ABSTRACT

Measurement of the Single Top Quark Cross Section in the Lepton Plus Jets Final State in Proton-Antiproton Collisions at a Center of Mass Energy of 1.96 TeV Using the CDF II Detector

Zhenbin Wu, Ph.D.

Advisor: Jay R. Dittmann, Ph.D.

We present a measurement of the single top quark cross section in the lepton plus jets final state using an integrated luminosity corresponding to 7.5 fb^{-1} of $p\bar{p}$ collision data collected by the Collider Detector at Fermilab. The single top candidate events are identified by the signature of a charged lepton, large missing transverse energy, and two or three jets with at least one of them identified as originating from a bottom quark. A new Monte Carlo generator POWHEG is used to model the single top quark production processes, which include s-channel, t-channel, and Wt-channel. A neural network multivariate method is exploited to discriminate the single top quark signal from the comparatively large backgrounds. We measure a single top production cross section of $3.04^{+0.57}_{-0.53}$ (stat. + syst.) pb assuming $m_{\rm top} = 172.5$ GeV/c^2. In addition, we extract the CKM matrix element value $|V_{tb}| = 0.96 \pm 0.09$ (stat. + syst.) ± 0.05 (theory) and set a lower limit of $|V_{tb}| > 0.78$ at the 95% credibility level.

Measurement of the Single Top Quark Cross Section in the Lepton Plus Jets Final State in Proton-Antiproton Collisions at a Center of Mass Energy of 1.96 TeV Using the CDF II Detector

by

Zhenbin Wu, B.S.

A Dissertation

Approved by the Department of Physics

_____

Gregory A. Benesh, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

_____

Jay R. Dittmann, Ph.D., Chairperson

_____

Kenichi Hatakeyama, Ph.D.

_____

Jonatan Lenells, Ph.D.

_____

Anzhong Wang, Ph.D.

_____

Zhenrong Zhang, Ph.D.

Accepted by the Graduate School
December 2012

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am ever in debt to my advisor, Dr. Jay R. Dittmann, for leading and supporting me along the path of experimental high energy physics. I am grateful for his consistent encouragement, patience, and thoughtfulness. Besides his involvement in my analysis, he also cares about my personal life and participated in the most important moments in my life.

I would like to thank Dr. Jan Lück and Manfredi Ronzani. I could not have finished this analysis without their help. Dr. Jan Lück helped me to start this analysis while I was searching for analysis topics. He taught and guided me through this analysis. I need to thank Manfredi Ronzani for working on this analysis together with me. It was a great pleasure to work with them and I enjoyed many productive discussions with them.

I want express my gratitude to Dr. Tom Junk. He is an expert in the single top analysis and much more. His deep knowledge about high energy physics really inspires me. I want to thank the WH/Single top analysis group at CDF: Dr. Jesus Vizan, Dr. Federico Sforza, Dr. Adrian Buzatu, Dr. Giorgio Chiarelli, Dr. Craig Group, Dr. Dominic Hirschbuehl, Dr. Homer Wolfe, and Dr. Weiming Yao. It was a pleasant experience to work within the group. I can't summarize how much I learned from them. I am proud to have worked with them on the WHAM package development.

I am grateful to other members of the Baylor High Energy Physics Group: Dr. Kenichi Hatakeyama, Dr. Azeddine Kasmi, Dr. Martin Frank, Dr. Sam Hewamanage, and Dr. Karen Bland. They never hesitate to share their knowledge with me. I would like to thank Dr. Karen Bland for her patient and selfless help with reviewing my thesis. I thank the professors that served on my dissertation committee:

Dr. Jay R. Dittmann, Dr. Kenichi Hatakeyama, Dr. Jonatan Lenells, Dr. Anzhong Wang, and Dr. Zhenrong Zhang.

The Physics Department of Baylor University gave me the opportunity to study in the USA. When I first came to Baylor, the Physics Department funded my work with a teaching assistantship. The United States Department of Energy funded my research when I started the analysis. I am very grateful for this funding support. I also want to thank the Tevatron group and the CDF Collaboration for providing the data and the service for this analysis.

I would like to thank my friends, Dr. Tingjun Yang, Qiuguang Liu, Yongqing Huang, and so many more that I can't list all their names. I can't imagine my life without their friendship.

Last but not least, I thank my family for their support and encouragement. I thank my parents for their support for all of my life. I am so thankful for my wife, Wenmei Guo, for her consistent encouragement and her love that keeps me warm all the time. At last, I thank my son Gavin Wu, who was recently born on Oct. 25th, for starting a new phase in my life.

DEDICATION

This work is dedicated to my wife, who is the love of my life.

CHAPTER ONE

Introduction

The standard model of particle physics is a successful theory that describes the interactions among fundamental particles. The top quark is the heaviest elementary quark in the standard model and is predominately produced together with its antiparticle partner, an antitop quark, in high energy particle collisions. The standard model predicts a rarer top quark process, called single top quark production, in which only one top quark is produced. Studies of single top quark production are important for understanding the standard model. However, an analysis of single top quark production is experimentally challenging. The rate of single top quark production (the signal) is very small compared to other standard model processes that have similar features in particle detectors (the backgrounds). Multivariate techniques are therefore used to discriminate between the single top signal and these backgrounds.

This thesis presents the measurement of the single top quark cross section in proton-antiproton collisions at the Fermilab Tevatron using data corresponding to an integrated luminosity of 7.5 fb^{-1}, collected by the CDF II detector. A brief introduction of the standard model and the top quark is provided in this chapter. Chapter 2 describes the experimental apparatus, which consists of the Fermilab Tevatron and the CDF II detector. In Chapter 3, the reconstruction and identification of detected particles are presented. Chapter 4 discusses the simulation of signal and background events using Monte Carlo techniques. The selection of candidate events for this analysis is described in Chapter 5, while Chapter 6 describes the prediction of signal and background rates corresponding to the 7.5 fb^{-1} dataset. Chapter 7 discusses the neural network multivariate technique exploited in this analysis for signal and background discrimination. The various systematic uncertainties consid-

ered in this analysis are presented in Chapter 8, and Chapter 9 presents the results of this analysis. The thesis concludes in Chapter 10, which provides a summary of the experimental results and their impact on the field of particle physics.

1.1 The Standard Model of Particle Physics

1.1.1 The Standard Model

Nature can be described by particle physics in terms of four distinct forces, characterized by different distance ranges and effects, listed in Table 1.1.

Table 1.1. The four fundamental forces in nature and some of their characteristics.

Force	Couples with	Effect	Range
strong	color charge	binds quarks and gluons	10^{-15} m
weak	weak charge	radioactive decay	10^{-17} m
electromagnetic	electric charge	interaction between charged particles	infinite
gravitation	mass	attraction of masses	infinite

Since the 1960s, the desire to develop a theory that unifies these forces led to a successful theoretical framework called the "Standard Model" (SM). The SM is a unified description of the strong, weak, and electromagnetic forces by quantum field theory with the introduction of a local symmetry SU(3) $\times$ SU(2) $\times$ U(1) to the Lagrangian. The Lagrangian is a mathematical description of a physical system in quantum mechanics. So far, the gravitational force is still not included in the SM framework; it is best described by the theory of general relativity. Since gravity is very weak (10^{-33} smaller in magnitude than the weak force), its effect is usually negligible in experimental measurements at the Tevatron.

The SM features a small set of fundamental particles, listed in Figure 1.1. A set of quantum numbers is introduced to describe the state of these elementary particles, e.g. electric charge, color charge, and spin. The elementary particles are divided into two groups based on their spin: *fermions* and *bosons*.

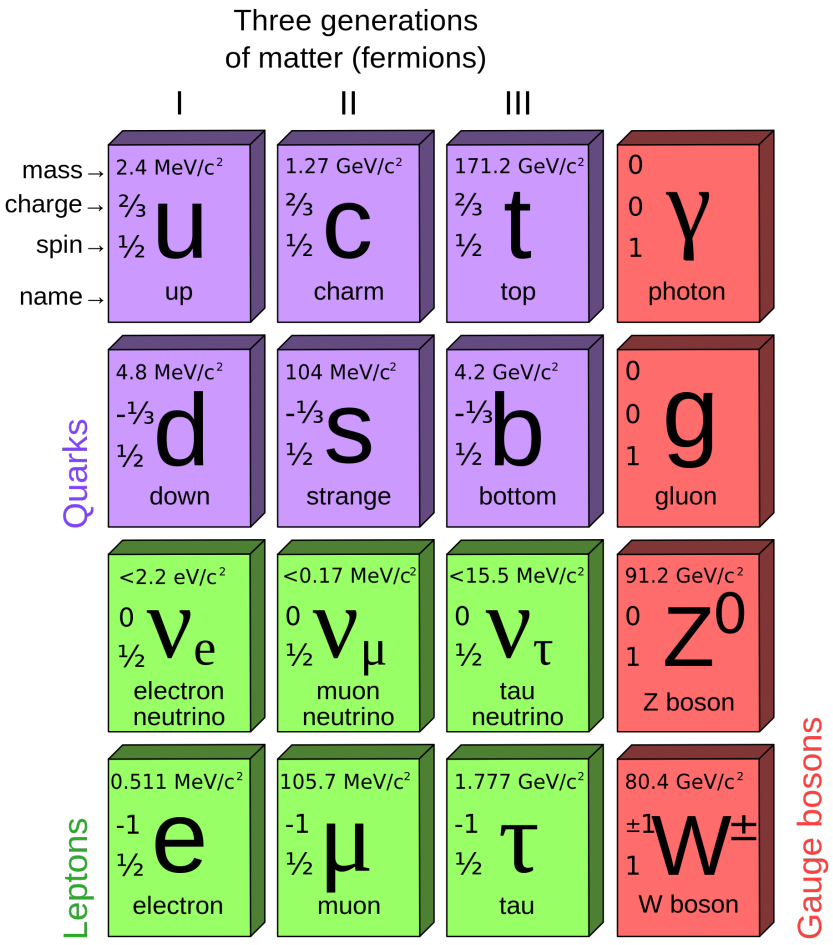


Figure 1.1: Table of elementary particles in the SM, with three generations of fermions on the left and gauge bosons in the rightmost column [1]. The SM also predicts the Higgs boson, which is not shown in this table.

1.1.1.1. *Fermions.* Particles with half-integer spin ($\frac{1}{2}$) follow Fermi-Dirac statistics and are called fermions. Fermions are ordered in three generations in the SM as shown in Figure 1.1. Only particles of the first generation can form stable matter, while the others can only be observed in high energy interactions. There are two types of fermions: *quarks* and *leptons*. Quarks carry fractional electric charge ($-\frac{1}{3}$ or $+\frac{2}{3}$) and interact by the strong, weak, and electromagnetic forces. Leptons carry integer electric charge (-1 or $+1$) and interact only by the weak and the electromagnetic forces.

There are six types of quarks, known as *flavors*: up (u), down (d), strange (s), charm (c), bottom (b), and top (t). The quarks are divided into two groups based on their masses: *light-flavor* quarks (u, d, and s) and *heavy-flavor* quarks (c, b, and t). Each quark has a corresponding antiquark, which is identical in all regards expect for its opposite electric charge. Due to the Pauli exclusion principle, an additional quantum number called the color charge is carried by quarks, which is not shown in Figure 1.1. There are three types of color: red, green, and blue, which are mediated through *gluons* via the strong interaction. Quarks cannot be observed as free particles. With the exception of the top quark, they are confined in bound states with integer electric charge called *hadrons*. Hadrons are classified into two families: *baryons*, which are made of three quarks, and *mesons*, which are made of one quark and one antiquark. Due to its enormous mass, the top quark has an extremely short lifetime. Thus, the top quark does not have time to form a hadron and immediately decays once it is produced.

Leptons interact only by the electromagnetic and weak forces. Thus, leptons have no color charge. There are six types of leptons that form three generations: the electron (e) and the electron neutrino (ν_e), the muon (μ) and the muon neutrino (ν_μ), and the tau (τ) and the tau neutrino (ν_τ). The charged leptons (electron, muon, and tau) have integer electric charge and interact by the electromagnetic and

weak forces. By contrast, neutrinos have no charge and thus only interact by the weak force. Therefore, neutrinos are able to travel long distances through matter. The SM originally assumed neutrinos to be massless. Experiments, however, have provided compelling evidence that neutrinos have non-zero mass [2]. All leptons have a corresponding antilepton with opposite charge. Antineutrinos, associated with neutrinos, also have neutral electric charge and half-integer spin. Whether the neutrino and its corresponding antineutrino are identical particles remains to be resolved. Due to the similarity of matter and antimatter particles, a reference to a matter particle is usually inclusive, e.g. "electron" usually refers to either an electron or a positron (the positively charged antiparticle of an electron).

1.1.1.2. *Gauge bosons.* The SM describes the strong, weak, and electromagnetic forces with gauge theories. The Lagrangian is invariant under local transformations in a gauge theory. With the introduction of local symmetry SU(3) $\times$ SU(2) $\times$ U(1) into the Lagrangian, the SM introduces additional gauge fields to compensate for the local transformations. In quantum field theory, the excitation of the gauge fields represents particles that transmit the forces, namely the gauge bosons. Bosons have integer spin and follow Bose-Einstein statistics.

The electromagnetic interactions are described by the theory of Quantum Electrodynamics (QED), based on the U(1) symmetry. The electromagnetic force is mediated by *photons* (γ) exchanged between electrically charged particles. From the invariance of the electrostatic potential, QED predicts that the photon is a massless, chargeless boson. The fact that the photon is massless leads to the infinite range of the electromagnetic force.

The weak force is mediated by massive gauge bosons (W^+, W^-, and Z^0) exchanged between particles (all quarks and leptons) of different flavors. The electrically charged $W^\pm$ bosons are referred to as the W boson and the neutral Z^0 boson

is referred to as the Z boson. In general, massive particles have both *left-handed* components (with spin oriented opposite to the direction of motion) and *right-handed* components (with spin oriented along the direction of motion) as illustrated in Figure 1.2. The W boson exclusively couples to the left-handed (right-handed) component of a fermion (an anti-fermion). The Z boson interacts with either left- or right-handed particles with different strengths. Due to the large mass of the W and Z bosons, the weak interaction has a very short range. Because the W boson carries electric charge, it must couple to the photon, which leads to the electroweak gauge theory that combines the weak and the electromagnetic forces in the SM [3–5].

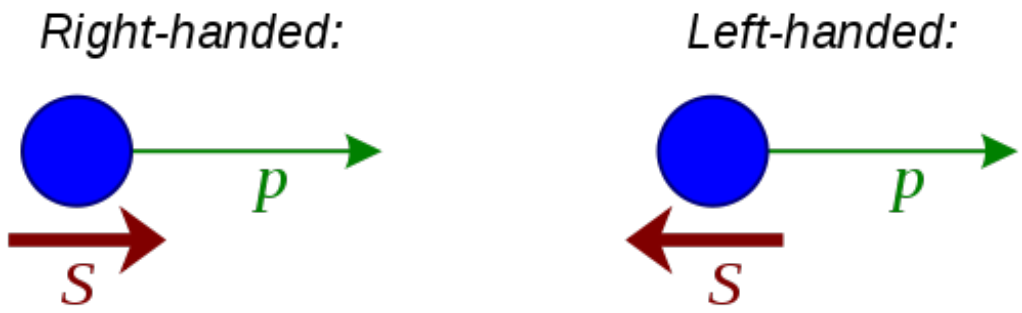


Figure 1.2: A particle is right-handed or left-handed based on the alignment of its spin (S) and momentum (p).

Gluons mediate the strong interaction; they are exchanged between color-charged particles (quarks). Gluons are massless and labeled by a combination of color and anticolor charge. Hence, gluons can not only couple to color-charged quarks but also to themselves, which is referred to as gluon self-interactions. All possible color combinations would lead to nine gluons. Since gluons in the color-singlet state do not exist, there are eight types of gluons. The strong interactions of colored quarks and gluons are described by the theory of Quantum Chromodynamics (QCD), corresponding to the SU(3) component of the SM local symmetry. QCD describes two peculiar properties of the strong interactions. The strength of the strong interaction decreases with increasing energy, which is termed asymptotic freedom,

allowing perturbative QCD theory in deep inelastic processes. The strength of the strong interaction increases at large distances, which is termed confinement. The binding energy between two quarks increases as they separate and it creates a new quark-antiquark pair when the energy is sufficiently large. In a proton-antiproton collision, as the resulting free quarks move away from the primary interaction, they fragment into additional partons, which immediately form colorless bound hadrons. This process is referred to as *hadronization*. This process repeats until the energy of the original energetic quark is expended, resulting in a collection of particles moving in roughly the same direction, called a *jet*.

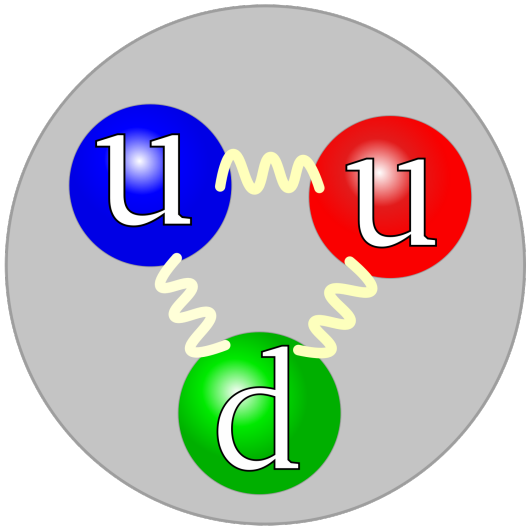


Figure 1.3: The quark structure of the proton. The wavy line represents the virtual gluons binding the valence quarks. The color assigned to each individual quark is arbitrarily chosen, but all three colors must be present within the proton.

In this analysis, we collect data from collisions between protons and antiprotons. A proton is a composite particle, consisting of two u quarks and one d quark as shown in Figure 1.3; the antiproton consists of the corresponding antiparticles. These constituents are called valence quarks and are bound by virtual (transient) gluons. The virtual gluons can split into quark-antiquark pairs, which are known as sea quarks. Quarks and gluons are also referred to as *partons*. The momentum of

the proton is shared by these valence quarks, sea quarks, and gluons (see Figure 1.4), which is described by the parton distribution functions.

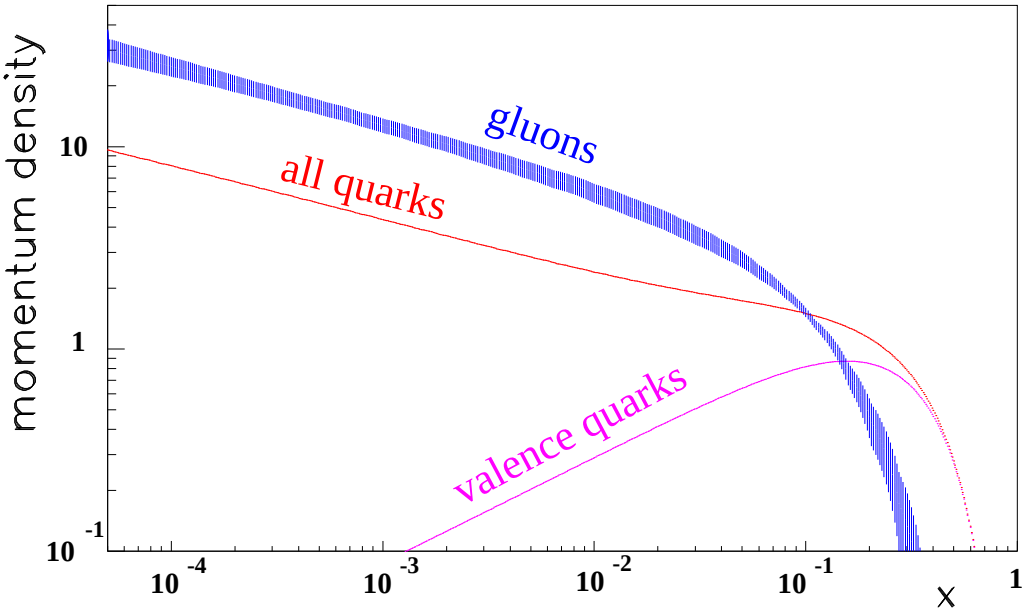


Figure 1.4: An illustration of the quark and gluon momentum densities in the proton as a function of the momentum fraction of each component in the proton (x) for $Q^2 = 20$ GeV^2 [6]. The integrated values of each component's momentum density gives the fraction of the parton momentum carried by that component within a proton. The difference between the "all quarks" curve and the "valence quarks" curve arises from the presence of sea quarks.

1.1.1.3. *Higgs boson.* If symmetry were well preserved in the SM, we would expect a symmetry among particle masses. However, the electron has a mass of 0.5 MeV/c^2, while the electron neutrino is almost massless; furthermore, the W and Z bosons are massive, while photons are massless. In the electroweak theory, the gauge theory alone predicts the W and Z bosons to be massless. To overcome this, the Higgs mechanism was introduced [7–10]. The Higgs mechanism, when applied to the electroweak theory through a complex scalar doublet field, leads to the generation of the W and Z masses but leaves the photon massless. In this process, known as spontaneous symmetry breaking, a scalar field called the Higgs field is introduced. The masses of fermions arise from the Yukawa coupling between

the fermion field and the Higgs field. The Higgs field can interact with itself, resulting in a matching particle called the Higgs boson (H). The Higgs boson is a detectable particle, providing a crucial test of the Higgs mechanism. Figure 1.5 summaries the interactions between particles described by the SM.

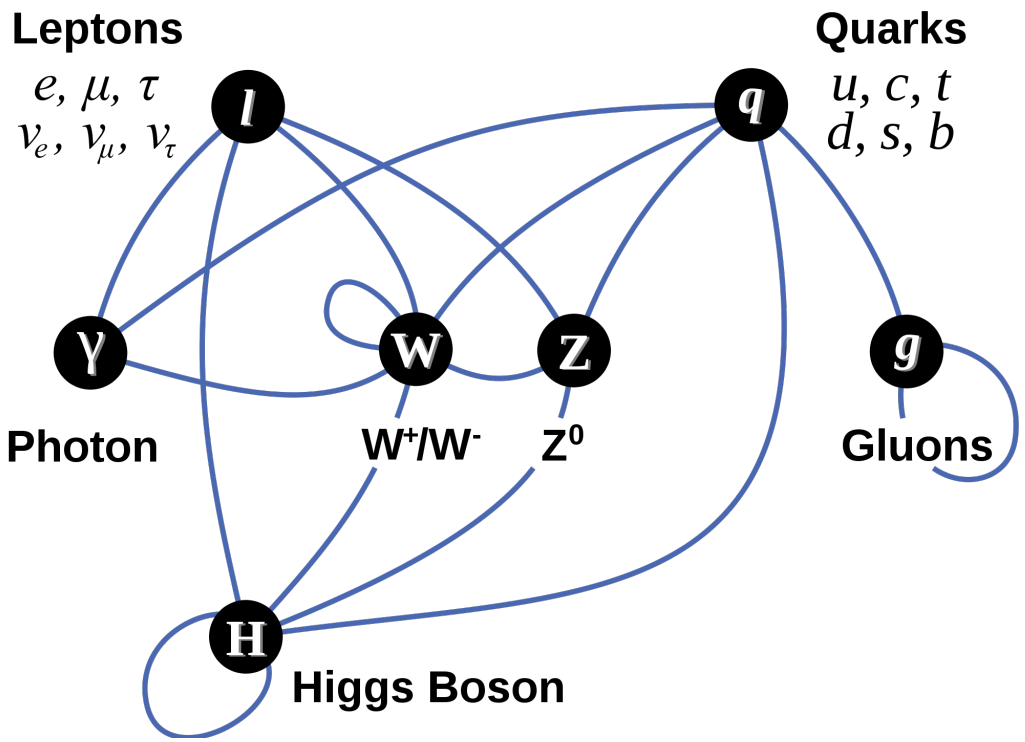


Figure 1.5: A diagram showing the interactions between particles described by the SM [11]. Each blue line signifies an interaction between the two connected particles.

Because of its important role in the SM and the challenges associated with searching for it, the Higgs boson is known in the media as the "God particle." On July 4, 2012, the CMS and the ATLAS experiments at the Large Hadron Collider (LHC) simultaneously announced the discovery of a previously unknown boson with a mass between 125 and 127 GeV/c^2, which is a highly possible candidate for the Higgs boson [12, 13].

1.1.2 Cabibbo-Kobayashi-Maskawa (CKM) Matrix

In electroweak interactions, flavor changing is possible through charged currents by the exchange of a W boson. The probability of transition from one flavor of quark to another flavor of quark is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [14, 15], a 3×3 unitary matrix. As in Eq. 1.1, the CKM matrix V operates on a vector of down-type quarks (d, s, and b), resulting in the weak interaction doublet partners of up-type quarks (d', s', and b'). $V_{q_1 q_2}$ is proportional to the probability of a transition from quark q_1 to quark q_2.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

The CKM matrix can be parameterized by three mixing angles and a CP-violating phase in the SM [15]. However, the values of the CKM matrix elements cannot be predicted by theoretical calculations and have to be measured by experiments. The most precise values of the CKM matrix elements are determined from a global fit that uses all measurements with SM constraints imposed [2]. These values are given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.000016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00026}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \tag{1.2}$$

1.1.3 Cross Section Calculation

In particle physics, the probability of an interaction is quantified as the *cross section* of the interaction. In addition to the cross section, the four momenta of the final-state particles can be calculated from the SM. The SM is a renormalizable quantum field theory, which means the quantitative features of interaction among particles can be calculated to arbitrary accuracy as a perturbative expansion.

The perturbative expansion can be conveniently represented by *Feynman diagrams*, which are graphical representations of expansion terms. A full set of Feynman rules is given in Ref. [16].

For a given interaction, a general approximation of the process can be represented by the *leading-order* (LO) term of the perturbative expansion, which can be expressed by the Feynman diagrams with the fewest number of vertices. A LO diagram that contains no closed *loop* is also referred to as a *tree-level* diagram, as shown in Figure 1.6(a). A better approximation of the process can be obtained by adding the next expansion term or terms, the *next-to-leading-order* (NLO) Feynman diagrams. NLO Feynman diagrams have at least one more vertex compared to LO diagrams and loop diagrams, resulting in a more precise, but difficult, calculation.

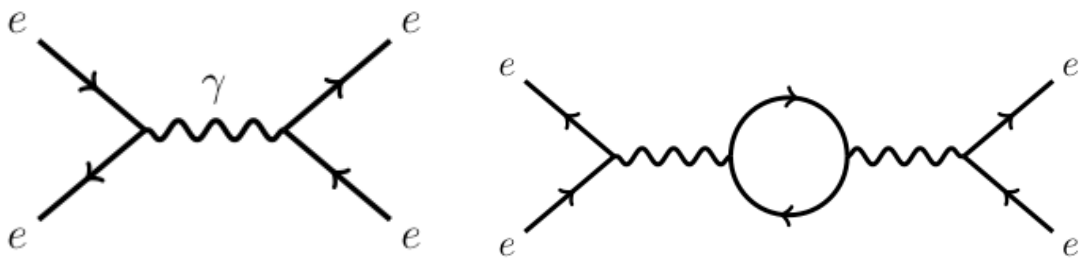


Figure 1.6. Feynman diagrams for $ee \to \gamma \to ee$ (a) at tree level and (b) with one loop.

The loop diagram requires an integral to be performed over the momenta of the particles in the loop. This integral often diverges in perturbation theory, usually called an *ultraviolet divergence* because it occurs at a large momentum scale. A process called *renormalization* is performed to deal with this problem and a *renormalization scale* is introduced to truncate the divergences. Another type of divergence, called *infrared divergence*, could appear in the calculation when the contributions of particles are at a very small momentum scale. This divergence is dealt with by introducing a *factorization scale* to truncate the integral when it reaches the nonperturbative region. Both renormalization and factorization scales are arbitrary

choices; in general they are chosen to be valid for sufficiently inclusive observables in processes. In some processes, the choice of renormalization and factorization scales affects the cross section calculation. A systematic uncertainty is then taken into account.

For a physics process with a given initial state changing into a given final state, the probability amplitude is obtained by summing up all possible related Feynman diagrams. In scattering theory, a matrix (called the *S-matrix*) is an operator mapping a free initial-state particle to a free final-state particle. The elements in the S-matrix are known as scattering amplitudes for all possible initial and final states, and usually referred to as *matrix elements*. In a QCD matrix element calculation, if a gluon becomes collinear (parallel) to another particle or the gluon momentum tends to zero ("soft"), the matrix element will have non-integrable divergences, which are referred to as the collinear and soft limits, respectively. The soft and collinear limits govern the typical structure of the events and determine which observables can be calculated in perturbative QCD. Hence, the SM calculation is performed in two steps. First, the matrix element is calculated by a perturbative QCD technique, producing a parton-level final state. Then, the resulting energetic partons fragment ("*shower*") and hadronize into a hadron-level final state, which is calculated using a parton shower technique. In actual experimental measurements, analyses place tight constraints on emissions of leptons and photons in the final state. These constraints veto a significant part of the integral over the soft-collinear divergence. In such cases, we can carry out a "resummation," which accounts for the dominant logarithmically enhanced terms from the known properties of matrix elements for multiple soft and collinear emissions.

1.2 Top Quark Production

The top quark is the heaviest quark within the SM, about 40 times heavier than the second-heaviest quark, the b quark. Due to its large mass, the top quark has an extremely short lifetime ($\tau \approx 0.5 \times 10^{-24}$ s), which is much smaller than the time scale for hadronization. As a result, the top quark decays immediately after it is produced, passing all its properties, including spin information, into its decay products. The top quark provides a unique opportunity to study a bare quark. Governed by the CKM matrix, the top quark can decay into Ws, Wd, and Wb final states. The Ws and Wd final states are heavily suppressed by the small value of their corresponding CKM matrix element, V_{ts} and V_{td}, respectively. V_{tb} is nearly one, indicating almost 100% of top quarks decay to the Wb final state. The top quark was discovered in 1995 by the CDF and D0 collaborations at the Tevatron in proton-antiproton collisions with a center-of-mass energy of 1.8 TeV [17,18]. Studies of top quarks have been performed thereafter.

1.2.1 Top Quark Pair Production

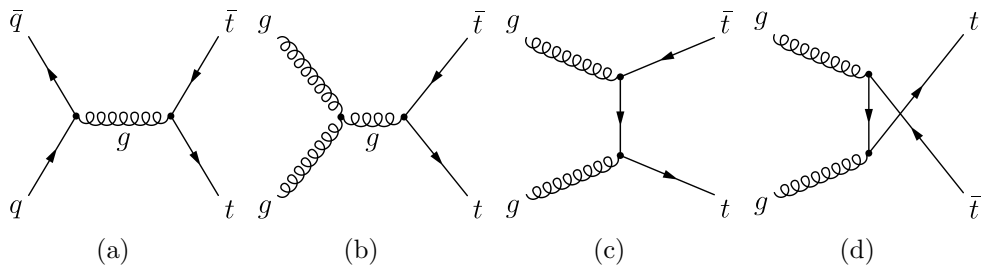


Figure 1.7: LO Feynman diagrams of top quark pair production: (a) quark-antiquark annihilation and (b)–(d) gluon fusion.

At the Tevatron, the top quark is predominately produced in a top-antitop pair ($t\bar{t}$), as shown in Figure 1.7. The next-to-next-to-leading-order (NNLO) approximation of the $t\bar{t}$ cross section at the Tevatron is 7.04 ± 0.44 pb [19], assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV/c^2. A large number of studies of the $t\bar{t}$ process

have been performed. The $t\bar{t}$ production cross section [20], the top quark charge [21], and the polarization of W bosons in top quark decay [22] have been measured precisely at CDF. The top quark mass relies on experimental measurements and is an essential input parameter in the SM. Various measurements of the top quark mass have been performed at both the CDF and D0 collaborations. A recent combination of top quark mass measurements from both the CDF and D0 collaborations yields a top quark mass of $m_{\text{top}} = 173.2 \pm 0.9$ GeV/c^2 [23]. This is the first result with an uncertainty less than 1 GeV/c^2. A large forward-backward asymmetry, which is beyond the prediction of the SM, has been observed in $t\bar{t}$ events at both the CDF [24] and D0 [25] experiments. More information on $t\bar{t}$ studies can be found in Ref. [26] from the CDF experiment and Ref. [27] from the D0 experiment.

1.2.2 Single Top Quark Production

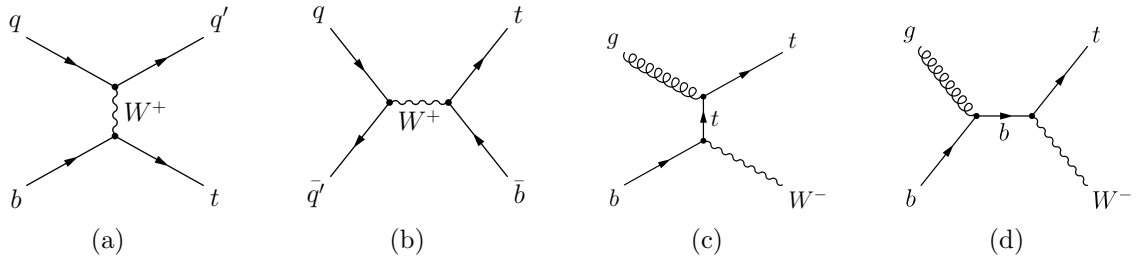


Figure 1.8: LO Feynman diagrams of single top quark production modes: (a) the t-channel, (b) the s-channel, and (c) (d) Wt-channel.

Besides $t\bar{t}$ production, the SM also predicts the production of a single top quark via the electroweak interaction. Distinguished by the mediated virtual W boson, there are three production modes: the t-channel process (Figure 1.8(a)), the s-channel process (Figure 1.8(b)) and Wt channel process (Figures 1.8(c) and 1.8(d)). At the Tevatron, the t- and s-channel processes are dominant; the Wt-channel has a very small predicted cross section compared to these dominant processes.

1.2.2.1. *t-channel production mode.* In t-channel production, a b quark scatters off a light quark by exchanging a virtual spacelike W boson, as shown in Figure 1.9. The next-to-next-to-next-to-leading-order (NNNLO) approximation of the t-channel cross section is 2.12 ± 0.22 pb at the Tevatron, assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV/c^2 [19].

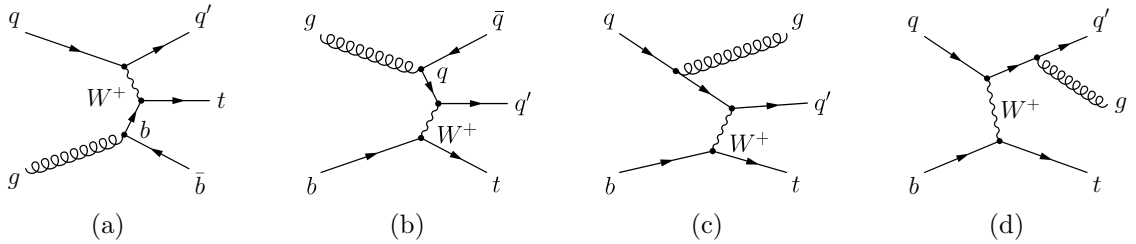


Figure 1.9: NLO Feynman diagrams of t-channel single top quark production: (a) W-gluon fusion, (b) initial-state gluon splitting, (c) initial-state gluon radiation, and (d) final-state gluon radiation.

1.2.2.2. *s-channel production mode.* In s-channel production, a virtual, timelike W boson is produced by quark fusion, as shown in Figure 1.10. At the Tevatron, the NNNLO approximation of the s-channel cross section is 1.06 ± 0.06 pb, assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV/c^2 [19].

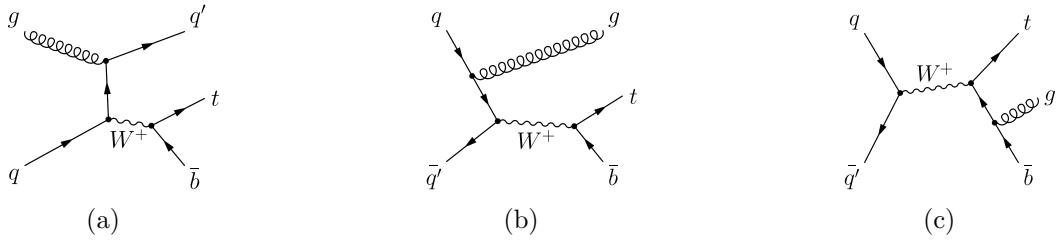


Figure 1.10: NLO Feynman diagrams of s-channel single top quark production: (a) initial-state gluon splitting, (b) initial-state gluon radiation, and (c) final-state gluon radiation.

1.2.2.3. *Wt-channel production mode.* In Wt-channel production, an on-shell W boson is produced in association with a top quark; this process is also referred to as associated production. The Wt-channel has a much smaller cross section compared with the s- and t-channels. At NNNLO accuracy, the predicted

cross section at the Tevatron is 0.22 ± 0.08 pb, assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV/c^2 [19].

(a) (b)

Figure 1.11: Feynman diagrams for Wt-channel single top production: (a) leading order and (b) next-to-leading order with initial gluon splitting.

1.2.3 Motivation for the Single Top Analysis

In the SM, the values of the CKM matrix elements rely on experimental measurements. The CKM matrix element $|V_{tb}|$ is the probability amplitude that a top quark couples to a b quark in a charged weak interaction. Assuming a 3×3 unitary CKM matrix from the SM, $|V_{tb}|$ is expected to be very close to one from the measurements of the other CKM matrix elements. The $|V_{tb}|$ value can be determined in the $t\bar{t}$ process by measuring the ratio of branching fractions

$$R = B(t \to Wb)/B(t \to Wq) = \frac{|V_{tb}|^2}{\sum_q |V_{tq}|^2} = |V_{tb}|^2 \tag{1.3}$$

where $q = b$, s, and d.

The cross section of single top quark production is directly proportional to the square of $|V_{tb}|$. Thus, a measurement of the single top quark cross section provides a unique opportunity for a direct measurement of $|V_{tb}|$ without the assumption of the unitarity of $\sum_q |V_{tq}|^2$. This allows a test of the unitarity of the CKM matrix. In the cross section measurement, we assume that the top quark decays to Wb 100% of the time. This assumption, however, does not constrain $|V_{tb}|$ to be near unity, but instead assumes $|V_{tb}|^2 \gg |V_{ts}|^2 + |V_{td}|^2$.

The number of generations of fermions in the SM is not constrained by theory. A fourth generation of fermions beyond the three established generations is still possible. If these additional fermions do exist, then the CKM matrix should be a 4×4 unitary matrix instead of 3×3 matrix. In this scenario, the 3×3 submatrix may not be necessarily unitary and $|V_{tb}|$ could be sizably smaller than one. From Ref. [28], precision electroweak constraints on possible values of $|V_{tb}|$ are presented in this extended scenario. The presence of a fourth generation could reduce the value of $|V_{tb}|$ and subsequently reduce the single top quark production cross section. Therefore a direct measurement of the single top cross section can provide a test of the existence of a fourth generation with no additional model dependence. However, it is possible to have a unitary 3×3 submatrix. A unit value of $|V_{tb}|$ does not exclude the fourth generation, so we can only find evidence of the existence of a fourth generation, but cannot exclude it.

Other new physics scenarios predict different values of σ_s and σ_t from the SM prediction [29, 30]. The SM predicts single top production via the electroweak charged current. A flavor changing neutral current (FCNC) is forbidden at tree-level and is highly suppressed at the one-loop level by the SM. New physics models allow the existence of FCNC coupling at tree level. For example, a flavor-changing Ztc coupling would manifest itself in $p\bar{p} \to t\bar{c}$ production, which would alter the measured value of σ_s or σ_t. Extensions of the electroweak symmetry of the SM predict an additional charged gauge boson W'. $W' \to t\bar{b}$ resonance production could enhance the σ_s measurement. Besides the neutral scalar Higgs boson in the SM, charged Higgs bosons are predicted by extensions of the SM Higgs sector. Charged Higgs boson exchange in the $p\bar{p} \to H^+ \to t\bar{b}$ process share the same final state as s-channel single top quark production. Several models beyond the SM predict that single top production is sensitive to charge-parity (CP) violation [31, 32].

Furthermore, single top production provides an additional channel for top quark property measurements. Since the top quark decays before hadronization, its polarization can be directly observed in the angular correlations of its decay products [33, 34]. The single top quark events provide a nearly pure polarization sample, which allows a probe of the spin projection of the top quark and the chirality of the W boson. A potential new physics model could modify the polarization fraction of the single top quarks.

1.2.4 Observation of Single Top Quark Production

In 2009, 14 years after the discovery of the top quark [18, 35], single top quark production was first observed simultaneously by the CDF [36, 37] and D0 [38] experiments with a total integrated luminosity of 3.2 fb^{-1} and 2.3 fb^{-1}, respectively. A combination of the CDF and D0 measurements of the single top production cross section yields a cross section of $2.76^{+0.58}_{-0.47}$ pb, assuming a top quark mass $m_{\text{top}} = 170$ GeV/c^2 [39].

In that observation analysis from CDF, four parallel analyses in the lepton plus jets sample were performed using four different multivariate methods to discriminate the signal from backgrounds: multivariate likelihood function [40,41], matrix element method [42, 43], artificial neural network [44, 45], and boosted decision tree [46]. Another analysis was performed in an orthogonal sample with the missing transverse energy plus jets signature [47]. All five analyses were combined by a neural network and reported an excess of 5.0 standard deviations, by which a observation of single top quark production was claimed. The combined analyses yielded a measurement of the single top production cross section of $2.3^{+0.6}_{-0.5}$ pb. In addition, they extracted the CKM matrix element value $|V_{tb}| = 0.91 \pm 0.11$ (stat. + syst.) ± 0.07 (theory) and set a lower limit of $|V_{tb}| > 0.71$ at the 95% credibility level.

CHAPTER TWO

Experimental Apparatus

The results of this thesis were obtained from data collected by the Collider Detector at Fermilab (CDF). The CDF experiment is located at the Tevatron proton-antiproton collider at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois. The Tevatron was the most powerful particle accelerator in the world until December 2009, when it was superseded by the Large Hadron Collider (LHC) at CERN. Constructed in a four-mile circumference tunnel buried 25 feet underground, the Tevatron accelerated and circulated beams of protons and antiprotons to nearly the speed of light in opposite directions. These beams collided inside two general purpose detectors, CDF and D0, at a center-of-mass energy ($\sqrt{s}$) of 1.96 TeV.

