phenomenology

The top quark is the heaviest known fundamental particle, with a measured mass, as of March 2009, of 173.1 ± 1.3 GeV/c^2 [15]. The extremely high mass of the top quark means that it also decays extremely quickly (approximately 0.5×10^{-24} s), despite the decay being mediated by the weak force. This rapid decay means that, unlike lighter quarks, the top quark does not hadronize before decaying but rather decays as a free quark. Consequently, this means that the properties of the top quark (mass, spin, charge, etc.) can be measured directly through reconstruction of its decay products.

Figure 2.4 shows the measurements contributing to the current world average for the top mass, which represents the most precise measurement so far. The previous version of this analysis with 3.2 fb^{-1} of integrated luminosity is the "CDF-II l+j" entry.

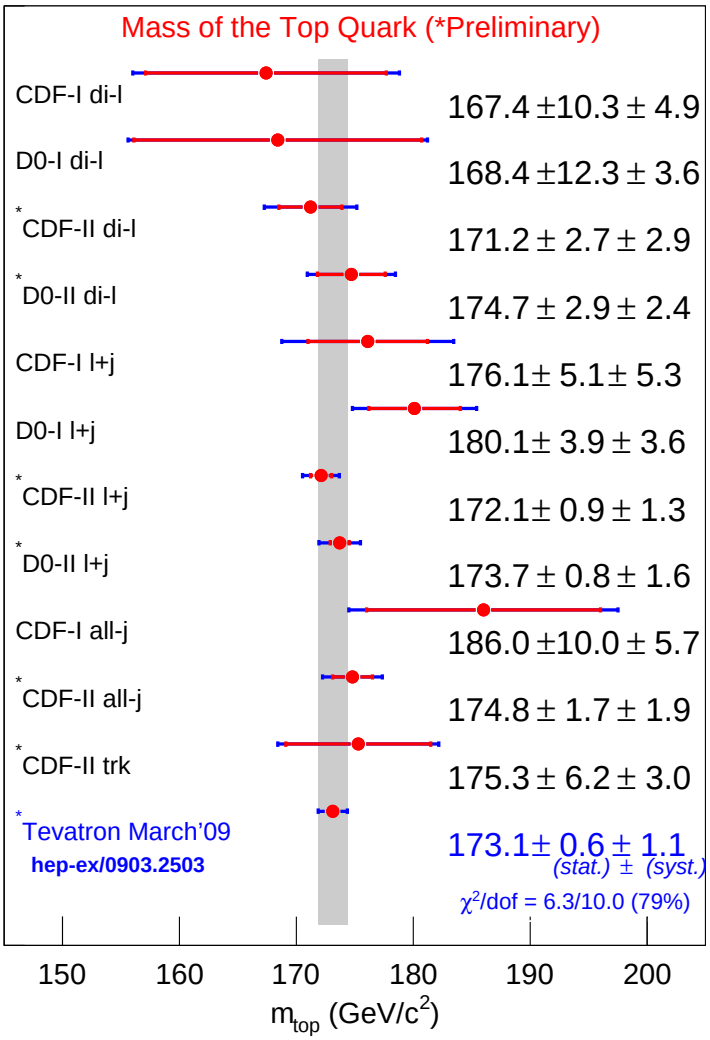

Figure 2.4: Individual top mass measurements contributing to the current world average. The version of this measurement with 3.2 fb^{-1} of integrated luminosity is the "CDF-II l+j" entry in this chart.

2.2.1 Top Quark Production

At the $\sqrt{s} = 1.96$ TeV center-of-mass energy of the Tevatron, the top quark is expected to be produced primarily in $t\bar{t}$ pairs; this production can proceed either by $q\bar{q}$ annihilation or gg fusion, as depicted in Figure 2.5. At the Tevatron energies, the next-to-leading order (NLO) prediction is that $15\% \pm 5\%$ of $t\bar{t}$ production will come from gg fusion. (Note that at the 14 TeV center-of-mass energies at the LHC, this ratios will be approximately reversed, as gluons are much more likely to carry sufficient energy to produce a $t\bar{t}$ pair.)

Measurement of the top quark pair production cross-section is an active area of research at the Tevatron; the latest CDF measurement [16] yields $\sigma_{t\bar{t}} = 7.02 \pm 0.63$ pb (assuming $m_t = 175$ GeV/c^2, as there is a slight variation of cross-section with mass).

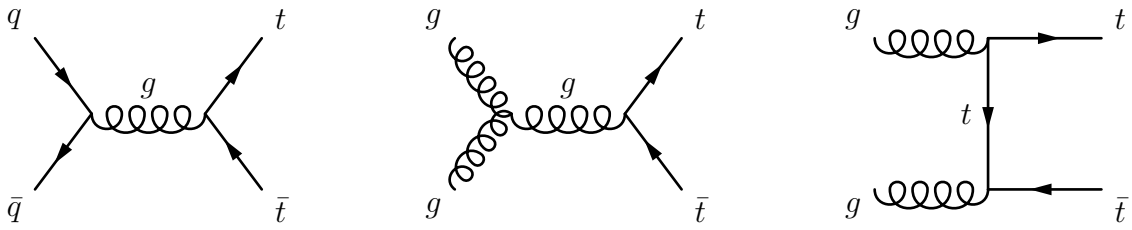

Figure 2.5: Feynman diagrams for $t\bar{t}$ production via $q\bar{q}$ annihilation (left) and gg fusion (center and right).

2.2.2 Top Quark Decay

The top quark decays weakly into a W boson and a down-type quark. As we saw in Equation 2.2, $|V_{tb}|$ is nearly unity[2]. Since the ratio $(t \to Wb)/(t \to Wq) = |V_{tb}|^2/(|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2)$, we can see that the quark produced will be a b quark

[2]Note that this measurement requires the assumption of the CKM matrix being unitary. Direct measurements of $|V_{tb}|$ not relying on this asssumption can be made, but this yields a much weaker constraint of $|V_{tb}| > 0.74$ [8].

nearly 100% of the time. Consequently, when measuring the top mass, we can assume that a $t\bar{t}$ pair will always decay into $W^+bW^-\bar{b}$; events can then be futher classified based upon the decay products of the W, as discussed in Section 4.1.

In addition to the top mass, many other measurements can be performed on the top decay results, including its width and lifetime, charge, and the properties of the Wtb coupling. Searches for new physics can also be performed by looking for decays not allowed or strongly suppressed in the SM such as $t \to \gamma c$.

2.3 Top Quark Mass and the Higgs

Measuring the top quark mass is not only important in itself, but also because it helps us constrain the Higgs mass. Specifically, radiative corrections to the W boson mass from Higgs and top loops allow us to relate the top, W, and Higgs masses. (Other quarks could also contribute, in principle, but as the top quark is much heavier than the other quarks, its effect is by far the dominant one.)

Figure 2.6 shows the relevant Feynman diagrams for these corrections; specifically, $\Delta m_W \sim m_t^2$ and $\Delta m_W \sim \ln m_H$ [17], so high-precision measurements of both the W boson and top quark masses are necessary to obtain good constraints on the Higgs mass.

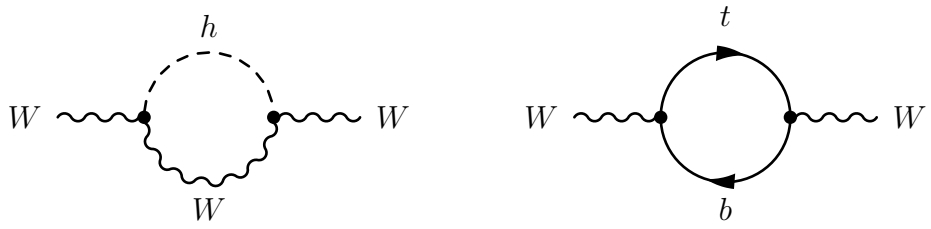

Figure 2.6: Feynman diagrams for the radiative corrections to the W boson mass arising from Higgs (left) and top (right) loops.

Figure 2.7 shows the current results of this fit as of March 2009 [18]. The left plot shows the contours in m_W and m_t space obtained from LEP2 and Tevatron

data. The range of Higgs masses is shown in the green band, with the upper edge of this band corresponding to the 114 GeV/c^2 limit set by direct searches at LEP. The right plot shows the $\Delta\chi^2$ of the fit as a function of m_H.

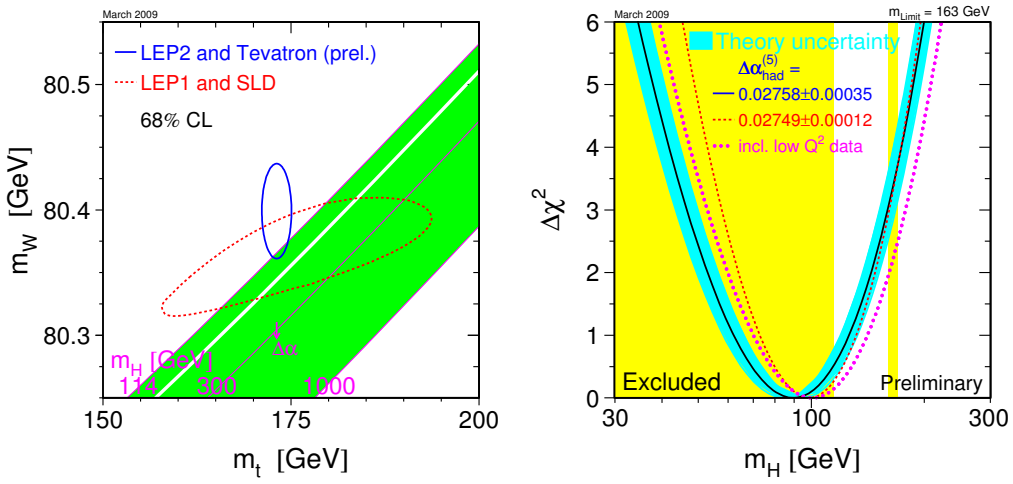

Figure 2.7: Results of the electroweak constraints on the Higgs mass as of March 2009. Left: Allowed contours in the m_W and m_t space. Right: Quality of fit as a function of m_H.

As we can see from these plots, the most-favored Higgs mass actually lies inside the excluded region, with only a small overlap between the region allowed by direct search results and the region favored by the electroweak fit. This has, naturally, fueled speculation of further physics beyond the Standard Model, such as the Minimal Supersymmetric Standard Model (MSSM), which features a two-Higgs doublet resulting in five final Higgs particles with somewhat different properties. In any case, the Higgs boson remains one of the most active areas of research in particle physics today, and measurement of the top quark mass is an important contribution to this effort.

Chapter 3

Experimental Setup

3.1 The Tevatron Accelerator

The Tevatron accelerator, located at Fermilab, is currently the only particle accelerator in the world with sufficient energy to produce top quarks (until the LHC begins physics operations). The Tevatron consists of the main Tevatron ring, with two large, general-purpose detectors, CDF and D0, located at interaction points on the ring, and a complex of accelerators which feeds into the main ring. The Tevatron began operations in 1985 with a center-of-mass energy of $\sqrt{s} = 1.8$ TeV, and has been regularly upgraded since then. Run II of the Tevatron, featuring substantial upgrades to the accelerator and both detectors, began operation in 2001 at a center-of-mass energy of $\sqrt{s} = 1.96$ GeV.

3.1.1 Initial Acceleration

Before reaching the Tevatron, protons go through a multi-stage chain of accelerators; the full complex is depicted in Figure 3.1. First, hydrogen gas is split into atoms and ionized to H^- ions. These ions are accelerated first in a Cockcroft-Walton accelerator to 750 keV and then further in a 130 m linear accelerator (the Linac) up to 400 MeV.

From the Linac, the electrons are then stripped from the H^- ions and the resulting protons are injected into the Booster. The Booster is a synchrotron 475 m in circumference which accelerates the protons from an energy of 400 MeV up to 8 GeV. Finally, from the Booster, the protons are sent to the next accelerator in the cycle, the Main Injector, a synchrotron 3 km in circumference which accelerates protons up to 150 GeV for injection into the Tevatron. The Main Injector also produces protons up to 120 GeV for use in antiproton production as described below or for fixed-target experiments.

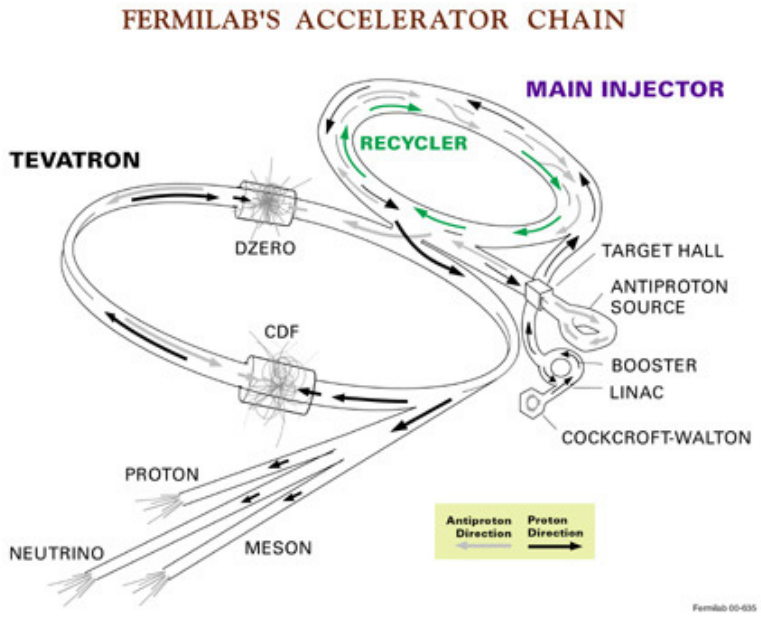

Figure 3.1: The chain of accelerators making up the Tevatron complex at Fermilab.

3.1.2 Antiproton Production

Antiprotons are produced by directing a beam of 120 GeV protons from the Main Injector onto a nickel target. The efficiency of antiproton production is approximately $2 \cdot 10^{-5}$ antiprotons produced for each proton incident on the target. The antiprotons are focused with a magnetized lithium lens and then separated from other collision products by passing the beam through a magnet which acts as a spectrometer. The antiprotons are then sent to the Debuncher, which uses stochastic cooling to reduce the variation in momentum of the antiprotons. After being cooled in the Debuncher, the antiprotons are then stored in the Accumulator, another synchrotron designed to store the antiproton beam for a period of many hours.

The newest addition to the antiproton production chain is the Recycler, a synchrotron sharing the tunnel with the Main Injector. Because the Recycler is much larger than the Accumulator, it can be used to store antiprotons and keep the Accumulator relatively empty; this improves the overall antiproton production

performance, since the "stacking rate" (i.e. the rate at which antiprotons can be stored) decreases with the number of antiprotons already stored. The Recycler has also implemented electron cooling, in which electrons are injected into antiproton bunches to reduce the size of the bunch, which allows for a more focused beam.

3.1.3 Tevatron Operation

The Tevatron is a synchrotron 1 km in radius which receives protons and antiprotons injected from the Main Injector at 150 GeV and accelerates them to a beam energy of 980 GeV. The protons and antiprotons are bunched into 36 bunches each, separated into 3 groups of 12 bunches. Within a group, each set of bunches is separated by 396 ns, with longer intervals between groups. Large superconducting dipole magnets operating at 4 K are used to steer the beams, and quadrupole magnets are located near the two interaction regions at B0 and D0 to reduce the transverse beam size before collision.

In normal operation, protons and antiprotons are injected into the Tevatron and circulate for up to a day in a "store", until beam losses and $p\bar{p}$ collisions have reduced the interaction rate sufficiently that the beams are dumped, at which point preparations for a new store begin. While a store is circulating in the Tevatron, the antiproton systems build up a new stack of antiprotons. If a store terminates abnormally (usually due to a "quench", when a superconducting magnet loses superconductivity, forcing the beam to be dumped), the Tevatron often must sit idle while the antiproton stack is built up again.

The Tevatron performance is measured by its luminosity, which measures the rate of particle collisions. The instantaneous luminosity is given by:

$$L_{\text{inst}} = \frac{f N_p N_{\bar{p}} N_b}{4\pi\sigma_x\sigma_y}, \tag{3.1}$$

where f is the frequency of bunch collisions, N_p and $N_{\bar{p}}$ are the numbers of protons and antiprotons per bunch, N_b is the number of bunches, and $4\pi\sigma_x\sigma_y$ represents

the transverse area of the beam, assuming a Gaussian distribution of particles with widths σ_x and σ_y in the x and y directions.

The integrated luminosity L is simply the integral of the instantaneous luminosity over a given time period; for a physical process with cross-section σ, the number of events expected to be produced is equal to $L\sigma$.

The performance of the Tevatron has improved dramatically since the beginning of Run II; Figure 3.2 shows the instantaneous and integrated luminosities at CDF as a function of time. The current record (recorded on 4/1/09) for instantaneous luminosity is $3.47 \cdot 10^{32}$ cm^{-2} s^{-1}. As of July 2009, more than 5 fb^{-1} of luminosity has been delivered and stored to tape at CDF.

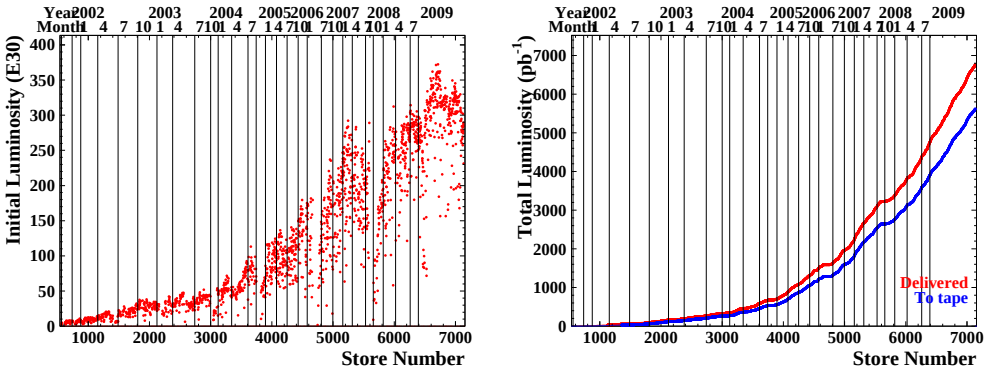

Figure 3.2: Instantaneous luminosity at the beginning of stores (left) and integrated luminosity (right) observed at CDF over the Run II period for the Tevatron.

3.2 The CDF Detector

The CDF detector is a general-purpose detector designed to measure the momenta and energy of objects produced in $p\bar{p}$ collisions. The detector is cylindrically symmetric about the beamline (with a few exceptions) and extends for several meters through the B0 interaction region.

Broadly speaking, the detector contains three main components. The innermost part of the detector is the tracking system, which contains a set of silicon strips and an open-cell drift chamber inside of a solenoid to measure the momenta of charged particles. Outside of the solenoid are the calorimeters, which measure the energies of electrons, photons, and hadrons. Finally, outside of the calorimeters are the muon chambers, which detect and measure the momenta of muons. These are described in more detail in the following sections. Figures 3.3 and 3.4 show the general layout of the CDF detector. (There are also some subsystems which are not used in $t\bar{t}$ analyses and thus not described here.)

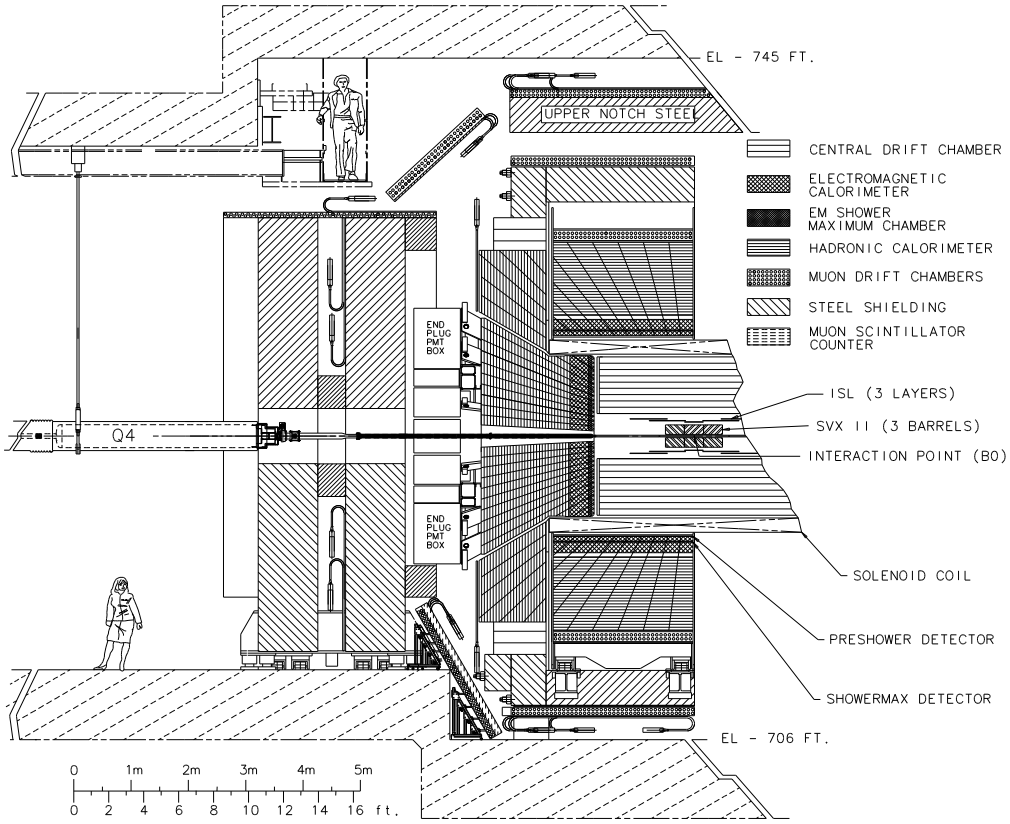

Figure 3.3: A side view of the CDF detector, with major detector subsystems indicated.

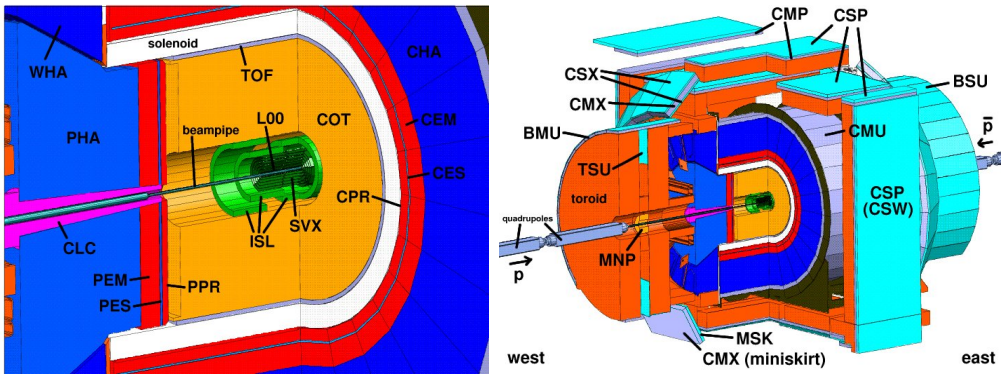

Figure 3.4: A cutaway view of the CDF inner detector (left) and whole detector (right), with the detector subsystems labeled. Note that not all of these systems are used for study of $t\bar{t}$ events.

3.2.1 CDF Coordinate System

CDF generally uses a combination of cylindrical and spherical coordinates to describe locations and directions, as follows:

- The z axis lies along the beamline, with the $+z$ direction defined as the direction of proton travel (east) and the $-z$ direction the direction of $\bar{p}$ travel (west). "Longitudinal" refers to components along the z axis, and "transverse" refers to components perpendicular to the z axis.
- x and y are not commonly used; the $+x$ direction is north, while the $+y$ direction is up (away from the ground).
- r is the radial distance from the beamline ($r = \sqrt{x^2 + y^2}$).
- θ is the polar angle from the beamline; $\tan\theta = r/z$.
- ϕ is the azimuthal angle around the beamline; $\tan\phi = y/x$.
- η is the pseudorapidity, defined as $-\ln(\tan(\theta/2))$.

Typically, the position of objects (such as detector components) is described with r, z, and ϕ coordinates. The direction of particles is usually measured in (η, ϕ)

space. The reason for the latter choice comes from the fact that the distribution of particles is expected to be roughly uniform in the rapidity Y, where

$$Y = \frac{1}{2} \log \frac{E + p_z}{E - p_z}$$

for a particle with energy E and z-momentum p_z. Since this is dependent on the particle mass, it is more convenient to use the purely geometric quantity η, especially since the mass of particles measured at CDF is generally negligible compared to their momentum, making Y and η equivalent.

For a particle with momentum p and energy E, we define the transverse momentum p_T and the transverse energy E_T as $p \sin\theta$ and $E \sin\theta$, respectively; the two-dimensional vector transverse momentum is simply the x and y components of the momentum vector: $\vec{p_T} = (p_x, p_y)$. The quantity ΔR is often used to measure distances in η–ϕ space; it is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.2 Central Tracking

The tracking consists of two major subsystems, the silicon microstrip detector, which lies closest to the beampipe for high-precision tracking and secondary vertex identification, and the central drift chamber (the Central Outer Tracker or COT), which provides general-purpose charged-particle tracking. These two systems are contained in a solenoid which provides a 1.4 T magnetic field to enable measurement of charged particle momenta.

The silicon detector [19] is, in turn, composed of three separate components. The main part of the silicon detector, SVXII, extends from $r = 2.1$ cm to $r = 17.3$ cm and covers $|\eta| < 2.0$. It consists of five layers (numbered 0 to 4) of double-sided silicon strips; each layer consists of twelve "ladders" arranged radially. The strips are aligned axially on one side (thus providing a measurement in r–ϕ coordinates); the strips on the other side are oriented at either a 90-degree angle (for layers 0, 1, and 3) or a 1.2-degree angle (for layers 2 and 4) to allow for stereo measurements in

the z direction. SVXII is divided into three barrels in z extending approximately 45 cm in both directions from the interaction point.

Inside the SVXII is an additional layer, Layer 00 (L00), mounted directly on the beampipe. This layer consists of single-sided silicon microstrips with improved resistance to radiation damage to improve the overall resolution of the silicon tracking. Lying outside of the SVXII is the Intermediate Silicon Layer (ISL), consisting of a single layer at $r = 22$ cm for $|\eta| < 1.0$ and two layers at $r = 20$ and $r = 28$ cm for $1.0 < |\eta| < 2.0$, using double-sided microstrips with small-angle stereo. The ISL serves to improve linking of tracks between the main silicon detector and the COT, as well as provide improved silicon tracking in regions not covered by the COT.

The overall resolution of the silicon systems in the $r - \phi$ plane is 11 μm, with an impact parameter resolution $\sigma(d_0) \approx 40$ μm; approximately 35 μm is due to the transverse size of the Tevatron interaction region. Figure 3.5 shows an end view of the silicon systems.

The COT [20] is an open-cell drift chamber with a total of 2520 cells organized into eight superlayers, lying outside the silicon detector and extending to a radius of $r = 137$ cm; it covers the range $|\eta| < 1.0$. Each cell contains a total of 12 sense wires and 13 potential wires; the even-numbered superlayers have wires oriented axially, while the odd-numbered superlayers are offset at an angle of ± 2 degrees for stereo measurements. Figure 3.6 shows the layout of the COT. The COT is filled with an argon-ethane mix; the drift time for electrons produced by passing particles to reach the sense wires is designed to be < 100 ns. The central tracker provides a resolution on the track transverse momentum given by $\sigma(p_T)/p_T \approx 0.1\% \cdot p_T/(\mathrm{GeV}/c)$.

The superconducting solenoid lies outside the COT and provides the magnetic field necessary for measuring charged particle momenta in the central tracking systems. It is made of an aluminum-stabilized NbTi conductor, and is capable of

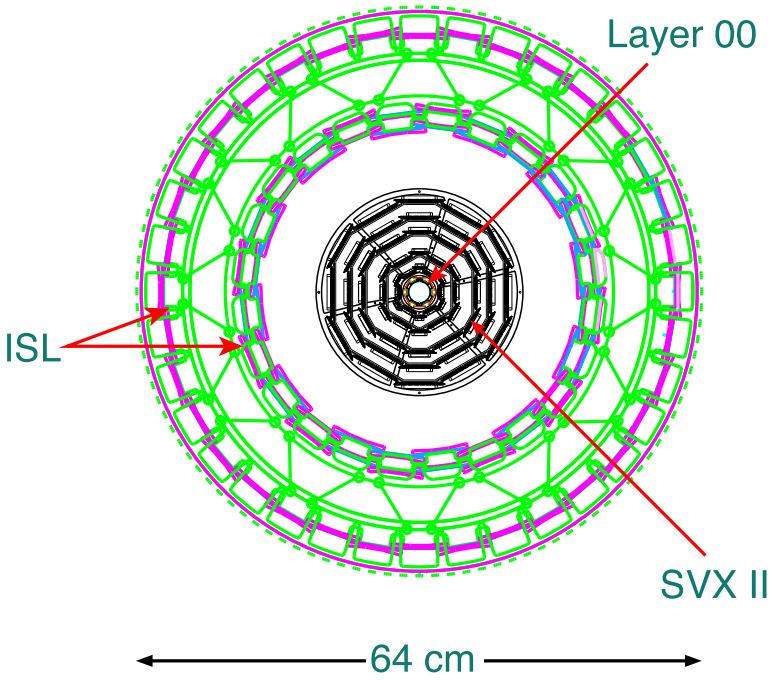

Figure 3.5: A cross-sectional view in the r–ϕ plane of the locations of the three components of the silicon detector, Layer 00, SVXII, and ISL.

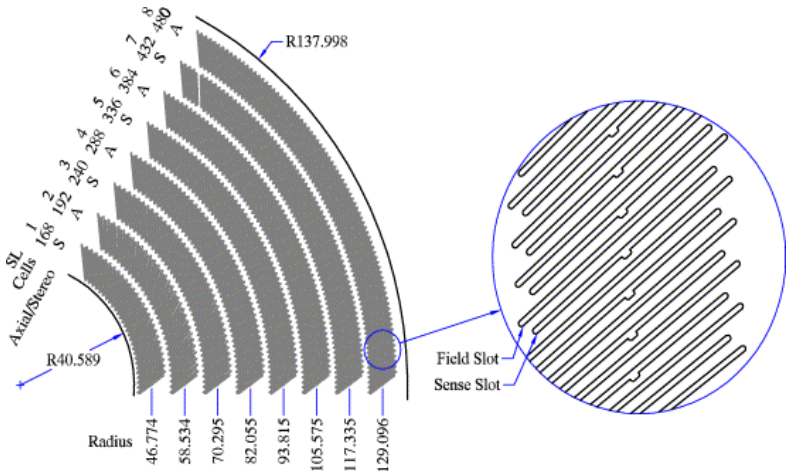

Figure 3.6: An end view of a 60-degree sector of the COT. The eight superlayers and the individual cells making up each superlayer are displayed.

operation up to 1.5 T; it is typically operated at a current of 4650 A which provides a field of 1.4 T in the $-z$ direction. The solenoid is contained inside a cryostat where it is cooled with liquid helium.

