Measurement of $t\bar{t}$ Helicity Fractions and Spin Correlation in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.96 TeV

by

David John Mietlicki

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2010

Doctoral Committee:

Professor Dante Eric Amidei, Chair
Professor Ratindranath Akhoury
Professor Keith Riles
Professor Bing Zhou
Assistant Professor Donald Malcolm Tulip

DEDICATION

To Laura, David, and Sarah

ACKNOWLEDGMENTS

The work presented here would not have been possible without the help of many others for whom I will always be grateful.

First, and most importantly, I would like to acknowledge my wife Laura. Without her love, support, and motivation, I would have been unable to complete this work. I am thankful every day that she is part of my life. I would also like to thank our children, David and Sarah, the greatest gifts I have ever received. They are my constant reminders of what is truly important in life.

I also want to thank my parents, Dave and Janet Mietlicki, for everything they have done for me. They were the ones to first instill in me a love of learning, and without their support I would not have been able to reach this point.

I am truly indebted to my advisor, Dan Amidei, for everything he has taught me and for all the guidance I have received from him. It is his knowledge, effort, and passion for discovery that helped bring this dissertation to fulfillment, and I look forward to working with him in the future as a collaborator and friend.

The work presented in this dissertation would not have been possible without the dedication of hundreds of members of the CDF collaboration. In particular, I would like to thank Alexei Varganov, fellow member of the Michigan CDF group. His efforts and ideas played an important role in the development of this measurement, and I am excited to continue working with him in the future.

I would also like to thank the other current and recent members of the Michigan CDF group: Glenn Strycker, Andrew Eppig, Clark Cully, Tom Wright, Monica Tecchio, Dave Gerdes, and Myron Campbell. They were an invaluable source of help and

support for me. I additionally want to thank Tom Schwarz and Nathan Goldschmidt, former University of Michigan graduate students and current CDF collaborators, for the help they have given me in all of my work at CDF.

Finally, I thank the members of my committee for their time and effort in helping me through this process.

CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER 1

Introduction

The current theoretical model of particle physics predicts that there are two basic types of fundamental particles: force carrying particles called bosons, and matter particles called fermions. These fermions are further divided into two groups, called leptons and quarks. Leptons, like the electron, interact via the electromagnetic and weak interactions, while quarks, such as the up and down quarks that are the fundamental constituents of protons and neutrons, also interact strongly through what is called the quantum chromodynamic (QCD) force. The top quark, which is the most massive known fundamental particle, was discovered in March 1995 by the CDF and D0 experiments at the Fermilab Tevatron [1]. Since that discovery, research into the properties of the top quark has been one of the most active fields in experimental particle physics.

The two most precisely measured properties of the top quark are its mass, $M_{top} = 173.1 \pm 1.3$ GeV/c^2 [2], and the production cross section for $t\bar{t}$ pairs at the Tevatron, $\sigma_{t\bar{t}} = 7.50 \pm 0.48$ pb[3]. Because of its very large mass, the top quark is expected to couple more strongly than other known particles to a number of hypothetical undiscovered physical processes [4]. These new couplings could be observed as deviations from the theoretically predicted values for various top quark properties, motivating a thorough investigation of the top quark and its interactions.

The best measurements of top quark properties currently come from the CDF and D0 experiments as they analyze data collected in $p\bar{p}$ collisions at the Tevatron

with a center of mass energy of 1.96 TeV. Top quark pair production by the QCD interaction at the Tevatron is dominated by the quark-antiquark annihilation process, $q\bar{q} \to t\bar{t}$, which accounts for 85% of the $t\bar{t}$ pairs produced, with the remaining 15% being produced via gluon fusion, $gg \to t\bar{t}$ [5]. Top quarks can also be produced singly at the Tevatron via the electroweak interaction, but this process is rarer than $t\bar{t}$ production and also much more difficult to see experimentally. This single top quark production was not observed at the Tevatron until 2009 [6].

Theoretically, the decay of a top quark is predicted to occur almost always via $t \to Wb$, creating a bottom quark and a W boson. This decay occurs with a very large total decay width of $\Gamma = 1.50$ GeV [7]. The overall top quark pair production and decay process at the Tevatron is given by $p\bar{p} \to t\bar{t} \to W^+bW^-\bar{b}$. The various possible final states are classified by the way in which the two W bosons in the event decay:

- Dilepton Channel: $t\bar{t} \to W^+bW^-\bar{b} \to (\bar{\ell}\nu_\ell b)(\ell'\bar{\nu_{\ell'}}\bar{b})$[1,2]
- Lepton Plus Jets Channel: $t\bar{t} \to W^+bW^-\bar{b} \to (q\bar{q'}b)(\ell\bar{\nu}_\ell\bar{b})$ or $(\bar{\ell'}\nu_{\ell'}b)(q''\bar{q'''}\bar{b})$
- All Hadronic Channel: $t\bar{t} \to W^+bW^-\bar{b} \to (q\bar{q'}b)(q''\bar{q'''}\bar{b})$

Approximately 12% of $t\bar{t}$ decays occur via the dilepton channel, while the lepton plus jets and all hadronic channels each account for approximately 44% of top pair decays. For the measurement described in this thesis, we focus on the lepton plus jets decay channel, and we additionally require that the lepton be either an electron or a muon. In Chapters 3 and 5, we will discuss how we select such events and how we use the final state decay products to reconstruct the momenta of the original top quarks respectively.

One of the important consequences of the large decay width of the top quark is that the top quark has a very short lifetime of approximately $0.5 \cdot 10^{-24}$ seconds [8].

[1]The ($'$) denotes a different type of lepton or quark

[2]We will sometimes use the notation $t \to \ell\nu b$ in cases where where the distinction between particle and antiparticle is unimportant

This short lifetime means that the top decays before hadronization caused by the QCD interaction can take place. In the lighter quarks, hadronization causes multiple changes to the quark spin before decay, resulting in most information about the initial spin of the quark being lost. However, since the top quark decays before hadronization occurs, information about its spin is passed on to the decay products [4]. Because of this, the top quark offers a unique opportunity in that we can directly observe the spin imparted to a top quark at production, something which is not possible for the lighter quarks.

The concept of spin was first introduced in 1924 by Wolfgang Pauli, although at the time he did not use the term "spin". It was introduced as an internal degree of freedom of the electron, in order to explain observations in atomic physics, but it soon became apparent that spin is an important property of all particles [9]. Although the spin of a particle is an internal property and does not correspond to a physical rotation, it was named spin because it does have some properties that are similar to a physical rotation - for example, it adds to a particle's total angular momentum, and causes a charged particle to interact with a magnetic field.

Leptons and quarks, including the top quark, have a total value for their spin of $\frac{1}{2}$. When measuring spin experimentally, a particular axis called the "spin quantization axis" must be chosen. The spin is measured along that axis and can have one of two values. If a top quark's spin is oriented in the same direction as the quantization axis, it has spin $+\frac{1}{2}$, and if it is oriented opposite the direction of the axis, it has spin $-\frac{1}{2}$. The choice of the quantization axis is also referred to as a choice of basis, and there are three common bases for measuring top quark spin in collider experiments:

- Helicity Basis: the spin quantization axis is the direction of motion of the top quark in the $t\bar{t}$ rest frame.
- Beamline Basis: the spin quantization axis is the direction of the incident colliding particles, called the "beamline".

- Off-Diagonal Basis: a hybrid basis where the spin quantization axis is in-between the direction of the axis in the helicity basis and the direction of the axis in the beam basis, with the exact choice of axis dependent on the kinematics of the event.

For this analysis, we work in the helicity basis, providing the first measurement of the top quark spin correlation in this basis. In the helicity basis, when a top quark's spin is oriented in the same direction as the quark's direction of motion in the $t\bar{t}$ rest frame, it is said to be a right-handed top quark (t_R), while when the spin is oriented against the quark's direction of motion, it is called a left-handed top quark (t_L).

1.1 The Spin Correlation

When looking at an individual top quark in a $t\bar{t}$ event, the helicity of that top quark is equally likely to be left-handed or right-handed. However, because the initial state producing the $t\bar{t}$ pair will have a particular angular momentum, which is conserved, there will be a correlation between the helicities of the two top quarks in an event. There are four possible helicity states for a top quark pair, depending on whether the top and antitop quarks are right- or left-handed: $t_R\bar{t}_R$, $t_L\bar{t}_L$, $t_L\bar{t}_R$, and $t_R\bar{t}_L$. Two of these states have top quarks with the same helicity and two have top quarks with opposite helicities.

Our analysis will measure the fraction of opposite helicity top quark pairs F_{OH}, as defined in Equation 1.1, where N_o and N_s are the numbers of events with opposite and same helicity pairs respectively.

$$F_{OH} = \frac{\sigma(t_L\bar{t}_R) + \sigma(t_R\bar{t}_L)}{\sigma(t_R\bar{t}_R) + \sigma(t_L\bar{t}_L) + \sigma(t_L\bar{t}_R) + \sigma(t_R\bar{t}_L)} = \frac{N_o}{N_o + N_s} \tag{1.1}$$

The same helicity fraction F_{SH} can be defined similarly, and it is apparent that $F_{OH} + F_{SH} = 1$.

As will be described in more detail in Chapter 6, $t\bar{t}$ production at the Tevatron

is dominated by the process $q\bar{q} \rightarrow t\bar{t}$, where the top quarks are produced via a gluon with spin 1, so the total angular momentum of the $t\bar{t}$ system must be 1. If the two top quarks have no relative orbital angular momentum, then conservation of angular momentum requires that the top quarks must spin in the same direction. In this case, because the two top quarks move in opposite directions in the $t\bar{t}$ rest frame, they will have opposite helicities. This is shown schematically in Figure 1.1. At the energy threshold for $t\bar{t}$ production (approximately 345 GeV), 67% of top quark pairs produced via $q\bar{q} \rightarrow t\bar{t}$ will have opposite helicities, and this fraction rises to 100% at very high energies because conservation of helicity demands that quark pairs with energies much larger than the quark masses must have opposite helicities. The small contribution from $gg \rightarrow t\bar{t}$ tends to produce top quarks with the same helicity, since in this case the total angular momentum of the $t\bar{t}$ system is 0. Integrating over the distributions for top quark momenta at the Tevatron, the theoretically predicted value for the opposite helicity fraction at leading order is $F_{OH} = 0.7$ [8, 10, 11]. If there were no top quark spin correlation, we would find $F_{OH} = 0.5$.

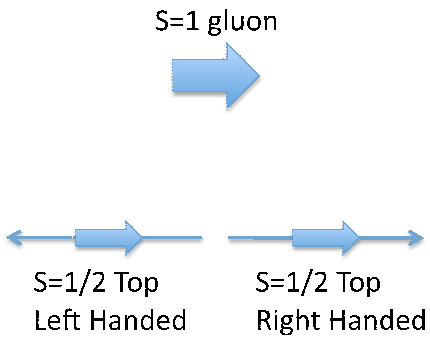

Figure 1.1. Schematic decay of a gluon with S=1 into opposite helicity top quarks

The correlation between top quark spins is often quantified by another variable, known as the top quark spin correlation coefficient κ, defined in Equation 1.2 [8].

$$\kappa = \frac{[\sigma(t_R\bar{t}_L) + \sigma(t_L\bar{t}_R)] - [\sigma(t_R\bar{t}_R) + \sigma(t_L\bar{t}_L)]}{\sigma(t_R\bar{t}_R) + \sigma(t_L\bar{t}_L) + \sigma(t_R\bar{t}_L) + \sigma(t_L\bar{t}_R)} = \frac{N_o - N_s}{N_o + N_s}. \tag{1.2}$$

The value of κ must be between -1 and +1, and if there were no top quark spin correlation, then we would measure $\kappa = 0$. Using the definition of F_{OH} and the fact that $F_{OH} + F_{SH} = 1$, we can rewrite κ in terms of F_{OH} as $\kappa = 2 \cdot F_{OH} - 1$. Thus, our measurement of F_{OH} is actually a measurement of κ as well. Using $F_{OH} = 0.7$, we can use this relation to find the theoretical prediction $\kappa = 0.4$ in the helicity basis. Another theoretical calculation of κ which includes next-to-leading order effects, such as extra gluon radiation in the $t\bar{t}$ production and decay process, predicts $\kappa = 0.352$ in the helicity basis at the Tevatron [12]. CDF also recently released a result measuring κ in the off-diagonal basis using dilepton decays, finding $\kappa = 0.32^{+0.55}_{-0.78}$ [13].

In measuring the spin correlation, we will study distributions of the helicity angles of the top quark decay products. The helicity angle of a decay product is defined to be the angle between the direction of motion of that decay product (in the top quark rest frame) and the direction of motion of the top quark (in the $t\bar{t}$ rest frame). As we will see in Chapter 6, the correlation of these angles with the top quark spin is given by the function $f(\cos(\theta_i)) = \frac{1}{2}(1 + \alpha_i \cos(\theta_i))$, where the subscript i specifies the decay product and the coefficient α_i determines the strength of the correlation between the decay product helicity angle and the top quark spin. In the leptonic decay $t \to \ell\nu b$, the lepton has the strongest correlation with the top quark spin, with $\alpha_\ell = 1$. For the hadronic decay $t \to udb$[3], it is the down-type quark that has this strong correlation $\alpha_d = 1$. The $t\bar{t}$ differential production cross section can be written in terms of the decay product helicity angles and the correlation coefficient κ as in Equation 1.3.

$$\frac{1}{\sigma}\frac{d^2\sigma}{d(\cos\theta_i^+)d(\cos\theta_j^-)} = \frac{1 - \kappa\alpha_i\alpha_j\cos\theta_i^+\cos\theta_j^-}{4} \tag{1.3}$$

By studying the distributions of helicity angles for decay products from both the top

[3]When we refer to a down quark, it always means either a down quark or a strange quark. Similarly, a reference to an up quark means either an up quark or a charm quark.

and antitop quarks in a $t\bar{t}$ event, we can determine the spin correlation of the original top quark pair.

The spin correlation measurement is important not only as a test of the theoretical prediction, but also as a possible factor in helping to understand new physical interactions. Known interactions that can produce top quark pairs are well understood, leading to accurate predictions for the $t\bar{t}$ spin correlation. Any deviation of the measured spin correlation from the theoretical predictions would indicate new top quark production processes. An example of such a process would be $t\bar{t}$ production via a Kaluza-Klein graviton in a Randall-Sundrum model [15], which would alter the spin correlation because unlike the gluon, the graviton has a total spin of 2. In particular, resonant production of such a graviton would produce large changes in the spin correlation as a function of the $t\bar{t}$ invariant mass compared to current theoretical predictions. Precision studies of the spin correlation and its behavior as a function of $t\bar{t}$ mass could provide evidence for or against this model and many other models of new physical processes.

A recent CDF measurement found an asymmetry in the direction that the top quark from a $t\bar{t}$ pair travels after it is produced that disagrees with the theoretical prediction by approximately 2σ [16]. Many models of new physical processes have been proposed to explain this discrepancy, and many of these models also predict that the new interactions will have an effect on the top quark pair spin correlation [17]. A precise measurement of the spin correlation would differentiate among these models, probing new physical interactions.

1.2 Measurement Plan

We plan to measure the top quark pair spin correlation in the helicity basis using events in the lepton plus jets decay channel. This is the first time that such a measurement has been made using this particular decay channel. Previous measurements

of the spin correlation have been performed [13, 14], but used the dilepton decay channel because the lepton has a strong correlation to the top spin and is easily identifiable. However, this channel has a small branching ratio, resulting in low counting statistics, and also suffers from difficult event reconstruction because of the presence of two undetectable neutrinos.

This first measurement in the lepton plus jets channel takes advantage of the larger branching ratio and the fact that lepton plus jets events can be fully reconstructed. We rely critically on the ability to identify the down quark, which has the same strong correlation to the top spin as the lepton. Because the QCD interaction causes the hadronization of light quarks into jets, a down quark jet measured in the detector is indistinguishable from a jet that originated from any other type of light quark. However, we will show in Chapter 5 that it is possible to identify the down quark reliably after reconstruction of the event kinematics.

In order to make our measurement, we will study distributions involving the decay product helicity angles θ_i. Because these helicity angles are correlated with the top spin, the helicity angles of decay products from both the top and antitop decays will contain information about the $t\bar{t}$ spin correlation. In order to extract this information, we will create a simulated model $t\bar{t}$ sample which contains only top quark pairs with the same helicity and another simulated model sample which contains only top pairs with opposite helicities. We will measure the opposite helicity fraction and spin correlation by determining the relative normalization of these two model samples which best describes our data.

In Chapter 2 we will describe the Tevatron and the CDF detector. Chapter 3 discusses the event selection for $t\bar{t}$ events in the lepton plus jets decay channel, while Chapter 4 describes our method for modeling the non-$t\bar{t}$ events which pass our selection. In Chapter 5 we will discuss the reconstruction of $t\bar{t}$ events and the identification of the down quark, as well as discussing the validation of our model.

Chapter 6 contains additional information about the production and decay of $t\bar{t}$ pairs, as well as describing the creation of the same helicity and opposite helicity simulated samples. Chapter 7 details our measurement method, followed by a discussion of our systematic uncertainties in Chapter 8. Finally, Chapter 9 presents the result of our measurement of the opposite helicity fraction F_{OH} and the correlation coefficient κ.

CHAPTER 2

Experimental Apparatus

This analysis uses data collected by the Collider Detector at Fermilab (CDF) experiment at the Fermilab Tevatron. Located just outside of Chicago, Illinois, Fermilab was founded in 1967 and is currently the largest laboratory for particle physics in the United States. The CDF experiment was built and is currently operated by a collaboration of several hundred scientists from around the world. The Tevatron collides protons and antiprotons at a center of mass energy of $\sqrt{s} = 1.96$ TeV, and it was only recently supplanted as the highest energy particle accelerator in the world by Large Hadron Collider (LHC), located at CERN in Geneva, Switzerland, which currently collides two proton beams at an energy of $\sqrt{s} = 7$ TeV.

2.1 The Tevatron

Built in 1983, the Tevatron accelerator was the first superconducting synchrotron in the world. It is located at the end of a chain of seven Fermilab accelerators which are used together to accelerate protons and antiprotons to a center of mass energy of 1.96 TeV [18]. The accelerator complex is shown schematically in Figure 2.1.

The acceleration process begins inside a Cockcroft-Walton accelerator, where H^- ions are created by adding an extra electron to hydrogen atoms. A static electric field is used to accelerate these ions to an energy of 750 keV, at which point they are grouped into bunches and enter a 500 foot long linear accelerator. Pulses of radio frequency electromagnetic waves in copper resonating cavities are used to accelerate

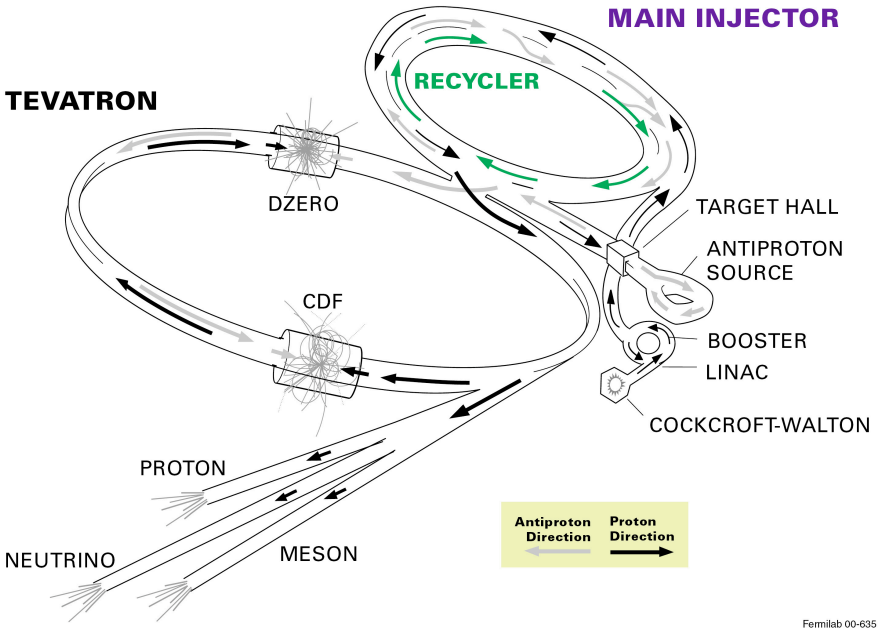

Figure 2.1. The Fermilab Accelerator Complex

the ion beam to an energy of 400 MeV. At the end of this linear accelerator, the ion beam passes through a carbon foil that strips the two electrons from each ion, leaving a beam of protons.

This proton beam then enters a circular synchrotron accelerator called the Booster. 500 feet in diameter, this accelerator collects protons from the linear accelerator in such a way as to create discrete proton bunches. These bunches are then accelerated, again using radio frequency (RF) cavities. The proton beam is directed along its circular path using dipole magnets. The beam circles the Booster approximately 20,000 times, reaching an energy of 8 GeV, at which point it is ready to be passed on to the next accelerator in the chain, the Main Injector. Only a portion of the protons that reach this stage are actually passed on to the Main Injector, while the rest are used for other experimental programs at Fermilab.

The Main Injector, another synchrotron, serves several purposes. At this point in the accelerator chain, it accelerates the proton beam from 8 GeV to 120 GeV.

Some of these 120 GeV protons are used for Fermilab's other experiments, while the rest are sent to the antiproton source to create antiprotons. These protons strike the antiproton target, a nickel disk, every 1.5 seconds and create a shower of new particles. The particles created are focused using a lithium lens, and then the negatively charged antiprotons, which have an average energy of 8 GeV, are separated from the other particles created by passing the shower of particles through a magnetic field. For every 100,000 protons which strike the target, only 1 antiproton is created, so a very large number of protons is needed to accumulate a usable antiproton beam.

Once the antiprotons have been filtered from the other particles, they enter an accelerator called the Debuncher. At this point, the antiprotons are accelerated to a uniform energy of 8 GeV. Because the antiprotons enter the Debuncher with a spread of energies, the result of this process is that, although the energy of the beam becomes uniform, the antiprotons are now spread out in space and no longer in bunches. The beam is then passed to the Accumulator, which is located in the same tunnel as the Debuncher. By making the transfers from the Debuncher to the Accumulator at discrete moments in time, a process called "stacking", the bunch structure is returned to the antiproton beam. The antiprotons are collected in the Accumulator over several hours, at which point they are transferred to the Recycler, a fixed energy storage ring located in the same tunnel as the Main Injector. It can take 12-24 hours for a sufficient number of antiprotons to become available for loading into the Tevatron.