In the following sections, we will describe the $p\bar{p}$ production and acceleration process (Figure 2.1) and the CDF II detector. Though both the Tevatron and the CDF detector are no long functioning, the present tense is often used to describe them in the following descriptions.

Figure 2.1: Aerial view of the Tevatron collider complex. The Tevatron collider resides in an underground tunnel about 25 feet beneath the largest circular ring. The oval ring to the left depicts the perimeter of the Main Injector tunnel. In this view, the CDF site is located at the one o'clock position along the Tevatron ring; the D0 site is at the five o'clock position.

2.1 Tevatron Accelerator Complex

The construction of the Tevatron was completed in 1983. The first collisions were initiated in 1985 at $\sqrt{s} = 1.8$ TeV. The data collected from 1988 to 1996 corresponded to an integrated luminosity of about 130 pb^{-1}; this period of time was called Run 0 and Run I. Data from Run I led to the discovery of the top quark in 1995 by the CDF and D0 experiments. Starting in 1996, the Tevatron and both experiments were upgraded for the next significant collider physics program, called Run II. During the five-year upgrade, the Tevatron accelerator complex was improved by increasing the center-of-mass energy to 1.96 TeV and enlarging the instantaneous luminosity, which is a measure of the collision rate. Run II started on March 5, 2001 and ended on September 30, 2011 with the Tevatron delivering a total integrated luminosity of 12 fb^{-1}.

The Tevatron is a synchrotron that accelerates protons and antiprotons. Protons are abundant and available in nature while antiprotons need to be produced and stored. In order to achieve the high beam energy obtained by the Tevatron, a series of accelerators must gradually increase the energy of particles. The accelerated beams of protons and antiprotons eventually collide at the site B0 and D0, where the collision events are measured by the CDF and D0 detectors respectively. Figure 2.2 illustrates the entire complex, which is now described in more detail.

2.1.1 Proton Source and Preacceleration

The first step of the acceleration chain is the production of protons, which begins at the proton source.

2.1.1.1. *Proton source.* The protons used in the Tevatron are extracted from a tank of pure hydrogen gas. Hydrogen gas is released into a magnetron composed of two electrodes: an oval-shaped cathode and a surrounding anode. An electric spark

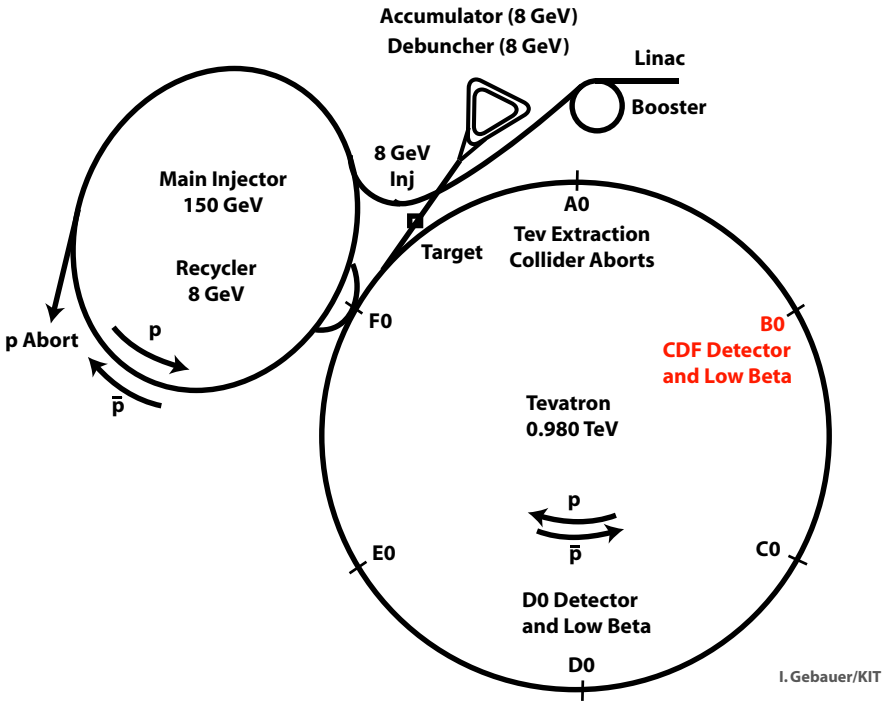


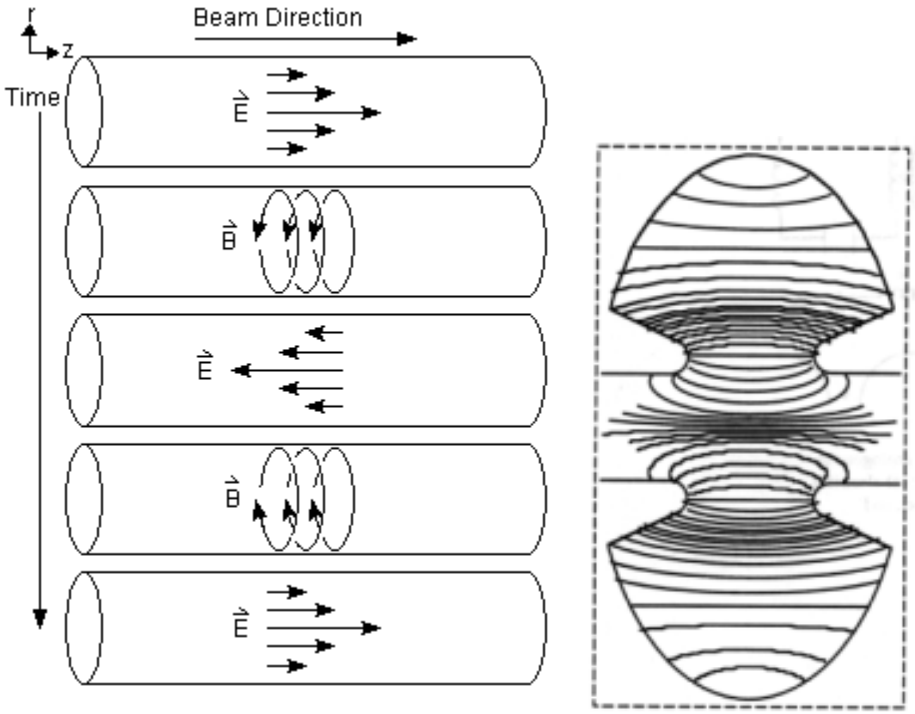
Figure 2.2. A diagram of the Tevatron accelerator complex.

ionizes the hydrogen into electrons and H^+ ions. The H^+ ions strike the cathode, which is made of cesium, and occasionally pick up two electrons to form H^- ions.

2.1.1.2. *Preaccelerator.* The H^- ions are sent to the Preaccelerator by an electrostatic extractor. The Preaccelerator accelerates the ions to an energy of 750 keV. The H^- ions are then passed through a device called the Electrostatic Chopper, which "chops" out a portion (a pulse) of the beam. Each pulse is then focused by magnets and transfered to the Linear Accelerator (Linac).

2.1.1.3. *Linac.* The Linac [48] is built in two sections. The first section consists of five drift tube cavities, which are made of copper and based on the design of the first proton linear accelerator built by Luis Alvarez [49]. The drift tubes use an alternating-current electric field (usually referred to as RF, which stands for radio frequency) as shown in Figure 2.3(a). The H^- ions are fed into the drift tube

at the correct time so that they are only exposed to the accelerating part of the RF field and shielded from the decelerating part. The RF method of acceleration, used by most accelerators in the Tevatron, causes the particles to group together into *bunches*. The drift tubes accelerate the H^- ion bunches to 116 MeV within 75 m. The second section has seven side-coupled cavities. Each cavity is made of alternating accelerating cells and coupling cells. The accelerating cells use a nose-cone field (Figure 2.3(b)) to accelerate the beam, which works like the RF field by shielding the beam from electric fields until the fields are in the direction of maximum acceleration. The coupling cells couple the energy between accelerating cells. The side-coupled cavities further accelerate the H^- ions to 400 MeV.



(a) (b)

Figure 2.3: A simplified example of the RF field in (a) a drift tube and (b) the nose-cone field used in the accelerating cells.

2.1.1.4. *Booster.* The Booster [50] is the third accelerator in the chain. It is circular in shape, with a radius of 74.47 m. It takes the H^- ions from the Linac, strips the electrons off to produce protons, and accelerates them to 8 GeV. A special set of magnets are used to bend the circulating protons in the ring. They can be also adjusted to extract the H^- ions from the Linac into a single beam that passes through carbon foils, where the electrons are stripped off. A similar set of magnets guide the resulting protons back into the ring. A set of fast kicker magnets extracts the proton beam from the Booster into the Main Injector.

2.1.2 Main Injector

From the Booster, protons are sent to the Main Injector [51], a circular synchrotron with a diameter of 1 km. It is composed by 344 dipole and 208 focusing quadrupole water-cooled magnets. It operates in several modes. It accelerates the 8 GeV protons from the Booster to 120 GeV, an energy at which the protons can then be sent to the Antiproton Source [52]. It also accelerates protons and antiprotons to 150 GeV for injection into the Tevatron. In addition, it sends protons to neutrino experiments and various other fixed-target experiments at Fermilab.

2.1.3 Antiproton Production and Storage

2.1.3.1. *Antiproton source.* Antiprotons need to be produced and stored in preparation for collisions in the Tevatron. The 120 GeV proton beams from the Main Injector are steered to a single cylinder of Inconel (a nickel-iron alloy), which is chosen as the best choice because it can withstand higher stresses caused by rapid beam heating. The Inconel fixed target is covered by a shell of beryllium to reduce target oxidation and damage. The collision of protons on the target results in a wide spray of all kinds of secondary particles. A lithium lens is placed directly behind the target to reduce the spreading of the secondary particles. The

secondaries are sent to a charge-mass spectrometer made of dipole magnets, from which 8 GeV negatively charged antiprotons are selected and sent to the Debuncher. The antiproton production is inherently inefficient; only 2 antiprotons are captured and stored for every 10^5 protons striking the target. It is the largest bottleneck for the Tevatron.

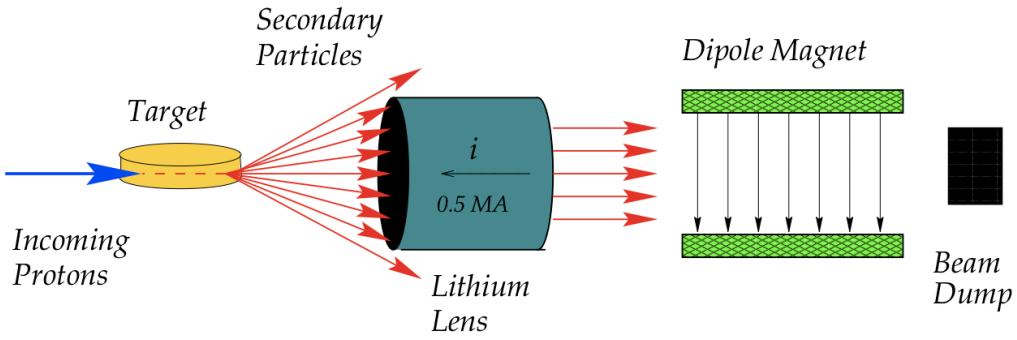


Figure 2.4: A diagram of antiproton production. The secondary particles from the collision of protons on the target are collected by the lithium lens and a dipole magnet. The 8 GeV antiprotons are sent to the Debuncher and all the rest are disposed by being sent to a beam dump.

2.1.3.2. *Debuncher.* The Debuncher is a triangular-shaped synchrotron with a mean radius of 90 m. The pulses of antiprotons from the Antiproton Source have a large spatial and momentum spread. The goal of the Debuncher is to reduce their momentum and transverse phase space for efficient transfer to the Accumulator. Using an RF manipulation called bunch rotation, it decreases the momentum spread of the antiprotons by rotating them in phase space, trading the momentum spread for time spread. This results in a continuous antiproton beam with roughly uniform momentum.

2.1.3.3. *Accumulator.* The Accumulator is also a synchrotron with a triangular shape, lying in the same tunnel as the Debuncher. As its name implies, it accumulates the antiprotons from the Debuncher. The antiprotons are "cooled" down in the Accumulator, where the transverse size and energy spread of the antipro-

ton beam is reduced without any accompanying beam loss. This is accomplished by stochastic cooling in which the momentum spread of a group of antiprotons is measured and then corrector magnets are adjusted to reduce the momentum spread of those particles. The cooling results in a denser antiproton beam, which will increase the resulting instantaneous luminosity of collisions in the Tevatron. When the Accumulator reaches its maximum optimal capacity, the antiproton bunches are injected into the Recycler.

2.1.3.4. *Recycler.* The Recycler [53] is a ring of permanent magnets along the ceiling of the Main Injector tunnel. The name comes from its originally proposed design: to recycle the antiprotons left over from the previous Tevatron store and merge them with new antiprotons. This plan was abandoned yet the name remained. The Recycler instead accepts and stores antiproton bunches from the antiproton source. The antiproton bunches are further cooled using stochastic cooling and electron cooling.

Electron cooling [54] is obtained by running a nearly monochromatic and parallel electron beam along with the antiproton beam at the same velocity. A heat exchange through Coulomb scattering between hot antiprotons and cold electrons causes the antiprotons to lose transverse momentum spread, reducing the antiproton beam size. The Recycler stores the antiproton beam at 8 GeV prior to injecting it into the Main Injector.

2.1.4 Tevatron Collider

The Tevatron tunnel has a diameter of 2 km and is buried 25 feet underground. Both protons and antiprotons circulate in the same beampipe, revolving in opposite directions due to magnetic fields of up to 4.2 T. The Tevatron uses more than 1000

superconducting magnets made of a niobium-titanium alloy, which are cooled to 4.2 K by liquid helium.

To prepare for $p\bar{p}$ collisions, 36 consecutive bunches of 150 GeV proton beams are injected from the Main Injector into the Tevatron. Then, every four out of 36 antiproton bunches are passed from the Recycler into the Main Injector, and then sent to the Tevatron with 150 GeV of energy. Both beams are accelerated to an energy of 980 GeV by eight accelerating RF cavities in the Tevatron. Both the proton and antiproton beams are split into three bunch trains; each contains 12 bunches separated in time by 396 ns. The gap between individual bunch trains, call an abort gap, is about 2.6 μs long. The beam configuration is shown in Figure 2.5. Proton and antiproton bunches are brought to collide at the B0 and D0 intersection points, where the CDF and D0 experiments are located. To initiate collisions, both beams are focused to a minimal transverse size in order to increase beam density. As collisions continue, both beams will spread out and cause the instantaneous luminosity to decrease. After a certain amount of time, the Tevatron will send the beams into abort gaps and start a new round of collisions. The period of time in which a set of beams are cycling in the Tevatron and producing collisions is called a *store*.

2.1.5 Tevatron Performance

The performance of the accelerator is often characterized by the center-of-mass energy $\sqrt{s}$ and the instantaneous luminosity L, which is defined by

$$L = n \cdot f \cdot \frac{N_p N_{\bar{p}}}{4\pi\sigma_x\sigma_y} \tag{2.1}$$

where n is the number of bunches, f is the bunch crossing frequency, N_p and $N_{\bar{p}}$ are respectively the number of protons and antiprotons per bunch, and σ_x and σ_y describe the transverse width of the beams at the interaction point. L is measured in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, typifying a particle flux. Accelerators can increase the instanta-

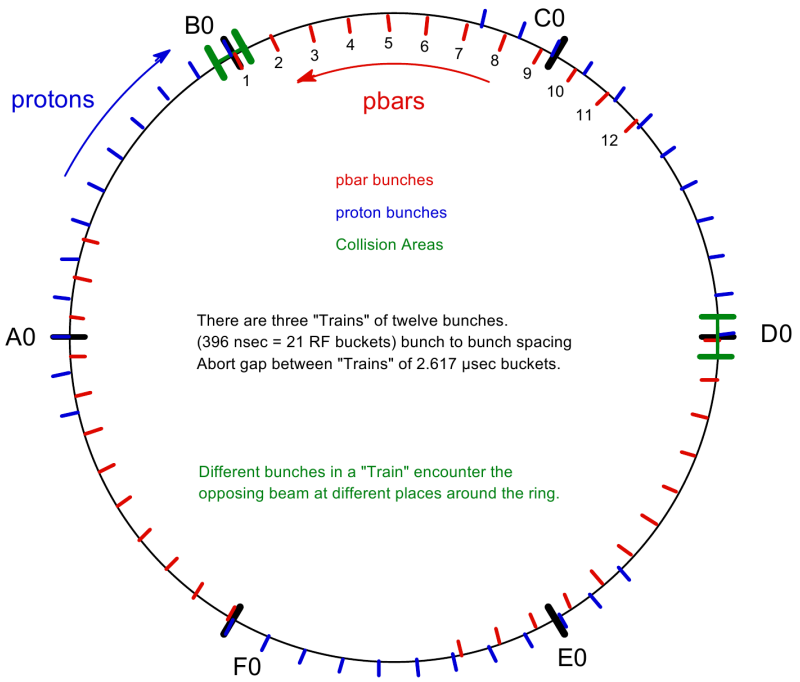


Figure 2.5: Beam structure in 36 × 36 mode. Proton bunches go clockwise and are shown as blue marks outside the ring. Antiproton bunches go counterclockwise and are shown as red marks inside the ring. The $p\bar{p}$ collisions occur every 396 ns at B0 and D0, where the CDF and D0 detectors are located, respectively.

neous luminosity by increasing the number of initial particles, increasing the collision frequency, or reducing the beam size.

The amount of data collected in experiments is measured by integrating the instantaneous luminosity over time, $\mathcal{L}_{\text{int}} = \int L\,dt$. For a given physics process with cross section σ, the number of expected events N in the collected data is $N = \sigma \cdot \mathcal{L}_{\text{int}}$. The cross section σ is given in cm^2, preferably quoted in picobarns (pb) in high energy physics, where 1 pb = 10^{-36} cm^2. The size of the collected data is normally given in units of inverse cross section, typically pb^{-1} or fb^{-1}, where 1 pb = 10^3 fb.

Until the shutdown on September 30, 2011, the Tevatron had delivered about 12 fb^{-1} (Figure 2.6). CDF recorded about 10 fb^{-1} of data, representing the final complete CDF dataset with a data-taking efficiency around 83% (Figure 2.7). This analysis uses 7.5 fb^{-1} of collected data at CDF.

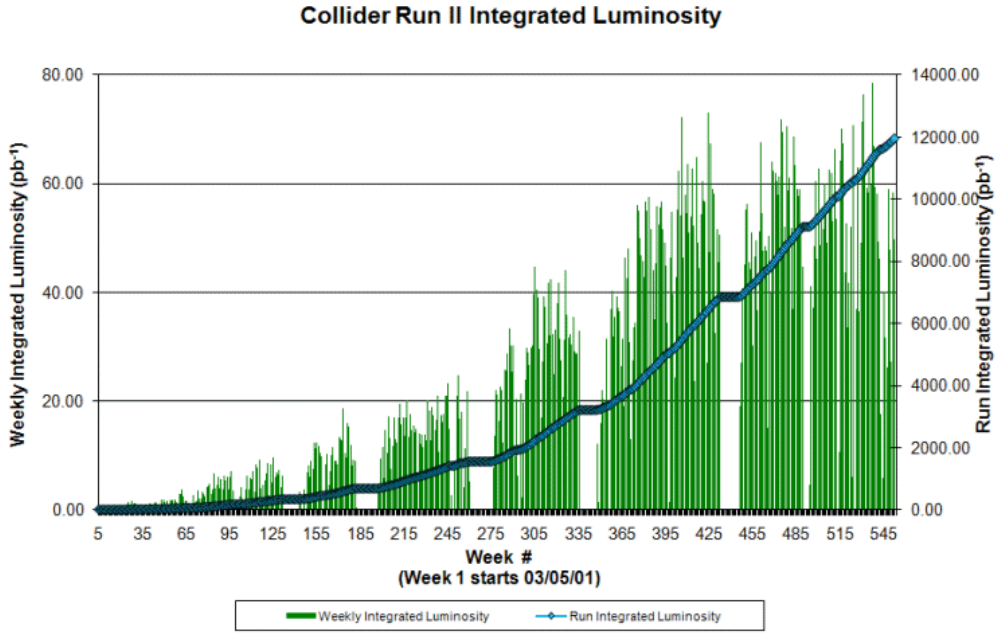


Figure 2.6: The Tevatron integrated luminosity in Run II. Empty periods are due to short Tevatron shutdowns.

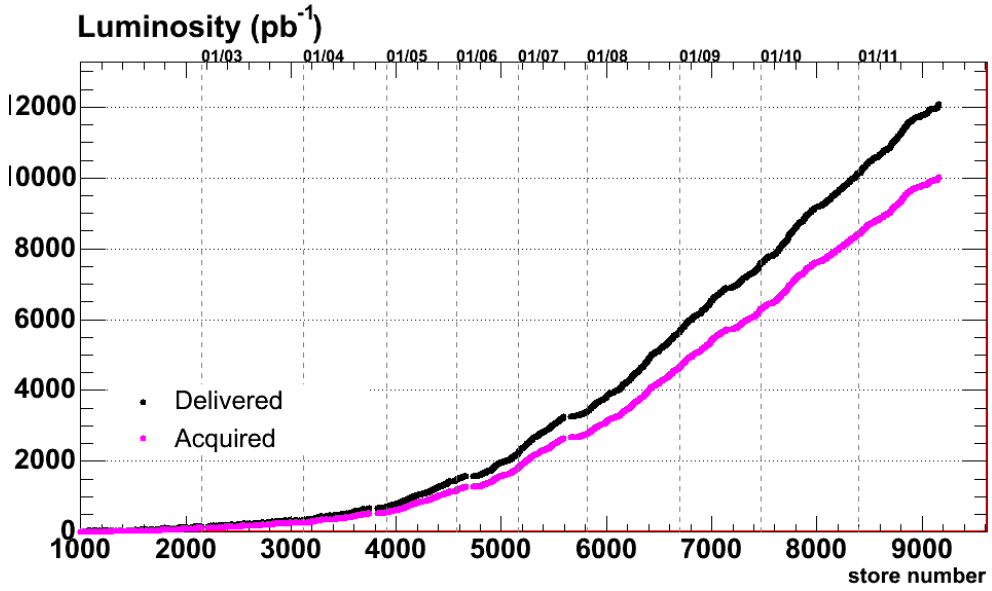


Figure 2.7: The black line is the delivered integrated luminosity by the Tevatron. The red line is the acquired integrated luminosity in Run II by the CDF detector. The final acquired integrated luminosity is about 10 fb^{-1}.

2.2 CDF Run II Experiment

The Collider Detector at Fermilab (CDF) is a general purpose solenoidal detector, which combines precision particle tracking with fast projective calorimetry and fine-grained muon detection [55]. The detector is designed to efficiently detect and measure particles produced in Tevatron collisions. It was operated and maintained by the CDF Collaboration, an international association of several hundreds of physicists from 60 universities and research laboratories. First collisions were produced and detected at CDF in October 1985, and data corresponding to over 120 pb^{-1} of integrated luminosity were recorded between 1988 and 1996. From 1996 to 2001, the detector was upgraded to accommodate an increase in the instantaneous luminosity and center of mass energy $\sqrt{s}$, and the result is referred to as the CDF II detector. On September 30, 2011, with the shutdown of the Tevatron, CDF completed 25 years of operation and became one of the longest lasting high energy physics experiments.

The CDF II detector is composed of several subdetectors, as shown in Figure 2.8. Tracking systems are contained in a superconducting solenoid. Outside the solenoid is the calorimetry system. Beyond the calorimetry system resides the muon system, which is the outermost part of the detector. With collisions occurring at each bunch crossing every 396 ns, a trigger system is used to extract the interesting physics events from an overwhelming number of other collisions. Both subdetectors and trigger system are described below. A detailed description of the CDF II detector can be found in Ref. [55].

2.2.1 CDF Coordinate System

Position and angle are expressed in a cylindrical coordinate system, accommodating the detector geometry as illustrated in Figure 2.9(a). The origin of the coordinate system is the center of the detector, which is known as B0. The $+z$ axis

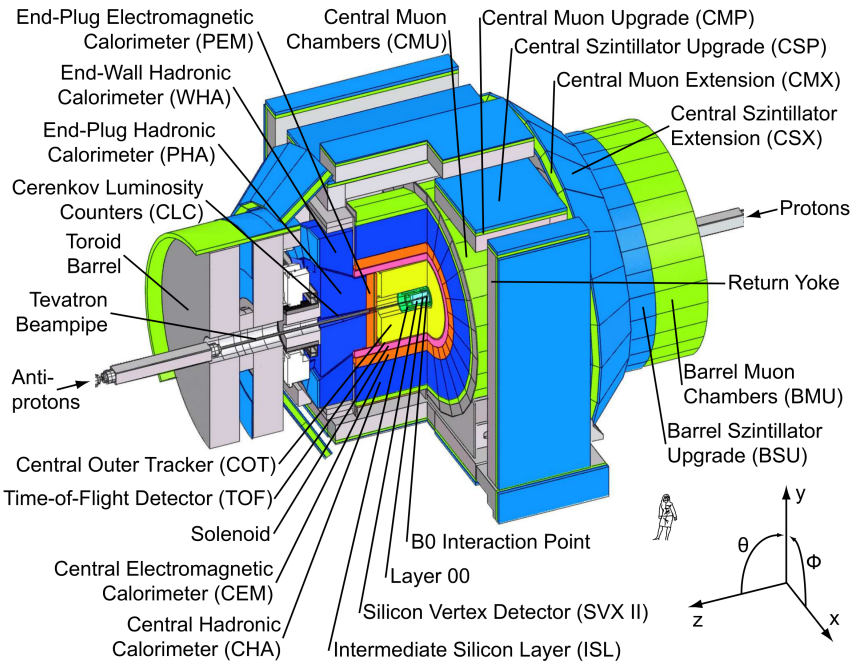


Figure 2.8: Isometric view of the main components of the CDF II detector; its coordinate system is also shown.

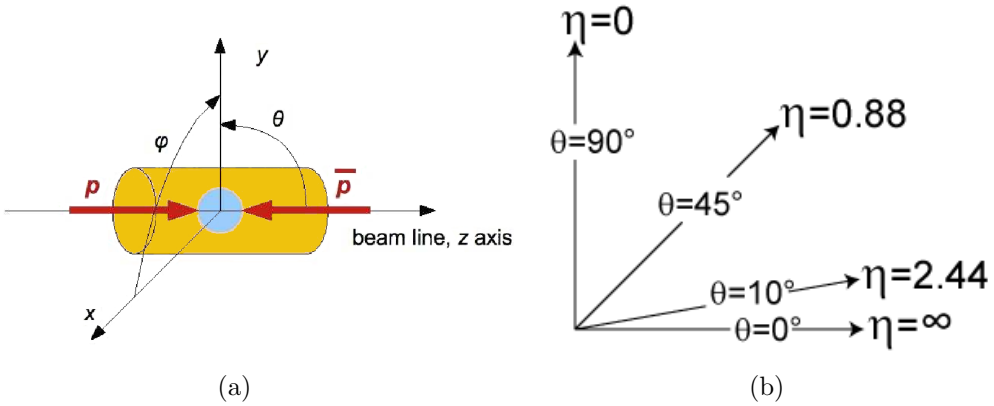
is directed along the proton beam. The azimuthal angle ϕ (around the z axis) is defined with $\phi = 0$ pointing outwards from the center of the Tevatron. The polar angle θ is defined with respect to the proton beam direction and the radius r is measured with respect to the beam axis. In this coordinate system, *longitudinal* means along the proton direction (the z axis) and *transverse* means within a perpendicular plane relative to the beam. The transverse momentum p_T and the transverse energy E_T of a particle are defined as $p_T \equiv p \sin\theta$ and $E_T \equiv E \sin\theta$, respectively.

In collider physics, instead of θ, we usually use *pseudorapidity* η defined as

$$\eta \equiv -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.2}$$

Pseudorapidity is closely related to the *rapidity* of a particle, which is defined as

$$Y \equiv \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right). \tag{2.3}$$



(a) (b)

Figure 2.9: (a) The CDF coordinate system with respect to beam pipe. (b) Values of η within the region $0^{\circ} < \theta < 90^{\circ}$.

Differences in rapidity are invariant under Lorentz transformations. Pseudorapidity is an approximation of rapidity when the energy of a particle is much larger than its mass, which is always the case at the Tevatron. Thus, a Lorentz-invariant quantity η is used to describe the polar position of a particle in the detector, independent of the reference frame (Figure 2.9(b)).

2.2.2 Tracking System

The CDF tracking system, installed close to the beam pipe, is designed for efficient and precise measurements of charged particle tracks. It is surrounded by a solenoid made of NbTi superconducting helium-cooled coils. The solenoid, 1.5 m in radius and 4.8 m in length, produces a 1.4 T magnetic field parallel to the beam-line. The trajectories of charged particles are bent in this field and the curvature of the reconstructed tracks allows a measurement of their charge and momentum. The tracking system, schematically displayed in Figure 2.10, consists of two parts: a barrel-shaped silicon microstrip system and an open-cell drift chamber (Central Outer Tracker) that surrounds the silicon system.

2.2.2.1. *Silicon detector.* The silicon detector uses silicon chips for particle detection. By applying a reverse voltage on the silicon chip, the semiconductor

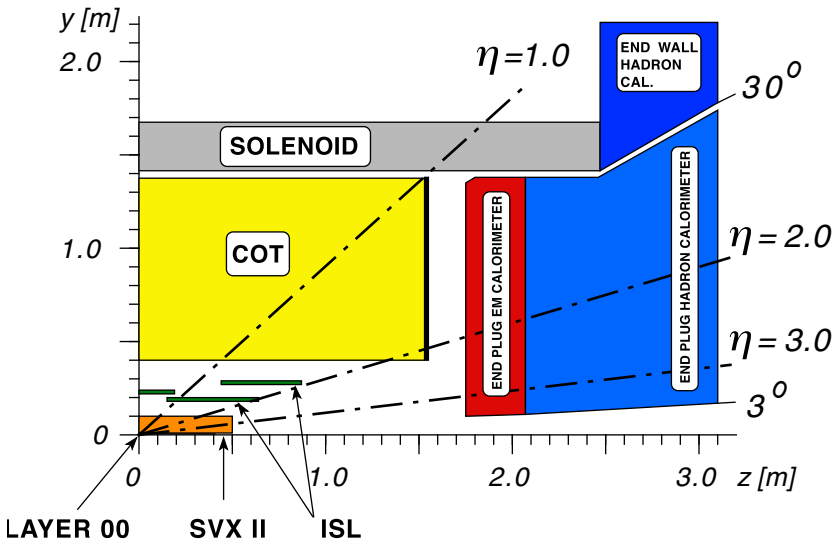
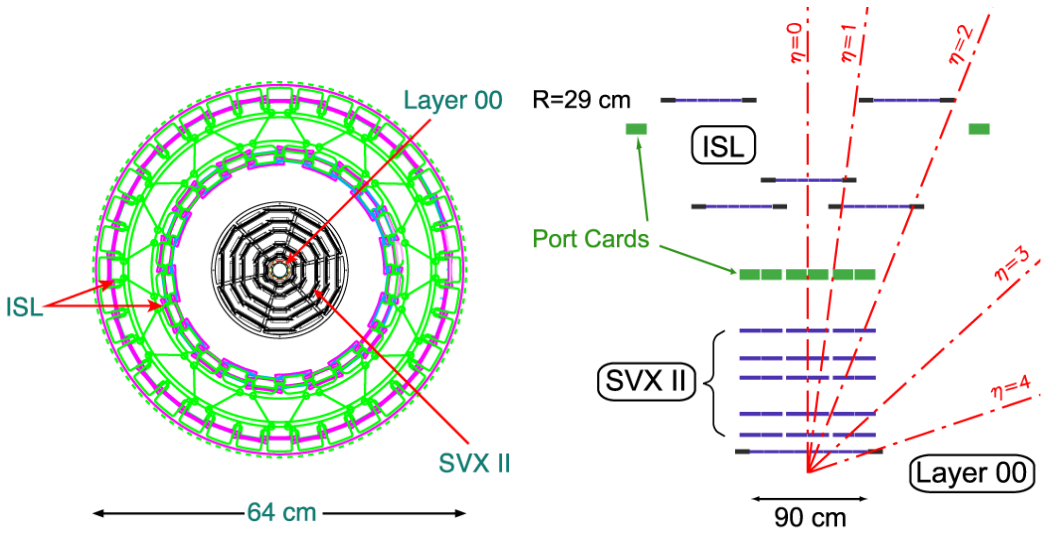


Figure 2.10: A quadrant view of the CDF II tracking system, the surrounding solenoid, and the forward calorimeters.

depletion region [56] is extended to include the entire strip of silicon. When a charged particle passes through a strip, it ionizes the silicon, creating electron-hole pairs consisting of electrons and positively charged ion pairs. The newly liberated electrons are pulled toward the more positive edge of the silicon by the electric field created by the applied voltage, and the holes go to the opposite direction. The resulting current is read out by readout chips mounted at the end of the sensors. Neutral particles do not ionize the silicon, thus they pass through the silicon detector without leaving a trace.

The silicon detector is comprised of three subdetectors: Layer 00 [57], the Silicon Vertex Detector (SVX II) [58], and the Intermediate Silicon Layers (ISL) [59]. The front and side views of the silicon detector are shown in Figure 2.11.

- *Layer 00* is a single-sided silicon microstrip detector attached directly to the beampipe. It consists of 12 sensors along the beam line, covering $|\eta| < 4.0$. To provide full coverage, they are arranged in two overlapping sub-layers at a radius of $r = 1.35$ cm and $r = 1.62$ cm. Layer 00 is designed to improve the precision of track measurements and the resolution of the position of secondary vertices.



(a) (b)

Figure 2.11. (a) Front and (b) side views of the silicon microstrip system.

- The *Silicon Vertex Detector (SVX II)* is the main part of the silicon detector. It is built in three cylindrical barrels with each barrel supporting five concentric layers of double-sided silicon microstrips. The SVX II detector extends from a radius of 2.4 to 10.7 cm and covers a region $|\eta| < 2$. The double-sided silicon layers have strips perpendicular to each other on the two sides, allowing for a simultaneous two-dimensional measurement of a particle's position. SVX II strips are axial (parallel to the z axis) on one side of a silicon layer for measurement of particles in the $r - \phi$ plane. For three-dimensional reconstruction, three of five layers have backside strips rotated by 90° with respect to the axial strips, and the other two layers have backside strips rotated by 1.2°. The SVX II detector provides high-resolution tracking information and is essential for resolving secondary displaced vertices (see Section 3.6.1).

- *Intermediate Silicon Layers (ISL)* are placed between the SVX II and the COT, covering the region $|\eta| < 2$. The ISL is divided into two regions. In the central region ($|\eta| < 1$), a single ISL layer is placed at a radius of 22 cm to provide enhanced linking of tracks between the SVX II and the COT. In the plug region ($1.0 < |\eta| < 2.0$), two layers of silicon are placed separately at radii of 20 cm and 28 cm. Since the COT coverage is incomplete in the plug region, the ISL improves track reconstruction in this region.

2.2.2.2. *Central outer tracker (COT).* The Central Outer Tracker [60] is a 3.1-m-long cylindrical drift chamber filled with a 50:50 mixture of argon and ethane gas. Its active volume covers a radial distance from 43.4 cm to 132.3 cm for $|\eta| < 1$. The COT is divided into eight concentric superlayers, with 2520 drift cells in total as shown in Figure 2.12. Each cell contains a wire plane containing 12 sense wires and 13 potential wires, with two additional shaper wires at either end as shown in Figure 2.13. The wire planes are separated by a field panel with stainless steel wires attached at either end. When voltage is applied to the wires, an energetic charged particle traversing the chamber ionizes the gas, forming electrons and oppositely-charged ions. These electrons and ions drift to wires at different electric potentials and create a current. The current is amplified and read out as a signal, called a *hit*. The positions of hits are then used to reconstruct the particle track. The hit position resolution is 140 μm and the resolution on track p_T is $\Delta p_T/p_T \sim 0.01\% \cdot p_T$, which slowly degrades with increasing $|\eta|$. The COT is organized into eight alternating stereo and axial superlayers. The axial superlayers lie parallel to the beamline, providing measurements in the $r - \phi$ plane. The stereo superlayers have a $\pm 3°$ offset from the beamline, which provides measurements along the z axis and enables three-dimensional reconstruction of tracks.

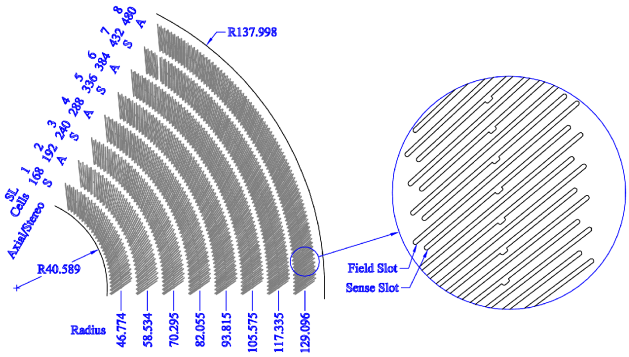


Figure 2.12: One-sixth of the COT endplate, showing drift cells organized into eight superlayers. For each superlayer, the total number of cells, the wire orientation (A = axial, S = stereo), and the average radius are provided.

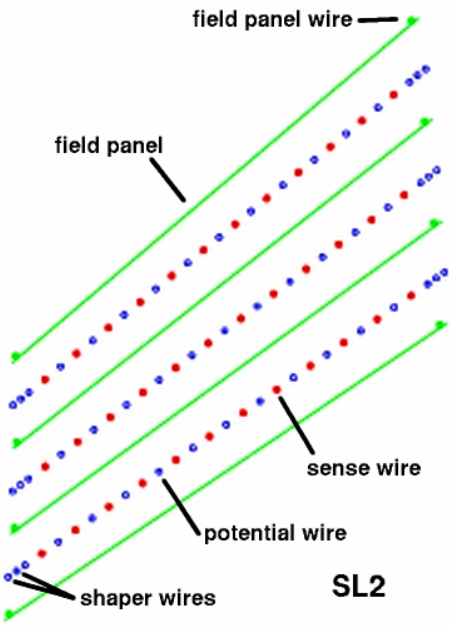


Figure 2.13. Cross-sectional view of a COT cell in superlayer 2.

2.2.2.3. *Time-of-flight detector (TOF).* The TOF detector [61] is a ring of plastic scintillators located outside of the COT. When charged particles pass through the scintillators, they create bursts of light that are collected by *photomulitplier tubes* (PMTs). The time difference between the bunch crossing and the arrival of a charged particle at the TOF can be measured with a resolution of about 100 ps. In this analysis, the timing information is used to recognize and reject cosmic rays.

2.2.3 Calorimeter System

While the tracking system determines the trajectory and momentum of a charged particle, the energy of a particle is measured by the calorimeter system. A particle passing through calorimeter material will lose energy through ionization and absorption. The CDF calorimetry system exploits the design of alternating layers of scintillator and absorber material. As a particle passes through the absorber material, it interacts with the atoms in the material (for electrons, this is called *bremsstrahlung*) and creates a *shower* of less energetic particles. The shower-

ing particles penetrate the scintillator, emitting light, which is collected and guided by fibers to PMTs. The response from a PMT is digitized and converted into a measurement of the deposited energy, based on the fact that the total amount of light observed is linearly correlated to the particle's energy. This process continues through numerous scintillator/absorber layers, until eventually all of the particle's energy is absorbed.

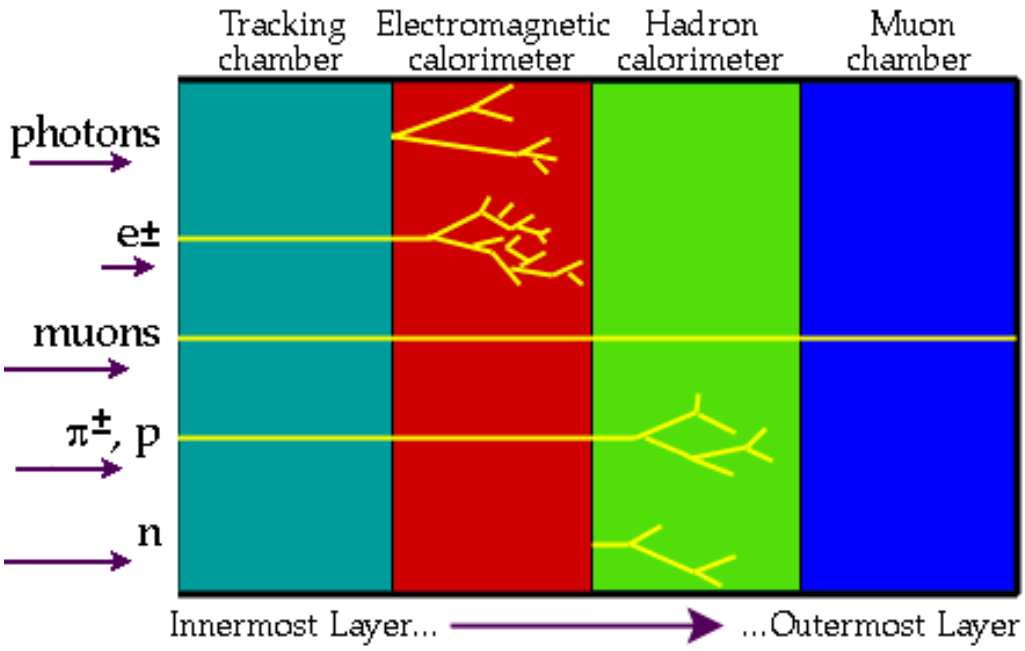


Figure 2.14: Diagram of the signal from various particles in the detector. The left side of the figure corresponds to the innermost tracking system of the detector; the right side of the figure corresponds to the outermost layer of the detector.

Since electrons have small masses and photons are massless, they radiate their energy much faster than heavy particles via electromagnetic interactions. Charged hadrons lose less energy in ionization. In addition, hadrons (both charged and neutral) lose energy through inelastic scattering with nuclei, which is mediated by the strong force. The calorimeter system is therefore built with two distinct type of calorimeters: lead-scintillator electromagnetic calorimeters and iron-scintillator hadronic calorimeters. Due to the slower radiation loss of hadrons, the hadronic calorimeter is thicker than the electromagnetic calorimeter and placed beyond the

electromagnetic calorimeter. The calorimeter system can measure the energy of particles as shown in Figure 2.14, including neutral particles, which leave the tracking system undetected with the exception of muons and neutrinos. Muons leave only minimum ionizing energy in the calorimeters while their energies are measured by the muon chambers, which will be described in the next section. Neutrinos travel through the detector undetected, but their transverse momenta can be determined from the transverse energy balance as discussed in Section 3.5.