3.2.3 Calorimetry

The calorimeters are located outside the solenoid and consist of alternating layers of absorbing material and scintillators. They are designed to absorb and measure the energy of electrons, photons, and hadrons. Overall, the calorimeter consists of two different layers. The innermost layer, the electromagnetic calorimeter, is designed to stop electrons and photons and uses lead as the absorbing material. The outer layer is the hadronic calorimeter, which is designed to measure charged and neutral hadrons and uses steel as the absorbing material. The calorimeter is also split into two halves, one covering $\eta > 0$ and the other $\eta < 0$.

The depth of the calorimeter is measured in radiation lengths X_0 for the electromagnetic calorimeters, where one radiation length corresponds to the distance over which a high-energy electron will be reduced to $1/e$ of its original energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon [8]. For the hadronic calorimeters, the depth is measured in interaction lengths λ, where λ is the mean distance a hadron travels before undergoing an inelastic interaction with a nucleus.

There are five separate subsystems in the calorimeter overall. The central electromagnetic calorimeter (CEM) [21] covers $|\eta| < 1.1$ and contains 31 layers of lead and scintillator. The calorimeter is organized into "towers" in a projective geometry in η and ϕ pointing towards the center of the detector, with ten groups in η on either side of the interaction point and 24 wedges in ϕ. (Two of the towers are removed for cryogenic services for the solenoid, resulting in a total of 478 towers.) Two photomultiplier tubes are used to read out the light from the scintillators for each tower.

The hadronic calorimeter in the central region consists of two separate overlapping systems, the central hadronic calorimeter (CHA) and endwall hadronic calorimeter (WHA) [22], which together cover the region $|\eta| < 1.2$, with a total of 12 groups in η on either side of the interaction point and 24 wedges in ϕ. The CHA towers are 32 layers of steel and scintillator, while the WHA towers total 15 layers.

The forward region $1.1 < |\eta| < 3.6$ is covered by the plug calorimeters. The plug electromagnetic calorimeter (PEM) [23] is made up similarly to the CEM, with 23 layers of lead and scintillator, organized into 12 tower groups in η on each side, while the plug hadronic calorimeter (PHA) [24] contains 23 layers of iron and scintillator organized into 11 tower groups in η on each side.

The electromagnetic calorimeters also contain additional systems located approximately six radiation lengths deep (which is where the electromagnetic shower produced by particles in the EM calorimeter is expected to reach its maximum) to measure the transverse profile of the showers to improve particle identification. In the CEM, this detector (the central electromagnetic shower maximum detector, or CES) consists of strip and wire chambers, while in the PEM this detector is made of two layers of crossed scintillating strips.

The resolution and depths of the various calorimeter subsystems is summarized in Table 3.1.

3.2.4 Muon Detection

We expect that the calorimeters will absorb nearly all particles produced in $p\bar{p}$ collisions. Neutrinos are, of course, expected to escape the detector completely undetected, but muons are also expected to pass through the calorimeters, as muons produced with energies typical of $p\bar{p}$ collisions are minimum-ionizing particles and lose little energy to bremsstrahlung. Consequently, CDF has additional drift chambers and scintillators located outside the calorimeter to detect and mea-

Table 3.1: Resolution for a single particle (e.g., an electron in the EM calorimeters, or a pion in the hadronic calorimeters), depth, and coverage of the different calorimeter subsystems. Note that because the CHA and WHA overlap, the depth that a given particle sees may be greater than the listed value for the individual detectors.

System	Resolution (%)	Coverage	Depth
CEM	$13.5/\sqrt{E_T/\text{GeV}} \oplus 2$	$\lvert\eta\rvert < 1.1$	$19X_0$
PEM	$16/\sqrt{E_T/\text{GeV}} \oplus 1$	$1.1 < \lvert\eta\rvert < 3.6$	$21X_0$
CHA	$50/\sqrt{E_T/\text{GeV}} \oplus 3$	$\lvert\eta\rvert < 0.9$	4.5λ
WHA	$75/\sqrt{E_T/\text{GeV}} \oplus 4$	$0.8 < \lvert\eta\rvert < 1.2$	4.5λ
PHA	$80/\sqrt{E_T/\text{GeV}} \oplus 5$	$1.2 < \lvert\eta\rvert < 3.6$	7λ

sure these muons.

There are three separate muon systems of importance to $t\bar{t}$ analyses. (There are also some additional systems at higher $|\eta|$ ranges which are not used for $t\bar{t}$ measurements and hence not discussed here.) The first is the central muon system (CMU) [25], which is located directly outside of CHA. The CMU consists of four layers of wire chambers operated in proportional mode, and covers a range of $|\eta| < 0.6$.

Although the calorimeter is designed to absorb all hadrons, there is still a non-negligible rate of "punch-through" of hadrons which traverse the calorimeter and reach the CMU. To reduce the background rate in the CMU due to this punch-through, an additional set of chambers was built to make the central muon upgrade (CMP). Unlike most of CDF, CMP is not cylindrically symmetric, as it uses the steel of the magnet return yoke as additional shielding; consequently, it forms a rough box around CDF. CMP covers the same $|\eta|$ range as CMU, and muons in

this region are required to record hits in both CMU and CMP in order to improve the signal-to-background ratio. These muons are thus called CMUP muons.

The central muon extension (CMX) covers the range $0.6 < |\eta| < 1.0$ and is shaped like a truncated cone outside of the CMP. Like CMU and CMP, it consists of wire chambers arranged in four layers. Due to the presence of the floor, the CMX geometry is slightly changed in the bottom quadrant, where it instead takes a fan arrangement (known as the "miniskirt").

In addition to the drift chambers, there are also scintillator tiles on the outside of the CMP and CMX for trigger and timing purposes. Because the drift time in the muon chambers can be quite long (much longer than the time between bunch collisions), these scintillator tiles can be used for fast triggering and correct association of hits in the muon chambers.

A track observed in the muon chamber is referred to as a "stub" and must be matched to a corresponding track in the COT in order to be identified as a muon. The muon identification algorithm is discussed further in Section 3.4.1.

3.3 Trigger and DAQ

Given the bunch crossing rate of 1.7 MHz at CDF, saving the full output of the detector for all events would be prohibitively impractical. However, only a small fraction of events actually contain interesting physics. The trigger system is a three-layer system which progressively reduces the event rate to a much smaller rate suitable for saving events to tape. Many different trigger paths are defined to collect events for the different analyses performed at CDF.

3.3.1 Level 1 Trigger

The Level 1 trigger is responsible for reducing the event rate to approximately 20 kHz; it use specialized hardware and reads out only some parts of the detector

which can be quickly read out. The three main components of the L1 trigger are the XFT (eXtremely Fast Tracker), which provides quick reconstruction of particle tracks with $p_T > 1.5$ GeV/c in the COT with a resolution of $\sigma(p_T)/p_T \approx 2\% \cdot p_T/(\text{GeV}/c)$ [26]. The XFT also passes tracks it finds to the extrapolation unit (XTRP), which extrapolates track trajectories into the CMU to look for matches with muon stubs. The L1 trigger also includes information from the calorimeter towers.

3.3.2 Level 2 Trigger

The Level 2 trigger reduces the event rate further to $\sim$ 300 Hz, and features two main pieces of dedicated hardware. The first is the silicon vertex tracker (SVT) [27], which uses the silicon tracking information to look for displaced secondary vertices. The second clusters the energy deposited in the calorimeters to improve the selection requirements for calorimeter-based triggers.

3.3.3 Level 3 Trigger

The Level 3 trigger runs a full event reconstruction, and features an output rate of $\sim$ 75-100 Hz. Events which pass the Level 3 trigger are saved to tape. The event reconstruction is carried out by a dedicated PC farm which uses an optimized version of the offline reconstruction code and utilizes the full detector information available to select events, including full three-dimensional track reconstruction.

3.3.4 Top Triggers

There are three different triggers used for collecting $t\bar{t}$ candidates. The CEM trigger selects high-E_T electrons in CEM, the CMUP trigger selects high-p_T muons in CMU and CMP, and the CMX trigger selects high-p_T muons in CMX.

Specifically, the CEM trigger requires, at level 1, a COT track with $p_T > 8$ GeV/c pointing to a CEM tower with $E_T > 8$ GeV and a ratio of electromagnetic

to hadronic energy $E_{\rm had}/E_{\rm em} < 0.125$. At level 2, the calorimeter clustering is performed, and the requirement is a cluster with $E_T > 16$ GeV matched to a $p_T > 8$ GeV/c track. Finally, at level 3, the final reconstructed energy of the electron is required to pass $E_T > 18$ GeV with a matching track of $p_T > 9$ GeV/c. The $E_{\rm had}/E_{\rm em} < 0.125$ requirement is also enforced at levels 2 and 3.

The CMUP trigger requires, at level 1, a track with $p_T > 4$ GeV/c matched to a stub in CMU and CMP hits consistent with the observed CMU hits. No additional requirements are made at level 2; level 3 requires a final reconstructed COT track of $p_T > 18$ GeV/c matched to stubs in CMU and CMP. The CMX trigger operates very similarly, but requires a track of $p_T > 8$ GeV/c at level 1 matched to a CMX stub.

3.4 Event Reconstruction

After an event has been accepted, a full event reconstruction is performed offline; at this stage, raw tracks and clusters are identified as electrons, muons, jets, and other objects.

3.4.1 Lepton Identification

While the triggers described in Section 3.3 apply basic selection requirements, the offline reconstruction allows us to apply more comprehensive electron and muon selection requirements to obtain very high-quality lepton identification [28]. The following describes the standard cuts used for electron and muon identification at CDF.

The final electron cuts, which are summarized in Table 3.2, are as follows:

- To be accepted as a tight CEM electron, the electron must have $|\eta| < 1.0$ and $E_T > 20$ GeV matched to a COT track passing various track quality cuts.

- To further reject showers from hadrons, the $E_{\rm had}/E_{\rm em}$ cut is tightened to $E_{\rm had}/E_{\rm em} < 0.055 + 0.00045 \cdot E/{\rm GeV}$. The additional linear term accounts for the fact that higher-energy electrons will leak more energy into the hadronic calorimeter.

- The ratio of the calorimeter energy to the COT track momentum, E/p ($= E_T/p_T$), is required to be at most 2.0 if $E_T < 100$ GeV. (Above 100 GeV this cut becomes unreliable and is not used.)

- $L_{\rm shr}$, a variable describing the match of the lateral profile shape to the expected shape for electrons derived from test-beam data, must be < 0.2.

- A χ^2 comparison is used to compare the shape of the profile measured in the CES to the expected profile and $\chi^2_{\rm strips} < 10.0$ is required.

- The distances between the extrapolated COT track and the position of the CES cluster, Δx in the r–ϕ plane and Δz in the r–z plane, are required to satisfy -3.0 cm $< Q \cdot \Delta x < 1.5$ cm and $|\Delta z| < 3.0$ cm, where Q is the sign of the charge.

- Conversions (pair production by interactions of a photon with the detector material) are rejected by searching for a track with opposite sign, $\Delta(xy) < 2$ mm, and $|\Delta(\cot\theta)| < 0.04$, where $\Delta(xy)$ is the distance in the r–ϕ plane at the point where the tracks are parallel. If such a track is found, the electron is assumed to come from a pair production and is rejected.

- To identify the electron as coming from a W decay (as opposed to a decay of a heavy-flavor hadron), the electron is required to be isolated from other sources of energy in the calorimeter. This is enforced by measuring other energy in a cone of radius $\Delta R = 0.4$ around the electron; the ratio of this other energy (not including the electron itself) to the electron energy must be less than 0.1.

Table 3.2: Selection requirements applied for electron identification.

Variable	Required Value
E_T	> 20 GeV
$\|\eta\|$	< 1.0
$E_{\text{had}}/E_{\text{em}}$	$< 0.055 + 0.00045 \cdot E/\text{GeV}$
E/p	< 2.0 if $E_T < 100$ GeV
L_{shr}	< 0.2
$Q \cdot \Delta x$	> -3.0 cm and < 1.5 cm
$\|\Delta z\|$	< 3.0
χ^2_{strips}	< 10.0
Conversions	Rejected
Isolation	< 0.1

The cuts used to identify muons, which are summarized in Table 3.3, are:

- To be identified as a tight CMUP or CMX muon, we require $p_T > 20$ GeV/c and $|\eta| < 1.0$.
- The energy deposited in the calorimeter is required to be consistent with a minimum ionizing particle. This is enforced by requiring $E_{\rm em} < 2$ GeV and $E_{\rm had} < 6$ GeV for $p < 100$ GeV/c and $E_{\rm em} < 2 + (p/({\rm GeV}/c) - 100) \cdot 0.0115$ and $E_{\rm had} < 6 + (p/({\rm GeV}/c) - 100) \cdot 0.0280$ for $p > 100$ GeV/c.
- The distance $|\Delta X|$ between the extrapolated COT track and the stub in the muon chambers is required to be < 3.0 cm in CMU, < 5.0 cm in CMP, and < 6.0 cm in CMX.
- For CMP and CMX, the extrapolated tracks are also required to lie at least 3 cm away along the wire axis from the edges of the muon chambers to avoid chamber edge effects.
- Muons originating from cosmic rays are identified by an algorithm which looks for tracks passing through the detector and rejected.
- Like electrons, an isolation cut of < 0.1 is applied to the muon to require it to come from W decay.

3.4.2 Jet Identification and Correction

As described earlier in Section 2.1.2, bare partons are, of course, not directly observed in our detector; rather, they produce a shower of collimated hadrons known as a "jet". By measuring the energy of a jet, we can thus get an estimate of the energy of the parton which originated it; this process involves a series of corrections to the raw measured jet E_T. However, this introduces several significant

Table 3.3: Selection requirements applied for muon identification.

Variable	Required Value
E_T	> 20 GeV
$\|\eta\|$	< 1
E_{em}	$< \max(2, 2 + (p/(\text{GeV}/c) - 100) \cdot 0.0115)$
E_{had}	$< \max(6, 6 + (p/(\text{GeV}/c) - 100) \cdot 0.0280)$
$\|\Delta X\|$	< 3.0 cm (CMU)/5.0 cm (CMP)/6.0 cm (CMX)
Distance from edge along z-axis	> 3.0 cm (CMP and CMX)
Cosmic rays	Rejected
Isolation	< 0.1

sources of systematic uncertainties, which in turn results in a potentially large source of uncertainty in our final measurement.

Jets are identified using a jet clustering algorithm known as JetClu [29]; briefly, JetClu identifies seed towers in the calorimeter with an energy of at least 1 GeV and adds groups of clusters in the electromagnetic and hadronic calorimeters within a cone of radius $\Delta R = 0.4$ to the seed. The centroid of the jet is then calculated by summing the η and ϕ positions of the towers weighted by their E_T and a new list of towers around the new center is calculated; this process is repeated iteratively until it converges. These measurements provide a resolution of approximately $\sigma(E_T) \approx 0.1 \cdot E_T + 1.0$ GeV.

After the raw jet E_T is calculated, a series of corrections is applied [30]. There is a total of five levels of corrections, numbered 1 through 7 (levels 2 and 3 correspond to corrections which are no longer used).

The relative correction (level 1) accounts for the η dependence of the calorimeter

response, primarily arising from the "central crack" at $\eta = 0$ where the two halves of the calorimeter meet and the region at $|\eta| \approx 1.1$ where the central and plug calorimeters meet. This correction is derived from dijet events where one jet is required to be in the central region of the central calorimeter, $0.2 < |\eta| < 0.6$, and assuming that the p_T of the two jets (in the absence of hard QCD radiation) should balance.

The multiple interactions correction (level 4) accounts for the different calorimeter response as a function of the number of different $p\bar{p}$ interactions occurring in a single bunch crossing; the number of interactions is measured by the number of vertices along the beamline, or z-vertices. This correction is obtained by taking minimum bias data and measuring the amount of energy in a randomly-selected cone in the central calorimeter region $0.2 < |\eta| < 0.6$ as a function of the number of observed z-vertices in the event.

The absolute correction (level 5) is designed to convert the measured jet energy into the energy of the underlying particle, accounting for the nonlinear response of the calorimeter. This correction is derived by using PYTHIA Monte Carlo dijet events and comparing the measured and true jet p_T in these events.

There are also two corrections which we do not actually use in defining the jet energies in our measurement, but for which the systematic uncertainties need to be taken into account. These are the corrections for the underlying event (level 6) and out-of-cone energy (level 7). The "underlying event" refers to the fact that partons in the parent $p\bar{p}$ system which do not take place in the main interaction can also be responsible for producing energy measured in the jets, while the out-of-cone energy refers to particles that escape the cone of $\Delta R = 0.4$. These are also measured using PYTHIA and HERWIG Monte Carlo dijet samples.

The systematic uncertainties for these corrections are summarized in Figure 3.7. These uncertainties total approximately 3% for high-p_T jets and can be up to 10% for lower-p_T jets; if taken directly as a systematic on our final measurement,

this would be by far our largest single source of uncertainty. However, since we use events which contain a W boson decaying to hadrons, we can use the known mass of the W boson to obtain an additional constraint to the jet energy scale (JES) in our measurement. This technique is discussed further in Section 5.2.

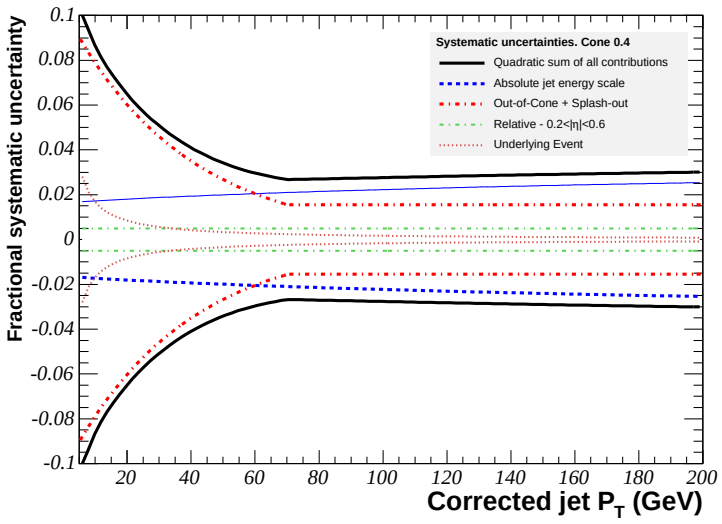

Figure 3.7: Systematic uncertainties on the measurement of jet energies as a function of jet p_T for central jets.

3.4.3 Missing E_T

Because neutrinos produced are expected to escape undetected, we can infer their presence by looking for imbalances in the energy in the detector. We define the missing E_T, $\not\!\!E_T$, as follows:

$$\not\!\!E_T = \left| \sum_{i \in \text{towers}} E_{T_i} \hat{n}_{T_i} \right|, \tag{3.2}$$

where the sum is performed over all calorimeter towers, $\hat{n}_{T_i}$ is the unit vector in the x–y plane pointing from the primary vertex to tower i, and E_{T_i} is the uncorrected E_T measured in that tower.

This raw measurement is corrected with two further corrections [31]. As muons only deposit a small fraction of their energy in the calorimeter, if isolated high-p_T muons are found in the event, the energy deposited by the muon in the calorimeter is subtracted and replaced by the muon $\vec{p_T}$. The $\not\!\!E_T$ is also corrected to take into account the corrections applied to the E_T of jets in the event.

3.4.4 Secondary Vertex Identification

When a b quark is produced in a $t\bar{t}$ decay, it hadronizes into a B hadron with a lifetime of ~ 1.5 ps. These hadrons are sufficiently energetic that they can travel several millimeters in the lab frame before decaying; thus, the particles produced in this b decay will form a vertex displaced from the primary vertex. Identification of these secondary vertices is thus of primary importance in finding b quarks and hence $t\bar{t}$ decays; the CDF silicon system described in Section 3.2.2 is primarily designed for this purpose.

The algorithm, known as SECVTX [32], operates as follows. Given a jet, the charged particle tracks in the jet are subjected to some quality cuts requiring a number of good-quality silicon hits and a good χ^2 for the fits of those silicon hits into a track. The algorithm then tries to reconstruct a secondary vertex with at least three tracks of $p_T > 0.5$ GeV/c, of which one must have $p_T > 1.0$ GeV/c, or a vertex with at least two tracks of $p_T > 1.0$ GeV/c. If a secondary vertex is located, then the two-dimensional decay length $L_{2\mathrm{D}}$ is calculated by projecting (in the r–ϕ plane) the vector from the primary vertex to the secondary vertex onto the jet axis. Figure 3.8 illustrates this process. A jet is considered tagged if $L_{2\mathrm{D}} > 7.5\sigma_{L_{2\mathrm{D}}}$, where $\sigma_{L_{2\mathrm{D}}}$, the uncertainty on $L_{2\mathrm{D}}$, is approximately 190 μm.

In order to measure the efficiency of the tagging algorithm for $t\bar{t}$ events, a two-step strategy is used. First, we measure the tagging efficiency in a sample enriched in heavy flavor; specifically, we select events with two back-to-back jets, where one of the jets is required to contain a low-p_T electron. We assume that the electron

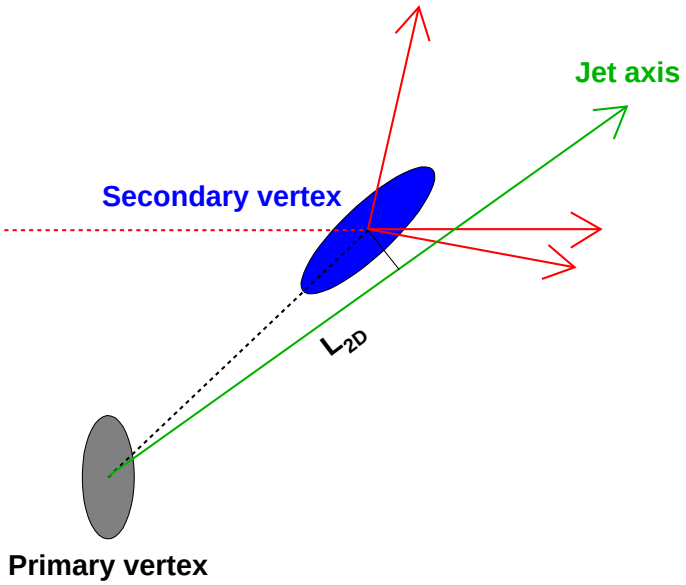

Figure 3.8: An illustration of a b-jet in the r–ϕ plane showing how L_{2D} is derived from the positions of the primary and secondary vertices. The value of L_{2D} is used to determine whether the jet originates from a b quark.

is produced by the decay of a heavy-flavor hadron, and so the other jet is also likely to originate from a heavy-flavor parton. Then, we can measure the tagging efficiencies in data and Monte Carlo. Because the Monte Carlo does not model the secondary vertex tagging precisely, there is a "scale factor" between the data and Monte Carlo efficiencies of approximately 0.95. The tagging efficiencies for jets in $t\bar{t}$ data can then be estimated by multiplying the tagging efficiencies obtained from $t\bar{t}$ Monte Carlo by this scale factor.

It is also important to know the mistag rate, that is, the rate of jets not originating from a heavy flavor quark which are nevertheless tagged by SECVTX. This is mostly caused by random combinations of tracks displaced from the primary vertex. We begin our estimate of the mistag rate by assuming that mistags are equally likely to have a negative L_{2D}, which can clearly not be produced by an actual physical process, as a positive L_{2D}, and thus the mistag rate of jets with a

positive L_{2D} is equal to the rate at which jets are tagged with a negative L_{2D} value; this rate can be measured directly from jet data samples. Because the mistag rate is not quite symmetric in L_{2D} because of long-lived particles and detector interactions, two corrections are applied to account for this asymmetry; these corrections are derived by fitting Monte Carlo templates of the signed vertex mass (the invariant mass of the tracks reconstructed in the vertex, multiplied by the sign of L_{2D}) for b, c, and light jets to the distribution in data.

The overall tagging efficiency and mistag rates as functions of jet E_T and η for b-jets in $t\bar{t}$ events are shown in Figure 3.9. As we can see, the efficiency for tagging b-jets is typically around 40% with a mistag rate of less than 2%.

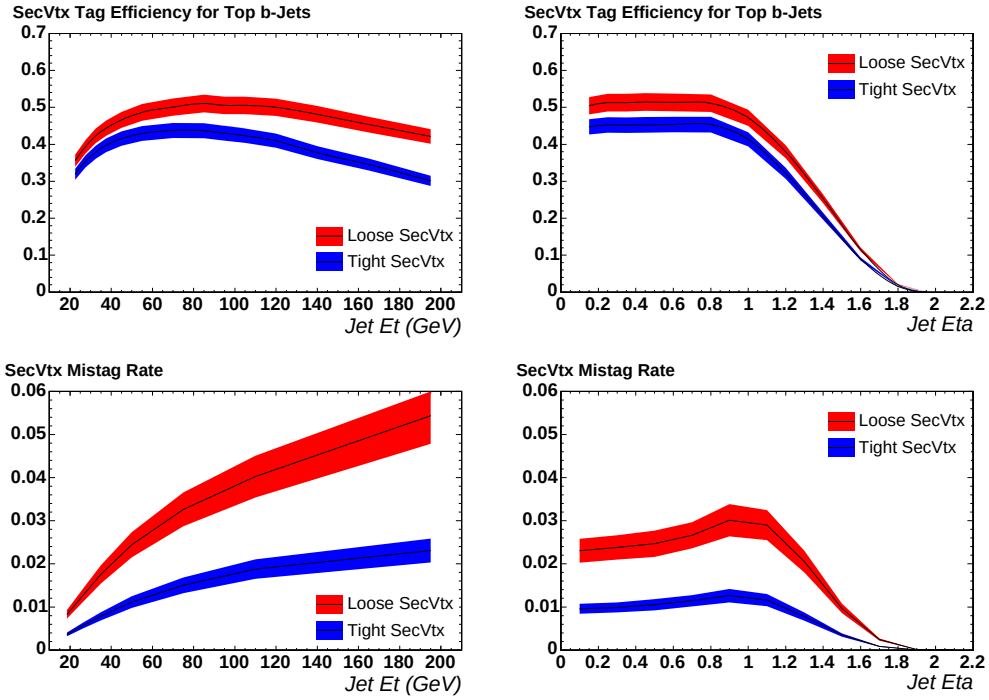

Figure 3.9: Tagging efficiency for b-jets in $t\bar{t}$ events (top) and mistag rates (bottom) as a function of jet E_T (left) and η (right). Our analysis uses the "tight" SECVTX tagging algorithm as described in the text.

Chapter 4

Event Selection

4.1 The Lepton + Jets Topology

As mentioned in Section 2.2, top quarks can be produced at the Tevatron either in $t\bar{t}$ pairs or singly. However, looking for $t\bar{t}$ pair events is experimentally favorable for two reasons: first, the cross section for $t\bar{t}$ production at the Tevatron is currently measured by CDF at 7.0 ± 0.6 pb [16], which is much larger than the currently-observed cross-section for single-top production of $2.3^{+0.6}_{-0.5}$ pb [33]; and second, the additional top quark in $t\bar{t}$ events means that the overall experimental signature is much cleaner. In fact, while the top quark was first discovered in $t\bar{t}$ production in 1995, the discovery of single-top production did not happen until this year (2009). For these reasons, we focus on $t\bar{t}$ events.

We can further classify $t\bar{t}$ events based on their decay properties. As previously noted in Section 2.2.2, in the Standard Model, top quarks decay into a W boson and a b quark almost 100% of the time, so the decay chain is always expected to begin $t\bar{t} \to W^+bW^-\bar{b}$. The W bosons can then decay two different ways:

- $W \to e\nu_e$, $\mu\nu_\mu$ or $\tau\nu_\tau$ ("leptonic" decay)
- $W \to q\bar{q}'$ ("hadronic" decay)

The $W \to \tau$ channel poses serious experimental difficulties, due to the subsequent decay of the τ (either into other leptons via $\tau \to \nu_\tau \ell \nu_\ell$ or via a hadronic decay such as $\tau \to \pi^-\pi^0\nu_\tau$) which produces at least one more neutrino and other objects, so events with $W \to \tau$ decay are typically discarded; thus, when we refer to "leptonic" decay, we mean specifically an electron or muon. We can classify the non-τ events as "dilepton" events, where both of the Ws decay leptonically, "all-hadronic" events, where both of the Ws decay hadronically, or "lepton+jets" events, where one of the Ws decays hadronically and the other leptonically.

The branching ratios for each of these three channels is shown in Figure 4.1. Each of these channels has its own advantages and disadvantages.

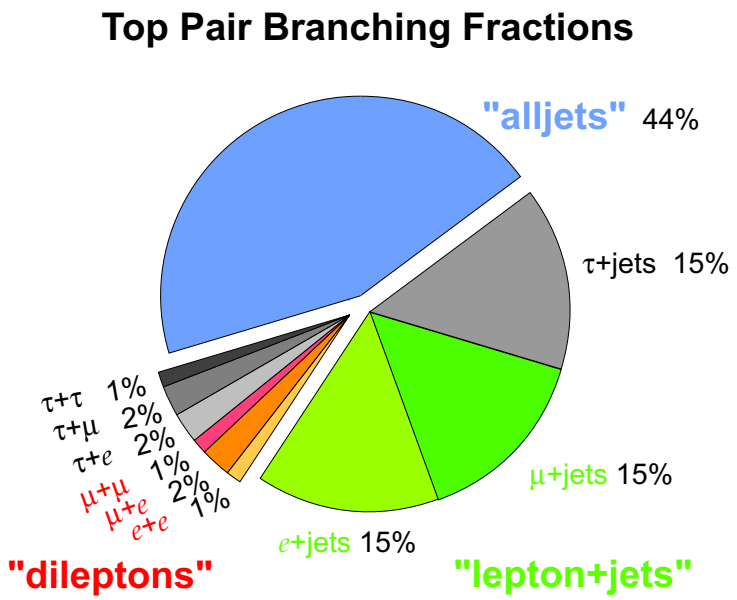

Figure 4.1: Branching ratios for the dilepton, lepton+jets, and all-hadronic channels in $t\bar{t}$ decay.

- The dilepton channel, because of the presence of two high-energy leptons, offers a good signal-to-background ratio. However, the branching ratio is the lowest for this channel.[1] Furthermore, because of the two neutrinos, the kinematics of the event are underconstrained.

- The lepton+jets channel represents a compromise: the experimental signature is still relatively clean thanks to the high-energy lepton, but the branching ratio is still reasonably large. With only one neutrino, the event can be kinematically reconstructed. Furthermore, the presence of the hadronic W allows for an in-situ technique to calibrate the jet energy scale, as mentioned in Section 3.4.2.

- The all-hadronic channel has the largest branching ratio, providing a large

[1] To compensate, dilepton analyses at CDF usually do not require a b-tag in their event selection, which actually means that their signal-to-background ratio is worse than in the lepton+jets channel.

sample of events. However, with no leptons present, the background from QCD events is very significant and sophisticated techniques are necessary to improve the signal to background ratio. The in-situ W techniques can also be applied here.

This measurement is performed in the lepton+jets channel, as it tends to offer the best balance between these competing goals; in fact, the best measurements of the top quark mass to date (including this measurement) have all been in the lepton+jets channel. Figure 4.2 shows a Feynman diagram for a typical lepton+jets event.