When enough antiprotons have been created, the Main Injector is used to accelerate proton and antiproton bunches from 8 GeV to 150 GeV and then inject these bunches into the Tevatron. With a radius of 1 kilometer and a 53.1 MHz RF used for acceleration, there are 1,113 RF wavelengths along the circumference of the Tevatron. The spacing between the proton and antiproton bunches is determined by the radio waves, resulting in 1,113 slots, called "buckets", in which the proton and antiproton

bunches are able to travel. At the beginning of a run, 36 bunches of protons with an energy of 150 GeV are injected into the Tevatron from the Main Injector, and 36 antiproton bunches are injected in the opposite direction. The proton bunches generally contain several times the number of particles that are in the antiproton bunches, since protons are much easier to obtain. The 36 bunches are injected in three groups of 12, called "trains". Within a train, there are 20 empty buckets between each bunch, and there are larger spaces between each train. If the beam becomes unstable for any reason, the large gap between trains allows time for abort switch magnets to turn on and abort the beam, directing it into a concrete block before it can do any damage to the accelerator or the detectors.

After all proton and antiproton bunches have been loaded into the Tevatron, the beams are accelerated once more from 150 GeV to 980 GeV. Large steel blocks called "collimators" are inserted close to the beam to remove any stray particles, and the transverse beam size is reduced to approximately 2 microns using quadrupole magnets on each side of both the CDF and D0 detectors. At 980 GeV, each bunch circles the 6 km circumference ring approximately 50,000 times each second. Collisions are initiated at the CDF and D0 experiments by electrostatic separators that rotate the helical proton and antiproton beams, creating direct collisions at only two points along the ring. The detectors are built around these two collision points.

Even though trillions of particles are contained in the proton and antiproton beams, only a few collisions - generally fewer than 10 - occur each time proton and antiproton bunches cross. The probability of a collision taking place at a given time can be quantified by the instantaneous luminosity, which is given by Equation 2.1 [7].

$$L = n \cdot f \cdot \frac{N_p \cdot N_{\bar{p}}}{\sigma_x \cdot \sigma_y} cm^{-2} s^{-1} \tag{2.1}$$

In this equation, n is the number of bunches, f is the revolution frequency, N_p and $N_{\bar{p}}$ are the number of protons and antiprotons in each bunch, and σ_x and σ_y are the

average width of a bunch in the transverse directions. The luminosity measures the number of protons that interact with antiprotons per cm^2 per second. For a process that has a given cross-section σ, the luminosity can be used as in Equation 2.2 to determine the number of times that process will occur. Once collisions have begun, the Tevatron can run 30 hours or even longer before it needs to be reloaded, although generally it is refilled approximately every 20 hours in order to keep the instantaneous luminosity as large as possible and maximize the number of collisions observed.

$$N = \int L \cdot \sigma dt \tag{2.2}$$

Because the cross sections for various processes will be constant at a given energy, we can define the integrated luminosity, a measure of the total amount of data collected, as $\mathcal{L} = \int L \cdot dt$. Then the total number of events that will take place for a given process can be written as $N = \mathcal{L} \cdot \sigma$. Since cross sections are often measured in units of femtobarns (fb) and the number of events is a unitless number, the integrated luminosity is measured in units of fb^{-1}. The dataset used in the analysis described here has an integrated luminosity of 4.3 fb^{-1}, which means that for a process that has a cross section of 1 fb, we would expect to observe 4.3 events.

2.2 The CDF Detector

When a collision takes place at the Tevatron, it can be observed and recorded by the CDF detector, a general purpose, longitudinally and cylindrically symmetric particle detector located within and around a solenoidal magnetic field [19]. CDF can measure the charge, momentum, and energy of various types of particles using several different detector systems, each of which is optimized for a particular type of particle or measurement. Figure 2.2 shows the various detector systems, each of which is discussed below [19].

Different types of particles interact differently with the CDF detector. CDF

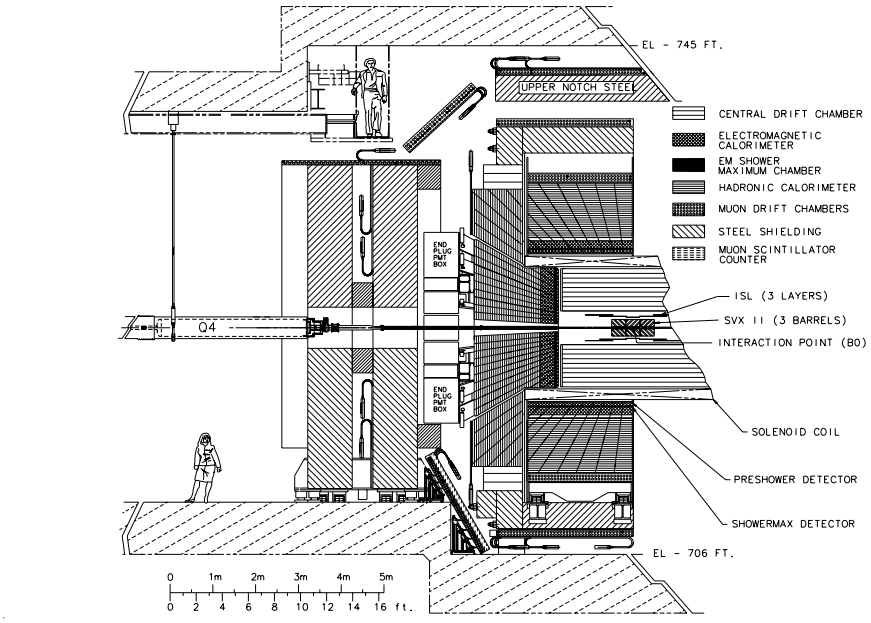

Figure 2.2. Diagram of the CDF Detector

has several basic layers of detector systems: tracking chambers, an electromagnetic calorimeter, a hadronic calorimeter, and muon chambers. There are several different types of objects that we need to identify using the detector: photons, electrons, muons, jets created by the hadronization of quarks, and neutrinos. The way that these particles interact with the detector is summarized in Figure 2.3 [20]. In the CDF detector, the tracking chambers are located inside the magnetic field, while the remaining components are outside of the magnet.

Because photons are neutral, they will not leave a track in the tracking chambers, but will deposit energy in the electromagnetic calorimeter. Electrons, being charged, will leave tracks in the tracking chambers as well as energy in the electromagnetic calorimeter. Muons are charged, and thus leave tracks in the tracking chambers, but they do not interact with most material, so they will not deposit much energy in either calorimeter system, and will travel all the way through the detector to leave hits in the muon chambers. Charged hadrons from quark jets will leave tracks in the tracking chambers and leave some energy in the electromagnetic calorimeter, but most of their energy will be deposited in the hadronic calorimeter. Neutral hadrons

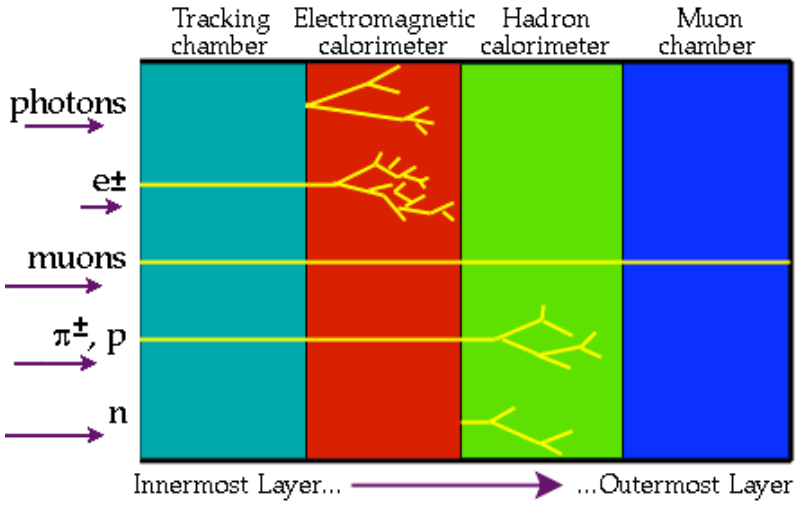

Figure 2.3. How Different Particles Interact with the Detector

in quark jets will not leave tracks, and will deposit little energy in the electromagnetic calorimeter, but will deposit most of their energy in the hadronic calorimeter. Finally, neutrinos do not interact with any part of the detector, and are detectable only by their absence - because they carry energy away but do not interact with the detector, they are detected as an imbalance of momentum left in the detector.

2.2.1 Coordinates and Definitions

The CDF detector can be described by a right-handed spherical coordinate system with the origin at the center of CDF and the z-axis pointing along the direction of the proton beam. The x-axis, from which the azimuthal angle ϕ is measured, points out radially from the center of the Tevatron in the plane of the Tevatron ring. The polar angle θ is measured from the z-axis, although it is often transformed to a new variable called pseudorapidity (η), defined in Equation 2.3.

$$\eta = -\ln(\tan(\frac{\theta}{2})) \tag{2.3}$$

The new variable η is convenient at a hadron collider because it is invariant under Lorentz boosts along the z-direction. This is useful because the momentum along the

z-direction of the two initial colliding partons will be different in each event, depending on what fraction of the proton (or antiproton) momentum each parton carries. For the same reason, it is also useful to define the transverse energy $E_T = E \sin\theta$ and the transverse momentum $P_T = P \sin\theta$.

2.2.2 The Silicon Tracking Detectors

The innermost detector system at CDF consists of the silicon tracking detectors. This detector system is critical to many analyses at CDF because it is able to measure track impact parameters with a resolution of approximately 30 microns. As will be discussed in Chapter 3, this allows for the tagging of jets originating from bottom quarks, providing a very useful way to differentiate signal events from background events in this analysis and many others.

The silicon detectors are made up of many silicon semiconductor p-n junctions. When a charged particle passes through the detector, it will cause ionization in the depletion regions of the semiconductors, allowing the particle to be detected and tracked. There are three separate detectors that make up the silicon tracking system: Layer 00 (L00), the Silicon Vertex Detector (SVX), and the Intermediate Silicon Layer (ISL). Together, these three systems cover a cylindrical area around the beam pipe with a radial range from 1.5 cm to 32 cm, and they can detect tracks from charged particles with $|\eta| \leq 2$.

L00 consists of a single-sided silicon microstrip detector and is attached directly to the beam pipe. It is surrounded by the SVX, which is made up of five layers of silicon microstrip ladders which are double-sided (each strip contains p-n junctions on both sides). Outside of the SVX is the ISL, which again consists of double-sided silicon ladders. There is one ISL layer that covers the central region of the detector and two additional ISL layers that provide tracking out to $|\eta| = 2$.

2.2.3 The Central Outer Tracker

The entire silicon tracking system is located inside CDF's other tracking detector, the central outer tracker (COT). The COT is an open cell drift chamber made up of eight concentric superlayers of tracking cells which use an argon-ethane mixture for the ionizing gas. Each cell spans the entire length of the COT and contains alternating potential wires, held at a voltage of 2 kV, and sense wires, held at a voltage of 3 kV.

When a charged particle enters the COT, it causes the argon-ethane mixture to become ionized. The electrons released during this ionization drift towards the sense wires, causing secondary ionization and releasing additional electrons. This charge is deposited on a sense wire, and the charge and its arrival time are recorded by a charge integrating amplifier/ADC and a fast time-to-digital converter. By knowing the electron drift velocity in the gas mixture, the location of the original ionization can be determined, and a track can be reconstructed. In this way, the COT is able to measure tracks for charged particles with $|\eta| \leq 1$.

Both the silicon detectors and the COT are enclosed inside a large superconducting solenoid magnet which creates a field of 1.4 T. This field causes the paths of charged particles to curve in the azimuthal plane. The amount of curvature depends on the transverse momentum of a particle, so the curvature can be used to measure this momentum. The silicon and COT combined can measure P_T with a resolution of approximately $\frac{\delta P_T}{P_T} = 0.0012 \cdot P_T$ [19].

2.2.4 The Electromagnetic and Hadronic Calorimeters

The electromagnetic (EM) calorimeters are located directly outside of the magnet and are used to measure energy deposition by particles that interact mainly via the electromagnetic interaction (photons and electrons). The central electromagnetic calorimeter (CEM) covers the range $|\eta| \leq 1.1$, while the plug electromagnetic calorimeter (PEM), covers $1.3 \leq |\eta| \leq 3.6$. Both electromagnetic calorimeter sys-

tems are made up of alternating layers of lead and plastic scintillator. When an electron enters the calorimeter, it will radiate photons due to deceleration caused by the electromagnetic fields of atomic nuclei (bremsstrahlung radiation). Each radiated photon can then convert into an electron-positron pair, which in turn causes more bremsstrahlung, and the shower continues to grow until the resulting particles do not have enough energy to emit any radiation.

The total energy in the shower (equivalent to the energy of the initial particle entering the calorimeter) is proportional to the amount of light produced in the shower, which is measured by the scintillator layers. The light is collected by the scintillators and transmitted to a photomultiplier tube, where the total amount of light produced is measured. In this way, the electromagnetic calorimeters are able to determine the energy of an incident photon or electron with a resolution of approximately $\frac{\delta E_T}{E_T} = 0.14/\sqrt{E_T}$ [19].

Several layers into the CEM, at the depth where the electromagnetic shower is expected to be at a maximum, there is a layer of wire chambers called the central electron strip (CES) detector. Similarly, there is a PES detector located inside the PEM. These detectors are used to localize the position of an electromagnetic shower inside the calorimeter, allowing better matching of showers to tracks observed by the COT.

Charged hadrons that pass through the electromagnetic calorimeter lose a small portion of their energy, but they do not cause a shower because they do not radiate photons. In order to measure the energy of hadrons, the hadronic calorimeters (HAD) are located just outside of the electromagnetic calorimeters. The HAD calorimeters cover the same range as the EM calorimeters, and are similar in design - the HAD calorimeters are made up of alternating layers of iron and scintillator. Hadrons entering the HAD calorimeters participate in nuclear interactions within the iron, creating a shower of new particles which can then undergo nuclear interactions

themselves. The showering hadrons then enter the scintillator layers and produce light, which is used to measure the energy in the shower. Because the shower in the HAD calorimeters depends on nuclear interactions, which occur less frequently than the bremsstrahlung radiation in the EM calorimeter, there are generally fewer particles in a shower in the HAD calorimeters, and the energy resolution is poorer - approximately $\frac{\delta E_T}{E_T} = 0.50/\sqrt{E_T}$ in the central HAD calorimeter [19].

2.2.5 The Muon Chambers

The final layer of the CDF detector, located beyond the HAD calorimeters, consists of the muon chambers. Muons interact via the electromagnetic force, but because they are 200 times more massive than the electron, they do not readily produce bremsstrahlung radiation. They therefore pass through the calorimeters without losing energy or creating a shower. Large steel slabs surround the hadronic calorimeters, designed to stop all particles except muons from passing through. Wire chamber drift cells are placed just outside of these slabs to detect muons that travel beyond the steel slabs. Similar to the COT, when a muon passes through a chamber, it ionizes the gas (which is argon-ethane, as is used in the COT). The ionization electrons then drift to high voltage wires in each chamber, and the detector records the passage of a muon.

The CDF detector contains several sets of muon chambers, each of which consists of four layers of drift cells with a slight azimuthal offset. Hits in these muon chambers define a track segment or "stub" which can then be matched to a reconstructed track in the COT. The central muon detector (CMU) and the central muon upgrade (CMP) cover the range $|\eta| \leq 0.6$. The central muon extension (CMX), located at the corners of the detector, covers the range $0.6 \leq |\eta| \leq 1.0$. Finally, the barrel muon upgrade (BMU) covers the region with $1.0 \leq |\eta| \leq 1.5$.

2.3 Data Acquisition at CDF

Collisions occur at the Tevatron every 396 nanoseconds. With 750,000 channels in the detector to be read out, it is impossible to record every collision on this timescale, so a "trigger" system must be used to decide which events to record. CDF uses a three stage trigger system that was designed to look at the 2 million collisions that occur each second and select only the approximately 75 most interesting ones to be recorded. This trigger system is shown schematically in Figure 2.4. It was designed so that the data acquisition (DAQ) system could check at least 95% of the collisions to determine whether or not to keep the event.

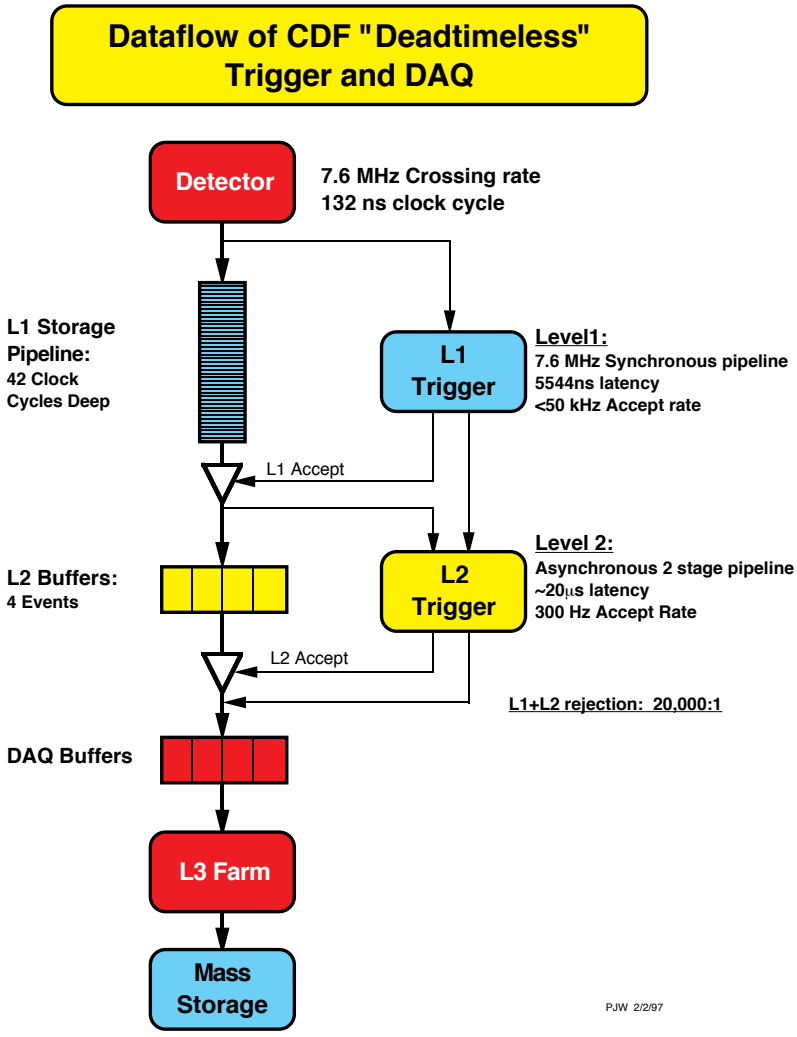

Figure 2.4. Schematic of the Data Acquisition System

The first trigger stage, the Level 1 selection, stores information for 42 consecutive bunch crossings in a hardware buffer. Each event is kept for 5500 ns, allowing the buffers to collect signals from all of the detector systems, including the COT and muon chambers, which have long drift times. The Level 1 hardware fits helical tracks to hits in the COT, matches extrapolated COT tracks to hits in the muon chambers, and finds the energy deposited in the calorimeter. The results of these operations are compared with the Level 1 selection criteria, as listed in a "trigger table", and a decision is made as to whether to pass the event on to the next level or reject it. The Level 1 system is designed to process each event in less than 4000 ns so that no event is lost from the buffer before a decision is made.

An event that passes the Level 1 selection is passed on to one of four Level 2 buffers. At this point, the silicon tracking system is read out and tracks are reconstructed to determine if there are any jets which could potentially be tagged as bottom quark jets. Information from the electromagnetic and hadronic calorimeters is used to identify possible electrons, photons, and jets. Processors then combine the information from various sources and make a more sophisticated decision than the Level 1 system by comparing each event to a set of criteria determined by the trigger table.

After an event is selected by the Level 2 system, the entire detector is read out and this information is sent to one of several hundred processors that make up the Level 3 system. This is the first time that the event is fully reconstructed using all the detector channels. The Level 3 processors can take up to a second to reconstruct an event and then make a final decision on whether or not to record that event, based on the types of particles present in the event and certain other properties of the event as a whole.

Once an event has been accepted by the Level 3 trigger system, it is written to a magnetic tape at Fermilab's computing center. From there, offline processing is performed to apply calibrations for variables that can change over time, such as

beam alignment, and a detailed validation of the data is performed. The data can then be used for analysis.

CHAPTER 3

Event Selection

Collisions at the Tevatron can result in many different physical interactions taking place, producing many possible final states, but this analysis is concerned with only those events that have a final state that could have been the result of the decay of a pair of top quarks. The event selection must be designed to separate these signal events from the background - the many other processes that occur at the Tevatron. In particular, we consider events where the final state has a signature consistent with a $t\bar{t}$ system decaying via the lepton plus jets decay channel, where one of the top quarks in the $t\bar{t}$ system decays leptonically via $t \rightarrow Wb \rightarrow \ell\nu b$ and the other top quark decays hadronically via $t \rightarrow Wb \rightarrow udb$. One reason for using this channel is that the detector is very good at identifying electrons and muons, which are rarely produced at hadron colliders. Using a final state containing a lepton helps greatly in reducing the number of background events.

The event selection that we use is the standard selection that has been developed at CDF for $t\bar{t}$ events in the lepton plus jets decay channel. We choose lepton plus jets events using a selection requiring one high energy charged lepton (from the leptonically decaying W boson), four high energy jets (clusters of particles resulting from the hadronization of the two bottom quarks and the two quarks from the hadronically decaying W boson), and a large amount of missing energy (from the neutrino, which passes through the detector without depositing any energy). We additionally require the charged lepton to be either an electron or a muon because of the difficulty of

correctly identifying tau leptons in the detector.

There are several steps in the event selection, the first of which is performed by the online trigger system as events are selected to be read out and written to tape. This online selection is primarily driven by the detection of a high momentum electron in the CEM calorimeter or a high momentum muon in the CMUP or CMX muon chambers. This method allows us to take advantage of the rarity of leptons at a hadron collider and immediately reduce the background. Once an event has been selected by the online trigger system, it is written to disk and the data is then processed offline. At this point, a number of calibration corrections are made, allowing tracks and energy showers to be reconstructed using the information taken directly from the detector. This reconstruction allows us to identify objects like electrons, muons, or jets. By looking at these objects, we can make a final determination as to whether an event should be selected as a $t\bar{t}$ candidate. The details of this selection are included below and described in [21].

3.1 Electron Identification

The detection of a high momentum electron is one of the primary signals of a possible $t\bar{t}$ event for the online trigger. The basic signal for an electron is a single track in our tracking chambers pointing to an energy deposit in the CEM calorimeter, with minimal energy passing through to the HAD calorimeter. Below we summarize the requirements for electron selection, both for the trigger system and the offline identification.

3.1.1 Trigger Requirements

The Level 1 trigger selection demands that several basic requirements for electron identification are met. At Levels 2 and 3, these requirements are made stricter, increasing the likelihood that the object passing the trigger requirements is a true

high momentum electron.

Level 1

At Level 1, only very basic trigger requirements indicative of an electron are imposed. We require that there be a cluster of energy in the CEM calorimeter ($|\eta| < 1.0$) with $E_T \geq 8.0$ GeV. We also demand that the ratio of energy deposited in the hadronic calorimeter to energy deposited in the electromagnetic calorimeter by this cluster be small [$E_{HAD}/E_{EM} \leq 0.125$], indicating an electromagnetic shower rather than a hadronic one. We finally require at least one COT track with $P_T \geq 8$ GeV/c.