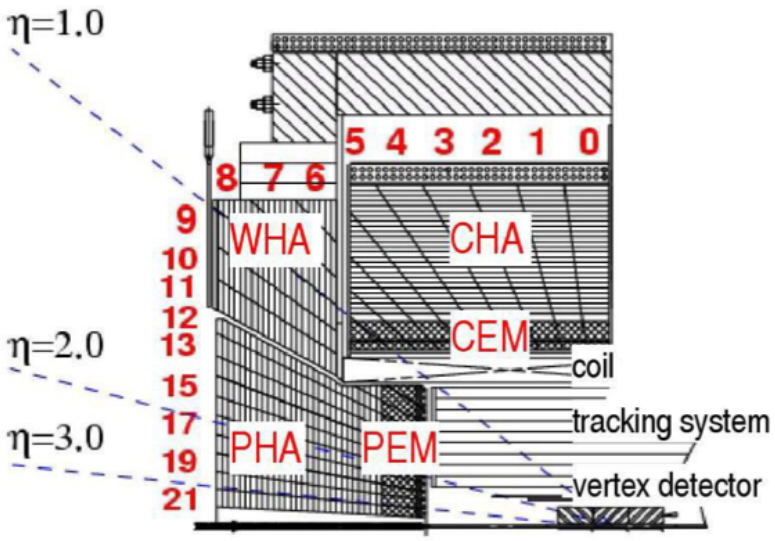


Figure 2.15. A quadrant view of the CDF calorimetry system.

Table 2.1: Overview of the calorimetry system. The depth is quoted in radiation lengths X_0 or hadronic interaction lengths λ_I.

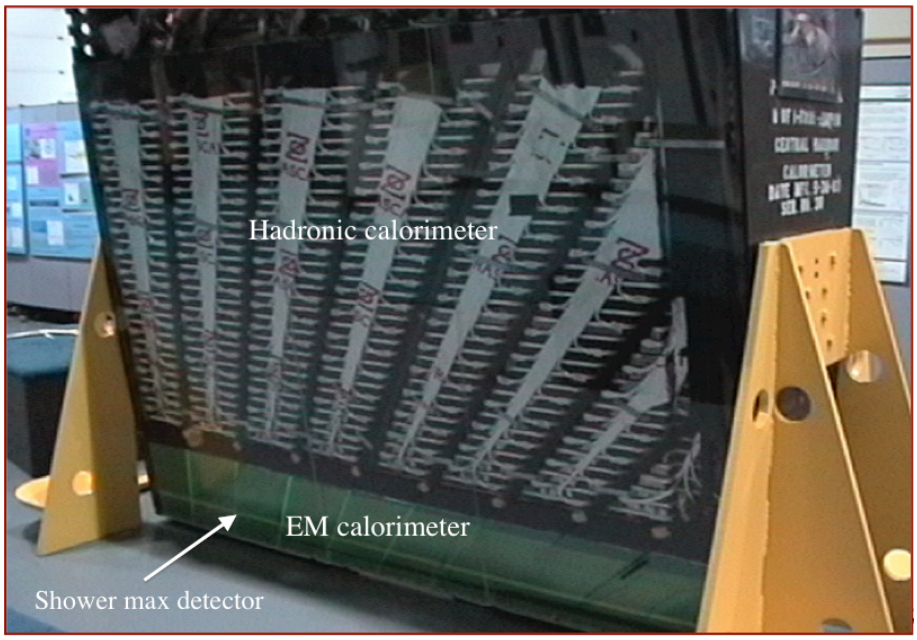
Calorimeter	η Range	Depth	Absorber
CEM	$\lvert\eta\rvert < 1.1$	$19X_0$	lead
PEM	$1.1 < \lvert\eta\rvert < 3.64$	$23.2X_0$	lead
CHA	$\lvert\eta\rvert < 0.9$	$4.7\lambda_I$	iron
WHA	$0.9 < \lvert\eta\rvert < 1.3$	$4.7\lambda_I$	iron
PHA	$1.3 < \lvert\eta\rvert < 3.64$	$6.8\lambda_I$	iron

The calorimeters are segmented into projective towers; each tower is a read-out unit covering a small range in the $\eta - \phi$ plane. The calorimeters consist of

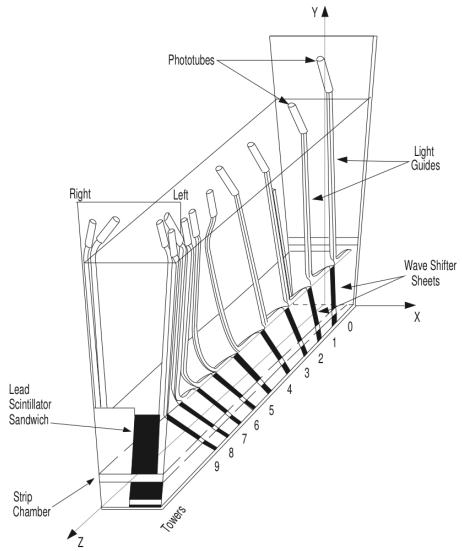
five sub-detectors as shown in Figure 2.15: the Central Electromagnetic calorimeter (CEM) [62], the Central Hadron calorimeter (CHA) [63], the End-Wall Hadron calorimeter (WHA) [63], the End-Plug Electromagnetic calorimeter (PEM), and the End-Plug Hadron calorimeter (PHA) [64]. Their properties are summarized in Table 2.1. Separated by the pseudorapidity, we describe the central and plug regions of the calorimeter system below.

2.2.3.1. *Central calorimeters.* The central calorimeter system is arranged in concentric layers that lie parallel to the beam line, covering a pseudorapidity of $|\eta| < 1.1$. Located directly outside the solenoid, the CEM is composed of 31 alternating layers of lead and scintillator. When particles traverse the CEM, the blue light emitted by the scintillator is collected by wave shifters embedded in the scintillator, which shift blue light to green light. The wave shifters are spliced to light-guide fibers, which carry the light out to two PMTs placed on both sides of each tower as shown in Figure 2.16(b). Outside of the CEM, the CHA has a similar structure with 32 alternating concentric layers of iron and scintillator. As illustrated in Figure 2.15, the CHA only covers the pseudorapidity range up to $|\eta| < 0.9$ due to its cylindrical design. In order to overlap the central electromagnetic calorimeter in pseudorapidity coverage, another hadronic calorimeter, the WHA, is placed along the endwall outside of the plug calorimeter, covering $0.8 < |\eta| < 1.2$. The WHA has a similar structure to the CHA, consisting of 15 layers of alternating iron and scintillator. However, its layers are oriented perpendicular to the beam, attached to the longitudinal endwall of the detector.

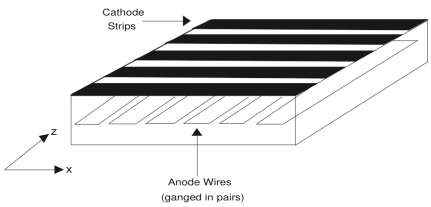
Within the CEM, a special layer is inserted at a point where the shower reaches the greatest number of particles in the electromagnetic calorimeter as shown in Figure 2.16(a). This layer is the central electromagnetic shower maximum chamber (CES). The CES is a series of wire and strip chambers as demonstrated in Fig-



(a)



(b)



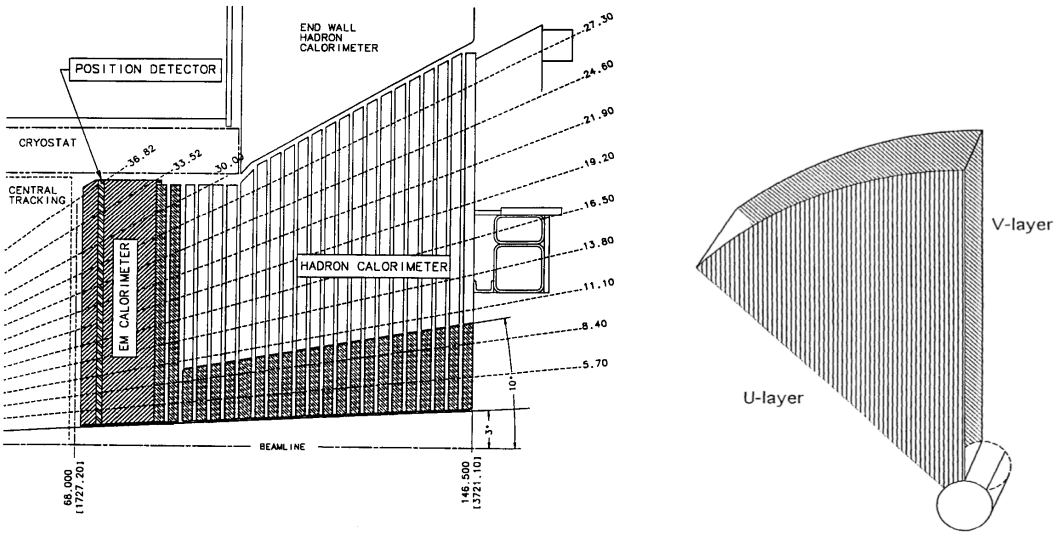
(c)

Figure 2.16: A central calorimeter wedge with labels to indicate the EM and hadronic calorimeters and shower maximum detector. (b) A detailed schematic of a CEM wedge. (c) A CES chamber, in which the z axis in the local $z - x$ plane is pointing in the same direction as incoming protons.

ure 2.16(c). The anode wires run along the direction of the z axis, providing position measurements in the $r - \phi$ plane. The cathode strips run in the direction of the local x axis, providing the z position of the shower in the CES. The wires and strips are more finely sectioned than the surrounding layers, providing more precise position measurements of particles than the coarsely segmented calorimeter. The CES information improves the EM cluster position resolution, contributing to the effective identification of electrons in this analysis.

2.2.3.2. *Plug calorimeters.* The plug calorimeter is named based on the fact that it resembles a giant plug that fits into the end of the central CDF detector. It is composed of an electromagnetic calorimeter (PEM) positioned within a hadronic calorimeter (PHA). The plug calorimeter allows the measurement of particles in the forward region of the detector with a pseudorapidity of $1.1 < |\eta| < 3.6$. The layers of the plug calorimeter are oriented perpendicular to the beam as shown in Figure 2.17(a). Designed similarly to the CEM, the PEM is composed of a series of 23 layers of alternating lead and scintillator. Placed next to the PEM, the PHA has 23 layers of alternating iron and scintillator. Because of the coverage of the WHA, the PHA only covers a pseudorapidity of $1.2 < |\eta| < 3.6$.

Like the CES, the plug electromagnetic shower maximum detector (PES) is inserted at the shower maximum point of the PEM. The PES detector is divided into eight 45° sectors; each sector contains two layers (called U and V) of 200 scintillator strips. The two scintillator layers U and V are aligned at +22.5° and –22.5° with respect to the radial direction to provide a two-dimensional position measurement. The fine position resolution of the PES improves the precision of position measurements in the plug calorimeter.



(a) (b)

Figure 2.17: (a) An elevation view of a quarter of the plug calorimeter. (b) The placement of the U and V strips in a 45° sector of the PES.

2.2.4 Muon System

Because most non-muon particles are absorbed by calorimeters or the steel absorber attached to the front of certain muon chambers, a muon system [65] is mounted outside of the calorimetry. The muon system is composed of a series of single-wire drift chambers, which detect passing charged muons and register the result as a muon *stub*. The drift chambers are paired with scintillators for timing information, and this is used to reduce the cosmic ray background. The muon chambers consist of four separate sub-detectors as shown in Figure 2.18, expanded in the $\eta - \phi$ plane.

2.2.4.1. *Central muon detector (CMU).* The CMU detector [66] is built directly outside the CHA calorimeter, covering the region $|\eta| < 0.6$. It is made of four radial layers of planar multi-wire drift chambers for muon detection. It is the oldest muon detector at CDF and has remained largely unchanged since Run I.

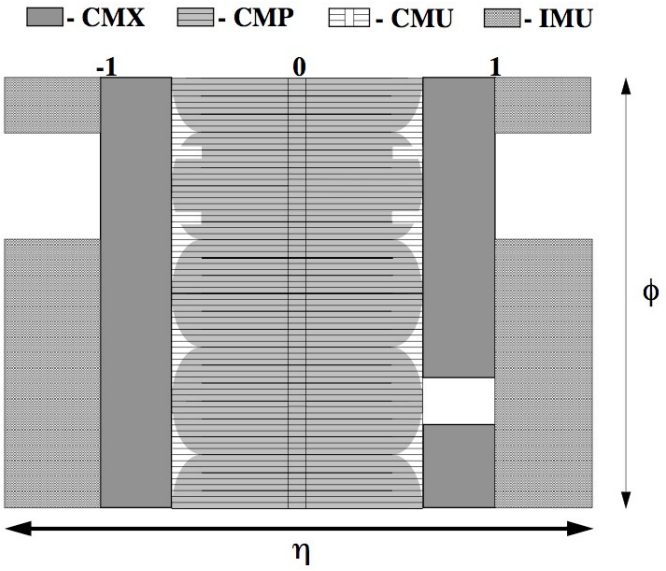


Figure 2.18: Diagram of the muon system coverage in the $\eta - \phi$ plane: the CMU, CMP, CMX, and BMU muon detectors. The BMU detector is referred to in this diagram as IMU.

2.2.4.2. *Central muon upgrade (CMP) and central scintillator upgrade (CSP).* The CMP was added during Run I to improve the CMU track purity. Since the CMU detector lacks shielding, it is easier for particles that are not muons to travel through the CMU, faking a muon signal. The CMP is placed behind large pieces of steel; this steel shielding wall is 60 cm thick and surrounds the central detector. It improves the CMP's signal-to-background ratio and increases the trigger efficiency. The CMP and CSP [67] consist of four layers of drift chambers and scintillators, which measure the momentum and timing of passing muons in the same η range as the CMU.

2.2.4.3. *Central muon extension (CMX) and central scintillator extension (CSX).* The Central Muon Extension (CMX) and Central Scintillator Extension (CSX) [67] extend the pseudorapidity coverage of the muon system to the region $0.6 < |\eta| < 1.0$. The CMX consists of three parts to provide full coverage as illustrated in Figure 2.19:

the arches, which fit in on the sides; the keystone, which sits on the top of the west side of detector; and the miniskirt, which goes beneath the detector. There is a gap on the east side of the detector for the cryogenic utilities for the solenoid.

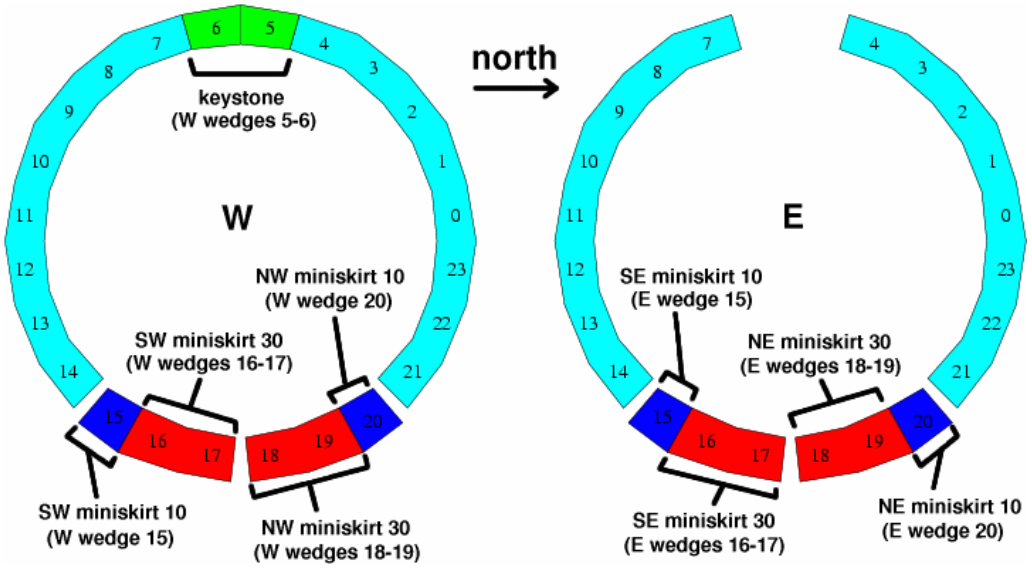


Figure 2.19. End view of the CMX wedge layer.

2.2.4.4. *Barrel muon detector (BMU) and barrel scintillator upgrade (BSU).* The Barrel Muon Detector (BMU) and Barrel Scintillator Upgrade (BSU) provides a measurement of the momentum and the timing of muons for $1.0 < |\eta| < 1.5$.

2.2.5 Luminosity Counter

The Cerenkov Luminosity Counters (CLC) [68] measure the instantaneous luminosity of the Tevatron beam collisions at the B0 interaction point. They are mounted close to the beampipe in the region $3.7 < |\eta| < 4.7$. The CLC is designed to detect the burst of Cherenkov radiation that results from charged particles coming from the interaction point at small angles. It is made of 48 Cherenkov light detectors filled with isobutane. The radiated Cherenkov light is collected and sent to the PMTs for readout.

2.2.6 Data Acquisition System

With collisions occurring every 396 ns, it is impossible to record every event. Moreover, only a very small set of collisions produce interesting physics processes. An elaborate system, the Data Acquisition System (DAQ), is designed to collect data fragments from the detector components and then select events for storage based on predetermined criteria. It is a three-level trigger system, illustrated in Figure 2.20. The first two trigger levels use special-purpose hardware based on a subset of detector components; the third level is implemented by software, which decides whether to record the event.

2.2.6.1. *Level 1.* The Level 1 trigger finds physics objects based on information from a subset of sub-detectors. It is composed of three parallel streams. Calorimeter trigger boards identify clusters of energy in the calorimeters, while muon trigger cards identify muon stubs from muon chambers. Additionally, the eXtremely Fast Tracker (XFT) [69][1] is designed to rapidly reconstruct tracks in the COT with high tracking efficiency using a pattern recognition technique. The resulting reconstructed tracks are matched to the energy depositions in the calorimeter towers, or hits in the muon chambers. Information from all three streams are sent to the Global Level 1 decision unit. With collisions occurring at a rate of 1.7 MHz, the Level 1 trigger reduces the rate down to about 40 kHz.

2.2.6.2. *Level 2.* The Level 2 trigger processes the time-ordered events passed by Level 1 with additional information from the CES and SVX II detectors. By exploiting information from the shower maximum detector, electron recognition and jet identification are improved at Level 2. Silicon information is processed by

[1] The Baylor high energy physics group was responsible for the maintenance of the XFT. We were responsible for monitoring its performance and responding to a pager when a problem occurred at any time. I shared the responsibility with my colleagues during my residence in Batavia.

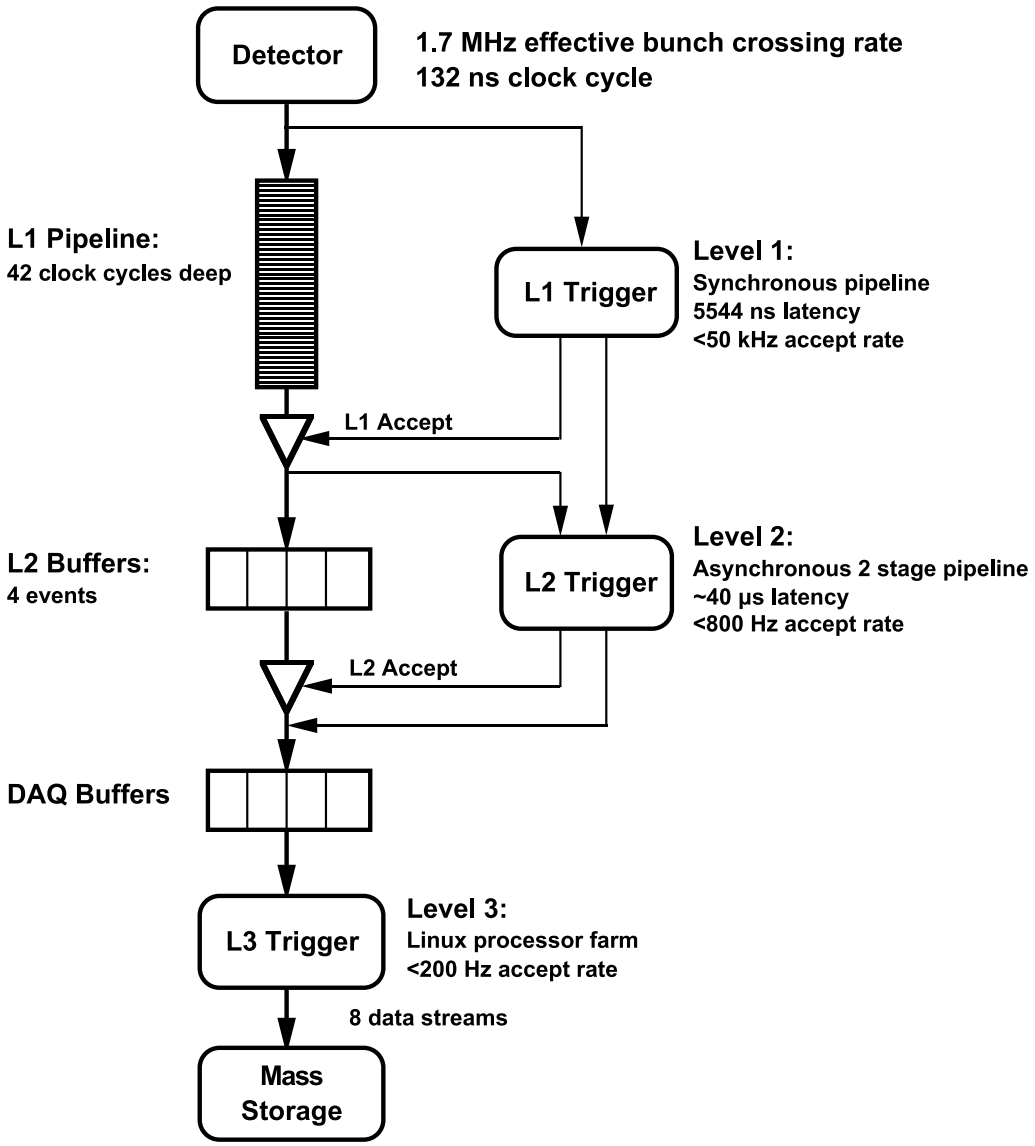


Figure 2.20. Functional block diagram of the CDF data acquisition system.

the Silicon Vertex Trigger (SVT) [70, 71]; the SVT identifies tracks with a large impact parameter, which is crucial for reconstructing jets from b quarks.

2.2.6.3. *Level 3.* The Level 3 trigger is implemented by software running on a Linux-based computer farm. Events passed at Level 2 are sent to an event builder, which assembles the Level 2 information into a Level 3-readable data format. Level 3 reconstructs the event with full detector information and makes further selection with high-level quantities. Accepted events are written to disk, ready to be further

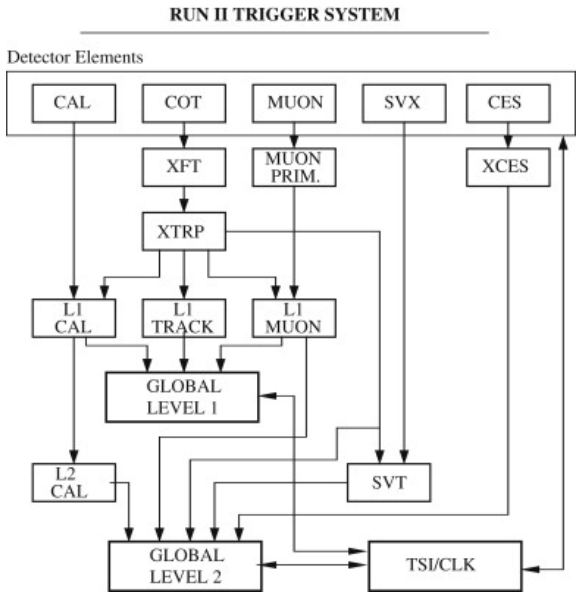


Figure 2.21: A schematic diagram showing the trigger system. The very top shows the detector elements, which are read out by the trigger hardware and then processed to form a Level 1 and Level 2 decision. If the event is accepted at Level 2, it is forwarded to the computing farms for the Level 3 decision.

processed offline. Level 3 reduces the event rate down to 100 Hz, a level that allows data to be stored on disk.

2.2.6.4. *Prescales.* The trigger system includes many different triggers, which are designed for different physics analyses and studies. Some triggers fire at a very high rate, thus it is necessary to *prescale* them in order to maintain the bandwidth. The prescale involves rejecting a fixed fraction of the triggered events. As the instantaneous luminosity of collisions increases, it is expected that the rate of a trigger also increases; it decreases as the instantaneous luminosity declines. Thus, *dynamic prescaling* is applied to some triggers, which changes the prescale on a trigger as instantaneous luminosity changes. Triggers with a dynamic prescale require dedicated modeling in the analysis.

2.2.7 Data Processing

The process of collecting collision data by the DAQ system is referred to as "online" data processing. Each continuous data-collecting period is called a *run*. The online data-taking system stores data from the Level 3 trigger to disk, and these data are referred to as *raw* data. For physics analysis, the raw data are later converted into an analytic data format and higher-level quantities are reconstructed. This is traditionally referred to as "offline" data processing.

A particular offline process, called *production*, splits the raw data into different datasets depending on the trigger path associated with each event; physics *objects* are then reconstructed in the data, where an object is an interesting physics quantity. The production process operates on a period of collected data. Data are subsequently organized by their period and trigger path at CDF.

From the production dataset, data are further stored into large arrays (ntuples) in `ROOT` [72], a format suitable for high-level analysis. According to the different types of information needed in the various analysis groups at CDF, several types of ntuples were developed. In this analysis we use the so-called `TopNtuples` [73], developed by the Top group at CDF.[2]

[2] I also joined the production and ntupling group, taking responsibility for producing production files and making ntuples for the whole CDF Collaboration.

CHAPTER THREE

Particle Identification and Reconstruction

As discussed in Chapter 2, the raw data, obtained either from Tevatron collisions or simulated Monte Carlo samples (see Section 4), are reconstructed offline as physics objects. First, information from subdetectors such as tracks in the tracking system (COT or silicon) or energy clusters in the calorimeters are collected and combined to form high-level detector objects. Afterwards, objects passing certain quality requirements are defined as physical object candidates. The resulting physics object candidates can be further identified and analyzed. Since this analysis is performed in W + jets samples, the physical objects of interest are a charged lepton (electron or muon), large missing transverse energy from an undetected neutrino, and two or three energetic jets; at least one jet must be identified as coming from a b quark. We will give a brief explanation of the identification and reconstruction of these physics objects in this chapter.

3.1 Track Reconstruction

The CDF tracking system measures the position of charged particle interactions with the detector material, which are called hits. The reconstruction of tracks is the task of identifying and combining all the hits along the trajectory of a particle's path inside the detector. Inside the CDF tracking system, charged particles moving in the solenoid's uniform axial magnetic field have a helical trajectory, and the energy loss by ionization in the detector material is negligible. The curvature of the helix depends on the momentum and the charge of the particle, and the helix points back to the origin of the particle. The reconstruction of the charged-particle trajectory is done in the offline reconstruction using a χ^2-based fit from the set

of spatial measurements of a candidate track. The helical fit takes into account non-uniformities of the magnetic field and scattering in the detector material.

3.1.1 Tracking Algorithms

Since the CDF tracking system consists of two dissimilar subdetectors, the COT and the silicon tracking system, the tracking algorithm varies for each subdetector. CDF exploits several tracking algorithms [74], each optimized for the information available in different detector regions. Since both detectors are 3D tracking devices, they support 3D tracking reconstruction algorithms.

The tracks in the COT are first reconstructed. This is because the active volume of the COT is larger and further from the beam line than the silicon tracking system, which leads to a lower track density and a higher number of isolated tracks, resulting in fewer combinations. Thus, the track reconstruction in the COT is purer and faster than in the silicon system. Due to the geometry of the COT, the reconstructed tracks are limited to the range $|\eta| \leq 1$ and $p_T > 0.5$ GeV/c. There are two different axial reconstruction algorithms in use for the COT. The first is a "segment linking" algorithm, adopted from Run I, which reconstructs and connects segments in the superlayers and fits them with a minimum χ^2 method to find the trajectory. The second algorithm, which is called "histogram linking," starts with a segment position from the outermost superlayer and the beam position. Then, it fills a histogram with hits from the other superlayers that, in curvature space, are within 1 cm from the reference circular road as illustrated in Figure 3.1. The hits along the track populate the same bin in this histogram, resulting in the identification of a 2D track [75]. After track reconstruction using axial layers, stereo information is used for 3D track reconstruction. First, the stereo segment linking algorithm matches stereo segments to existing axial tracks. The matched stereo segments are fit in the $r - z$ plane: if the χ^2 fit yields $\chi^2 < 100$ and the track z_0 is less than 175 cm, the

stereo segment linking is a success. A second algorithm, stereo hit linking, recovers the stereo information for axial tracks that failed the stereo segment linking algorithm. It uses the z vertex seeds produced by the segment-linked tracks and scans for the best hit combination. The set of COT tracks from the above algorithms might include numerous duplicates. All the duplicate tracks are compared, and the best track is kept.

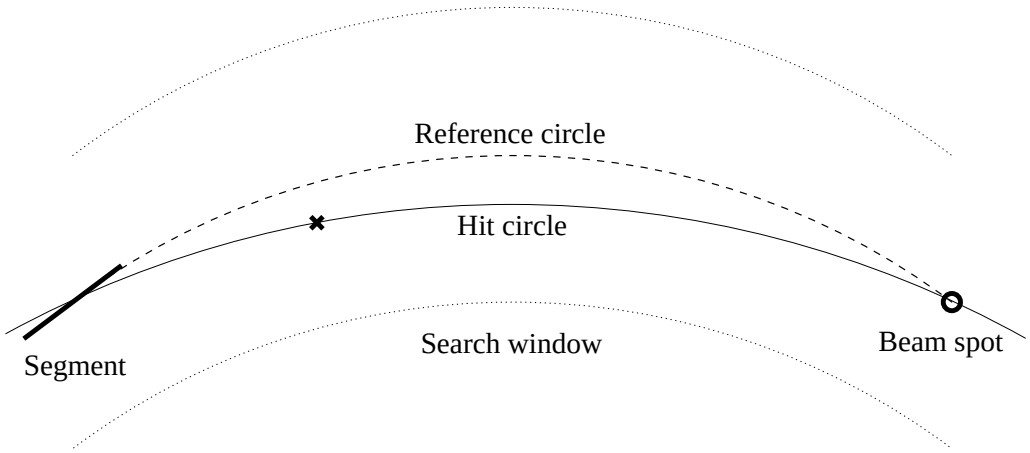


Figure 3.1. Illustration of the histogram linking algorithm, taken from Ref. [75].

The silicon track reconstruction has three different phases. The first method, called "Outside-In," extrapolates the tracks found in the COT into the silicon detector. Axial and stereo silicon hits are subsequently added to the track (from outside in) by a progressive fit. The second phase ("Silicon-stand-alone") consists of stand-alone pattern recognition in the silicon subdetectors. Hits already assigned to another track are not considered in order to reduce the number of combinations. The silicon-stand-alone technique enables track-finding in the forward region up to about $|\eta| \lesssim 2.8$, which is not covered by the COT. In the last phase, silicon tracks are extrapolated from the previous Silicon-stand-alone algorithms into the COT, which is called "Inside-Out" tracking [76]. The inside-out algorithm recovers COT hit information from particles passing through fewer COT superlayers than are re-

quired by the COT pattern recognition algorithm. It also improves the efficiency in the forward region ($1.2 \leq |\eta| \leq 1.8$) where the COT coverage is limited.

A special algorithm is used for identifying tracks from electrons in the forward region, called the "Phoenix" algorithm [77]. Similar to the outside-in tracking algorithm, it combines calorimeter information (an energy cluster in the PEM) with silicon stand-alone hits. The position of the PEM cluster and the primary vertex are used to construct *seed* tracks. For each seed, two hypotheses about the charge of the particle are considered by computing the curvature for both an electron and a positron corresponding to the deposited energy. The charge of the particle is determined from the best fit.

3.1.2 Primary Vertex

Since the interaction region of the particle beams has a substantial volume, an accurate determination of the primary interaction point of the hard scattering is crucial for this analysis. Knowledge of the position of the primary vertex affects the measured kinematic properties of the event objects. Furthermore, the identification of long-lifetime particles heavily depends on the primary vertex, as in secondary vertex tagging (see Section 3.6). Because of multiple proton-antiproton interactions in the event, it is possible to have more than one primary vertex. The primary vertices are reconstructed from seed vertices [78]. Tracks with $|z_{\rm trk} - z_{\rm vtx}| < 1$ cm, $|d_0| < 1$ cm, and $|d_0/\sigma_{d_0}| < 3$ are ordered in decreasing p_T and used to identify a common vertex based on a χ^2 fit. The tracks with $\chi^2 > 10$ are iteratively removed until either no track fails the χ^2 cut or the number of tracks associated with the vertex falls below a minimum quantity. The primary vertex with the best fit value is chosen as the primary vertex of the hard scattering in the event.

3.2 Calorimeter Clustering

Particles passing through the calorimeter deposit energy in the calorimeter towers. Collecting these towers to form clusters is important for particle identification and jet reconstruction.

For a given event, towers with measured E_T larger than 100 MeV are called clusterable towers; clusterable towers that have $E_T > 3$ GeV are called seed towers. The calorimeter clustering algorithm starts with a seed tower; adjacent clusterable towers are added to form a cluster. The resulting cluster size varies by calorimeter subdetectors and ranges from a minimum of one tower up to a maximum size of 3×3 towers.

The EM clustering in the event starts from the largest-E_T seed tower and proceeds until the lowest-E_T seed tower is considered [79]. The sum of the energies of the towers in the cluster is defined as the total energy of the cluster. Thus, the total energy in a EM calorimeter cluster is the EM energy (E_{EM}) of the cluster, and the total energy in a hadronic calorimeter cluster is the hadronic energy (E_{HAD}) of the cluster. The final calorimeter energy of the reconstructed object is $E = E_{\mathrm{EM}} + E_{\mathrm{HAD}}$.

Once a cluster is defined, the position of the cluster is defined by the energy-weighted mean of the towers in the cluster. By matching the cluster position with the cluster in the shower maximum detector (reconstructed with a similar algorithm), the precision of the cluster position is improved since the shower maximum detector has better position resolution. This gives us the final cluster position.

3.3 Lepton Reconstruction

3.3.1 Electrons

An electron typically manifests itself in the detector by producing a track in the tracking system and depositing most of its energy in the electromagnetic calorimeter. Thus the electron candidate is a physics object with a high-p_T isolated track, which

is matched to an energy deposit (a cluster) in the electromagnetic calorimeter. We identify central electrons from an energy cluster in the CEM; these are referred to as CEM electrons. In the forward region, electrons are identified by a cluster in the PEM, which has to be matched to a Phoenix track; thus we refer to forward electrons as PHX electrons.

3.3.2 Muons

Since muons have a relatively long mean lifetime of 2.2 $\times 10^{-6}$ s, they can travel through the detector before decaying. Muons pass through the calorimeter system as minimum ionizing particles (MIP), which refers to the fact that their mean energy loss rate through matter is minimal. They can be identified by a track inside the COT, a track segment ("stub") in the muon chambers, and an energy deposit consistent with a MIP in the EM and HAD calorimeters.

3.4 Jet Reconstruction

An isolated parton (quark or gluon) produced in the hard collision will quickly undergo *parton showering* and then recombine with quarks and antiquarks spontaneously created from the vacuum to form colorless hadrons (*hadronization*). The resulting hadrons will further decay as they travel through the detector material and form a wide spray of particles in the detector. A characteristic energy profile is deposited in the calorimeters; the collimated cluster of these stable particles is collectively known as a *jet*, which approximately retains the total momentum and direction of the initial parton (see Figure 3.2).

3.4.1 Jet Algorithm

The reconstruction of jets is carried out by a jet clustering algorithm using the energy deposited in the electromagnetic and hadron calorimeter towers. The role of the algorithm is to associate these energy depositions into jets such that the

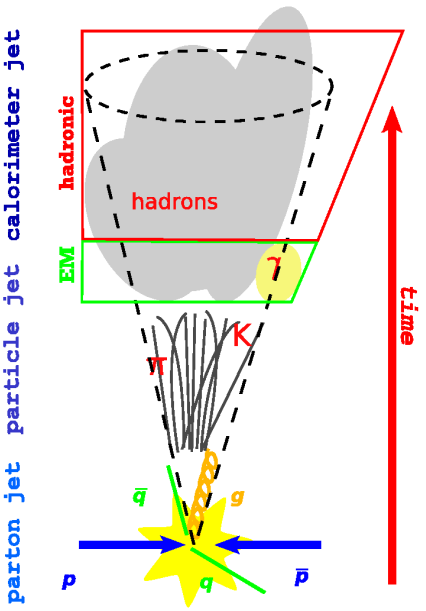


Figure 3.2: Schematic picture representing the hadronization process from an initial parton to its final calorimeter energy cluster, which is called a jet.

kinematic properties of the jets can be related to the corresponding properties of the energetic partons produced in the hard scattering process [80]. In this analysis, we use the JETCLU cone algorithm for jet clustering.

A cone algorithm forms jets by associating together all calorimeter towers within a circle of radius R in $\eta \times \phi$ space, where the center of the jet is defined as $(\eta_{\text{jet}}, \phi_{\text{jet}})$ and the size of the jet cone R is defined as

$$R = \sqrt{(\eta_{\text{tower}} - \eta_{\text{jet}})^2 + (\phi_{\text{tower}} - \phi_{\text{jet}})^2} \tag{3.1}$$

In this analysis, we use a cone size of $R = 0.4$. The algorithm starts with a list of seed towers (calorimeter towers with $E_T \geq 1$ GeV) that are sorted by decreasing E_T. We let $E_{T_i} = E_i \cdot \sin\theta$ denote the transverse energy deposited in i^{th} tower with respect to the primary vertex, where E_i is the total energy measured in the electromagnetic and hadronic components of that tower. Starting from the highest-E_T seed tower, a precluster is built by adding all the adjacent seed towers within the cone of radius R. The E_T-weighted centroid is next calculated to form a new cone of

radius R. Then, towers within the new cone that have E_T larger than 100 MeV are clustered and a new centroid is formed. This process of preclustering iterates until the list of towers assigned to the cluster is stable, that is, when the E_T-weighted cluster centroid is aligned with the geometric axis of the cone, which is called a "stable solution." During these iterations, no clustered towers are ever removed from the cluster, even if they are far way from the center of the cone. This is called "ratcheting." Once a stable solution is found, the next unused seed tower is used to form the next precluster. The clustering continues on the list of seed towers until the last seed tower on the list is clustered. All of the clusters become the jets in the event; towers that have E_T less than 100 MeV are called unclustered energy in the event. At the end of this process, it is possible that some towers are contained in two different jets. Two jets that have more than 75% of their towers in common are combined into a single jet; otherwise towers in two different jets are assigned to the nearest jet. The final jet energy E_{jet} is computed from the final list of N_{tow} towers:

$$E_{\text{jet}} = \sum_{i=1}^{N_{\text{tow}}} (E_i^{\text{EM}} + E_i^{\text{HAD}}) \tag{3.2}$$

where E_i^{EM} and E_i^{HAD} are the energies of the i^{th} tower in the electromagnetic and hadron calorimeters, respectively.

3.4.2 Jet Energy Corrections

In many high energy physics analyses, jets are used to estimate the energies of partons resulting from the underlying physics process. For various reasons, the energy reconstructed by the cone algorithm isn't an adequate measure of the parton's energy. Several correction factors have been developed to better estimate the original parton energy by comparing the jet energy response between data and Monte Carlo simulation for various physics processes [81]. Since these corrections are divided into discrete levels, a subset of them can be applied according to the needs of an analysis.

3.4.2.1. *Level-1: η-dependent correction.* Due to the geometry of the CDF calorimeter, the calorimeter response is not uniform in pseudorapidity. This results from the difference in the performance of the calorimeter in the central and forward regions, and it also results from hardware non-uniformities due to the presence of cracks between sections of the calorimeter. The Level 1 (L1) correction is calibrated by using the balance of transverse energy in dijet events to scale jet energies outside the $0.2 < |\eta| < 0.6$ region to that inside of this region. The $0.2 < |\eta| < 0.6$ region is where the CEM and CHA located, which are the best understood calorimeters of CDF and far away from cracks. The correction also includes a E_T-dependence, which is due to the fact that the response of the central and forward calorimeters depends on the E_T of the jet. And it includes a data run dependence, which uses a different set of corrections depending on the time periods of the data. The final L1 correction function can be parameterized as:

$$f_{L1}(\eta, E_T, r) = \beta^{-1} \tag{3.3}$$

where β is a factor that depends on the average momentum balancing fraction between a jet selected only in the central part of the calorimeter ($0.2 < |\eta| < 0.6$) and the other jet within $|\eta| < 3.6$, as shown in Figure 3.3.

3.4.2.2. *Level-4: Multiple $p\bar{p}$ interactions correction.*[1] At higher instantaneous luminosity, more than one $p\bar{p}$ interaction occurs in the same bunch crossing (referred to as pileup in Section 4.1.6). The energy from these pileup events may fall into the jet clustering cone of the hard interaction and must be subtracted from the jet before comparing with theory predictions. Minimum-bias events, which are triggered by activity in the CLC and usually have small number of hard scatterings, are used to obtain the correction. The number of additional reconstructed vertices

[1] Level 2 and 3 are deprecated and not discussed.