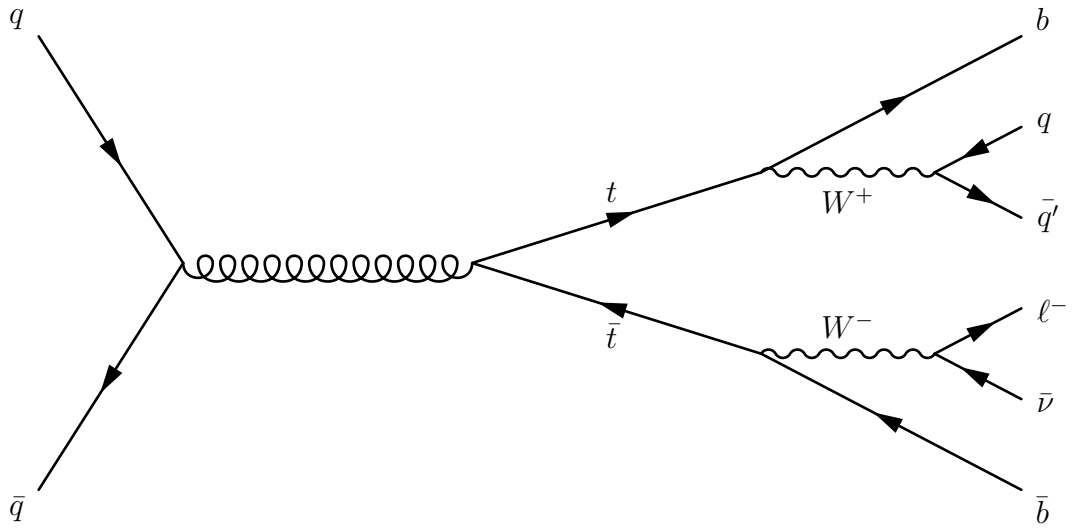

Figure 4.2: Feynman diagram for a typical lepton+jets event.

A note on terminology: In lepton+jets events, we refer to the $t \to Wb \to \ell\nu b$ side of the event as the "leptonic side", and the $t \to Wb \to q\bar{q}'b$ side of the event as the "hadronic side". By extension, the t and b quarks on the hadronic side are referred to as the "hadronic t" and "hadronic b", while the t and b quarks on the leptonic side are the "leptonic t" and "leptonic b". It should be noted that these terms do not refer to the decay of the t or b themselves, but simply of the decay of the W on that side of the event. We also often refer to the two quarks produced by the hadronic W as "light" quarks, although of course some of these are actually

c quarks.[2]

4.2 Event Signature

Events in the lepton+jets channel, in total, produce four quarks (two b and two light quarks), a lepton, and a neutrino. We thus would expect to see in our detector four high-E_T jets, of which two would originate from b quarks, one high-p_T lepton, and $\not\!\!E_T$ from the neutrino. Consequently, we apply the following cuts:

- For the jets, we require exactly four jets with a corrected E_T (as described in Section 3.4.2) of at least 20 GeV in the region $|\eta| < 2.0$. At least one of the jets must be tagged as a b-jet using the algorithm described in Section 3.4.4. (The event may contain any number of jets with $E_T < 20$ GeV.)

- For the lepton, we require at least one identified electron with $E_T > 20$ GeV or one identified muon with $p_T > 20$ GeV/c in the central region of the detector ($|\eta| < 1.0$), using the identification requirements described in Section 3.4.1. Events with more than one lepton passing these requirements are rejected.

- For the neutrino, we require $\not\!\!E_T > 20$ GeV.

There are several non-$t\bar{t}$ processes which can mimic this signature, and thus need to be considered as potential background. The three main sources of background are:

- "W + heavy flavor" (or "W + HF"), events where a W boson is produced in association with jets containing heavy flavor ($b\bar{b}$, $c\bar{c}$, or c) and other jets. As these events contain a real W and a correct b-tag, this is an unavoidable background source. However, the production of real heavy flavor is relatively low, so this background source does not overwhelm our signal.

[2]There is also a very small number of b quarks produced by $W \to cb$ decay.

- "W + light", events where a W boson is produced in association with jets not containing heavy flavor, but where one of the jets is mistagged as a b-jet. The amount of this background is naturally dependent on the mistag performance of our tagging algorithm. Since W + light production is much more common than W + HF, the mistag rate needs to be low in order to keep this background source low.

- "Non-W QCD", QCD events where no true W boson is produced but both a fake lepton and fake $\not\!\!E_T$ are created (the former from the misidentification of a jet as a lepton and the latter usually from a jet escaping into a non-instrumented region of the calorimeter). Since QCD events are produced in much greater quantities than W-containing events, this background is potentially a very large source but is reduced significantly by the selection cuts.

These first two categories are often referred to collectively as "W + jets" backgrounds. There are also some smaller sources of background, which contribute much less than the above three but should still be noted:

- Single top production. These events contain a real W and two real b quarks, but the cross-section is lower for single top production to begin with, and these events also need two other high-energy jets to mimic the $t\bar{t}$ signature, so these are a relatively smaller contribution. (These also tend to have different kinematic properties from $t\bar{t}$ events, as the second b jet tends to have lower momentum.)

- Diboson production (WW, WZ, or ZZ). Like single top, these require extra jets to mimic the $t\bar{t}$ signature and also have a relatively small cross-section.

- $Z \to \ell\ell$ + jets. If a Z decays into two leptons (one of which is misidentified or missed) and is produced in conjunction with several other jets, it can

reproduce the $t\bar{t}$ signature. While the cross-section for single Z production is relatively high, the acceptance for these events to pass the $t\bar{t}$ cuts is quite low.

4.3 Data Samples

The data used in this analysis is collected by the CEM, CMUP, and CMX triggers described in Section 3.3, and then the selection criteria described in Section 4.2 are applied. This analysis uses data collected by the CDF II detector over a period from February 4, 2002 to March 21, 2009, with a total of 4.3 ± 0.3 fb^{-1} of integrated luminosity.

The condition of the detector has changed over time; to take this into account, the data is divided into 24 run periods, each with appropriate efficiencies and corrections. Table 4.1 shows the number of events observed in each run period.

Note that we separate events with 1 b-tag from events with 2 (or more) b-tags. The latter sample has a much lower background rate, but of course also fewer events.

4.4 Background Estimation

As mentioned above, a certain fraction of the events passing our selection cuts will not be true $t\bar{t}$ events, but rather background. Since the background events do not contain any useful information on the top mass (the single top events, of course, do contain some information but are still not useful if reconstructed as $t\bar{t}$ events), it is necessary to adopt a strategy to minimize their effect on the reconstructed top mass, so we need to know the expected contributions of each of the above background types to the overall observed total. We use a method originally developed for the $t\bar{t}$ cross-section measurement [34] (where precise background knowledge is even more important), known as "Method II For You". Below is a summary of the

Table 4.1: Data samples used in this analysis, corresponding to a total integrated luminosity of 4.3 fb^{-1}. "Tagged events" is the total number of events passing our selection cuts.

Period	Dates	Lum. (pb^{-1})	Tagged events	1-tag	$\geq$ 2-tag
0	2/4/02-8/22/04	318	70	56	14
1-4	12/7/04-9/4/05	360	77	59	18
5-7	9/5/05-2/22/06	258	53	42	11
8	6/9/06-9/1/06	166	24	17	7
9	9/1/06-11/22/06	153	33	26	7
10	11/24/06-1/30/07	243	51	37	14
11	1/31/07-3/30/07	230	43	31	12
12	4/1/07-5/13/07	155	23	19	4
13	5/13/07-8/4/07	268	42	33	9
14	10/28/07-12/3/07	31	9	9	0
15	12/5/07-1/27/08	156	30	25	5
16	1/27/08-2/27/08	101	12	11	1
17	2/28/08-4/16/08	183	24	22	2
18	4/18/08-7/1/08	305	55	45	10
19	7/1/08-8/24/08	207	32	27	5
20	8/24/08-10/4/08	227	33	21	12
21	10/12/08-1/1/09	436	63	54	9
22	1/2/09-2/10/09	266	42	37	5
23	2/15/09-3/21/09	201	22	19	3
Total	2/4/02-3/21/09	4264	738	590	148

technique used; for full details, please consult the preceding paper.

The basic principle of Method II is to use Monte Carlo simulation to estimate the background contributions from sources where the Monte Carlo is well-understood and models the data well; for channels where the Monte Carlo is known not to model the data contributions well, it employs a data-based approach.

Method II operates on both the "pretag" sample, the sample derived from applying all of the $t\bar{t}$ selection requirements except for the tagging requirement, and the "tagged" sample, which includes the tag requirement, using them together to arrive at a final answer. The method begins with the assumption that the number of observed events in both the pretagged and tagged samples is equal to the sum of the known contributions:

$$N_{\text{obs}} = N_{t\bar{t}} + N_{W+\text{jets}} + N_{\text{non-}W\text{ QCD}} + N_{\text{single top}} + N_{\text{diboson}} + N_{Z+\text{jets}}. \qquad (4.1)$$

The first step is to estimate the contribution for samples well-modeled by the Monte Carlo. This includes the single top, diboson, and Z + jets backgrounds, as well as the $t\bar{t}$ contribution itself. For each of these samples (note that the two production channels for single top, s-channel and t-channel, are computed separately), the estimated total is calculated by multiplying the theoretical cross section by the integrated luminosity and the event selection efficiency derived from Monte Carlo:

$$N_i = \sigma_i \cdot \epsilon_i \cdot L_{\text{int}}, \text{ for } i = t\bar{t}, \text{ single top, diboson, and } Z + \text{jets}, \qquad (4.2)$$

where σ_i is the theoretical cross-section, ϵ_i is the event selection efficiency, and L_{int} is the integrated luminosity. These contributions are calculated for both the pretag and the tagged sample. The latter includes an efficiency for the tagging, which is corrected by the scale factor as discussed in Section 3.4.4. (Note that the $t\bar{t}$ contribution for this calculation is determined by assuming $m_t = 172.5$ GeV/c^2 with a cross-section of 7.4 pb.)

The second step is to estimate the fraction of non-W QCD events. Because of the complexity of modeling QCD events in Monte Carlo, this contribution cannot

be accurately estimated using a Monte Carlo–based approach. Instead, Method II uses data sidebands to estimate the non-W QCD contribution. For electrons, the data sideband used is an "antielectron" sample, which consists of electrons that fail two of the selection requirements discussed in Section 3.4.1, and for muons, the data sideband used consists of events with muons which pass all of the selection requirements except for the isolation requirement. In general, the non-W QCD events have a lower $\not\!\!E_T$ distribution than the W + jets events, so this can be used to separate the two.

In the pretag sample, the non-W QCD fraction is estimated by performing a fit to the observed $\not\!\!E_T$ distribution using two templates: the W + jets $\not\!\!E_T$ distribution derived from Monte Carlo, and the non-W QCD $\not\!\!E_T$ distribution derived from the data sideband. The tagged sample is also fit similarly. After the non-W QCD fraction is calculated, the number of expected non-W QCD events is calculated using this fraction:

$$N_{\text{QCD}} = f_{\text{QCD}} \cdot N_{\text{obs}}, \tag{4.3}$$

where f_{QCD} is the non-W QCD fraction as calculated above.

Now, for both the pretag and tagged samples, these contributions are subtracted from the total, leaving the remainder as the number of W + jets events:

$$N_{W+\text{jets}} = N_{\text{obs}} - N_{t\bar{t}} - N_{\text{single top}} - N_{\text{diboson}} - N_{Z+\text{jets}} - N_{\text{non-}W\text{ QCD}} \tag{4.4}$$

The final step is thus to divide up the number of W + jets events into the W + HF and W + light contributions. We know that the Monte Carlo simulation models the relative contributions of W + HF and W + light to the W + jets total well, but the overall normalization is unreliable, which is why the method adopts this approach rather than simply estimating the W + jets contribution from Monte Carlo. Hence, Monte Carlo simulation is used to calculate f_{HF}, the fraction of W + jets events containing heavy flavor, and then the tagging efficiency (including the scale factor defined in Section 3.4.4) is used to estimate the number of W +

HF events in the tagged sample:

$$N_{W+\mathrm{HF}}^{\mathrm{tag}} = N_{W+\mathrm{jets}}^{\mathrm{pretag}} \cdot f_{\mathrm{HF}} \cdot \epsilon_{\mathrm{tag}}. \tag{4.5}$$

Similarly, the expected number of tagged W + light events is calculated by applying the mistag rate as mentioned in Section 3.4.4:

$$N_{W+\mathrm{light}}^{\mathrm{tag}} = (N_{W+\mathrm{jets}}^{\mathrm{pretag}} - N_{W+HF}^{\mathrm{pretag}}) \cdot \epsilon_{\mathrm{mistag}}. \tag{4.6}$$

Table 4.2 shows the expected number of background events calculated using this method for the 4.3 fb^{-1} sample. Overall, in our sample of 738 events, we expect a total of 173.1 ± 50.0 background events.

4.5 Monte Carlo Samples

We use a variety of Monte Carlo simulated samples in constructing and evaluating our method. For $t\bar{t}$ signal events, we use events generated at a range of top masses from 160 to 184 GeV/c^2 using the PYTHIA Monte Carlo generator [35]. We also crosscheck our analysis using $t\bar{t}$ signal events generated with the HERWIG generator [36]. There are also a number of samples needed to evaluate our systematic uncertainties where various parameters have been changed to reflect our uncertainty in modeling a particular aspect of $t\bar{t}$ events; these are discussed more fully in Chapter 9.

For the background, the W + jets samples are simulated using the ALPGEN generator with the parton showering performed by PYTHIA [37]. These samples are generated with a specific number of partons in the matrix element; for instance, the $W + b\bar{b}$ contribution includes samples with $W + b\bar{b} + 0p$, $W + b\bar{b} + 1p$, and $W + b\bar{b} + \geq 2p$. In order to get the total $W + b\bar{b}$ contribution, we must add up these three subsamples. However, we must avoid double-counting, since it is possible that additional jets can be created during the parton shower, so overlaps are removed using a jet-parton matching algorithm. Similarly, because heavy flavor can be

Table 4.2: Expected backgrounds for the 4.3 fb^{-1} W+4 tight jet sample used.

Background	1 tag	$\geq$ 2 tags
non-W QCD	34.6 $\pm$ 29.8	2.2 $\pm$ 2.8
W+light mistag	29.4 $\pm$ 10.7	0.6 $\pm$ 0.3
diboson (WW, WZ, ZZ)	7.5 $\pm$ 0.8	0.7 $\pm$ 0.1
$Z \rightarrow \ell\ell$ + jets	5.4 $\pm$ 0.8	0.5 $\pm$ 0.1
Sum of above 3	42.2 $\pm$ 10.7	1.7 $\pm$ 0.3
$W + b\bar{b}$	39.0 $\pm$ 16.3	7.5 $\pm$ 2.7
$W + c\bar{c}$	24.0 $\pm$ 9.0	1.1 $\pm$ 0.4
$W + c$	12.1 $\pm$ 4.5	0.6 $\pm$ 0.2
Single top s-channel	2.7 $\pm$ 0.3	0.9 $\pm$ 0.1
Single top t-channel	3.5 $\pm$ 0.4	0.9 $\pm$ 0.1
Sum of above 5	81.4 $\pm$ 27.3	11.0 $\pm$ 3.2
Total background	158.2 $\pm$ 49.6	15.0 $\pm$ 5.7
Predicted top signal ($\sigma = 7.4$ pb)	440.5 $\pm$ 75.6	163.9 $\pm$ 25.7
Events observed	590	148

produced in the parton shower of W + light events, we also must remove heavy flavor using a jet-based heavy flavor algorithm [34]. After the overlap removal, the subsamples are then added together, with each subsample weighted by the cross-section for that subprocess times the acceptance. For the bins with lower parton multiplicity (e.g., $W + b\bar{b} + 0p$ or $W + 2p$), the acceptance for our cuts is very small, so for the sake of simplicity (and to avoid the statistical problems associated with using a sample with only a few events in it), we typically discard these lower-multiplicity bins and use only the two highest parton multiplicity subsamples (e.g., $W + b\bar{b} + 2p$ and $W + b\bar{b} + 1p$ for the $W + b\bar{b}$ channel).

The single-top samples are simulated with a top mass of 175 GeV/c^2 using the MadGraph/MadEvent [38] package along with PYTHIA for the parton shower and hadronization. For the sake of simplicity, since the contribution from the diboson and Z + jets backgrounds is small, we do not use separate MC samples for these, but rather increase the W + light total to include these events. All Monte Carlo samples are then simulated using the CDF II detector response simulation package based on the GEANT3 framework [39].

As Monte Carlo is not able to adequately model the non-W QCD background, we use instead a data sideband for our non-W events. Specifically, we select events from the electron and muon data using the same selection requirements as for the $t\bar{t}$ sample, except that we require the lepton to fail the isolation requirement discussed in Section 3.4.1 (i.e. Isolation > 0.1).

To validate that our Monte Carlo samples correctly model the parameters of interest in the data, we compare a variety of quantities between the two; this comparison is shown in Appendix A.

Tables 4.3 and 4.4 summarize the main Monte Carlo (and data-based QCD) samples used in this analysis. Table 4.5 shows how the W+light, $W + b\bar{b}$, $W + c\bar{c}$, and $W + c$ samples are subdivided into the individual parton multiplicity subsamples. Note that this computation is actually performed for single-tag and multiple-

tag results separately; we also show the results before and after the lowest parton multiplicity bins are discarded to show that this simplification has a very small effect on our result.

Note that the Monte Carlo samples in Tables 4.3 and 4.4 are generated with a simulated luminosity profile matching CDF run periods 0 through 8. We also use another sample, `ttop25`, generated with a luminosity profile covering CDF run periods 0 through 19 at a top mass of 172.5 GeV/c^2, to improve the match of our luminosity profile to data.

4.6 Bad Signal

In addition to the background discussed above, there is one further class of undesirable events that needs mentioning. These events are what we call "bad signal" events — events where the underlying physical process is indeed a $t\bar{t}$ event, but where the final objects that we observe in our detector do not all come directly from the $t\bar{t}$ decay. (The opposite of a bad signal event, a $t\bar{t}$ event where the objects we see in our detector do all come from the $t\bar{t}$ decay, is, naturally, called a "good signal" event.) There are several possibilities for how a bad signal event can arise, including:

- Extra jets can be produced from gluons radiated in the event, either from an initial state particle ("initial state radiation" or "ISR") or from a final state particle ("final state radiation" or "FSR"), while a jet from the $t\bar{t}$ decay is lost or fails to pass the E_T or η cut.

- In a $t\bar{t}$ dilepton event, one of the leptons can be lost or misidentified; conversely, in a $t\bar{t}$ all-hadronic event, one of the jets can be misidentified as a lepton and another jet lost.

- A $W \rightarrow \tau$ decay can occur, producing a signature which looks like a lepton

Table 4.3: Monte Carlo samples used for modeling the $t\bar{t}$ signal at a variety of m_t values. The numbers listed are the full numbers of events passing the selection cuts; note that our integration does not necessarily use all available events.

Sample name	Description	Events (pretag)	Tagged events
`ttkt60`	PYTHIA $m_t = 160.0$ GeV/c^2	18720	11768
`ttkt62`	PYTHIA $m_t = 162.0$ GeV/c^2	18899	12009
`ttkt64`	PYTHIA $m_t = 164.0$ GeV/c^2	19208	12291
`ttkt66`	PYTHIA $m_t = 166.0$ GeV/c^2	39218	24704
`ttkt68`	PYTHIA $m_t = 168.0$ GeV/c^2	39468	24982
`ttkt70`	PYTHIA $m_t = 170.0$ GeV/c^2	40716	25882
`ttkt72`	PYTHIA $m_t = 172.0$ GeV/c^2	40884	26112
`ttkt74`	PYTHIA $m_t = 174.0$ GeV/c^2	41384	26343
`ttkt75`	PYTHIA $m_t = 175.0$ GeV/c^2	166385	105805
`ttkt76`	PYTHIA $m_t = 176.0$ GeV/c^2	41213	26210
`ttkt78`	PYTHIA $m_t = 178.0$ GeV/c^2	42542	27138
`ttkt80`	PYTHIA $m_t = 180.0$ GeV/c^2	42996	27316
`ttkt82`	PYTHIA $m_t = 182.0$ GeV/c^2	43651	27844
`ttkt84`	PYTHIA $m_t = 184.0$ GeV/c^2	43677	27907

Table 4.4: Samples used for the background modeling.

Sample name	Description	Events (pretag)	Tagged events
btop1w	ALPGEN $W \rightarrow e + b\bar{b} + 1p$	768	428
btop2w	ALPGEN $W \rightarrow e + b\bar{b} + 2p$	19452	8793
btop6w	ALPGEN $W \rightarrow \mu + b\bar{b} + 1p$	686	374
btop7w	ALPGEN $W \rightarrow \mu + b\bar{b} + 2p$	17017	7584
ctop1w	ALPGEN $W \rightarrow e + c\bar{c} + 1p$	739	98
ctop2w	ALPGEN $W \rightarrow e + c\bar{c} + 2p$	19849	2674
ctop6w	ALPGEN $W \rightarrow \mu + c\bar{c} + 1p$	650	108
ctop7w	ALPGEN $W \rightarrow \mu + c\bar{c} + 2p$	18139	2461
stopw2	ALPGEN $W \rightarrow e + c + 2p$	5035	571
stopw3	ALPGEN $W \rightarrow e + c + 3p$	95795	11274
stopw7	ALPGEN $W \rightarrow \mu + c + 2p$	4467	461
stopw8	ALPGEN $W \rightarrow \mu + c + 3p$	83991	9969
ptop3w	ALPGEN $W \rightarrow e + 3p$	2169	85
ptop4w	ALPGEN $W \rightarrow e + 4p$	45196	1917
ptop8w	ALPGEN $W \rightarrow \mu + 3p$	1989	68
ptop9w	ALPGEN $W \rightarrow \mu + 4p$	39868	1672
stop00	MadEvent s-channel single top	5914	3679
stopm0	MadEvent t-channel single top	3169	1853
bhel_ni	Non-isolated electron data	701	143
bhmu_ni	Non-isolated muon data	504	175

Table 4.5: A table illustrating the process by which the contributions for the W + jets background samples are broken down into the individual subsamples by parton multiplicity. Note that the W + light number includes the diboson and Z + jets contributions. The "before discard" column indicates the initial division, while the "after discard" column shows the division after the lowest parton multiplicity bins are discarded.

Process	σ (pb)	Acceptance (%)	Before discard	After discard
W + light (43.9 events)				
$W \to e/\mu + 0p$	1800	0	0	—
$W \to e/\mu + 1p$	225	0	0	—
$W \to e/\mu + 2p$	35.3	0.00015	1.0	—
$W \to e/\mu + 3p$	5.59	0.0076	7.9	8.1
$W \to e/\mu + \geq 4p$	1.03	0.18	35.0	35.8
$W + b\bar{b}$ (46.6 events)				
$W \to e/\mu + b\bar{b} + 0p$	2.98	0.0015	1.2	—
$W \to e/\mu + b\bar{b} + 1p$	0.888	0.026	5.9	6.0
$W \to e/\mu + b\bar{b} + \geq 2p$	0.287	0.54	39.5	40.5
$W + c\bar{c}$ (25.0 events)				
$W \to e/\mu + c\bar{c} + 0p$	5.00	0.00023	0.3	—
$W \to e/\mu + c\bar{c} + 1p$	1.79	0.0052	2.5	2.5
$W \to e/\mu + c\bar{c} + \geq 2p$	0.628	0.13	22.2	22.5
$W + c$ (12.7 events)				
$W \to e/\mu + c + 0p$	17.1	0	0	—
$W \to e/\mu + c + 1p$	3.39	0.00086	0.6	—
$W \to e/\mu + c + 2p$	0.507	0.026	2.8	2.9
$W \to e/\mu + c + \geq 3p$	0.083	0.53	9.3	9.7

+ jets event but which is missing the additional neutrino produced in the subsequent decay of the τ.

- A jet from a single parton can split into two jets in the detector, or jets from two different partons can merge into a single jet.

Figure 4.3 shows one possible way in which such a "bad signal" event can arise. Bad signal events are particularly problematic for our analysis — they do contain some useful information about the top mass (although some information is, of course, lost), but treating them as normal $t\bar{t}$ events will naturally not produce meaningful results. However, these events may still have a peak in the top mass likelihood and reduce the accuracy and precision of our measurement.

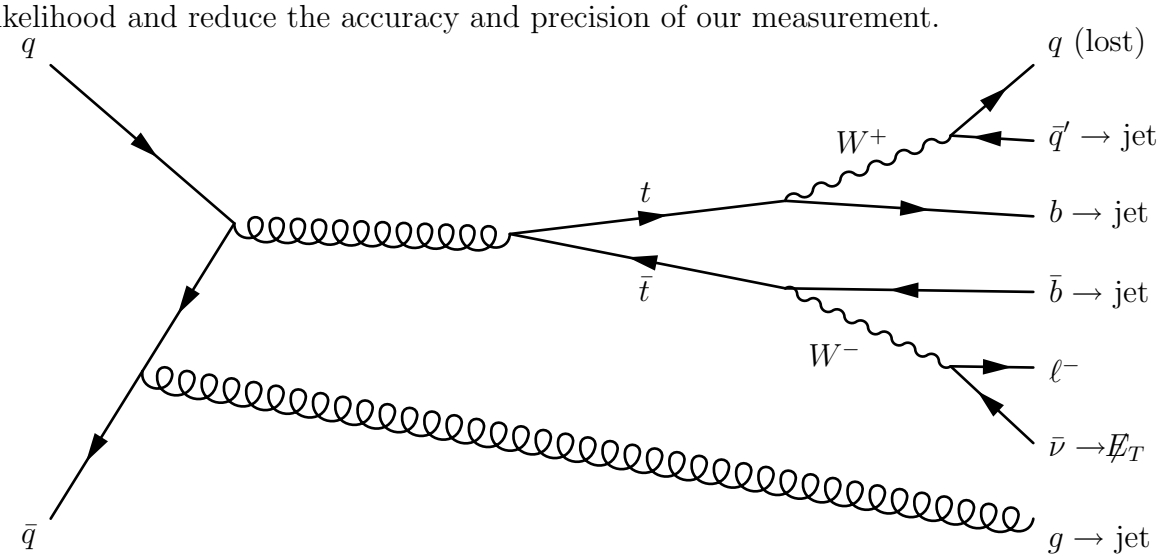

Figure 4.3: Feynman diagram for one possible way in which a "bad signal" event can be created, in this case from initial state radiation (ISR).

Our Monte Carlo $t\bar{t}$ simulated signal events naturally include both good signal and bad signal; to identify good signal events in Monte Carlo, we apply two cuts. First, we examine the tree-level process in the Monte Carlo; a good signal event is required to be lepton + jets at tree level (i.e., one W must decay into an electron or muon and the other into a $q\bar{q}'$ pair). Second, we match the jets with the partons using a χ^2 matching; a good signal event must have a good match. Specifically,

we consider all 24 permutations of assignments of jets to partons, and for each permutation we compute a χ^2 defined by:

$$\chi^2 = \sum_{i=1}^{4} \left(\frac{\eta_j^i - \eta_p^i}{\sigma_\eta} \right)^2 + \left(\frac{\phi_j^i - \phi_p^i}{\sigma_\phi} \right)^2 , \tag{4.7}$$

where η_j^i and ϕ_j^i are the η and ϕ values for the four jets, and η_p^i and ϕ_p^i are the η and ϕ values for the four partons. σ_η and σ_ϕ are the resolution in the η and ϕ measurements, which we take to be both equal to 0.05 in this calculation. This χ^2 is calculated for each permutation, and we require for a good match that the best χ^2 is less than 200 and the difference between the best χ^2 and the second-best χ^2 is at least 150.

Using these criteria, we observe that approximately 35% of events in our $t\bar{t}$ signal Monte Carlo are bad signal. This is clearly a substantial fraction, and so we need to ensure that our method is able to deal with bad signal events in addition to background events.

Chapter 5

Signal Likelihood Calculation

5.1 Introduction to the Matrix Element Method

The idea of a matrix element calculation to measure the top mass has existed since the '80s [40], but is only within the past decade or so that the technique has become computationally feasible. As a consequence, it has become extremely popular in recent years, with many analyses in both CDF and D0 using some form of matrix element technique [6]. However, our analysis includes several features not seen in other matrix element analyses which improve our overall result.

The basic concept behind the matrix element method is quite straightforward. Suppose we had perfect knowledge of all of the particles involved in a given $t\bar{t}$ event, except for the quantity of interest, the top pole mass m_t. In this case, it would be straightforward to evaluate the matrix element for $t\bar{t}$ production and decay for a variety of possible m_t values and thus build a likelihood curve of observing that event in the detector as a function of m_t. The likelihood curves for a set of events can then be multiplied together to obtain an overall likelihood for m_t, and the peak of this total likelihood curve taken as the measured value of m_t.

In practice, of course, a large number of these variables are poorly measured or not measured at all. The way we deal with this problem is simply to integrate over all of the unknown variables, using appropriate priors to reflect our knowledge of the expected values of the variables, as well as "transfer functions" which connect these variables with the known quantities measured in our detector. In its simplest form, we can write this integral as:

$$L(\vec{y}|m_t) = \int |M(m_t, \vec{x})|^2 \mathrm{TF}(\vec{y}|\vec{x}) P(\vec{x}) d\vec{x}, \tag{5.1}$$

where $\vec{x}$ represents the (unmeasured) true parton-level momenta, $\vec{y}$ represents the measured quantities in our detector, M is the matrix element for the $t\bar{t}$ lepton + jets process, P represents the prior distributions of the quantities in $\vec{x}$, and the transfer functions TF give the probability of observing an object with momentum $\vec{y}$ in our detector given a parton with momentum $\vec{x}$.

Of course, the complexity of a $t\bar{t}$ event means that the phase space $\vec{x}$ over which this integral must be evaluated is quite large, and hence the calculations required to obtain the likelihood can be very time-consuming. Other analyses, including the previous version of this analysis [7], made a large number of assumptions about the kinematics of the event to reduce the dimensionality of the integral and hence the computation time required. Of course, these assumptions are not completely perfect, which causes an inevitable loss of resolution in the result. In this analysis, we use a new, quasi–Monte Carlo integration technique, described in Section 5.11, which significantly reduces the time required to integrate an event, allowing us to reduce the number of assumptions made and correspondingly improving the precision and accuracy of our result.

5.2 The JES Parameter

While we have presented the likelihood above as a function of one parameter, there is no reason that the same technique cannot be used to obtain a joint likelihood of two or more parameters. We take advantage of this fact by introducing a second parameter into our likelihood, Δ_{JES}, thus obtaining a two-dimensional likelihood. As we saw in Section 3.4.2, the systematic uncertainty on the jet energy scale (JES) can be quite substantial, and, if we were to convert this into a systematic uncertainty on our top mass measurement by the normal way (shifting the jets by their uncertainty and measuring the resulting difference in the top mass), the resulting systematic uncertainty would be by far the single largest source of uncertainty in our result. However, each $t\bar{t}$ event contains not only information about m_t, but also information about JES; specifically, because the two light quarks produced in the hadronic W decay must add up to a W boson, we can use the known W mass to obtain information about the JES in a given event. Thus, by introducing this additional parameter, we can effectively convert the JES systematic uncertainty into a smaller statistical uncertainty.

The parameter that we use in our 2-D likelihood is Δ_{JES}, which parameterizes the shift in units of the jet systematic uncertainty. For instance, $\Delta_{\text{JES}} = 1$ means that all jets are shifted upwards by $+1\sigma$, which may be different for different jets in an event. (In past versions of the analysis, we used simply a flat multiplicative factor for our JES parameter, but we have found that using Δ_{JES} gives superior performance.) More formally, if a jet with a given p_T and η has a JES systematic uncertainty of $\sigma_{\text{JES}}(p_T, \eta)$, then for a given Δ_{JES} value, the jet p_T is multiplied by $\text{JES} = 1 + \Delta_{\text{JES}}\, \sigma_{\text{JES}}(p_T, \eta)$.