Level 2

At Level 2, the trigger requirements are strengthened slightly. There must be a cluster of energy in the CEM calorimeter with $E_T \geq 16.0$ GeV and $E_{HAD}/E_{EM} \leq 0.125$, and we require that one of the COT tracks with $P_T \geq 8$ GeV/c found at Level 1 points to this cluster.

Level 3

The trigger requirements are again made more stringent at Level 3, where we require a cluster of energy in the CEM calorimeter with $E_T \geq 18.0$ GeV and $E_{HAD}/E_{EM} \leq 0.125$. Additionally, we again check that the matching track found at Level 2 has $P_T \geq 8$ GeV/c. Also, the lateral shower profile L_{shr} is calculated, comparing the sharing of the shower energy across neighboring calorimeter towers to what is expected according to test beam data. There is a cut on the χ^2 of the comparison requiring $L_{shr} < 0.4$.

3.1.2 Offline Selection

Once an event has passed the trigger level cuts, it is recorded to tape. However, because of the speed with which the trigger decision must be made, there is no guarantee that the object which passed the online electron identification is truly an

electron. Once an event has been recorded and reconstructed offline, much more information is available than in the online selection, so additional offline cuts are implemented which are generally stronger and more detailed. The criteria for selecting electrons offline are listed below, and any electrons that pass these selection cuts are referred to as "tight" electrons.

- Because of the large mass of the top quark compared to the electron, electrons from a lepton plus jets decay of a $t\bar{t}$ pair will be energetic. Thus, we require $E_T \geq 20$ GeV.
- Conservation of energy requires $P_T = E_T$ for an ideal electron, but due to energy losses and possible mis-measurements, this equality is not always realized in the detector. Our offline selection first requires $P_T \geq 10$ GeV/c. Then, if $E_T < 100$ GeV we also require $E/P < 2.0$, while if $E_T > 100$ GeV we require $P_T \geq 50$ GeV/c.
- The majority of the energy from an electron shower is contained in the electromagnetic calorimeter, so we can reduce the number of jets misidentified as electrons by requiring that the ratio of the energy measured in the HAD calorimeter to the energy measured in the CEM calorimeter be small: $E_{HAD}/E_{EM} < 0.055 + 0.00045 \cdot E$. We must include a scale factor that is linear in the electron energy because very high energy electrons are more likely to deposit some energy in the hadronic calorimeter.
- We require the electron track found in the COT to have hits in at least 3 COT axial segments and at least 2 COT stereo segments as a quality cut on the track.
- We want to make sure that the electron track originates from a region with good COT coverage, so we require that z_0, the z-intercept of the electron track, is within the range $|z_0| < 60$ cm.
- We require Δx, the distance in the x-coordinate between the COT track position extrapolated to the CES and the actual hits in the CES chamber, to fall in

the range $-3.0cm < Q \cdot \Delta x < 1.5$ cm. This cut depends on the electron charge because the algorithm for matching a COT track to the CES is charge dependent.

- We require Δz, the distance in the z-coordinate between the COT track position extrapolated to the CES and the actual hits in the CES chamber, to satisfy $|\Delta z| < 3.0$ cm.
- Based on test beam data and simulation, we are able to make a prediction for the profile of hits in the CES for a real electron. We perform a χ^2 test comparing the actual profile of hits in the CES to the prediction, and we require $\chi^2 < 10$.
- We enforce a tighter requirement on the lateral shower profile than the one imposed by the online selection, demanding $L_{shr} < 0.2$.
- Generally the energy that an electron deposits in the calorimeter will be very collimated because of the large electron energy. To take advantage of this fact, we examine the energy in the region surrounding the electron and impose an "isolation" requirement. Jets that happen to pass the other electron selection requirements are filtered out by requiring that energy deposited in a cone around the electron is small compared to the electron's energy. In order to impose this cut, we look at the energy in a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$ around the electron and define the variable I_l, which we demand to be less than 0.1.

$$I_l = \frac{\sum E_T^{\Delta R<0.4} - E_T^{electron}}{E_T^{electron}} < 0.1 \tag{3.1}$$

- When high energy photons pass through the detector, they can interact with the detector and convert into electron-positron pairs. The electrons from these conversions are real electrons and might pass the other selection cuts, but they did not originate from a $t\bar{t}$ event. These conversions appear as oppositely charged tracks traveling in the same direction and having a total invariant mass near zero (the photon mass). If an electron is determined to have such a conversion

partner, it is rejected.

3.2 Muon Identification

A high momentum muon is another one of the primary signals of a possible $t\bar{t}$ lepton plus jets event for the online trigger. Muon identification takes place by linking a high momentum track in the COT tracking chamber to a stub in one of the muon chambers, with minimal energy deposited in the hadronic and electromagnetic calorimeters. There are two different sets of muon chambers used to identify muons, the CMUP and the CMX muon chambers, and there are slightly different online trigger requirements for the two sets of detectors.

3.2.1 CMUP Trigger Requirements

Level 1

As with electrons, only basic trigger requirements consistent with muons are required at Level 1. We require a muon stub in the CMU detector with $P_T \geq 6.0$ GeV/c, as well as a track with $P_T \geq 4$ GeV/c in the COT which can be extrapolated to the muon stub. Additionally, we require a muon stub to be present somewhere in the CMP detector.

Level 2

The only additional requirement imposed at Level 2 for the identification of a CMUP muon is that the COT track found at Level 1 must have $P_T \geq 8$ GeV/c.

Level 3

At Level 3, we require that the track in the COT found in Levels 1 and 2 must be able to be extrapolated to stubs in both the CMU and CMP muon detectors. Additionally, we require that this track must have $P_T \geq 18$ GeV/c.

3.2.2 CMX Trigger Requirements

Level 1

For a muon detected by the CMX muon chambers, at Level 1 we require a muon stub in the CMX detector with $P_T \geq 6.0$ GeV/c. Also, there must be a COT track with $P_T \geq 8$ GeV/c and hits in at least four of the COT superlayers.

Level 2

There are no additional trigger requirements imposed at Level 2 for muons detected by the CMX detector.

Level 3

At Level 3, we require that the COT track found at Level 1 must be able to be extrapolated to the CMX stub. Additionally, this COT track is required to have $P_T \geq 18$ GeV/c.

3.2.3 Offline Selection

As with electrons, stronger cuts can be made offline than those made at the trigger level because of the additional information available. This helps to reduce backgrounds and extract the $t\bar{t}$ lepton plus jets signal. Muons that pass the offline selection criteria are referred to as "tight" muons. The offline selection criteria for muons are itemized below.

- Like electrons, muons in a top quark decay will be very energetic due to the large top mass, so we require $P_T \geq 20$ GeV/c.
- Muons are minimum ionizing particles, rarely interacting with other materials, so they are not expected to deposit significant energy in the calorimeters. In the electromagnetic calorimeter, we demand the energy deposited E_{EM} to be less than the maximum of either 2.0 GeV or $2.0 + 0.0115 \cdot (P - 100.0)$ GeV. Because the hadronic calorimeter contains more material and muons will there-

fore deposit more energy there, we require E_{HAD} to be less than the maximum of either 6.0 GeV or $6.0 + 0.0115 \cdot (P - 100.0)$ GeV. There is a linear correction term in both cases because the energy deposited in the calorimeters increases with the muon momentum.

- For an event with CMU hits, we demand that the extrapolated COT track must be within $\Delta x_{CMU} < 3.0$ cm of the hits in the CMU chamber.
- For an event with CMP hits, we require the extrapolated COT track to be within $\Delta x_{CMP} < 5.0$ cm of the hits in the CMP chamber.
- For an event with CMX hits, the extrapolated COT track is required to be within $\Delta x_{CMX} < 6.0$ cm of the hits in the CMX chamber.
- The distance of closest approach of the muon track to the z-axis is called the impact parameter, d_0. If the muon track has hits in the silicon detector, we require $|d_0| < 0.02$ cm. If there are not any hits in the silicon detector, we require $|d_0| < 0.2$ cm. This cut rejects cosmic rays as well as tracks that are mis-reconstructed such that they do not appear to originate from the primary vertex.
- We again want to ensure that the muon track originates from a region with good COT coverage, so we require that z_0, the z intercept of the muon track, falls in the range $|z_0| < 60$ cm.
- We require hits in at least 3 COT axial segments and at least 2 COT stereo segments as a quality cut on the muon track found in the COT.
- For muons detected in the CMX chambers, we require that the COT exit radius be less than 140 cm. This ensures that the track left by the muon in the COT had passed through enough of the COT superlayers to be considered a good track.
- As was done for the electron case, we impose an isolation requirement that examines the energy in the region close to the muon. Jets that contain "punchthrough"

(energetic particles that escape the hadronic calorimeter and are detected by the muon chambers) or muons created by the decays of other particles are rejected by requiring that energy deposited in the calorimeters in a cone around the muon is small. We require that I_l, the ratio of the energy in a cone around the muon to the muon momentum, be less than 0.1 .

$$I_l = \frac{\sum E_T^{\Delta R<0.4}}{P_T^{muon}} < 0.1 \tag{3.2}$$

- The CDF detector is capable of detecting cosmic ray muons, which are characterized by two back-to-back tracks separated by an angle very close to 180 degrees. When timing information from the detector is used in the reconstruction, it looks as if one of these tracks appears to be traveling towards the center of the detector, rather than away from it. These characteristics of cosmic rays are used to reject muons from such events.

3.3 Jet Identification and Correction

When partons are produced in a collision event, they are not observed as individual particles in the detector. Instead, the QCD interaction creates a shower of many particles in a process called hadronization. This particle shower is called a jet, and it is these jets which are actually observed by the detector, rather than the original partons. A jet is identified by a large number of tracks in the tracking chambers and a large deposit of energy in the electromagnetic and hadronic calorimeter systems. To determine the total energy present in a jet and its direction, a "clustering" algorithm is used to combine the energies deposited in neighboring calorimeter towers into a single calorimeter cluster [22].

Clustering occurs by searching for a tower with deposited E_T of at least 3 GeV and then calculating the total energy deposited in all towers within $\Delta R < 0.4$ of this original tower. A new cluster center is calculated as the weighted centroid of the

energy in the original cluster, and the clustering is repeated. This process continues until the cluster center converges to a constant point, at which point the final jet energy is determined.

Because there are known and well-understood issues related to the response of different parts of the calorimeter, as well as the fact that there can be other sources for the energy deposited in the calorimeters, several corrections (summarized below) must be made.

- Level 1 (η Dependence): This correction is applied to the raw energy deposited in the calorimeter in order to account for the fact that the calorimeter response is non-uniform as a function of η.
- Level 2: This correction is no longer in use.
- Level 3: This correction is no longer in use.
- Level 4 (Multiple Interactions): At a $p\bar{p}$ collider, there can be multiple $p\bar{p}$ interactions during the same bunch crossing, and any extra interactions will increase the amount of energy deposited in the calorimeter. This correction subtracts the average extra energy deposited by these additional interactions.
- Level 5 (Absolute): The calorimeter does not provide complete coverage around the interaction region. Instead, there are gaps between calorimeter towers where particles can travel undetected and carry away part of the energy of a jet. There is also some non-linearity in the calorimeter response which depends on the amount of energy deposited. The Level 5 correction accounts for these effects using the total P_T of tracks in the COT that match the cone of the jet. It is at this stage that we expect jets in the data to best match jets in simulated events, and thus jets are corrected to Level 5 in this analysis for the purpose of making selection cuts.
- Level 6 (Underlying Event): Some of the energy deposited in the calorimeter results from interactions between the spectator partons - those partons in the

proton and antiproton that are not directly involved in the collision. In some analyses, the additional energy resulting from these spectator partons needs to be subtracted from the total jet energy.

- Level 7 (Out Of Cone): A cone size of $\Delta R < 0.4$ is used for clustering, but this choice is arbitrary and in some cases the jet will contain additional energy that is outside of this cone. This correction uses simulated events to model jets and correct for any energy that might leak outside of the cone.

In this analysis, all jets are corrected up to Level 5 for the selection and up to Level 7 for all other purposes. We define "tight" jets as those which have a Level 5 corrected energy $E_T > 20.0$ GeV and also have $|\eta| < 2.0$. We require at least four of these tight jets in our event selection.

3.3.1 Secondary Vertex Tagging

Top quarks almost always decay to a W boson and a bottom quark, but many processes that are backgrounds to our selection, such as the production of a W boson plus light flavor jets, do not contain heavy bottom quarks in their final states. Because it can only decay through the weak interaction and the CKM mixing angle is small, the bottom quark has a rather long lifetime compared to other heavy objects - approximately 1.5 ps [7]. This long lifetime means that the bottom quark can travel a fairly long distance before it decays - if a bottom quark has a momentum of $P = 65$ GeV/c (a reasonable value for a quark coming from top decay) it will travel a distance of approximately $\gamma\beta\tau c = 500\mu$m before decaying. This distance is large enough that it can be observed by the silicon detector: there will be some tracks in the jet that form a secondary vertex some distance away from the primary vertex of the jet. CDF's "SecVtx" algorithm is used to identify these secondary vertices, a process referred to as "b-tagging" [23]. Our event selection requires at least one of the four observed jets to be tagged as a bottom jet in order to reduce the number of background events

that pass the selection. The tagging algorithm checks every jet in an event for the presence of secondary vertices created by high-quality tracks. The quality of a given track is determined by looking at the number of silicon hits and the χ^2 of the helical track fit, as well as the momentum of the particle that made the track. If a secondary vertex is found, the length L_{xy} between the primary vertex and the secondary vertex is calculated. A jet is considered to be tagged if $L_{xy}/\sigma_{L_{xy}} > 3.0$, where $\sigma_{L_{xy}}$ is the estimated uncertainty on L_{xy}.

3.4 Missing Energy

In $p\bar{p}$ collisions at the Tevatron, the colliding particles always have zero total momentum in the plane that is transverse to the beamline (although not along the beamline, since the initial partons in the collision do not necessarily have equal and opposite p_z). Conservation of momentum therefore requires that the total transverse momentum of all final state particles must be zero as well. In $t\bar{t}$ events in the lepton plus jets decay channel, the neutrino, which cannot be observed by our detector, has a significant momentum, causing the total transverse momentum of all the particles observed in the detector to be unbalanced. The size of this conservation of momentum "violation" is called the missing transverse energy, $\not\!\!E_T$. For an ideal detector, this missing energy would be exactly equal to the transverse momentum of the neutrino.

To calculate the $\not\!\!E_T$ [21], we start with the negative of the vector sum of the transverse energy observed in all calorimeter towers, with the vectors broken into x- and y-components. The negative sign results from the fact that, in order to conserve momentum, the $\not\!\!E_T$ must point directly opposite the total transverse momentum of all observed particles in the event.

$$\not\!\!E_{T,x}^{raw} = -\sum_{towers} E^{tower} \cdot \cos(\phi_{tower}) \tag{3.3}$$

$$\not\!\!E_{T,y}^{raw} = -\sum_{towers} E^{tower} \cdot \sin(\phi_{tower}) \tag{3.4}$$

Two corrections must be made to this raw value for the $\not\!\!E_T$. Some events contain muons, which are observable but do not deposit energy in the calorimeter, which means they are not accounted for in our raw $\not\!\!E_T$ calculation. For these events, the total energy in the event must be corrected by including the momentum of the muon track. Additionally, we must also include the effect of the jet corrections by replacing the raw jets observed in the detector with jets that are corrected up to Level 5. Applying these two corrections gives new values for the x- and y-components of the missing transverse energy.

$$\not\!\!E_{T,x}^{corr} = \not\!\!E_{T,x}^{raw} - P_T^{muon} \cdot \cos(\phi_{muon}) + \sum_{jets} E_{raw}^{jet} \cdot \cos(\phi_{jet}) - \sum_{jets} E_{L5}^{jet} \cdot \cos(\phi_{jet}) \tag{3.5}$$

$$\not\!\!E_{T,y}^{corr} = \not\!\!E_{T,y}^{raw} - P_T^{muon} \cdot \sin(\phi_{muon}) + \sum_{jets} E_{raw}^{jet} \cdot \sin(\phi_{jet}) - \sum_{jets} E_{L5}^{jet} \cdot \sin(\phi_{jet}) \tag{3.6}$$

The magnitude and direction of the final $\not\!\!E_T$ vector can now be determined from the x- and y-components.

$$\not\!\!E_T = \sqrt{(\not\!\!E_{T,x}^{corr})^2 + (\not\!\!E_{T,y}^{corr})^2} \tag{3.7}$$

$$\phi_{\not\!\!E_T} = \tan^{-1}(\not\!\!E_{T,y}^{corr} / \not\!\!E_{T,x}^{corr}) \tag{3.8}$$

During the offline selection, we impose the requirement $\not\!\!E_T \geq 20.0$ GeV.

3.5 Additional Selection Cuts

At this point, we have identified all of the final state objects in a lepton plus jets $t\bar{t}$ event: a high momentum lepton, four high energy jets from the four final state quarks, and a large $\not\!\!E_T$ from the neutrino. However, there are a few additional cuts to be made that consider the event as a whole, rather than the individual objects, in order to further reduce the number of background events that pass the selection.

3.5.1 Dilepton Veto

We expect $t\bar{t}$ events in the lepton plus jets decay channel to contain exactly one lepton. If a second tight lepton is present in the event, it is either the result of a $t\bar{t}$ event decaying in the dilepton channel or of a different process entirely, so any such event with a second tight lepton fails our selection.

3.5.2 Z Veto

Many possible background processes contain a Z boson which can then decay to two leptons. If both leptons are identified as tight leptons, then these events would be removed by the dilepton veto. However, sometimes the second lepton might not meet the requirements for a tight lepton or could be mistaken for a jet. In order to remove these background processes, events fail our selection if a tight lepton and any one of a few other objects in the event have a total invariant mass that is consistent with a Z boson ($76 < M_{\ell\ell} < 106$ GeV/c^2). For a tight electron, the second object could be an isolated electromagnetic object (electron or photon), a jet with 95% or more of the total energy deposited in the electromagnetic calorimeter, or an isolated single track with opposite charge. If the tight lepton is a muon, then the second object could be another isolated muon or an isolated single track with opposite charge.

3.5.3 Primary Vertex Reconstruction

The primary vertex for an event is the point where the initial collision takes place and the top quark pair is created. The position of the primary vertex along the beamline is used in the jet clustering algorithm, and it also allows us to check whether the lepton and the jets in an event came from the same vertex. The location of the primary vertex along the z-axis is determined by calculating the sum of the z-intercepts of all tracks with a common point of origin, weighted by their uncertainties:

$$z_{Vertex} = \frac{\sum_{tracks} z_0^{track}/\Delta^2_{track}}{\sum 1/\Delta^2_{track}} \tag{3.9}$$

An event is rejected if the z-intercept of the lepton is not within 5 cm of the primary vertex, reducing the number of events where the lepton and the jets are created in separate interactions.

3.6 Event Selection Summary

Below we include a summary of the event selection for $t\bar{t}$ events in the lepton plus jets decay channel. In the dataset used for this analysis, which has an integrated luminosity of 4.3 fb^{-1}, we select 1,001 candidate $t\bar{t}$ events.

- We require one tight lepton (CEM, CMUP, or CMX) as our first selection criteria. The selection of this lepton drives the online trigger requirements for the dataset used in this analysis.
- We require the presence of at least four tight jets, since there are four final state partons when a top quark pair decays via the lepton plus jets channel, each of which should create a jet. We additionally require that at least one of these jets be tagged as a bottom jet.
- Because of the presence of an undetectable neutrino in our final state, we require $\not\!\!E_T \geq 20$ GeV.

- We finally require a dilepton veto, a Z veto, and a primary vertex check in order to eliminate background events which have all of the necessary final state particles but did not originate from a $t\bar{t}$ decay.

CHAPTER 4

Modeling the Expected Signal and Background

Our event selection criteria are optimized for selecting $t\bar{t}$ events in the lepton plus jets decay channel. However, there are several other physical processes with final states that are similar to the $t\bar{t}$ lepton plus jets final state. Of the 1,001 events in our dataset which pass the event selection cuts, we expect some of them to be the result of these other background processes.

In order to determine the spin correlation in real $t\bar{t}$ events, it is important to understand and account for the effect of all the background processes which contribute to the selected events. These processes include: electroweak diboson production, Drell-Yan production (Z+jets), electroweak production of single top quarks, QCD production of fake $t\bar{t}$ events, W-boson plus heavy flavor jets production (W+HF), and W-boson plus light flavor jets production (W+LF). Each of these processes will be discussed in more detail below. Some of these processes are very well understood and can be modeled fully using simulated events, while others are not as well understood theoretically and are modeled using real data collected by the CDF detector [21, 24].

4.1 Event Generation and Simulation

Most of our backgrounds are modeled by simulating events produced by each particular background process. This simulation takes place over several steps. First, the initial conditions of the $p\bar{p}$ collision must be modeled, and the momenta of the colliding particles are fed into an event generator. This event generator performs the

matrix element calculation for the various processes and produces output "events" which fully specify the momenta of the various outgoing particles. The hadronization of any outgoing partons into jets must then be simulated in a process called parton showering. Finally, the results of this simulation are passed through a simulation of the detector response, called CDFSim. The output of this final step is a set of data banks identical to those which would be read out from the detector, as if a given simulated event actually occurred inside the detector. The various event generators used for modeling both our signal and background processes are listed below.

- PYTHIA [25]: PYTHIA is a leading order event generator which also includes parton showering. It is used for $t\bar{t}$ modeling and also for some of our background samples.
- HERWIG [26]: HERWIG is a leading order matrix element event generator which also includes parton showering. It is particularly suited for modeling the emission of low-energy gluons. Because HERWIG includes $t\bar{t}$ spin correlation effects in the event generation, while PYTHIA does not, it is useful in generating the signal samples used to test our measurement method.
- ALPGEN [27]: ALPGEN is a leading order matrix element generator that does not include parton showering. It is especially useful for modeling the production of electroweak vector bosons in association with energetic partons, and is used for many of our simulated samples for processes which include vector bosons. Because it does not include parton showering, PYTHIA is used to simulate the parton shower in events generated by ALPGEN.
- MadEvent [28]: MadEvent is a leading order matrix element generator that includes color flow and spin polarization effects, but does not include parton showering. It is used here to model the background processes involving the production of single top quarks. Again, PYTHIA is used to simulate the parton shower in events generated by MadEvent.