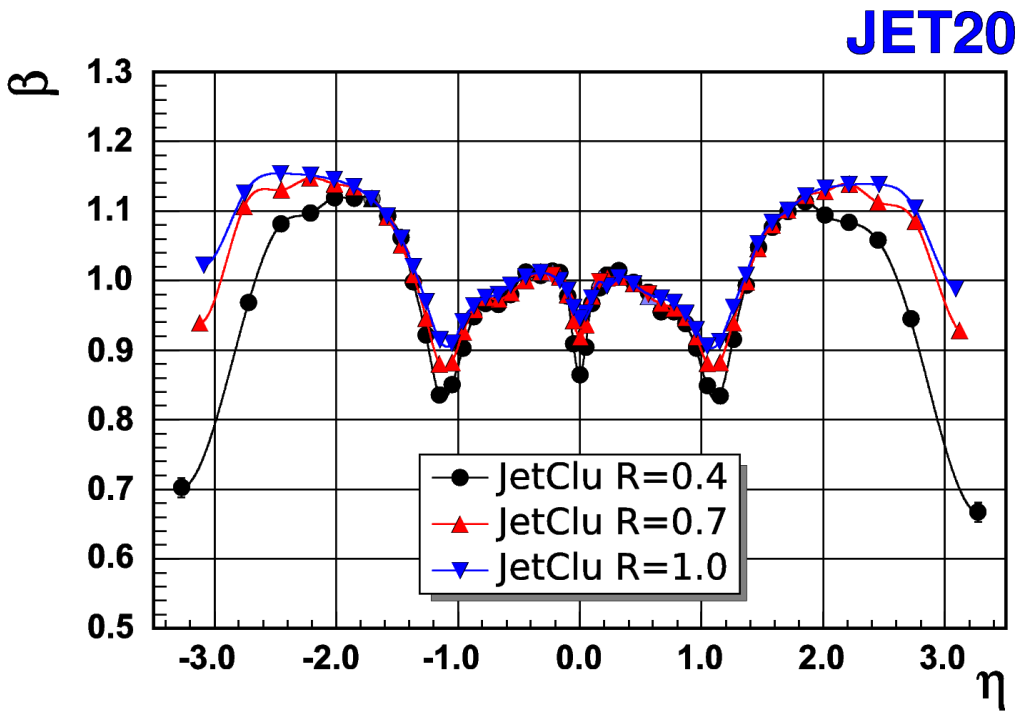


Figure 3.3: η-dependence of the β factor for different jet cone sizes R measured in the dijet component of a jet20 sample (a data sample containing jets with $E_T > 20$ GeV).

in the event gives a hint of the number of additional interactions in the same bunch crossing. In general, the number of reconstructed vertices follows a Poisson distribution whose mean increases with instantaneous luminosity. In this correction, the transverse energy within a random cone was measured in minimum-bias data and parameterized as a function of the number of vertices in the event.

3.4.2.3. *Level-5: Absolute jet energy scale correction.* After applying the η-dependent corrections and the multiple $p\bar{p}$ interactions correction, the measured jet energy needs to be further corrected for any non-linearity and energy loss in the uninstrumented regions of each calorimeter. This correction is determined from inclusive dijet events with jets corrected by all the corrections described above. The absolute energy scale correction is derived by comparing *calorimeter* Monte Carlo (MC) jets, which are reconstructed from simulated detector calorimeter towers, with

hadron-level jets, which are jets clustered from MC-simulated stable particles before the detector simulation using the same CDF clustering algorithm. The hadron-level jets include particles from the "underlying event" and initial-state radiation (see Section 4.1.5 and 4.1.7), which may fall within the jet cone. They also include particles that might not appear in the calorimeter response, due to interactions with material in front of the calorimeter or the magnetic field. The calorimeter jets are matched to hadron-level jets in $\eta - \phi$ space within $\Delta R < 0.1$. The transverse momenta of the calorimeter jet and the hadron-level jet, p_T^C and p_T^H respectively, are then calculated. The jet energy correction is defined as a function of $\mathcal{P}(p_T^C|p_T^H)$, the probability to measure p_T^C given p_T^H. Figure 3.4 shows the correction factor f_{L5} for different cone sizes as a function of jet p_T. Since this correction deals with the difference between the final jet energy measurement in the calorimeter and the parton-level measurement, it is also referred to as the absolute jet energy scale correction. In this analysis, we use corrections up through Level 5 (L5) for our jets.

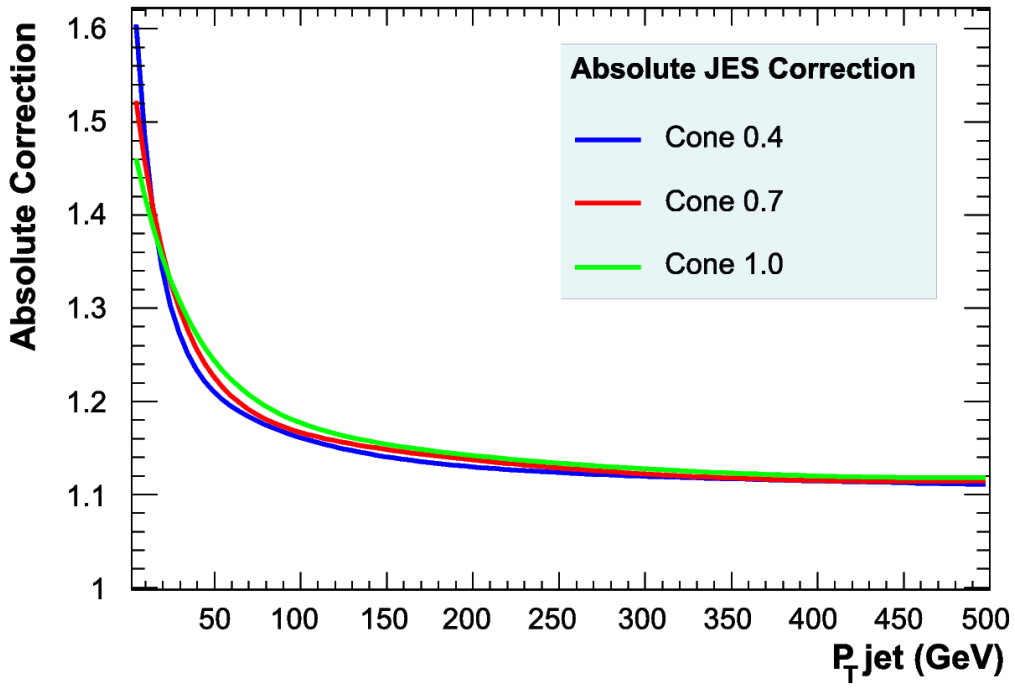


Figure 3.4. Absolute jet energy scale correction (f_{L5}) for different jet cone sizes.

3.4.2.4. *Level-6: Underlying event correction.* In order to determine the energy of the original parton rather than the corresponding jet energy, the energy of the underlying event (see Section 4.1.5) needs to be subtracted. This correction is obtained from minimum-bias events with only one vertex. Since the L5 absolute correction already includes the underlying event, a further correction factor 1.6 is applied to obtain the final correction.

3.4.2.5. *Level 7: Out-of-cone correction.* Due to our fixed cone size, the jet clustering may not include all the energy from the initiating partons. Some particles from the partons generated during fragmentation may fall outside the cone, and their energies must be added to the calorimeter jet to get the true parton-level jet energy. Correction functions of the measured jet p_T for different cone sizes are derived. The Level 7 (L7) jet energy correction was only used for some sideband studies in this analysis, but not used in the primary analysis.

3.5 Missing Transverse Energy ($\not\!\!E_T$)

Neutrinos cannot be detected by CDF since they pass through the detector without interacting. However, an energetic neutrino from the hard scattering will manifest itself as missing energy in the overall energetic balance of the event. This is usually referred to as *missing transverse energy*. While it is impossible to know the exact longitudinal momentum of the colliding partons, the transverse component of the interacting partons is assumed to be zero. The missing transverse energy of an event, usually written as $\not\!\!E_T$, is calculated with a vector sum over the transverse energies of the calorimeter towers:

$$\vec{\not\!\!E}_T \equiv -\sum_i E_T^i \hat{n}_i \tag{3.4}$$

where i denotes the number of calorimeter towers that lie within $|\eta| < 3.6$ and have energy E_T^i above 0.1 GeV, and $\hat{n}_i$ is a unit vector that points from the primary vertex to the center of the i^{th} calorimeter tower.

After the missing transverse energy is calculated from Equation 3.4, further corrections are needed. Since muons pass through the calorimeters as minimum ionizing particles (MIP), a correction is required for events containing muons: by extrapolating the muon track to the calorimeters, the transverse energy corresponding to the MIP is removed for Equation 3.4; the transverse energy of all muons is then added to the total event E_T balance. Because various jet energy corrections affect the total E_T in the calorimeters, the $\not\!\!E_T$ is corrected by the transverse energy difference between raw jets and the Level 5 correction jets. The final formula for the missing transverse energy becomes:

$$\vec{\not\!\!E}_T^{\,L5} = -\sum_{i=1}^{N_{tow}} E_T^i \cdot \hat{n}_i + E_T^{MIP} \cdot \hat{n}_\mu - \vec{E}_{T_\mu} + \sum_{j=1}^{N_{jets}} \vec{E}_{T_j}^{RAW} - \sum_{j=1}^{N_{jets}} \vec{E}_{T_j}^{L5}. \tag{3.5}$$

3.6 Jet b-tagging

A b quark produced in the hard scattering will immediately hadronize into a B hadron, which can be a b meson (B^0, $B^{\pm}$, B_s^0) or a b baryon (like Λ_B). Since B hadrons can only decay through weak interactions, they have a relatively long lifetime. Because the B hadron carries most of the transverse momentum of the original b quark, it has a large Lorentz boost and travels a transverse distance of several millimeters before decaying. This produces a secondary vertex inside the b jet that is displaced from the primary interaction point, as illustrated in Figure 3.5. The decay of the B hadron produces a subjet with a large impact parameter d_0, which is the distance in $r - \phi$ space from the primary vertex to the point of a track's closest approach. The associated uncertainty of the impact parameter σ_{d_0} includes both the uncertainty on the track's position and the beam line position. The impact parameter can be reconstructed by the silicon detector with a precision

of about 50 μm. This enables the reconstruction of both the primary and secondary vertices of the event. With these special features, we are able to determine whether a jet originates from a bottom quark. This process is called "b-tagging." This b-tagging procedure is implemented at CDF using the secondary vertex reconstruction algorithm, SECVTX [82].

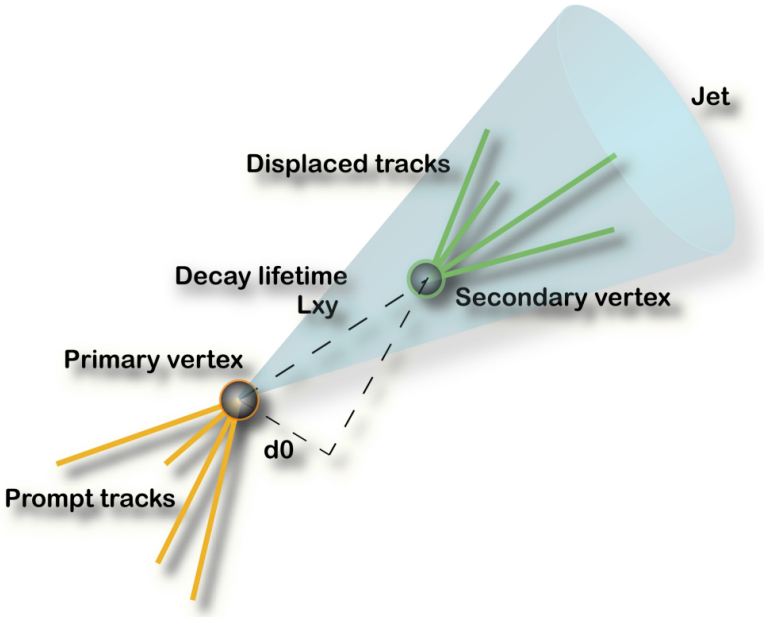


Figure 3.5. Schematic view of displaced tracks forming a secondary vertex.

3.6.1 SECVTX *Algorithm*

The SECVTX algorithm relies on the displacement of secondary vertices relative to the primary vertex to identify b jets. On a per-jet basis, secondary vertex tagging operates on tracks within the jet cone. In order to remove poorly reconstructed tracks, a set of cuts involving the transverse momentum, the number of silicon hits, the quality of those hits, and the chi-squared per degree of freedom χ^2/ndf of the final track fit are applied to the tracks. A jet is defined as "taggable" if it has at least two good tracks; only a taggable jet can produce a displaced vertex. The SECVTX algorithm uses a two-pass approach to find secondary vertices. The first pass attempts to reconstruct a secondary vertex with at least three tracks

with $|d_0/\sigma_{d_0}| > 2.5$ and $p_T > 0.5$ GeV/c, while at least one of these tracks has $p_T > 1$ GeV/c. In case the first pass fails, a second pass is performed with a tighter track requirement ($p_T > 1$ GeV/c and $|d_0/\sigma_{d_0}| > 3$) that attempts to reconstruct a two-track vertex with at least one track having $p_T > 1.5$ GeV/c.

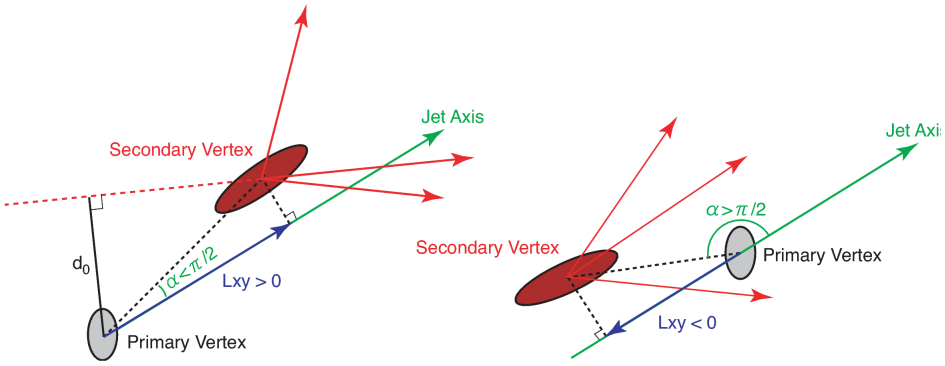


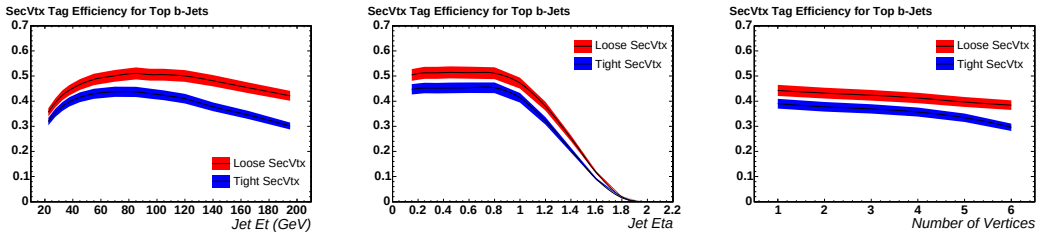
Figure 3.6: Left: A true reconstructed secondary vertex. Right: A negative SECVTX tag, which is a falsely reconstructed negative secondary vertex.

Once a secondary vertex is identified in a jet, the two-dimensional decay length of the secondary vertex L_{xy} is calculated by the SECVTX algorithm. L_{xy} corresponds to the projection onto the jet axis, in the (r, ϕ) plane, of the vector pointing from the primary vertex to the secondary vertex. The sign of L_{xy} is defined relative to the jet direction, specified by the angle α between the jet axis and the SECVTX vector: it is positive if $\alpha < \pi/2$ and negative if $\alpha > \pi/2$, as shown in Figure 3.6. Jets arising from the decay of a B hadron are expected to have secondary vertices with large positive L_{xy} due to the long lifetime of the B hadron. Due to the finite tracking resolution of the CDF tracking system, the SECVTX algorithm can falsely reconstruct a negative secondary vertex from random mis-measured tracks. A cut on the transverse decay length significance $|L_{xy}/\sigma_{xy}| > 7.5$ is required to reduce the background from false secondary vertices. The remaining negative tags are useful for estimating the rate of fake b-tags in the sample.

The SECVTX algorithm defines two types of b-tag: "tight" or "loose" tags. The tight b-tag requires $|d_0/\sigma_0| > 3.5$ in pass 2 and a tighter vertex χ^2 fit. A loose tag uses the same method with less stringent requirements ($|d_0/\sigma_0| > 3.0$ in pass 2 and a looser vertex χ^2 fit). We use only tight b-tags in this analysis.

3.6.2 SECVTX *Tag Efficiency*

The b-tag efficiency in the data is calibrated using the method detailed in Ref. [82]. By taking advantage of the characteristic semileptonic decays of B hadrons, the tagging efficiency can be estimated with two different methods, the electron method and the muon method. The electron method uses dijet events in which one jet is tagged with the SECVTX algorithm and the other jet has a high-momentum electron. The fraction of the electron jets passing the SECVTX tag is used to calculate the SECVTX tagging efficiency of b jets that contain electrons. The muon method is similar to the electron method, but it uses jets containing a high-energy muon. Both methods gives similar results, and we use the tag efficiency obtained from the electron method in this analysis. As shown in Figure 3.7, the tagging efficiency does not show a noticeable kinematic independence on jets. The drops in tagging efficiency in Figure 3.7(b) are due to the low track reconstruction efficiency in the forward region.



(a) (b) (c)

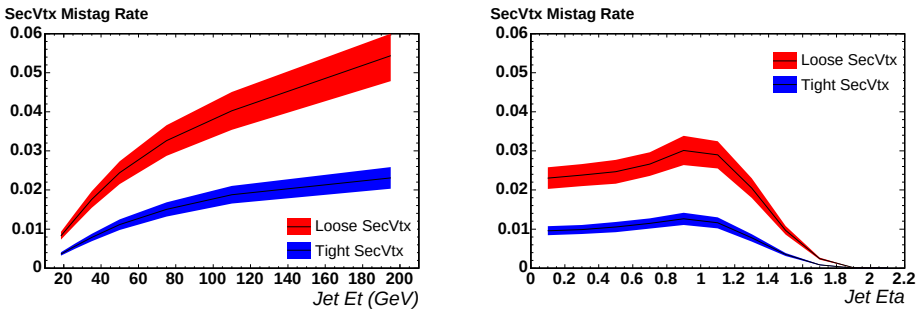
Figure 3.7: SECVTX b-tagging efficiency as a function of (a) jet E_T, (b) jet η, and (c) number of vertices, for both the loose and tight SECVTX b-tagger.

Simulated MC events tend to overestimate the SECVTX efficiency compared to data, thus a "scale factor" (Φ) is applied to the tagging efficiency to account for the data/MC difference. The b-tag efficiency from data is compared with the efficiency from Monte Carlo, and the ratio of the efficiencies is taken as the scale factor. The same scale factor is assumed for the tagging rate of the charm quarks. The scale factor 0.95 ± 0.04 [83] is applied to b-tagged jets matched with heavy flavor quarks in the MC events in this analysis. The systematic uncertainty on the scale factor is twice as large in the double b-tagged events.

3.6.3 Mistag Matrix

A jet that does not result from the fragmentation of a heavy quark, yet has a SECVTX secondary vertex, is called "mistag." Caused mostly by random overlap of tracks that are displaced from the primary vertex, mistags are mainly light flavor jets. Although the rate of mistags is very low after the $|L_{xy}/\sigma_{xy}| > 7.5$ requirement, there is still a substantial contamination in the tagged sample due to the large production rate of light jets. Because the SECVTX algorithm exploits a symmetric requirement on L_{xy}, a good estimate of the positive mistag rate due to resolution effects can be obtained from the negative tag rate. Considering that some heavy flavor jets can be negative tagged, a correction term is needed, resulting in an "asymmetry" estimation of the mistags.

Unlike real b-tags, the mistag rate has a strong dependence on kinematic variables. The rate of mistags is measured in inclusive jet trigger data samples, parameterized as a function of five variables: E_T, η, ϕ, the number of tracks of the jet, and the scalar-summed E_T of all jets in the event. A five-dimensional *mistag matrix* is constructed as a function of these variables, divided into four to eleven bins. A systematic uncertainty (10%) is assigned to the mistag matrix, taking into account



(a) (b)

Figure 3.8: The rate of mistags for the loose and tight SECVTX b-tagger as a function of (a) jet E_T and (b) jet η. The tight b-tagger has a lower mistag rate, thus this analysis uses the tight SECVTX tag.

the jet trigger bias and the difference from various jet samples. The systematic uncertainty is doubled in the double b-tagged events.

CHAPTER FOUR

Monte Carlo Simulation

For the measurement of the single top quark cross section, an understanding of the efficiencies and kinematic properties of reconstructed collision events is essential. A deep knowledge of the physics processes and the detector response is therefore necessary, which can be learned by comparing the reconstructed data with simulated Standard Model physics processes. The simulated physics processes are obtained from Monte Carlo (MC) event generators, which produce random collision events according to the probability density function of the phase space for a given process and the matrix element of that process in theoretical calculations. The resulting simulated particles are then passed to a CDF detector simulation that emulates the detector response, and produces the "real life" collision event as collected from the CDF detector. In this chapter, we will describe the signal model of single top production and all the background MC models.

4.1 Monte Carlo Event Generation

In high energy physics, we either try to validate known SM physics processes or search for new physics. For each of these pursuits, we usually compare predictions of SM physics processes with what is observed in the real collision data from the detector. The theoretical calculation of a process usually depends on complicated multi-dimensional integrals, which are challenging for analytic calculations. In classical numerical integration, the numerical quadrature rules are the best method for one-dimensional integrals. However, the efficiency of numerical quadrature rules decreases rapidly with the number of dimensions [84]. The Monte Carlo method is a preferred method for integrals in high dimensions. By generating inputs randomly from a probability distribution over the allowed phase space for the physics

process, it numerically evaluates each point in the phase space and derives a result by summing up the whole phase space.

In addition to this mathematical challenge, another complexity of high energy physics calculations comes from multiparticle production. In the first-order approximation, all processes can be described by interactions between the fundamental particles of nature, i.e., quarks, leptons, and gauge bosons. With a proton and an antiproton moving towards each other, a parton from each of the two particles can collide, creating a hard scattering process. This interaction sometimes produces one or more short-lived resonant particles, like the Z^0 or $W^{\pm}$ gauge boson, which then decays to a number of outgoing particles.

However, corrections to this relatively simple hard process make the theoretical calculations much more complex. First, there are higher-order corrections to the hard process that may involve a combination of loop graphs and the soft parts of bremsstrahlung graphs. The necessary perturbative calculations are usually very difficult. Second, there are bremsstrahlung-like modifications, called "parton showers," in which a single particle gives rise to many particles in the final state via radiation. Bremsstrahlung corrections are universal; the exact calculations may be carried out order by order in perturbation theory, but they rapidly become prohibitively complicated. Third, quarks and gluons are confined. For leptons and colorless bosons, the perturbative theory is sufficient for their short-distance interaction. For quarks and gluons, however, it must be complemented with an approximation of the structure of incoming hadrons and the hadronization process.

Hadronization is the process by which colored particles are transformed into jets containing colorless hadrons, photons, and leptons [85]. After hadronization, the hard interaction of the incoming beams results in the production of up to hundreds of outgoing particles. Due to the failure of perturbation theory to describe hadronization, MC event generation is commonly performed in two main stages.

The computation of a hard process is performed by matrix element event generators, which include finite higher-order corrections. The bremsstrahlung and hadronization effects are approximated by parton showering programs. The combination of matrix element generators and showering programs produces MC events that model the Tevatron collision events.

4.1.1 Matrix Element Event Generators

We first describe several matrix element event generators used in this analysis.

4.1.1.1. ALPGEN. ALPGEN [86] is designed for the generation of multiparton hard processes in hadronic collisions at leading-order accuracy, with an emphasis on final states with large jet multiplicities. For large jet multiplicities, the complexity of the matrix element evaluation requires a sophisticated approach. ALPGEN generates events in two steps. In the first step, it performs the parton-level calculation of the matrix elements related to the selected hard process using the ALPHA algorithm [87], generating weighted events. In the second step, it performs the unweighting: using the knowledge of the maximum weight of the sample and the weight of each individual event, unweighted events are generated. ALPGEN allows the veto of shower emissions in regions of phase-space that are already covered by the parton-level configurations, following the so-called MLM prescription [86].

For events with multiple jets in the final state, the jets may be produced in either the matrix element or parton showering stages, creating significant ambiguity in the division of phase space of jet production and the potential for "double counting." Thus, a proper jet-matching algorithm is crucial to ensure proper phase space coverage. This algorithm should dictate which jets should be produced in the hard process and which in the parton shower. The MLM matching [86], as implemented in ALPGEN, allows the event evolution to proceed without restriction and vetoes

events whose hard jets do not match the parton-level quarks or gluon produced in the hard process. In the case that two hard jets are merged by the jet algorithm, a mismatch between the observed and expected jets will trigger a veto unless exactly one of the jets originates from the MC hard process and the other from the parton shower. On the other hand, jets radiated during the parton shower may escape the veto if they are sufficiently close to a jet from the hard process so as to be merged according to the jet algorithm.

4.1.1.2. MADEVENT. MADEVENT [88] is a leading-order multi-purpose event generator which is powered by MADGRAPH [89]. Given a physics process, MADGRAPH automatically generates the amplitudes for all the relevant subprocesses and produces the mapping for the integration over the phase space. This process-dependent information is passed to MADEVENT, and stand-alone code is produced to calculate cross sections and to obtain unweighted events using a multi-channel integration method, which is called the Single-Diagram-Enhanced method [88]. Once the events have been generated, they are passed to parton showering programs and eventually turned into physical states.

4.1.1.3. POWHEG. POWHEG is a framework for interfacing next-to-leading-order (NLO) theoretical calculations with a parton shower program. It was first suggested in Ref. [90] and was described in great detail in Ref. [91]. POWHEG can produce parton-level events with positive (constant) weight at NLO accuracy.[1] Further information about the POWHEG implementation is given in Section 4.2.

[1] These features make POWHEG outstanding compared with MC@NLO [92]. To use MC@NLO at CDF, we need to filter out negative weight events and provide an interface with the HERWIG [93] program.

4.1.2 Showering and Hadronization Event Generators

The hard scattering events simulated by the matrix element event generators are passed to one of several showering programs for parton showering. The parton showering is based on the successive random generation of gluon emissions. The ordering of the emissions is based on certain choices: mass, transverse momentum, or angle. Each emission is generated at a scale lower than the previous one until the cutoff scale is reached. Thereafter, a hadronization model is used to convert the resulting parton into hadrons. We describe the parton showering program PYTHIA used at CDF below.

4.1.2.1. PYTHIA. PYTHIA [85] is a multi-purpose event generator and parton showering program. It contains a coherent set of hard-scattering physics models and a library of hard processes and models for initial- and final-state parton showers, multiple parton-parton interactions, beam remnants, string fragmentation, and particle decays. PYTHIA has a set of utilities and interfaces to external programs. The version of PYTHIA used in this analysis is 6.216, which uses mass-ordered evolution for the initial- and final-state parton showering algorithms. The hadronization of PYTHIA is based solely on the Lund string model [94, 95].

4.1.3 Detector Simulation

A good description of the response of the CDF detector to final-state particles is crucial for the event simulation. The CDF simulation framework mainly consists of two parts, the detector simulation and the trigger simulation.

The detector simulation uses the GEANT3 package [96] to model the detector response of tracks of particles passing through matter. The CDF geometry [97] package provides geometrical and material information of the CDF II detector for GEANT3. In order to properly evaluate the performance of silicon tracking, which is

based on the charge deposition of particles traversing the silicon strips, three charge deposition models are implemented in the silicon simulation: geometric, parametric, and physical. The geometric model is based purely on geometry, while the parametric model is based on the physical effects simulated by GEANT3 [98]. The physical model simulates the physics of charge deposition from first principles [99]. The simulation of the COT detector is based on the GARFIELD package [100], a generic drift chamber simulation program; the default GARFIELD parameters are tuned to describe the CDF data. The calorimeter simulation is based on the GFLASH package [101], a fast simulation of electromagnetic and hadronic particle showers. GFLASH quickly generates a random particle shower and computes the energy deposited in the calorimeter volume. While the GFLASH parameters are tuned using electron test beam data for the electromagnetic calorimeter, they are tuned using high-p_T pion test beam data and $p\bar{p}$ collision data for the hadronic calorimeter [102]. A detailed description of the CDF II detector simulation can be found in Ref. [103].

The trigger simulation emulates the various CDF trigger level decisions. It uses simulated detector data and produces trigger data in an identical manner to that which is used for real data. Because we use a data-driven, turn-on curve technique to model the trigger response in this analysis, we do not use the trigger simulation.

4.1.4 Parton Distribution Functions (PDF)

The hard process in a simulated MC event is the collision of partons from an incoming proton and antiproton in the CDF detector. In the center-of-mass frame of the incoming particle beams, each proton or antiproton beam has a total energy $E = 980$ GeV. The partons that participate in the hard interaction only carry a portion of this energy. Their four-momentum is based on the parton distribution function (PDF) $f_i(x, Q^2)$. This function gives the probability of finding a parton with a flavor i (quark or gluon) in the proton/antiproton that carries a fraction x

of the proton/antiproton momentum, with Q being the energy scale of the hard interaction [104]. Since QCD does not predict the parton content of the proton, the shapes of the PDFs are determined by fitting to experimental observables from data in various processes using the DGLAP evolution equation [105]. Currently, the determination of PDFs is carried out by several groups, namely MSTW, CTEQ, NNPDF, HERAPDF, AB(K)M and GJR. For most of our MC samples, we apply the default PDF set that is used at CDF, CTEQ5L [106]. For the POWHEG samples, we use CTEQ6.1 [107] instead. This is because POWHEG is a NLO generator, and we need to use a NLO-fitted PDF set. Also, the only available tuning for CDF at NLO accuracy is Tune QW, as described in Section 4.1.5, which is tuned based on CTEQ6M (the central value of CTEQ6.1). An illustration of the CTEQ6.1 parton distribution functions is shown in Figure 4.1.

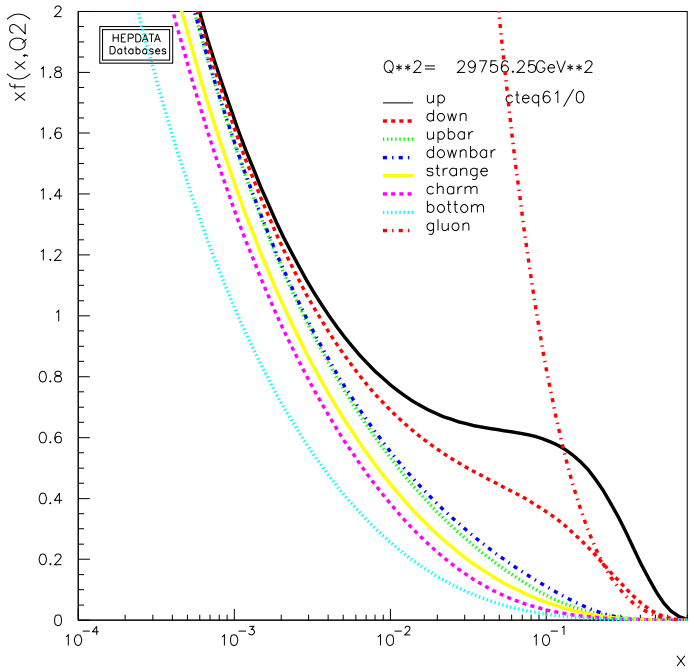


Figure 4.1: An illustration of the CTEQ6.1 parton distribution functions for $Q = 172.5$ GeV. Plotted is the product of the longitudinal momentum fraction x and the distribution function $f(x, Q^2)$ versus x.

4.1.5 Underlying Event

Protons and antiprotons are composite particles. The hard interaction occurs between a parton in the proton and a parton in the antiproton. For a proton beam, for example, a u quark could enter the hard interaction, which will leave behind a beam remnant containing the remaining u and d valence quarks, the gluons, and the sea quarks inside the proton. Since the u quark only carries a fraction of the proton's energy according to the PDF, the beam remnant will take the remaining energy. The probability for interactions among the energetic beam remnant is not zero. Hard or semi-hard scattering can occur between a different pair of partons than those participating in the primary interaction; these are referred to as multiple parton interactions (MPI). Both the beam remnants and MPI contribute to the final event activity, which is called the underlying event (Figure 4.2). The underlying event is an unavoidable background for hadron collider physics. A good understanding of it is important for precise measurements at the Tevatron.

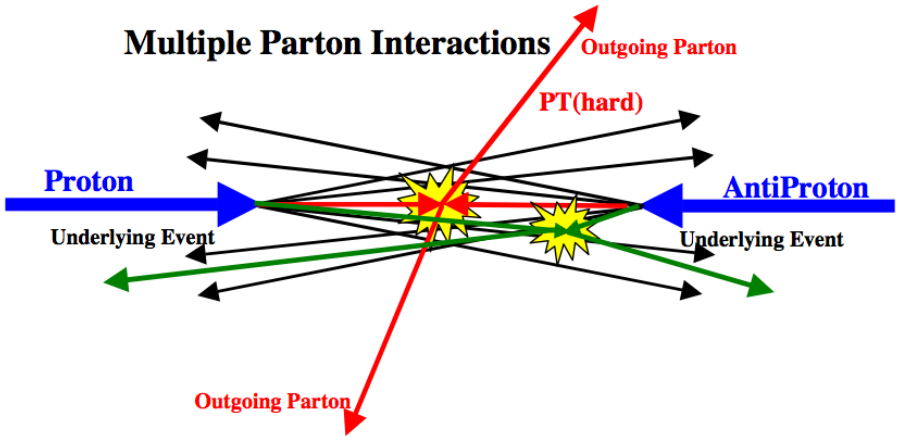


Figure 4.2: An illustration [108] of a hard scattering interaction between a pair of partons from the proton and antiproton, indicated by the red lines pointing towards one another; the resulting high p_T (hard) particles are indicated with red lines pointing outward, away from the collision vertex. All other objects shown in black are the soft $p\bar{p}$ (beam) remnants. A second pair of partons are shown having interacted, indicated by the green arrows moving away from the second vertex. Both beam remnants and multiple parton interactions contribute to the underlying event of the hard scattering process.

Our understanding of the underlying event is still primitive. The description of the underlying event is provided by phenomenological models within PYTHIA, by tuning the PYTHIA parameters in such a way that the MC modeling fits the CDF data. PYTHIA Tune A [109] was tuned to the underlying event measurement in CDF Run I [108]. However, Tune A does not fit the CDF Run I Z boson p_T distribution very well. An improved PYTHIA tuning, Tune AW, fits the Z boson p_T distribution as well as the "underlying event" at the Tevatron [110]. The CDF default tuning is therefore Tune AW using CTEQ5L, which is at LO accuracy. In order to use POWHEG at NLO accuracy, we use Tune QW, which is tuned to Tune AW using CTEQ6.1M in order to maintain consistency.

4.1.6 Pileup

In collisions at high instantaneous luminosity, there is a non-negligible probability that a single bunch crossing may produce more than one interaction; these are called pileup events. In the MC event generation, the luminosity profile of a set of data is passed to PYTHIA, which calculates the instantaneous luminosity of the collected data per bunch crossing. Multiplied by the cross section for pileup processes, PYTHIA obtains the average number of collisions per beam crossing. More than one collision in a beam crossing results in more than one reconstructed vertex in an event. We can compare the number of reconstructed vertices N_{vtx} simulated in MC samples to that which is observed in the data and look for discrepancies due to the synchronization between MC generation and data taking. A reweighting of the N_{vtx} distribution of MC samples to the data is applied in order to model the instantaneous luminosity distribution of the data.

4.1.7 Initial- and Final-State Radiation

In processes that contain colored (charged) objects in the initial or final state, gluon (photon) radiation may give large corrections to the overall topology of events. Hard emission of this kind is important in determining the event structure at the Tevatron. Initial-state radiation (ISR) is the radiation from an initial-state particle before the collision, and is observed as additional objects in the detector that are not produced by the primary interaction. Final-state radiation (FSR) is radiation from a final-state particle in the hard process. The initial- and final-state radiation is shown in Figure 4.3.

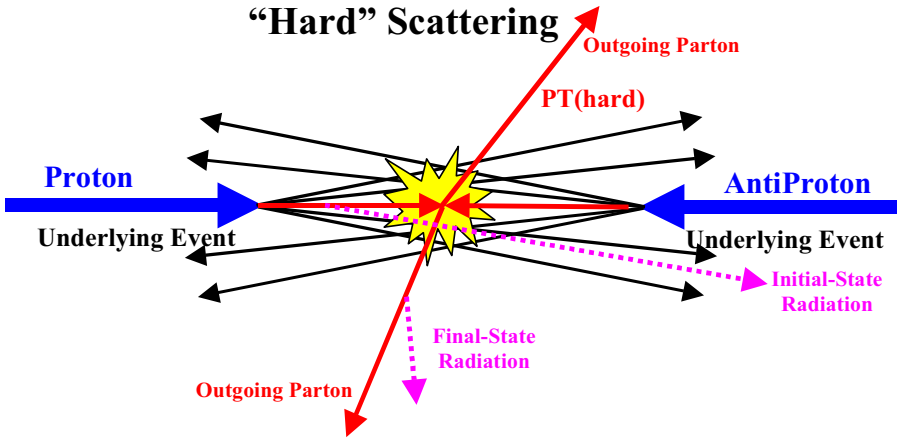


Figure 4.3: An illustration of a hard scattering interaction between a pair of partons from the proton and antiproton, indicated by the red lines pointing towards one another; the resulting high p_T (hard) particles are indicated with red lines pointing outward, away from the collision vertex. Two pink lines indicate the initial-state radiation from the incoming parton inside the proton and the final-state radiation from the outgoing parton. All other objects shown in black are the soft $p\bar{p}$ (beam) remnants.

4.2 Signal Monte Carlo

CDF Run I analyses used PYTHIA to generate single top s- and t-channel MC events. However, it was pointed out that the single top t-channel production modeled by PYTHIA at LO accuracy does not adequately represent the expected distributions of observable jets [111]. Also, the PYTHIA generator did not incorporate

the spin of the top quark in the modeling of top quark decay. Since the top quark is almost 100% polarized in its rest frame, a correct MC description is important for the discrimination of single top quark events against background events. To bypass these problems, the CDF Run II single top analysis uses MADEVENT for single top modeling. MADEVENT calculates the top quark decay at parton level with the spin of the top quark included. It provides the opportunity to generate two different t-channel modes, which can be matched to better model the NLO theoretical expectations; this matching is described in Section 4.2.1. Since such a manual matching procedure is artificial and requires large manpower, we choose POWHEG as the single top MC event generator for this analysis.

4.2.1 t-channel

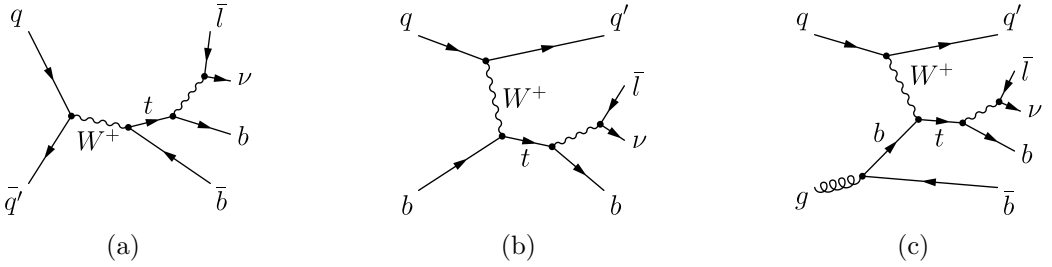


Figure 4.4: Feynman diagrams of single top quark production modes: (a) the s-channel, (b) the t-channel $2 \to 2$ process, and (c) the t-channel $2 \to 3$ process.

The LO t-channel process is a $2 \to 2$ process with a b quark in the initial state, as shown in Figure 4.4(b). In the MC simulation, the energy of the initial-state b quark is calculated from the parton distribution function. Since flavor is conserved in the strong interaction, a $\bar{b}$ quark must be present in the event, which is called the spectator b quark. The charge conjugate processes are implied for antitop quark production.

For LO parton shower programs like PYTHIA, modeling of the t-channel process starts with the LO $2 \to 2$ diagram using a b-quark PDF. Then, the initial state

is created through backward evolution using the DGLAP scheme [112–114]. This method only models well the low-p_T region of the transverse momentum spectrum of the spectator b quark, while the high-p_T tail is underestimated. In addition, the pseudorapidity distribution of the spectator b quark is biased towards higher pseudorapidities when compared with NLO theoretical predictions.

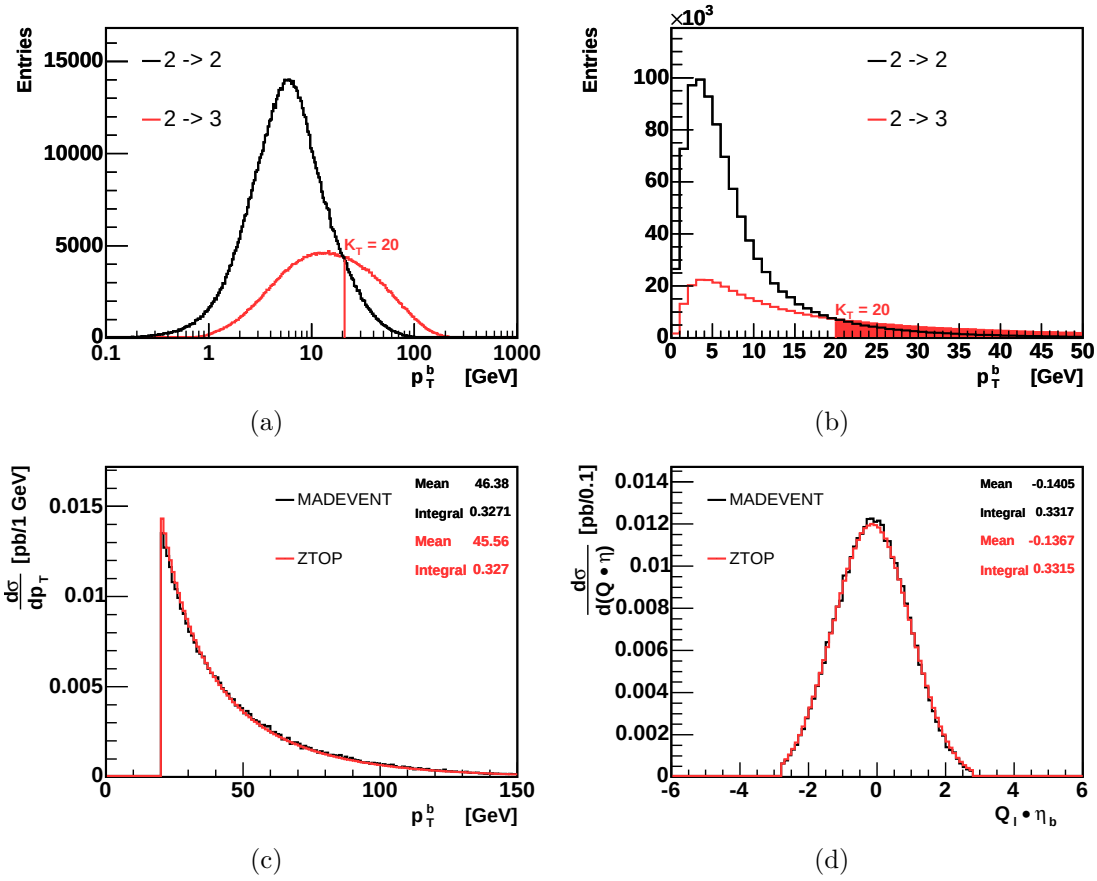


Figure 4.5: Matching of t-channel events in the $2 \to 2$ and the $2 \to 3$ processes [37]. The p_T distributions of the spectator b quark are shown (a) on a logarithmic scale, and (b) on a linear scale. The ratio of $2 \to 2$ to $2 \to 3$ events is adjusted such that the rate of spectator b quarks with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.8$ matches the NLO theoretical prediction of 16.7%. The fraction of these events is illustrated in (b) by the shaded area. The matched MADEVENT sample reproduces both the rate and the shape of the differential ZTOP (c) p_{T} and (d) $Q_\ell \cdot \eta$ cross section distributions of the spectator b quark.