Because the Δ_{JES} is what is known in the statistical literature as a "nuisance parameter", i.e., one that we are not primarily interested in measuring, we need to determine some way to eliminate it in order to obtain a final m_t measurement. This issue is discussed further in Section 7.1. (Note, however, that we do make a final measurement of Δ_{JES} as well as m_t; this helps as a cross-check on the measured Δ_{JES} to make sure that its value is reasonable.)

5.3 Full Signal Likelihood

The full expression for the signal likelihood for a single event is slightly more complicated than the formula shown in Equation 5.1. In its full form, the signal likelihood used in our integration is given by

$$L(\vec{y} \mid m_t, \Delta_{\text{JES}}) = \frac{1}{N(m_t)} \frac{1}{A(m_t, \Delta_{\text{JES}})} \sum_{i=1}^{24} w_i L_i(\vec{y} \mid m_t, \Delta_{\text{JES}}) \tag{5.2}$$

with

$$L_i(\vec{y} \mid m_t, \Delta_{\text{JES}}) = \int \frac{f(z_1) f(z_2)}{FF}\ \text{TF}(\vec{y} \mid \vec{x}, \Delta_{\text{JES}})\ |M(m_t, \vec{x})|^2\ d\Phi(\vec{x}) \tag{5.3}$$

where the variables are defined as follows:

- As mentioned above, $\vec{y}$ represents the quantities measured in our detector (lepton and jet momenta), and $\vec{x}$ represents the parton-level momenta.

- $N(m_t)$ is a global normalization factor to account for the varying cross-section and phase space as a function of m_t.
- $A(m_t, \Delta_{\text{JES}})$ is an acceptance factor to account for the fact that the detector acceptance for $t\bar{t}$ events varies with m_t and Δ_{JES}.
- $\sum_i$ represents the sum over the 24 possible permutations of assigning partons to jets.
- w_i is the weight assigned to each permutation, determined by the presence of b-tags in the event.
- L_i is the likelihood for a single permutation, calculated as shown in Equation 5.3.
- $f(z)$ are the parton distribution functions (PDFs) describing the momenta carried by the incoming partons.
- The flux factor FF is the normalization factor for the PDFs.
- The transfer functions $\text{TF}(\vec{y}|\vec{x}, \Delta_{\text{JES}})$ represent the probability of seeing a jet with momentum $\vec{y}$ given the parton-level momentum $\vec{x}$ and a JES shift Δ_{JES}. (For the sake of conciseness, we only write the TF once, although actually what is meant in Equation 5.3 is the product of the four transfer functions for each jet.) The transfer functions also include a normalization factor discussed in Section 5.6.3.
- $|M(m_t, \vec{x})|$ is the matrix element for $t\bar{t}$ production and decay in the lepton + jets channel.
- $\Phi(\vec{x})$ is the phase space being integrated over, including the appropriate Jacobian for the change of variables into the variables used in our integration and the priors.

These components are discussed in more detail in the following sections.

5.4 Matrix Element

The matrix element is, naturally, the heart of the matrix element method; it gives the amplitude for a $t\bar{t}$ production and decay with parton momenta $\vec{x}$ to occur. We use the matrix element expression developed by Kleiss and Sterling [41], which is a leading-order expression including both $q\bar{q}$ and gg production and includes full spin correlations.

5.5 Parton Distribution Functions

The PDFs $f(z)$ give the probability that the two incoming partons have a given momentum. We use the CTEQ5L PDFs [42], with appropriate weights for q, $\bar{q}$, and g contributions. The flux factor FF essentially acts as a normalization factor for the PDFs; it is given by

$$\mathrm{FF} = \sqrt{(z_1 \cdot z_2)^2 - m_{z_1}^2 m_{z_2}^2} \text{ with } \mathrm{FF} = |z_1||z_2| \text{ when } m_{z_1} = m_{z_2} = 0, \qquad (5.4)$$

where z denotes the z-component of the momentum for the two incoming partons.

5.6 Transfer Functions

The transfer functions are one of the most important components of any matrix element–based analysis, as they connect the parton-level quantities used in the matrix element with the reconstructed quantities observed in our detector. In principle, transfer functions can be used for all of the objects observed in our detector — jets, leptons, and $\not\!\!E_T$. However, we only apply our transfer functions to the jets. We assume that the lepton momentum is well-measured in our detector (essentially, that the transfer function for lepton momentum is a delta function), and we do not use the value of the measured $\not\!\!E_T$ in our likelihood calculation at

all, since it does not contain sufficient information to be useful in our top mass calculation (this is essentially equivalent to setting the TF for $\not\!\!E_T$ to be identically 1).

The transfer functions are built by taking Monte Carlo events, matching the jets to the partons using the χ^2 algorithm described in Section 4.6, and then building distributions of the quantities of interest as described below. (Note that the jets are corrected to level 5, as described in Section 3.4.2.) These then give the probability $\mathrm{TF}(\vec{y}|\vec{x})$ desired. The transfer functions are built separately for b and light jets, as well as in four different $|\eta|$ bins reflecting the different regions of the calorimeter: $|\eta| < 0.2$, $0.2 \leq |\eta| < 0.85$, $0.85 \leq |\eta| < 1.4$, and $|\eta| \geq 1.4$.

For the sake of simplicity, we factorize our transfer functions into separate parts for the magnitude of the momentum and the angles. That is, for a given jet with observed momentum $\vec{y}$ and a parton with momentum $\vec{x}$, the value of the transfer function is given by

$$\mathrm{TF}(\vec{y}|\vec{x}) = \mathrm{TF}_{\mathrm{mom}}(p_j|p_p) \cdot \mathrm{TF}_{\mathrm{ang}}(\eta_j, \phi_j|\eta_p, \phi_p), \tag{5.5}$$

where the j subscripts indicate the jet quantities and the p subscripts the parton quantities.

5.6.1 A Note on Proto-Jets

Before proceeding into the discussion of transfer functions, it is important to first understand how generators such as HERWIG and PYTHIA carry out their hadronization process. First, a "tree-level" process is generated, in which the quarks are on their mass shell (negligibly small mass for light quarks, and 4.95 GeV/c^2 for b quarks). The expression for our matrix element also is designed for this "tree-level" calculation. Then, these quarks are adjusted in such a way that they acquire mass (while conserving 4-momentum) to take into account the parton shower effects. We call these adjusted partons, which later hadronize, "proto-jets", as they are

the actual partons that produce the final state jets.

It is, of course, possible to build our transfer functions between the tree-level partons and the observed jets, ignoring the proto-jet level entirely. However, we find in our studies that the properties of the final state jets depend significantly on the proto-jet properties; specifically, larger proto-jet masses tend to produce wider distributions in the resulting jet momenta. Consequently, taking this fact into account will result in a more accurate modeling in the transfer function.

As a result, we build our transfer functions between the proto-jet momenta and the observed jets as a function of the proto-jet mass. (The proto-jet mass then becomes another integration variable, as discussed in Section 5.10.) To recover the tree-level partons from the proto-jets, we reverse the procedure that HERWIG uses to create proto-jets from partons, so we can thus recover the tree-level kinematics from the proto-jets in our integration to evaluate the matrix element.

5.6.2 Momentum Transfer Functions

The momentum transfer functions are constructed as a function of the variable $u = p_T^{\text{jet}}/p_T^{\text{parton}}$ and are dependent on the variables p_T^{parton} and the mass of the proto-jet m described above. A three-dimensional distribution in u, p_T^{parton}, and m is thus constructed and smoothed using local orthogonal polynomial expansion [43], which can be viewed as an extension of the standard kernel density estimation (KDE) technique [44] to polynomials of higher degree. The transfer function is then saved on a grid constructed to contain a reasonable number of points in each bin. For speed purposes, the transfer function values are not parameterized, but rather saved as a lookup table and then interpolated as necessary during the integration. Figure 5.1 shows a sample momentum transfer function for a light quark.

While the transfer functions are calculated for $\Delta_{\text{JES}} = 0$, they can naturally be used for other Δ_{JES} values as follows. The magnitude of the momentum ob-

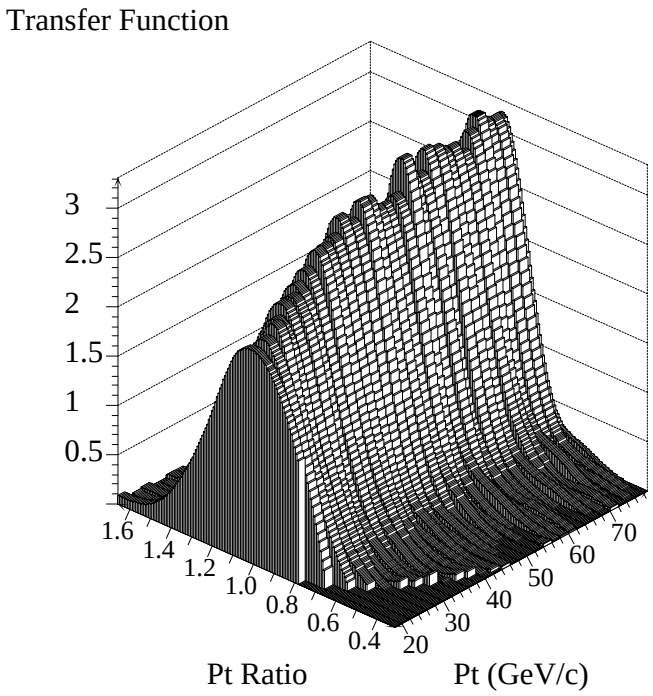

Figure 5.1: Sample momentum transfer function for a given fixed proto-jet mass as a function of the p_T ratio and the parton p_T. The sharp edge at low parton p_T and ratio values is due to the 20 GeV cut for jets.

served in the detector y can be related to the momentum used in calculating the transfer functions y_0 by $y_0 = \mathrm{JES} \cdot y$, where $\mathrm{JES} = 1 + \Delta_{\mathrm{JES}}\, \sigma_{\mathrm{JES}}(p_T, \eta)$. Thus, the probability distribution of y is given by

$$p(y) = p(y_0)\frac{dy_0}{dy} = p(y \cdot \mathrm{JES})(\mathrm{JES} + y\frac{d\,\mathrm{JES}}{dy}). \tag{5.6}$$

The latter term becomes important when the JES depends on p_T, as it does in this analysis (note that if JES were simply a flat multiplicative factor, this term would be zero).

5.6.3 Transfer Function Normalization

In order for our probability calculation to be sensibly normalized, the transfer functions must be themselves normalized.[1] That is,

$$\int_Y \mathrm{TF}(\vec{y}|\vec{x})d\vec{y} = 1 \text{ for all } \vec{x}, \tag{5.7}$$

where Y indicates the space of all possible $\vec{y}$ values. Unfortunately, this is a very difficult condition to satisfy, since applying our cuts to the Monte Carlo samples mean that the whole region of Y is not covered. (Even if we took our Monte Carlo sample without any cuts, jets simply cannot be reconstructed at very low momenta, so our sample would still not cover the full space Y.) There are two possible approaches to addressing this problem. One is to extrapolate the transfer functions into the region below our cuts. However, this results in the transfer functions being very sensitive to the way the extrapolation is performed — while the transfer functions are never actually used for momenta below the cuts, by definition, the more of the transfer function that lies in the extrapolated region, the less that will actually lie in the observed region.

[1] Of course, like all other components of the integration, the transfer functions can be multiplied by an overall constant factor without any effect on the likelihood; what matters is the relative normalization for different values of parton p_T and mass.

The approach used in this analysis, thus, is to build the transfer functions only in the subspace Y' which represents values of $\vec{y}$ passing all cuts. In this case, to ensure correct normalization of the transfer functions, we must impose the condition

$$\int_{Y'} \mathrm{TF}(\vec{y}|\vec{x}) d\vec{y} = \epsilon(\vec{x}), \tag{5.8}$$

where $\epsilon(\vec{x})$ represents the probability for a parton with momentum $\vec{x}$ to produce a jet that passes all cuts. We thus build this probability $\epsilon(\vec{x})$ from Monte Carlo as a function of the proto-jet p_T and mass and smooth it using local quadratic logistic regression, an extension of local linear logistic regression as described in [45].

Figure 5.2 shows this normalization for a light quark over a range of proto-jet p_T and mass values.

5.6.4 Angular Transfer Functions

The angular transfer functions are built as a function of $\Delta\eta = \eta_{\mathrm{jet}} - \eta_{\mathrm{parton}}$ and $\Delta\phi = \phi_{\mathrm{jet}} - \phi_{\mathrm{parton}}$. The distribution is built as a function of $\Delta\eta$, $\Delta\phi$, and m and then smoothed into a probability distribution using KDE.

Figure 5.3 shows a sample angular transfer function for a given η bin and m value.

5.7 Normalization

The normalization factor ($N(m_t)$ in Equation 5.2) compensates for the fact that the $t\bar{t}$ cross-section is not constant as a function of m_t; along with the acceptance term, it is required to ensure that the overall likelihood is properly normalized. We calculate this normalization factor using a Monte Carlo integration which integrates the matrix element, PDFs, and flux factor (essentially, all the components of Equation 5.3 except for the transfer functions) over the phase space defined by the

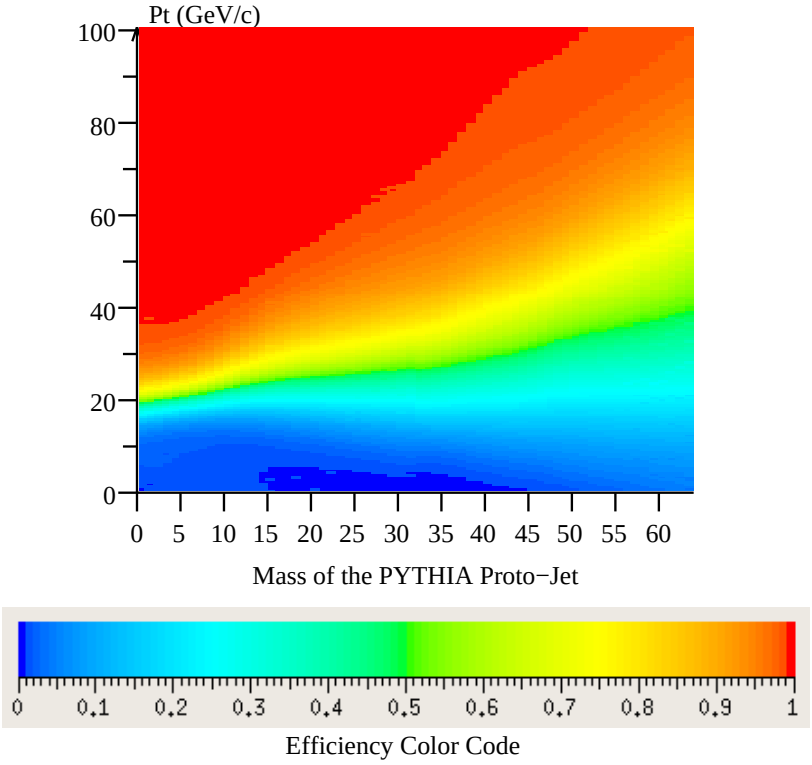

Figure 5.2: Transfer function normalization as determined by the efficiency to reconstruct a jet passing the E_T cut of 20 GeV as a function of the proto-jet p_T and mass.

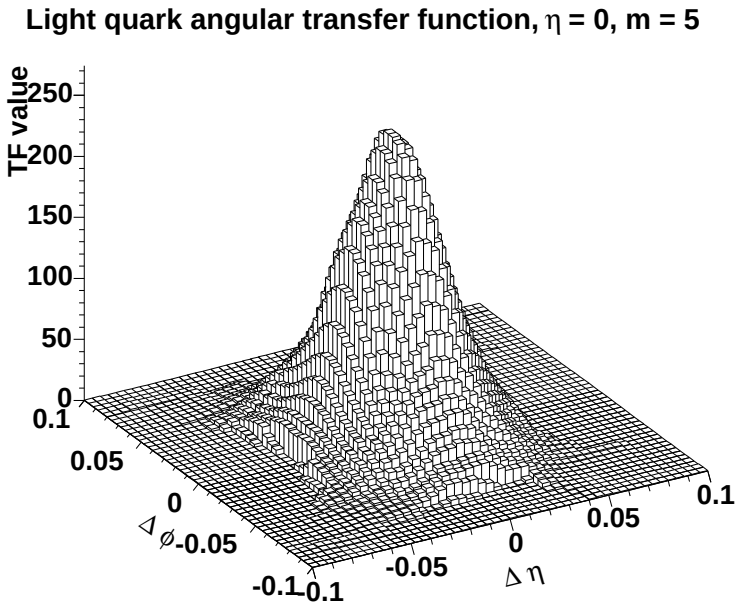

Figure 5.3: Sample angular transfer function as a function of $\Delta\eta$ and $\Delta\phi$ with the proto-jet mass fixed at 5 GeV/c^2 in the $\eta = 0$ bin.

two initial state and six final state particles. As the normalization factor can be multiplied by an arbitrary constant without changing its effect on the likelihood, we do not keep track of all constant factors in this integration, but we can compare it to the $t\bar{t}$ cross section from HERWIG by scaling it appropriately. This comparison is shown in Figure 5.4; we note that there is good agreement between our result and the HERWIG value. Since our calculation does not include everything included in HERWIG (e.g. ISR/FSR), we do not expect the agreement to be perfect.

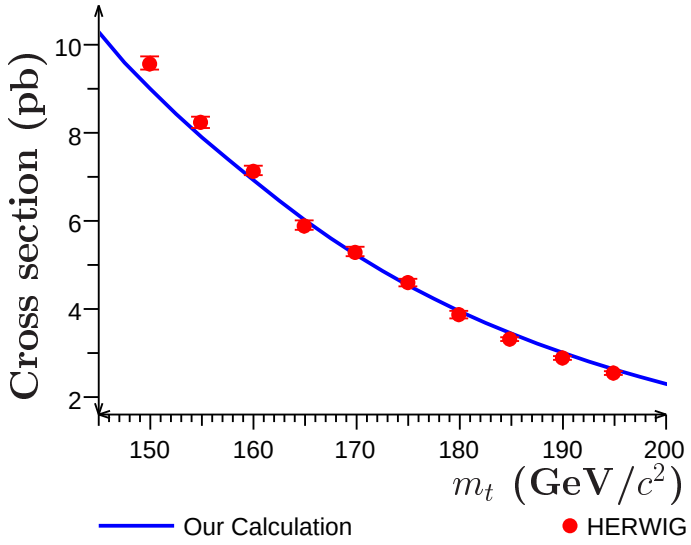

Figure 5.4: $t\bar{t}$ cross-section produced by our calculation (blue line) with appropriate scaling to compare with the $t\bar{t}$ cross-section in HERWIG (red dots).

5.8 Acceptance

The acceptance factor $A(m_t, \Delta_{\text{JES}})$ is the other term (along with the normalization) required to correctly normalize the overall likelihood. The acceptance simply gives the probability that an event with a given m_t and Δ_{JES} value passes the selection cuts and is observed in our detector.

Perhaps the most obvious way to calculate this term is to simply take the simulated Monte Carlo samples at a wide variety of top masses m_t and calculate the acceptance for each. However, this approach is less than ideal for our analysis for two reasons. First, there is the simple matter of limited statistics available in the simulated samples, which can result in statistical fluctuations in the calculated acceptance. The second, more serious, issue is that the simulated samples include a large number of processes (basically, the cases that lead to "bad signal" events discussed in Section 4.6) which are not included in our model, and so the mismatch between the simulation and our model means that applying an acceptance term derived from the simulation may not lead to a correct normalization for our likelihood.

Instead, we generate our own set of Monte Carlo events using HERWIG in which the full detector simulation is not performed — only the parton-level description of the event is created. Because the full detector simulation is quite lengthy, this allows us to create much larger samples of events, and we can ensure that the events generated do adhere to our model. We then take the partons created and smear them by our transfer functions to obtain objects which should behave like final-state jets, and apply a set of cuts to these smeared partons designed to recreate the effect of the normal cuts:

- The four smeared partons must have a $E_T > 21$ GeV and $|\eta| < 2$.[2]
- The lepton must have a $p_T > 20$ GeV/c and $|\eta| < 1$.
- The smeared partons must be separated from each other by a minimum ΔR of 0.5, as if two partons were closer than that distance, they would merge into a single final-state jet.

[2]The cut is 21 GeV instead of the 20 GeV used in our actual analysis for historical reasons, as this calculation dates from an older version of our analysis which had slightly different cuts. The difference is negligible, however.

- The lepton must be separated from each of the partons by a minimum ΔR of 0.5, to account for the lepton isolation cut.

- The difference between the visible transverse momentum and the $t\bar{t}$ transverse momentum must be at least 20 GeV, to account for the $\not\!\!E_T$ requirement. The visible transverse momentum is calculated by summing the $\vec{p_T}$ of the four smeared partons and the lepton.

We then calculate the acceptance by simply dividing the number of events which pass the above cuts as a function of m_t and $\Delta_{\rm JES}$ by the total number of events in the sample. Figure 5.5 shows the resulting 2-D acceptance.

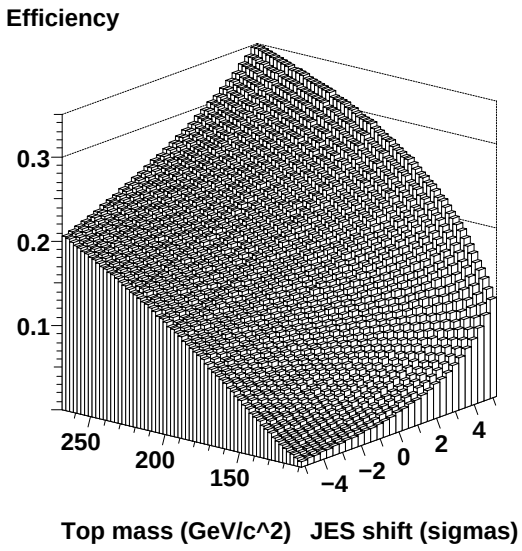

Figure 5.5: Acceptance term used in our integration as a function of m_t and $\Delta_{\rm JES}$.

5.9 Permutation Weights

The likelihood is calculated separately for each of the 24 possible permutations of assignments of jets to partons; these individual likelihoods are then summed to get a total likelihood for the event. Each permutation is weighted by a weight w_i determined by how the permutation corresponds to the b-tagging information in the event as follows:[3]

Each jet in an event has a certain probability of being tagged, which we denote as P_b for a b-jet, P_c for a c-jet, and P_l for a light (u, d, or s) jet. These probabilities are parameterized as functions of the jet E_T and η by fitting the efficiencies for tagging b-jets and mistags determined by the CDF b-tagging group with fifth-degree polynomials in E_T and η. These fits are shown in Figure 5.6. The probability that a charm jet is tagged is taken to be 22% of the corresponding b-tag probability (i.e., $P_c(E_T, \eta) = 0.22 P_b(E_T, \eta)$).

For a given permutation, we compute the tagging probability for each parton. Assuming that the two decays $W \to u\bar{d}$ and $W \to c\bar{s}$ each occur 50% of the time, this tagging probability P_{tag} is thus P_b for a b-parton, $0.5P_c + 0.5P_l$ for a u-type parton, and P_l for a d-type parton. Then, we add a weight for each jet equal to P_{tag} if the jet assigned to that parton in this permutation is tagged, or $1 - P_{\text{tag}}$ if the jet is not tagged. The overall weight for the permutation is then the product of the four individual weights for each jet-parton pair.

[3] In earlier versions of this analysis, we only considered permutations where the b partons were assigned to b-tagged jets; this reduced the number of permutations we needed to consider. However, this produced very poor results in $t\bar{t}$ events which contained a mistagged jet, so we adopted this approach.

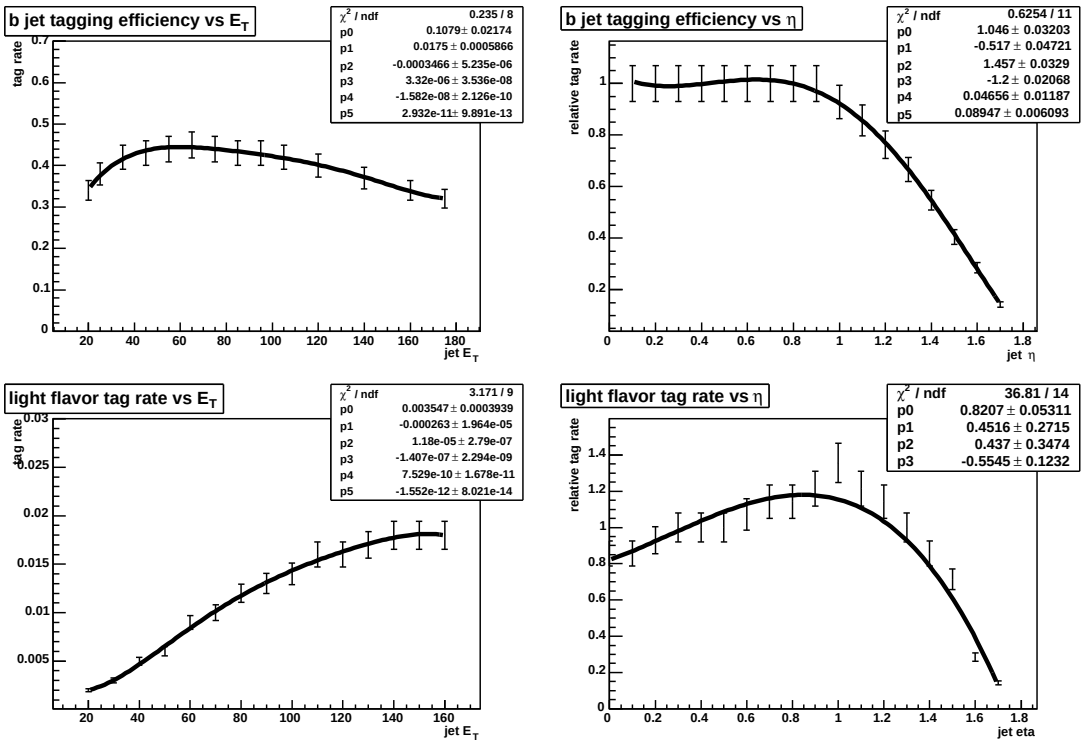

Figure 5.6: Fits used for parameterizing the b-tag and mistag probability as a function of E_T and η. The total b-tag or mistag probability is obtained by multiplying the efficiency as a function of E_T by the relative efficiency as a function of η.

5.10 Phase Space and Variables

5.10.1 Integration Variables

A total of 32 variables is needed to fully describe the kinematics of a $t\bar{t}$ production and decay event, four for each of the two incoming partons and six final state particles (two light quarks, two b quarks, the lepton, and the neutrino). This number is reduced by the following constraints and assumptions:

- Energy-momentum conservation allows us to eliminate four variables.
- We assume that the charged lepton and neutrino mass are perfectly known (i.e., that their uncertainties are negligible, as they are much smaller than other sources of uncertainty in our measurement).
- We assume that the momentum for the charged lepton is perfectly measured.
- We assume that the masses of the initial partons are perfectly known.
- Finally, we neglect the effects of the individual transverse momenta of the initial partons, and model only the transverse momentum of the total $t\bar{t}$ system. This eliminates two more degrees of freedom (since $\vec{p_T}$ is a two-dimensional vector).

After all these eliminations, we are left with a 19-dimensional phase space. In our previous analysis [7], we were forced to make more assumptions to further reduce the dimensionality of the integral to seven dimensions. The gain in integration speed provided by Quasi–Monte Carlo integration [46], however, allows us to integrate over all 19 remaining dimensions.

The choice of which variables to use is determined by speed considerations. Because we use importance sampling techniques to improve the numerical convergence, we wish to choose integration variables which allow us to apply this

technique in a straightforward manner. Consequently, we select the following 19 variables:

- $M^2_{t,\mathrm{lep}}$ and $M^2_{t,\mathrm{had}}$, the squared mass of the leptonic and hadronic top quarks[4]
- $M^2_{W,\mathrm{lep}}$ and $M^2_{W,\mathrm{had}}$, the squared mass of the leptonic and hadronic Ws
- $\beta \equiv \log \frac{p_q}{p_{\bar{q}'}}$, the logarithm of the ratio of the magnitudes of the momenta of the two partons from the hadronic W
- $\vec{p_T}(t\bar{t})$, the transverse momentum of the $t\bar{t}$ system
- $m_{1\ldots4}$, the masses of the four proto-jets as defined in Section 5.6.1
- $\eta_{1\ldots4}$ and $\phi_{1\ldots4}$, the angles of the four proto-jets

Quasi–Monte Carlo integration is used for 18 of these variables. The leptonic W mass $M^2_{W,\mathrm{lep}}$ requires special treatment to avoid phase space singularities, and so it is integrated over a fixed grid; this procedure is discussed more fully in Appendix B.

In order to create a sensible final likelihood, we must include the appropriate prior distributions for our integration variables. These priors are not included explicitly; rather, we use the prior distributions as our distributions for importance sampling, which thus implicitly includes the effects of the priors in our integration. The priors used are as follows:

- For the top and W masses, a Breit-Wigner distribution centered at the top pole mass m_t for the top quark and at the known W mass for the W boson is used.

[4]It is important not to confuse these with m_t, the top pole mass. The value of m_t is fixed at a given integration point and scanned over to calculate a total $L(m_t, \Delta_{\mathrm{JES}})$ curve; M_t, on the other hand, represents the top mass in a particular event and is integrated over.

- For the β parameter ($\equiv \log p_q/p_{\bar{q}'}$), a flat distribution from -1.7 to 1.7, centered around the value of the same ratio for the assigned jets, is used.

- For the $\vec{p_T}(t\bar{t})$ distribution, we use a distribution built from Monte Carlo samples at different masses.

- For the proto-jet masses, the distribution of proto-jet masses from Monte Carlo is used.

- For the proto-jet angles, no explicit prior is used; rather, the angular transfer function provides the distribution we sample in.

For a given integration point, values for the integration variables are selected, and then the kinematic equations to reconstruct the rest of the event are solved. Since we now allow for arbitrary jet masses, the general solution for the leptonic b jet momentum requires solving a 8th-order polynomial, which we solve using the Jenkins-Traub algorithm [47]. Because more than one solution may be possible, we compute the integrand for up to four different solutions and sum the likelihood over these different solutions.[5]

5.10.2 Phase Space and Variable Transformation

By Fermi's Golden Rule, since we are obtaining an amplitude for a scattering process, we need to include the appropriate factors for the phase space available to the final products of the process; in our case (neglecting constants), this is simply:

$$\prod_{i=1}^{4} \frac{d^3\vec{p}_i dE_i}{(2\pi)^3} \cdot \frac{d^3\vec{p}_\ell}{(2\pi)^3 2E_\ell} \cdot \frac{d^3\vec{p}_\nu}{(2\pi)^3 2E_\nu}, \tag{5.9}$$

[5]The limit of four is chosen simply to reduce the integration time; cases where there are more than four real solutions are rare. If there are more than four solutions, then we take the four in which the solved b momentum is closest to the momentum for the assigned jet; because the transfer function would have a small value for values far away from the observed value, these would only contribute a very small likelihood in any case.

where $\vec{p}_i$ and E_i represent the three-momentum and energy of the four partons.