4.2 $t\bar{t}$ Modeling

In this analysis, $t\bar{t}$ events are our signal, so it is important that we have an accurate model of the $t\bar{t}$ production process. We create this model using simulated events. Having models of both the signal and background will allow us to compare simulated signal+background distributions to the data which passes our selection so that we can ensure that our understanding of the data is accurate. We model $t\bar{t}$ production using both PYTHIA and HERWIG, allowing us to perform cross checks in multiple simulated samples. The HERWIG $t\bar{t}$ sample is especially important in this analysis because during generation, the spin correlation effect that we are measuring in this analysis is included, while PYTHIA does not include this effect.

$$N_{p\bar{p}\to X} = \sigma_{p\bar{p}\to X} \cdot \epsilon \cdot \int dt \cdot L \tag{4.1}$$

For $t\bar{t}$ production and several of our background processes, we estimate the number of events that pass our event selection using the cross section for the process, the integrated luminosity of the data sample, and the overall selection efficiency for the chosen process, which is derived from our simulated samples. The expected number of events for a given process can be determined using Equation 4.1, where σ is the cross section (measured to be $\sigma_{t\bar{t}} = 7.50 \pm 0.48$ pb[3]), ϵ is the selection efficiency, and $\int dt \cdot L$ is the integrated luminosity. For an integrated luminosity of 4.3 fb^{-1}, the number of predicted $t\bar{t}$ events is

$$N^{t\bar{t}} = 808.31 \pm 105.33$$

4.3 Electroweak Diboson Production

The electroweak interaction can produce events with pairs of vector bosons - WW, WZ, and ZZ events. Some of these events are able to pass our selection criteria

Process	σ
WW	13.25 ± 0.25 pb
WZ	3.96 ± 0.06 pb
ZZ	1.58 ± 0.02 pb

Table 4.1. Cross Sections for Diboson Production

because they contain a real lepton and neutrino (from the decay of a W boson) as well as multiple jets (from the decay of the second boson). These diboson backgrounds are simulated using the PYTHIA event generator. The number of events expected in our sample is again calculated using Equation 4.1. The cross sections for the various diboson processes are shown in Figure 4.1 [7].

Using this information, we find the total predicted number of diboson events that will pass our selection to be

$$N^{Diboson} = 10.29 \pm 1.01$$

4.4 Drell-Yan Production

The electroweak Drell-Yan production of a Z boson can be accompanied by several jets. In this case, particularly when the boson decays via $Z \rightarrow \tau\tau$, the final state of such an event could pass our event selection, thanks to the presence of a real lepton and neutrino (from a τ lepton decay) as well as multiple jets. We model this Z+jets production using the ALPGEN event generator, with the parton showering performed by PYTHIA. The overall normalization is determined in the same way as for the electroweak diboson production. The contribution from this particular background process is fairly small (the cross section for $Z \rightarrow \tau\tau$ is $\sigma = 13.0 \pm 1.5$ pb [7]), but it is non-negligible. The total predicted number of Drell-Yan events in our sample of candidate $t\bar{t}$ events is

$$N^{Z+Jets} = 7.30 \pm 1.04$$

Process	σ
Single Top T-Channel	1.98 ± 0.08 pb
Single Top S-Channel	0.88 ± 0.05 pb

Table 4.2. Cross Sections for Single Top Quark Production

4.5 Electroweak Single Top Production

This analysis studies top quarks produced in pairs by the QCD interaction, but the electroweak force can produce individual top quarks one at a time. When this occurs, these single top quarks form a background for our analysis. There are two electroweak processes that can produce single top quarks at the Tevatron. The first, called S-channel production, occurs when a virtual W boson decays via $W^* \rightarrow tb$. The second process, T-channel single top production, occurs when a W boson interacts with a bottom quark in the "sea" of quarks inside the proton, producing a top quark. Figures 4.1 and 4.2 depict S-channel and T-channel single top production respectively.

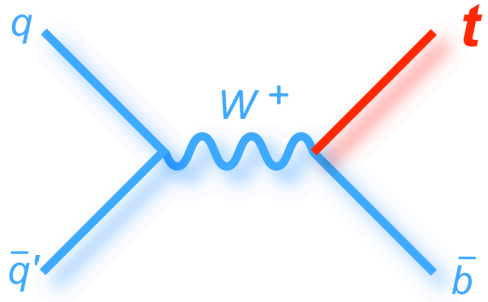

Figure 4.1. S-channel Single Top Production

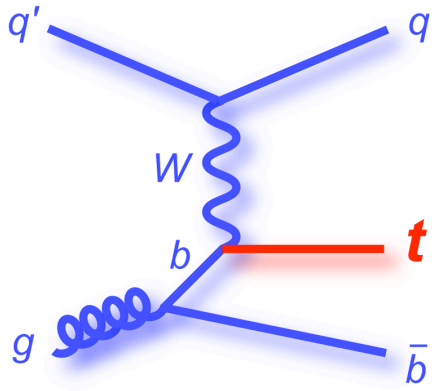

Figure 4.2. T-channel Single Top Production

Because these events contain one real top quark, as well as extra jets, they are often able to pass our event selection. We model single top production using the MadEvent event generator, with the parton showering performed by pythia. The normalization for the single top quark contribution to the total background is found using Equation 4.1, with the cross sections given in Table 4.2 [7]. The cross sections are small, but because these events look very much like $t\bar{t}$ events, they still make a

sizable contribution to our overall background, with the predicted number of single top quark events passing our selection given by

$$N^{SingleTop} = 9.45 \pm 0.64$$

4.6 QCD

The vast majority of collisions at the Tevatron involve only the QCD interaction. Without any electroweak process taking place, such events will not involve any W bosons. Because of this, they will not include any real leptons or neutrinos with large energy, so our event selection removes an overwhelming majority of these events. However, these events can occasionally pass our event selection by containing both a "fake lepton" - a jet that coincidentally meets all of our lepton selection cuts - as well as a significant amount of missing transverse energy, usually due to the mis-measurement of jet energies. The probability that any individual QCD event will meet both these requirements and pass our event selection is very small, but because there are so many QCD events produced at the Tevatron, they form one of the largest backgrounds for the $t\bar{t}$ events we are studying.

Only QCD events with very specific characteristics will be able to produce both a fake lepton and fake missing transverse energy. Because these things require both mis-identification of a lepton and mis-measurement of energy in the detector, it is difficult to verify that our detector simulation properly models these effects. Therefore, rather than using simulated samples to model this background, we use the data itself. To do so, we work with the "jet electron" dataset, which is created using selection criteria similar to what is used for this analysis, except that there is no $\not\!\!E_T$ cut and, instead of tight leptons, we select jet electrons. These jet electrons are defined as jets containing at least four tracks in the COT, but which deposit most of their energy in the electromagnetic calorimeter. The jet electrons meet the requirements for electron

identification, except for the fact that they contain multiple tracks and are thus unlikely to be real electrons. It is assumed that QCD events in the jet electron dataset are similar kinematically to QCD events where there is a jet that actually does get identified as a tight electron.

In order to determine the normalization of the QCD background, a fit is performed to the $\not\!\!E_T$ distribution in the data using the jet electron sample as one template and a W+jets simulated sample as a second template. This determines the expected QCD fraction in the data, and by then applying our $\not\!\!E_T$ cut, we can determine the number of expected QCD events that will pass our selection. This prediction is

$$N^{QCD} = 47.79 \pm 33.32$$

4.7 W + Heavy Flavor

The W + heavy flavor background processes are those which produce a real W boson as well as heavy quarks ($W + b\bar{b}$, $W + c\bar{c}$, and $W + c$). Such events occur fairly often, and they are also able to pass our event selection easily. They contain a real W boson which can decay leptonically to produce a real lepton and missing transverse energy. They also contain multiple jets, and due to the presence of real bottom quarks, our b-tagging requirement does not significantly reduce this background. As a result, these events form the largest portion of our overall background model.

The W + heavy flavor background component is modeled using ALPGEN, with the parton showering performed by PYTHIA. Because it is difficult to determine theoretically the cross section and absolute normalization for the production of W bosons in association with various numbers of jets, we do not estimate the normalization for this background process using Equation 4.1. Instead, we start with the pretag data sample, which is the sample we find when we impose all of our selection cuts except for requiring at least one jet to be tagged as a b jet. Starting from this sample, the

expected number of W + heavy flavor events in our signal sample can be calculated using Equation 4.2.

$$N^{W+HF} = (N_{pretag} - N_{QCD} - N_{Diboson} - N_{Z+Jets} - N_{SingleTop} - N_{t\bar{t}}) \cdot f_{HF} \cdot \epsilon_{tag} \quad (4.2)$$

Starting with the pretag sample, the number of predicted QCD, diboson, Drell-Yan, single top, and $t\bar{t}$ events is subtracted, giving us an estimate of the number of events containing a W boson and jets. Simulated ALPGEN samples are used to calculate the fraction f_{HF} of these events with heavy flavor quarks. The result is then multiplied by the b-tagging efficiency ϵ_{tag}. In this way, while simulated samples are used to find the heavy flavor fraction f_{HF}, the absolute normalization is determined directly from the pretag data. The total number of W + heavy flavor events predicted using this method is

$$N^{W+HF} = 103.71 \pm 33.19$$

4.8 W + Light Flavor

The final background component is made up of processes which produce a real W boson in association with light quarks. As with the W + heavy flavor background, the W + light flavor processes produce events with a real W boson, which can decay into a lepton and neutrino, as well as multiple jets. However, in this case, there are no real bottom jets in the event, so any event that passes the selection cuts will do so only because one of the jets is incorrectly tagged as a bottom jet (a "mistagged" jet).

We again use ALPGEN, with the parton showering performed by PYTHIA, to model the W + light flavor background. In order to determine the appropriate normalization

of this background, we first consider the secondary vertex that is the basis of our b-tagging algorithm. If we let the distance from the primary vertex to the secondary vertex be T, the path of the decaying particle be $\vec{d}$, and the direction of the jet be $\hat{j}$, then for a real secondary vertex, we should have $T = \vec{d} \cdot \hat{j} > 0$ because the decaying particle must be traveling in the same direction as the jet. If we consider a sample of jets with no measurable lifetime (like those in the W + light flavor sample), the distribution of T will be symmetric around 0, but not exactly 0 because of the finite detector resolution. Thus a mistagged jet is equally likely to have $T > 0$ or $T < 0$. Since no real process can produce a secondary vertex with $T < 0$, the number of such vertices in our data can be used to estimate the number of mistagged jets with $T > 0$ in our candidate events.

The rate of secondary vertex tags with $T < 0$ is measured in the data and parameterized by six jet variables: jet E_T, the number of good SVX tracks in the jet, the total E_T for all jets in the event, the jet η, the number of reconstructed primary vertices in the event, and the z-coordinate of the jet's primary vertex [24]. Using these variables, we can determine N^{mis}, the number of mistagged events predicted in the pretag sample. We can then multiply this number by the estimate for the fraction of W + light flavor events in the pretag sample, as in Equation 4.3, to determine the expected number of W + light flavor events to pass the selection.

$$N^{W+LF} = N^{mis} \cdot \frac{N_{pretag} - N_{t\bar{t}} - N_{QCD} - N_{W+HF} - N_{Diboson} - N_{Z+Jets} - N_{SingleTop}}{N_{pretag}} \tag{4.3}$$

The final result of this estimate is that we expect W + light flavor to be our third largest background component, with a predicted number of events given by

$$N^{W+LF} = 36.34 \pm 8.83$$

4.9 Overall Background Model

Our overall background model is determined by summing the predictions for all of the various background processes. The various backgrounds and their respective normalizations are shown in Table 4.3, giving a total background prediction of 215±48 events. If we add in the expected $t\bar{t}$ events, our prediction for the total number of observed events becomes 1023.18 ± 115.70, which is in good agreement with the 1,001 events that pass the selection cuts. The signal to background ratio is approximately 3.7:1.

Process	Events
W + HF Jets	103.71 ± 33.19
Non-W (QCD)	47.79 ± 33.32
Mistags (W+LF)	36.34 ± 8.83
WW/WZ/ZZ	10.29 ± 1.01
Single Top	9.45 ± 0.64
Z+Jets	7.30 ± 1.04
Total Prediction	214.87 ± 47.88

Table 4.3. Summary of Background Predictions

4.10 Validation of Model Prediction

Now that our signal and background models have been simulated and the expected normalizations have been calculated, it is important to check that these models provide accurate representations of our data. In order to perform this validation, we will compare the prediction of our background and signal models to the data in various distributions. In all of these plots, and those in following chapters, the blue histograms represent our background model, the yellow histograms represent our predicted $t\bar{t}$ signal, and the points represent the data itself.

In lepton plus jets $t\bar{t}$ events, we expect a single high energy lepton and four high energy jets, so we first look at these five objects. Figure 4.3 shows the distribution of the transverse momentum of the lepton, while Figures 4.4 through 4.7 show the

transverse energies for the four jets, in decreasing order. In all cases, our model agrees well with the observed data.

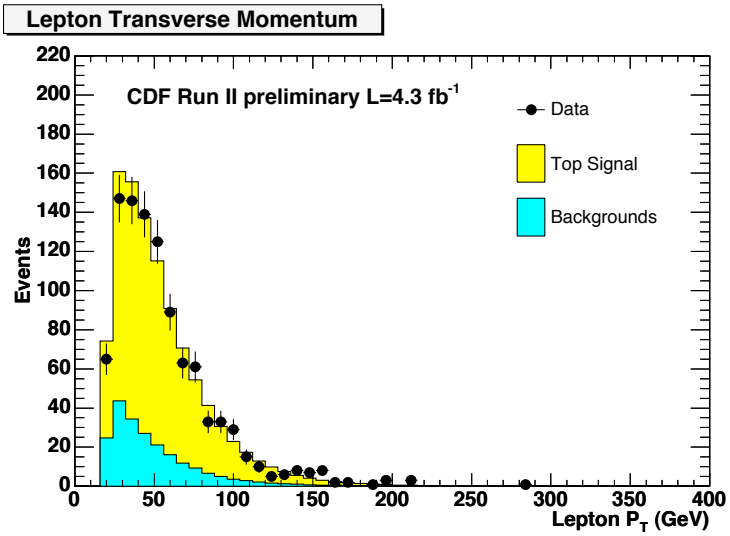

Figure 4.3. Lepton Transverse Momentum

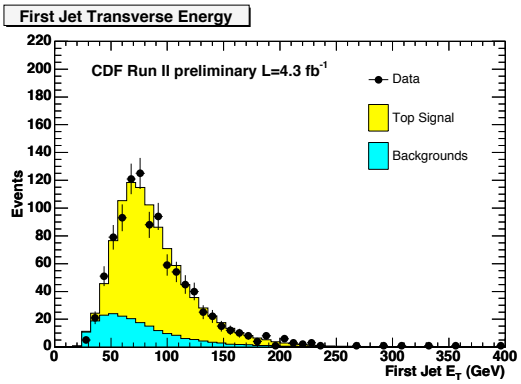

Figure 4.4. Transverse Energy of First Jet

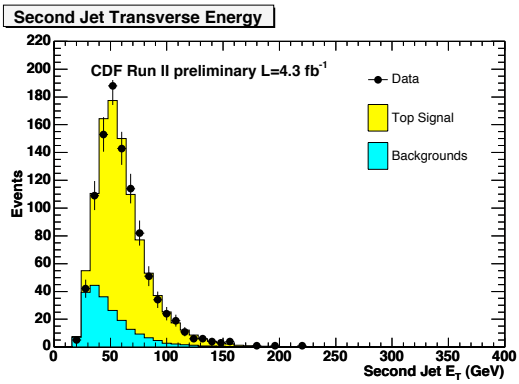

Figure 4.5. Transverse Energy of Second Jet

We also expect lepton plus jets events to have a large missing transverse energy due to the presence of a neutrino. The distribution of this $\not\!\!E_T$ is shown in Figure 4.8. Because $t\bar{t}$ events are expected to be very energetic in general, we also consider the scalar sum of the transverse energy of all objects in the event, called the H_T. The distribution of this H_T is shown in Figure 4.9. For both the $\not\!\!E_T$ and H_T distributions, the data is well modeled by our signal and background predictions.

Finally, we consider the number of jets in each event in Figure 4.10 and the num-

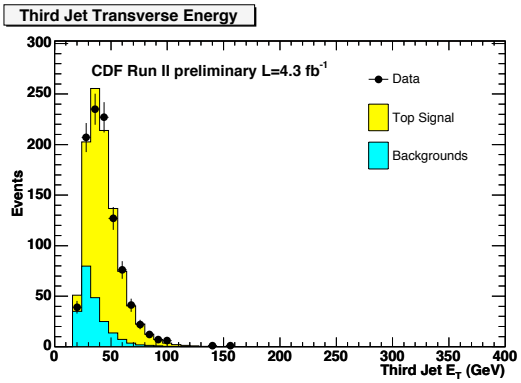

Figure 4.6. Transverse Energy of Third Jet

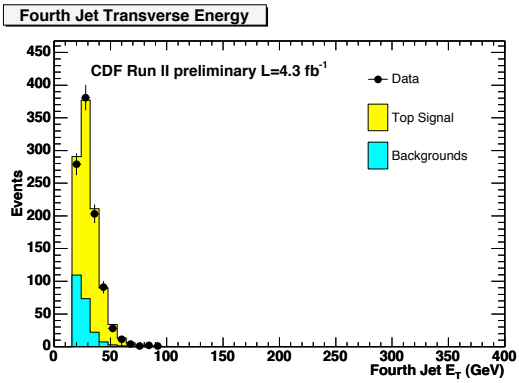

Figure 4.7. Transverse Energy of Fourth Jet

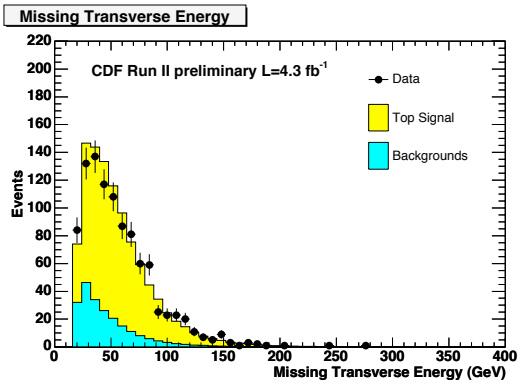

Figure 4.8. Missing Transverse Energy

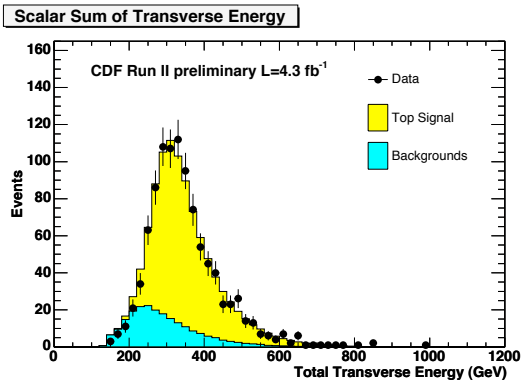

Figure 4.9. Scalar Sum of Transverse Energy

ber of jets that are b-tagged in Figure 4.11. Due to a peculiarity in the weighting algorithm used for b-tagged jets in our simulated samples, the signal and background events observed in the 0 tag bin in Figure 4.11 are actually part of the 1 tag bin for these b-tagged events which pass our selection. In the distributions of both the number of jets and the number of b-tags, the data and our model are in good agreement.

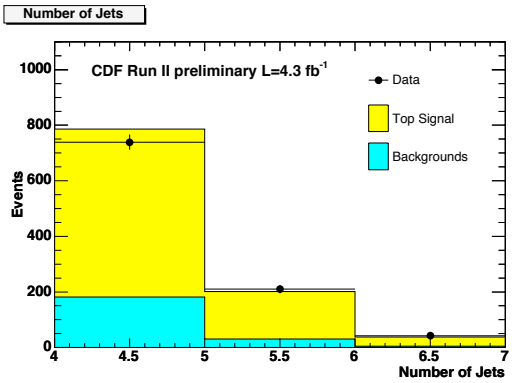

Figure 4.10. Number of Jets

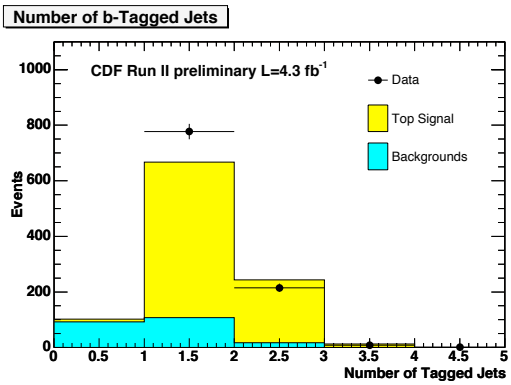

Figure 4.11. Number of b-Tagged Jets

CHAPTER 5

Reconstructing $t\bar{t}$ Events

We have identified $t\bar{t}$ candidate events with a lepton, four jets, and large missing energy, as described in Chapter 3. In this lepton plus jets decay channel, one top, called the leptonic top, decays via $t \to Wb \to \ell\nu b$, and the other, the hadronic top, decays via $t \to Wb \to udb$. Our measurement will require us to use these identified decay products to calculate decay angles, as measured in the top quark and $t\bar{t}$ rest frames. In order to do this, we must be able to fully reconstruct the $t\bar{t}$ kinematics for each event, matching each final state jet with the appropriate top decay product and using simple constraints to determine the t and $\bar{t}$ momentum vectors [29].

5.1 Kinematic Reconstruction

The final state of a $t\bar{t}$ event in our selected decay channel contains a lepton, a neutrino, two bottom quarks, one up-type quark, and one down-type quark. There are several difficulties in trying to match the objects observed in the detector to these final state partons. The lepton and the quark jets can be directly observed in the detector, but for the neutrino, we can measure only the transverse component of its momentum, which is the cause of the $\not\!\!E_T$. The detector can tag jets from bottom quarks, but otherwise it cannot determine which jet is created by a given type of quark, so another problem in reconstructing the event from the available information is determining the correct assignment of jets to partons from all of the possible combinations.

Our reconstruction assumes that the four highest energy jets in the event come

from the four final state quarks. Ignoring any b-tagging information for now, there are 24 possible combinations for matching the four quarks to the four observed jets. However, interchanging the up and down quarks from the hadronically decaying W boson has no effect on the overall kinematics for the t and $\bar{t}$, so the number of unique combinations is reduced from 24 to 12.

The $\not\!\!E_T$ can be used to determine the components p_x and p_y of the neutrino momentum. Unfortunately, we cannot know the portion of the proton and antiproton momenta carried by the initial colliding partons, so we do not know the total momentum of the event along the beam direction, p_z. This means we cannot directly measure the z-component of the neutrino momentum. It is possible, however, to calculate the neutrino p_z from other information available in the event. Because the lepton and neutrino are the only two decay products of the leptonically decaying W, when their momenta are added, they should form a particle whose mass is consistent with that of the W boson. This equality involves a quadratic term, and thus solving it produces two possible solutions for the neutrino p_z. The kinematic reconstruction considers both possible solutions, and therefore the number of possible combinations for the event kinematics is increased again from 12 to 24.