With MADEVENT, such problems can be avoided by performing a matching of both the $2 \to 2$ and $2 \to 3$ matrix element processes. The $2 \to 2$ process from a b-quark PDF is expected to dominate in the low-p_T region of the transverse momen-

tum spectrum of the spectator b quark. The $2 \to 3$ process contains an initial-state gluon splitting into $b\bar{b}$ as illustrated in Figure 4.4(c). The spectator b quark in the $2 \to 3$ process is produced directly in the hard scattering evaluated in the matrix element calculation, which is therefore suitable to describe the high-p_T tail in the spectator b quark p_T distribution. The $2 \to 3$ process describes the most important NLO contribution to t-channel production. In order to correctly model the NLO predictions, a joint MC sample is created by matching the p_T spectrum of the spectator b quark to the differential cross section distribution predicted by the ZTOP program [111]. The ZTOP program provides NLO s- and t-channel theoretical distributions of single top quark production within the geometrical *acceptance* of a given detector. The matched t-channel sample is created such that its spectator b quark p_T spectrum consists of $2 \to 2$ events for p_T values below a cutoff K_T, and $2 \to 3$ events for p_T values above K_T. This is illustrated in Figure 4.5. The ratio of the number of $2 \to 2$ and $2 \to 3$ MC events is adjusted for the continuity of the p_T spectrum of the spectator b quark at K_T. The ratio is also adjusted so that the percentage of events with a detectable spectator b-quark jet (with $p_T > 20$ GeV/c and $|\eta| < 2.8$) matches 16.7%, predicted by ZTOP. With these matching procedures, the ratio $R = 2.1$ and cutoff $K_T = 20$ GeV/c were obtained assuming a top quark mass of 175 GeV/c^2. As a result, all detectable spectator b quarks with $p_T > 20$ GeV/c are simulated using the $2 \to 3$ process while $2 \to 2$ samples are used to describe the spectator b quarks with $p_T < 20$ GeV/c^2. The matched MADEVENT MC events faithfully represent the NLO single top quark production prediction, and the matching procedure for the t-channel sample takes the main NLO effects into account [37]. As shown in Figure 4.5(c) and 4.5(d), both the falling p_T spectrum of the spectator b quark and the slightly asymmetric shape of the $Q \times \eta$ distribution are well modeled when compared with the ZTOP prediction.

Even though the MADEVENT matching sample provides a well modeled t-channel sample, the manual matching procedure is time consuming and artificial. With the availability of the single top process in POWHEG, such a procedure can be avoided. Since POWHEG is an NLO MC generator, we can obtain a more accurate modeling of single top production than with MADEVENT. With the next-to-leading-order accuracy of POWHEG, we can ensure the correct fraction of $2 \to 2$ and $2 \to 3$ processes in the whole physical phase space.

The NLO calculations of t-channel single top production in ZTOP and POWHEG are based on the $2 \to 2$ scattering process, where a b-quark appears in the initial state, which is called the five-flavor (5F) scheme. In the 5F scheme, the mass of the initial b quark has to be neglected in the matrix element calculation. The 5F scheme can greatly simplify the calculation, leading to more compact results. Effects related to the spectator b quark only enter at NLO as the $2 \to 3$ process. An alternative approach is to consider the $2 \to 3$ scattering process as a Born process, keeping a finite b mass as in Ref. [115] performed by the MCFM program. The b quarks do not enter in the QCD evolution of the PDF and the strong coupling, which is called the four-flavor (4F) scheme. The 4F approach involves more calculations of the NLO correction due to the inclusion of an additional parton in the final state. However, the features associated with the kinematic description of the spectator b quark can be properly investigated at NLO accuracy. The production cross section from this approach remains unchanged, but a larger fraction of events have a high-p_T spectator b within the detector acceptance.

With the introduction of a new MC generator for the single top analysis (POWHEG), it is important to validate the modeling of single top quark events. We compare the kinematic distributions of the primary partons obtained from POWHEG against those from the MADEVENT sample and the theoretical differential cross sec-

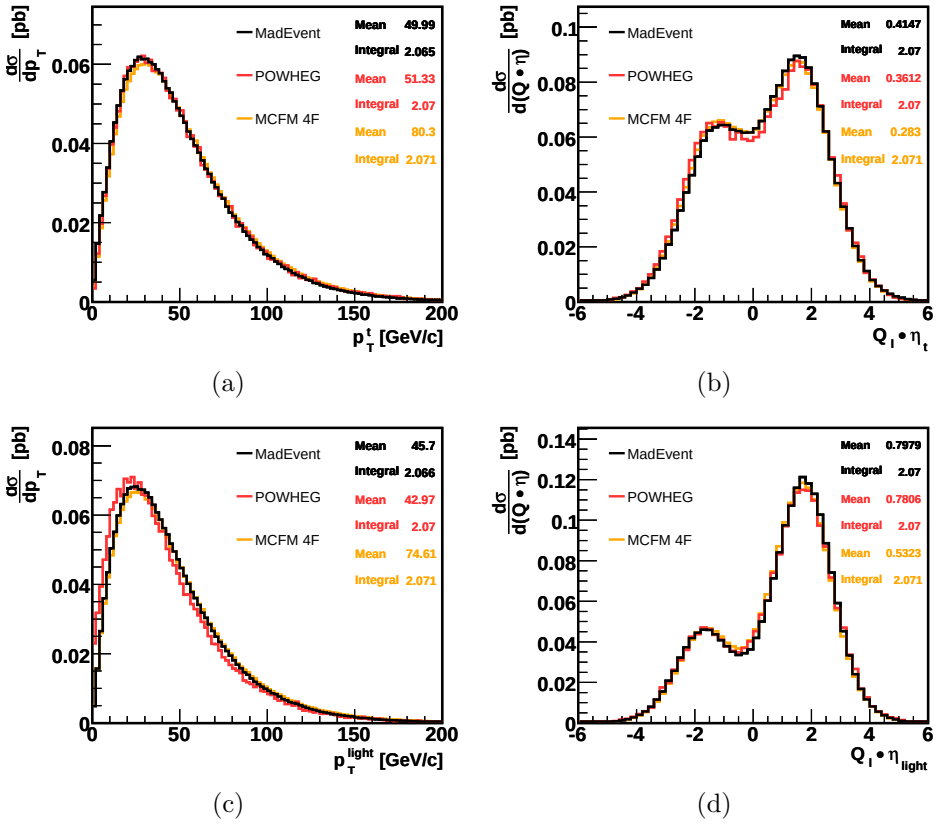
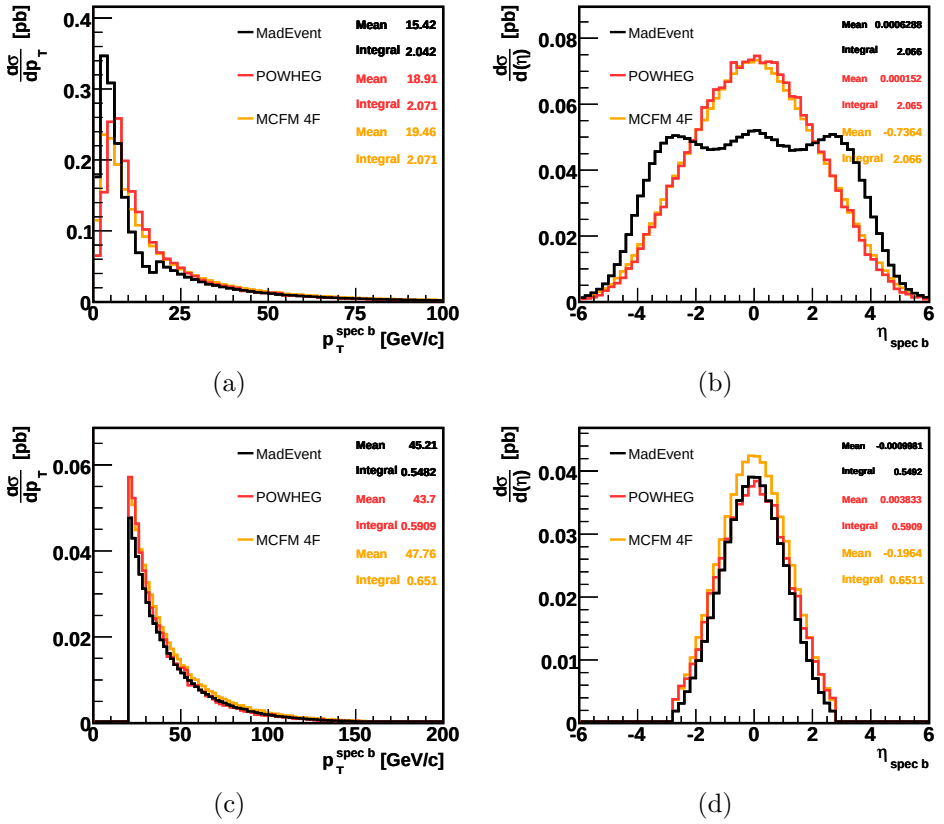


Figure 4.6: Validation of POWHEG t-channel events with the matched MADEVENT sample and the MCFM 4F prediction. The assumed top quark mass is 172.5 GeV/c^2. The top quark p_T and lepton charge multiplied by top quark pseudorapidity ($Q_l \times \eta_{top}$) distributions are shown in (a) and (b). The light jet p_T and lepton charge multiplied by the light flavor jet pseudorapidity ($Q_l \times \eta_{light}$) distributions are shown in (c) and (d). All distributions are from the parton-level MC simulation, without any selection requirements. The rates of the distributions are normalized to the total theoretical cross section.

tion distribution from the MCFM 4F calculation.[2] From Figures 4.6 and 4.7, we find good agreement of the top quark and light jet p_T and $Q \times \eta$ distributions when comparing with MCFM and MADEVENT samples at the parton level. From Figures 4.7(a) and 4.7(b), POWHEG shows an almost identical distribution as the MCFM prediction; the MADEVENT sample is slightly different. The discontinuity in the p_T distribution and the multiple peaks in the η distribution come from the manual matching pro-

[2] Recently POWHEG includes a NLO calculation of single top t-channel production in the 4-flavor scheme [116]. The comparison between the 4- and 5-flavor scheme shows small differences in the acceptance, which is within the theoretical uncertainty. In this analysis, we use the 5-flavor scheme calculation from POWHEG.

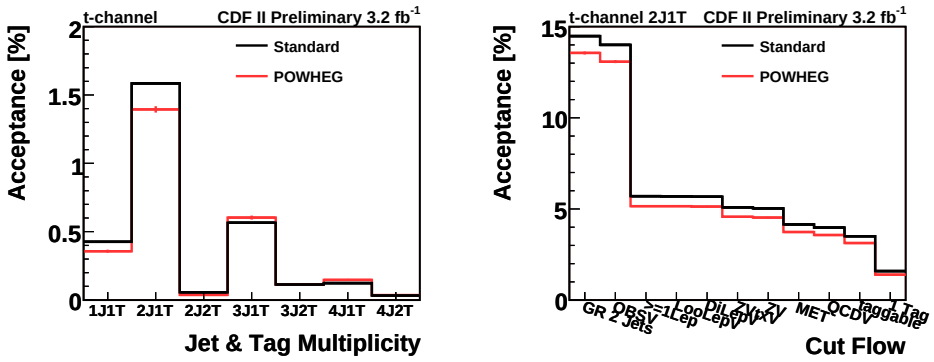


(a) (b)

(c) (d)

Figure 4.7: Validation of POWHEG t-channel events with the matched MADEVENT sample and the MCFM 4F prediction. The assumed top quark mass is 172.5 GeV/c^2. The p_T and η distributions of the spectator b quark at the parton level are shown in (a) and (b). The p_T and η distributions of the spectator b quark at the parton level but within the CDF acceptance are shown in Figure (c) and (d). The rates of the distributions are normalized to the total theoretical cross section.

cedure. However, within the CDF acceptance, as in Figures 4.7(c) and 4.7(d), the comparisons among them are very good. Since we only perform the analysis within the CDF acceptance, the discrepancy of MADEVENT observed at the parton level is not a problem. As we can see, the POWHEG distributions agree with the MCFM prediction for the t-channel process. The acceptance comparisons between POWHEG and the standard MADEVENT samples are shown in Figure 4.2.1. Relative to the standard MADEVENT acceptance, the acceptance predicted by POWHEG is smaller for events with one reconstructed jet that has a b-tag (1J1T) and for events with two reconstructed jets, one of which has been b-tagged (2J1T). The POWHEG acceptance



(a) (b)

Figure 4.8: (a) The acceptance comparison of the POWHEG sample and the standard MADEVENT sample in jet and tag multiplicity. (b) The cut flow of the acceptance comparison of POWHEG and the standard MADEVENT sample in the 2 jet 1 tag channel. The rate is normalized to the acceptance of 2 jet 1 tag bin in (a). The cut flow contains a series of analysis requirements as detailed in Chapter 5.

is larger than the standard acceptance for events with 3 jets and 1 b-tag (3J1T) and events with 4 jets and 1 b tag (4J1T). This is expected from the POWHEG sample due to the NLO effects; we expect a higher jet multiplicity due to the inclusion of NLO corrections. The cut flow of acceptance in Figure 4.8(b) shows the acceptance changes for different consecutive cuts in the 2J1T channel. The overall acceptance of POWHEG is lower than the standard MADEVENT sample as shown in Figure 4.2.1. Both samples respond to each acceptance cut similarly, however.

4.2.2 s-channel

The s-channel production mode comes from $q\bar{q}$ annihilation as Figure 4.4(a) illustrates. A calculation from the ZTOP program predicts at least one jet should be visible within the detector acceptance in about 90% of the s-channel events. In most cases, this jet is the leading jet in the event (the jet with the highest p_T), and is the b-quark jet not coming from the top decay; this jet is referred to as spectator b jet in this thesis, similar to the t-channel spectator b jet. According to Ref. [111], no significant NLO corrections affect the kinematic distribution of s-channel events.

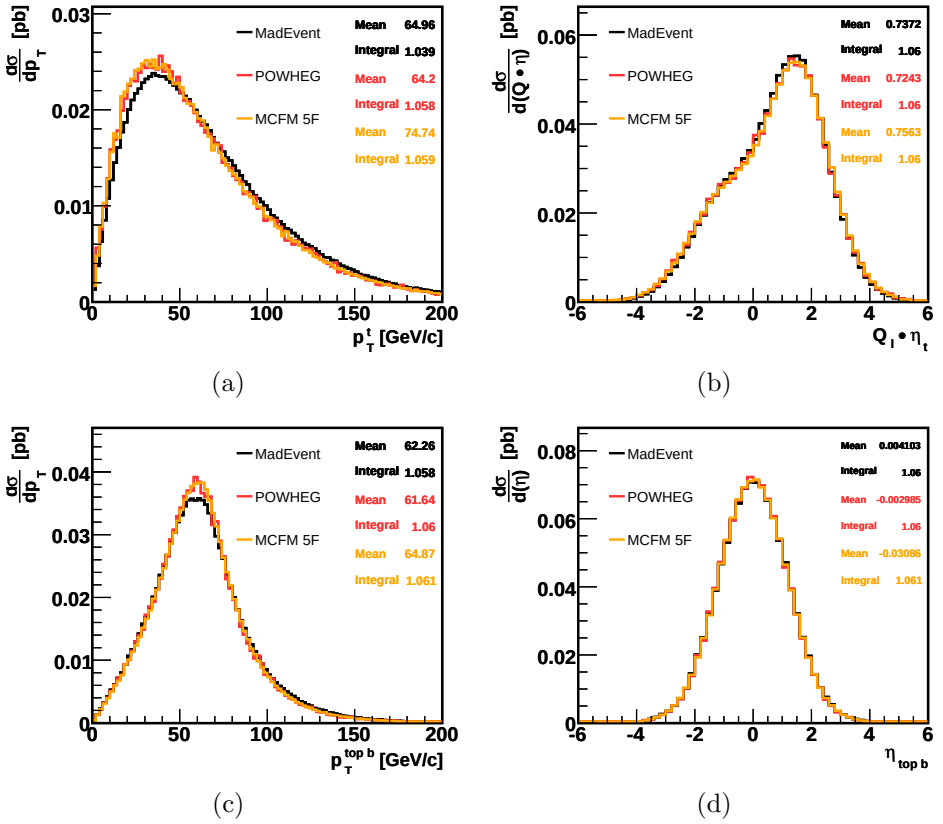


Figure 4.9: Validation of POWHEG s-channel events with the MADEVENT sample and MCFM 5F prediction. The assumed top quark mass is 172.5 GeV/c^2. The p_T and $Q_{lepton} \times \eta_{top}$ distributions of the top quark are shown in (a) and (b). The p_T and η distributions of the b quark from top decay are shown in (c) and (d). The rates of the distributions are normalized to the total theoretical cross section.

We compare the kinematic distributions of the s-channel events obtained from POWHEG against both the MADEVENT sample and the theoretical differential cross section distribution from MCFM. From Figure 4.9, we find that the top quark properties before the CDF acceptance cuts are well modeled by POWHEG when comparing with the MCFM prediction. The MADEVENT prediction, on the other hand, tends to produce a harder top quark (top quark with higher than average p_T). In Figure 4.10, the p_T distribution of the spectator b quark from POWHEG is softer than the MCFM and standard MADEVENT predictions. POWHEG tends to have more radiation, compared to MADEVENT, through additional NLO corrections. The spectator b jet is expected to be more sensitive to these corrections and to be softer in the LO calcula-

tion. The differences between the POWHEG and MCFM distributions, however, points to the clustering algorithm exploited within MCFM. The MCFM calculation [117] uses a midpoint algorithm to cluster and calculate the spectator b-jet kinematic distributions. Because we do not have other information from the MCFM calculations (MCFM only produces kinematic distributions) and MADEVENT and MCFM have similar distributions, we studied the additional radiation effects by comparing the POWHEG and MADEVENT samples. k_T clustering [118] is performed on all the hadron-level particles (excluding the leptons) in both the POWHEG and MADEVENT samples and the resulting jets are compared. The comparison of the hadron-level jets excludes the impact of detector response and different jet clustering algorithms on the final detectable jets. We found that the most energetic jet (mostly the b jet from top decay) and second-most energetic jet (mostly the spectator b jet) of POWHEG and MADEVENT agree very well. The POWHEG sample, however, has harder and more third-most energetic jets (mostly jets from NLO radiation) compared to the MADEVENT sample. This discrepancy between POWHEG and MCFM is taken into account as a systematic uncertainty, as Section 8.1.4 describes. The acceptance comparisons between POWHEG and the standard MADEVENT samples are shown in Figure 4.11. The overall acceptance of POWHEG is lower than the standard MADEVENT sample as expected due to NLO corrections. Both samples respond to each acceptance cut similarly, however.

4.2.3 Wt-channel

The Wt-channel involves the production of a top quark in association with a W boson. At the Tevatron, the Wt-channel contribution is negligible and was ignored in the previous single top analyses. However, with the good performance of the Tevatron and the CDF detector, we have more than doubled the dataset used in the previous analyses. Also, with a better understanding of the CDF data from CDF

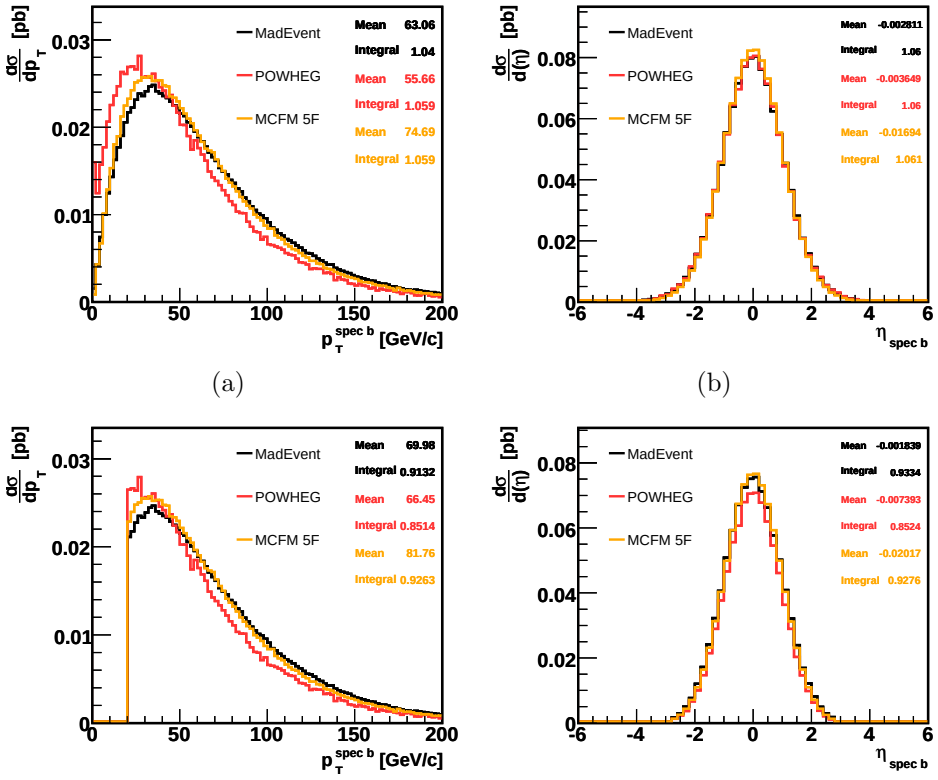


Figure 4.10: Validation of POWHEG s-channel events with the MADEVENT sample and MCFM 5F prediction. The assumed top quark mass is 172.5 GeV/c^2. The p_T and $Q_{lepton} \times \eta_{top}$ distributions of the spectator b quark are shown in (a) and (b). The same distributions at parton level within the CDF acceptance are shown in (c) and (d). The rates of the distributions are normalized to the total theoretical cross section.

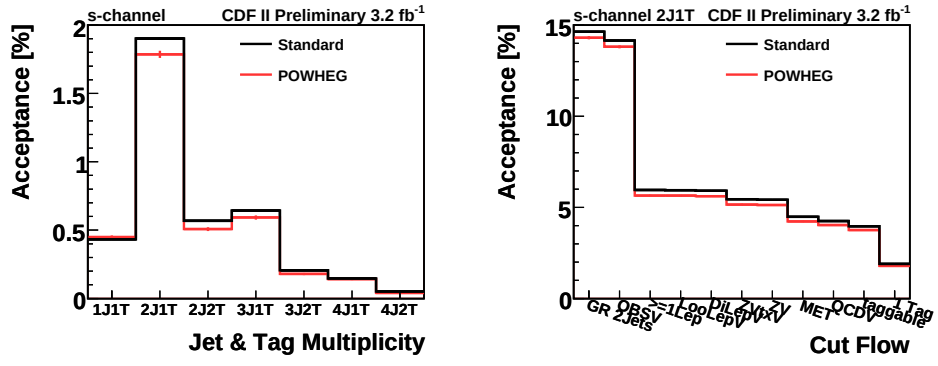


Figure 4.11: (a) The acceptance comparison of the POWHEG sample and the standard MADEVENT sample in jet and tag multiplicity. (b) The cut flow of the acceptance comparison of POWHEG and the standard MADEVENT sample in the 2 jet 1 tag channel. The rate is normalized to the acceptance of 2 jet 1 tag bin in (a).

Higgs boson searches, the expected ability to observe this process has increased. It is also important for us to start considering such a process in the single top analysis.

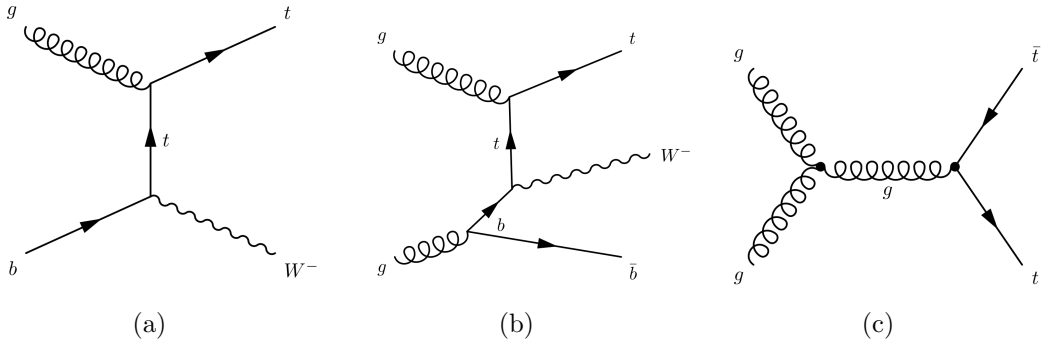


Figure 4.12: Feynman diagrams of single top quark Wt-channel production modes: (a) leading order and (b) next-to-leading order with initial gluon splitting. (c) The gluon fusion $t\bar{t}$ production Feynman diagram.

The LO Feynman diagram for Wt-channel production is provided in Figure 4.12(a). As it shares a similar initial state as the t-channel, the NLO correction of the Wt-channel introduces an initial gluon splitting as shown in Figure 4.12(b). However, the theoretical NLO corrections to the Wt-channel are not well defined due to interference effects with the $t\bar{t}$ process. The $t\bar{t}$ process in Figure 4.12(c) has a final state that corresponds to $t\bar{t}$ production, followed by the decay $\bar{t} \to W^-\bar{b}$. The consequence of this is the problem of interference between Wt and $t\bar{t}$ production: the top quark propagator goes on shell when the invariant mass of the system made by the W-boson and the outgoing $\bar{b}$-quark ($m_{W\bar{b}}$) approaches m_t. The interference effects are present when summing up all the Wt-channel diagrams at the amplitude level. At this level, this interference between the lowest-order $t\bar{t}$ production (followed by a decay) and the NLO real corrections to the Wt process become numerically non-negligible. The perturbative expansion for the NLO corrections to the single top Wt-channel loses its meaning.

Several approaches for this problem have been introduced [119–121]. POWHEG used the same strategy as first described in the corresponding MC@NLO publication [122], which was presented in [109]. To deal with the interference problem, two

definitions for the NLO corrections are considered in POWHEG, which are known as *Diagram Removal* (DR) and *Diagram Subtraction* (DS). These definitions rely on the observation that a meaningful definition of the Wt-channel process is possible only through cuts on the final-state objects in the analysis. If interference effects with $t\bar{t}$ are negligible after these cuts are applied, then it is possible to consider Wt-channel as a well defined process. Since cuts act differently in different phase space regions, one needs to quantify the interference between Wt and $t\bar{t}$ locally in that phase space. In this analysis, since the cross section of $t\bar{t}$ is almost 30 times that of the Wt-channel, our analysis is not sensitive to this interference phase space of the Wt-channel signal. Thus, it is possible to implement such approaches in our analysis.

The DR method drastically considers the top quarks as undetectable particles. It would remove any inference issue by eliminating the $t\bar{t}$ contribution from the Wt-channel Feynman diagrams before squaring the amplitude. Such a method is not gauge invariant. It has been shown, however, that the impact of gauge dependence in the DR calculation is numerically negligible [122]. The DS method gives a prescription for removing the contamination from $t\bar{t}$-like contributions, while keeping the top quark as a final-state particle. By building a gauge invariant subtraction term that exactly cancels the $t\bar{t}$ contribution when the interference region is approached, it keeps the full squared amplitude and remains gauge invariant. Since this is the first time we introduce Wt-channel samples at CDF, we cannot validate it with previous samples. However, we compare the samples from the two methods and find that they are almost identical within the CDF acceptance. This is shown in Figure 4.13.

The small difference between these two methods is taken as the MC systematic uncertainty for the Wt-channel process as suggested by POWHEG. The relative dif-

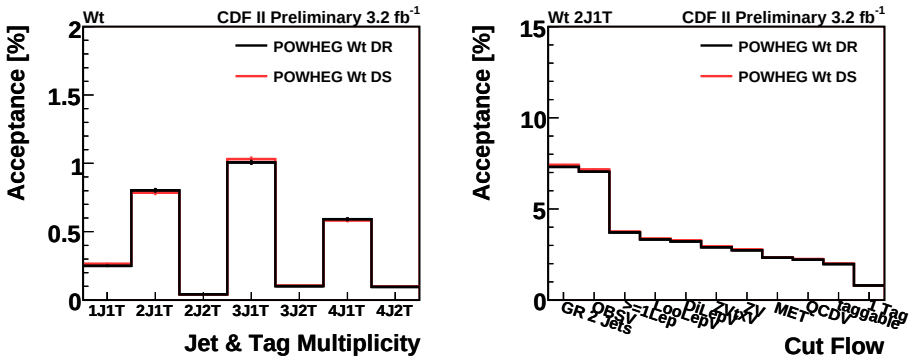


Figure 4.13: The acceptance comparison of DS and DR samples in the POWHEG Wt-channel process. Both samples respond similarly to the CDF acceptance.

ference between them can be interpreted as a measure of the theoretical uncertainty in the definition of genuine NLO corrections to the Wt-channel process [109].

4.2.4 POWHEG *Validation with* $t\bar{t}$

In order to better understand the POWHEG generator, we also study a $t\bar{t}$ sample from POWHEG. Since the $t\bar{t}$ process has been carefully studied for decades at CDF and has led to a precise measurement of the top mass, it is well understood. Also, the $t\bar{t}$ process provides various top kinematics and features that have been well studied at CDF. The POWHEG single top processes share the same implementation and framework as the $t\bar{t}$ process. Comparing the $t\bar{t}$ sample from POWHEG with the standard CDF sample from PYTHIA will help us understand our single top sample. Both samples are made assuming a top mass of 172.5 GeV/c^2.

As shown in Figure 4.15, the comparison of the $t\bar{t}$ sample from POWHEG shows good agreement with the standard CDF PYTHIA sample. The POWHEG sample produces a lower acceptance in events with a small number of jets and a higher acceptance in events with a larger number of jets than the standard sample. This is expected when comparing the NLO calculation to the LO calculation. With the successful implementation of the POWHEG $t\bar{t}$ process at CDF, it is used in several analyses at CDF, such as the recent $t\bar{t}$ forward-backward asymmetry measurement [24].

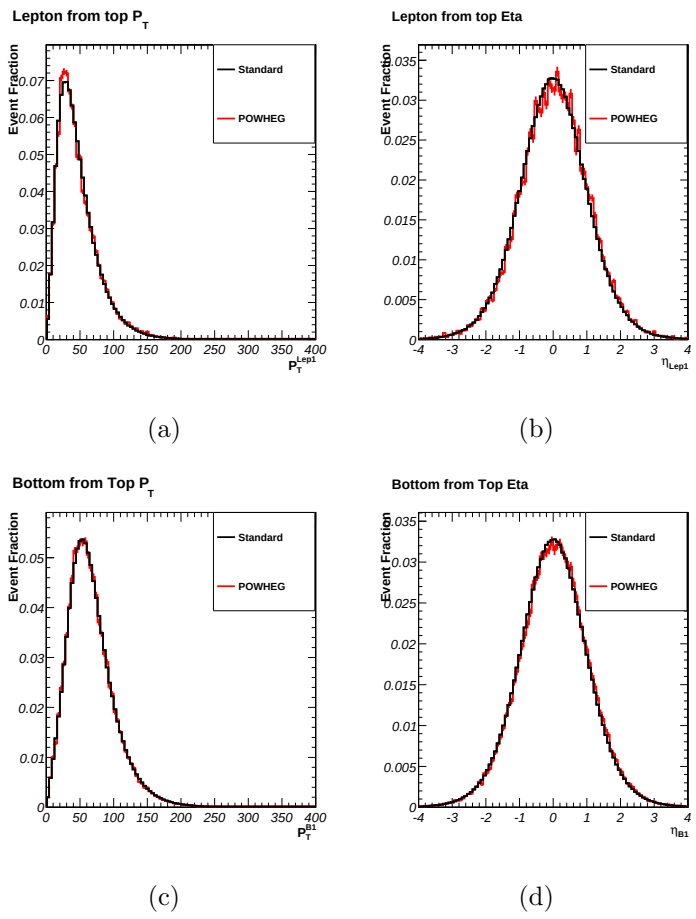


Figure 4.14: Comparison of the POWHEG $t\bar{t}$ sample with the standard CDF PYTHIA sample. The (a) p_T and (b) η distribution of the lepton from the top quark decay at parton level. The (c) p_T and (d) η distribution of the b quark from the top quark decay at parton level.

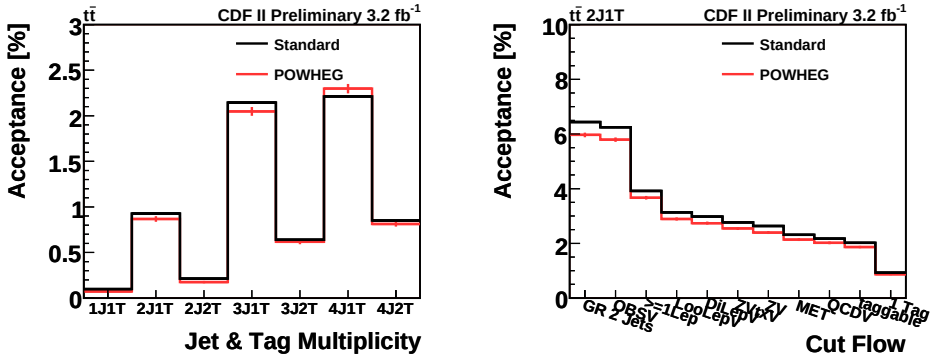


Figure 4.15: (a) The acceptance comparison of the POWHEG $t\bar{t}$ sample and the standard CDF PYTHIA sample in jet and tag multiplicity. (b) The cut flow of the acceptance comparison of POWHEG and the standard sample in the 2 jet 1 tag channel. The rate is normalized to the acceptance of 2 jet 1 tag bin in (a).

One of the challenges of this analysis is that we are measuring the cross section of a small signal among many large backgrounds. The final state of a single top quark signal event — a charged lepton, large missing transverse energy from the undetected neutrino, and two or three jets with one or more from b-quark decays — is also the final state of the $Wb\bar{b}$ process, which has a much larger cross section. Other processes, such as $t\bar{t}$ or $Wc\bar{c}$, also mimic the single top quark signature due to either the misreconstruction of an object or the failure to reconstruct particles in the expected final state. In order to accurately measure the single top quark production cross section, a detailed understanding of the rates and kinematic properties of such background processes is crucial.

Table 4.1: MC event generators and parton showering software programs used in this analysis to simulate events for signal and background processes.

Process	Event Generator	Parton Showering
Single Top	POWHEG	PYTHIA
$t\bar{t}$	PYTHIA	PYTHIA
Diboson	PYTHIA	PYTHIA
W + jets	ALPGEN	PYTHIA
Z + jets	ALPGEN	PYTHIA

The background processes to single top production include $t\bar{t}$ production, W + jets production, Z + jets production, diboson production, and QCD multijet events. Except QCD multijet, the event generators used to model the signal and the different backgrounds in the current analysis are summarized in Table 4.1. For QCD events, it is difficult to simulate all the QCD backgrounds with MC programs. For a QCD event lacking a leptonically decaying W boson to fake the lepton + jets signature, it must have a jet that manages to fake a lepton or a real lepton from a heavy flavor quark decay. Simultaneously, the energies of the jets must be badly measured so that the event has a large $\not\!\!E_T$. Even though the probability of these events occurring is very small, with the extremely high cross section of QCD

processes, it is one of the dominant backgrounds. Considering the variety of QCD processes and the low fake rate, a reliable model of the QCD background requires a tremendous number of MC events, which is not practical at CDF. We exploit a different method for QCD modeling as described in Section 6.3.

4.3.1 Top Quark Pair Production

We use PYTHIA to model the $t\bar{t}$ production contribution assuming $m_{\rm top} = 172.5$ GeV/c^2. The $t\bar{t}$ events can be classified by the decay mode of the W bosons from the top quark decay. We define lepton + jets events as events in which one W boson decays leptonically and the other decays hadronically. This is shown in Figure 4.16(a). If both W bosons decay leptonically, the event is called a *dilepton* event (Figure 4.16(b)). Events in which both W bosons decay hadronically are named *all-hadronic* events. The dilepton and all-hadronic events can pass the single top selection for various reasons: either one lepton is lost in dilepton events, or one jet is lost and one jet is misidentified as a lepton in all-hadronic events.

In the previous analysis [37], because the MC efficiencies for lepton identification and b-tagging differ from that observed in the data, a scale factor ε_{corr} is applied to the $t\bar{t}$ efficiencies estimated from the MC samples. The scale factors are functions of the number of leptonically decaying W bosons and of the number of b-tagged jets.

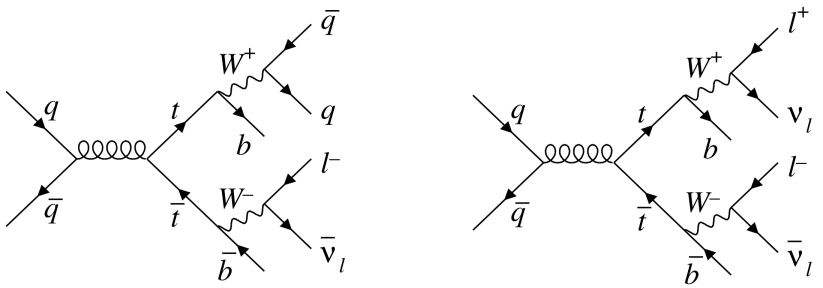
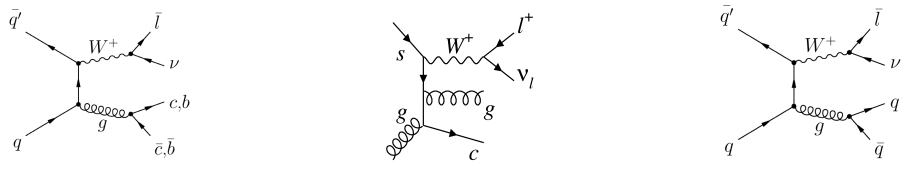


Figure 4.16: LO Feynman diagrams for top quark pair production and decay, which are irreducible background processes for this analysis.

From our studies, the scale factor for the number of leptonically decaying W bosons in the $t\bar{t}$ sample is negligible. We therefore do not apply this scale factor to the $t\bar{t}$ MC sample.

4.3.2 W + Jets Production

The largest background process in the single top analysis is the associated production of a leptonically decaying W boson with two or more jets. The cross section for W + jets production is much larger than the single top signal, and its kinematic properties can be quite similar to the signal events if more than one jet is from a b quark. This final state can be identical to that in single top production. Depending on the response to the SECVTX algorithm, W + jets events are separated into two categories. W + heavy flavor (W + HF) production refers to the events with a W boson accompanied by heavy flavor quarks (c or b quarks). It contains $Wb\bar{b}$ and $Wc\bar{c}$ as shown in Figure 4.17(a), and Wcj, which is shown in Figure 4.17(c). Because c quarks have a significantly long lifetime, charm production can lead to the reconstruction of a real secondary vertex. The W + light flavor processes (W + LF), as shown in Figure 4.17(c), contain jets originating from light flavor partons. Due to its large production rate, W + LF processes make a significant contribution to the background when jets are mis-tagged as b-jets. At CDF, the W + jets samples are generated by ALPGEN and showered by PYTHIA.

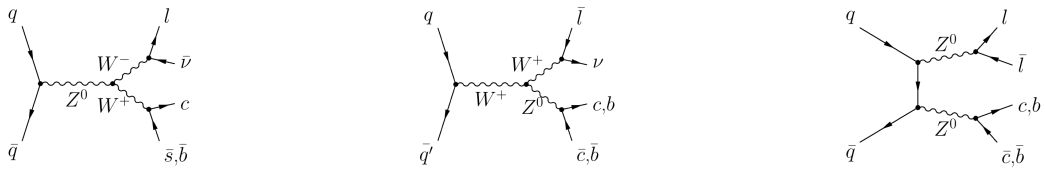


(a) (b) (c)

Figure 4.17: Feynman diagrams for several W + jets production and decay modes: (a) $Wb\bar{b}$ and $Wc\bar{c}$, (b) Wcj, and (c) W + LF.

4.3.3 Diboson Production

The diboson production modeled by ALPGEN contains WW (Figure 4.18(a)), WZ (Figure 4.18(b)), and ZZ (Figure 4.18(c)) processes. The WW and WZ processes contribute to the selected data sample due to their final states including a lepton, a neutrino, and heavy-flavor jets. ZZ events mimic a lepton + jets signature when one of the two leptons is lost, faking a neutrino through the missing transverse energy of the lepton.



(a) (b) (c)

Figure 4.18. LO Feynman diagrams for the (a) WW, (b) WZ, and (c) ZZ processes.

4.3.4 Z + Jets Production

Z + jets production contains Z + heavy flavor and Z + light flavor, and the Z boson decays to leptons. Even though its production rate is large, the selection of only one lepton in this analysis strongly suppresses its contribution to the background. The main contribution comes from $Zb\bar{b}$ and $Zc\bar{c}$ production and from $Z \to \tau\bar{\tau}$, as shown in Figure 4.19. Similar to ZZ production, the Z + heavy flavor modes only contribute when one of the two leptons is lost.