Note that the phase space factor for the partons is slightly different from the more standard form used for the charged lepton and neutrino. This is because the lepton and neutrino mass are taken as constant, while the parton masses are allowed to vary in our integral.

Because of our previous assumption that the lepton momentum is well-measured, the integration over $\vec{p}_\ell$ is eliminated. We can also rewrite the three-momenta in terms of their magnitude p_i and angles Ω_i:

$$\prod_{i=1}^{4} \frac{p_i^2 dp_i d\Omega_i dE_i}{(2\pi)^3} \cdot \frac{d^3\vec{p}_\nu}{(2\pi)^3 2E_\nu}. \tag{5.10}$$

At this point, we need to transform the variables $p_1 \ldots p_4$, $E_1 \ldots E_4$, $p_{\nu x}$, $p_{\nu y}$, and $p_{\nu z}$ into the variables actually used in our integration. This requires computing the Jacobian matrix; for the full form of this Jacobian, see [48].[6]

5.11 Quasi–Monte Carlo Integration

A popular way for performing integration in physics problems is Monte Carlo integration, in which points in the integration space are randomly selected.[7] Monte Carlo integrations typically exhibit $O(1/\sqrt{N})$ convergence, where N is the number of integration points. In our method, we use, for the first time in a top mass measurement, a different method of integration, Quasi–Monte Carlo (QMC) integration [49]. QMC integration employs quasi-random sequences, which results in an overall improved convergence, thus allowing us to integrate over a more complicated phase space and still compute the integral in a reasonable amount of time.

[6] This Jacobian only includes the first seven variables in Section 5.10.1; the angles are already present in the phase space factor in Equation 5.10, but changing from dE to dm requires an appropriate term.

[7] More precisely, of course, any random numbers generated by a computer are pseudo-random.

In QMC, as in regular Monte Carlo integration, we approximate our integral by evaluating it at a series of points; that is,

$$\int f(\vec{x})d\vec{x} \approx \frac{1}{N}\sum_{i=1}^{N} f(\vec{x}_i). \tag{5.11}$$

(For the sake of simplicity in the following discussion, we assume the integration space has been transformed into the s-dimensional unit cube I^s.) What distinguishes QMC from regular Monte Carlo integration is how the set of points $\vec{x}_i$ is chosen. In regular Monte Carlo integration, these points are chosen randomly, which naturally means that there will be some regions of phase space which have more points than average and some regions with fewer points than average. In QMC integration, we use a quasi-random sequence of points to determine $\vec{x}_i$, which is a deterministic sequence defined by its uniform coverage [50].

A quasi-random sequence has a low discrepancy. There are many ways to define discrepancy, but intuitively speaking, a low-discrepancy sequence is one in which equal subvolumes of the integration space contain as close to equal numbers of points as possible. One common measure of discrepancy is the star discrepancy:

$$D^* = \sup_{B\in J^*}\left|\frac{1}{N}(\text{number of points of } \vec{x}_i \text{ in } B) - \text{volume of } B\right| \tag{5.12}$$

where J^* is the family of all subintervals of I^s of the form $\prod_{k=1}^{s}[0, u_i)$ for $u_i \in [0, 1)$ (that is, the set of all subrectangles with one vertex at the origin).

Some popular low-discrepancy sequences are the Halton, Faure, and Sobol sequences [49, 50]; in this analysis, we use an 18-dimensional Sobol sequence as our pseudo-random sequence. Figure 5.7 shows a sample Sobol sequence compared to a random sequence.

Note that, in high dimensions, the Sobol sequence can exhibit patterns in 2-dimensional cross-sections, so we also apply an additional "scrambling" which randomizes the digits of the Sobol sequence to reduce these patterns while preserving its low-discrepancy properties [51].

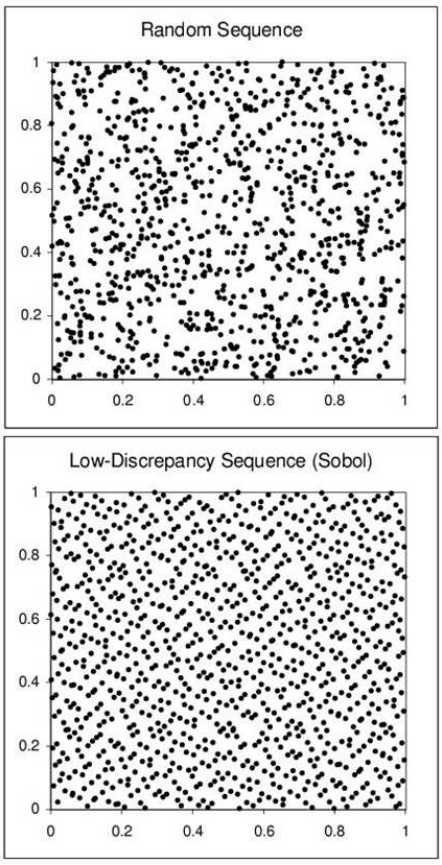

Figure 5.7: A comparison of a 2-dimensional random sequence (top) and a Sobol quasi-random sequence (bottom).

Given the application of a low-discrepancy sequence, the Koksma-Hlawka inequality [50] gives the following bound on the error between the quasi–Monte Carlo integration value and the actual value:

$$b_{s,N}(\log N)^s/N \text{ for } N \geq 2, \tag{5.13}$$

where $b_{s,N}$ is a bounded constant for $N \geq 2$. For large dimensions s, then, QMC integration can exhibit significantly better convergence than ordinary Monte Carlo integration procedures. As this is an upper bound, the actual convergence of QMC integration can be better than implied by the above equality [52, 53]. As an illustration, we have performed a small study of QMC integration performance applied to the calculation of the muon lifetime in first-order perturbation theory, where we can compare it to the known theoretical answer. Figure 5.8 shows the results of the study; it appears in this case that the integral converges roughly as $O(1/N)$.

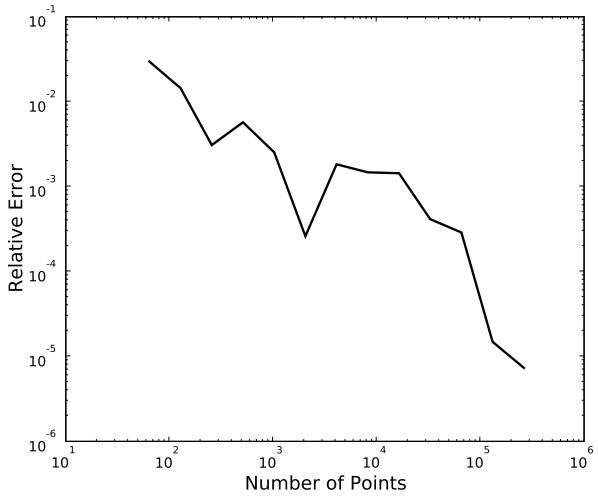

Figure 5.8: QMC integral convergence applied to the problem of the muon lifetime. A Sobol sequence is used to perform the 5-dimensional integration, and the result is compared with the exact theoretical answer.

5.12 Distributed Integration Framework

The actual integration is carried out using a distributed method similar (albeit on a smaller scale) to that used by analyses such as SETI@home or Folding@home. The framework consists of three parts:

- The **director** oversees the entire integration. All masters and workers connect to the director first; the director assigns workloads and workers to the masters as necessary.

- The **masters** are responsible for workloads, typically 1000-event chunks. The director assigns workers to masters; the master, in turn, assigns and transmits events to the workers, and receives the completed integration results from the workers in turn. When the master finishes a workload, it returns to the director to obtain a new one.

- The **workers** run on the CDF Central Analysis Farms (CAFs) and are responsible for the actual integration. Workers first contact the director, which assigns them to a master; the worker then receives an event from the master and performs the integration. When the integration concludes, the worker sends the results to the master and receives a new event.

The worker carries out the likelihood integration on a grid consisting of 32 points in 2 GeV/c^2 intervals in m_t from 143 to 207 GeV/c^2, and 26 points in 0.4 σ intervals in Δ_{JES} from -5.2 to 5.2 σ. The integration proceeds until it reaches a predetermined target precision of 0.1 relative precision, or until it reaches a time limit of 2 hours elapsed time. Typically about 30% of events time out before reaching the integration precision, and the mean time required to integrate a single event is $\sim$ 80 minutes/event.

The workers use a number of techniques to speed up the integration:

- Before beginning the integration, points on the leptonic side which will not result in solvable kinematics are identified and excluded.

- As the integration proceeds, if a permutation is identified as having a significantly lower likelihood, the worker stops working on that permutation to focus on the higher-likelihood permutations.

- Because the Δ_{JES} value affects only the jet-level quantities, not the parton-level quantities, the only part of the integrand affected by Δ_{JES} is the transfer functions. Thus, at a given point, we only need to compute the rest of the integrand once, and then compute the transfer function value for each Δ_{JES} value, thus allowing us to process many Δ_{JES} points in a relatively quick time.

If a worker fails to return an integrated curve, usually due to worker node crashes or a worker running out of CAF time without being able to report that fact to the master, then the master will retry the event up to a maximum of three tries. Typically, $\sim 0.5\%$ of events will fail to be processed, but this is small enough that we neglect it. (For data, we run on the more reliable CAFs, so we get 100% of events returned in this case.)

Due to time considerations, we do not integrate every single Monte Carlo event in our samples. Rather, for an input $t\bar{t}$ sample, we typically integrate either 8k or 16k events, depending on time availability and the size of the original sample.

After the integration has concluded, we apply the acceptance and normalization terms to the resulting curves. The curves are also smoothed by convoluting with a Breit-Wigner and interpolated onto a finer grid with 320 points in m_t at 0.2 GeV/c^2 intervals and 150 points in Δ_{JES} at 0.0667 σ intervals; studies have been run to optimize these values.

Chapter 6

Background Handling

6.1 Motivation

So far, we have been considering events under the assumption that they are really $t\bar{t}$ lepton + jets events. Of course, as discussed in Section 4.2, a certain fraction of non-$t\bar{t}$ events will pass our selection cuts, and, as discussed in Section 4.6, even some $t\bar{t}$ events are "bad signal" and hence will not produce meaningful results when analyzed using the technique described in the preceding chapter. Consequently, we need a method to deal with these events.

In principle, it is possible to construct a background likelihood in much the same way as the signal likelihood; this is the approach taken by some other matrix element analyses [54]. However, such an approach requires a large amount of time to develop the background likelihood calculation. For the sake of simplicity, then, we adopt a simpler approach, consisting of three parts.

First, we construct a discriminant which uses an artificial neural network to identify the probability of an event being background given some kinematic and topological variables in the event. Secondly, we use the discriminant to estimate the portion of the total likelihood contributed by background events, and then subtract off this part to recover the likelihood from signal events. Finally, we apply an additional cut to our sample on the peak log-likelihood value of individual event curves; this removes a significant fraction of bad signal and background events from our sample while removing only a small fraction of good signal events, resulting in a pronounced improvement in our resolution.

These three steps are described in the following sections, 6.2–6.4.

6.2 Neural Network Discriminant

The first step is to create a variable to distinguish between signal and background events. To this end, we construct a neural network which uses various kinematic and topological variables describing the event to create a probability that the

event is background. We train the neural network with ten inputs, summarized in Table 6.1: the p_T for each of the four leading jets; the E_T of the lepton; the missing E_T, $\not\!\!E_T$; H_T, the scalar sum of the jet transverse energies, lepton transverse energy, and missing E_T; and three variables describing the shape of the event: the aplanarity, defined as 3/2 the smallest eigenvalue of the momentum tensor $\Theta_{ab} = \sum_i p_a^i p_b^i / \sum_i |\vec{p_i}|^2$, where a and b are indices for the three axes x, y, and z; $D_R = \Delta R_{ij}^{\min} \cdot \min(p_z^{(i,j)})/p_T^\ell$, where $\Delta R_{ij}^{\min}$ is the smallest ΔR between any pair of jets; and $H_{\rm TZ} = \sum_{i=2}^4 |p_T^i| / (\sum_{i=1}^4 |p_z^i| + |p_z^\ell| + |p_z^\nu|)$, the ratio of the scalar sums of the transverse momenta for the three non-leading jets to the longitudinal momenta of all six final objects in the event. The network itself is built using the JETNET neural network package, version 3.5 [55] with the RootJetnet interface to ROOT.

Table 6.1: The ten variables used in the neural network discriminant. The first seven describe the kinematics of the event, while the last three are topological variables describing the shape of the event.

Variable	Definition
$p_T^{1\cdots4}$	p_T for each of the four leading jets
E_T^ℓ	Charged lepton E_T
$\not\!\!E_T$	The missing E_T
H_T	Scalar sum of p_T for four jets, lepton p_T, and $\not\!\!E_T$
Aplanarity $= 3/2 \cdot Q_1$	Q_1: smallest eigenvalue of the momentum tensor
$D_R = \Delta R_{ij}^{\min} \cdot \min(p_z^{(i,j)})/p_T^\ell$	$\Delta R_{ij}^{\min}$: smallest ΔR between any pair of jets
$H_{\rm TZ}$	Ratio of transverse to longitudinal momenta

We train the neural network using $t\bar{t}$ signal events with a mass of 170 GeV/c^2 and $W + b\bar{b} + 2p$ events for the background; we then cross-check the neural network

with other signal masses and background types to make sure that the output shape is not significantly dependent on the signal mass present or background composition.

Figure 6.1 shows the neural network output for a variety of different samples. Now, given an event with a neural network value of q, we can compute the background fraction for this event as $f_{\mathrm{bg}}(q) = B(q)/(B(q)+S(q))$, where the background and signal distributions are normalized to their overall expected fractions. Figure 6.2 shows the final distributions, including their normalization, used for calculating f_{bg}.

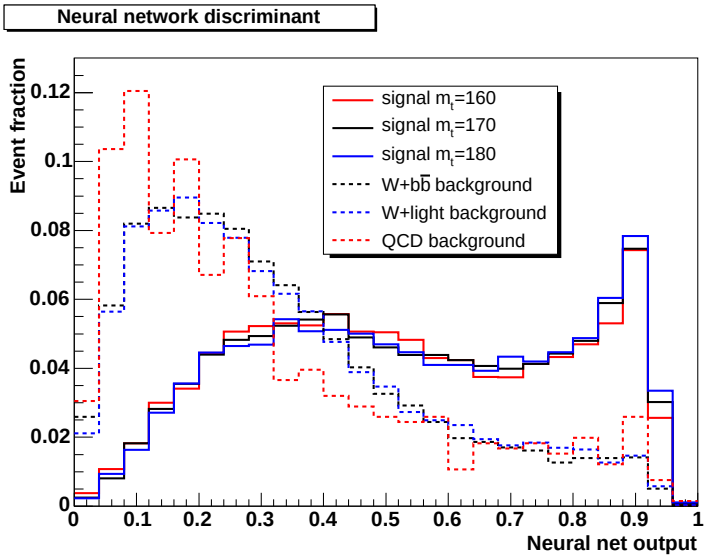

Figure 6.1: The distributions of our discriminant variable for signal and background Monte Carlo events. The solid lines indicate signal events at various masses, while the dashed lines indicate various types of backgrounds.

6.3 Modified Likelihood

When we combine the likelihood curves from all events, our total likelihood will naturally contain likelihoods from signal events and background events. However,

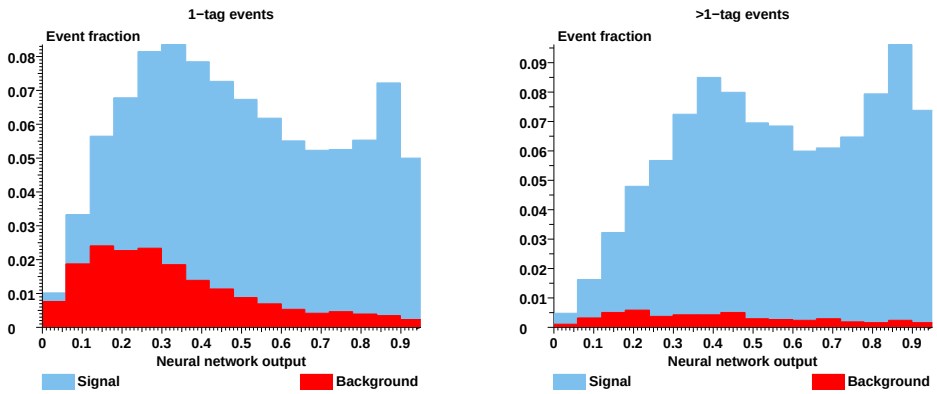

Figure 6.2: Background templates for single-tag (left) and multiple-tag (right) events, showing the calculation of $f_{\rm bg}(q)$. Note that the signal is stacked on top of the background.

only the signal events will contain meaningful information about m_t. Thus, we want to remove the contribution due to background events from the total likelihood to recover the likelihood from signal events. (Note that there is not a separate matrix element for background processes — the likelihood for all events, signal and background, is calculated under the assumption that the event is a $t\bar{t}$ signal event.) Consequently, we compute from Monte Carlo simulation the average likelihood for background events and subtract out the expected contribution due to background events from the total likelihood:

$$\log L_{\rm mod}(m_t, \Delta_{\rm JES}) = \sum_{i \in \text{events}} [\log L(\vec{y_i}|m_t, \Delta_{\rm JES})] - n_{\rm bg} \log \overline{L_{\rm bg}}(m_t, \Delta_{\rm JES}), \quad (6.1)$$

where $L_{\rm mod}$ is the modified total likelihood for a given set of events, $L(\vec{y_i}|m_t, \Delta_{\rm JES})$ is the likelihood for an individual event, $n_{\rm bg}$ the expected number of background events, and $\overline{L_{\rm bg}}(m_t, \Delta_{\rm JES})$ is the average likelihood for a background event as computed in Monte Carlo simulation. This calculation is performed separately for 1-tag and >1-tag events, as the background fractions and $\overline{L_{\rm bg}}(m_t, \Delta_{\rm JES})$ are different for the two subsamples.

We can rewrite Eq. 6.1 in terms of the individual per-event background fraction to obtain our final modified likelihood L_{mod}:

$$\log L_{\mathrm{mod}}(m_t, \Delta_{\mathrm{JES}}) = \sum_{i \in \mathrm{events}} [\log L(\vec{y_i}|m_t, \Delta_{\mathrm{JES}}) - f_{bg}(q_i) \log \overline{L_{\mathrm{bg}}}(m_t, \Delta_{\mathrm{JES}})], \quad (6.2)$$

where $f_{\mathrm{bg}}(q_i)$ is the background fraction given the discriminant variable q_i for a given event, as derived in the previous section. Equations 6.1 and 6.2 are equivalent if the number of background events in the data is equal to the expected background contribution. However, the advantage of using Eq. 6.2 is that if there are more or fewer background-like events in our data than expected, the average value of $f_{\mathrm{bg}}(q_i)$ will be correspondingly higher or lower, thus compensating for the difference.[1]

This method has the advantage of simplicity; the disadvantage is that, for any given sample of events, the average background likelihood in that sample will naturally show some fluctuation around $\overline{L_{\mathrm{bg}}}(m_t, \Delta_{\mathrm{JES}})$, so our subtraction procedure will not be completely accurate. This results in a slight increase in our pull widths, as discussed in Section 7.6.

6.4 Likelihood Cut

As mentioned in Section 4.6, bad signal events make up approximately 35% of our expected $t\bar{t}$ signal. While the above method is designed to handle background events, it does not deal with bad signal events, so we need a separate method to handle them. We observe that, if we look at the values of the peaks of the log-likelihood of the curves, these peaks tend to be lower for bad signal and background

[1] Originally, the motivation for introducing the per-event background discriminant f_{bg} was that it could be used to modify the likelihood of individual curves so that they would contribute less to the total likelihood if f_{bg} were greater. However, in our studies, we found that the improvement by decreasing the contribution of real background events was outweighed by the loss in resolution caused by decreasing the contribution of signal events incorrectly identified as likely to be background, so this idea was not used in our final method.

events than for good signal.[2] Figure 6.3 shows the distribution of this peak value in Monte Carlo events. From examining this figure, we choose a cut value of 10, which eliminates a substantial fraction of bad signal and background while retaining nearly all of our good signal. (Our studies show that our results are not particularly affected if the cut value is changed to other values near 10.) Table 6.2 shows the efficiency of this cut for good signal events, bad signal events, and background events.

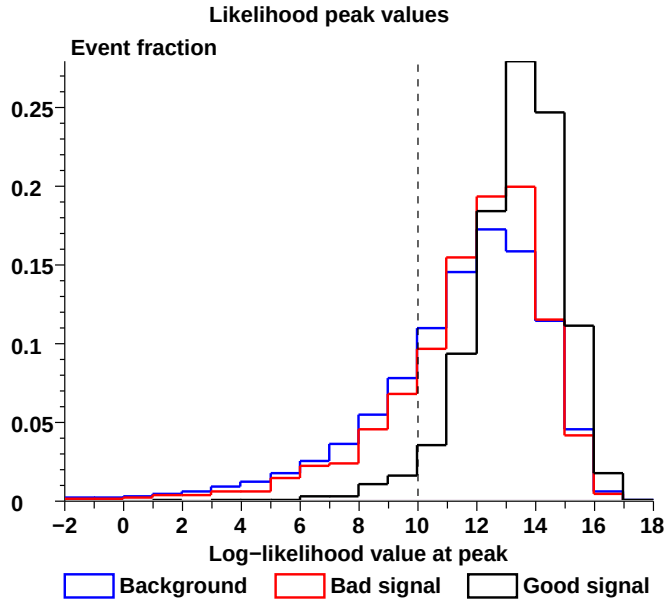

Figure 6.3: Distributions of the peak of the event log-likelihood for good signal, bad signal and background. The cut at 10 is illustrated by the dashed line. The top mass sample used here is at 172 GeV/c^2.

Overall, as we will see in Section 7.6, the likelihood cut results in a substantial improvement in our resolution despite the decreased number of events in our sample.

[2]We make no effort to apply an overall normalization to the likelihood curve, so the absolute value of this peak is not in itself meaningful, but the relative differences between peaks from different events is meaningful.

Table 6.2: Efficiency of the log-likelihood cut used in our analysis at a value of 10, given a top mass sample of 172 GeV/c^2. The background value and uncertainty are obtained by summing appropriately across the different background types. (Note that the uncertainties displayed are simply the binomial uncertainty. There is also a slight variation on the efficiency for good signal and bad signal with respect to top mass of about 1%.)

Type of event	Total	1-tag	>1-tag
Good signal	96.8% ± 0.2%	96.5% ± 0.2%	97.5% ± 0.3%
Bad signal	79.7% ± 0.5%	80.1% ± 0.6%	78.5% ± 1.1%
Background	73.5% ± 0.4%	73.6% ± 0.4%	73.6% ± 1.3%

Chapter 7

Method Testing and Calibration

7.1 Top Mass Extraction

Given an ensemble of events, we calculate the total log-likelihood by summing the log-likelihoods of each of the individual events, and then apply the likelihood modification procedure described in Section 6.3. Since we are interested in measuring the top mass, we need to eliminate the nuisance parameter Δ_{JES}. There are two main ways to perform this elimination — the profile likelihood and the marginalized likelihood. In the profile likelihood method, we select, for each point along the m_t axis, the point along the Δ_{JES} axis for which the likelihood is maximized:

$$L_{\mathrm{prof}}(m_t) = \max_{j \in \Delta_{\mathrm{JES}}} L(m_t, j). \tag{7.1}$$

The alternative method, the marginalized likelihood, integrates the likelihood along the Δ_{JES} axis with some given prior. While, in our studies, the difference between the profile and the marginalized likelihood results is small, we choose the profile likelihood, as it does not make any prior assumptions about the Δ_{JES} variable.

After reducing the 2-D curve to a 1-D likelihood, we can then extract a measured mass and uncertainty from this curve. The measured mass m_{meas} is simply determined by the position of the peak of the likelihood curve, while the uncertainty is determined by the standard technique of descending one-half unit of log-likelihood from this peak:

$$\begin{aligned} \log L(m_{\mathrm{meas}} + \sigma_+) &= \log L(m_{\mathrm{meas}}) - 0.5 \\ \log L(m_{\mathrm{meas}} - \sigma_-) &= \log L(m_{\mathrm{meas}}) - 0.5 \end{aligned}$$

For the sake of simplicity, we then symmetrize the individual positive and negative errors into a single symmetrized error $\sigma = (\sigma_+ + \sigma_-)/2$.

For a measurement in which the true top mass is known, we can also define the pull as the ratio of the actual error of the measurement to the measured

uncertainty:

$$\text{pull} = \frac{m_{\text{meas}} - m_{\text{true}}}{\sigma}. \tag{7.2}$$

While Δ_{JES} is not the parameter of interest in our final measurement, it is still valuable to measure it as a sanity check. We can perform the Δ_{JES} measurement in exactly the same way, eliminating m_t using the profile method and extracting a measured Δ_{JES} and uncertainty using the same method.

7.2 Pseudo-Experiment Procedure

Before we can perform a measurement on the data, it is necessary for us to test our technique on samples with known top masses so that we can verify that it returns correct top mass values and uncertainties for those values, and if not, to calibrate the method to appropriately correct these measured values. We do this by running ensembles of "pseudo-experiments" (PEs) on Monte Carlo samples with known values of m_t and Δ_{JES}.

A single PE is constructed by selecting a sample of Monte Carlo events. Events are selected from the $t\bar{t}$ signal sample being tested and the various background samples described in Section 4.5, with the contribution of each sample to the total being equal, on average, to its expected contribution from the background calculation described in Section 4.4. The average total number of events in each PE is set equal to the observed number of events in data. (Since the number of expected signal and background events does not add up exactly to the observed number of events in data, we use the estimated background numbers for the background, and then take the difference between the number of observed events and the total background contribution to be the signal contribution.) However, for each PE, the contribution from each sample is Poisson-fluctuated around the expected average, so the total number of events in any given PE may be more or less than the

expected number.

For a given top mass and Δ_{JES}, we perform an ensemble of 2000 PEs. For a single PE, we extract a measured mass and uncertainty using the procedure described above. We then define four quantities of interest from the PE ensemble as follows:

- Measured mass: The measured mass for an ensemble of PEs is simply the mean of the measured mass for the individual PEs, $\langle m_{\text{meas}} \rangle$.
- Bias: The bias is the difference between the measured mass and the true mass of the sample, $\langle m_{\text{meas}} \rangle - m_{\text{true}}$.
- Range: The range is the standard deviation (or RMS) of the distribution of measured masses of the PE ensemble, $\sqrt{\langle m_{\text{meas}}^2 \rangle - \langle m_{\text{meas}} \rangle^2}$.
- Pull width: The pull width (often referred to simply as "pull") is the standard deviation (RMS) of the distribution of the individual PE pulls.

A perfect analysis would have a bias of 0 (indicating that it correctly measures the input mass) and a pull width of 1 (indicating that it correctly estimates its own uncertainties). However, imperfections in the analysis mean that these conditions are not exactly satisfied. By measuring the bias and pull width on Monte Carlo samples, however, we can use these measured quantities to calibrate our analysis so that we can correctly measure the top mass and its uncertainty on data.

We also call the range the "expected uncertainty", as it provides a measurement of the expected uncertainty of an actual measurement, including any corrections necessitated by a non-unit pull width. (Note that in the limiting case where σ is constant, the range is by definition equal to σ times the pull width.) It is important to note that if the measured top mass is corrected by a multiplicative factor, as we will see in the next section, the range must also be corrected by this same factor.

It is of course also important to know the error on the bias, range, and pull for an ensemble of PEs. Because the number of PEs performed times the number of

events per PE is much greater than the number of events in our sample, a single MC event thus appears many times in our ensemble of PEs. Consequently, estimating the error on the bias using the standard formula $\sigma/\sqrt{N_{\mathrm{PE}}}$ will incorrectly underestimate the error (and similarly for the error on the range and pull width). Instead, we use the bootstrap method to evaluate the errors on our measured mass, expected error, and pull width. For a full description of the bootstrap method, see [56]; to briefly summarize how it is used in our analysis, we construct N = 50 "pseudo-samples" from our original sample by randomly choosing events (with replacement[1]) from the original sample. The pseudo-samples have a size equal to the original sample. We then perform our normal PE procedure on each pseudo-sample to obtain 50 different mass, expected error, and pull width measurements. By examining the spread of these distributions, we can thus obtain the errors on these measurements. We perform these bootstrap studies on one sample and assume that the results are valid for all input samples of the same size; since some of our samples contain 8k events and some 16k events, we perform the bootstrap twice for each number of events. The results of these studies for 8k events are shown in Figure 7.1. Based on these results, we assign an error of 0.28 to our bias measurements, 0.031 to our pull width measurements, and 0.035 to our expected uncertainty measurements for a sample size of 8k events and 0.16, 0.022, and 0.024 for the bias, pull, and expected uncertainty for a sample size of 16k events.

For the bias, at least, one can also obtain an estimate of the uncertainty by slightly modifying the $\sigma/\sqrt{N}$ formula so that the N used is the number of unique PEs — that is, the total number of events available in the sample divided by the number of events per PE (for this calculation, we use the number of signal events, since they tend to be the limiting factor). The uncertainty on the bias obtained in this method agrees well with the uncertainty obtained from the bootstrap method. When we have a sample for which we have processed a different number of events

[1]That is, a given event from the original sample can appear more than once in a single pseudo-sample.

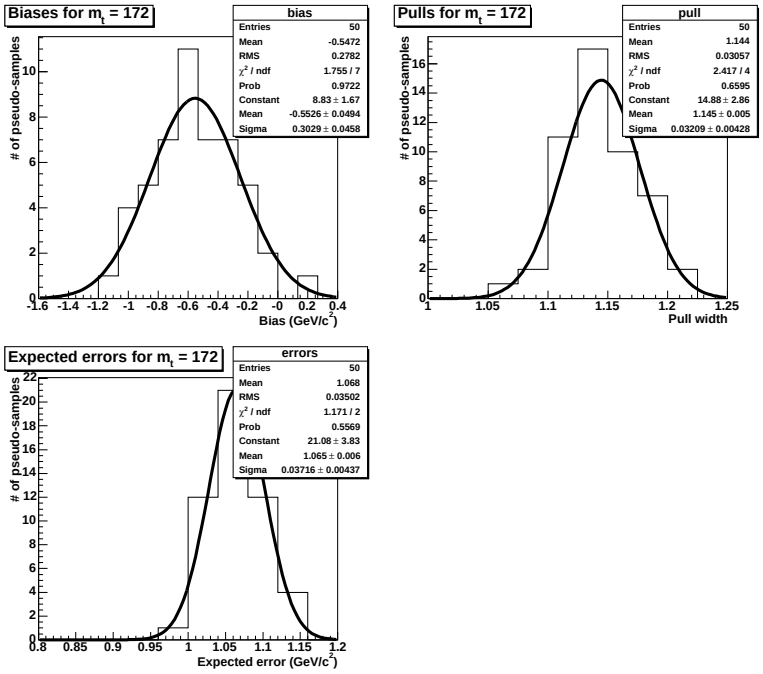

Figure 7.1: Results of our bootstrap studies for the $m_t = 172$ GeV/c^2 sample with 8k events. Top left: distribution of biases for the ensemble of pseudo-samples. Top right: distribution of pull widths for the ensemble. Bottom left: distribution of range (expected errors) for the ensemble.

than 8k or 16k, we can use this simpler method to calculate the uncertainty on the bias rather than redoing the whole bootstrap analysis. (Note that this method does not produce reasonable results for the range or pull width, but typically we are only interested in the bias anyway when making systematic measurements.)