To determine the correct combination, we evaluate how closely each possible combination matches the expectation that it is the final state of a $t\bar{t}$ decay. We expect such a final state to have four basic characteristics:

- The lepton and neutrino should be the decay products of a W boson, $W \to \ell\nu$
- Two of the jets should be the decay products of another W boson, $W \to jj$
- The lepton, neutrino, and a third jet should be the decay products of a top quark, $t \to \ell\nu j$
- The two jets from the W boson and the fourth jet should also be the decay products of a top quark, $t \to jjj$

The degree to which each of the 24 combinations is compatible with these require-

ments is evaluated using a χ^2 test, with the χ^2 function given in Equation 5.1.

$$\begin{aligned}\chi^2 &= \sum_{i=\ell,jets} \frac{(p_T^{i,meas} - p_T^{i,fit})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,meas} - p_j^{UE,fit})^2}{\sigma_j^2} \\ &+ \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{\ell\nu} - M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj} - M_{fit})^2}{\Gamma_t^2} + \frac{(M_{b\ell\nu} - M_{fit})^2}{\Gamma_t^2}\end{aligned} \tag{5.1}$$

The first term in Equation 5.1 is a sum over the lepton and jet transverse energies, while the second term is a sum over the "unclustered" energy, which consists of all the measured energy in the event that is not part of the actual $t\bar{t}$ decay. When evaluating the χ^2 function, small corrections to the energies of each object are made by varying them within the experimental resolution. This improves the energy resolution and increases the probability of finding the correctly matched combination.

The final four terms in Equation 5.1 are the constraints imposed by the assumption that the event is a $t\bar{t}$ event. M_{jj} is the mass of the two jets that we expect to be consistent with the W boson mass, $M_W = 80$ GeV/c^2. The mass of the lepton and neutrino, $M_{\ell\nu}$, also needs to be consistent with M_W. $M_{b\ell\nu}$ and M_{bjj} are the invariant masses of the decay particles from the leptonically and hadronically decaying top quarks. The kinematic reconstruction can be used to measure a top quark mass M_{fit}, but instead we include the constraint $M_{fit} = 172.5$ GeV/c^2 to help improve the reconstruction. The denominators in these constraints are the W boson and top quark decay widths, but do not include a factor for the experimental resolution of the lepton and jets because this resolution has already been accounted for in the first term of Equation 5.1.

The package MINUIT [30] is used to minimize the χ^2 function for each of the 24 possible combinations. At this point we use the b-tagging information in each event to require that any b-tagged jets also be matched to one of the original bottom quarks. The combination that meets this requirement and has the smallest χ^2 value

is then chosen as the best candidate for describing the true $t\bar{t}$ final state.

5.2 Down Quark Identification

The kinematic reconstruction described in the previous section is used to match the objects observed in the detector to the various $t\bar{t}$ decay products and to reconstruct the original $t\bar{t}$ system. However, up to this point in the reconstruction, no effort has been made to individually identify the down-type and up-type quarks from the hadronically decaying W boson. Because the down quark is strongly correlated with the spin of the original top quark, it is essential for this analysis to know which quark jet corresponds to the down quark. Information coming directly from the detector itself cannot be used to distinguish between a down quark and an up quark, so a different method for down quark identification must be found.

We use a probabilistic method outlined in [10] which relies on conservation of angular momentum. Due to the V-A electroweak couplings in the top quark decay chain, angular momentum conservation tends to force the down quark to travel close to the bottom quark from the hadronically decaying top quark. This is shown schematically in one-dimension for the decay $t \rightarrow W^+b \rightarrow udb$ in Figure 5.1. The top quark decays to a bottom quark and a positive W boson, both of which must be left-handed (their spin must be opposite their direction of motion) because of the V-A coupling. When the positive W boson decays by the electroweak interaction into an up quark (which must be left-handed) and an antidown quark (which must be right-handed), conservation of angular momentum forces the down quark to travel in the same direction as the bottom quark.

Once both jets from the hadronically decaying W boson are identified, our algorithm chooses the jet closest to the bottom quark jet, as measured in the rest frame of the W boson, to be identified as the down quark jet. Using the true kinematic information for the top decay products, this algorithm correctly identifies the down

quark at a rate of approximately 60%. Figures 5.2 and 5.3 illustrate the effectiveness of this algorithm. Figure 5.2 shows the difference $\alpha_{db} - \alpha_{ub}$, where α_{db} and α_{ub} are the angle between the bottom quark and down quark momenta and the angle between the bottom quark and up quark momenta respectively. In this plot, 61% of the entries have negative values. Figure 5.3 shows α_{db} and α_{ub} individually. The red histogram, coinciding with α_{db}, is clearly shifted towards smaller angles.

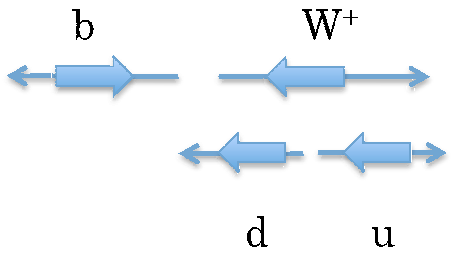

Figure 5.1. Schematic depiction of angular momentum conservation in top decay

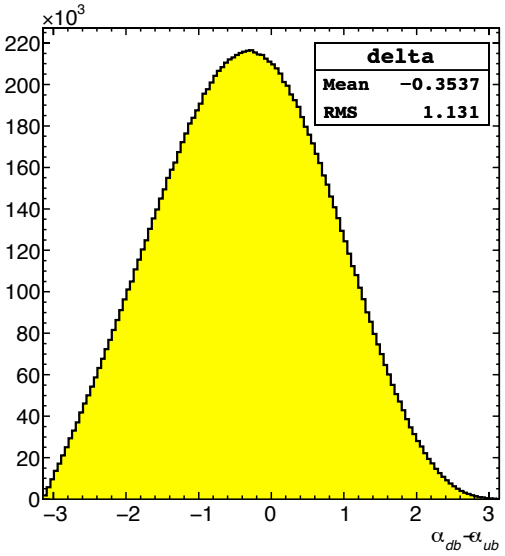

Figure 5.2. Difference between bottom-down angle and bottom-up angle

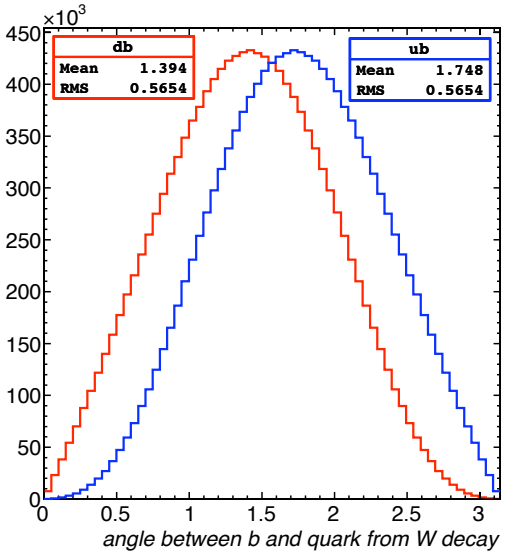

Figure 5.3. Angle between bottom-down and bottom-up quarks in W rest frame

Detector effects, misidentification of other jets in the event, and the presence of jets that did not originate from the $t\bar{t}$ pair can dilute the effectiveness of our down quark identification. In a simulated $t\bar{t}$ sample created using the event generator PYTHIA

and then subjected to the full detector simulation and kinematic reconstruction, the effects mentioned result in the down quark being correctly identified in 76,235 out of 183,091 events (41.6%). This identification is not perfect, but the strong correlation of the down quark with the spin of the original top quark makes up for this identification inefficiency, as will be seen in Chapter 7.

5.3 Reconstruction Validation

It is important to verify that the event reconstruction is performed adequately and that the results are modeled well. The plots in this section compare the data to our signal+background prediction and show that there is reasonable agreement in all cases. Because this measurement will be based on the kinematics of the lepton, the down-type quark, and the bottom quark from the hadronically decaying top quark, we focus on distributions of reconstructed variables involving these three decay products.

We first consider the transverse energy and the pseudorapidity for the down quark, in order to check our down quark identification. These distributions are shown in Figures 5.4 and 5.5 for the jet identified as the down quark, and it can be seen that the model provides a good description of the data.

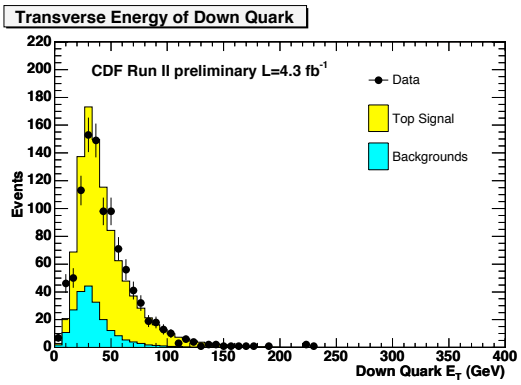

Figure 5.4. Down Quark Transverse Energy

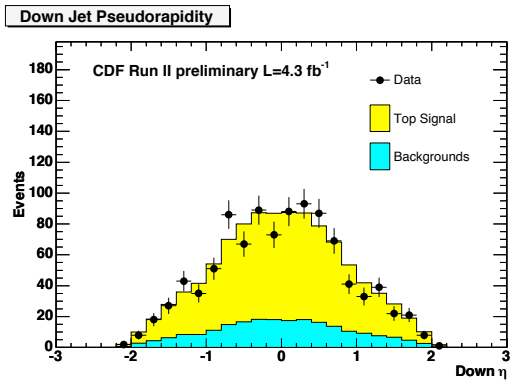

Figure 5.5. Down Quark Pseudorapidity

Because the measurement depends on the full reconstruction of the top quarks, we also show, as an additional check, the transverse momentum of the reconstructed

top quarks in Figure 5.6 and the top quark pseudorapidity in Figure 5.7. Again there is good agreement between the data and our prediction.

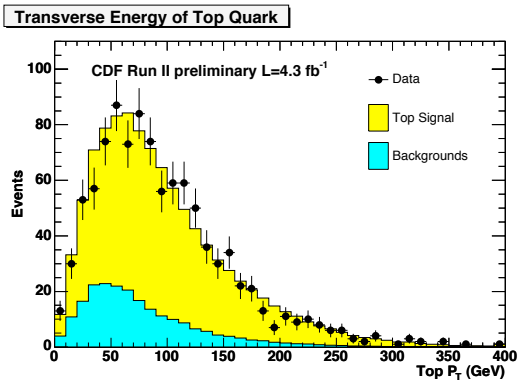

Figure 5.6. Top Quark Transverse Momentum

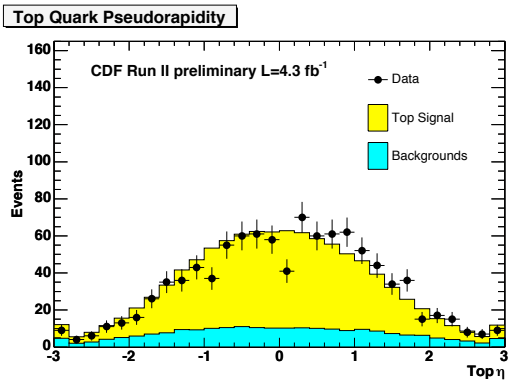

Figure 5.7. Top Quark Pseudorapidity

We also take a look at the momentum of the fully reconstructed $t\bar{t}$ system. Figure 5.8 shows the transverse momentum of the $t\bar{t}$ system, while Figure 5.9 shows the $t\bar{t}$ momentum along the z-direction, with both distributions being well modeled by our prediction.

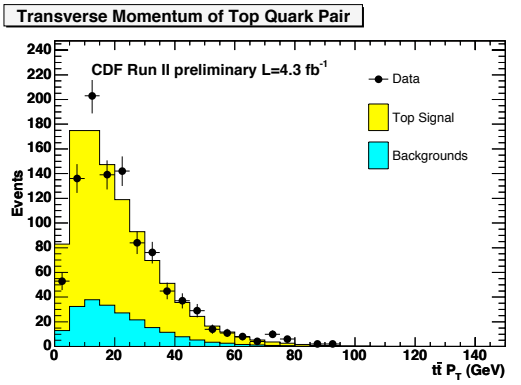

Figure 5.8. Transverse Momentum of $t\bar{t}$ System

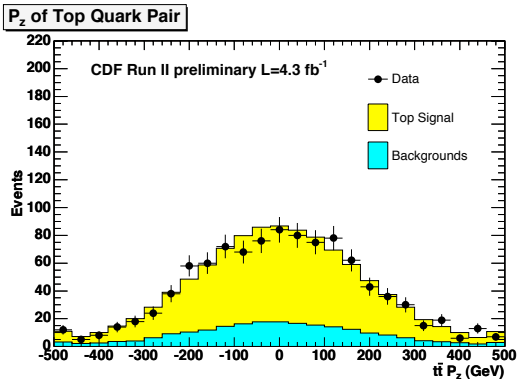

Figure 5.9. p_z of $t\bar{t}$ System

We are especially interested in the angular correlations between the decay products, so we next consider the variable ΔR, the angular separation between two objects in η-ϕ space as defined in Equation 5.2.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{5.2}$$

In particular, we look at ΔR between the lepton and the down quark in Figure 5.10 and ΔR between the lepton and the bottom quark in Figure 5.11, and we can see that the data is well modeled by our signal and background predictions.

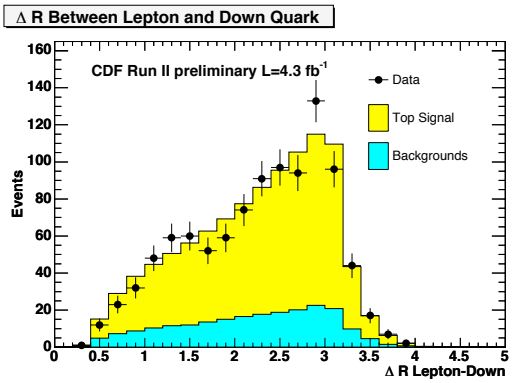

Figure 5.10. ΔR between the lepton and the down quark

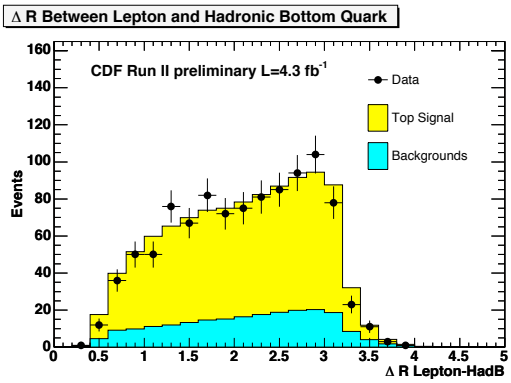

Figure 5.11. ΔR between the lepton and the bottom quark

We finally consider the helicity angle θ_i of the various decay products, defined to be the angle between the momentum of a decay product in the top rest frame and the direction of the top momentum in the $t\bar{t}$ rest frame. Our final result will be based on the product of the cosines of these helicity angles, but here we consider the cosines of the helicity angles for individual decay particles. Because the helicity of an individual top quark in a $t\bar{t}$ event is random and the correlation is only apparent when both the top and the antitop helicities are considered, the spin correlation will not be visible in any of these individual helicity angles. Figure 5.12 shows the cosine of the lepton helicity angle, $\cos(\theta_\ell)$, while Figures 5.13 and 5.14 show $\cos(\theta_b)$ and $\cos(\theta_d)$ respectively. For completeness, we also include $\cos(\theta_u)$ in Figure 5.15, although this variable will not be used further. As with our previous plots, we see good agreement between the data and the model.

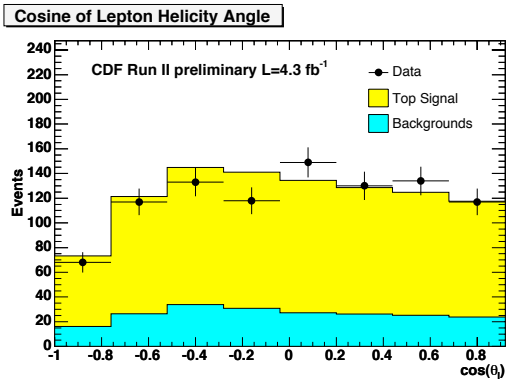

Figure 5.12. $\cos(\theta_\ell)$ after reconstruction

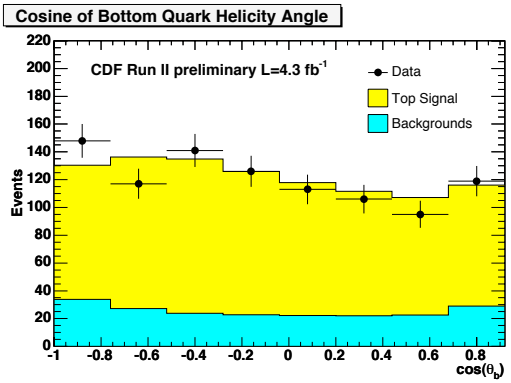

Figure 5.13. $\cos(\theta_b)$ after reconstruction

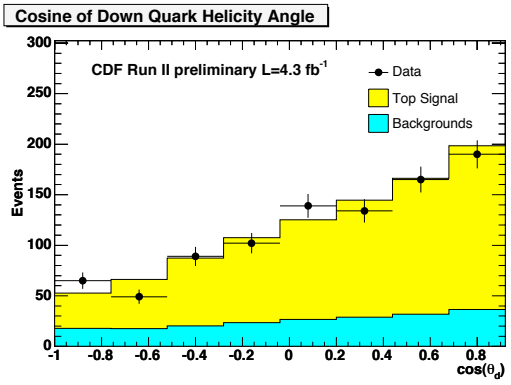

Figure 5.14. $\cos(\theta_d)$ after reconstruction

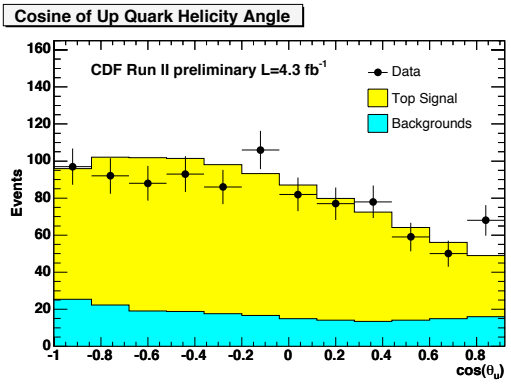

Figure 5.15. $\cos(\theta_u)$ after reconstruction

CHAPTER 6

Top Quark Spin and Helicity

The large mass of the top quark forces it to decay very quickly after it is produced - before de-polarizing QCD effects cause information about the top spin to be lost. This decay proceeds via the weak interaction, which passes information about the top quark spin to the decay products. This is one reason that the top quark is unique among quarks - its spin at production can be directly studied using information from its decay products.

6.1 Top Quark Pair Production

The spin correlation that exists between two pair-produced top quarks is dependent on the initial particles in the collision which created the $t\bar{t}$ pair and on the energy of the $t\bar{t}$ pair. At the Tevatron, the production process $q\bar{q} \rightarrow t\bar{t}$ dominates, accounting for 85% of $t\bar{t}$ events, with the remaining events being produced via $gg \rightarrow t\bar{t}$. In both cases, the majority of $t\bar{t}$ events are produced with relatively low energy, within ~ 100 GeV of the minimum energy threshold for $t\bar{t}$ production of approximately 345 GeV.

For this measurement, we work in the helicity basis, which is defined to be the basis where the top quark spin polarization axis is along the direction of motion of the top quark in the $t\bar{t}$ rest frame. A top quark that has its spin oriented along its direction of motion is said to be right-handed, while a top quark that has its spin oriented against its direction of motion is said to be left-handed.

6.1.1 Production by $q\bar{q} \rightarrow t\bar{t}$

When events are produced by $q\bar{q} \rightarrow t\bar{t}$, the interaction is mediated by a gluon with J=1. Close to the energy threshold, the $t\bar{t}$ pair does not have any orbital angular momentum and thus it will be in a $^3S_1$ spin state with three eigenstates [8]:

$$| ++\rangle$$

$$\frac{1}{\sqrt{2}}[| +-\rangle + | -+\rangle]$$

$$| --\rangle$$

In the helicity basis, since the two quarks move oppositely in the $t\bar{t}$ rest frame, they will have opposite helicities if they have the same spin direction. Thus, at threshold, two of the three possible states for the $t\bar{t}$ pair - 67% of events - will have opposite helicities. For $t\bar{t}$ pairs at energies far above threshold, helicity conservation requires that 100% of events have opposite helicity top quark pairs. For interpolating between $t\bar{t}$ pairs with energies near threshold and those with energies far above threshold, Equation 6.1 can be used as an approximation [8].

$$\frac{\sigma(t_L\bar{t}_R) + \sigma(t_R\bar{t}_L)}{\sigma(t_R\bar{t}_R) + \sigma(t_L\bar{t}_L)} = 2\frac{M_{t\bar{t}}^2}{4m_t^2} \tag{6.1}$$

The exact energy dependence of the helicity fractions can be determined by looking at the matrix elements for the production of same helicity and opposite helicity top quark pairs via $q\bar{q} \rightarrow t\bar{t}$. These matrix elements, summed over the spins of the initial quarks and over the color degrees of freedom in the initial and final states, are given in Equations 6.2 and 6.3 [10].

$$\sum_{RR,LL} |\mathcal{M}(q\bar{q} \rightarrow t\bar{t})|^2 = 8g^4(1-\beta^2)\sin^2\theta^* \tag{6.2}$$

$$\sum_{LR,RL} |\mathcal{M}(q\bar{q} \to t\bar{t})|^2 = 8g^4(1 + \cos^2\theta^*) \tag{6.3}$$

In these equations, g is the QCD coupling strength, θ^* is the top quark scattering angle with respect to the proton beam, and $\beta = v/c$ is the top quark velocity, which depends on the $t\bar{t}$ pair production energy and is limited to be between 0 and 1. As can be seen from Equation 6.2, the production of $t\bar{t}$ pairs with the same helicity is suppressed for high energy $t\bar{t}$ pairs ($\beta \to 1$).

6.1.2 Production by $gg \to t\bar{t}$

For the 15% of events produced by $gg \to t\bar{t}$, a $t\bar{t}$ pair produced at threshold will be in a $^1S_0$ spin state, with only one eigenstate:

$$\frac{1}{\sqrt{2}}[|+-\rangle - |-+\rangle]$$

In this case, all $t\bar{t}$ pairs will have the same helicity. As above, helicity conservation requires that top quark pairs produced via $gg \to t\bar{t}$ with energy far above threshold must have opposite helicities. The exact energy dependence can again be seen by looking directly at the matrix elements for $gg \to t\bar{t}$ production, as given in Equations 6.4 and 6.5, with the angular factor $\mathcal{Y}$ defined in Equation 6.6 [10].

$$\sum_{RR,LL} |\mathcal{M}(gg \to t\bar{t})|^2 = \frac{16}{3}g^4\mathcal{Y}(1-\beta^2)(1+\beta^2+\beta^2\sin^4\theta^*) \tag{6.4}$$

$$\sum_{LR,RL} |\mathcal{M}(gg \to t\bar{t})|^2 = \frac{16}{3}g^4\mathcal{Y}\beta^2\sin^2\theta^*(1+\cos^2\theta^*) \tag{6.5}$$

$$\mathcal{Y} = \frac{7+9\beta^2\cos^2\theta^*}{(1-\beta^2\cos^2\theta^*)^2} \tag{6.6}$$

In Equation 6.5, we can see that the production of top quark pairs with opposite

helicities is suppressed for low energies near threshold ($\beta \to 0$), while Equation 6.4 shows that at high energies ($\beta \to 1$), same helicity $t\bar{t}$ production is suppressed.