(a) (b)

Figure 4.19: LO Feynman diagrams for Z + jets production in (a) heavy flavor and (b) light flavor.

CHAPTER FIVE

Event Selection

In this analysis, we are interested in the production of a single top quark that decays to a b quark and a W boson, where the W boson then decays leptonically. As shown in Figure 5.1, this production and decay process yields a final-state topology that contains two or three jets in addition to a charged lepton and a neutrino from the W boson decay.

The requirement that the W boson decays into a charged lepton (e or μ) and a neutrino (ν_e or ν_μ) reduces the QCD background. We do not select events in which the charged lepton is a τ particle due to the fact that the τ identification efficiency at CDF is very low. However, a small fraction of these τ events, in which the τ decays leptonically, is included in our event selection even if the τ is not directly identified. Our selection also requires two or three energetic jets in the final state, where at least one of those jets is b *tagged*, i.e. is identified by the SECVTX algorithm as coming from a hadronized b quark.

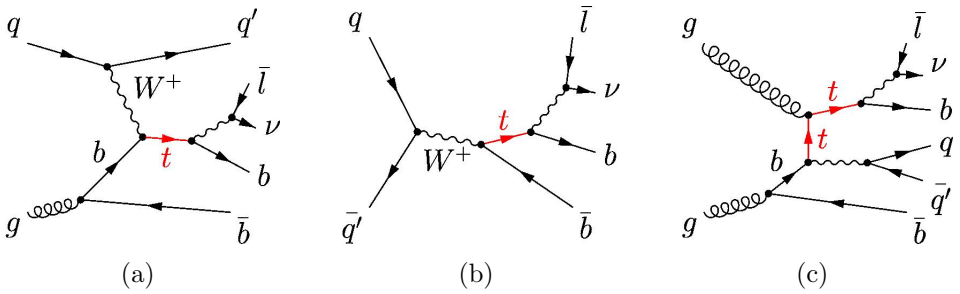


Figure 5.1: Feynman diagrams for (a) NLO t-channel, (b) LO s-channel, and (c) NLO Wt-channel single top production with successive $t \to Wb$ and $W \to l^+\nu_l$ leptonic decays. Charge conjugation is applied for single $\bar{t}$ production.

5.1 Data Sample and Trigger

The signature of the leptonically decaying W bosons is the appearance of a charged lepton (electron or muon) and large missing transverse energy ($\not\!\!E_T$). Thus, it is sensible to trigger on the presence of leptons during the data collection. We use data taken from the high-p_T central electron data stream (*bhel*), high-p_T central muon data stream (*bhmu*), high-p_T plug electron stream (*bpel*), and high-$\not\!\!E_T$ data stream (*emet*). Each data stream contains one or more triggers. Each trigger is composed of requirements in all three trigger levels in the CDF trigger system, with strict quality requirements imposed. We use CDF collision data taken from February 2002 through March 2011, corresponding to a total integrated luminosity of 7.5 fb^{-1}.

5.1.1 Central (CEM) Electrons

Central (CEM) electrons need to pass the ELECTRON_CENTRAL_18 trigger, which requires a COT track with $p_T > 9$ GeV/c matched to an energy cluster in the CEM calorimeter with $E_T > 18$ GeV. The shower profile of such a cluster, measured by the shower maximum detector, has to be consistent with expectations obtained using test-beam electrons. The ratio of energy deposited in the hadronic calorimeter to the energy in the EM calorimeter ($E_{\rm HAD}/E_{\rm EM}$) must be less than 0.125.

5.1.2 Forward (PHX) Electrons

Forward (PHX) electron candidates have $1.2 < |\eta| < 2.0$; they are also called plug electrons (see Section 3.1.1) and are collected by the MET_PEM trigger. This trigger requires a transverse cluster with an energy deposition greater than 20 GeV in the PEM and $E_{\rm HAD}/E_{\rm EM} < 0.125$, since no tracking is available for this class of electron. The trigger also fires on the presence of large missing energy. Due to the geometry of the PEM, the MET_PEM trigger has higher background from "beam splash" (particles from elastic collisions and collision remnants), which tends

to happen in the forward region. In addition, the lack of COT tracking information reduces the sample purity relative to the central region of the detector. Thus, tighter selection cuts are applied for PHX electrons, as described in Section 5.2.1.2.

Since the MET_PEM triggers on raw $\not\!\!E_T$ during data collection, the corresponding data sample would be biased with respect to the corrected $\not\!\!E_T$, as described in Section 3.5. For example, the corrected $\not\!\!E_T$ might be lower than the $\not\!\!E_T$ requirement of the trigger. In order to compensate for this bias, all Monte Carlo PHX events are sculpted by a *turn-on* function to match the data. An additional correction to the plug electron energy is also applied. The turn-on function is parameterized for both the corrected missing transverse energy and plug electron energy. The parametrized turn-on curve gives a weight for MC events with an electron in the plug region.

5.1.3 Central (CMUP) Muons

Muons collected by the MUON_CMUP18 trigger are identified as central (CMUP) muons. This trigger requires a track in the COT with $p_T > 18$ GeV/c that has to match track segments in both central muon chambers (CMU and CMP) simultaneously.

5.1.4 Forward (CMX) Muons

Forward (CMX) muon candidates have $0.65 < |\eta| < 1.0$ and are collected by the MUON_CMX18 trigger, which requires a COT track segment matched to hits in the CMX muon chambers. While CMUP muons are identified in two muon subdetectors, CMX candidates are reconstructed from only one hit in the CMX detector, which has no steel between itself and the calorimeter. Thus, the CMX trigger is not as pure as the CMUP trigger. Even with the requirement that the timing signal

from CSX scintillators be consistent with particles coming from Tevatron collisions, the CMX trigger still gives a relatively high rate of fake muon particles.

5.1.5 Extended (EMC) Muons and ISOTRK

Extended Muon Coverage (EMC) is an inclusive category of muon candidates used to increase acceptance for events that do not fire a muon trigger. EMC muons are collected from $\not\!\!E_T$ + jets trigger paths, containing six mutually exclusive sub-categories: CMU, CMP, BMU, CMIO, SCMIO, and CMXNT. The previous single top analysis [37] only used one trigger (MET2J) for collecting EMC events. In this analysis, we employ a novel trigger combination that uses the MET2J, MET45, and METDI triggers [123] for collecting EMC events. We also add a new category (ISOTRK) of muon candidates into the EMC. In the following, we describe the $\not\!\!E_T$-based triggers used to collect EMC muons. Detailed descriptions of the EMC muons and ISOTRK are given in Sections 5.2.1.3 and 5.2.1.4.

There are three $\not\!\!E_T$-based triggers: MET2J, MET45, and METDI. The MET2J trigger contains 4 versions, based on how the trigger has evolved over time. Only one version is active in a given run. The trigger fires if there is large missing transverse energy ($\not\!\!E_T > 35$ GeV) as well as two jets ($E_T > 25$ GeV, $\Delta R > 1.0$) in the event. The second version added the requirement that one of the jets is central ($|\eta^{\text{jet}}_{\text{det}}| < 0.9$). The third version added the requirement that when the instantaneous luminosity is above a certain threshold the trigger will be prescaled. The fourth version added the dynamic prescale.

The MET45 trigger comes in two versions (MET40 and MET45) for different running periods due to the higher instantaneous luminosity. They both require higher $\not\!\!E_T$ values than the MET2J trigger as their names suggested. The MET45 trigger, however, does not require any jet information.

The METDI trigger comes in only one version. It is very similar to MET2J, requiring large transverse missing energy and two jets. The trigger requirements were designed especially to improve the Higgs boson signal and it is more efficient than MET2J.

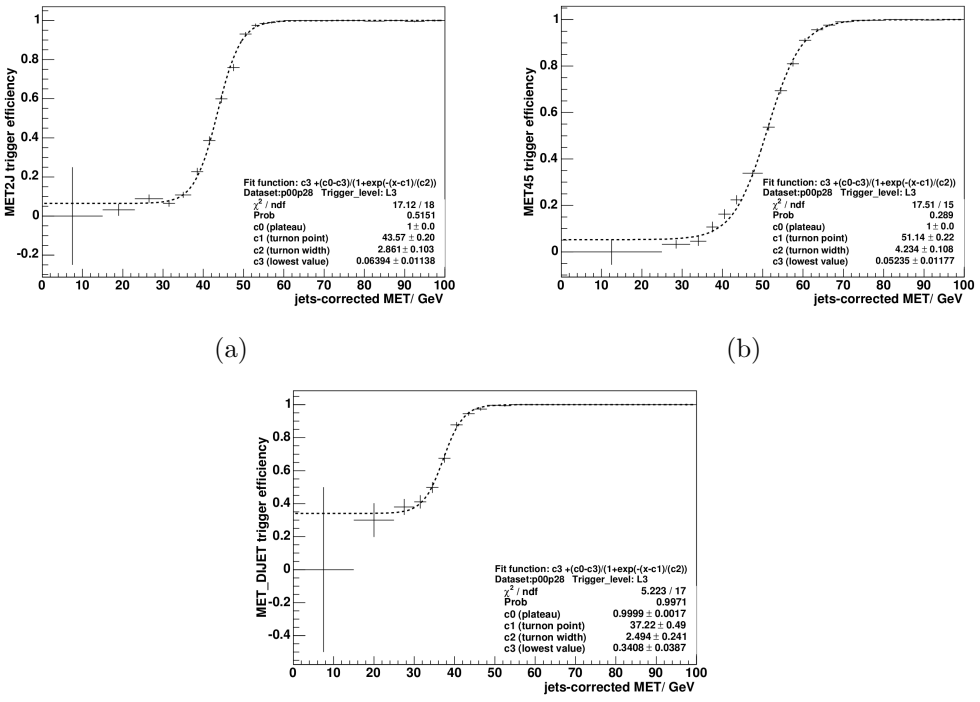


Figure 5.2: The (a) MET2J, (b) MET45, and (c) METDI trigger turn-on curves, parametrized as a function of $\not\!\!E_T$, calculated at Level 3 of the trigger, which is closest to the $\not\!\!E_T$ value in the original collected data. These figures are taken from [123].

In order to combine these triggers and parametrize their efficiency, we use a sigmoid (turn-on) curve as function a of $\not\!\!E_T$, without correcting for the muon momenta. The resulting turn-on curves for the three triggers are shown in Figure 5.2. A sophisticated method [124] is developed for combining these triggers. For an event in which more than one trigger fired, the probability of passing a specific trigger is computed (called the "trigger weight") for each trigger fired. The computation also takes into account the prescale factor for each trigger. The event is defined as belonging to the trigger with the largest "trigger weight."

For events collected by one of the triggers, further requirements are imposed on the reconstructed physics objects of Section 3 to increase the purity of the samples. "Tight" and "loose" physics object candidates are defined by the strength of the quality requirements upon the reconstructed objects. We describe these requirements below.

5.2.1 Lepton Selection

In order to gain as many single top event candidates as possible, we exploit all lepton categories available at CDF in this analysis. Based on the trigger, we distinguish electrons (CEM, PHX), high-quality muons (CMUP, CMX), extended muons (EMC), and isolated tracks (ISOTRK). Their distributions in the $\eta - \phi$ plane are shown in Figure 5.3.

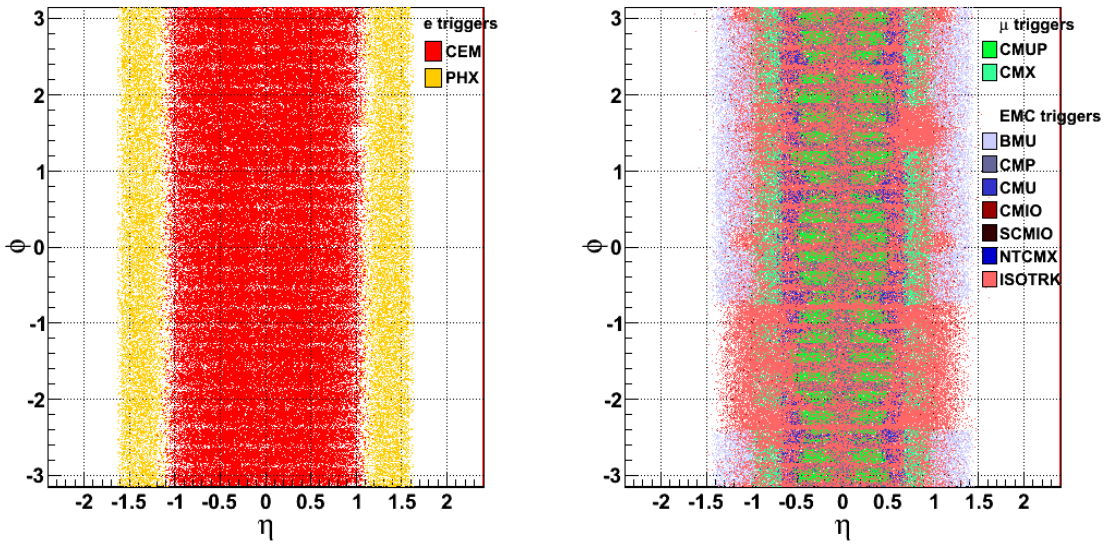


Figure 5.3: The leptonic acceptance distribution in the $\eta - \phi$ plane, showing the coverage of the detector provided by each lepton category.

5.2.1.1. *Central electrons.* The largest sample of candidate events in this analysis comes from those with central (CEM) electrons. A CEM candidate has an

energy cluster in the central calorimeter matched to an extrapolated track from the COT. The selection cuts are given in Table 5.1 and a description of each variable is provided below.

Table 5.1. Selection requirements for central electrons.

Variable	Selection Requirement
Geometry	Fiducial in CEM
E_T	> 20 GeV
p_T	> 10 GeV/c
$E_{\rm HAD}/E_{\rm EM}$	$< 0.055 + 0.00045 \times$ E
Isolation	< 0.1
E/p	< 2 if $p_T < 50$ GeV/c
$Q \times \Delta x$	< 1.5 cm and > -3.0 cm
$\lvert\Delta z\rvert$	< 3.0 cm
$\chi^2_{\rm strip}$	< 10
$L_{\rm shr}$	< 2.0
COT track quality	Pass COT track requirements
Conversion veto	Pass conversion veto

- *Geometry* The electron candidate must have a cluster in the fiducial region of the CEM, which means the well-instrumented region that can be triggered on with good efficiency. A cluster near a crack in the calorimeter will be removed. The fiducial region of the CEM is $|\eta| < 1.1$.

- E_T The transverse energy of the cluster, which is required to be larger than 18 GeV by the trigger. We require it to be larger than 20 GeV for the analysis to ensure that the trigger is fully efficient.

- p_T The transverse momentum of the associated track. A requirement of $p_T > 10$ GeV/c removes low-momentum electrons that might come from bremsstrahlung.

- $E_{\rm HAD}/E_{\rm EM}$ The ratio between the energy deposited in the hadronic calorimeters and the energy deposited in the electromagnetic calorimeters. The threshold is energy dependent. High-energy leptons can leave more energy in the hadronic calorimeter. This requirement reduces the jet contamination in the sample.

- *Isolation* Within a cone size of $\Delta R = 0.4$ around the electron cluster in the CEM, the ratio of the other clusters' energy to the electron energy (isolation) is required be smaller than 0.1.

- E/p The ratio between the energy of the calorimeter cluster and the track momentum. For true electrons, this ratio should be nearly one. For high-

momentum tracks ($p_T > 50$ GeV/c), the requirement of $E/p < 2$ is removed since it is unlikely for a jet to fake a high-momentum electron.

- $Q \times \Delta x$ Q is the measured electric charge of the electron candidate (± 1). Δx is the signed difference in the x direction between the track and the calorimeter cluster when the track is extrapolated to the position of the central shower maximum detector (CES). The x direction is defined to be perpendicular to both the particle beam (z axis) and the radial vector pointing to the calorimeter tower. The asymmetry of this requirement is due to the trajectory of particles in the detector. For a negative product, a particle traverses a larger part of the calorimeter in adjacent towers, which results in more radiation and a less precise final position.

- $|\Delta z|$ The absolute value of the difference between the z coordinate between the CES cluster and the track that is extrapolated to the position of the shower max.

- χ^2_{strip} The χ^2 fit result of the measured shower profile in the CES compared to an electron baseline shower profile acquired from electron test beam data.

- L_{shr} Short for "lateral energy sharing." It is a quantity that measures the difference between the lateral sharing of energy of towers in a calorimeter cluster and the expectation from the test beam data. It is defined as:

$$L_{\text{shr}} = \frac{0.14 \sum_i (E_i - E_i^{\text{exp}})}{\sqrt{(0.14\sqrt{E_i})^2 + \sum_i (\Delta E_i^{\text{exp}})^2}} \tag{5.1}$$

  where E_i is the measured energy in the i^{th} tower and E_i^{exp} is the expected value for the energy in the same tower according to a parametrization based on test beam data. The index i runs over all the towers adjacent to the seed tower.

- *COT track quality* The COT tracks must have high quality: each track must have at least five hits in each segment and there must be at least three axial superlayer segments and two stereo superlayer segments. The z component of the track where the track intersects the beam line should also be within 60 cm of the center of the detector.

- *Conversion veto* An energetic photon traveling through detector material can convert into an electron-positron pair, which is referred to as a photon *conversion*. An electron or positron from a photon conversion is discriminated from an electron from the hard scattering process. The conversion veto looks for an opposite-sign COT track near the track of the first identified electron. It also rejects events in which the cotangent of the polar angle between the two tracks is less than 0.4.

The identification (ID) efficiency of these selection cuts is measured in the data using a $Z \to e^+e^-$ sample where both the electron and positron are detected in the central calorimeter. A cut around the Z boson mass is applied to improve the sample purity; the background is estimated by the number of events in which two electrons of the same sign pass the event requirements. While one of the electrons in the pair is required to pass the CEM trigger and identification cuts of Table 5.1, the other one is examined for whether it also passes the identification cuts of Table 5.1. The fraction of data events in which both electrons pass the identification cuts leads to a central electron selection efficiency ϵ_{data} of about 96%.

The same procedure is done in the MC simulation to measure the MC efficiency ϵ_{MC}. A scale factor $SF = \epsilon_{\text{data}}/\epsilon_{\text{MC}}$ is then applied to events in the MC sample when modeling the data identification efficiency. The scale factor for central electrons is 0.976.

5.2.1.2. *Forward electrons.* A PHX candidate is an energy cluster in the plug calorimeter that is matched to a track reconstructed by the Phoenix algorithm. Thus, it uses only the tracking information from the silicon detector without the COT. Due to this and the lower resolution of the PEM, the PHX sample is less pure than the CEM sample, and tighter selection requirements are applied. The selection requirements are summarized in Table 5.2 and discussed here.

- *Geometry* The energy cluster must lie in the fiducial region of the plug calorimeter, which requires that the pseudorapidity of the cluster be within the range $1.2 < |\eta| < 2.0$.
- E_T The transverse energy of the electron must be larger than 20 GeV, which is larger than the trigger E_T requirement and reduces fake electrons.
- $E_{\text{had}}/E_{\text{EM}}$ The ratio of hadronic to electromagnetic energy, as described in the section on central electrons 5.2.1.1.
- χ^2_{PEM} The χ^2 requirement is like that described for CEM electrons, but measured for electrons in the PEM.

Table 5.2. Selection requirements for forward electrons.

Variable	Selection Requirement
Geometry	Fiducial in PEM
E_T	> 20 GeV
$E_{\rm HAD}/E_{\rm EM}$	< 0.05
$\chi^2_{\rm PEM}$	< 10
$N_{\rm silicon}$	≥ 3
$\Delta R_{\rm PEM-PES}$	< 3.0 cm
Isolation	< 0.1
E_5/E_9	> 0.65
$\lvert\eta\rvert$	< 1.6

- $N_{\rm silicon}$ Number of hits in the silicon detector associated with the track. This requirement improves the quality of the silicon track.
- $\Delta R_{\rm PEM-PES}$ The distance between the reconstructed cluster in the PEM calorimeter and the reconstructed cluster in the PES detector.
- E_5/E_9 The energy ratio of the central five strips to all nine strips of the PES energy cluster. For true electrons, the energy cluster tends to have a narrow energy distribution, which can be observed from the PES measurements by comparing the energy of the five strips at the center of the cluster (E_5) to the energy of all nine strips (E_9) for both layers (U and V layers) of the PES.
- $|\eta|$ We observed a high production rate of PHX electrons in the high $|\eta|$ region in the collision data, which might come from beam splash or pileup. We require the PHX electron to have $|\eta| < 1.6$, which is a region that is well understood.

The estimation of the ID efficiency and scale factor of PHX electrons is done in a similar way as in the CEM, expect one electron is triggered in the central region while the other is required to be in the plug region. A scale factor of 0.919 is applied to the MC samples.

5.2.1.3. *Muons.* Since the CDF II detector has a complex muon system, various types of muon candidates are reconstructed that bear the name of the corresponding muon detector that records them. A common set of identification cuts

applied to all muons is shown first in Table 5.3. Additional cuts for the different muon categories are shown later in Table 5.4.

Table 5.3. Common selection requirements for all muons.

Variable	Selection Requirement
Geometry	Fiducial to the sub-detectors
p_T	> 20 GeV/c
$E_{\rm EM}$	$< 2 + \max(0, (p - 100) \cdot 0.0115)$ GeV
$E_{\rm HAD}$	$< 6 + \max(0, (p - 100) \cdot 0.028)$ GeV
$\lvert d_0 \rvert$	< 0.2 cm (with silicon hits); < 0.02 cm without silicon hits
$\chi^2_{\rm track}$	< 2.3
Isolation	< 0.1
COT track quality	Pass COT track requirements

The following explains the muon-specific variables that have not already been described:

- $E_{\rm EM}$ The energy deposited in the electromagnetic calorimeter. Since muons pass through the calorimeter systems as MIP, the first term ($E_{\rm EM} < 2$ GeV) removes particles with enough energy to "punch through" the calorimeter. These particles (often pions) will leave more energy in the calorimeter than muons. The second term, which depends on momentum p, accounts for the rise of ionization energy from a true muon when its momentum is large.

- $E_{\rm HAD}$ The energy deposited in the hadronic calorimeter. The purpose of each term is similar to $E_{\rm EM}$. The threshold difference relative to the $E_{\rm EM}$ requirement is due to the difference in the material and thickness of the hadronic calorimeter.

- $|d_0|$ The absolute value of the impact parameter of the track. It is defined as the distance between the beam line and the position of the vertex reconstructed by the track in the $r - \phi$ plane. By requiring that the track points back to the beam line, this cut reduces the pion or kaon background from long-lived particles decaying in flight. The cut is tighter for lower quality tracks, which have no silicon hits.

There are eight different muon types used in this analysis. Because they are detected from different muon sub-detectors or triggered by different triggers, each of them has slightly different cuts for the specific characteristics and geometry of the subdetector or the trigger. The different selection requirements are summarized

in Table 5.4. The common variable Δx in the table is the distance between the extrapolated COT track and the matched stub. ρ_{COT} is the radial distance from the beam pipe to where the reconstructed track crosses the end plane of the COT.

Table 5.4. Specific selection requirements for different muon types.

Muon Type	Selection Requirement
CMUP	$\lvert\Delta x_{\mathrm{CMU}}\rvert < 3$ cm, $\lvert\Delta x_{\mathrm{CMP}}\rvert < 5$ cm
CMX	$\lvert\Delta x_{\mathrm{CMX}}\rvert < 6$, $140 < \rho_{\mathrm{COT}} < 180$ cm
CMU	$\lvert\Delta x_{\mathrm{CMU}}\rvert < 3$ cm
CMP	$\lvert\Delta x_{\mathrm{CMP}}\rvert < 5$ cm
BMU	$\lvert\Delta x_{\mathrm{BMU}}\rvert < 9$ cm
CMXNT	$\lvert\Delta x_{\mathrm{CMX}}\rvert < 6$, $\rho_{\mathrm{COT}} > 180$ cm
CMIO	$E_{\mathrm{EM}} + E_{\mathrm{HAD}} > 0.1$ GeV
SCMIO	$E_{\mathrm{EM}} + E_{\mathrm{HAD}} > 0.1$ GeV

- *CMUP* Stubs are required in both the CMU and CMP central muon chambers. The redundancy of the two chambers ensures a very pure sample with high efficiency. CMUP muons fall in the region $|\eta| < 0.6$.
- *CMX* Stubs are required in the CMX subdetector within the region $0.6 < |\eta| < 1.0$.
- *CMU/CMP* For each of these subdetectors, a stub must lie in the corresponding fiducial region, with no stub in the other subdetector.
- *BMU* A muon has to have a BMU stub and a track that points to the fiducial region of the BMU chamber.
- *CMXNT* Stubs are required in the CMX detector, but in a region ($|\Delta x_{\mathrm{CMX}}| < 6$ and $\rho_{\mathrm{COT}} > 180$ cm) that cannot be used in the trigger due to the tracking limitations of the trigger. The CMXNT muons have the same quality requirements as the normal CMX muons.
- *CMIO* CMIO muons must have an isolated track matched to a calorimeter cluster, but not matched to the fiducial region in any muon chambers. The requirement of additional energy strengthens the minimum ionizing energy requirement in the calorimeter.
- *SCMIO* The SCMIO requirements are similar to that for the CMIO muons except that a matched stub is required in a non-fiducial region of the muon subdetector. The same quality requirements as CMIO muons apply to these muons.

The ID efficiencies of the CMUP and CMX selections are measured with $Z \to \mu^+\mu^-$ samples using a similar method as CEM electrons. The resulting ID efficiencies in the data are 87.6% and 90.2%, respectively. The MC scale factor is 0.892 for CMUP and 0.948 for CMX.

The remaining muon types are selected from the $\not\!\!E_T$ + jets trigger path, as described in Section 5.1.5. They are categorized as extended muon coverage (EMC) from the previous analysis [37]. Their ID efficiencies and scale factors are parameterized by the turn-on curve function. As in the previous analysis, we group CEM, PHX, CMUP, and CMX into the tight lepton category (TLC), which are high quality lepton candidates that pass stringent selection cuts. Both EMC and TLC categories are used in the analysis.

5.2.1.4. *ISOTRK.* In order to increase the acceptance, a new lepton category named ISOTRK was introduced from the Higgs boson search group at CDF. If a reconstructed event contains a high quality, high-p_T isolated track within $|\eta| < 1.2$, it can be accepted as an "isotrk" lepton. Since the isolated track is not required to match to a calorimeter tower or a muon stub, we can recover real charged leptons that arrive in non-instrumented regions of the detector. These recovered events are not included in any of the previously described EMC categories from the $\not\!\!E_T$ + jets trigger path.

On an event-by-event basis, a good quality track candidate needs to pass the criteria detailed in Table 5.5. The isolation of each good quality track candidate is defined by the "track isolation":

$$\text{Track isolation} = \frac{p_T(\text{candidate})}{p_T(\text{candidate}) + \sum p_T(\text{trk})} \tag{5.2}$$

where $\sum p_T(\text{trk})$ is the sum of the p_T of all surrounding good quality tracks within a cone of $\Delta R = 0.4$ around the candidate track, and the difference in z vertex position is less than 5 cm. The ISOTRK selection cuts are given in Table 5.6.

Table 5.5: Good quality tracks are required to have p_T larger than 0.5 and at least 20 and 10 COT axial and stereo hits, respectively.

Variable	Selection Requirement
p_T	> 0.5 GeV/c
COT axial hits	≥ 20
COT stereo hits	≥ 10

Table 5.6. Event selection requirements for central electrons.

Variable	Selection Requirement
p_T	> 20 GeV/c
$\|\eta\|$	< 1.2
z_0	≤ 60 cm
Track isolation	≥ 0.9
COT axial hits	≥ 24
COT stereo hits	≥ 20
χ^2	$> 10^{-8}$
Impact parameter d_0	< 0.2 cm (with); < 0.02 cm (without silicon hits)

The isolation requirement is necessary in order to avoid the track being part of a hadron jet that originated from quarks or gluons. Additional vetoes are applied to ensure the ISOTRK corresponds to a charged lepton produced from the decay of a W boson; this lepton category remains orthogonal to any other lepton types used in the analysis. The following vetoes are applied sequentially:

- *Tight Jet Veto* The ISOTRK is required to have an angular separation of $\Delta R > 0.4$ from any tight jet in the event.
- *Two Track Veto* If a lepton is already identified in the event, then the event is vetoed.
- *Tight Lepton Veto* If there is a lepton already identified in the event, then the good quality track candidate will not be processed as ISOTRK.

Though the ISOTRK category may include electrons, muons, or tau leptons, it is collected into the EMC category in this analysis. It recovers real charged leptons that arrive in non-instrumented regions of the calorimeter or muon chambers, as illustrated in Figure 5.3(b) for muon candidates. By adding ISOTRK leptons, the

acceptance in the EMC signal region is increased by 15%. A detailed study [125] concluded that ISOTRK charged lepton candidates are muon candidates in 85% of cases, electron candidates in 7% of cases, and tau lepton candidates in 8% of cases.

5.2.2 Jet Multiplicity Selection

The jets reconstructed by the JETCLU algorithm are corrected up to the Level 5 jet correction as described in Section 3.4. The corrected jet candidates are required to have $E_T > 20$ GeV. Single top t-channel signal events feature forward jets, and we therefore select jets up to $|\eta^{\text{jet}}_{\text{det}}| < 2.8$. Events from $\not\!\!E_T$ + jets triggers have further jet requirements based on the specific trigger requirements described in Section 5.1.5. From Figure 4.8(a) and Figure 4.11(a), the majority of single top signal events contain two or three jets, with one or two jets tagged by the SECVTX algorithm as a b-quark jet. Thus we define four signal channels: two jets, one tagged (2J1T); two jets, two tagged (2J2T); three jets, one tagged (3J1T); and three jets, two tagged (3J2T).

5.2.3 Missing Transverse Energy

The presence of the neutrino coming from W-boson decays is inferred by the amount of missing transverse energy ($\not\!\!E_T$). The raw $\not\!\!E_T$ is corrected for the presence of muons and jet energy corrections as described in Section 3.5. We require $\not\!\!E_T > 25$ GeV to reduce the QCD multijet background.

Since the missing transverse energy is calculated in the transverse plane, the z-component of the neutrino momentum remains unknown. Based on the assumption that the neutrino and lepton are formed by a real W boson; however, we can further reconstruct the neutrino's properties. The neutrino momentum can be solved from the kinematic constraint that the lepton and neutrino four-momenta are consistent

with a W boson mass equal to 80.45 GeV/c^2 as follows.

$$p^W = p^l + p^\nu \tag{5.3}$$

$$m^W = \sqrt{(p^l + p^\nu)^2} = 80.45 \text{ GeV}/c^2 \tag{5.4}$$

By solving the quadratic equation, the z component of p^ν is chosen by the smallest $|p_z^\nu|$ solution. In the case where both solutions are complex, by varying the x and y component of the $\not{E}_T$, a kinematic fit is performed on the equation to find the closest solution for $|p_z^\nu|$ [126]. Thus, the neutrino energy can be reconstructed as $E(\mu) = \sqrt{\not{E}_T{}^2 + p_z^{\mu\,2}}$.

5.3 Background Vetoes

After the event selection, vetoes are further imposed to reduce certain types of known background events.

5.3.1 Dilepton Veto

To ensure that the event contains exactly one lepton, events are rejected that have more than one lepton, including loose leptons. Loose leptons are defined as leptons that have passed all lepton quality cuts except the isolation requirement. This veto greatly reduces the $t\bar{t}$ dilepton background.

5.3.2 Z Boson Veto

We reject events in which the trigger lepton candidate can be paired with a jet or high-momentum isolated track pointing in the opposite direction, forming an invariant mass within the Z boson mass window, defined as $76 \leq m_Z \leq 106$ GeV/c^2. This veto reduces most of the Z + jets background events.

5.3.3 *z Vertex Requirement*

The reconstructed primary vertex of the event is required to be within ±60 cm from the center of the detector. This maximizes the coverage of the detector for particles emerging from the collision region.

5.3.4 *Cosmic Veto*

Cosmic rays from the outer atmosphere can pass through the detector leaving a muon-like signal. This veto rejects events identified as coming from cosmic rays by taking advantage of the event topology and the characteristic timing of cosmic ray events. Because cosmic rays travel through the detector from top to bottom, they often appear as a track punching through the detector. The timing of a cosmic ray track with respect to the bunch crossing could be different from a track coming from the hard scattering. In addition, the TOF timing system can distinguish cosmic rays by comparing the time of flight of the two tracks. This veto only applies to data since we do not simulate the cosmic ray background in MC production.

5.3.5 *Single Top QCD Veto*

This veto is introduced to reduce the QCD background, which is a relatively large background. Since our knowledge of this background is primitive, our strategy is to remove it as much as possible. The QCD events do not contain W bosons, thus we also call them "non-W" events in the following text. Based on the assumption that $\not\!\!E_T$ comes from an escaping neutrino from W decay, we identify the non-W background as events in which the $\not\!\!E_T$ comes from lost or mismeasured jets and the lepton was misidentified as a jet. As a result, the $\not\!\!E_T$ in non-W events mostly points close to the lepton candidate's direction. Furthermore, the transverse mass of the W boson is expected to be small; this quantity is defined as:

$$m_T^W = \sqrt{2\left(p_T^\ell \not\!\!E_T - p_x^\ell \not\!\!E_T{}^x - p_y^\ell \not\!\!E_T{}^y\right)} \tag{5.5}$$

where p_T^ℓ is the transverse momentum of the lepton; $\not\!\!E_T{}^x$ and $\not\!\!E_T{}^y$ are the x and y components of the $\not\!\!E_T$, respectively. Since it is easier for an energetic jet to mimic an electron rather than a high-p_T muon, we expect more non-W background from the electron sample. We require $m_T^W > 20$ GeV for electrons and $m_T^W > 10$ GeV for tight muons. In the case of EMC muons, m_T^W is required to be above 20 GeV due to a larger QCD contamination in the current analysis. (The previous analysis applied a cut at 10 GeV for loose muons except SCMIO, for which a cut of $m_T^W > 20$ GeV was applied.)

To further remove non-W events, a variable called MET-significance ($\not\!\!E_{T,\mathrm{sig}}$) is defined as:

$$\not\!\!E_{T,\mathrm{sig}} = \frac{\not\!\!E_T}{\sqrt{\sum_{\mathrm{jets}} C_{\mathrm{JES}}^2 \cos^2\left(\Delta\phi_{\mathrm{jet},\vec{\not\!\!E}_T}\right) E_{T,\mathrm{jet}}^{\mathrm{raw}} + \cos^2\left(\Delta\phi_{\vec{E}_{T,\mathrm{uncl}},\vec{\not\!\!E}_T}\right) \sum E_{T,\mathrm{uncl}}}} \tag{5.6}$$

where C_{JES} is the jet energy correction factor [81], $E_{T,\mathrm{jet}}^{\mathrm{raw}}$ is a jet's energy before corrections are applied, $\vec{E}_{T,\mathrm{uncl}}$ refers to the vector sum of the transverse components of calorimeter energy deposits not included in any reconstructed jets, and $\sum E_{T,\mathrm{uncl}}$ is the scalar sum of these unclustered energies. $\Delta\phi_{\mathrm{jet},\vec{E}_{T,\mathrm{uncl}}}$ is the angle between a jet and $\vec{\not\!\!E}_T$ projected in the $r-\phi$ plane. The angle between the projections in the $r-\phi$ plane of $\sum \vec{E}_{T,\mathrm{uncl}}$ and $\vec{\not\!\!E}_T$ is denoted as $\Delta\phi_{\vec{E}_{T,\mathrm{uncl}},\vec{\not\!\!E}_T}$. The $\not\!\!E_{T,\mathrm{sig}}$ is an approximate significance of the dispersion in the measured $\not\!\!E_T$, approximated by the denominator, in events with no true $\not\!\!E_T$.

Since the non-W background has a higher contamination from electron samples than muon samples, additional cuts are required for CEM and PHX events as listed below. All the energies are measured in GeV.

- CEM events are required to pass the triangular cut

$$\not\!\!E_{T,\mathrm{sig}} > 3.5 - 0.05 \cdot m_T^W, \tag{5.7}$$

  which is shown in Figure 5.4(c). In addition, they should have

$$\not\!\!E_{T,\mathrm{sig}} > \begin{cases} -7.6 + 3.2 \cdot |\Delta\phi(l,\mathrm{jet})| & \text{1-jet bin} \\ 2.5 - 3.125 \cdot |\Delta\phi(\not\!\!E_T,\mathrm{jet}_2)| & \text{2-jets, 3-jets bin} \end{cases} \tag{5.8}$$

where jet_2 is the second-most energetic jet.

- PHX candidate events must have:

$$\not\!\!E_{T,\text{sig}} > 2; \quad \not\!\!E_{T,\text{sig}} > 45 - 30 \cdot |\Delta\phi(\not\!\!E_T, \text{jet})|. \tag{5.9}$$

  The last requirement is applied for all the jets in the event.

This sequence of requirements substantially reduces the amount of contamination from non-W events as shown in Figure 5.4. This procedure is called the "Single Top QCD Veto" since it is designed specifically for the single top analysis [127]. This veto is applied to both data and MC events in this analysis.

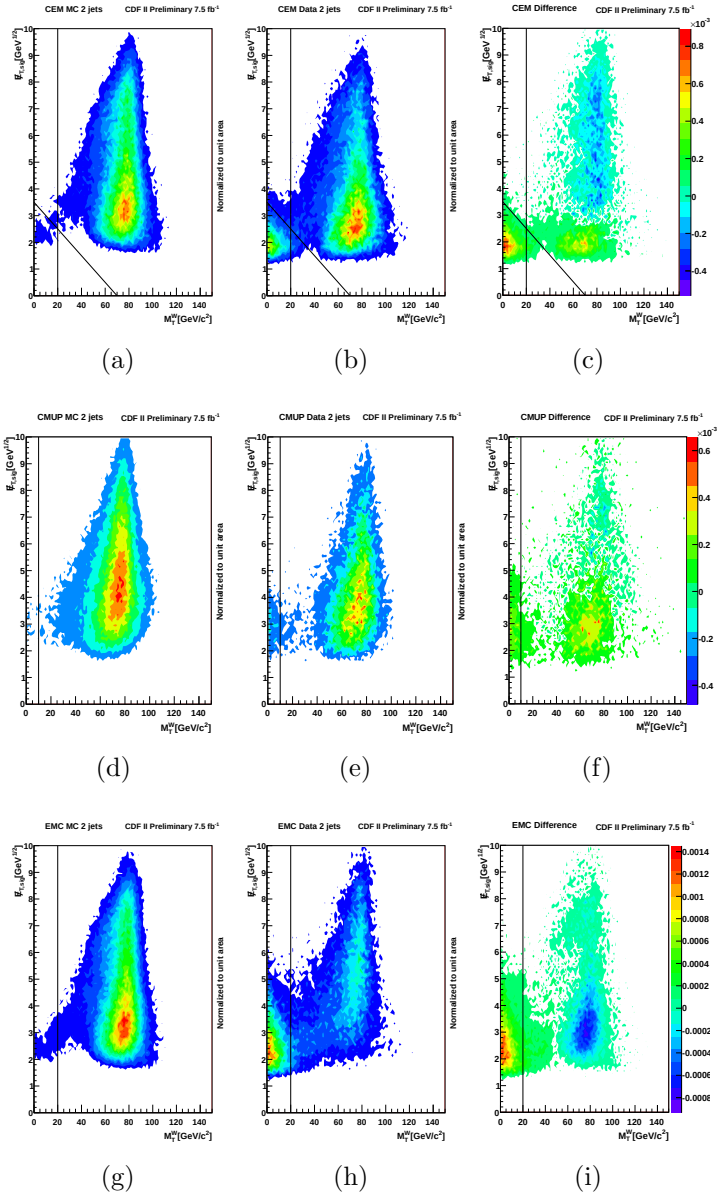


Figure 5.4: Scatter plots representing the single-top-QCD-veto variables $\not\!\!E_{T,\mathrm{sig}}$ vs. m_T^W for pretag events in the 2-jet bin. (a)–(c) shows this for CEM events, (d)–(f) for CMUP events, and (g)–(i) for EMC events. The first column shows simulated events from a W + jets MC sample. The second column shows events from selected data events. In the third column, the MC and data distributions are subtracted, showing the presence of the non-W background in the low-W mass region, which is rejected by the cuts represented by the vertical and angled lines.

CHAPTER SIX

Background Estimation

A precise cross section measurement of a desired physics process strongly relies on the accurate estimation of its background physics processes. In this chapter, we will describe the background modeling and estimation in this analysis. Based on our selected event signature, we use a sophisticated method called "Method II for you" [128], developed at CDF.

The single top lepton + jets data sample can be divided into events with exactly two or three jets, with one or two of these being b-tagged jets. This gives a total of four categories. Each category can be further divided based on the lepton type. We separate our lepton types into the CEM, PHX, CMUP, CMX, and EMC categories. Thus, we have a total of 20 categories, also called channels. Since each category has its own characteristic features, separating out each category enables a more detailed study of these characteristic features and it improves the sensitivity of the analysis. Thus, each category has a separate estimation of the sample composition. Based on our selection in the reconstructed lepton + jets sample, the data sample can be broken down into several regions. The data sample before applying the SECVTX algorithm is called the *pretag* sample. The data sample after applying the SECVTX algorithm (b-tagging), is called the *tag* sample. Events that contain zero b-jets, but have at least one taggable jet, are called the untagged sample. We define events with two or three jets as our *signal* region. Other events make up the *sideband* region, which is used as a control sample for background studies or validation.

6.1 Method II Background Estimation

Method II is a method containing a sequence of procedures that are used to calculate the normalization of physics processes in the tagged lepton + jets data samples, based on the assumption that all the possible physics processes contributing to the data are known. The normalized predictions (rates) should give good estimations of how much the physics processes contribute to the data samples. Method II has successfully been used in the $t\bar{t}$ cross section measurement [82] and the previous single top analysis that led to the first single top observation [36]. As described in Section 4.1, the selected lepton + jets data sample contains the following physics processes: single top quark production, $t\bar{t}$ production, diboson production, W + jets and Z + jets production, and the non-W (QCD) processes. These processes are treated by a variety of methods to predict their contribution to the signal region of the data sample.