Note that when we are performing PEs on samples with known top mass and Δ_{JES}, we reduce the range of the likelihood curves to a ± 12 GeV/c^2 range around the true top mass and a $\pm 2\sigma$ range around the true Δ_{JES} value. This improves the speed of the PEs, and is more than large enough that all of the individual PE results will still fall in this range.

7.3 Fully Realistic Pseudo-Experiment Results

The most important PEs run in testing and calibrating our method are what we call "fully realistic" PEs; that is, PEs which use all the parts of our analysis machinery, including the background subtraction and likelihood cut described in Chapter 6, and where the composition of the events in the PE is the same as the composition of events expected in the data. These PEs are run with a mean of 640.5 events per PE, which is the total number of events expected to pass the likelihood cut. (This number is obtained by taking the expected number of events from Table 4.2 before the cut for each type of background and multiplying for the efficiency of the likelihood cut for that type of background.)

7.3.1 Results in m_t with fixed Δ_{JES}

In our first set of PEs, we vary the input top mass while holding the input Δ_{JES} fixed at 0. For these tests, we use a total of 14 input masses between 160 GeV/c^2 and 184 GeV/c^2. These results are shown in Figure 7.2.

As we can see, while the response of the measured top mass to the input mass is reasonably linear, the bias and slope are not quite 0 and 1, respectively,

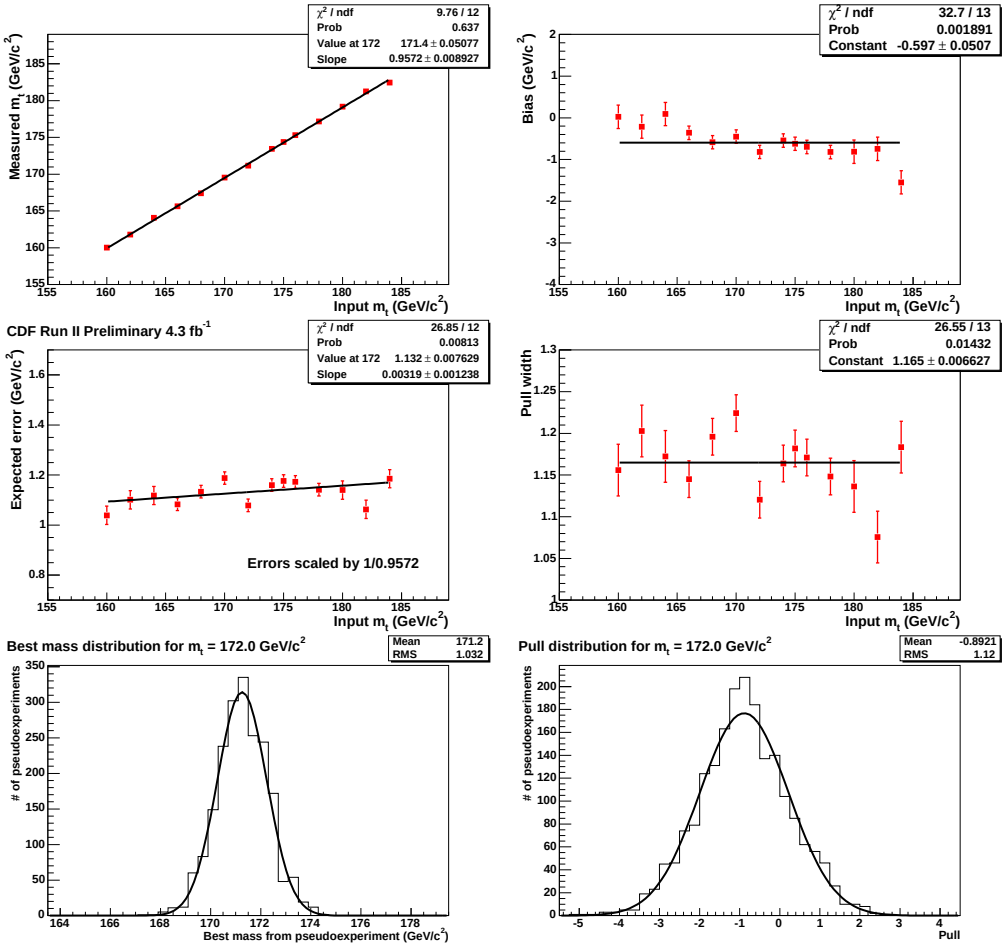

Figure 7.2: Pseudo-experiment results using fully simulated signal and background events after applying the likelihood cut of 10. Calibration has not been applied. The expected number of events is 640.5 after the likelihood cut has been applied. Here, 14 mass values between 160 and 184 GeV/c^2 have been used. Top left: reconstructed vs. input top mass; top right: bias vs. input top mass; center left: expected uncertainty vs. input top mass; center right: pulls vs. input top mass; bottom left: distribution of mass from individual pseudo-experiments for m_t = 172 GeV/c^2; bottom right: distribution of pulls for m_t = 172 GeV/c^2.

indicating that we need to apply a calibration to the measured m_t to obtain the correct value. Similarly, the pull width is not quite 1, indicating that we need to apply a calibration to the measured uncertainty to obtain the correct value. The distributions of PE masses and pulls look reasonably Gaussian, indicating that the PE ensembles are stastically well-behaved.

7.3.2 Results in Δ_{JES} with fixed m_t

Next, we proceed to measure the output Δ_{JES} in samples with a fixed m_t of 172 GeV/c^2 and vary the input Δ_{JES} value among five different possibilities: -1 σ, -0.5 σ, 0 σ, +0.5 σ, and +1 σ. These results are shown in Figure 7.3.

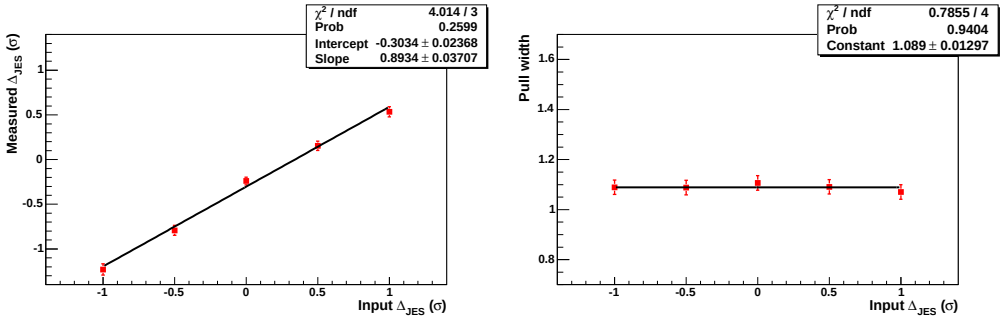

Figure 7.3: Pseudo-experiment results in measuring the Δ_{JES} value using the fully realistic case. Left: reconstructed Δ_{JES} vs. input Δ_{JES}. Right: pull width for the Δ_{JES} measurement vs. input Δ_{JES}.

Again, we see that a calibration is needed to account for the bias, slope, and pull width in the Δ_{JES} measurement.

7.3.3 Results with varying m_t and Δ_{JES}

Finally, we need to see if the calibration constants themselves vary with input m_t and Δ_{JES}, so we vary both of these parameters. Specifically, we examine three

different top masses (166, 172, and 178 GeV/c^2) with the five different Δ_{JES} values and look at the resulting measured m_t and Δ_{JES}. The results of these PEs are shown in Figure 7.4.

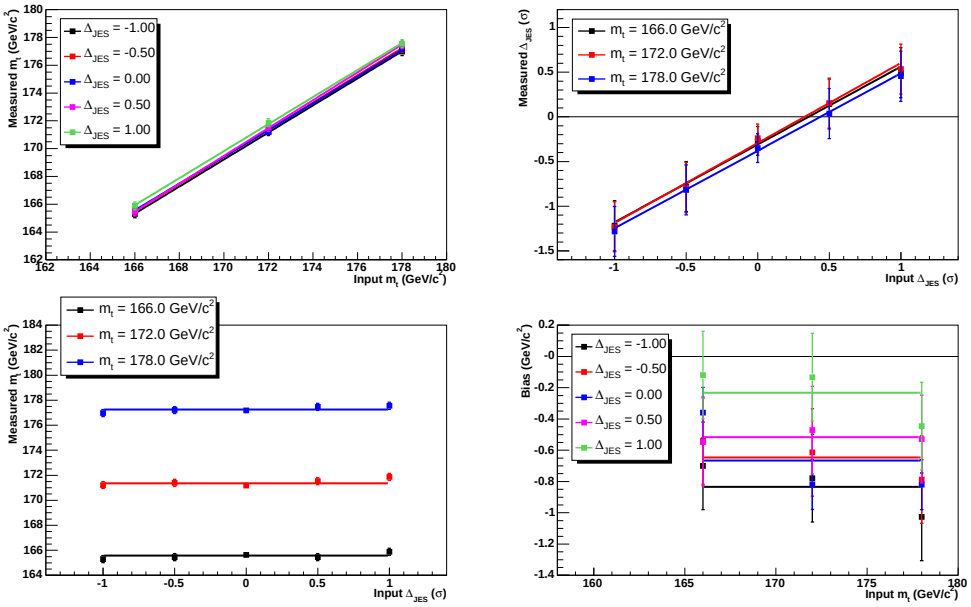

Figure 7.4: Study of JES linearity. Top left: Reconstructed vs. input top mass at the five different Δ_{JES} points; top right: reconstructed vs. input Δ_{JES} at the three different mass points; bottom left: reconstructed top mass vs. input Δ_{JES} for the three different top masses; bottom right: bias vs. input mass for the five different Δ_{JES} values.

As we can see from the bottom two plots, the bias in m_t does not appear to be a constant, but rather varies with the input Δ_{JES} as well. Thus, we cannot simply calibrate m_t and Δ_{JES} independently using the results from the PEs shown in Figures 7.2 and 7.3, but must also account for this cross-dependence.

7.4 Calibration of the Method

In the most general case, the measured mass (which we write in terms of $\Delta m = (m_t - 172\ \mathrm{GeV}/c^2)$, since our fits are centered around $m_t = 172\ \mathrm{GeV}/c^2$) and measured Δ_{JES} depend on both the true mass and Δ_{JES}. That is, assuming a linear dependence in our calibration,

$$\Delta m_{\mathrm{meas}} = a \times \Delta m_{\mathrm{true}} + b \times (\Delta_{\mathrm{JES}})_{\mathrm{true}} + c \tag{7.3}$$

$$(\Delta_{\mathrm{JES}})_{\mathrm{meas}} = d \times (\Delta_{\mathrm{JES}})_{\mathrm{true}} + e \times \Delta m_{\mathrm{true}} + f \tag{7.4}$$

Inverting this, we obtain

$$\Delta m_{\mathrm{true}} = d/(da - be) \times \Delta m_{\mathrm{meas}} - b/(da - be) \times (\Delta_{\mathrm{JES}})_{\mathrm{meas}} + (bf - dc)/(da - be) \tag{7.5}$$

$$(\Delta_{\mathrm{JES}})_{\mathrm{true}} = a/(da - be) \times (\Delta_{\mathrm{JES}})_{\mathrm{meas}} - e/(da - be) \times \Delta m_{\mathrm{meas}} + (ec - af)/(da - be) \tag{7.6}$$

However, in our case we can simplify the calibration procedure somewhat. First, we notice from the upper right plot of Figure 7.4 that the dependence of the measured Δ_{JES} on the true top mass is negligible, so we can set $e = 0$ in the above equation.

To determine a and c, we look at PEs where the input Δ_{JES} is 0 (thus eliminating the effect of the b term), as shown in Figure 7.2. From these fits, we can directly obtain $a = 0.957$ and $c = -0.597$. Similarly, to determine d and f, we examine the output Δ_{JES} for pseudo-experiments with varying values of input Δ_{JES} and a fixed $m_t = 172\ \mathrm{GeV}/c^2$, as shown in Figure 7.3. From these fits we obtain $d = 0.893$ and $f = -0.303$.

With these PEs, we can perform 1-D calibrations on the mass and the $\Delta_{\rm JES}$ variables ignoring the dependence of m_t on $\Delta_{\rm JES}$. Figures 7.5 and 7.6 show the resulting m_t and $\Delta_{\rm JES}$ measurements after these 1-D calibrations are applied; we can see that for the cases of $\Delta_{\rm JES}$ fixed at 0 and m_t fixed at 172 GeV/c^2, respectively, the 1-D calibrations lead to a bias of 0, slope of 1, and pull width of 1, as desired.

However, as we have seen, there is still a dependence of the top mass on the $\Delta_{\rm JES}$. To obtain the relevant parameter (the b/a in the above equation), we fit the observed shift in m_t after the 1-D calibrations are applied against the input $\Delta_{\rm JES}$. We observe, as shown in Figure 7.7, that a shift of $\Delta_{\rm JES} = 1$ yields a shift in the top mass of 0.28. Hence, we have that $b/a = 0.28$.

Hence, our final calibration formula is:

$$\Delta m_{\rm calib} = (\Delta m_{\rm meas} + 0.597)/0.957 - 0.28 \times (\Delta_{\rm JES})_{\rm calib} \quad (7.7)$$

$$(\Delta_{\rm JES})_{\rm calib} = ((\Delta_{\rm JES})_{\rm meas} + 0.303)/0.893 \quad (7.8)$$

We also calibrate the measured uncertainties using the slopes measured above and the pull widths. From Figures 7.2 and 7.3 we obtain a pull width of 1.165 for the mass measurement and 1.089 for the $\Delta_{\rm JES}$ measurement. Hence, we have

$$(\sigma_m)_{\rm calib} = (\sigma_m)_{\rm meas} \times 1.165/0.957$$

$$(\sigma_{\Delta_{\rm JES}})_{\rm calib} = (\sigma_{\Delta_{\rm JES}})_{\rm meas} \times 1.089/0.893$$

Figure 7.8 shows the full effects of the 2-D calibration for all of our Monte Carlo samples.

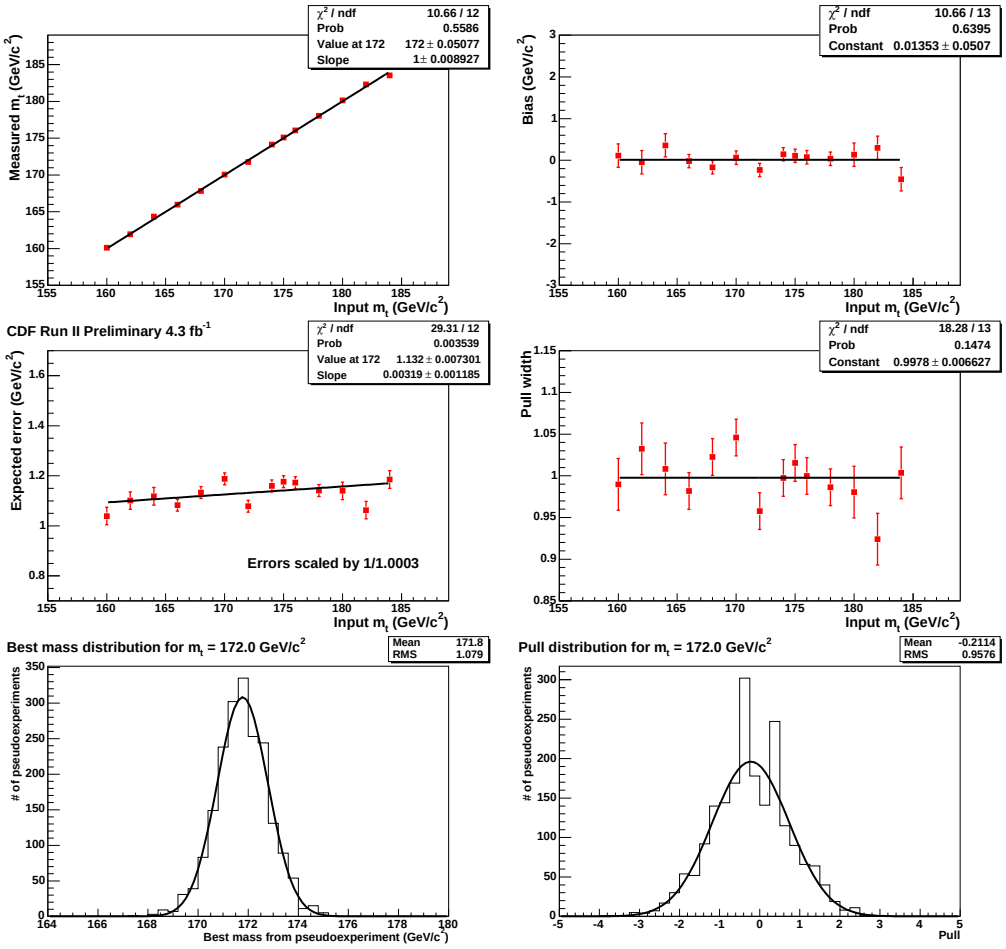

Figure 7.5: Pseudo-experiment results in the fully realistic case after the 1-D calibration has been applied. Top left: reconstructed vs. input top mass; top right: bias vs. input top mass; center left: expected uncertainty vs. input top mass; center right: pulls vs. input top mass; bottom left: distribution of mass from individual pseudo-experiments for $m_t = 172$ GeV/c^2; bottom right: distribution of pulls for $m_t = 172$ GeV/c^2.

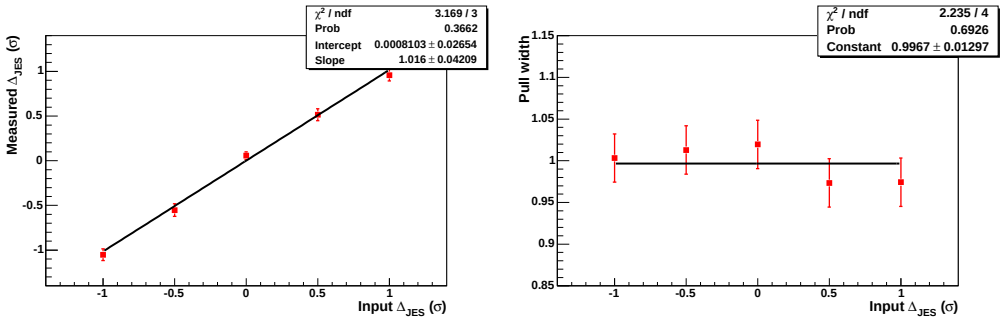

Figure 7.6: Pseudo-experiment results in measuring the Δ_{JES} value using the fully realistic case after the 1-D calibration has been applied. Left: reconstructed Δ_{JES} vs. input Δ_{JES}. Right: pull width for the Δ_{JES} measurement vs. input Δ_{JES}.

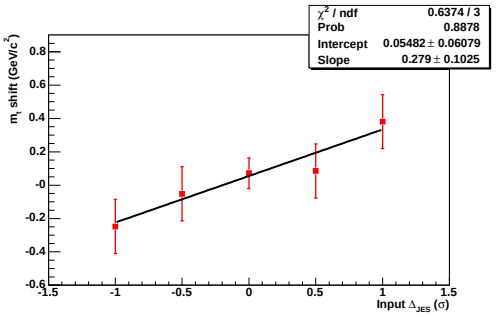

Figure 7.7: Final calibration: Shift in m_t vs. input Δ_{JES} after the individual mass and Δ_{JES} calibrations have been applied.

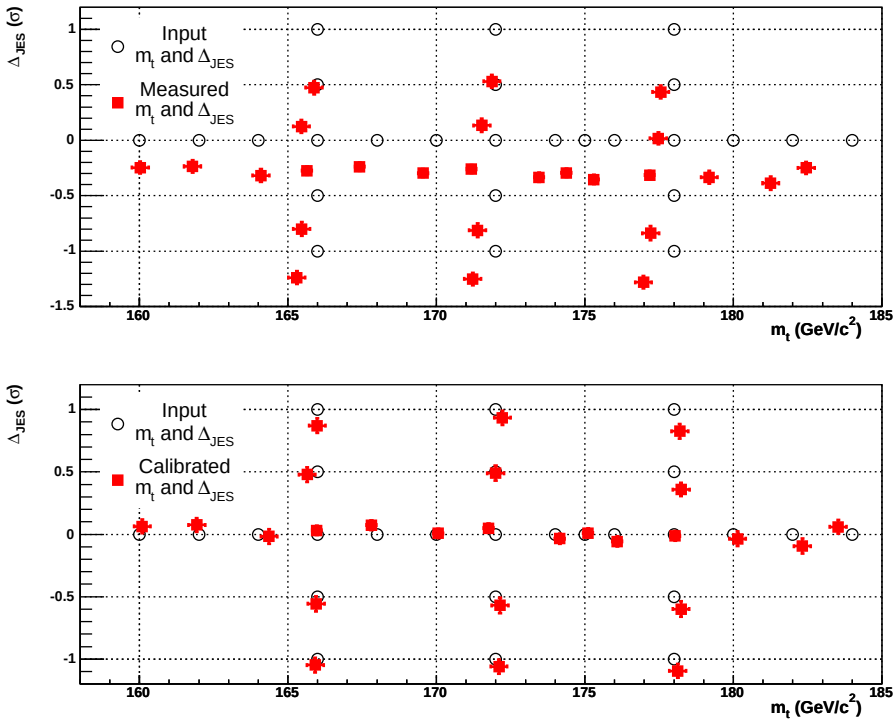

Figure 7.8: Top: Measured m_t and $\Delta_{\rm JES}$ before applying the 2-D calibration. Bottom: Measured m_t and $\Delta_{\rm JES}$ after applying the 2-D calibration.

7.5 Blind Samples

As a check to ensure that our method is unbiased after the calibration is applied, we also test our analysis on ten "blind samples", samples prepared by the CDF Top Mass Group conveners where the true top mass is hidden from us. (These samples do not have a Δ_{JES} shift, however, so we assume $\Delta_{\text{JES}} = 0$ when measuring them.)

The results, as plotted by the conveners, are shown in Figure 7.9 and Figure 7.10. The blind sample biases are consistent with 0 and the pulls with 1, so we conclude that our calibration works successfully for $\Delta_{\text{JES}} = 0$.

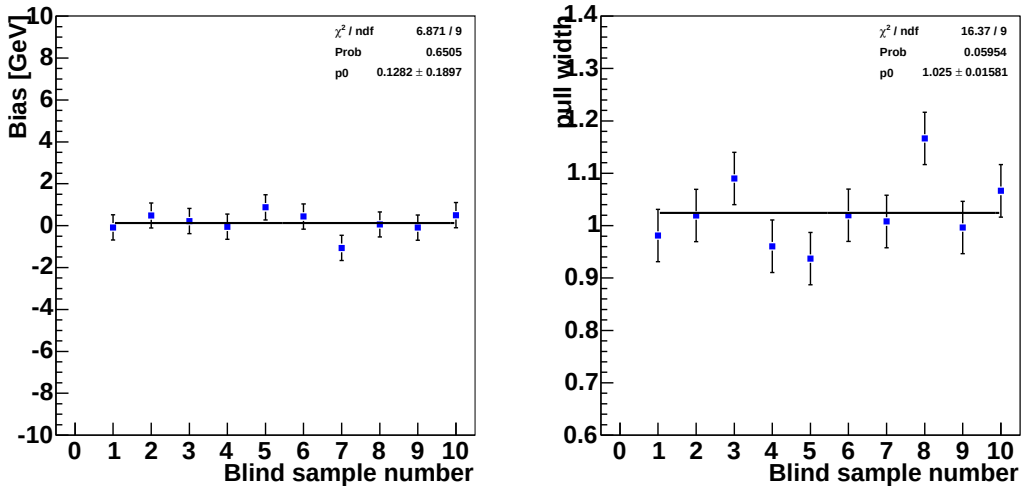

Figure 7.9: Measured bias (left) and pull width (right) for the ten blind samples. The numbering of the samples is randomized to avoid identification.

7.6 Other PE Results

As we have noted, since we know that our model is not completely perfect, we do not expect the bias and pull width to be perfectly 0 and 1, respectively. However, we can run PEs with other configurations of events to see if we can identify which events create the most problems for our analysis, to identify areas where we can

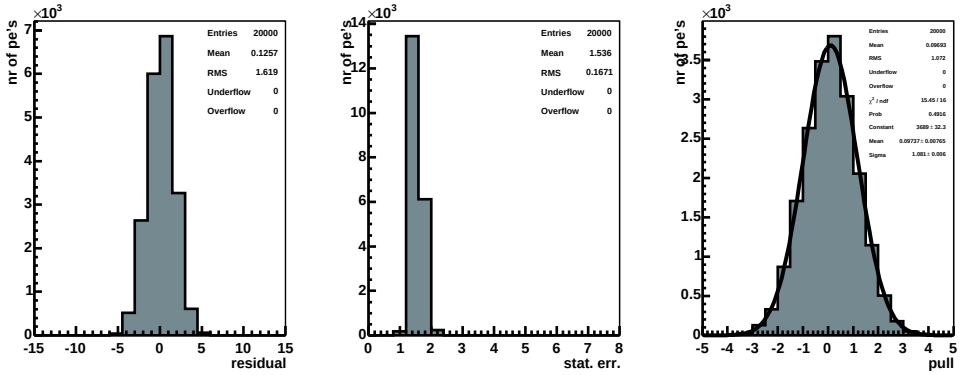

Figure 7.10: Distributions for bias (left), expected uncertainty (center), and pull (right) for individual PEs across the ten blind samples.

improve in future versions. Similarly, we can also run PEs in which various parts of our analysis are not used, so that we can quantify the improvements obtained by using them.

Table 7.1 shows some of these results. The various configurations used are:

- Signal only: In this setup, the analysis is performed using signal events only without any background. The background handling procedure is not used. We use 513.4 events per PE in this case, which is the expected number of signal events passing our likelihood cut in our sample.

- Good signal only: This is as above, except the bad signal events are also removed, leaving only good signal events in the PE. In this case, the likelihood cut is not used. This uses 368.1 events per PE, which is the expected number of good signal events.

- No background handling: This uses the fully realistic signal+background setup, but the background handling procedure described in Section 6.3 is not used. This allows us to examine the improvement obtained by using the modified likelihood.

- No likelihood cut: This uses the fully realistic setup, but without the likelihood cut described in Section 6.4. This allows us to isolate the improvement obtained by using this cut. In this case, 738 events per PE are used.

Table 7.1: Summary of PE results for different PE configurations to isolate the effects of various types of events and features of our analysis.

Input used	Avg. bias	Slope	σ at 172 GeV/c^2	Avg. pull
Fully realistic	-0.60 ± 0.05	0.957 ± 0.009	1.13 ± 0.01	1.17 ± 0.01
Good signal only	0.04 ± 0.05	0.976 ± 0.009	0.99 ± 0.01	0.95 ± 0.01
Signal only	-0.56 ± 0.05	0.962 ± 0.009	1.02 ± 0.01	1.05 ± 0.01
No bg handling	-1.90 ± 0.05	0.925 ± 0.009	1.11 ± 0.01	1.16 ± 0.01
No likelihood cut	-0.67 ± 0.05	0.993 ± 0.008	1.27 ± 0.01	1.35 ± 0.01

As we can see, when we restrict our sample of events to only good signal, which are the events that our model is actually designed to describe, we obtain a bias consistent with 0 and a pull width actually slightly less than unity, although the slope is still slightly smaller than 1; overall, we conclude that our model performs well on good signal events. Figure 7.11 displays the plots of the PE results in the good signal only case. Going from good signal only to all signal results in a worse resolution, bias, slope, and pull width, indicating that the presence of bad signal events which our model is ill-equipped to handle is a main contributor to the imperfect results we obtain. Furthermore, going again from the all signal only case to the fully realistic case again results in worse performance across the board, indicating that the imperfections in our background method do hurt our final result.

The likelihood cut can be seen to improve our resolution and pulls substantially, despite the fact that it reduces the size of the event sample; this indicates that the likelihood cut does a good job at removing the poorly-behaved bad signal and background events. In contrast, the background handling has almost no effect on the resolution, although it does dramatically improve the bias and slope. This may be due to the fact that our discriminant simply does not discriminate well enough for us to see a substantial improvement by using it.

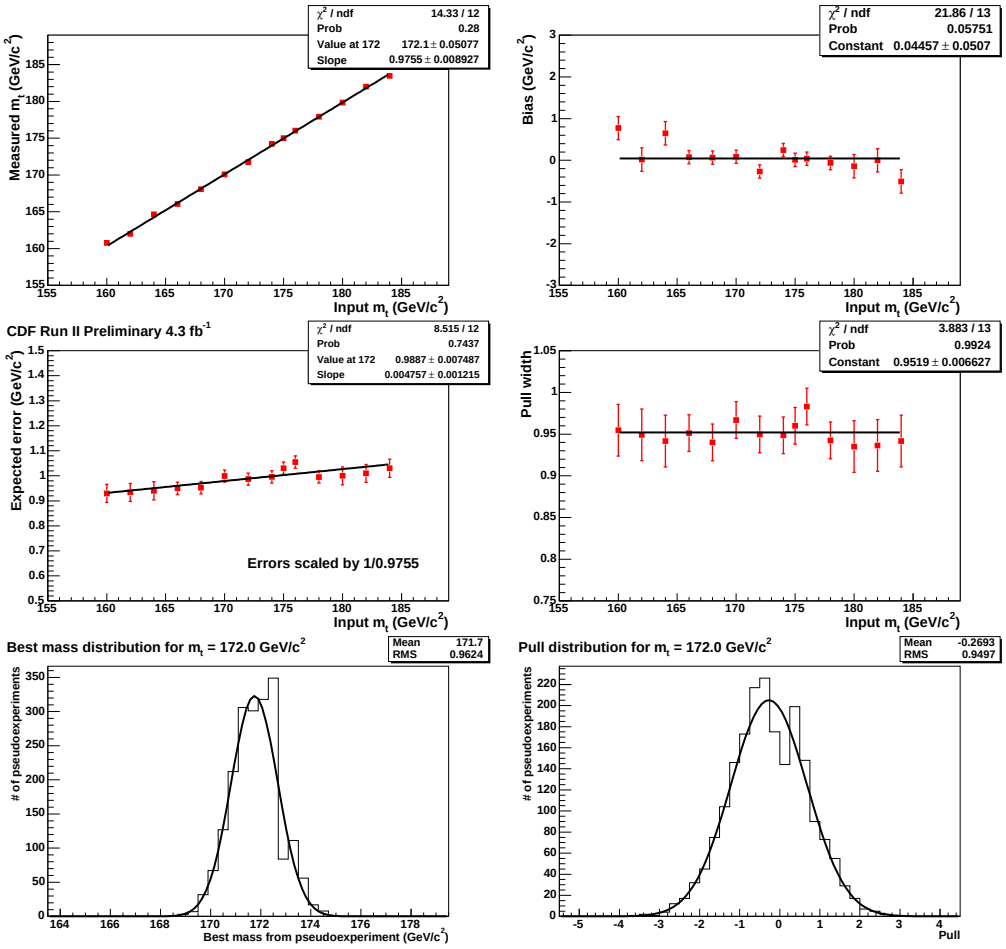

Figure 7.11: Pseudo-experiment results using only good signal events. The expected number of events is 368.1, with no likelihood cut used. Top left: reconstructed vs. input top mass; top right: bias vs. input top mass; center left: expected uncertainty vs. input top mass; center right: pulls vs. input top mass; bottom left: distribution of mass from individual pseudo-experiments for $m_t = 172$ GeV/c^2; bottom right: distribution of pulls for $m_t = 172$ GeV/c^2.