6.1.3 Total Spin Correlation in the $t\bar{t}$ System

Knowing the form of the matrix elements for $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$, the theoretical prediction for the $t\bar{t}$ spin correlation can be determined. Accounting for the relative fractions of $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ at the Tevatron and integrating over the expected $t\bar{t}$ energy distributions, we expect to find an opposite helicity fraction $F_{OH} = 0.70$, corresponding to a spin correlation coefficient of $\kappa = 0.40$ [8, 11].

6.2 The Decay of a Spinning Top Quark

Because of its large mass, the lifetime of a top quark before it decays by the spin-preserving electroweak interaction is only approximately $(1.5 \text{ GeV})^{-1} = 4.4 \cdot 10^{-25}$ seconds, much smaller than the average time for the QCD interaction to cause a spin flip, $(1.3 \text{ MeV})^{-1} = 5.1 \cdot 10^{-22}$ seconds [8]. This means that any angular correlation that exists between the parent top quarks will be passed on to their decay products. By examining these decay products, we can learn about the spins of the original top quarks. The spin of a top quark affects only the angular portion of the matrix element for top decay. The resulting kinematics of the decay products can be specified by only two angles: the angle $\psi_l{}^*$ between the lepton and the W momentum direction in the W rest frame and the angle θ_l between the lepton and the top spin direction in the top rest frame [31]. For a top quark that decays hadronically, the down-type quark from the W boson decay takes the place of the lepton. The angular part of the top decay matrix element can be written in terms of these two angles as in Equations 6.7, 6.8, and 6.9, with x being the ratio of the top quark mass to the W boson mass. The coefficients h_-, h_0, and h_+ determine the fraction of left handed, longitudinal, and right handed W bosons respectively, while P represents the magnitude of the top

quark polarization.

$$F(\cos(\theta_l), \cos({\psi_l}^*)) = \frac{8}{3} M_W^4 (x^2 - 1)(1 + \frac{x^2}{2}) g(\cos({\psi_l}^*)) f(\cos(\theta_l)) \tag{6.7}$$

where

$$g(\cos({\psi_l}^*)) = h_- \frac{3}{8}(1 - \cos({\psi_l}^*))^2 + h_0 \frac{3}{4}(1 - \cos^2({\psi_l}^*)) + h_+ \frac{3}{8}(1 + \cos({\psi_l}^*))^2 \tag{6.8}$$

$$f(\cos(\theta_l)) = \frac{1}{2}(1 + P \cos(\theta_l)) \tag{6.9}$$

In this matrix element, a right-handed top quark has $P = +1$, while a left-handed top quark has $P = -1$, with the sign of P being reversed for antitop quarks.

6.3 Top Quark Decay in HERWIG

We plan to measure the fraction of top quarks with opposite helicity using a template fitting method, with which we will determine the relative normalization of two templates - one with same helicity top pairs and one with opposite helicity top pairs - that best describes our data. This requires that we create simulated $t\bar{t}$ template samples with top pairs having either the same helicity or opposite helicities. We will implement the ability to do this using the event generator HERWIG. In simulating top quark decay, HERWIG makes use of the fact that the kinematics of all the decay products can be determined from the two angles ${\psi_l}^*$ and θ_l by defining the lepton direction first and then determining the kinematics of all other decay products from that lepton direction. HERWIG performs the three body decay $t \to l + \nu + b$ via the following process [26]:

- Choose $m^2_{b\nu}$ according to a phase-space*(V-A) reweighting

- In top rest frame, perform the two body decay $t \rightarrow l + (b\nu)$
- In $b\nu$ rest frame, perform the decay $b\nu \rightarrow b + \nu$
- Boost to the top rest frame, determine $m_{l\nu}$, and make a cut based on a comparison to the Breit-Wigner distribution for the W mass
- Boost all decay products to the lab frame

Early in the decay process, the direction of the lepton in the top rest frame is chosen, and once this choice is made, all other decay product kinematics are determined according to the matrix element above. If no spin correlations are included, this lepton direction follows an isotropic distribution, but by changing this distribution, spin correlation effects can be introduced.

6.4 Creation of Polarized Top Quark Samples

In standard HERWIG with no top spin correlations, the W helicity parameters are set to their theoretically predicted values, $h_- = 0.3$, $h_0 = 0.7$, and $h_+ = 0.0$, and the polarization magnitude P is set to 0. In a modified version of HERWIG, first introduced by Gene Guillian [31] and reintroduced for this measurement, we implement the option for the user to control these parameters directly. This is accomplished by using these parameters, as chosen by the user, to define the distribution of the lepton direction in the top quark rest frame. Because the lepton direction is used by HERWIG to determine the rest of the kinematics for the top quark decay, the remaining kinematics for the decay of a top quark with spin are accounted for automatically.

For the creation of our helicity template samples, we keep the W helicity parameters at their predicted values, but then set $P = \pm 1$ in the helicity basis to create samples with right- and left-handed top quarks. We use this modified event generator to create simulated $t\bar{t}$ samples, with top quark mass 172.5 GeV/c^2, for the four possible $t\bar{t}$ helicity states, shown in Table 6.1.

We know from Equation 6.9 that the helicity angle of the lepton (or down-type

Helicity State	P_{top}	$P_{antitop}$
$t_R\bar{t}_R$	+1	-1
$t_L\bar{t}_L$	-1	+1
$t_L\bar{t}_R$	-1	-1
$t_R\bar{t}_L$	+1	+1

Table 6.1. Helicity Basis Values for P

quark) from a top quark decay follows the distribution $1 + P\cos(\theta_l)$, where $P = \pm 1$, with the sign depending on whether we have a right- or left-handed quark and whether it is a top or an antitop quark. The helicity angles for the remaining top quark decay products follow a similar distribution, given in Equation 6.10 [10].

$$f(\cos(\theta_i)) = \frac{1}{2}(1 + \alpha_i \cos(\theta_i)) \tag{6.10}$$

The coefficient α_i, for $i = \ell, \nu, d, u, b$, determines the strength of the correlation between the original top quark spin and the angular distribution of the given decay product. For the lepton from a right-handed top quark, $\alpha_l = 1$, corresponding to the value $P = 1$ which was used as input when creating the polarized samples. The values of α_i for the remaining decay products are shown in Table 6.2. The down quark in the decay $t \to udb$ plays exactly the same role as the lepton in $t \to \ell\nu b$, making the identification of the down jet a critical part of this measurement in the lepton plus jets channel.

Decay Product	α_i
lepton	1.0
down quark	1.0
bottom quark	-0.41
neutrino	-0.31
up quark	-0.31

Table 6.2. Top decay product correlation coefficients for right-handed top quarks. The signs of these coefficients are reversed for left-handed top or right-handed antitop decay products.[10]

Figures 6.1 through 6.6 show the true distributions, before any event selection or reconstruction, for the helicity angles θ_i for the various decay products for right- and

left-handed tops and antitops in our four template samples. The expected correlations are present in all cases: for example, in Figure 6.1, the green curve corresponds to the sample $t_L\bar{t}_L$, with the expected slope of -1 for a lepton coming from the left-handed top quark. This can be compared to Figure 6.2, where the the green $t_L\bar{t}_L$ curve has the expected slope of +1 for a lepton from the left-handed antitop quark. Similar correlations can be seen in Figures 6.3 and 6.4, where the slopes are ± 0.41 corresponding to the expected bottom quark correlation, and in Figures 6.5 and 6.6, where the neutrino correlation gives slopes of ± 0.31. In all cases, the sign of the slope is determined by whether the decay product is coming from the top or the antitop quark, and whether that quark was right- or left-handed.

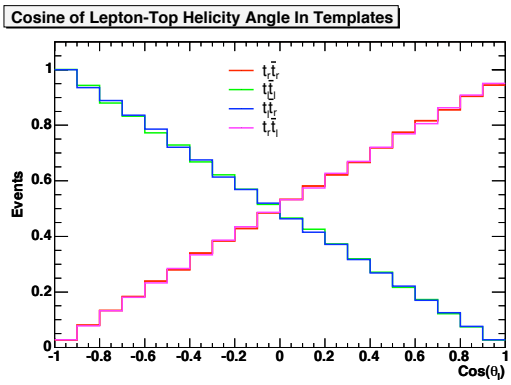

Figure 6.1. True $\cos(\theta_\ell)$ for Top Decays in Helicity Frame

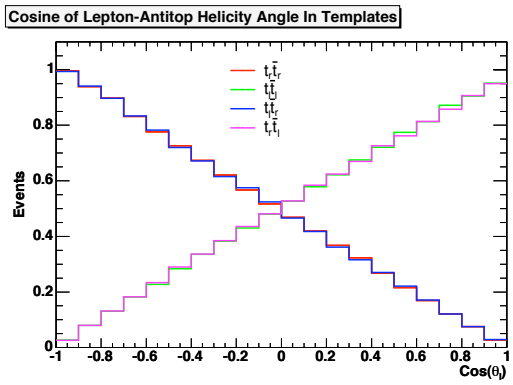

Figure 6.2. True $\cos(\theta_\ell)$ for Antitop Decays in Helicity Frame

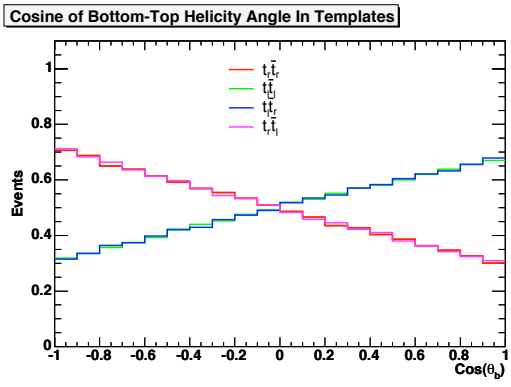

Figure 6.3. True $\cos(\theta_b)$ for Top Decays in Helicity Frame

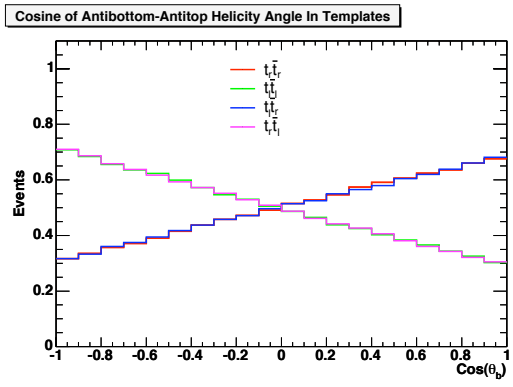

Figure 6.4. True $\cos(\theta_b)$ for Antitop Decays in Helicity Frame

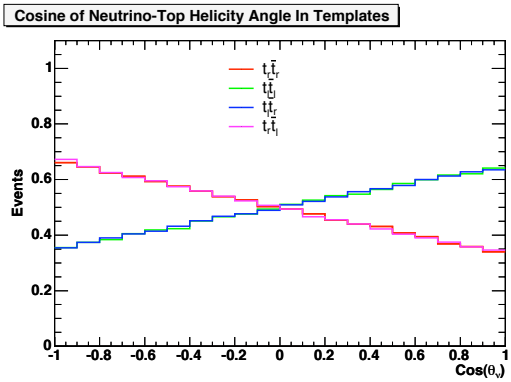

Figure 6.5. True $\cos(\theta_\nu)$ for Top Decays in Helicity Frame

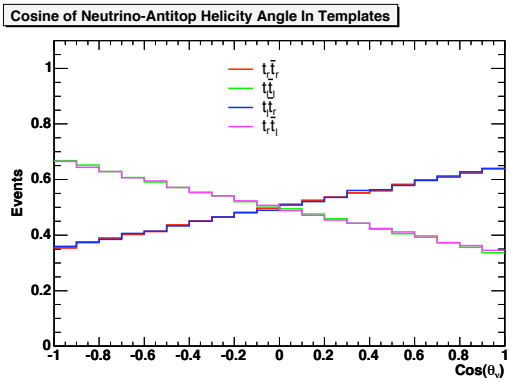

Figure 6.6. True $\cos(\theta_\nu)$ for Antitop Decays in Helicity Frame

6.5 Same Helicity and Opposite Helicity Combinations

We now have simulated samples corresponding to the four possible $t\bar{t}$ helicity states. The next step is to combine these into two templates, one for same helicity top quark pairs and one for $t\bar{t}$ pairs with opposite helicity. In order to do this, we must take advantage of two symmetry principles. There are three discrete symmetries that are important in particle physics, resulting from three discrete transformations:

- Parity (P), which transforms all spatial coordinates (x,y,z) into (-x,-y,-z), producing a mirror image.
- Charge Conjugation (C), which transforms particles into the corresponding antiparticles
- Time Reversal (T), which transforms the time coordinate t into $-t$.

According to current theoretical understanding, the strong interaction which produces top quark pairs conserves parity P and also conserves CP, the product of the parity symmetry transformation with the charge conjugation symmetry transformation. The parity transformation P acting on a top quark with spin will change the direction of motion of the quark, but will not change the direction of spin. Due to the definition of right- and left-handed helicities, the parity transformation therefore causes a change in handedness, so that, for example, a left-handed top quark t_L transforms into a right-handed top quark t_R. This means that the parity transformation

transforms $t_L\bar{t}_R$ into $t_R\bar{t}_L$, and parity conservation therefore requires that the cross-section for producing $t\bar{t}$ pairs in the helicity state $t_L\bar{t}_R$ is equal to the cross section for producing $t_R\bar{t}_L$ pairs, $\sigma(t_L\bar{t}_R) = \sigma(t_R\bar{t}_L)$. Similarly, the combined CP transformation not only switches the handedness of a quark, but also transforms a quark into an antiquark, so that, for example, we have the transformation of t_R into $\bar{t}_L$. Thus, CP transforms $t_R\bar{t}_R$ into $t_L\bar{t}_L$, and CP conservation requires $\sigma(t_R\bar{t}_R) = \sigma(t_L\bar{t}_L)$. Using these equalities, we can define the same helicity (SH) and opposite helicity (OH) samples to be the symmetric sum of the corresponding components, as in Equations 6.11 and 6.12. These samples are explored further in Chapter 7.

$$SH = t_R\bar{t}_R + t_L\bar{t}_L \tag{6.11}$$

$$OH = t_L\bar{t}_R + t_R\bar{t}_L \tag{6.12}$$

CHAPTER 7

Measuring the Helicity Fractions

As we have seen, the spin of a top quark is correlated with the angular distributions of that quark's decay products. Thus, by considering the angular correlations among the top decay products, we can determine the spin correlation of the original $t\bar{t}$ pair. In particular, we will consider distributions involving cosines of the top decay product helicity angles, $\cos(\theta_i)$, as defined in Chapter 6. Using our same helicity and opposite helicity samples, we will create models of $\cos(\theta_1) \cdot \cos(\theta_2)$ distributions, where particle 1 and particle 2 are decay products of the top and antitop quarks respectively. We will then combine these same helicity and opposite helicity models and determine the relative normalization that best matches our data.

7.1 Distributions Used in Measurement

In order to determine a spin correlation between the two quarks in a $t\bar{t}$ pair, we need to consider information from the decay products of both the leptonically and hadronically decaying top quarks. Because the helicity angles of the lepton and the down-type quark have the strongest correlation with the spin of the original top quark, we first consider distributions involving these two decay products. The first and most natural thing to consider is a two-dimensional distribution with $\cos(\theta_\ell)$ on one axis and $\cos(\theta_d)$ on the other. If we assume that the down quark can always be identified correctly, then it is completely equivalent to a lepton, so our two-dimensional distribution is equivalent to $\cos(\theta_\ell^+)$ vs. $\cos(\theta_\ell^-)$ in a dilepton event. This distribution,

with the angles being calculated using the true momenta of the decay products, is shown in Figure 7.1 for our opposite helicity basis sample and in Figure 7.2 for our same helicity basis sample, with darker colors indicating a higher concentration of events. As can be seen, there is a clear distinction between the two samples.

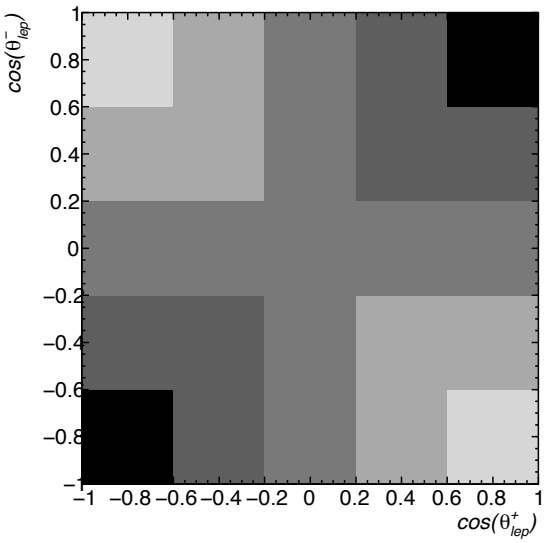

Figure 7.1. Distribution of $\cos(\theta_\ell^+)$ vs. $\cos(\theta_\ell^-)$ in OH basis sample

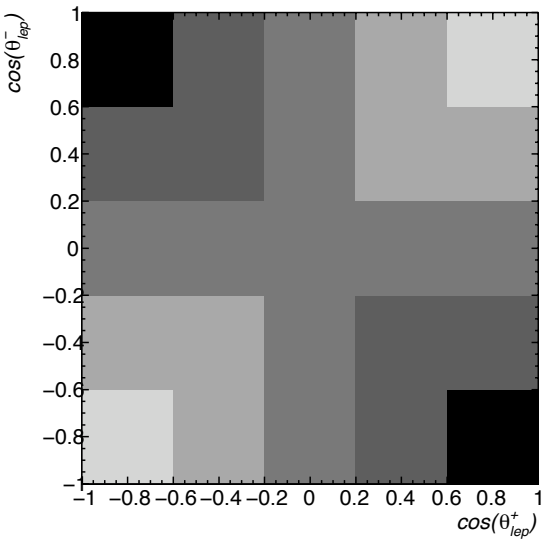

Figure 7.2. Distribution of $\cos(\theta_\ell^+)$ vs. $\cos(\theta_\ell^-)$ in SH basis sample

The angles in the case of opposite helicity $t\bar{t}$ tend to be either both large or both small, while for same helicity pairs one angle is small and one is large. This orthogonality allows us to make a simplification by considering a one-dimensional product, rather than the two-dimensional distributions shown in Figures 7.1 and 7.2. The helicity angle bilinear $\cos(\theta_\ell) \cdot \cos(\theta_d)$ is our distribution of choice. We will see in Section 7.2 that this choice will not cause a loss in sensitivity.

If we assume that we can always correctly identify the down-type quark, then $\cos(\theta_\ell) \cdot \cos(\theta_d)$ is equivalent to $\cos(\theta_\ell^+) \cdot \cos(\theta_\ell^-)$. In Figures 7.3, we see this distribution, calculated using the true decay product momenta, in our opposite helicity and same helicity basis samples as well as a HERWIG sample which does not include spin correlation effects. The opposite helicity basis sample tends towards positive values of

$\cos(\theta_\ell^+) \cdot \cos(\theta_\ell^-)$, while the same helicity basis sample tends towards negative values.

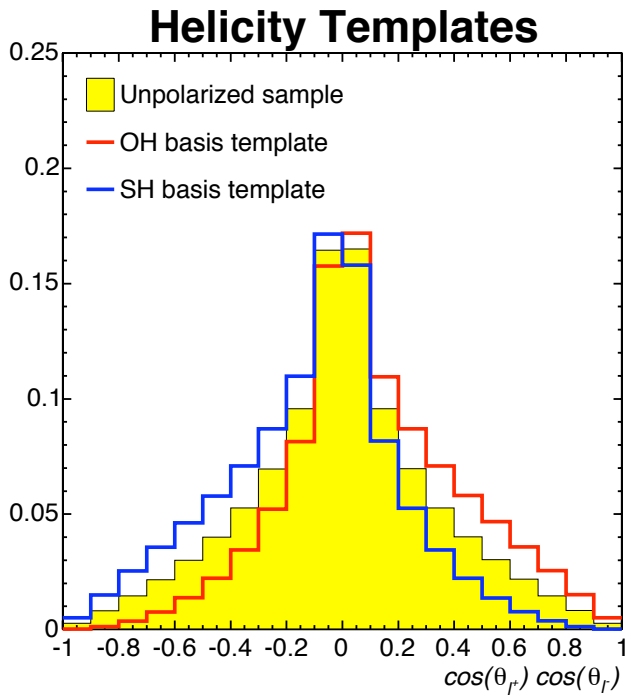

Figure 7.3. Distribution of $\cos(\theta_\ell^+) \cdot \cos(\theta_\ell^-)$ in OH and SH basis samples

As was discussed in Chapter 5, however, the down-type quark cannot always be reliably identified. Using the algorithm outlined in [10], we can correctly identify the true down quark in only 60% of events. If we recreate Figure 7.3 but use our algorithm for choosing the down quark, rather than having it always be correctly identified, then the results are shown in Figure 7.4. The algorithm for guessing the down quark reduces the asymmetry between the same helicity and opposite helicity samples, but there is still an easily visible difference between the two, and it is upon this difference that our measurement is based.

Because of the sensitivity loss due to the down identification, we also consider the information contained in other correlations in the event. In particular, we consider $\cos(\theta_\ell) \cdot \cos(\theta_b)$, where $\cos(\theta_b)$ is the cosine of the helicity angle for the bottom quark from the hadronically decaying top quark. Although the bottom quark is less strongly

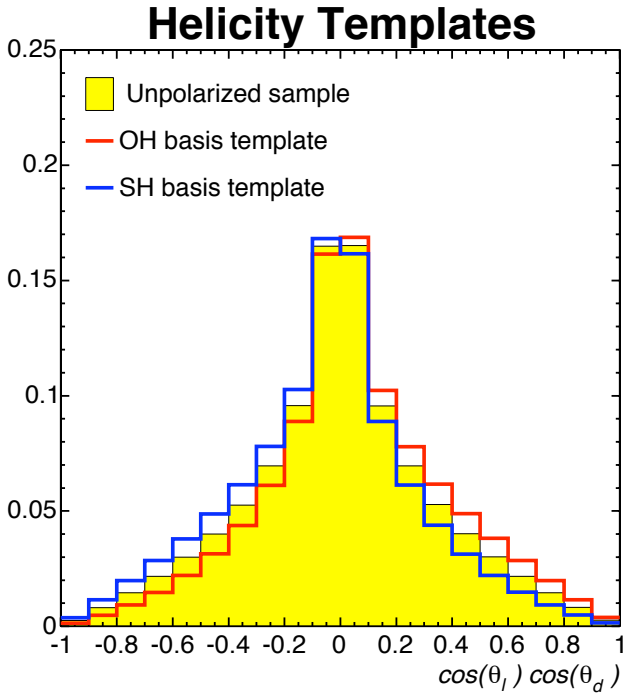

Figure 7.4. Distribution of $\cos(\theta_\ell) \cdot \cos(\theta_d)$ in OH and SH basis samples

correlated to the top quark spin than the down quark is, it is more easily identified because it can be tagged. Thus, we will explore the sensitivity of both one-dimensional distributions, $\cos(\theta_\ell) \cdot \cos(\theta_d)$ and $\cos(\theta_\ell) \cdot \cos(\theta_b)$. In addition, we will also explore the two-dimensional distribution $\cos(\theta_\ell) \cdot \cos(\theta_d)$ vs. $\cos(\theta_\ell) \cdot \cos(\theta_b)$, due to the possibility that the additional information it contains compared to either 1-D distribution will allow us to increase our sensitivity.