Due to the failure of the MC simulation to model the whole data sample well, Method II uses a data-driven method to model part of the backgrounds that are not as well understood (like the non-W background, as in Section 4.3). It divides all the physics processes into three categories: fully modeled by MC predictions, fully modeled by data-driven techniques, and modeled by a combination of the two methods. Well understood processes, such as $t\bar{t}$, are purely based on MC simulations, scaled to higher-order cross section predictions. For poorly understood backgrounds like non-W, we use a data-driven method to model the kinematic distributions (shapes) and estimate the background rate. The non-W models are derived from QCD-enriched sideband data (data selected outside of our signal region). We normalize them by fitting the corresponding $\not\!\!E_T$ distribution to the data sample as discussed in Section 6.3. For the W + jets processes, we combine the Monte Carlo simulation with a data-driven method. We obtain kinematic distributions from MC samples. The

normalization of these W + jets processes is carried out separately for the W + HF and the W + LF samples. These procedures are explained below.

6.2 Monte Carlo-Based Event Yield Estimate

For those processes whose theoretical cross section are well understood — $t\bar{t}$, diboson, and Z + jets — we use samples of simulated MC events to estimate their contributions in the data. We also use the MC prediction to describe the expected single top signal. For each simulated process, the number of expected events in the data (N) is given by

$$N = \sigma \cdot \epsilon_{\text{evt}} \cdot \epsilon_{\text{tag}} \cdot \int L\ dt \tag{6.1}$$

where σ is the theoretical predicted cross section of the respective process at the Tevatron as listed in Table 6.1 and $\int L\ dt$ is the integrated luminosity of the data used in this analysis (7.5 fb^{-1}).

Table 6.1: The theoretical cross sections and branching ratios (BR) used for MC-based processes. The cross sections are calculated at NLO (or higher) accuracy assuming a top quark mass of 172.5 GeV/c^2. The BR depends on how samples are generated: some of them are generated inclusively with all decay modes (we then use unity as the BR value) and others are generated exclusively with leptonic decays (we then use the leptonic BR for this choice).

EW processes	Cross Section (pb)	Branching Ratio
$t\bar{t}$	7.04 ± 0.44	1.0
Single Top (s-ch)	1.06 ± 0.06	0.324
Single Top (t-ch)	2.12 ± 0.22	0.324
Single Top (Wt-ch)	0.22 ± 0.08	1.0
WW	11.60 ± 0.70	1.0
WZ	3.46 ± 0.30	1.0
ZZ	1.51 ± 0.20	1.0
Z + jets	787.4 ± 85.0	1.0

The event detection efficiency ϵ_{evt} can be further broken down into five factors:

$$\epsilon_{\text{evt}} = \epsilon_{\text{evt}}^{\text{MC}} \cdot \text{BR} \cdot \epsilon_{\text{z}_0} \cdot \epsilon_{\text{trigger}} \cdot \epsilon_{\text{leptonID}}. \tag{6.2}$$

- $\epsilon_{\rm evt}^{\rm MC}$ is the event selection efficiency obtained as the ratio of the number of selected MC events (pretag sample) to the total number of MC simulated events.

- BR is the branching ratio of the W boson leptonic decay. Some MC samples are generated such that the W boson must decay leptonically. For these samples, the branching ratio is set to BR = 0.324. Inclusive samples have a BR set to 1.0. This is shown in Table 6.1.

- $\epsilon_{\rm z_0}$ is the z-vertex cut efficiency in data for a z vertex within $|z| \leq 60.0$ cm. The quoted luminosity 7.5 fb^{-1} includes the fully luminous region in z measured by the CLC detector. The z-vertex cut limits the event efficiency to a portion of the fully luminous region of $p\bar{p}$ collisions [129]. An efficiency of $(97.20 \pm 0.05)\%$ is measured using minimum-bias data collected in the same period as the 7.5 fb^{-1} data sample.

- $\epsilon_{\rm trigger}$ is the trigger efficiency obtained from the data for each trigger. Different leptons come from different trigger paths; some of the triggers have a prescale, and some events trigger on $\not\!\!E_T$ or EM cluster energy. We therefore apply a different trigger efficiency value for each trigger according to its features. For kinematic-dependent triggers (like the $\not\!\!E_T$-based trigger), we use the turn-on curve to sculpt the MC samples with the kinematic-dependent features of the data. On an event-by-event basis, each corresponding kinematic variable is evaluated to obtain a trigger weight. A combination of these weights is calculated as the event trigger weight. For CEM electrons, a turn-on curve is constructed based on the calorimeter cluster and its associated track [130]. Since PHX electrons are triggered by raw $\not\!\!E_T$ and a PEM cluster, their turn-on effects are included in the PHX trigger weight [131]. For CMUP and CMX muons, a constant trigger weight of $(87.7 \pm 0.2)\%$ and $(90.2 \pm 0.2)\%$ is applied, respectively [132]. For EMC leptons, due to the complex trigger combination [133], a turn-on curve is applied as described in Section 5.1.5.

- $\epsilon_{\rm leptonID} = \epsilon_{\rm leptonID}^{\rm data}/\epsilon_{\rm leptonID}^{\rm MC}$ is the ratio of the lepton identification efficiencies calculated in the data and MC simulation, also called the *lepton identification scale factor*. For each period of data, the lepton identification scale factors were evaluated by the Joint Physics Group at CDF. Almost all the lepton ID scale factors are flat (no kinematic dependence) with respect to time [134]. We apply the period-dependent lepton ID scale factor to the MC prediction in order to better model the data.

The b-tagging selection efficiency $\epsilon_{\rm tag}$ in Eq. 6.1 is a bit more complicated. This is because the MC simulation tends to overestimate the tagging efficiency of jets originating from heavy flavor quarks and underestimate the number of mistagged

light flavor jets. A tagging scale factor and mistag matrix must be included in the tagging efficiency calculation. The tagging efficiency is obtained from the probability of a MC event being tagged ($P_{\text{event}}^{\text{tag}}$), which is given by

$$P_{\text{event}}^{\text{tag}} = 1 - \prod_{i}^{\text{jets}} (1 - p_{\text{tag}}^{i}) \tag{6.3}$$

Each detector-level jet in the event is first matched to a parton-level object within $\Delta R < 4$ in $\eta - \phi$ space. If the simulated b-tagged jet is matched to a heavy flavor quark, p_{tag}^{i} is the tagging scale factor. If the simulated b-tagged jet is matched to a light flavor quark or gluon, p_{tag}^{i} is the mistag probability. The p_{tag}^{i} value is zero for jets that are not b-tagged. The summation runs over all the available tight jets in the event. The tagging efficiency can then be calculated as

$$\epsilon_{\text{tag}} = \frac{\sum_{j}^{\text{events}} P_{j}^{\text{tag}}}{N_{\text{pretag}}} \tag{6.4}$$

where N_{pretag} is the number of pretag events in the MC sample and the summation is over the number of pretag events.

6.3 Modeling of Non-W (QCD) Background

As shown in Figure 6.1, non-W (QCD) multijet events can mimic the $W \to l\nu$ signature. A real lepton can exist from a semileptonic b decay in direct $b\bar{b}$ production (Figure 6.1(a)) or a fake lepton can be reconstructed from the misidentification of a jet in strong gluon production (Figure 6.1(b)); $\not\!\!E_T$ can arise from the mismeasured energy of jets. Even after the single top QCD veto, this background is still significant due to its huge cross section. Because it is not feasible to model non-W events with MC simulation, we use a data-driven method that is referred to as a non-W model. This method selects data samples from the sideband regions with less stringent (looser) selection requirements. These samples are dominated by non-W events with similar kinematic features as the non-W contribution to the signal sample. It is hard for fake leptons from non-W events to pass the lepton selection requirements.

By taking advantage of this, we can obtain a non-W enriched sample by loosening or reversing lepton selection requirements. A variety of non-W models are considered and a final set are presented below, according to the characteristic features of each lepton category.

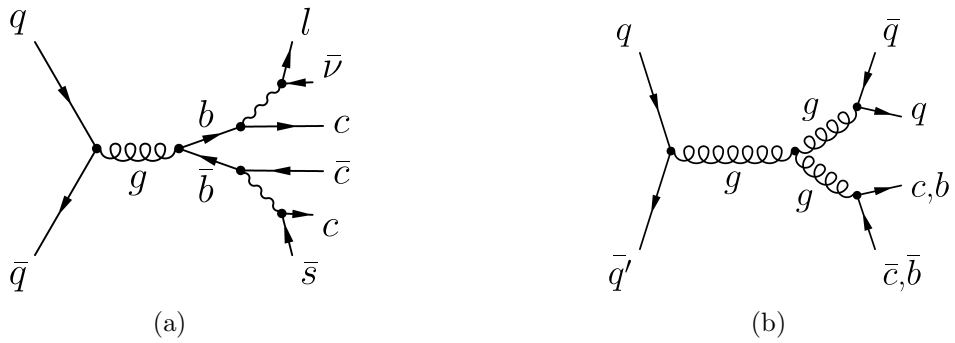


Figure 6.1: Feynman diagrams of two QCD multijet production modes: (a) direct $b\bar{b}$ production and (b) strong gluon production.

6.3.1 Anti-CEM Model

The anti-CEM model is a subset of the anti-electron model [135]. The anti-electron model requires the electron-triggered event to pass the electron selection from Section 5.2.1, except for any two or more of the five requirements, which are specified for electron identification. For CEM electrons, these identification requirements are:

- $E_{\rm HAD}/E_{\rm EM} \leq 0.055 + (0.00045 \cdot E)$ GeV
- $\chi^2 < 10.0$
- $L_{\rm shr} < 0.2$
- $|\Delta z| < 3.0$ cm
- $-3.0 < Q \cdot \Delta x < 1.5$ cm

Because these identification requirements serve primarily to reduce QCD multijet events, anti-CEM events are QCD-enriched and their kinematic properties still resemble real W-boson events. In this analysis, we improve the anti-CEM modeling

by applying an additional $\not\!\!E_T$ correction. Since the anti-CEM is not a real electron in the event, during the reconstruction, the anti-CEM is reconstructed as a loose jet (a reconstructed jet with less stringent requirements). By manually setting the loose jet as an anti-CEM, we artificially produce a fake electron and the loose jet is removed from the jet list in the event. Thus, within the anti-CEM modeling, the anti-CEM electron always has a loose jet matched in $\eta - \phi$ space, which was previously ignored in the $\not\!\!E_T$ calculation. Since anti-CEM events are mostly multijet events, the mismeasured $\not\!\!E_T$ should be corrected for all QCD jets, including the matched jet. From a sideband study, we found that only a fraction of the L5 energy correction of the matched jet should be taken into account in the $\not\!\!E_T$ correction. By studying the zero-tag sideband sample, we found that applying a correction corresponding to 5% of the energy of the matched jet best models the non-W background in our signal sample, especially the $\Delta\phi(\not\!\!E_T, \text{jets})$ distribution.

The corrected anti-CEM model does a remarkable job of describing the kinematic properties of the non-W sample, even for central muon events. Thus, the exact same model is used to model non-W events for the CMUP and CMX categories. The anti-CEM events are assigned to the CMUP or CMX categories according to the pseudorapidity of the anti-CEM electrons.

6.3.2 Anti-PHX Model

Similar to the anti-CEM model, the anti-PHX model is derived from MET_PEM triggered events in which an electron candidate passes all the PHX selection requirements (see Section 5.2.1.2) except any two or more of the identification requirements: $E_{\text{HAD}}/E_{\text{EM}}$, E_5/E_9, PEM 3×3 Fit Tower, and χ^2_{PEM}.

6.3.3 Non-Isolated Muon Model

The kinematic distributions of reconstructed objects in the EMC sample are different from those in the TLC (CEM, PHX, CMUP, and CMX) samples due to the $\not\!\!E_T$-based trigger requirement. Thus, a separate model must be used to model the non-W background in the EMC data. We collect events from the $\not\!\!E_T$ + jets trigger path that pass all EMC selection requirements expect for the isolation requirement. This is dubbed the "non-isolated" muon model. These non-isolated leptons are typically leptons contained inside of jets. Jets that contain energetic leptons are more likely to pass the lepton identification cuts. Since this sample comes directly from the $\not\!\!E_T$ + jets trigger path, it models the low $\not\!\!E_T$ region of the EMC sample much better than the anti-electron model. We require the isolation value to be greater than 0.2 to further improve the modeling.

6.4 Non-W (QCD) Background Estimate

We use different non-W models to model the kinematic features of non-W events in the data. The normalization of these non-W models is determined separately by fitting a kinematic distribution to the data. Missing transverse energy is the first choice because it is expected that QCD multijet events dominate the low $\not\!\!E_T$ region (since QCD multijet events do not contain a true neutrino). The pretag data sample without the $\not\!\!E_T$ cut contains a large sample of QCD multijet events. We require that at least one jet in the pretag events is taggable. By keeping the normalization of the MC-based processes fixed, the W + jets and the non-W samples are fit to the $\not\!\!E_T$ spectrum of the pretag data using a binned likelihood method. This procedure is performed separately for each lepton category. The full $\not\!\!E_T$ spectrum is used in the fit except for PHX events; for the PHX category, we merge $\not\!\!E_T$ below 15 GeV into one bin because we are not confident of the trigger modeling in that region. The fraction of the non-W events is then calculated in the pretag sample for

$\not\!\!E_T > 25$ GeV. The pretag *QCD fraction* is essential for Method II since it provides the starting point for the normalization of W + HF and W + LF as described later in Section 6.5. The total number of QCD events in the pretag sample ($N_{\rm QCD}^{\rm pretag}$) is given by

$$N_{\rm QCD}^{\rm pretag} = F_{\rm QCD}^{\rm pretag} \cdot N^{\rm pretag} \tag{6.5}$$

where $F_{\rm QCD}^{\rm pretag}$ is the fraction of non-W events obtained from the fit and $N^{\rm pretag}$ is the number of pretag data events. The results of the $\not\!\!E_T$ fits in the pretag samples for this analysis are shown in Figure 6.2 and Figure 6.3 for the 2-jets and 3-jets bins, respectively.

After the W + jets normalization is calculated, the same procedure is done in the b-tagged samples to determine the QCD fraction in the b-tagged channels. The requirement of at least one b-tagged jet reduces the QCD contribution to a negligible number of events, since only a few events pass the secondary vertex requirement. Therefore we manually set all the taggable jets in the sample to be b-tagged jets in order to enrich the statistics of the fit. The total number of QCD events in the tagged sample ($N_{\rm QCD}^{\rm tag}$) is given by

$$N_{\rm QCD}^{\rm tag} = F_{\rm QCD}^{\rm tag} \cdot N^{\rm tag} \tag{6.6}$$

where $F_{\rm QCD}^{\rm tag}$ is the fraction of non-W events obtained from the fit and $N^{\rm tag}$ is the number of b-tagged data events. The results of the $\not\!\!E_T$ fits in the single SECVTX (SVT) tagged channels are shown in Figure 6.4 and Figure 6.5 for the 2J1T and 3J1T bins, respectively. And the results of the $\not\!\!E_T$ fits in the double SECVTX tagged channels are shown in Figure 6.6 and Figure 6.7 for the 2J2T and 3J2T bins, respectively.

The inclusion or omission of the single top contribution to these fits has a negligible impact on the non-W fractions in both the pretag and tagged samples, which makes the QCD fraction statistically independent of our signal. Because of

the uncertainties in the modeling of the $\not\!\!E_T$ shapes, the estimation method of fitting, and the tagging rates, a conservative systematic uncertainty of $\pm 40\%$ is assigned to the normalization of the non-W background. The 40% uncertainty also takes into account the differences in the results obtained by fitting variables other than $\not\!\!E_T$, as well as by changing the histogram binning, varying the fit range, and using alternative non-W models.

6.5 W + Jets Background Estimate

The W + jets rate cannot be easily determined from its predicted cross section for several reasons. First, this background is described by a large number of possible Feynman diagrams and these diagrams contain many instances of gluon splitting, which causes an infrared divergence when the angle between split quarks is small. The infrared divergence is handled by applying a factorization scale, which defines the energy at which two jets are merged. This factorization scale has a large impact on the cross section calculation since its value is a mathematical artifact and has no intrinsic physical meaning. Thus, the leading-order cross section from the ALPGEN generator cannot be trusted blindly. Furthermore, jets may fail to be detected or pass our event selection. In this case, a W + 3 jets event may be mis-identified as a W + 2 jets event. A direct quoting of the ALPGEN cross sections cannot be applied in the normalization. The uncertainty in the W + jets cross section can easily overwhelm the size of the single top signal. Thus we need an alternative method to extract the W + jets contribution.

Based on the assumption of Method II, we consider all remaining backgrounds in the pretag sample that are not from QCD processes or the MC-based processes of Section 6.2 as coming solely from the W + jets process. Thus, the normalization of the W + jets samples can be calculated as

$$N_{W+\text{jets}}^{\text{pretag}} = N^{\text{pretag}} \cdot (1 - F_{\text{QCD}}^{\text{pretag}}) - N_{\text{t}\bar{\text{t}}+\text{EW}}^{\text{pretag}} - N_{\text{ST}}^{\text{pretag}} \tag{6.7}$$

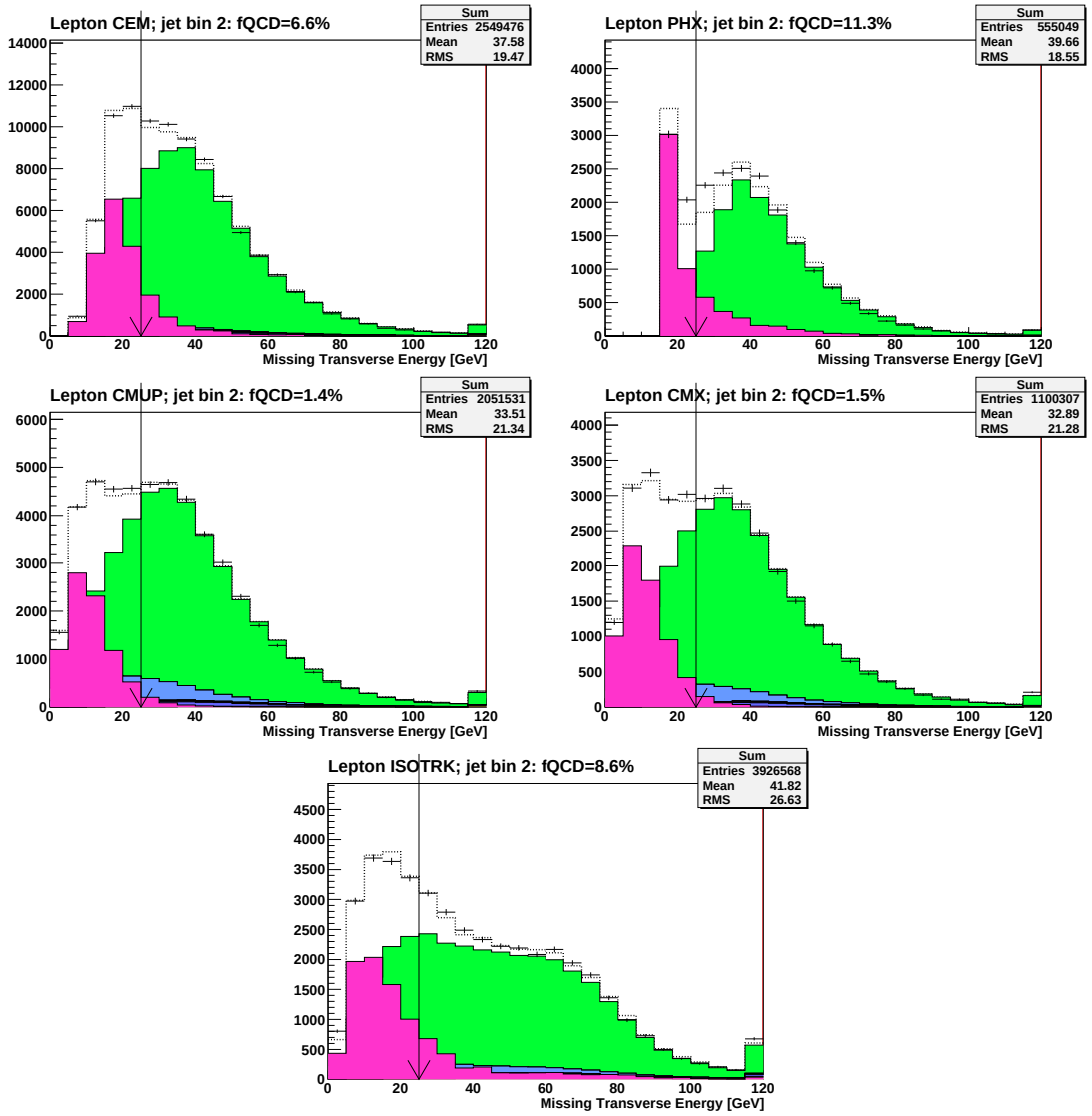


Figure 6.2: QCD fraction estimation for the pretag 2-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

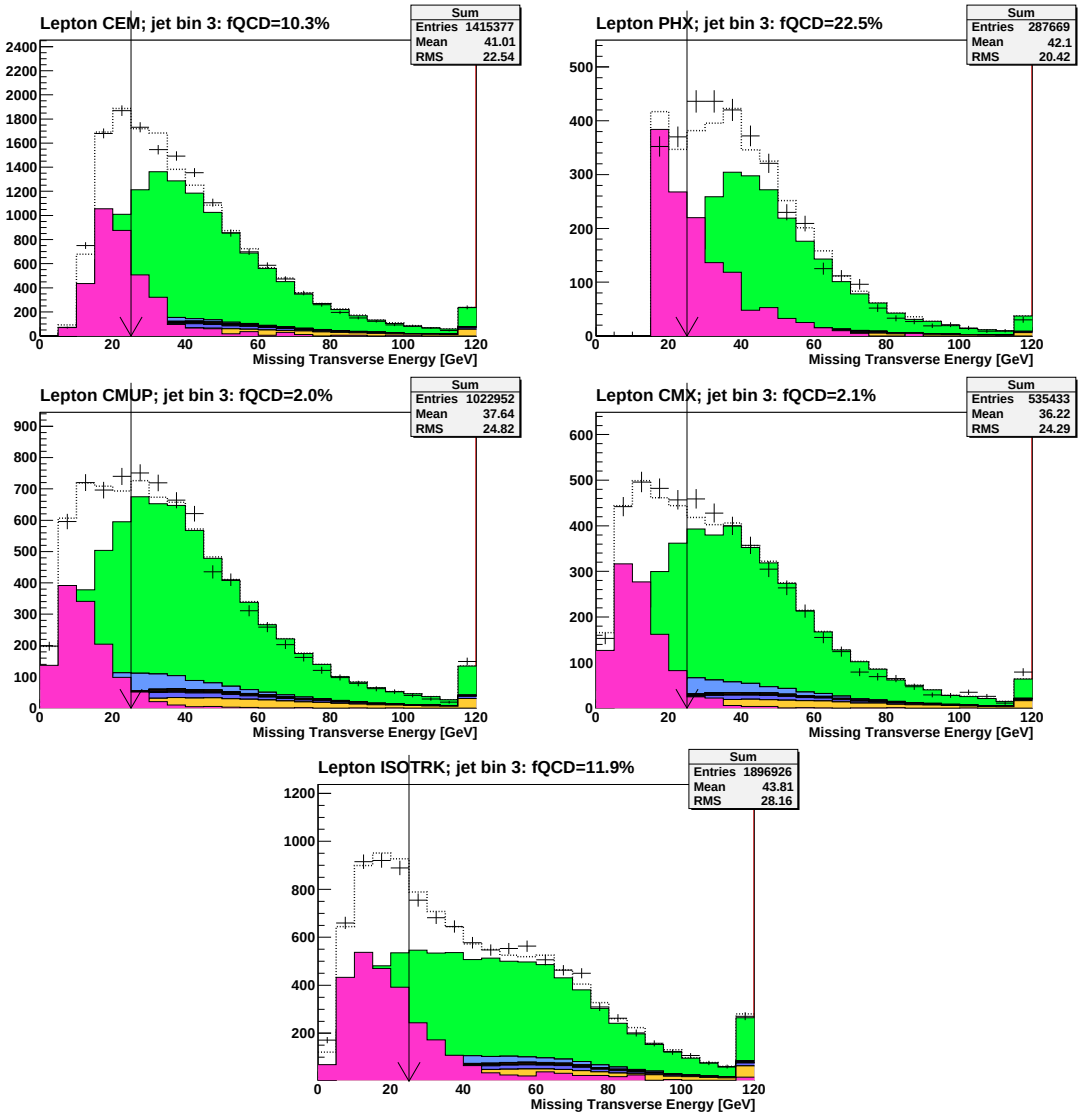


Figure 6.3: QCD fraction estimation for the pretag 3-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

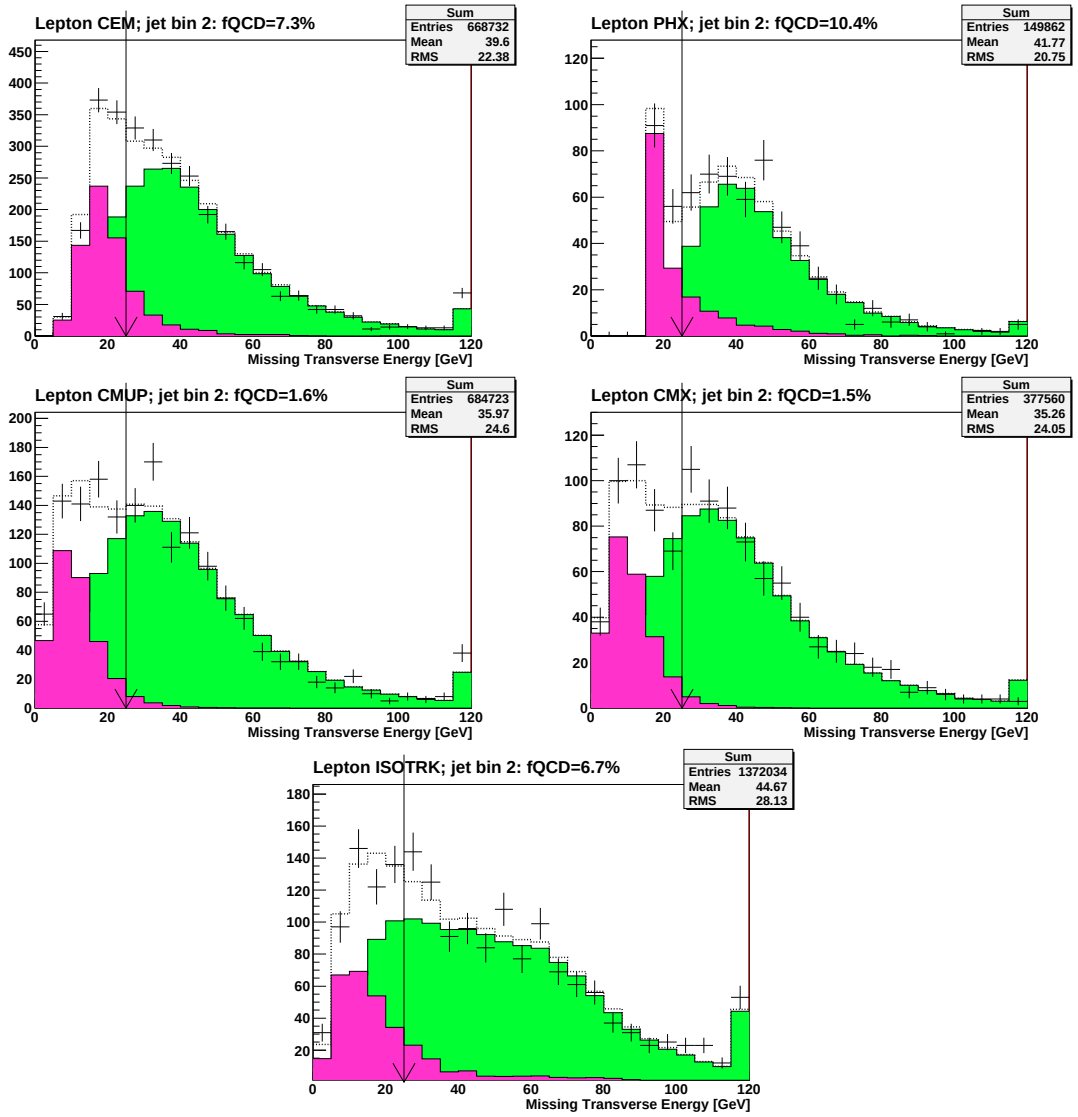


Figure 6.4: QCD fraction estimation for the one-SVT-tagged 2-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

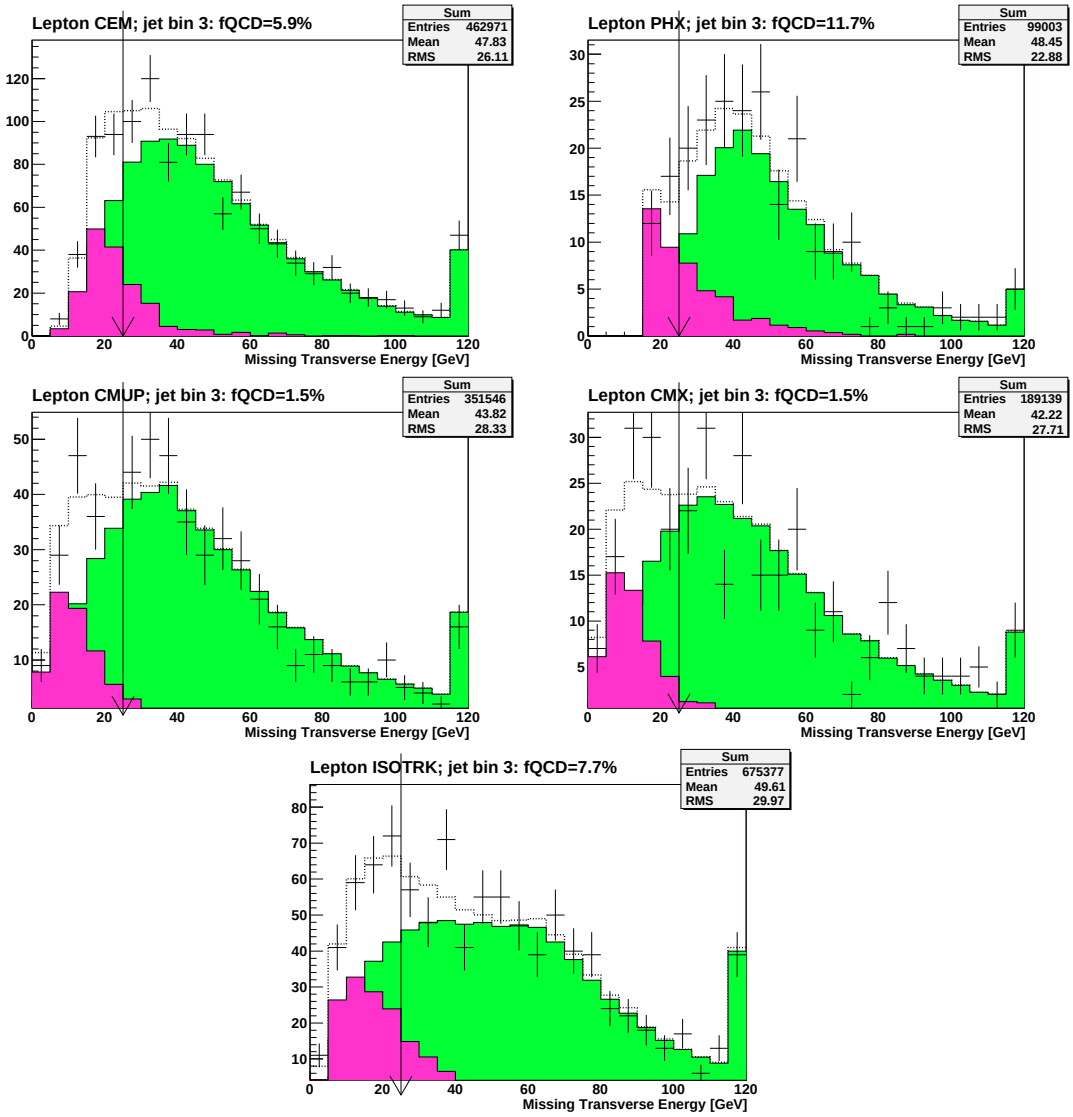


Figure 6.5: QCD fraction estimation for the one-SVT-tagged 3-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

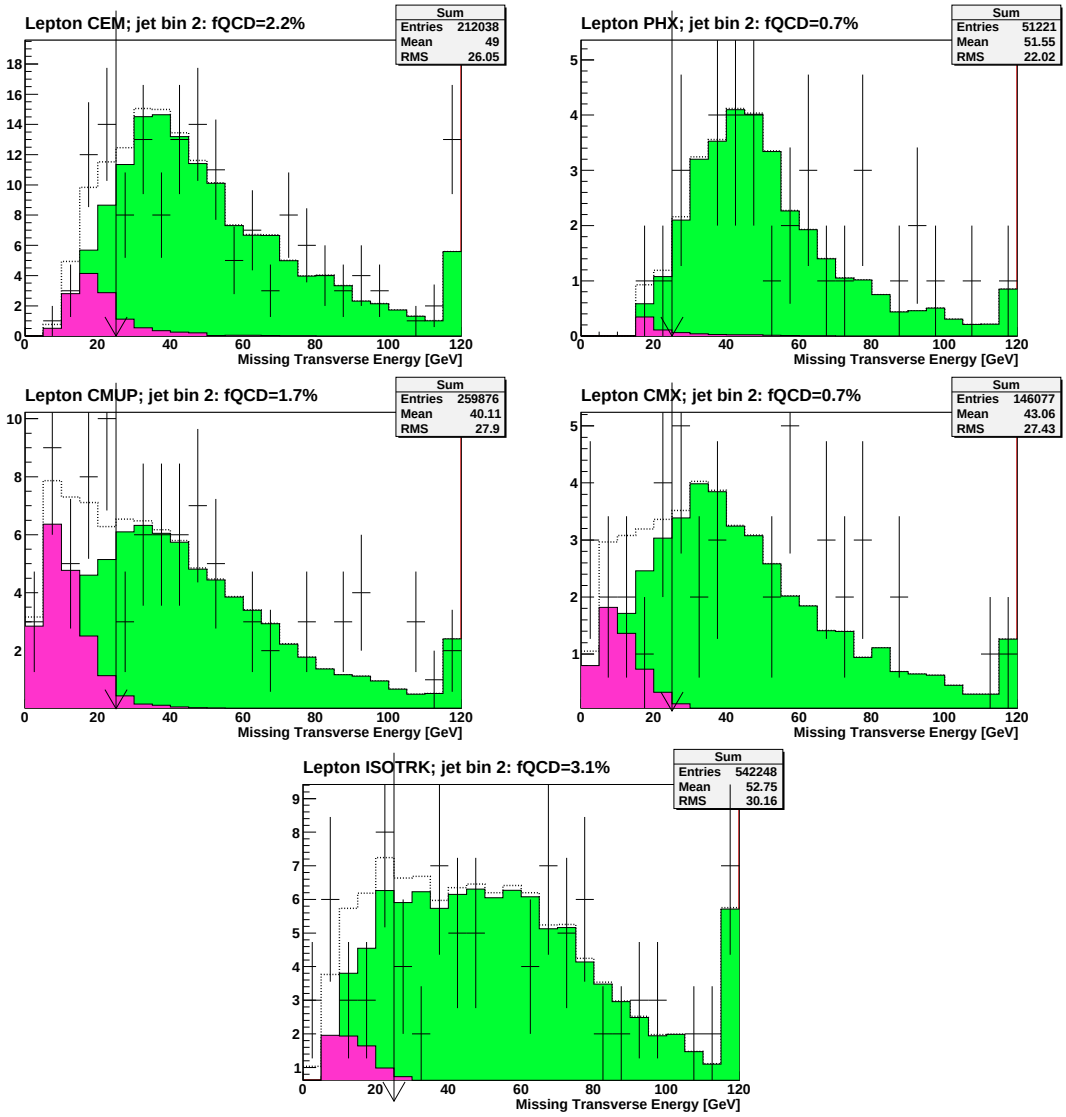


Figure 6.6: QCD fraction estimation for the double-SVT-tagged 2-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

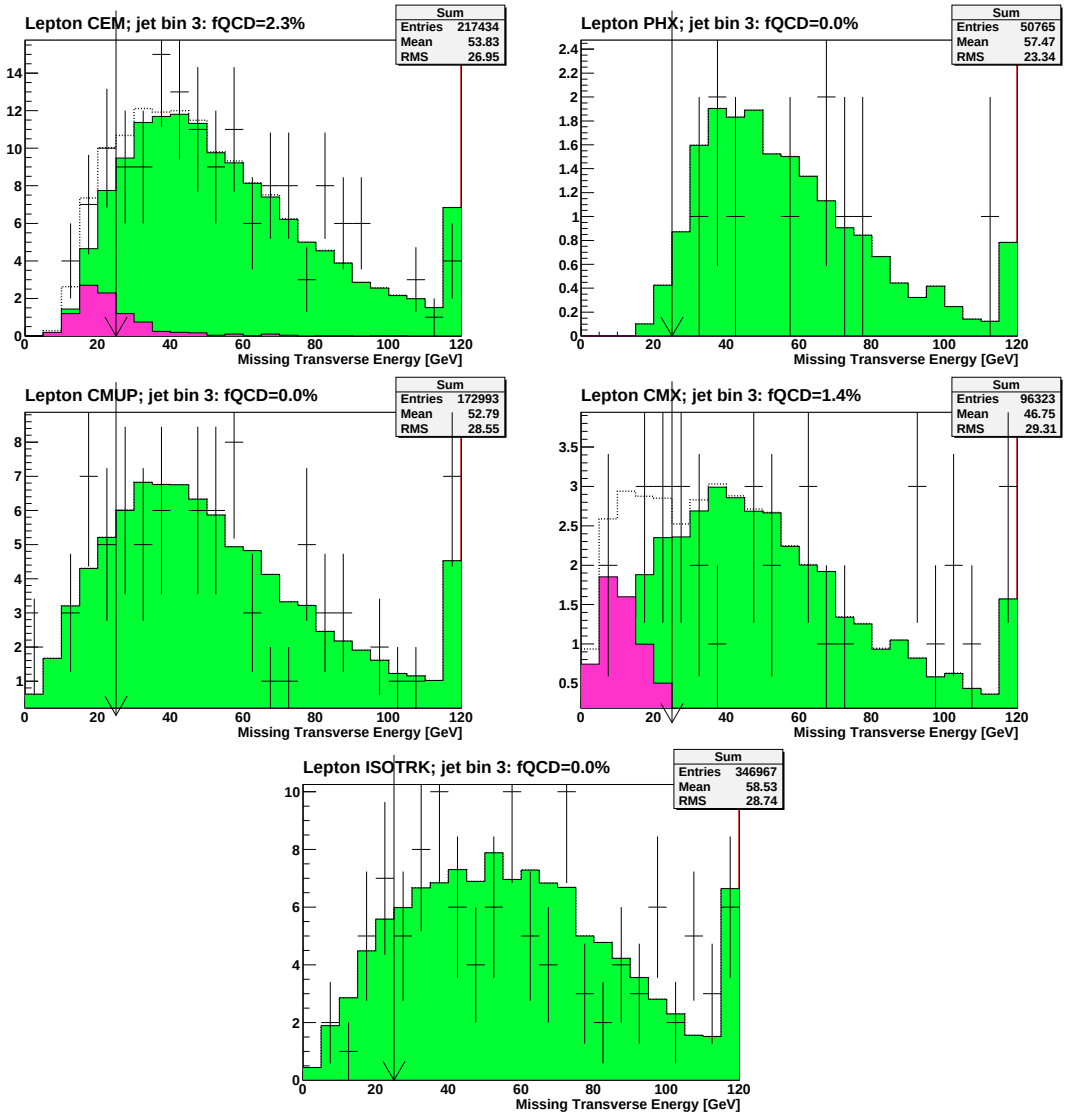


Figure 6.7: QCD fraction estimation for the double-SVT-tagged 3-jets sample. The x axis represents the fully corrected $\not\!\!E_T$ and the y axis represents the number of events. The QCD background is represented in pink and the W + jets backgrounds are in green. The MC-based processes are normalized and presented in blue (EW processes) and yellow ($t\bar{t}$). The dashed line represents the sum of all the backgrounds and the points represent the data. The vertical black arrow pointing at $\not\!\!E_T = 25$ GeV represents the $\not\!\!E_T$ cut we applied in this analysis. The resulting QCD fraction (labeled as fQCD) in the cut sample is shown in the plot title. The figures represent the CEM, PHX, CMUP, CMX, and EMC (labeled as ISOTRK in the figure) charged lepton categories.

where $N^{\text{pretag}}_{W+\text{jets}}$ is the predicted number of the pretag W + jets events, $N^{\text{pretag}}_{\text{t}\bar{\text{t}}+\text{EW}}$ is the MC-based predicted number of the pretag $t\bar{t}$ and electroweak (diboson, Z + jets) events, and $N^{\text{pretag}}_{\text{ST}}$ is the MC-based predicted number of the pretag single top events.

In the tagged sample, where our single top signal lies, the W + jets sample is broken down into W + HF and W + LF categories. The calculation of their normalization is treated differently for each as described below.

6.5.1 W + Heavy Flavor Estimate

W + HF events contain the $Wb\bar{b}$, $Wc\bar{c}$, and Wcj processes, which produce heavy flavor jets. W + LF events can also have heavy flavor jets from gluon splitting into a heavy flavor quark pair during parton showering. Thus, the number of W + HF events in the data sample is given by

$$N_{W+\text{HF}} = N^{\text{pretag}}_{W+\text{jets}} \cdot f_{\text{HF}} \cdot \epsilon_{\text{tag}} \cdot K_{\text{HF}} \tag{6.8}$$

where f_{HF} is the fraction of events in the total W + jets MC samples that contain jets matched to heavy flavor quarks. This fraction is calculated as

$$f_{\text{HF}} = \frac{N_{W+b\bar{b}} + N_{W+cj/c\bar{c}} + N_{W+\text{LF}}}{\sum N_{W+\text{jets}}} \tag{6.9}$$

where $N_{W+b\bar{b}}$, $N_{W+cj/c\bar{c}}$, and $N_{W+\text{LF}}$ are the number of events that contain heavy flavor jets in $Wb\bar{b}$, $Wc\bar{c}$ and Wcj, and the W + LF processes, respectively. $N_{W+\text{jets}}$ is the total number of pretag W + jets events, and ϵ_{tag} is the b-tagging efficiency as previously described. Both f_{HF} and ϵ_{tag} are calculated for the $Wb\bar{b}$, $Wc\bar{c}$, and Wcj MC samples separately, defining the rates of each process. The K-factor K_{HF} is a scale factor that is a correction to the MC heavy flavor fraction.