Chapter 8

Data Result

As mentioned in Section 4.3, we observe a total of 738 events passing our initial selection cuts, 590 single-tag and 148 with two or more tags. After applying the additional likelihood cut, 630 events remain, of which 493 are single-tag and 137 multiple-tag. We compute the likelihood for each of these events, apply the background subtraction procedure in Section 6.3, use the profile likelihood method to obtain a 1-D measurement, and then apply the calibration in Section 7.4, thus extracting a measured mass and uncertainty of:

$$m_t = 172.64 \pm 1.13 \text{ GeV}/c^2 \tag{8.1}$$

As noted in Section 5.2, this statistical uncertainty includes the uncertainty in m_t alone as well as the uncertainty due to the Δ_{JES} parameter. To separate the uncertainty due to these two causes, we reduce the 2-D likelihood to 1-D by taking the likelihood in the $\Delta_{\text{JES}} = 0$ bin (essentially, thus, as if we had never used the Δ_{JES} parameter). This results in an uncertainty of 0.86 GeV/c^2. We thus subtract in quadrature to conclude that the remaining uncertainty of 0.73 GeV/c^2 is due to the Δ_{JES} and report a final measurement of:

$$m_t = 172.64 \pm 0.86 \text{ (stat.)} \pm 0.73 \text{ (JES) GeV}/c^2 \tag{8.2}$$

The measured value of Δ_{JES}, after the full calibration, is:

$$\Delta_{\text{JES}} = 0.30 \pm 0.23 \ \sigma \tag{8.3}$$

This measurement is consistent with zero, as we would hope.

Figure 8.1 shows the 2-D likelihood for single-tag, multiple-tag, and all events. These plots include the full 2-D calibration. These plots show that the total likelihood is well-behaved over the full m_t and Δ_{JES} range used in our integration. Figure 8.2 shows the contours corresponding to a 1-σ, 2-σ, and 3-σ statistical uncertainty around the peak, as well as the 1-D profile likelihood obtained from the 2-D curve.[1]

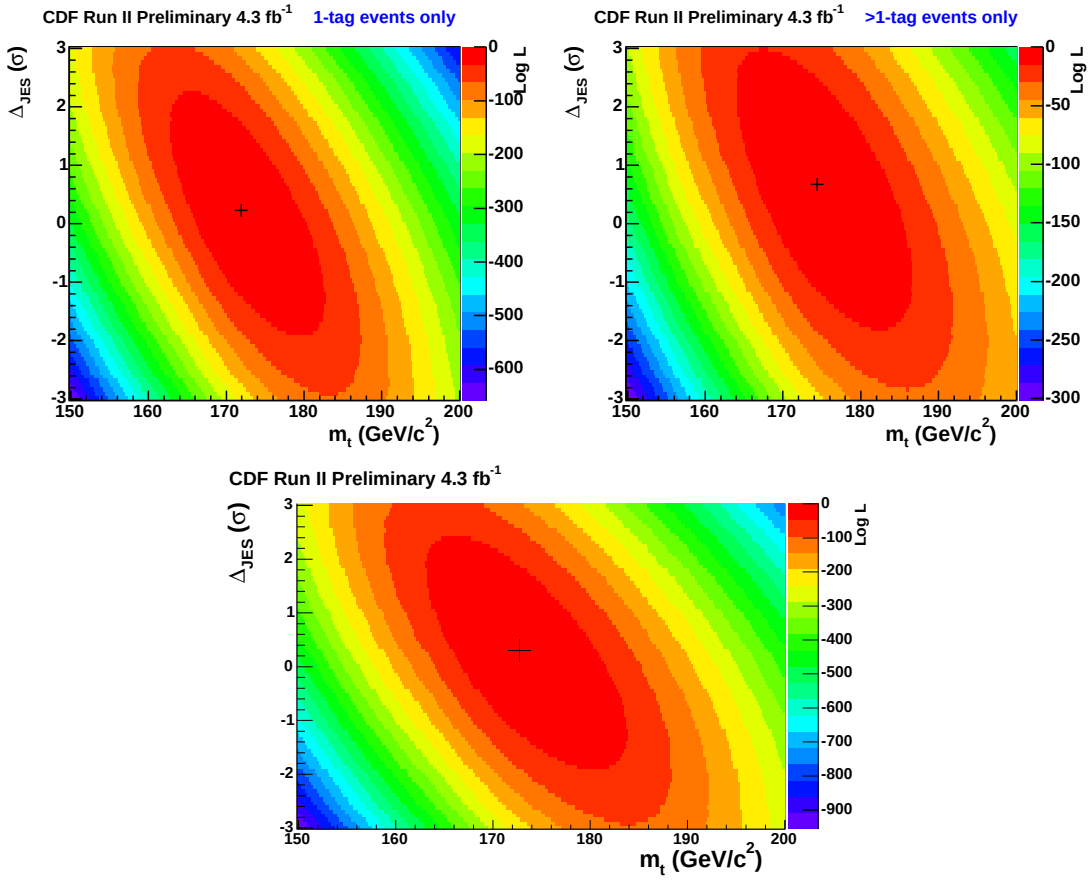

Figure 8.1: Fully calibrated 2-D likelihood for the data events. Top left: 1-tag events. Top right: >1-tag events. Bottom: Combined 1-tag and >1-tag. The marker shows the point of maximum likelihood.

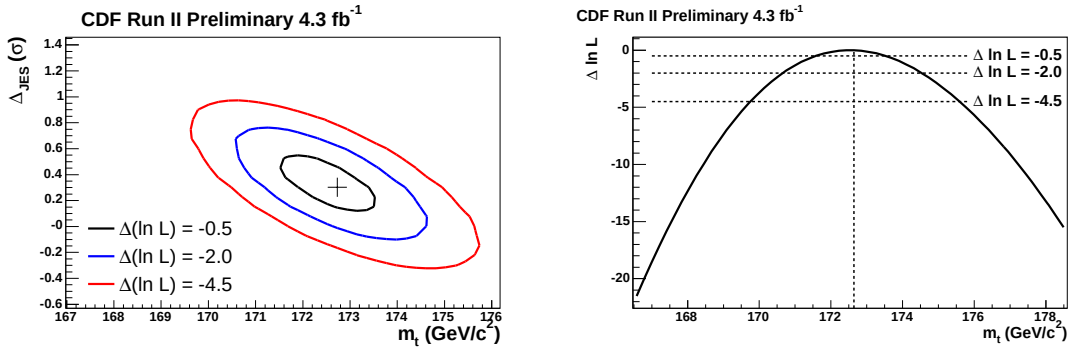

Figure 8.2: Left: Contours of 1-σ, 2-σ, and 3-σ statistical uncertainty around the peak of the likelihood after all calibration has been applied. Right: 1-D likelihood in m_t after the profile method has been used to eliminate Δ_{JES} from the 2-D likelihood curve.

Figure 8.3 shows the expected statistical uncertainty from an ensemble of PEs at a top mass of 172 GeV/c^2, with the uncertainty from the data measurement shown as the black arrow. 52% of pseudo-experiments show a lower uncertainty than measured in the data, indicating that our expected uncertainty is almost exactly what we expected from PEs.

As our likelihood cut method assumes that the distribution of the log-likelihood peaks for data follow the same shape as for Monte Carlo, we plot the log-likelihood value of the likelihood curve at its peak for each data event against a Monte Carlo distribution derived by taking the distribution for each individual subsample and adding them up according to their expected fraction. The results are shown in Figure 8.4; the vertical line indicates the value of the likelihood cut used. A K-S test indicates a confidence level of 0.73, showing good agreement between the data and the Monte Carlo.

[1]Note that, because we always use the profile measurement to obtain a m_t and Δ_{JES} measurement separately, the $\Delta(\log L)$ values used to obtain these contours are those appropriate to a 1-D measurement. Thus, it is correct to view the left and right edges of this oval as defining the 1-D uncertainty on m_t, but it is not correct to view the oval itself as defining the 2-D joint uncertainty in m_t and Δ_{JES}.

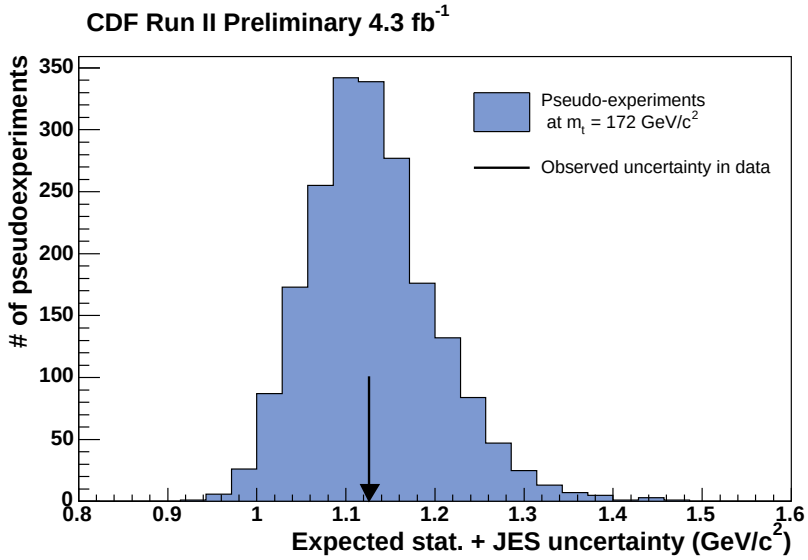

Figure 8.3: Expected statistical uncertainty on the 2-D profile likelihood method from Monte Carlo PEs at $m_t = 172$ GeV/c^2. The actual uncertainty observed in the data measurement is indicated by the black arrow. All uncertainties have been scaled by the average pull width of 1.165 and corrected by 1/.957 to account for the non-unit slope of the calibration.

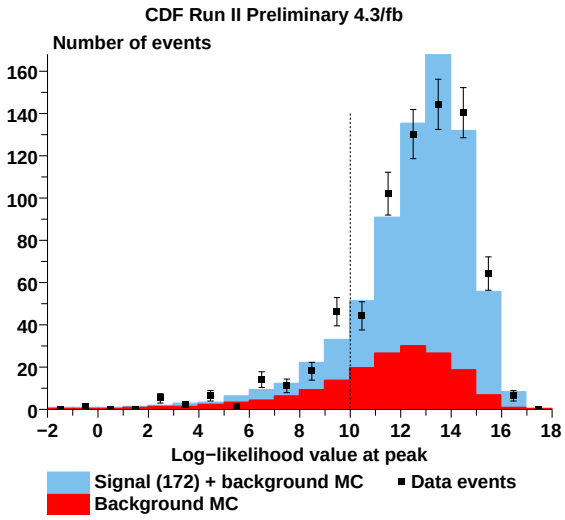

Figure 8.4: Comparison of the log-likelihood value of the peak of individual likelihood curves for data and Monte Carlo. The vertical line indicates the likelihood cut at 10. The confidence level returned by a K-S test is 0.73.

8.1 Data Subsamples

As a crosscheck, both to check the validity of our method and to make sure that the data itself is consistent, we try dividing the data into different subsamples and measuring the resulting top mass and Δ_{JES} in each subsample. The results are in Table 8.1. These results have been corrected using the same calibration as used in the main measurement.

Table 8.1: Data crosschecks on various subsamples of the data. The number of events listed is the number passing the likelihood cut out of the total number passing the other selection cuts.

Subsample	Events	Measured m_t (GeV/c^2)	Measured Δ_{JES} (σ)
All data	630/738	172.64 ± 1.13	0.30 ± 0.23
electron events	364/422	172.83 ± 1.49	0.38 ± 0.31
muon events	266/316	172.02 ± 1.66	0.30 ± 0.35
single-tag events	493/590	171.83 ± 1.29	0.23 ± 0.27
multiple-tag events	137/148	174.42 ± 2.37	0.68 ± 0.48
positive leptons	287/339	173.50 ± 1.75	0.23 ± 0.35
negative leptons	343/399	171.77 ± 1.47	0.45 ± 0.32
periods 0-7	173/200	169.17 ± 1.95	0.75 ± 0.43
periods 8-12	152/174	173.96 ± 2.46	-0.67 ± 0.48
periods 13-19	172/204	173.08 ± 2.00	0.97 ± 0.44
periods 20-23	133/160	174.36 ± 2.98	0.15 ± 0.56
First 3.2 fb^{-1}	497/578	172.20 ± 1.19	0.38 ± 0.26

While we do observe some differences between the samples, none of them is

statistically particularly significant, and all of them are consistent with patterns observed by other CDF groups. We also cross-check our method by comparing our result on the first 3.2 fb^{-1} of data to our previous result from that dataset, 172.14 $\pm$ 1.19 GeV/c^2, and note good agreement.

Chapter 9

Systematic Uncertainties

So far, we have only considered the statistical uncertainties of our measurement. However, there are many potential sources of systematic uncertainty as well. Many of these arise from areas where there is an inherent uncertainty in how well the Monte Carlo models the data, and so even if our analysis performs completely perfectly on Monte Carlo, we need to account for this potential mismodeling.

Table 9.1 lists all of the systematics evaluated in this measurement. We assume that the individual systematics are uncorrelated and Gaussian[1] and add them in quadrature to obtain a final systematic uncertainty of 1.11 GeV/c^2. These systematics are individually discussed in the following sections.

Broadly speaking, there are two different types of systematics. The first is where we can associate the uncertainty with a well-defined 1σ uncertainty in a given parameter. In this case, we vary that parameter by its uncertainty and measure the resulting change in the top mass. In this case, we define the systematic uncertainty as $|m_+ - m_-|/2$. (This assumes that m_+ and m_- lie on opposite sides of the unshifted mass. If they lie on the same side, we define the systematic uncertainty as half of the largest of $|m_+ - m_0|$, $|m_- - m_0|$, and $|m_+ - m_-|$.) The second is where we do not have a well-defined 1σ shift in a parameter. In this case, we generate one or more alternatives which we hope cover the space of possibilities and take the largest difference between the nominal mass and the various alternatives as our systematic shift.

There are three ways in which the systematics are measured. One is to generate an entirely new Monte Carlo signal or background sample in which the parameter of interest has been varied. In this case, the two samples are statistically independent, so we take the larger of the difference or the uncertainty on the difference as our actual systematic. For this reason, we often use more than the standard 8k or 16k events in these systematics, so that the uncertainty on the difference is reduced. In this case, the uncertainty on the mass measurement is calculated using the

[1]As we will see, not all of the systematics are necessarily actually Gaussian, but we have to make this assumption to sensibly combine them.

method described in Section 7.2. A second way is to take the standard `ttkt72` sample for $m_t = 172$ GeV/c^2 and to modify it to vary the quantity of interest. In this case, because the samples are completely correlated[2], the uncertainty on the difference is zero, so to reduce the time required to process these samples, we use fewer events, typically 5k. The third type of systematic does not change the underlying Monte Carlo sample at all, but rather attaches weights to the individual events. Then, when we run the PEs, the probability of selecting an event to be used is proportional to its individual weight. This does not require redoing the integration, so we can use as many events as we have.

The pseudo-experiment procedure used for systematic samples is the same as that used for the regular samples; signal and background are used in their expected proportions and then the full 2-D calibration described in Section 7.4 is applied to the result. (Note that, in most cases, only the signal sample is changed; the background sample is only changed for the residual JES and background Q^2 systematics.) Since we are only generally interested in the systematic shift of m_t, we do not compute the systematics for the Δ_{JES} measurement; however, since the 2-D calibration requires that both m_t and Δ_{JES} be measured, we do include the measured Δ_{JES} value as well.

9.1 Calibration

Since we use the calibration constants derived in Section 7.4 to correct our final result, any uncertainty in these constants naturally translates into an uncertainty on that result. Our quoted uncertainty includes two sources: first, the uncertainty on our fitted bias of -0.60 ± 0.05 GeV/c^2, which results directly in an uncertainty on our measurement, and second, the uncertainty in the slope of the m_t response with respect to Δ_{JES}. Since we measure a Δ_{JES} of 0.30 ± 0.23 with a slope of 0.28 ± 0.10, we obtain that the uncertainty due to this term is $\pm$ 0.09 GeV/c^2. (There

[2]Or, in the case of the residual JES, very highly correlated.

Table 9.1: Total list of sources of systematic uncertainty and their resulting measured uncertainty in the m_t measurement.

Systematic source	Systematic uncertainty (GeV/c^2)
Calibration	0.10
MC generator	0.56 ± 0.24
ISR and FSR	0.26 ± 0.13
Residual JES	0.52
b-JES	0.38
Lepton p_T	0.18
Permutation weighting	0.01
Multiple hadron interaction	0.10
PDFs	0.17
Background: fraction	0.33
Background: composition	0.36
Background: average shape	0.03
Background: Q^2	0.08 ± 0.07
Background: MC statistics	0.05
Gluon fraction	0.00
Color reconnection	0.32 ± 0.25
Total	1.11

is also a potential third source of uncertainty from the slope calibration constant, but the uncertainty due to this source is negligible compared to the preceding two.) Adding these in quadrature yields a total uncertainty of 0.10 GeV/c^2.

9.2 Monte Carlo generator

Most features of our analysis (e.g., the transfer functions) are derived from the PYTHIA Monte Carlo generator, and the samples used for the testing and calibration are also all derived from PYTHIA. Thus, it is important to check that our results are consistent with other Monte Carlo generators. We thus compare HERWIG and PYTHIA samples generated at a top mass of 172.5 GeV/c^2. (Note that to reduce the uncertainty on the difference, we use 24k events.) The results are shown in Table 9.2. We take the difference of 0.56 GeV/c^2 as our systematic uncertainty.

Table 9.2: Generator systematics. All samples have a nominal m_t of 172.5 GeV/c^2.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)	Δm_t (GeV/c^2)
PYTHIA (`ttop25`)	172.90 ± 0.17	0.03 ± 0.04	—
HERWIG (`dtops0`)	172.34 ± 0.17	0.38 ± 0.04	-0.56 ± 0.24

9.3 ISR and FSR

The amount of initial state radiation (ISR) and final state radiation (FSR) present in the Monte Carlo is subject to a systematic uncertainty. To evaluate the uncertainty on the amount of radiation present, we use a control sample of Drell-Yan events to measure the amount of ISR present in events as a function of the dilepton invariant mass. We can then extrapolate this quantity and its uncertainty to $t\bar{t}$ mass scales and, since the physical processes governing ISR and FSR are the same,

assume that this uncertainty also applies to FSR, thus obtaining an estimate of the uncertainty on the PYTHIA parameters governing ISR and FSR [57].

Two PYTHIA samples are then generated, one featuring increased ISR and FSR and one featuring decreased ISR and FSR (note that ISR and FSR are increased or decreased together). We integrate 24k events from these samples; the results are shown in Table 9.3.

Table 9.3: Systematics from ISR/FSR. All samples are generated with PYTHIA with a nominal m_t of 172.5 GeV/c^2.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)	Δm_t (GeV/c^2)
Nominal (`ttop25`)	172.90 ± 0.17	0.03 ± 0.04	—
more ISR/FSR (`dtops1`)	172.39 ± 0.18	0.26 ± 0.04	-0.51 ± 0.25
less ISR/FSR (`dtops2`)	172.50 ± 0.17	0.11 ± 0.04	-0.40 ± 0.24

We take as our systematic half the difference between the largest pair, which in our case is the nominal and more ISR/FSR sample, yielding a systematic of 0.26 ± 0.13 GeV/c^2.

9.4 Residual JES

Our 2-D likelihood method, by definition, is designed so that the measured top mass should not change if the jets are all shifted by their uncertainty, and this uncertainty depends on the p_T and η for each jet. However, as we saw in Section 3.4.2, the jet corrections (and their corresponding uncertainties) are composed of several different components, each of which has different behavior with respect to p_T and η. Consequently, a shift of 1σ in, for instance, the level 5 correction alone may produce behavior which will not be corrected for in our 2-D method. Note

that, although our jets are only corrected to level 5, we need to account for the systematic uncertainties in all levels of the jet corrections.

To evaluate a systematic due to this effect, we take our normal 172 GeV/c^2 signal sample and shift the jets up or down by the uncertainty for each individual jet correction level. (Note that, for this test, we integrate only 5k signal events for each shift. These 5k events are not necessarily identical for each shift, since we allow events to enter or leave the samples if the shift changes their jet energies appropriately, but they are highly correlated.) We also cross-check by shifting the background samples as well; however, shifting all of the background samples at all of the individual shift levels would be prohibitively time-consuming, so as a compromise, we shift all of the background samples for levels 5 and 7, which are are two largest systematics; for the other four levels, we use only the $W + b\bar{b}$ background so that we only need to shift that one sample.

Table 9.4 shows the results. For the case where only signal is shifted, we measure shifts of 0.06, 0.02, 0.40, 0.10, 0.18, and 0.14 GeV/c^2 for levels 1, 4, 5, 6, 7, and 8, respectively. Adding these shifts in quadrature yields a total error of 0.48 GeV/c^2. For the case where signal and background are shifted together, we obtain shifts of 0.37 and 0.28 for levels 5 and 7 (using all backgrounds) and 0.08, 0.02, 0.14, and 0.16 for levels 1, 4, 6, and 8, respectively (using $W + b\bar{b}$ only). We take the conservative estimate of using the higher shift for each level and add them in quadrature to obtain a total of 0.52 GeV/c^2, which we take as our final uncertainty due to residual JES.

9.5 b-jet energy scale

So far, we have assumed that the JES is the same for all jets. However, there is an additional uncertainty arising from relative differences between b and light jets. (Note that the jet systematic uncertainties are predominantly determined using light jets.) We identify three sources of uncertainty: one due to the uncertainty in

Table 9.4: Systematics for the residual JES uncertainty. All samples are based on 5k events from the `ttkt72` sample with a nominal mass of 172.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)
Nominal	172.28 ± 0.37	-0.01 ± 0.08
Level 1 $+1\sigma$	172.38 ± 0.38	0.17 ± 0.08
Level 1 -1σ	172.27 ± 0.37	-0.18 ± 0.08
Level 4 $+1\sigma$	172.30 ± 0.38	0.01 ± 0.08
Level 4 -1σ	172.30 ± 0.37	-0.04 ± 0.08
Level 5 $+1\sigma$	172.79 ± 0.38	0.48 ± 0.08
Level 5 -1σ	172.00 ± 0.36	-0.54 ± 0.08
Level 6 $+1\sigma$	172.23 ± 0.37	0.08 ± 0.08
Level 6 -1σ	172.43 ± 0.37	-0.12 ± 0.08
Level 7 $+1\sigma$	172.16 ± 0.37	0.75 ± 0.07
Level 7 -1σ	172.52 ± 0.38	-0.79 ± 0.08
Level 8 $+1\sigma$	172.15 ± 0.41	0.13 ± 0.08
Level 8 -1σ	172.42 ± 0.37	-0.15 ± 0.08
Level 5 s+b $+1\sigma$	172.71 ± 0.37	0.51 ± 0.08
Level 5 s+b -1σ	171.97 ± 0.37	-0.55 ± 0.08
Level 7 s+b $+1\sigma$	172.05 ± 0.36	0.75 ± 0.07
Level 7 s+b -1σ	172.60 ± 0.37	-0.79 ± 0.08
Nominal $Wb\bar{b}$ bkgd only	172.19 ± 0.38	-0.01 ± 0.08
Level 1 s+$Wb\bar{b}$ $+1\sigma$	172.22 ± 0.37	0.17 ± 0.08
Level 1 s+$Wb\bar{b}$ -1σ	172.27 ± 0.37	-0.22 ± 0.08
Level 4 s+$Wb\bar{b}$ $+1\sigma$	172.17 ± 0.38	0.02 ± 0.08
Level 4 s+$Wb\bar{b}$ -1σ	172.18 ± 0.37	-0.04 ± 0.08
Level 6 s+$Wb\bar{b}$ $+1\sigma$	172.06 ± 0.37	0.08 ± 0.08
Level 6 s+$Wb\bar{b}$ -1σ	172.33 ± 0.37	-0.13 ± 0.08
Level 8 s+$Wb\bar{b}$ $+1\sigma$	172.05 ± 0.38	0.12 ± 0.08
Level 8 s+$Wb\bar{b}$ -1σ	172.37 ± 0.37	-0.16 ± 0.08

the semileptonic decay ratio, which we evaluate by reweighting the `ttkt75` sample to vary this ratio by $\pm 1\sigma$; one due to the uncertainty in the b-fragmentation modeling, which we evaluate by varying the parameters used in the Bowler fragmentation model [58] in the PYTHIA Monte Carlo generator, using two different sets of parameters derived from SLD and LEP results [59]; and one due to uncertainty in the calorimeter response[3], which we evaluate by shifting the E_T of jets identified as b jets in the `ttkt72` sample by 1% and then multiplying by 0.2, as the expected uncertainty from this source is 0.2%. The results are shown in Table 9.5.

For the uncertainty due to the semileptonic fraction, we take half of the measured diffference, or 0.10 GeV/c^2; for the uncertainty due to the b fragmentation model, we take the single largest shift, or 0.35 GeV/c^2; and for the uncertainty due to the calorimeter response, we follow the procedure above to obtain an uncertainty of 0.12 GeV/c^2. Adding these in quadrature yields an overall uncertainty of 0.38 GeV/c^2.

9.6 Lepton p_T

There is also a systematic uncertainty on the lepton energy scale, although it is much simpler than the uncertainty for the jet energy scale; it has been measured to be a systematic uncertainty of $\pm 1\%$ on the lepton p_T. We thus evaluate this systematic by taking the nominal `ttkt72` sample, shifting the lepton energies by $\pm 1\%$, and measuring the resulting top mass; we take half of the resulting difference, 0.18 GeV/c^2, as our systematic.

[3]Specifically, because the charged particle fraction and momentum spectrum of b-jets is different from that of light jets, the calorimeter response may be different.

Table 9.5: Systematics for the b-JES.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)
Nominal $m_t = 175$ GeV/c^2	175.00 ± 0.26	0.01 ± 0.07
Semileptonic fraction $+1\sigma$	174.93 ± 0.26	-0.02 ± 0.07
Semileptonic fraction -1σ	175.12 ± 0.26	0.04 ± 0.07
LEP b fragmentation model	174.65 ± 0.26	0.07 ± 0.07
SLD b fragmentation model	174.69 ± 0.26	0.06 ± 0.07
Nominal $m_t = 172$ GeV/c^2	171.93 ± 0.26	0.01 ± 0.06
b-jets +1%	172.84 ± 0.38	0.00 ± 0.08
b-jets -1%	171.67 ± 0.38	-0.06 ± 0.08

Table 9.6: Systematics for the lepton P_T uncertainty. All samples use 5k events from the `ttkt72` sample with a nominal mass of 172.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)
Nominal	172.28 ± 0.37	-0.01 ± 0.08
Lepton p_T +1%	172.46 ± 0.38	-0.01 ± 0.08
Lepton p_T -1%	172.10 ± 0.37	0.02 ± 0.08

9.7 Permutation weighting

The permutation weights w_i, as described in Section 5.9, are derived from the fits to the tagging probabilities as measured in data, so we need to assign a systematic uncertainty due to the uncertainty in these fits. We estimate that the main source of uncertainty is the ratio of charm tags to b-tags, which is nominally 22%, so we vary this ratio by its relative uncertainty of 15%. We take half the resulting difference, 0.01 GeV/c^2, as our systematic.

Table 9.7: Systematics for the permutation weighting uncertainty. All samples use 5k events from the `ttkt72` sample with a nominal mass of 172 GeV/c^2.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)
Nominal	172.28 ± 0.37	-0.01 ± 0.08
Charm ratio $+1\sigma$	172.28 ± 0.39	-0.02 ± 0.08
Charm ratio -1σ	172.29 ± 0.37	0.00 ± 0.08

9.8 Multiple hadron interaction

Multiple hadron interaction (also known as "pileup") represents the systematic associated with our modeling of multiple $p\bar{p}$ interactions in a single event. We consider two sources of uncertainty due to this effect. First, we consider the fact that most of our Monte Carlo samples are generated to match the earlier run range of the detector. However, as the average luminosity observed in the detector has increased over time, the average number of interactions in our Monte Carlo samples is now lower than the average number of interactions observed in data. To estimate this effect, we take a Monte Carlo sample generated with a higher average number of interactions, `utop75`, and divide it into subsamples with 1, 2, 3, and 4 or more

vertices, and then measure the top mass in each subsample separately. The results are shown in Figure 9.1; from a linear fit, we obtain the slope of m_t vs. number of vertices to be 0.20 $\pm$ 0.16 GeV/c^2/vertex. Given that the average number of vertices in the data is 2.16 and in the normal signal + background MC is 1.53, we multiply the difference by this slope to assign an uncertainty of 0.13 GeV/c^2 due to this source.

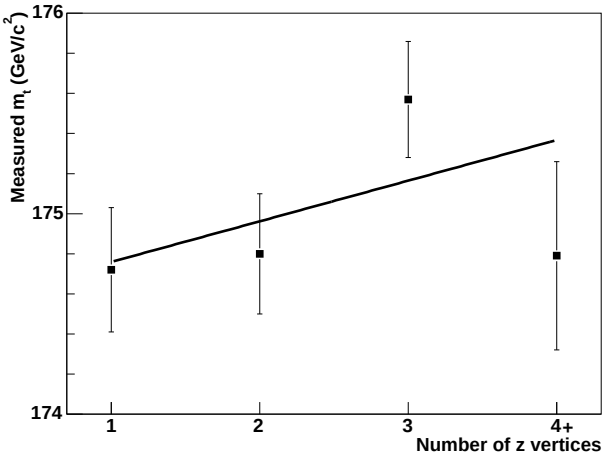

Figure 9.1: Measured top mass vs. number of vertices, as measured by separating the `utop75` sample into independent subsamples.

Second, we consider the modeling of additional interactions in an event. The level 4 jet correction is intended to correct for the effect of additional interactions on the jet response, but since this correction is derived from minimum bias events, we consider the possibility that it may not correctly model $t\bar{t}$ events. Studies show that the uncertainty can be estimated by scaling up the level 4 jet systematic uncertainty by a factor of 2.3, and then correcting this further to account for the different number of vertices in data and MC. Using the observed L4 systematic of 0.02 GeV/c^2 yields an uncertainty of 0.10 GeV/c^2. We take the larger of these two sources, which in this case yields 0.13 GeV/c^2 as our systematic uncertainty.

9.9 Parton distribution functions

Our analysis uses parton distribution functions to give the probability distribution for the momentum of the incoming partons. Naturally, there is some uncertainty in the determination of these PDFs, which results in a systematic uncertainty in our results.

We examine a total of 46 different PDF sets. In addition to the default CTEQ5L set, we also examine two sets of MRST [60] PDFs, MRST72 and MRST75, where MRST72 uses the same value of α_s as CTEQ5L and MRST75 a different value; CTEQ6L and CTEQ6L1, which use different α_s calculations, and finally CTEQ6M and variation of 20 different eigenvectors in CTEQ6M up and down. The results for all of these sets are shown in Figure 9.2.

We use the guidelines established by the CDF Top Group to obtain our final systematic, in which we consider the difference between CTEQ5L and MRST72 (to evaluate the uncertainty due to the PDF set used), MRST72 and MRST75 (to evaluate the uncertainty due to the α_s value used), and each pair of eigenvectors in CTEQ6M added in quadrature (to evaluate the uncertainty in the PDF set itself). The CTEQ5L-MRST72 difference is ignored, as it is smaller than the uncertainty from the CTEQ eigenvectors, and we add the remaining variations in quadrature to obtain our final uncertainty of 0.17 GeV/c^2.