Our final measurement will also require us to account for the fact that some background events pass our event selection and will have an effect on the data distributions that we are using. We must ensure that our model for these background events is accurate. The various background processes that could pass our event selection were discussed in Chapter 4. Figure 7.5 shows the $\cos(\theta_\ell) \cdot \cos(\theta_d)$ distribution for each of the background components, as well as their relative normalizations. In order to again verify our background model, we consider the anti-tagged sample, which corresponds to the data that passes all of our selection cuts except that it has no b-tagged

jets. This is a background dominated sample, and thus is a good place to check our background modeling. Figure 7.6 shows the $\cos(\theta_\ell) \cdot \cos(\theta_d)$ distribution in the anti-tagged sample, and we see good agreement between the data and model.

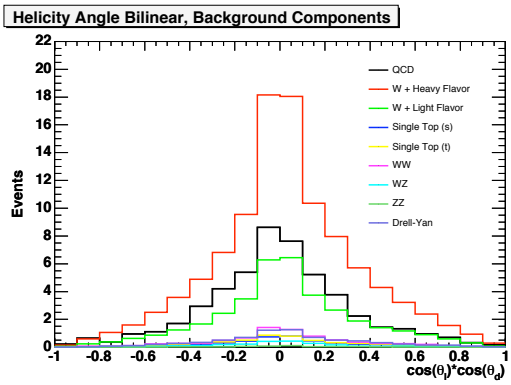

Figure 7.5. $\cos(\theta_\ell) \cdot \cos(\theta_d)$ for the various background components

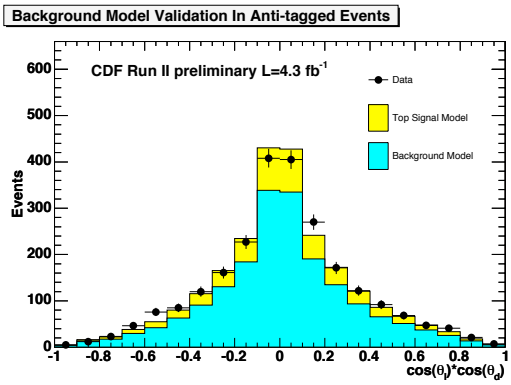

Figure 7.6. Test of the background modeling in the anti-tagged sample

7.2 Measurement of the Opposite Helicity Fraction

We will make our measurement by fitting our data distribution using the sum of the same and opposite helicity templates, allowing us to determine the relative fraction of opposite helicity events in our $t\bar{t}$ sample. We perform this fit by optimizing the binned likelihood ratio [32], equivalent to finding the minimum of the function

$$\chi^2 = 2\sum_i \left[t_i - n_i + n_i ln(\frac{n_i}{t_i}) \right] \tag{7.1}$$

where n_i and t_i are respectively the data and model content of the i^{th} bin of the histogram used for the fit. This can be generalized to multidimensional histograms by looping the index i over all bins in every dimension. We use the MINUIT package [30] to find the minimum of the function and extract the measured parameters.

The predicted model content t_i is related to our helicity basis templates by

$$t_i = (F_{OH} t_i^o + F_{SH} t_i^s) \tag{7.2}$$

where t_i^o and t_i^s denote the i^{th} bin content of our opposite helicity and same helicity templates respectively. The opposite and same helicity fractions, F_{OH} and F_{SH}, are further constrained according to the equation

$$F_{OH} + F_{SH} = 1 \tag{7.3}$$

although we do not individually require both F_{OH} and F_{SH} to take on physical values between 0 and 1.

Equation 7.2 applies to a sample which is composed purely of actual $t\bar{t}$ signal events, but we must also consider the background events that pass our selection and are present in our data sample. In order to do so, we use our background prediction, as developed in Chapter 4, as an additional template in our fit. Including this background template requires us to modify the model prediction t_i in Equation 7.2 so that it becomes

$$t_i = (F_{OH}t_i^o + F_{SH}t_i^s)S_t + b_i S_b, \tag{7.4}$$

Here, b_i is the content of the i^{th} bin in our background distribution. S_t and S_b are nuisance scale factors which are included in order to account for uncertainties affecting the overall normalization of the number of signal and background events respectively. There are various systematic effects that cause these uncertainties, including uncertainties in acceptance and luminosity, as well as linear scale factors accounting for differences in theoretical modeling and experimental and simulation efficiencies. These are equivalent to the systematic uncertainties that would arise in a $t\bar{t}$ production cross section measurement [33]. The largest individual uncertainty included in S_t is a 6.1% luminosity uncertainty, and the overall uncertainty on S_t is 9.0%. The largest individual uncertainty included in S_b is a 19.3% uncertainty on the normalization of W plus heavy flavor events, and the overall uncertainty on S_b is 20.8%. The inclusion of these nuisance scale factors amounts to Gaussian-constraining the

number of background events to the normalization found in Chapter 4 of 215 ± 48 events.

The inclusion of these nuisance parameters requires us to expand our expression in Equation 7.1, so that it becomes

$$\chi^2 = 2\sum_i \left[t_i - n_i + n_i ln(\frac{n_i}{t_i}) \right] + \sum_{k=t,b} \left(\frac{S_k - 1}{\sigma_k} \right)^2 \tag{7.5}$$

with t_i as defined in Equation 7.4. This is the final form of the function used in our fit, and it is this function that MINUIT will minimize in order to extract the opposite helicity fraction F_{OH}.

We now want to test our fitting procedure to ensure that it performs the measurement as expected. In order to do so, we perform our fit in a simulated sample with a known value for F_{OH}. We create a simulated $t\bar{t}$ sample using the event generator HERWIG that has a known true value $F_{OH} = 0.7292$. We scale the signal sample to 786 total events, we add to it our background prediction normalized to 215 events, so that the total is equal to the 1,001 candidate events in the data, and then we perform our fit in the simulated sample. The results of this fit are shown in Figure 7.7, and the measured opposite helicity fraction

$$F_{OH} = 0.763 \pm 0.291$$

is consistent with the true value of F_{OH} in the HERWIG sample.

Having established our fitting procedure, we can now also verify that no sensitivity is lost using the helicity angle bilinears, as mentioned in Section 7.1. Table 7.1 shows the expected statistical uncertainty on the opposite helicity fraction F_{OH} in a fit to a simulated sample containing 10,000 events, comparing the results for the one-dimensional fit of $\cos(\theta_\ell) \cdot \cos(\theta_d)$ to the two-dimensional fit of $\cos(\theta_\ell)$ vs. $\cos(\theta_d)$, as well as for various numbers of bins. As can be seen, there is virtually no change

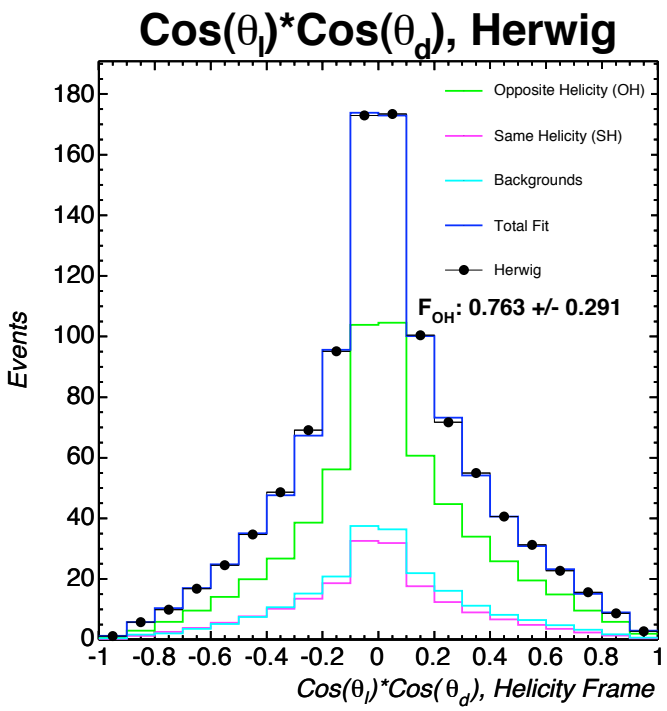

Figure 7.7. Measurement of F_{OH} using $cos(\theta_\ell) \cdot cos(\theta_d)$ distribution in HERWIG

in the statistical uncertainty when going from 2-D to 1-D, but the uncertainty does decrease when the number of bins is increased.

Histogram binning	Statistical uncertainty ΔF_{OH} for $N = 10,000$ sample				
	N_{bin} =2	N_{bin} =4	N_{bin} =5	N_{bin} =10	N_{bin} =20
1-D, N_{bin}	±2.0%	±1.6%	±1.5%	±1.4%	±1.4%
2-D, $N_{bin} \times N_{bin}$	±1.9%	±1.5%	±1.5%	±1.4%	±1.4%

Table 7.1. Comparison of 1-D versus 2-D template fit statistical uncertainties

7.3 Sensitivity Studies

With our fitting procedure working properly, we must now return to the question of which distribution to use when making the measurement. There are three possibilities, described in Section 7.1, from which this choice must be made: $\cos(\theta_\ell) \cdot \cos(\theta_d)$, $\cos(\theta_\ell) \cdot \cos(\theta_b)$, and the two-dimensional combination $\cos(\theta_\ell) \cdot \cos(\theta_d)$ vs. $\cos(\theta_\ell) \cdot \cos(\theta_b)$. In order to make a choice, we evaluate the expected size of the statisti-

cal uncertainty that will result from each distribution, giving us an estimate of the experimental sensitivity.

This sensitivity will be measured by performing a series of measurements (called "pseudo-experiments") using simulated data samples with known values of F_{OH}. Each pseudo-experiment analysis is prepared in the following way:

- Require that each simulated sample will have a size equal to the number of reconstructed candidate events in the data (1,001)
- Normalize the background template according to our prediction from Chapter 4 for the total number of events (215±48)
- Choose an input value for the opposite helicity fraction F_{OH}
- Create a signal distribution S from the OH and SH templates via $S = F_{OH} \cdot OH + (1 - F_{OH}) \cdot SH$
- Normalize the signal distribution to the expected number of top pair events $(1001 - 215 = 786)$
- Add the signal and background distributions together
- Poisson fluctuate each bin to create a simulated data distribution
- Perform the fit using this simulated data

In each pseudo-experiment, we record two values: the measured opposite helicity fraction F_{OH}^{meas} and the pull $P = (F_{OH}^{meas} - F_{OH}^{true})/\sigma$, where σ is the statistical uncertainty. We repeat this procedure 10,000 times, as if we were to perform this CDF measurement in 10,000 different samples, each with an integrated luminosity of 4.3 fb^{-1}, with each sample being an independent Poisson fluctuation around the true distribution determined by the input value of F_{OH}. The width of the F_{OH}^{meas} distribution will give us the expected statistical uncertainty of our measurement, while the pull distribution will identify any bias in the fitting procedure. For an unbiased fit, the pull distribution should have a mean of 0 and a width of 1.

We first study fits using the $\cos(\theta_\ell) \cdot \cos(\theta_d)$ distribution. The lepton and down-

type quark are the decay products that are most strongly correlated with the spin of the original top quark, but this strong correlation is diminished by the fact that the down quark is not always identified correctly. We performed 10,000 pseudo-experiments using this distribution, with an input opposite helicity fraction of 0.70. The measured values of F_{OH} are shown in Figure 7.8, while the pull distribution is shown in Figure 7.9. The pull distribution has mean 0 and width 1, as expected for an unbiased fit. The mean measured result is close to 0.7, indicating that the fit can accurately reproduce the true value of F_{OH}. The width of the distribution of the measured F_{OH} values indicates that, using the $\cos(\theta_\ell) \cdot \cos(\theta_d)$ distribution, we can expect a statistical uncertainty of 0.292.

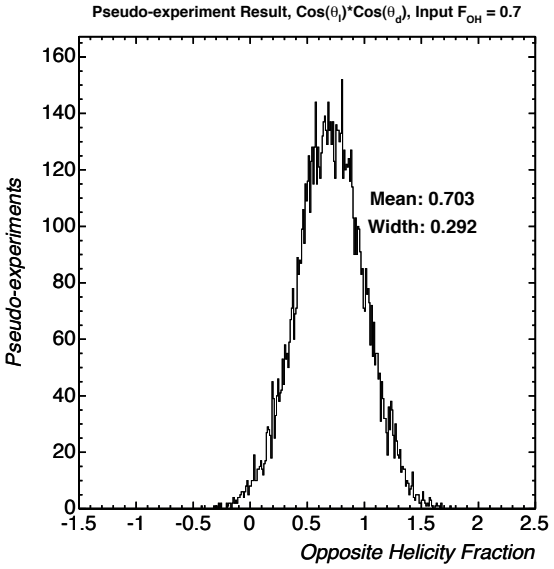

Figure 7.8. Pseudo-experiment results using $\cos(\theta_\ell) \cdot \cos(\theta_d)$

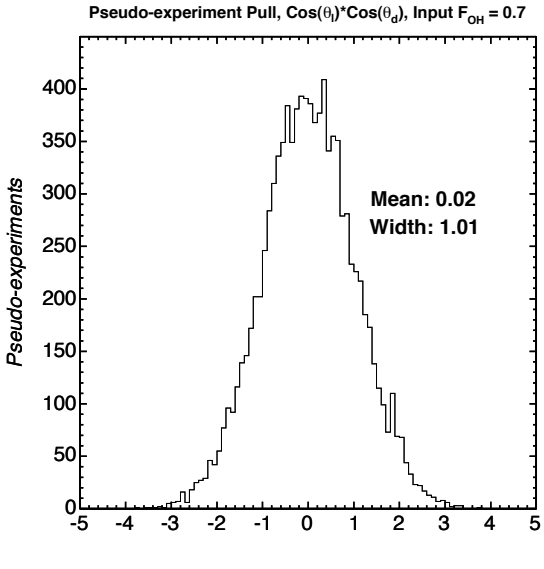

Figure 7.9. Pull distribution using $\cos(\theta_\ell) \cdot \cos(\theta_d)$

Next we consider the $\cos(\theta_\ell) \cdot \cos(\theta_b)$ distribution. The bottom quark is more weakly correlated with the top quark spin than the down quark, but it can be identified more efficiently. We again performed 10,000 pseudo-experiments using this distribution, with an input opposite helicity fraction of 0.70. The measured values of F_{OH} and the pull distribution are shown in Figures 7.10 and 7.11 respectively. The pull distribution has mean 0 and width 1, verifying that we have an unbiased fit.

The mean measured result is again close to 0.7, while width of the distribution of the measured F_{OH} values indicates that, using the $\cos(\theta_\ell) \cdot \cos(\theta_b)$ distribution, we can expect a statistical uncertainty of 0.357.

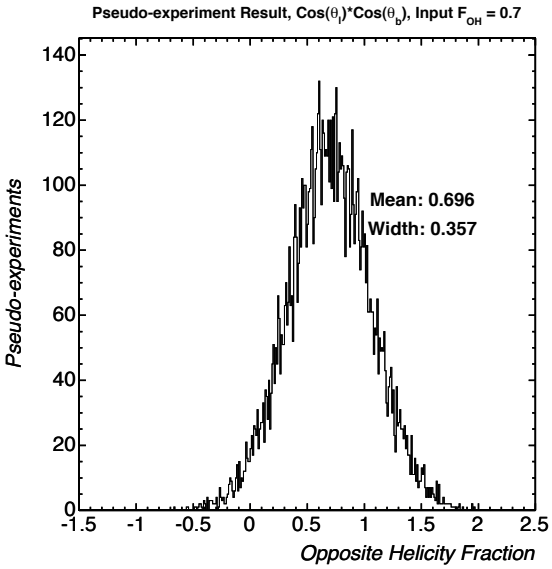

Figure 7.10. Pseudo-experiment results using $\cos(\theta_\ell) \cdot \cos(\theta_b)$

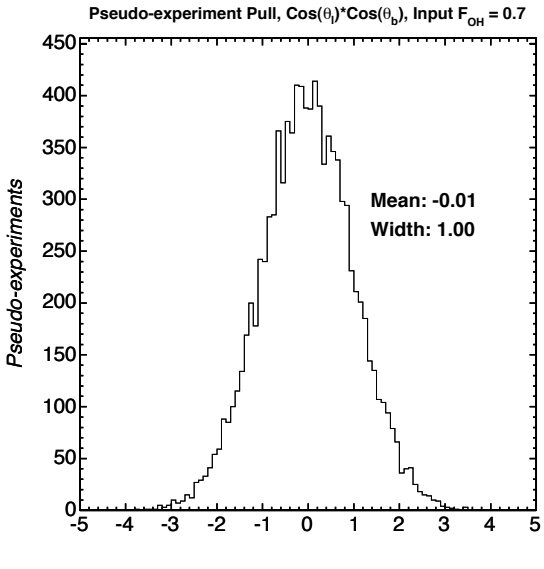

Figure 7.11. Pull distribution using $\cos(\theta_\ell) \cdot \cos(\theta_b)$

Finally, we look at the two-dimensional distribution $\cos(\theta_\ell) \cdot \cos(\theta_d)$ vs. $\cos(\theta_\ell) \cdot \cos(\theta_b)$. By including the information from both of the one-dimensional distributions, we hope to see improved sensitivity and a decrease in the expected statistical uncertainty. Once again, we performed 10,000 pseudo-experiments using this distribution, with an input opposite helicity fraction of 0.70. Figures 7.12 and 7.13 show the measured values of F_{OH} and the pull distribution respectively. The pull distribution once more shows an unbiased fit, with mean 0 and width 1. The mean measured value of F_{OH} is close to the expected 0.7, and the width of the distribution of the measured F_{OH} values indicates an expected statistical uncertainty of 0.232. The two-dimensional distribution does show the desired improvement in the statistical uncertainty, so we choose this distribution to be the one that will be used in the final measurement. The expected uncertainties for the three distributions are summarized in Table 7.2.

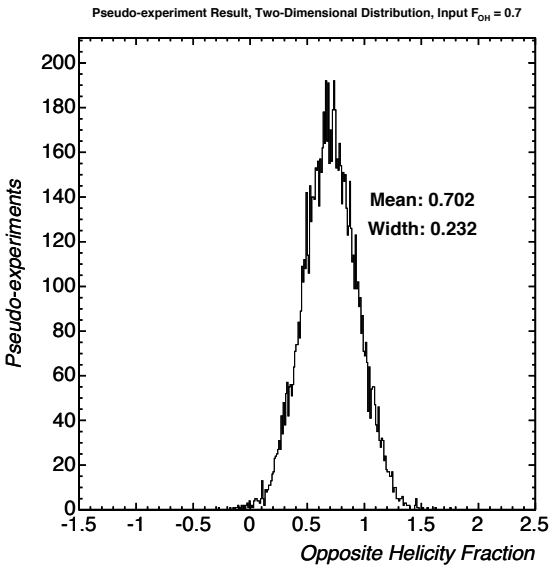

Figure 7.12. Pseudo-experiment results using 2D fit distribution

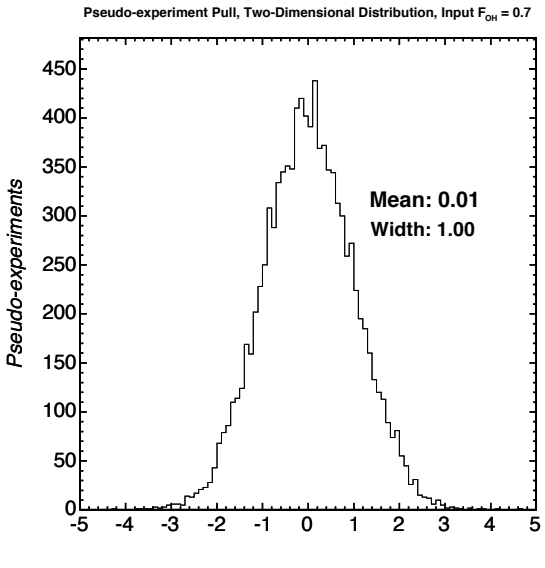

Figure 7.13. Pull distribution using 2D fit distribution

Distribution	Expected Uncertainty
$\cos(\theta_\ell) \cdot \cos(\theta_d)$	0.292
$\cos(\theta_\ell) \cdot \cos(\theta_b)$	0.357
Two-Dimensional	0.232

Table 7.2. Expected Statistical Uncertainties

Our final two-dimensional distribution contains information from three of the six $t\bar{t}$ decay products in the lepton plus jets channel - the lepton, the down quark and the bottom quark from the hadronically decaying top quark. There are potentially three additional decay products carrying information about the top pair spin correlation - the neutrino, the up quark, and an additional bottom quark. Various distributions involving these other decay products were tested as well, but none provided any improvement in the expected statistical uncertainty of the result.

7.3.1 Linearity Test

There is one final check of our fitting procedure that must be performed. Thus far, we have shown that the fit result is unbiased when the input value of F_{OH} is 0.7. Now, we want to verify that the technique works and is unbiased for any value of F_{OH}, so we check various input opposite helicity fractions. This is done using the same pseudo-experiment procedure that was used for the sensitivity check. Because the simulated data samples used in the pseudo-experiments can be tuned to have any true value of F_{OH}, we perform 10,000 pseudo-experiments with various input values of F_{OH}, and then verify that the measured opposite helicity fractions match the input values. The results of this linearity check are shown in Table 7.3. As can be seen, the response of our fitter is accurate at all F_{OH}, with the output result always closely matching the input value.

Input F_{OH}	Measured F_{OH}
0.00	0.003 ± 0.234
0.10	0.099 ± 0.236
0.20	0.199 ± 0.237
0.30	0.301 ± 0.238
0.40	0.404 ± 0.238
0.50	0.499 ± 0.238
0.60	0.601 ± 0.237
0.70	0.698 ± 0.235
0.80	0.802 ± 0.235
0.90	0.903 ± 0.235
1.00	1.001 ± 0.234

Table 7.3. Check of the linearity of the fitting procedure

CHAPTER 8

Systematic Uncertainties

There are multiple effects, related to detector response, interaction modeling, and other issues, which can contribute to the uncertainty of our measurement in a way that has not yet been taken into account. We divide these systematic uncertainties into two basic categories: background-based systematic effects and signal-based systematic effects. We consider first the two background-based uncertainties, and then we look into the signal based uncertainties. Each uncertainty is calculated in its own way but the same general procedure is followed in each case. First, we create a simulated data sample from the sum of our background and signal predictions, where the systematic parameter of interest has been varied from the nominal value. We then perform our measurement in this new sample, and compare the result to the same measurement in the nominal sample. The comparison of these two values yields the uncertainty due to the chosen systematic effect.