9.10 Background

In addition to the systematics previously discussed, which mostly derive from our modeling of the $t\bar{t}$ signal, there are also a variety of sources of systematic uncertainties in our background model. We identify five sources and evaluate the systematic uncertainty from them as follows:

The first uncertainty is the uncertainty on the background fraction used in our analysis. This uncertainty is measured by shifting the total percentage of

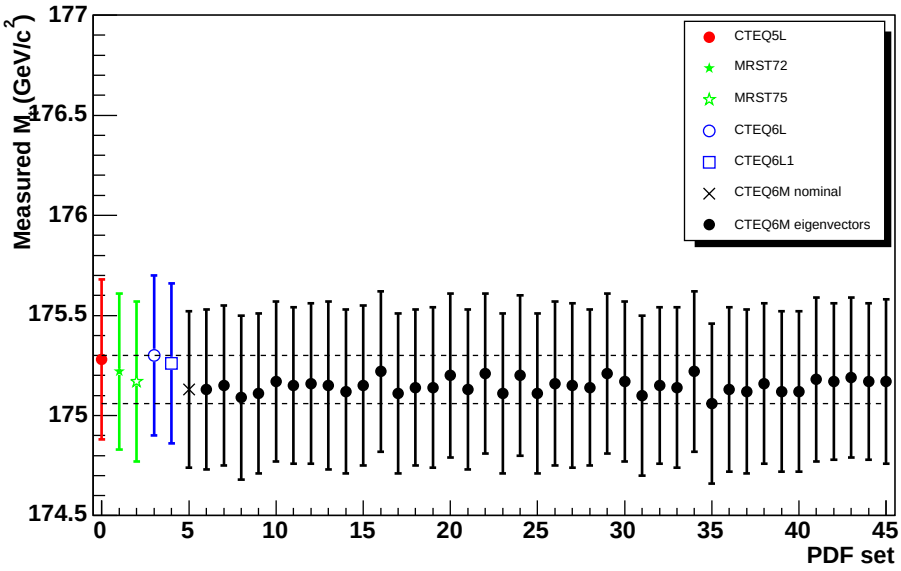

Figure 9.2: Variation of measured top mass with different sets of parton distribution functions (PDFs).

background present in our PEs up and down by 1σ. This uncertainty includes three sources:

1. The uncertainty on the background estimate listed in Table 4.2.

2. The uncertainty in the background fraction corresponding to a 1σ variation of Δ_{JES}. We calculate this uncertainty by shifting our samples by 1σ, calculating the ratio of acceptance for signal and background, and recomputing the background fraction appropriately.

3. The binomial uncertainty on the acceptance of our likelihood cut.

These three sources are summarized in Table 9.8; the total of these in quadrature is thus our uncertainty on the total background fraction. We then perform PEs with the shifted background fraction and take half the resulting difference.

The second source of uncertainty is the uncertainty in the background composition. For the background fraction, the overall fraction is changed but the relative

Table 9.8: Sources of uncertainty in the background fraction used in our PEs.

	1-tag	>1-tag
Nominal background fraction	26.54%	10.36%
Estimate uncertainty	± 6.90%	± 3.68%
JES uncertainty	± 3.41%	± 1.29%
Likelihood cut uncertainty	± 0.16%	± 0.18%
Total uncertainty	± 7.70%	± 3.90%

composition of the background remains the same, so we also would like to assess the potential systematic due to the relative composition. To do this, we take the maximally conservative approach of running PEs where the background events are entirely of a single type: either $W+b\bar{b}$, $W+c\bar{c}$, $W+c$, W+light, single-top, or QCD events. We then take the largest difference as our uncertainty due to background composition.

The third source of uncertainty is to the average background shape, the $\overline{L_{\rm bg}}(m_t, \Delta_{\rm JES})$ used in the background subtraction method discussed in Section 6.3. To assess a potential systematic from this source, we divide the background events into two independent subsamples, one with odd-numbered events and one with even-numbered events. We then build the background likelihood curve using events only from one sample and perform the PEs with events only from the other subsample and take the resulting largest difference as our systematic uncertainty.

The next source of uncertainty comes from the Q^2 scale used to generate the hard scatter process in W + jets events. ALPGEN controls this Q^2 scale through two parameters, `ktfact` and `qfact`, so we generate samples in which they have been doubled to 2.0 or decreased to 0.5. Note that this also changes the cross-section times acceptance values which we use to weight the $W + b\bar{b} + 2p$ vs. $W + b\bar{b} + 1p$

sample contributions, so we recalculate the relative fractions appropriately.

Finally, we have to consider an uncertainty due to the limited background MC statistics. The calibration systematic covers this uncertainty for the signal, but the background is not yet accounted for. To evaluate this systematic, then, we use the bootstrap method where only the background samples are bootstrapped, and obtain an uncertainty of 0.05 GeV/c^2. We thus take this as our uncertainty due to limited statistics.

(Note that since the background Q^2 involves separate samples, we have to calculate the appropriate error on the difference. Using the above error of 0.05 on the mass due to the background alone, we thus obtain an error of 0.07 for the difference between two different measurements.)

All of the background systematic results are summarized in Table 9.9.

Based on these results, we assign a systematic of 0.33 GeV/c^2 for background fraction, 0.36 GeV/c^2 for background composition, 0.03 GeV/c^2 for the average background shape, and 0.08 GeV/c^2 for the background Q^2 scale, in addition to the 0.05 GeV/c^2 already mentioned for the limited background MC statistics.

9.11 Gluon fraction

HERWIG and PYTHIA are both leading-order MC generators, so $t\bar{t}$ events in these samples are approximately 95% produced from $q\bar{q}$ annihilation and 5% produced from gg fusion. However, NLO expectations are closer to 15% $\pm$ 5% gg production. To check for a potential systematic due to this discrepancy, we run pseudo-experiments where $q\bar{q}$ and gg events have been reweighted so that the sample is expected to contain 80% $q\bar{q}$ events and 20% gg events (we choose the maximal gg percentage to be conservative). These results are shown in Table 9.10; we observe no difference in this case, so we take no systematic uncertainty from this source.

Table 9.9: Background systematics.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)	Δm_t (GeV/c^2)
Nominal $m_t = 172$	172.02 ± 0.26	0.02 ± 0.07	—
Background frac. +1 σ	171.64 ± 0.26	-0.03 ± 0.07	-0.38
Background frac. -1 σ	172.30 ± 0.26	0.09 ± 0.06	0.28
100% $W + b\bar{b}$ bg	171.92 ± 0.26	0.02 ± 0.07	-0.10
100% $W + c\bar{c}$ bg	171.69 ± 0.26	0.01 ± 0.06	-0.33
100% $W + c$ bg	172.34 ± 0.26	-0.11 ± 0.07	0.32
100% W+light bg	172.38 ± 0.26	0.13 ± 0.07	0.36
100% single-top bg	172.32 ± 0.26	0.26 ± 0.07	0.30
100% QCD bg	171.91 ± 0.26	-0.10 ± 0.07	-0.11
odd $\overline{L_{\mathrm{bg}}}$, even events	171.97 ± 0.26	0.02 ± 0.06	-0.03
even $\overline{L_{\mathrm{bg}}}$, odd events	172.02 ± 0.26	0.06 ± 0.06	0.02
`ktfact` = `qfact` = 2.0	171.95 ± 0.26	0.03 ± 0.06	-0.05 ± 0.07
`ktfact` = `qfact` = 0.5	171.92 ± 0.26	0.03 ± 0.06	-0.08 ± 0.07

Table 9.10: Systematics for gluon fraction, evaluated by reweighting the `ttkt72` sample for the nominal $m_t = 172$ GeV/c^2 sample.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)	Δm_t (GeV/c^2)
Nominal	171.93 ± 0.26	0.01 ± 0.06	—
80% $q\bar{q}$, 20% gg	171.93 ± 0.26	0.03 ± 0.07	0.00

9.12 Color reconnection

One important effect so far not considered in our analysis is the effect of color reconnection; that is, the effect of the color connection of the $t\bar{t}$ decay products to the beam remnants (the quarks in the initial $p\bar{p}$ system not involved in the primary interaction). While the PYTHIA samples used in our analysis are generated using PYTHIA 6.2 and do not include this effect, the latest PYTHIA version, PYTHIA 6.4, includes options to perform this color reconnection [61].

To measure the systematic uncertainty, we compare the nominal `ttop25` sample with two new samples generated with PYTHIA 6.4. Tune Apro is an updated version of the standard Tune A, while Tune ACRpro is the same with the color reconnection turned on. We show the results in Table 9.11.

The difference between the Tune Apro and Tune ACRpro tunes is 0.32 ± 0.25 GeV/c^2, which we take as our current color reconnection systematic.

Table 9.11: Color reconnection studies. All samples have a nominal m_t of 172.5 GeV/c^2 and use 24k integrated events.

Sample	Meas. m_t (GeV/c^2)	Meas. Δ_{JES} (σ)
Nominal 6.2 (`ttop25`)	172.90 ± 0.17	0.03 ± 0.04
6.4 Tune Apro (`ctopsd`)	172.77 ± 0.18	0.05 ± 0.04
6.4 Tune ACRpro (`otop46`)	172.45 ± 0.18	0.17 ± 0.04

Chapter 10

Conclusion

In this thesis, I have presented the results of a very precise top quark mass measurement. Using a total of 738 events before and 630 events after the likelihood cut, we have obtained a measured value of:

$$\begin{aligned} m_t &= 172.6 \pm 0.9 \text{ (stat.)} \pm 0.7 \text{ (JES)} \pm 1.1 \text{ (syst.) GeV}/c^2 \\ &= 172.6 \pm 1.6 \text{ (total) GeV}/c^2. \end{aligned}$$

This measurement was performed using a matrix element method, which has proven itself to produce the most precise top mass measurements. However, the enhancements included in this analysis, specifically the use of Quasi–Monte Carlo integration to allow us to integrate over a higher-dimensionality phase space, have resulted in a measurement superior to other matrix element analysis and one that is currently the single most precise top mass measurement in the world.

Of course, top mass measurement is an extremely competitive field, and we are always looking for ways to improve. As we saw in Section 7.6, the bad signal events present a serious problem to our model, as they result in a reduced resolution despite the fact that they still contain useful information. We are currently actively researching a method to handle these bad signal events by integrating over the $t\bar{t}$ jet that is lost when one of these events takes place, and we hope that this will result in a method which will allows us to recover the information about the top mass still remaining in bad signal events.

Other potential areas of improvement include the background treatment. While our neural network shows promising results, to date, we have not seen a significant improvement from using it, suggesting that it needs to be more effective at discriminating signal from background in order to produce a gain in resolution. The possibility also exists that we could replace the current background method with a full background likelihood, which would be a more accurate model than the current method.

Overall, we are pleased to have demonstrated the success of our method in measuring the top quark mass, and hope to improve upon it in the future.

Bibliography

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

[2] F. Abe *et al.* (CDF Collaboration), Phys. Rev Lett. **74**, 2626 (1995).

S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett **74**, 2632 (1995).

[3] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **122**, 103 (1983).

P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B **122**, 122 (1983).

G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **126**, 398 (1983).

P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B **129**, 130 (1983).

[4] CDF Collaboration, CDF Notes 9196, 9301, and 9427 (unpublished).

[5] V.M. Abazov *et al.* (D0 Collaboration), Nature **429**, 638 (2004).

[6] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 092002 (2006).

T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **99**, 182002 (2007).

A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 032009 (2006).

V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett **101**, 182001 (2008).

[7] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **79**, 072001 (2009).

[8] C. Amsler *et al.* (Particle Data Group), Phys. Lett. B **667**, 1 (2008).

[9] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998).

[10] S. Fanchiotti, B. Kniehl, and A. Sirlin, Phys. Rev. D **48**, 307 (1993).

[11] A. H. Hoang and I. W. Stewart, Nucl. Phys. Proc. Suppl. **185**, 220 (2008) [arXiv:0808.0222].

[12] P. W. Higgs, Phys. Rev. Lett **12**, 132 (1964).

[13] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Westview Press, 1995).

[14] Tevatron New Phenomena and Higgs Working Group, Fermilab publication FERMILAB-PUB-09-060-E (2009) [arXiv:0903.4001].

[15] Tevatron Electroweak Working Group, Fermilab publication FERMILAB-TM-2427-E (2009) [arXiv:0903.2503].

[16] CDF Collaboration, CDF note 9448 (unpublished).

[17] W. J. Marciano, arXiv:hep-ph/0411179.

[18] LEP Electroweak Working Group, http://lepewwg.web.cern.ch/LEPEWWG/

[19] SVXII: A. Sill *et al.*, Nucl. Instrum. Methods A **350**, 1 (2004).

L00: C.S. Hill, Nucl. Instrum. Methods A **447**, 1 (2000).

ISL: A. Affolder *et al.*, Nucl. Instrum. Methods A **453**, 84 (2000).

[20] T. Affolder *et al.*, Nucl. Instrum. Methods A **526**, 249 (2004).

[21] L. Balka *et al.*, Nucl. Instrum. Methods A **267**, 272 (1988).

[22] S. Bertolucci *et al.*, Nucl. Instrum. Methods A **267**, 301 (1988).

[23] M. Albrow *et al.*, Nucl. Instrum. Methods A **480**, 524 (2002).

[24] G. Apollinari *et al.*, Nucl. Instrum. Methods A **412**, 515 (1998).

[25] G. Ascoli *et al.*, Nucl. Instrum. Methods A **268**, 33 (1988).

[26] E. J. Thomson *et al.*, IEEE Trans. on Nucl. Science **49**, 1063 (2002).

[27] W. Ashmanskas *et al.*, Nucl. Instrum. Methods A **518**, 532 (2004).

[28] A. Abulencia *et al.* (CDF Collaboration), J. Phys. G Nucl. Part. Phys **34**, 2457 (2007).

[29] F. Abe *et al.*, Phys. Rev. Lett **68**, 1104 (1992).

[30] A. Bhatti *et al.*, Nucl. Instrum. Methods A **566**, 375 (2006).

[31] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 092002 (2006).

[32] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005).

[33] T. Aaltonen *et al.* (CDF Collaboration), arXiv:0903.0885, submitted to Phys. Rev. Lett.

[34] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 072005 (2005).

D. Sherman, Ph. D. thesis, Harvard University, 2007.

[35] PYTHIA version 6.216, T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Computer Phys. Commun. **135**, 238 (2001) [arXiv:hep-ph/0010017].

[36] HERWIG version 6.510, G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001) [arXiv:hep-ph/0011363]; arXiv:hep-ph/0210213.

[37] ALPGEN version 2.10 prime, M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, JHEP **0307**, 001 (2003) [arXiv:hep-ph/0206293]. Note that PYTHIA v6.325 is used for the showering.

[38] MadGraph/MadEvent v4, J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. Rainwater, T. Stelzer, JHEP **0709**, 028 (2007) [arXiv:0706.2334].

[39] E. Gerchtein and M. Paulini, in *Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, March 2003*, pp. TUMT005 (2003) [arXiv:physics/0306031].

[40] K. Kondo, J. Phys. Soc. Jpn. **57**, 4126 (1988).

R. H. Dalitz, G. R. Goldstein, and K. Sliwa, Phys. Rev. D **47**, 967 (1993).

R. H. Dalitz and G. R. Goldstein, Proc. Royal Soc. London A **445**, 2803 (1999).

[41] R. Kleiss and W. J. Stirling, Z. Phys. C **40**, 419 (1988).

[42] H. L. Lai *et al.*, Eur. Phys. J. C **12**, 375 (2000).

[43] B. Abdous and E. Bensaid, J. Nonparametr. Stat. **13**, 77 (2000).

[44] D. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization* (Wiley, 1992).

[45] T. J. Hastie, SLAC publication SLAC-PUB-3160 (1983).

[46] G. Takhtamyshev, B. Vandewoestyne, and R. Cools, Math. Comput. Simul. **73**, 309 (2007).

[47] M. A. Jenkins and J. F. Traub, ACM T. Math. Software **1**, 178 (1975).

[48] J. C. Freeman, Ph.D. thesis, UC Berkeley, 2007, FERMILAB-THESIS-2007-59. The Jacobian calculation is on pages 67-69.

[49] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1992).

[50] W. Morokoff and R. E. Caflisch, SIAM J. Sci. Stat. Comp. **15**, 1251 (1994).

[51] H. S. Hong and F. J. Hickernell, ACM T. Math. Software **29**, 95 (2003). We use the Faure-Tezuka algorithm described.

[52] I. Sloan and H. Wozniakowski, J. Complexity **14**, 1 (1998).

[53] R. Kleiss and A. Lazopoulos, Comput. Phys. Commun. **175**, 93 (2006).

[54] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **99**, 182002 (2007). D0 Collaboration, D0 note 5877-CONF.

[55] C. Peterson, T. Rögnvaldsson, and L. Lönnblad, Comput. Phys. Commun. **81**, 185 (1994).

[56] B. Efron, Ann. Statist. **7**, 1 (1979).

B. Efron and R. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall/CRC, 1994).

[57] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 032003 (2006). Note that we use a procedure slightly different from the one described in this article; we use samples in which the ISR and FSR have been shifted together by a $\pm 1\sigma$ variation. The determination of the 1σ uncertainty is the same as described in this article.

[58] M. Bowler, Z. Phys. C **22**, 155 (1984).

[59] Y. Peters, K. Hamacher and D. Wicke, Fermilab publication FERMILAB-TM-2425-E (2009).

[60] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **4**, 463 (1998).

[61] D. Wicke and P. Z. Skands, arXiv:0807.3248 (2008).

Appendix A

Validation of the Data

Since this analysis is necessarily reliant on Monte Carlo simulation, we make a variety of validation plots to check that various properties of the observed data do indeed agree well with the Monte Carlo. The following plots show some quantities of interest plotted for the data against the Monte Carlo for events which pass our $t\bar{t}$ selection requirements, where the Monte Carlo expectation is derived by taking the samples described in Section 4.5 added together with the expected fractions described in Section 4.4. The confidence level returned by a K-S test is indicated on the plots. The $t\bar{t}$ signal sample used here is `ttop25`, which has a top mass of 172.5 GeV/c^2 and a luminosity profile matching periods 0 through 19 of CDF data.

Figure A.1 shows a comparison of the jet E_T corrected to level 5 (as described in Section 3.4.2) for the four tight jets, as well as for the fifth jet, if present. While we do not actually use a fifth jet in this analysis, we include this plot for comparison purposes. Figure A.2 shows a comparison of the $\not\!\!E_T$, the total number of tight and loose jets, the number of z vertices, the jet E_T for jets with a b-tag, and the lepton p_T. Again, while we do not use loose jets (which are defined as jets with $E_T > 12$ GeV/c^2 and $|\eta| < 2.4$), we include them here for comparison purposes. Finally, Figure A.3 shows a comparison between the leading jet E_T and the lepton p_T for

events containing an electron and events containing a muon.

In general, the plots indicate a good level of agreement between the data and the Monte Carlo. The main exception is the number of z vertices. This is due to the fact that, while the $t\bar{t}$ signal sample is designed to match the luminosity in periods 0-19 of CDF operation, the most recent data has a higher luminosity and hence a higher number of $p\bar{p}$ interactions in a given event. Potential systematics for this difference are addressed by the multiple hadron interaction systematic in Section 9.8.

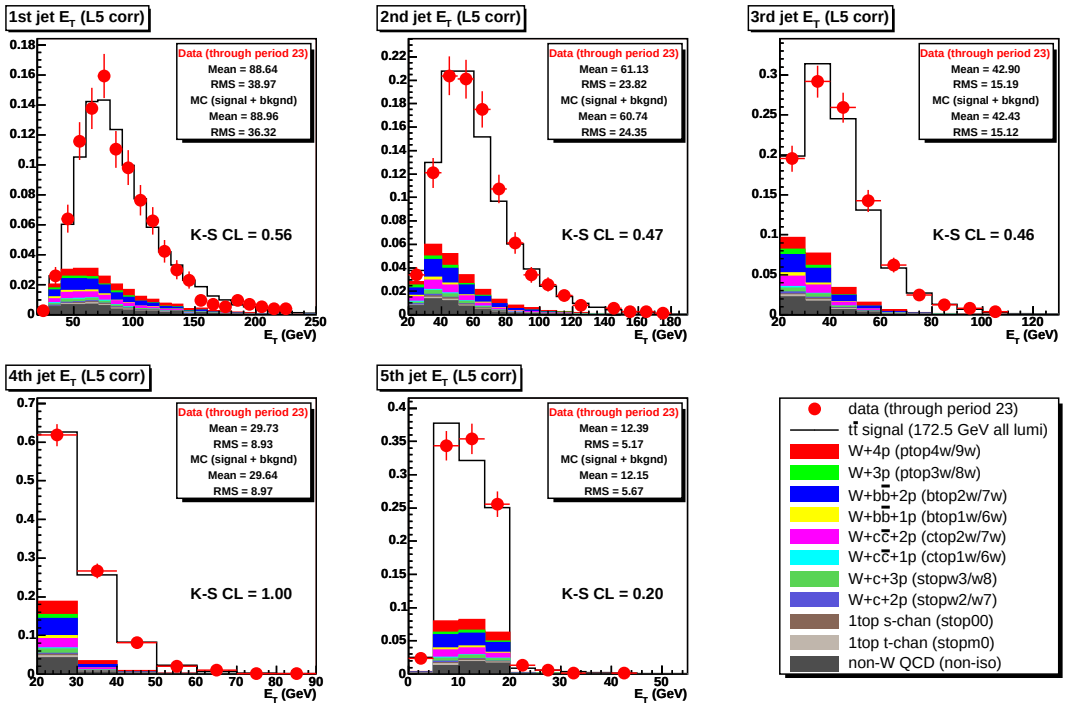

Figure A.1: Comparison of jet energies between data and Monte Carlo for events passing our selection cuts.

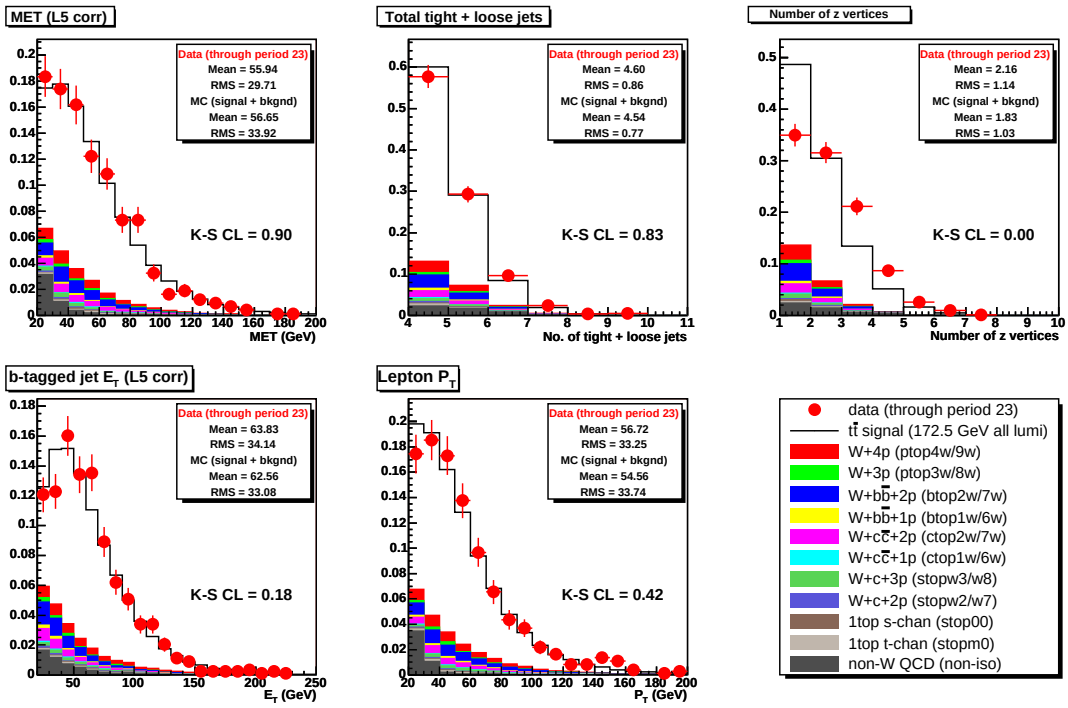

Figure A.2: Comparison of other quantities of interest between data and Monte Carlo for events passing our selection cuts.

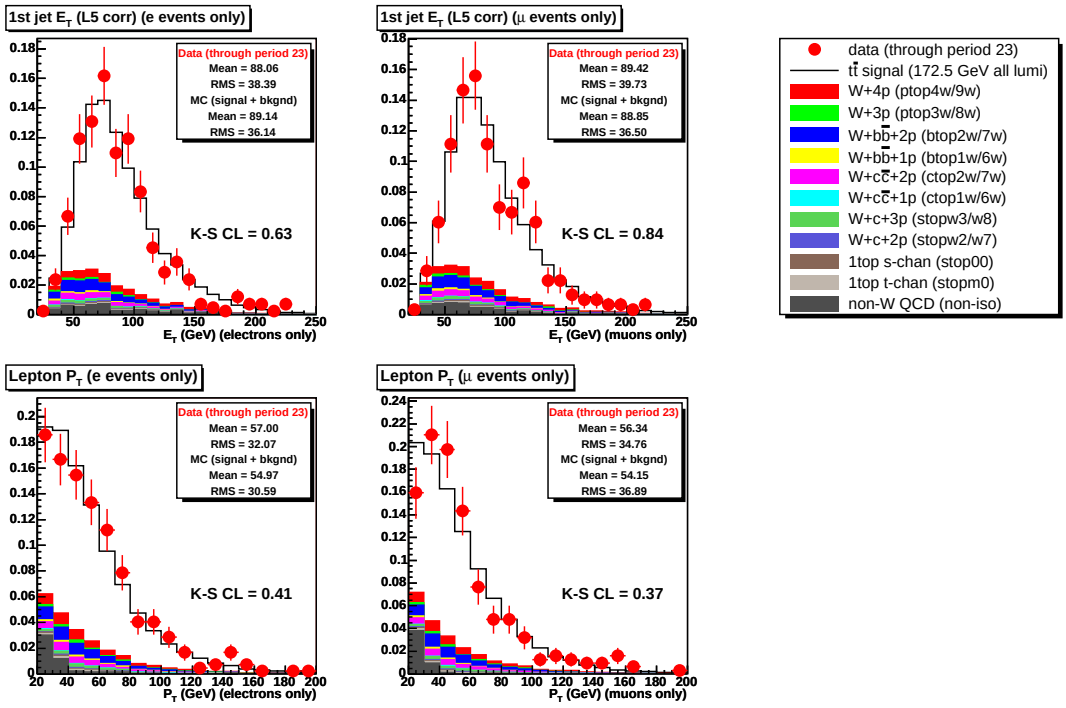

Figure A.3: Comparison of leading jet E_T and lepton p_T for events containing an electron (left) and events containing a muon (right).

Appendix B

Avoiding Phase Space Singularities

Because solving the kinematic equations to determine the parton momentum requires an 8th-degree polynomial, the space of allowed solutions can be rather complex. Specifically, when considering the allowed values of the squared leptonic W mass $M^2_{W,\mathrm{lep}}$ vs. the neutrino z-momentum $p_{\nu z}$, we can encounter problems in the Jacobian for transforming between these two variables.

Figure B.1 illustrates (in somewhat simplified form) the basic problem. We can see that $\frac{\partial M^2_W}{\partial p_{\nu z}}$ goes to zero at a point. Thus, if we transform our integral from depending on $p_{\nu z}$ to depending on M^2_W, the Jacobian for this transformation will go to zero at that point, and so, since the Jacobian is in the denominator, the integrand will become infinite. (Of course, the full Jacobian is slightly more complicated, but the problem is still the same.) Similarly, using $p_{\nu z}$ as the integration variable will also fail in some cases, such as the one illustrated in Figure B.2.

To avoid these problems, the integration code switches between the two integration variables as necessary: normally M^2_W is used, but when the distance to the minimum M^2_W point becomes less than two grid points, the integration switches to using $p_{\nu z}$ instead (with the Jacobian changing appropriately). The code is also

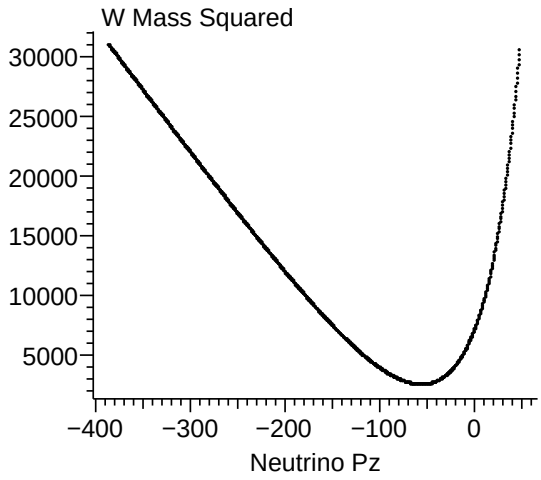

Figure B.1: One possible set of solutions to the kinematic equations in the space of $M^2_{W,\mathrm{lep}}$ vs. $p_{\nu z}$.

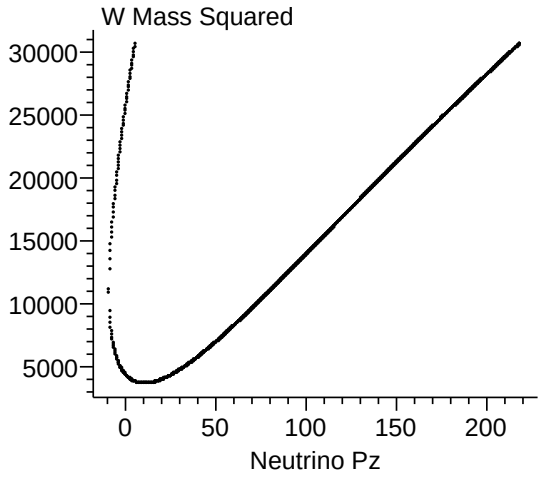

Figure B.2: A set of solutions in which both $\frac{\partial M^2_W}{\partial p_{\nu z}}$ and $\frac{\partial p_{\nu z}}{\partial M^2_W}$ are zero at different points.

designed to handle particularly complicated cases such as that shown in Figure B.3. In this way, we can still use the importance sampling technique as much as possible, thus reducing our integration time, but also avoid the singularities in the integration.

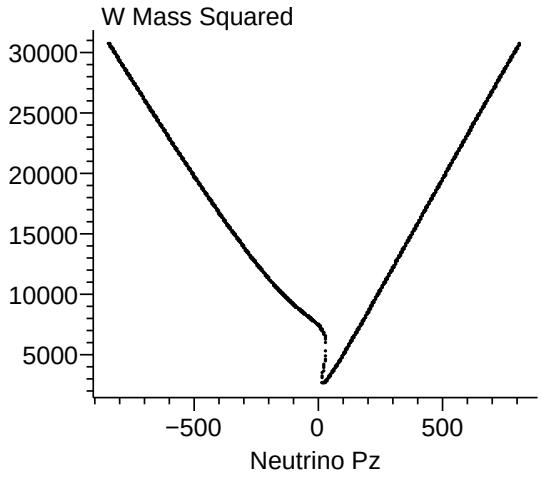

Figure B.3: A set of solutions which requires very careful treatment of possible singularities in the Jacobian.