8.1 Background Size

Our background model predicts that there will be 215 ± 48 events in our dataset containing 1,001 events. In order to evaluate the systematic uncertainty due to our

Variation	F_{OH}
Nominal	0.7039 ± 0.0173
$+1\sigma$	0.7037 ± 0.0183
-1σ	0.7040 ± 0.0164

Table 8.1. Opposite helicity fractions under variations of background size

Variation	F_{OH}
Nominal	0.7039 ± 0.0173
QCD shape	0.7489 ± 0.0200
W+HF shape	0.7077 ± 0.0166

Table 8.2. Opposite helicity fractions under variations of background shape

uncertainty in the background normalization, we create simulated datasets by adding our background prediction to a $t\bar{t}$ signal sample created by the event generator HERWIG. In the nominal case, the background prediction is normalized to 215 events, and then we also create simulated samples where this normalization is scaled by $+1\sigma$ and -1σ. The results of our measurement of the opposite helicity fraction F_{OH} in each of these samples are shown in Table 8.1. As our systematic uncertainty, we use one-half of the largest difference between measurement results, equal to 0.0015.

8.2 Background Shape

There will also be an uncertainty due to the shape of our background model. To evaluate this uncertainty, we start with a nominal simulated sample which is the same as above: a HERWIG signal sample added to our background prediction. The nominal background model is then replaced with background models where the shapes are given by the shapes of our two largest background components: QCD fakes and W+heavy flavor jets. The results of performing our measurement of F_{OH} in these samples are given in Table 8.2. Again, we use one-half of the largest difference between measurements for the uncertainty, yielding a total of 0.0225 as our background shape uncertainty.

8.3 Initial and Final State Radiation

For our first signal-based systematic uncertainty, we consider initial and final state radiation (ISR/FSR). When a pair of top quarks are produced, the incoming partons will occasionally radiate a gluon before the collision. Additionally, the partons from

Variation	F_{OH}
Nominal	0.7325 ± 0.0229
$+1\sigma$	0.7707 ± 0.0354
-1σ	0.7118 ± 0.0363

Table 8.3. Opposite helicity fractions under ISR/FSR variations

the top quark decays can also radiate gluons. These radiated gluons can sometimes be energetic enough to produce additional jets, which can create measurement error because sometimes these extra jets are identified as belonging to the top decay by our reconstruction procedure. The branching ratios for these gluon radiation effects are not well understood, and the amount of ISR/FSR radiation that takes place in our simulated $t\bar{t}$ samples is treated as a tunable parameter.

In order to evaluate the systematic uncertainty in our measurement due to ISR/FSR, we start with a nominal sample made up of the sum of our background prediction and a $t\bar{t}$ signal sample created using the event generator PYTHIA. This signal sample is reweighted to have a true opposite helicity fraction of $F_{OH} = 0.7$. We then replace this nominal signal sample with two samples which have the amount of ISR/FSR shifted by $+1\sigma$ and -1σ respectively. Again, these samples were created using PYTHIA and reweighted to have a true opposite helicity fraction of $F_{OH} = 0.7$. Performing our measurement in each of these simulated samples gives the results shown in Table 8.3. Once again we take one-half of the largest difference, 0.0295, to be our uncertainty due to ISR/FSR.

8.4 Jet Energy Scale

When jets are measured in the detector, the energy in a given jet is spread over several sections of the calorimeter. To measure the total energy in a jet, we must integrate over these sections using a jet clustering algorithm, but once this is performed there are still several corrections that must be applied, as described in Chapter 3. These corrections are necessary because different calorimeter systems in the detector have

different responses to various particles, the response of the calorimeter as a function of the true particle energy is non-linear, there are areas within the detector - such as spaces between calorimeter sections - where there is no calorimeter coverage, and sometimes the energy from a jet is spread out beyond the area included in the jet clustering algorithm. The uncertainty associated with these corrections is included as the jet energy scale (JES) systematic uncertainty. The individual uncertainties that the JES uncertainty takes into account include:

- Relative Correction: uncertainty in the corrections made to account for the η dependent calorimeter response.
- Underlying Event Correction: uncertainty in the corrections accounting for the energy associated with the spectator partons in an event.
- Absolute Correction: uncertainty in the correction of the jet energy measured in the calorimeter accounting for any non-linearity in the detector response and for energy loss in the regions of the detector where there is no calorimeter coverage.
- Out-of-Cone Correction: uncertainty in correcting back to the particle-level energy by accounting for leakage of radiation outside of the jet clustering cone.
- Splash-Out Correction: uncertainty in the correction made to account for energy leakage beyond that which is accounted for by the out-of-cone correction.

An incorrect accounting of these effects would cause an over or under correction of the amount of energy in jets in different parts of the detector. This could then lead to incorrect jet assignments in our event reconstruction, causing a possible bias in the measurement. To evaluate the size of this effect, we start with a nominal sample consisting of our background prediction added to a signal sample made of $t\bar{t}$ events generated using HERWIG. We then create two additional samples where the signal sample is replaced by a sample where the jet energy scale corrections are scaled by $+1\sigma$ and -1σ respectively. The results of performing our measurement in these samples are shown in Table 8.4. To calculate the uncertainty, we again take one-half

Variation	F_{OH}
Nominal	0.7039 ± 0.0173
$+1\sigma$	0.6520 ± 0.0173
-1σ	0.7365 ± 0.0174

Table 8.4. Opposite helicity fractions under JES variations

of the largest difference, giving 0.0423.

8.5 Parton Distribution Function

Parton distribution function (PDFs) are empirically determined functions describing the momentum distributions of the partons and gluons inside a proton or antiproton. In our simulated samples, the choice of which PDF to use determines the energy spectrum of the $t\bar{t}$ system, which affects the overall kinematics of the $t\bar{t}$ decay and could have an effect on our measurement. To study this effect due to the uncertainty in the PDF, we compare samples consisting of our background prediction added to simulated $t\bar{t}$ samples made with PYTHIA and using 46 different PDF sets: CTEQ6M, CTEQ5L, MRST72, MRST75, CTEQ6L, CTEQ6L1, CTEQ6M $\pm$ 20 (eigenvectors for orthogonal PDF sets). The differences between the CTEQ6M $\pm$ 20 eigenvectors are calculated and added in quadrature. We then separately determine the difference between MRST72 and CTEQ5L. The larger of these 2 values is then added in quadrature with the difference between the results for the MRST72 and MRST75 sets, and the result is taken as our PDF systematic uncertainty. This procedure yields an uncertainty of 0.0071.

8.6 Parton Shower

The simulation of $t\bar{t}$ events is expected to correctly model the entire event, from production through decay and up to the final state particles seen by the detector. The actual production and decay process for top quark pairs is well understood theoretically and can be implemented in event generators, but the parton showering process,

Variation	F_{OH}
ALPGEN+PYTHIA	0.6973 ± 0.0239
ALPGEN+HERWIG	0.7029 ± 0.0235

Table 8.5. Opposite helicity fractions under variation of the parton showering

Variation	F_{OH}
Tune Apro	0.7813 ± 0.0228
Tune ACRpro	0.7900 ± 0.0229

Table 8.6. Opposite helicity fractions for different color reconnection models

in which partons from the decay fragment into jets, is not as well understood, and different event generators include different implementations of this parton showering. In order to evaluate the magnitude of the uncertainty in our measurement due to parton showering, we consider two different simulated samples consisting of a simulated $t\bar{t}$ sample added to our background prediction. These two simulated $t\bar{t}$ samples both use the event generator ALPGEN to simulate the $t\bar{t}$ production and decay, but then the parton showering is simulated by PYTHIA and HERWIG respectively. The results of performing our measurement in these two samples are shown in Table 8.5. We use the difference between these two results, 0.0056, as the parton shower uncertainty.

8.7 Color Reconnection

In hadron collisions, there can be non-trivial effects due to interactions with the beam remnants and the underlying event. These effects, called color reconnection effects, are difficult to calculate theoretically, and event generators rely on comparisons to previously collected data in order to model them. To account for any error this modeling might introduce into our measurement, we compare two samples made up of a $t\bar{t}$ signal that was simulated using different tunes for the color reconnection effects added to the background prediction. We use tunes Apro and ACRpro for the PYTHIA event generator. Our systematic uncertainty, 0.0087, is determined by taking the difference in the measured results for F_{OH}, shown in Table 8.6.

Variation	True	Measured
PYTHIA	0.5051 ± 0.0009	0.5607 ± 0.0160
HERWIG Sample 1	0.5056 ± 0.0023	0.5679 ± 0.0395
HERWIG Sample 2	0.5081 ± 0.0024	0.4611 ± 0.0404

Table 8.7. Opposite helicity fractions at truth level and after reconstruction

8.8 Generator Dependence

One thing that we need to check is the behavior of our fitting technique in a sample where there is a null correlation ($F_{OH} = 0.50$). To do this, we consider a simulated sample consisting of our background prediction added to a signal sample created using PYTHIA, which does not include spin correlation effects and is expected to have an opposite helicity fraction of 0.50. The result of performing our measurement in this sample is $F_{OH} = 0.5607 \pm 0.0160$. It appears that there might be a small possible bias in this sample, so we repeat the fit using two independent HERWIG samples which were created to have $F_{OH} = 0.50$ for the signal. The results of these fits are shown in Table 8.7. In the "null" HERWIG samples, we see fluctuations both above and below 0.50, but both of these fluctuations are less statistically significant than the one seen in the PYTHIA sample. In order to account for any possible effect due to the difference between event generators seen in these null samples, we include a systematic uncertainty of 0.0600, the size of the deviation of the PYTHIA sample from 0.50.

8.9 Total Systematic Uncertainty

All of the systematic uncertainties that we have calculated are listed in Table 8.8. To determine the total systematic uncertainty, these uncertainties are added in quadrature, yielding $\sigma_{syst} = 0.0832$.

Systematic	Uncertainty
Generator Dependence	0.0600
JES	0.0423
ISR/FSR	0.0295
Background Shape	0.0225
Color Reconnection	0.0087
PDF	0.0071
Parton Shower	0.0056
Background Size	0.0015
Total Uncertainty	0.0832

Table 8.8. Summary of Systematic Uncertainties

CHAPTER 9

Measurement Result

Having studied our sensitivity and evaluated our systematic uncertainties, we now can perform our measurement in 4.3 $\text{fb}^{-1}$of CDF data. We have a total of 1,001 candidate $\text{t}\bar{\text{t}}$ lepton plus jets events in this dataset, all of which pass the selection described in Chapter 3. Following the methods described in Chapter 4, we predict 215 background events, so we expect the remaining 786 events to be $\text{t}\bar{\text{t}}$ signal events. Having constructed same helicity and opposite helicity templates, as well as a predicted background template, we now measure the opposite helicity fraction in the data using the method described in Chapter 7.

9.1 One-Dimensional Fits

Although our final result will be found using the two-dimensional fit of $\cos(\theta_\ell)\cdot\cos(\theta_d)$ vs. $\cos(\theta_\ell)\cdot\cos(\theta_b)$, we first consider the two bilinears separately in one-dimensional fits. This allows us to test our experimental procedure in the data and compare to the expected sensitivities estimated previously. Performing a fit to the one-dimensional distribution of $\cos(\theta_\ell)\cdot\cos(\theta_d)$, binned into 20 bins, we find a result of

$$F_{OH} = 0.59 \pm 0.29 \quad ,$$

where the uncertainty is due to statistics only. The error matches our expected sensitivity of 0.29.

We now perform our fit using the one-dimensional distribution of $\cos(\theta_\ell)\cdot\cos(\theta_b)$,

again binned into 20 bins. The result of this fit is

$$F_{OH} = 1.02 \pm 0.35 \quad ,$$

where again the uncertainty is statistical only, and it is close to our expected uncertainty of 0.36. For this distribution, the data prefers an unphysical value for F_{OH} in which the fraction of $t\bar{t}$ events where the quarks have opposite helicities is larger than 100%, although F_{OH} is consistent with 1.0 within the statistical uncertainty.

9.2 Final Two-Dimensional Fit

Our final result comes from the two-dimensional fit where one axis consists of the bilinear $\cos(\theta_\ell)\cdot\cos(\theta_d)$ and the other axis contains $\cos(\theta_\ell)\cdot\cos(\theta_b)$. We use histograms with 10 bins along each axis, for a total of $10 \cdot 10 = 100$ bins. Our expected sensitivity using this two-dimensional distribution is a statistical uncertainty of 0.23. Including our evaluated systematic uncertainty of 0.08, we find the opposite helicity fraction to be

$$F_{OH} = 0.80 \pm 0.25_{\text{stat}} \pm 0.08_{\text{syst}} \quad .$$

This first measurement in the lepton plus jets channel agrees well with the theoretical prediction $F_{OH} = 0.70$ [8, 11], within our large statistical uncertainty.

We can convert this to the spin correlation coefficient κ using $\kappa = 2 \cdot F_{OH} - 1$. This yields a spin correlation coefficient, measured in the helicity basis, of

$$\kappa = 0.60 \pm 0.50_{\text{stat}} \pm 0.16_{\text{syst}} \quad .$$

In Figures 9.1 and 9.2 we see the results of our two-dimensional fit, projected along the $\cos(\theta_\ell) \cdot \cos(\theta_d)$ and $\cos(\theta_\ell) \cdot \cos(\theta_b)$ axes respectively. The data points are compared to the sum of the background, same helicity, and opposite helicity

templates, with F_{OH} being determined by our fit result, and there is seen to be good agreement between the data and our measured result.

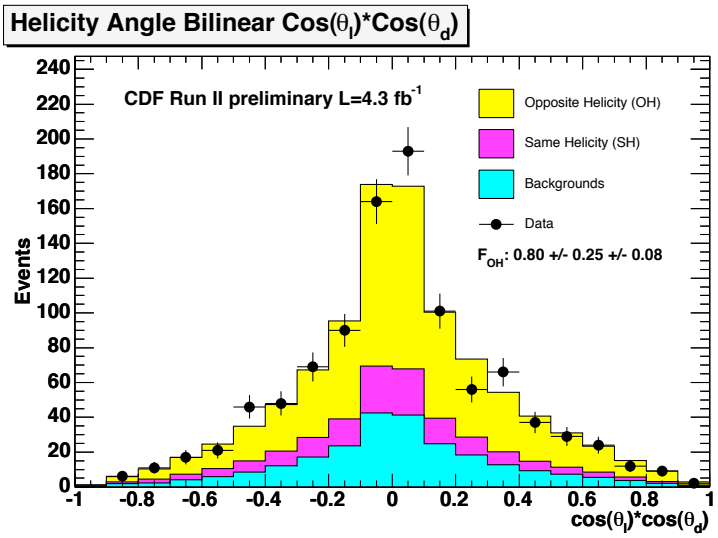

Figure 9.1. Distribution of $\cos(\theta_\ell) \cdot \cos(\theta_d)$ in Data Compared to Final Fit Result

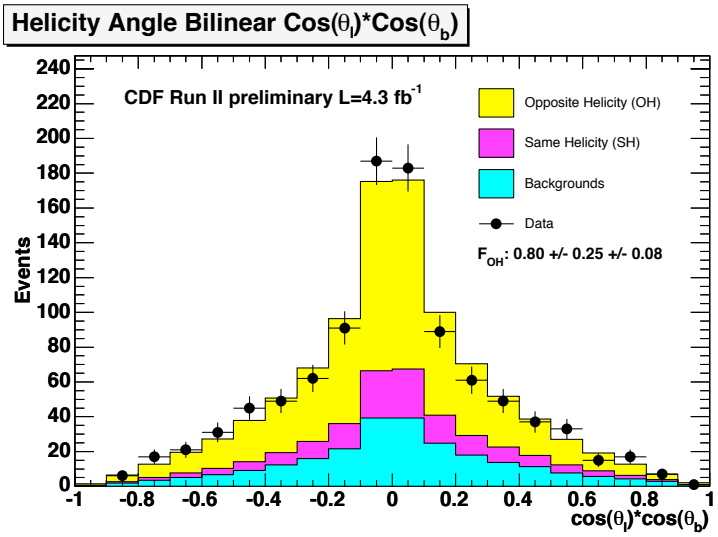

Figure 9.2. Distribution of $\cos(\theta_\ell) \cdot \cos(\theta_b)$ in Data Compared to Final Fit Result

Currently, our results agree within 1σ with the theoretical predictions for F_{OH} and κ, but our statistical uncertainty is large and we are also within approximately 1σ of a null result, $F_{OH} = 0.5$ and $\kappa = 0.0$. The large $t\bar{t}$ datasets that will be produced at the LHC will provide enough statistics to measure these values with precision, but

until those datasets start to become available, the Tevatron continues to produce data quickly and efficiently. By taking advantage of this new data, taking steps to improve the measurement method, and combining with results in other decay channels, it may be possible for the Tevatron to produce strong evidence for the existence of the $t\bar{t}$ spin correlation before large amounts of LHC data are ready to be analyzed.

CHAPTER 10

Conclusion

We have performed a measurement of $t\bar{t}$ helicity fractions and the $t\bar{t}$ spin correlation in data corresponding to an integrated luminosity of 4.3 fb^{-1}, collected during $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. This is the first time such a measurement has been made in the lepton plus jets decay channel. Taking advantage of the large statistics available and the ability to fully reconstruct events in this decay channel, this measurement of the spin correlation is the most precise yet made.

The fraction of top quark pairs produced with opposite helicities is measured to be

$$F_{OH} = 0.80 \pm 0.25_{\text{stat}} \pm 0.08_{\text{syst}}$$

This is consistent with the theoretical prediction of $F_{OH} = 0.70$ at the 1σ level. Converting this result using $\kappa = 2 \cdot F_{OH} - 1$, we find the value of the spin correlation coefficient in the helicity basis to be

$$\kappa = 0.60 \pm 0.50_{\text{stat}} \pm 0.16_{\text{syst}}$$

This measurement is currently limited by the statistics of the data sample. The amount of data available is expected to at least double to 10 fb^{-1} by the end of the current Tevatron run, so if the uncertainty scales as $\sqrt{N}$, the statistical uncertainty on F_{OH} would be reduced to 0.18. With additional improvements and through combination with measurements in other channels, it is possible that by the end of the

Tevatron lifetime there could be evidence for the existence of spin correlations at the level of a 3σ deviation from a null result. The dominant $t\bar{t}$ production mechanism at the LHC will be different, causing a shift in the expected central value, but the larger datasets that will be produced there will also offer an opportunity for many precise measurements of the spin correlation, and the method developed here will be useful in those analyses.

BIBLIOGRAPHY

[1] F. Abe *et al.*, "Observation of the Top Quark in $p\bar{p}$ Collisions with the CDF Detector at Fermilab", Phys. Rev. Lett. **74**, 2626 (1995).

[2] Tevatron Electroweak Working Group, "Combination of CDF and D0 Results on the Mass of the Top Quark", arXiv:hep-ex/0903.2503v1 (2009).

[3] T. Aaltonen *et al.* (CDF Collaboration), CDF Conference Note 9913 (2009).

[4] M. Beneke *et al.*, "Top Quark Physics", CERN-TH-2000-100 (2000).

[5] M. Cacciari *et al.*, JHEP **0404**, 68 (2004).

[6] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 092002 (2009).

[7] S. Edielman *et al.* "Review of Particle Physics", Phys. Lett. B **592**, 1 (2004).

[8] T. Stelzer and S. Willenbrock, Phys. Lett. B **374**, 169 (1996).

[9] W. Pauli, "The Connection Between Spin and Statistics", Phys. Rev. **58**, 716 (1940).

[10] G. Mahlon and S. Parke, Phys. Rev. D **53**, 4886 (1996).

[11] G. Mahlon and S. Parke, Phys. Lett. B **411**, 173 (1997).

[12] W. Bernreuther *et al.*, "Investigation of Top Quark Spin Correlations at Hadron Colliders", CERN-PH-TH-2004-206 (2004).

[13] T. Aaltonen *et al.* (CDF Collaboration), CDF Conference Note 9824 (2009).

[14] The D0 Collaboration, D0 Conference Note 5950 (2009). http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/TOP/T84/

[15] M. Arai, N. Okada, K. Smolek, V. Simak, "Top Quark Spin Correlations in the Randall-Sundrum Model at the CERN Large Hadron Collider", arXiv:hep-ph/0701155 (2007).

[16] T. Aaltonen *et al.* (CDF Collaboration), CDF Conference Note 9724 (2009).

[17] D.W. Jung *et al.*, "Model Independent Analysis of the Forward-Backward Asymmetry of Top Quark Production at the Tevatron", arXiv:hep-ph/0912.1105v1 (2009).

[18] D. McGinnis *et al.*, "The Run II Handbook", Fermilab National Accelerator Laboratory (2000). http://www-ad.fnal.gov/runII

[19] D. Amidei *et al.*, "The CDF II Detector Technical Design Report", FERMILAB-PUB-96-390-E, Fermilab National Accelerator Laboratory (1996).

[20] The CDF Collaboration, CDF Detector Pictures (2005). http://www-cdf.fnal.gov/events/detector_pictures.html

[21] T. Schwarz, "Measurement of the Front Back Asymmetry in Top-Antitop Quark Pairs Produced in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV" (Doctoral Dissertation), 2006.

[22] G. Blazey *et al.*, "Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop", arXiv:hep-ex/0005012 (2000).

[23] T. Affolder *et al.*, Phys. Rev. D **64**, 032002 (2001).

[24] T. Aaltonen *et al.* (CDF Collaboration), "Observation of Single Top Quark Production and Measurement of $|V_{tb}|$ with CDF", arXiv:hep-ex/1004.1181 (2010).

[25] T. Sjostrand, L. Lonnblad, and S. Mrenna, "PYTHIA 6.2 Physics and Manual", Comput. Phys. Commun. **101**, 232 (1997).

[26] G. Corcella *et al.*, "HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons", J. High Energy Phys. **01**, 010 (2001).

[27] M. Mangano *et al.*, "ALPGEN, A Generator for Hard Multi-parton Processes in Hadronic Collisions", J. High Energy Phys. **07**, 001 (2003).

[28] F. Maltoni and T. Stelzer, "MadEvent: Automatic Event Generation with MadGraph", J. High Energy Phys. **02**, 027 (2003).

[29] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 032003 (2006).

[30] F. James, "Minuit: Function Minimization and Error Analysis Reference Manual", Computing and Networks Division, CERN (1998). http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html

[31] E. Guillian, "Top Quark Decay Kinematics in Full Reconstructed $t\bar{t}$ Events in the e or $\mu + \not\!\!E_T + \geq 4$ Jet Decay Channel" (Doctoral Dissertation), 1999.

[32] K. Hagiwara *et al.* (Particle Data Group), Phys. Rev. D **66**, 010001 (2002).

[33] D. Acosta *et al.*, Phys. Rev. D **72**, 052003 (2005).