It was found that the MC simulation does not properly predict the heavy flavor fraction in the data. An additional factor needs to be applied to match the MC predictions to the data. The K-factor was derived from an independent sample, the W + 1 jet 1 b-tagged sample, which has high statistics and is dominated by

W + jets events. Two flavor-sensitive variables are used to evaluate the K-factor: the secondary vertex mass and a neural-network jet flavor separator as described in Section 7.2.1. As the $Wb\bar{b}$ and $Wc\bar{c}$ come from the same diagram with b and c quarks interchangeable (as shown in Figure 4.17(a)), a similar K-factor is expected. By free floating the $Wb\bar{b}$ and $Wc\bar{c}$ components, a K-factor of 1.4 ± 0.4 is obtained from a fit in the $W + 1$ jet sample. The K-factor for the Wcj component is obtained from a direct measurement of the Wc fraction using lepton charge correlations as in Ref. [136]. The measurement agreed with the MC predictions, thus we set the K-factor of the Wcj process to be 1.0 ± 0.3. The 30% uncertainty in the K-factor covers the differences between the two fitting variables and also approximates our uncertainty in extrapolating this fraction to W + 2 and 3 jets events [37].

6.5.2 W + Light Flavor Estimate

The last background to consider is the W + LF process. W + LF events that contain heavy flavor jets are removed from the MC samples, a process called "heavy flavor removal." The resulting sample has no heavy flavor jets. However, some of these events can still be tagged by the SECVTX algorithm by mistakenly reconstructing a secondary vertex when poorly reconstructed tracks seem to cross each other near the origin. We therefore also call the W + LF processes in the tagged sample the "mistag" background.

The number of expected W + LF events is computed analogously to the W + HF case with a data-driven method

$$N_{\mathrm{W+LF}} = N^{\mathrm{pretag}}_{\mathrm{W+jets}} \cdot (1 - f_{\mathrm{HF}} \cdot K_{\mathrm{HF}}) \cdot \epsilon_{\mathrm{mistag}} \tag{6.10}$$

where $\epsilon_{\mathrm{mistag}}$ is the predicted fraction of mistag events in the data sample. $\epsilon_{\mathrm{mistag}}$ is obtained by summing up the mistag probability of events in the pretag data sample. The mistag matrix used in our analysis was validated for the entire 7.5 fb^{-1} dataset.

We use the W + LF ALPGEN samples to predict the kinematic features of W + LF events, where each event is weighted by its mistag probability.

6.6 Signal and Background Predictions

Tables 6.2 and 6.3 show the signal and background yield estimates in the single and double SVT tagged categories, respectively. The uncertainties include systematic uncertainties on the lepton ID scale factor, the quality primary vertex requirement, the trigger efficiency, the b-tagging scale factor, the QCD estimate, the mistag matrix, and the K-factor. They also include the uncertainty on the integrated luminosity of the data and the cross section uncertainties for MC-based processes. Further discussion of the systematics uncertainties is given in Chapter 8.

Table 6.2: Summary of the predicted number of signal and background events with exactly one b tag from Method II, with systematic uncertainties included. The three single top production channels are listed below the backgrounds, and the total number of observed data events passing the event selection is shown at the bottom.

Process	W + 2 jets	W + 3 jets
$t\bar{t}$	473.9 ± 49.1	1067.2 ± 108.6
WW	147.7 ± 20.6	48.3 ± 6.6
WZ	52.9 ± 5.8	14.4 ± 1.6
ZZ	1.7 ± 0.2	0.7 ± 0.1
Z + jets	118.3 ± 15.5	45.6 ± 5.9
$Wb\bar{b}$	1452.1 ± 436.9	434.1 ± 131.1
$Wc\bar{c}$	766.1 ± 233.0	254.5 ± 77.4
Wcj	583.2 ± 177.4	127.7 ± 38.8
W + LF	1459.3 ± 148.5	432.6 ± 47.0
non-W	315.7 ± 126.3	141.5 ± 56.6
t-channel	192.8 ± 25.3	84.0 ± 10.6
s-channel	127.6 ± 11.3	42.8 ± 3.8
Wt-channel	16.2 ± 4.3	25.7 ± 6.8
Total Prediction	5707.4 ± 876.6	2719.1 ± 292.9
Observed	5533	2432

Table 6.3: Summary of the predicted number of signal and background events with exactly two b tags from Method II, with systematic uncertainties included. The three single top production channels are listed below the backgrounds, and the total number of observed data events passing the event selection is shown at the bottom.

Process	W + 2 jets	W + 3 jets
$t\bar{t}$	98.3 ± 14.5	284.2 ± 41.8
WW	1.1 ± 0.3	1.2 ± 0.3
WZ	8.8 ± 1.3	2.4 ± 0.4
ZZ	0.3 ± 0.0	0.1 ± 0.0
Z + jets	4.8 ± 0.7	2.7 ± 0.4
$Wb\bar{b}$	182.9 ± 56.1	64.7 ± 19.8
$Wc\bar{c}$	10.2 ± 3.2	7.0 ± 2.2
Wcj	7.8 ± 2.4	3.5 ± 1.1
W + LF	7.4 ± 1.5	5.4 ± 1.1
non-W	6.8 ± 3.5	3.4 ± 3.2
t-channel	5.9 ± 1.0	14.7 ± 2.4
s-channel	32.3 ± 4.4	11.6 ± 1.6
Wt-channel	0.7 ± 0.2	2.3 ± 0.6
Total Prediction	367.3 ± 65.7	403.1 ± 52.6
Observed	335	355

CHAPTER SEVEN

Neural Network Discriminant

The measurement of the single top quark cross section presents substantial experimental challenges. It suffers from a low Standard Model production rate and a large set of kinematically similar backgrounds. As presented in Tables 6.2 and 6.3, the systematic uncertainty from the background prediction is more than twice the size of the expected signal. Thus, simply counting the number of selected events will not yield a precise measurement of the single top quark cross section. Further separation of the signal from the backgrounds is required. In this case, multivariate techniques can be used to discriminate between signal and background processes by exploiting the kinematic differences of signal and background events. A region of phase space can be found where the signal to background ratio is strongly enhanced so that the background uncertainty does not bury the signal. In this chapter, we will describe the artificial neural network (NN) used in this analysis for signal discrimination.

7.1 Neural Network Event Classifier

In order to separate the signal from background processes, we exploit an artificial NN technique using the NeuroBayes® package [137,138]. The output from a NN is a continuous probability density function that classifies the probability of a given event being signal or background. Events that are classified to be signal-like are used to measure the signal cross section while events classified to be background-like are used to constrain the background processes. Templates are constructed from the NN output for each signal and background process considered. With the inclusion of the uncertainty parameters listed in Chapter 8, these templates will be fit

to the background and signal + background predictions in a manner described in Chapter 9.

7.1.1 Neural Network Technique

For an event i in a set of events, there exists a vector of measurements $\vec{x}_i$ that are correlated to the variable t_i, which is distributed according to a probability density function $f(t)$. Thus, an overall probability density function $f(t, \vec{x})$ is defined by all the events in the whole phase space. Assuming that this probability density function can be estimated from a large but finite number of examples, the NN aims to achieve a smooth estimate of the conditional probability density function $f(t|\vec{x})$ for a given measurement vector $\vec{x}_i$. We use a three-layered, feed-forward NN from the NeuroBayes® package. As illustrated in Figure 7.1, it consists of one input layer with input nodes for each input variable, one hidden layer with an arbitrary number of hidden nodes, and one output node, which gives a continuous output within the interval [–1, 1].

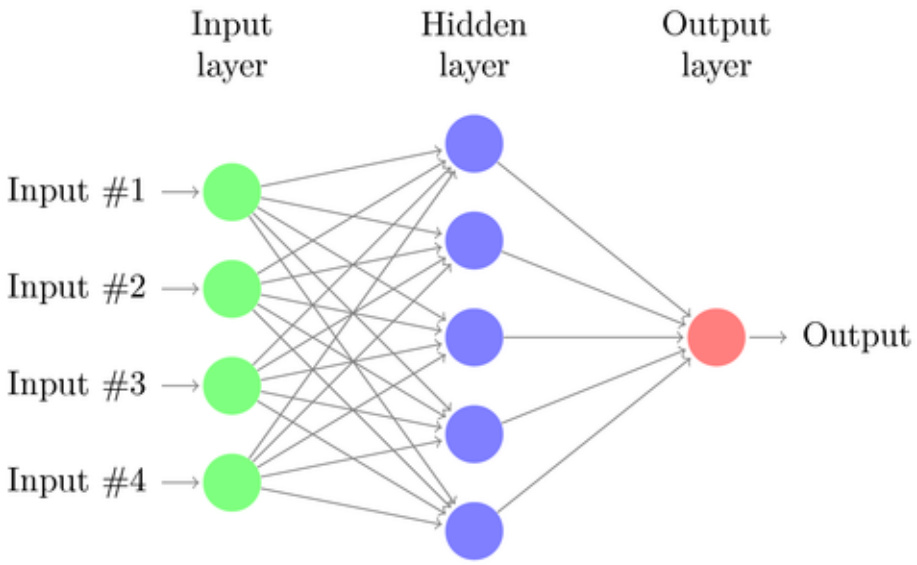


Figure 7.1. A schematic diagram of a three-layer NN.

First, all the input variables are processed by a robust automatic *preprocessing* procedure in NeuroBayes®. All input variables are normalized and decorrelated such that the covariance matrix of the new set of input variables obtained is given by a unit matrix. Outliers far away from the bulk of the values for each variable are treated by a Bayesian regularization scheme. The significance of each input variable is computed automatically at the end of the preprocessing. The correlation matrix for all preprocessed input variables is calculated, including the correlation of all variables to the target. One by one, the variables are omitted to determine the loss of correlation to the signal output. The variable with the smallest loss of correlation is discarded leading to an $n-1$ dimensional correlation matrix. This procedure is repeated to find the least important of the remaining variables in the reduced correlation matrix. The significance of each variable is calculated by dividing the loss of correlation caused by its removal by the square root of the sample size. After the preprocessing, the list of input variables is provided, sorted by their significance. It is possible to cut on the significance to remove those variables highly correlated with other variables.

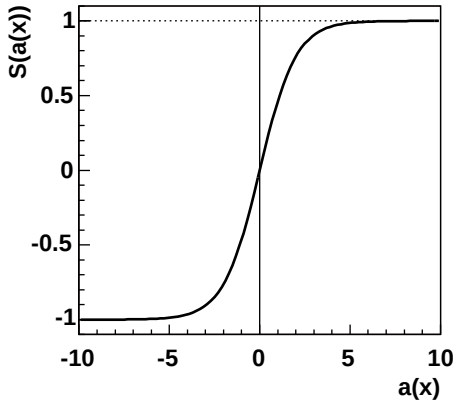


Figure 7.2. The transfer sigmoid activation function S(a(**x**)) as given by Eq. 7.1.

The output of each node is determined by a symmetric sigmoid function

$$S(a\,(\mathbf{x})) = \frac{2}{1+e^{-a(\mathbf{x})}} - 1 \tag{7.1}$$

where $S(x)$ is a transfer function that maps $[-\infty, +\infty]$ to the interval [–1, 1] as shown in Figure 7.2. It is most sensitive to a relatively small range around $a(\mathbf{x}) = 0$. For very large or very small values of $a(\mathbf{x})$, a saturation effect is reached.

The input to the sigmoid function is a biased weighted sum defined as

$$a_k(\mathbf{x}) = \sum_j \omega_{jk} x_j + \mu_{0,k}. \tag{7.2}$$

For each node k, the value of the node j in the previous layer x_j is weighted by an arbitrary weight ω_{jk}. The bias $\mu_{0,k}$ is implemented as the threshold of node k. It shifts the mean of the weighted sum of the input distribution to the linear part of the sigmoid functions where the sigmoid function is most sensitive.

The output of the NN is therefore calculated by

$$O = S\left(\sum_{k=0}^{M} \omega_k^{2\to3} \cdot S\left(\sum_{j=0}^{D} \omega_{jk}^{1\to2} x_j + \mu_{0,k}\right)\right) \tag{7.3}$$

where D is the number of input nodes and M is the number of hidden nodes. x_j is the input value from the node j in the input layer and $\omega_{jk}^{1\to2}$ is the weight of the node j in the input layer while $\omega_k^{2\to3}$ is the weight of the node k in the hidden layer.

The training of a NN is the task of minimizing the difference between the target output and the one obtained from the NN. By using the iterative back-propagation algorithm, the weights are adjusted until the minimum difference between the target and NN outputs is reached. The difference is calculated by an error function. We use the entropy error (E_D) function as in Eq. 7.4 because of its advantage in classification problems and Bayesian regularization. The E_D function is

$$E_D = \sum_j \omega_j E_D^j = \sum_j \omega_j \sum_i \log\left(\frac{1}{2} \cdot (1 + T_{ji} \cdot O_{ji} + \epsilon)\right), \tag{7.4}$$

where T_{ji} is a binary number to classify event i as signal (+1) or background (–1) for output node j, and O_{ji} denotes the output for event i in node j. In the case of a completely wrong classification like $O_{ji} = 1$ for $T_{ji} = -1$, it leads to an

infinitely large E_D. In order to avoid such numerical problems at the beginning of the training, a small regularization constant ϵ is introduced. The constant is reduced in each training iteration and is zero after a few iterations.

During the training process, the back-propagation algorithm calculates how the error depends on the weights, inputs, and outputs. The adjustment of the weight $\Delta\omega_{ij}$ is obtained by a *gradient descent* method: $\Delta\omega_{ij} = -\eta\frac{\partial E_D}{\partial\omega_{ij}}$. It is proportional to the gradient of the error function where the constant η is the step width adapted individually for each weight during the training.

The generalization ability of a NN depends on a balance between the information in the training examples and the complexity of the network. The complexity of the NN is normally implied by its number of free parameters such as the number of weights and the number of thresholds. If the network is very complex and there is little information in the training set, the network will be over-fitting the data. In the opposite situation, the network will be under-fitting the data. Regularization techniques are used to improve the general performance of neural networks, which is achieved by *weight decay* regularization in NeuroBayes®. Weights are prevented from growing too large (unless it is really necessary) by adding a term to the error function that penalizes large weights. This leads to a new term, $-\lambda\omega_i$, in the gradient descent formula $\Delta\omega_{ij} = -\eta\frac{\partial E_D}{\partial\omega_{ij}} - \lambda\omega_{ij}$. λ is a decay parameter governing how strongly large weights are penalized. By exploiting the Bayesian approach, NeuroBayes® uses Bayesian statistics to incorporate *a priori* knowledge in the regularization. The Bayesian Regularization provides objective criteria for the determination of the decay parameters and the regularization function. During the training, connections are pruned away for weights that become insignificant. Thus, the architecture of the NN is changed and the number of free parameters is lowered. In this way, pruning improves the signal-to-background ratio.

7.1.2 Neural Network Training

For the measurement of the single top quark cross section, eight NN classifiers are trained. As previously described, we separate our samples into four jet-tag categories: 2J1T, 2J2T, 3J1T, and 3J2T. According to the characteristics of the lepton triggers, we further separate the jet-tag categories by lepton type: TLC and EMC. Thus, a total of eight NN categories are studied in this analysis.

For the training of a NN, it is necessary to arrange the composition of a training sample to have large statistics and a reasonable signal and background ratio. Due to a different background composition in different categories, we make separate training samples for each category. All training samples are composed in such a way that the relative signal process contributes 50% to the total number of events. Due to the different event topology of the s- and t-channels, we consider the t-channel process as signal in a certain channel while the s-channel process is unknown for that NN and vice versa. The relative fractions of all considered background processes are given by the predictions from Method II (see Tables 6.2 and 6.3), except for QCD events, whose event properties are hard to model with MC simulation. Our studies show that even if we train without QCD processes, the final NN can correctly distinguish the QCD events as background-like.

Table 7.1: Composition of the training samples used to train the neural networks for TLC lepton types.

Category	2J1T	2J2T	3J1T	3J2T
t-channel	50.0%	—	50.0%	50.0%
s-channel	—	50.0%	—	—
$t\bar{t}$	4.2%	14.4%	21.3%	37.8%
$Wb\bar{b}$	14.7%	29.4%	9.3%	9.2%
$Wc\bar{c}$	7.8%	1.6%	5.4%	1.0%
Wcj	6.1%	1.3%	2.8%	0.5%
W + LF	14.3%	1.1%	9.1%	0.7%
Diboson	2.0%	1.6%	1.3%	0.5%
Z + jets	0.9%	0.6%	0.8%	0.3%

Table 7.2: Composition of the training samples used to train the neural networks for EMC lepton types.

Category	2J1T	2J2T	3J1T	3J2T
t-channel	50.0%	—	50.0%	50.0%
s-channel	—	50.0%	—	—
$t\bar{t}$	6.2%	18.0%	24.0%	39.5%
$Wb\bar{b}$	13.1%	25.5%	8.1%	7.6%
$Wc\bar{c}$	7.0%	1.5%	4.8%	0.9%
Wcj	4.6%	1.0%	2.1%	0.4%
W + LF	14.9%	1.2%	8.4%	0.7%
Diboson	2.0%	1.4%	1.3%	0.5%
Z + jets	2.1%	1.3%	1.3%	0.5%

A measurement of the single top quark cross section is limited by the statistics of the collected data and our understanding of various systematic uncertainties. Either by increasing the number of selected events or constraining the systematic uncertainties, we can achieve a more precise measurement. However, improvement of certain systematic uncertainties requires extensive studies that are impossible in our case. Hence we exploit a different path by training systematic-constrained neural networks.

In our NN training, instead of composing our training samples with only the standard background samples (Tables 7.1 and 7.2), we compose a systematic-mixed training sample that includes additional events from different systematically-varied background samples. We exploit a region of phase space where the NN is not sensitive to certain systematic variations. In addition, we increase the NN purity by increasing the statistics of the training sample.

We include the simulated events with variations in jet energy scale (*JES*) and the factorization and renormalization scale (Q^2) in our training sample (Table 7.3), since these two systematic uncertainties contribute most to the uncertainty of our cross section measurement. We expect to gain about 3% improvement in our cross section measurement by using the systematic-constrained NN.

Table 7.3: Composition of the training samples with systematically-varied background samples.

Category	Standard	JES Plus	JES Minus	Q^2 More	Q^2 Less
t-channel	X	X	X		
s-channel	X	X	X		
$t\bar{t}$	X	X	X		
$Wb\bar{b}$	X	X	X	X	X
$Wc\bar{c}$	X	X	X	X	X
Wcj	X	X	X	X	X
W + LF	X	X	X	X	X
Diboson	X	X	X		
Z + jets	X	X	X		

7.2 Definition of Input Variables

In order to separate signal from background events in the NN, we need to exploit event properties that differ in signal and background processes. Single top quark events have distinctive energy and angular properties while background events also have distinctive features. By exploiting these characteristic features, we select various variables to feed into the NN training. The input variables can be assembled into three groups: those that are directly measured in the detector, those that are reconstructed from the measured values, and a few that are calculated by advanced algorithms and are based on the measured values. More than one hundred input variables are investigated in each category. However, most of the new variables are either highly correlated to the variables used in the previous analysis or not significant to the NN discriminant. Thus, we use the same variables as the previous analysis [37] to maintain consistency.

For the single top quark discriminant, it is important to reconstruct the top quark correctly. Correct reconstruction of the top quark will improve the discrimination between the top quark processes and the W + heavy flavor processes. The top quark four-momenta is calculated from the reconstructed W boson and the b-

tagged jet. The W boson is reconstructed from the measured tight lepton and the reconstructed neutrino. The p_z component of the neutrino is calculated as described in Section 5.2.3. $Q_\ell \cdot \eta$ is the product of the charge of the lepton Q_ℓ and the jet pseudorapidity η. For events with more than one b-tagged jet, the b-tagged jet with larger $Q_\ell \cdot \eta$ is assigned to be the jet from the top quark decay.

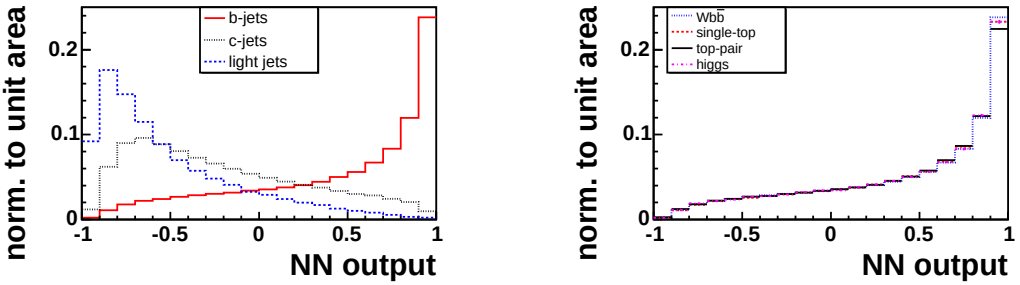
7.2.1 Jet Flavor Separator

From Section 3.6, we identify b-quark jets by requiring a reconstructed secondary vertex. Due to the long lifetime and mass of charm hadrons and the false reconstruction of secondary vertices in light jets, a sizable fraction of the b-tagged jets have no real B hadrons in them. By taking advantage of the longer lifetime ($\tau \approx 1.6$ ps) and the larger mass ($m \approx 5$ GeV/c^2) of B hadrons, we can further separate the tagged jets without B hadrons from those containing B hadrons by extending the secondary vertex requirement using reconstructed quantities that differentiate the two classes of jets. The SECVTX algorithm is merely a set of binary decisions based on a few requirements that are applied to the track information resulting from the long B-hadron lifetime. A NN is trained on simulated SECVTX jets to make full use of all discriminating quantities of the B hadron.

In general, the invariant mass of the tracks associated with the reconstructed vertex from a B hadron decay is larger than for those jets that do not contain a B hadron. The number of tracks and the transverse decay length ($\Delta L_{xy}/\sigma_{xy}$) are also larger for B hadron vertices. In addition, tracks from charged particles from a B hadron decay have larger impact parameters and higher transverse momentum relative to the direction of the jet. The semileptonic B hadron decay also increases the number and p_T of electrons and muons in b-tagged jets.

By using the NeuroBayes® package, the NN is trained with simulated events of single top quark production and the main background processes. Processes with

a secondary vertex from B hadron decays such as single top quark, $t\bar{t}$, and $Wb\bar{b}$ processes are treated as signal events. Processes containing no b quarks but have charm and light flavor jets are treated as background, namely $Wc\bar{c}$, Wcj, and W + LF processes. For jets containing a B hadron, the output of the network accumulates near +1; jets without a B hadron produce an output close to –1. The resulting NN provides a measurement of the probability of a true b quark being present in a SECVTX-tagged jet as shown in Figure 7.3(a). It is normally referred to as the jet flavor separator or KIT flavor separator (named after the group at the Karlsruhe Institute of Technology that developed it) [139]. The network has a similar shape for different physics processes containing a b quark as shown in Figure 7.3(b). This indicates that the jet flavor separator is sensitive to the properties of b-quark jets and is process independent.



(a) (b)

Figure 7.3: (a) The NN output from the jet flavor separator for b, c, and light flavor jets. (b) The output distribution for b jets from several processes. Both plots are extracted from Ref. [140].

Since the NN jet flavor separator is trained with Monte Carlo samples, it is important to validate the modeling of its input and output distributions with the collected data. In the b-enriched region, a comparison of MC jets with B hadron decays and double-SECVTX-tagged dijet events in the data is performed. In this method, in order to purify the b content of the sample, one jet is SECVTX-tagged and required to have a high-momentum electron. The other SECVTX-tagged jet is

used to compare the distributions of the jet flavor separator output for both the MC events and the dijet events. It was found that the distributions of these jets are well matched [141].

For the mistag-enriched region, negative-tagged dijet data samples are tested with MC samples. A discrepancy between the data and the MC prediction is observed in the output shape, thus a correction function is applied to MC samples to match the output distribution from the data. The correction function is parameterized by the transverse energy and the number of tracks of the jet, along with the sum of the transverse energies of the event. In this analysis, the correction function is applied to the light-flavored and charm jets in MC samples. The uncorrected jet flavor separator outputs are used to evaluate the systematic uncertainty as described in Section 8.2.1.

For the W + LF process, the prediction of its contribution is given by Method II as described in Section 6.5.2. We use the W + LF MC samples to predict the shape distribution, where each event is weighted by the mistag matrix. However, with the requirement of at least one b-tagged jet in the signal region, the number of W + LF MC events will be significantly reduced such that they can't correctly predict the W + LF shape distributions, because most of the W + LF MC events have no secondary vertex and jet flavor separator output value. To overcome this problem, a random jet flavor separator value is assigned to each taggable jet in the event based on the distribution in the light-flavor data templates.

For non-W multijet events, which have a lower number of b-tagged jets, it is more complicated to obtain a jet flavor separator distribution because the flavor composition of the jets in non-W events is poorly known. Since a non-W event must contain a fake lepton from a mis-identified jet or a real lepton from a heavy-flavor jet in addition to mismeasured $\not\!\!E_T$, the flavor composition of non-W events passing the selection relies on the details of the detector response, which is difficult

to model. This is resolved by constraining the flavor fraction with QCD-enriched sideband data. Templates for three flavors are constructed and fit to the jet flavor separator distribution of b-tagged jet data in the $15 < \not\!\!E_T < 25$ GeV region. This gives an approximate flavor composition of 45% for b quark jets, 40% for c quark jets, and 15% for light-flavor jets. Each non-W event is randomly assigned a flavor according to the flavor composition mentioned above and a jet flavor separator value from the corresponding flavor distribution. To take into account the uncertainty in the fit and the uncertainty in the extrapolation to the signal sample, an alternative flavor composition of 60% b quark jets, 30% c quark jets, 10% light-flavor jets is chosen [142]. This alternative flavor composition is the most b-like scenario on the flavor measurement and is used as a systematic uncertainty.

7.2.2 List of Input Variables

The input variables for all NN classifiers are listed in Table 7.4. The variables are separated into four jet-tag categories. For the TLC and EMC categories, we use the same input variables, but train with different training samples. The input variables in each category are described below, ordered by their significance in the NN discriminant of the TLC lepton.

7.2.2.1. *Input variables for the 2J1T category.* In total, 14 input variables are used in the 2J1T NN discriminant to distinguish t-channel signal events from background events. They are ordered by decreasing significance as in the TLC NN.

- $M^{\ell\nu b}$: the reconstructed top quark mass is built out of the charged lepton, the reconstructed neutrino, and the b-tagged jet, as illustrated in Figure 7.4(a). For top quark processes, this variable peaks around 170 GeV/c^2 while the remaining background processes culminate in the smaller values.
- *Jet Flavor Separator*: the NN output of the jet flavor separator for the b-tagged jet. It is used to separate processes with B hadrons from processes with charm or light flavor quarks due to mis-reconstructed secondary vertices

Table 7.4: Input variables used for training the different neural networks. The symbol X indicates that the variable is used in the corresponding category.

Input Variable	2J1T	2J2T	3J1T	3J2T
$M_{\ell\nu b}$	X	X	X	
$M_{\ell\nu bb}$		X		X
$M_T^{\ell\nu b}$	X	X	X	X
M_{jj}	X	X	X	X
M_T^W	X	X		
$E_T^{b_{\text{top}}}$		X	X	
$E_T^{b_{\text{other}}}$				X
$\sum E_{\text{T}}^{jj}$			X	X
E_T^{light}	X			X
p_T^{ℓ}	X			
$p_T^{\ell\nu jj}$			X	X
H_T	X		X	
$\not{E}_T$		X		
$\not{E}_{T,\text{sig}}$			X	
$\cos\theta_{\ell j}^t$	X		X	X
$\cos\theta_{\ell W}^W$	X			
$\cos\theta_{\ell W}^t$	X			
$\cos\theta_{jj}^t$		X		X
$Q \times \eta$	X		X	X
η_ℓ		X		
η_W	X	X		
$\sum \eta_j$	X		X	
$\Delta\eta_{jj}$			X	X
$\Delta\eta_{t,\text{light}}$			X	
$\sqrt{\hat{s}}$				X
Centrality				X
Jet Flavor Separator	X	X	X	

(see Figure 7.4(d)). This variable is a powerful tool for differentiating the signal from the W + LF and $Wc\bar{c}$ and Wc backgrounds.

- M_{jj}: the invariant mass of the two jets. It peaks at higher values for top quark events compared to the backgrounds. For diboson processes, the peak is near the W and Z boson mass as shown in Figure 7.4(f). This is expected since both jets come from the hadronic decay of the W or Z boson. For W + jets and QCD events, the invariant dijet mass is much lower due to the fact that the two jets come from a gluon. In top quark events, both jets are expected to come from different physics objects, so the corresponding distribution has a larger dijet mass.

- $Q \times \eta$: the product of the charge of the lepton and the pseudorapidity of the light-quark jet [143, 144]. As can be seen in Figure 7.4(i), this variable has a very distinctive feature for signal events. This characteristic feature comes from the t-channel single top quark events. A proton consists of two u and only one d valence quark, thus it is more likely that a single top quark in the t-channel is produced by an initial-state u-quark. The initial b quark is a sea quark carrying a much smaller fraction of the proton momentum. Since the light quark arises from the virtual W-boson exchange of the initial valence quark, it has a stronger tendency of propagating along the direction of the valence quark. Hence, the light-quark jet in the t-channel propagates in the proton direction, which is equivalent to positive pseudorapidity. By charge conservation, the sign of the lepton charge determines the sign of the top quark's charge, which is correlated with the sign of the pseudorapidity of the light-flavor jet. The resulting $Q \times \eta$ lies in the positive forward region (higher values of $|\eta|$). A smaller fraction of single top production in the t-channel is initiated by the initial d quark originating from the proton, resulting in the opposite charge-η correlation. Thus, the asymmetry of the t-channel is induced by the parton distribution function of the proton. The charge conjugate applies for antitop-quark production. $t\bar{t}$ and W + jets backgrounds lack such a correlation, resulting in a symmetric distribution.

- $M_T^{\ell\nu b}$: the transverse mass of the reconstructed top quark (see Figure 7.5(a)). It is calculated as

$$\mathrm{M}_{\mathrm{T}}^{\ell\nu b} = \sqrt{\left(p_{\mathrm{T}}^{\ell} + p_{\mathrm{T}}^{\nu} + p_{\mathrm{T}}^{b}\right)^2 - \left(p_x^{\ell} + p_x^{\nu} + p_x^{b}\right)^2 - \left(p_y^{\ell} + p_y^{\nu} + p_y^{b}\right)^2} \qquad (7.5)$$

- $\cos\theta_{\ell j}^{t}$: the cosine of the angle between the charged lepton and the light-quark jet (produced in the single top production process) in the top quark's rest frame. The top quarks in the s- and t-channel processes are produced 100% polarized along the direction of the down-type quark in the top quark rest frame. Due to the short lifetime of the top quark, the W boson from top quark decay carries the polarization information of the top quark. Because of the V–A angular dependence of the W boson vertex, the $\cos\theta_{\ell j}^{t}$ variable

tends to be positive for t-channel events. As shown in Figure 7.5(d), this variable separates the signal from almost all background processes.

- E_T^{light}: the transverse energy of the light-quark jet, which is the untagged jet in this category. Figure 7.5(g) shows this distribution, which is much harder for top quark processes.
- $\cos\theta_{\ell W}^{W}$: the cosine of the polar angle between the charged lepton and the reconstructed W boson in the W-boson rest frame.
- η_W: the pseudorapidity of the reconstructed W boson.
- M_T^W: the transverse mass of the reconstructed W boson.
- $\sum \eta_j$: the scalar sum of the pseudorapidities of the two jets.
- p_T^{ℓ}: the transverse momentum of the charged lepton.
- H_T: the scalar sum of the transverse energies of the charged lepton, the reconstructed neutrino, and all the jets in the events.
- $\cos\theta_{\ell W}^{W}$: the cosine of the polar angle between the charged lepton in the W-boson rest frame and the reconstructed W boson in the top quark rest frame.

7.2.2.2. *Input variables for the 2J2T category.* The two-jet, two-tag NN is trained with 11 input variables to distinguish s-channel signal events from background events. In this category, the b-tagged jet with larger $Q_\ell \cdot \eta$ is assigned to be the jet from the top quark decay. The variables are ordered by decreasing significance as in the TLC NN.

- $M_{\ell\nu bb}$: the invariant mass of the charged lepton, the reconstructed neutrino, and the two b-tagged jets in the event.
- M_{T}^{W}: the transverse mass of the reconstructed W boson.
- $M_{\mathrm{T}}^{\ell\nu b}$: the transverse mass of the reconstructed top quark.
- $\cos\theta_{jj}^{t}$: the cosine of the angle between the two b-tagged jets in the top quark rest frame.
- $M_{\ell\nu b}$: the reconstructed top quark mass.
- *Jet Flavor Separator*: the sum of the neural-network output values of the jet flavor separator for both b-tagged jets.

- η_W: the pseudorapidity of the reconstructed W boson.
- $\not\!\!E_T$: the transverse energy of the reconstructed neutrino.
- M_{jj}: the invariant mass of the two b-tagged jets.
- $E_{\mathrm{T}}^{b_{\mathrm{top}}}$: the transverse energy of the b-tagged jet from the top quark decay.
- η_ℓ: the pseudorapidity of the charged lepton.

7.2.2.3. *Input variables for the 3J1T category.* The three-jet, one-tag NN is trained with 18 input variables to distinguish t-channel signal events from background events. The variables are ordered by decreasing significance as in the TLC NN.

- $Q \times \eta$: the product of the charge of the lepton and the pseudorapidity of the light-quark jet, which is the most energetic non-b-tagged jet in the event.
- *Jet Flavor Separator*: the neural-network output of the jet flavor separator of the b-tagged jet.
- H_T: the scalar sum of the transverse energies of the charged lepton, the reconstructed neutrino, and the three jets in the event.
- $M_{j_1 j_3}$: the invariant mass of the most and third-most energetic jets.
- $M_{\ell\nu b}$: the reconstructed mass based on the charged lepton, the reconstructed neutrino, and the b-tagged jet, which corresponds to a top quark mass.
- $p_T^{\ell\nu jj}$: the transverse momentum sum of the charged lepton, the neutrino, and all the jets in the event.
- $M_{j_2 j_3}$: the invariant mass of the second- and third-most energetic jets.
- $\cos\theta_{\ell j}^t$: the cosine of the angle between the charged lepton and the most energetic light-quark jet in the top quark rest frame.
- $\Delta\eta_{jj}$: the difference in pseudorapidity between the two most energetic jets.
- $\sum \eta_j$: the scalar sum of the pseudorapidities of the three jets in the event.
- $\not\!\!E_{T,\mathrm{sig}}$: the significance of the missing transverse energy $\not\!\!E_T$ as introduced in Section 5.3.5.
- $M_{j_1 j_2 j_3}$: the invariant mass of the three jets in the event.

- $\sum E_T^{j_2 j_3}$: the sum of the transverse energies of the second- and third-most energetic jets.
- E_T^b: the transverse energy of the b-quark jet from the top quark decay.
- $M_T^{\ell\nu b}$: the reconstructed top quark transverse mass based on the charged lepton, the reconstructed neutrino, and the b-tagged jet.
- $\Delta\eta_{t,\text{light}}$: the difference in pseudorapidity between the reconstructed top quark and the most energetic light-quark jet.
- $\sum E_T^{j_1 j_3}$: the transverse energy of the first- and third-most energetic jets.
- $\sum E_T^{j_1 j_2}$: the transverse energy of the two most energetic jets.

7.2.2.4. *Input variables for the 3J2T category.* The three-jet, two-tag NN is trained with 15 input variables to distinguish t-channel signal events from background events. In this category, the b-tagged jet with larger $Q_\ell \cdot \eta$ is assigned to be the jet from the top quark decay. The variables are ordered by decreasing significance as in the TLC NN.

- $Q \times \eta$: the product of the charge of the lepton and the pseudorapidity of the light-quark jet.
- $p_T^{\ell\nu jj}$: the transverse momentum sum of the system, which is composed of the reconstructed top quark and the remaining two jets in the event.
- $M_{j_1 j_2}$: the invariant mass of the two most energetic jets.
- $\cos\theta_{\ell j}^t$: the cosine of the angle between the charged lepton and the light-quark jet in the top-quark rest frame.
- $\sum E_T^{j_1 j_2}$: the transverse energy of the two most energetic jets.
- $M_{j_1 j_3}$: the invariant mass of the first- and third-most energetic jets.
- $\Delta\eta_{j_2 j_3}$: the difference in pseudorapidity between the second- and third-most energetic jets.
- $E_T^{b_{\text{other}}}$: the transverse energy of the b-tagged jet that is not from the top-quark decay, which is chosen from the b-tagged jet with smaller $Q \times \eta$ value between the two b-tagged jets.
- $M_T^{\ell\nu b}$: the reconstructed top quark transverse mass based on the charged lepton, the reconstructed neutrino, and the b-tagged jet, which is considered to come from top quark decay.

- $\Delta\eta_{j_1 j_2}$: the difference in pseudorapidity between the two most energetic jets.
- E_T^{light}: the transverse energy of the least energetic jet.
- *Centrality*: the sum of the transverse energies of the two most energetic jets, divided by $\sqrt{\hat{s}}$.
- $\sqrt{\hat{s}}$: the energy of the center-of-mass system of the hard interaction.
- $\cos\theta^t_{j_1 j_2}$: the cosine of the angle between the two most energetic jets in the top quark rest frame.

For a NN training, it is crucial that all input variables are correctly constructed and consistent with the data. Since we trained the NNs with the MC training samples and then use the resulted NNs to classify the data events, we need to check the modeling of the input variables by comparing their MC-predicted distribution with the distributions of the data samples. The MC-predicted distribution is built by stacking the distribution of each MC process, whose integrated area is scaled to the prediction given by Chapter 6. Figures 7.4–7.7 show the distributions of the input variables for the TLC NN with two jets and one b tag. The shapes of both distributions are compared and reasonable agreement is found in the signal regions and the zero-tag sideband region. Discrepancies like the negative region of Figure 7.4(k) is considered and covered by a systematic uncertainty as discussed in Section 8.2.5. Because the jet flavor separator relies on the SECVTX tags, there is no jet flavor separator distribution in the zero-tag sideband region as shown in Figure 7.4. Input variables in each category (2J1T, 2J2T, 3J1T, and 3J2T) have been checked for TLC and EMC lepton types [145]. They show good agreement and the discrepancies are covered by systematic uncertainties listed in Chapter 8.

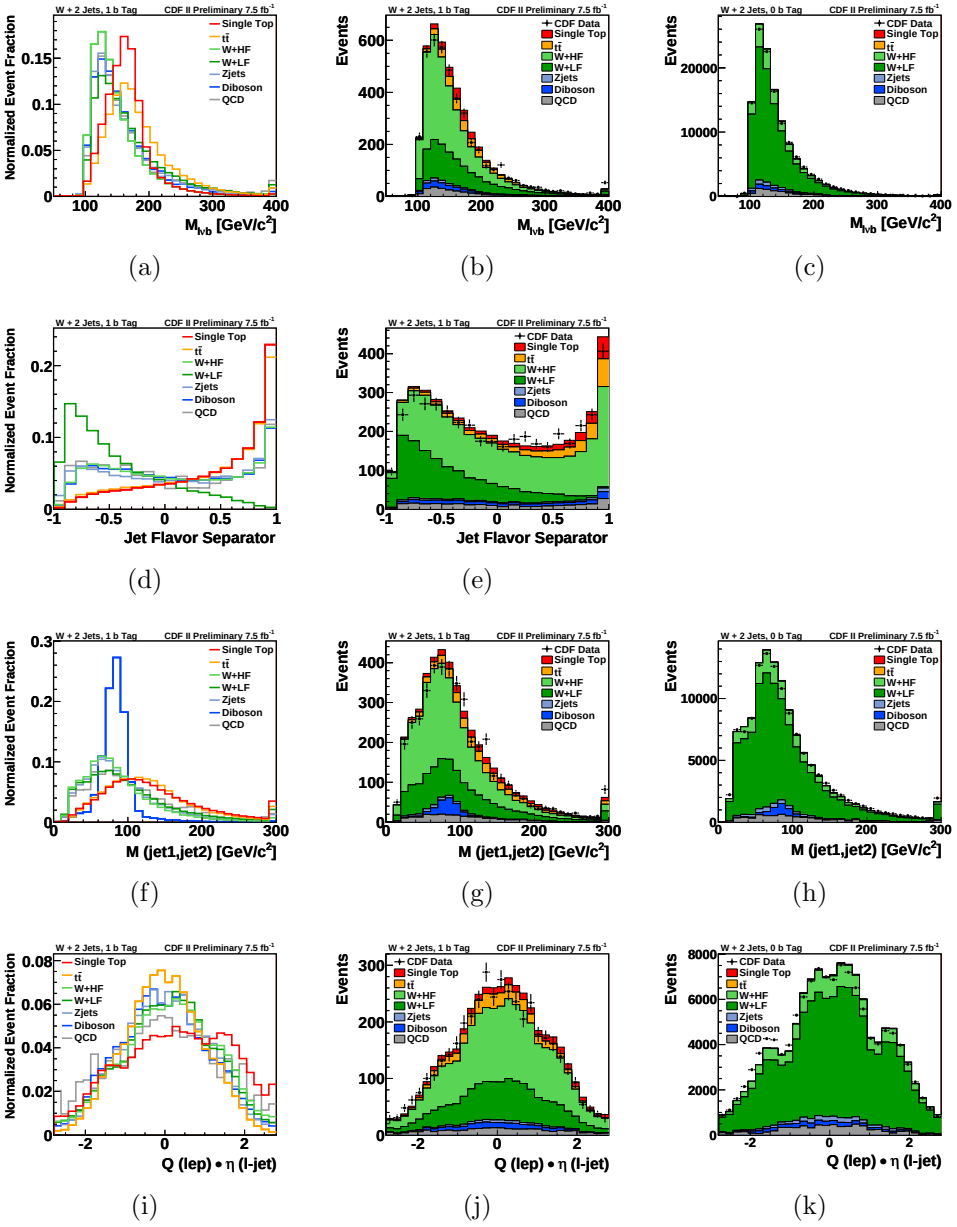


Figure 7.4: Shape comparison (first column) and MC modeling validation (second column) in the 2J1T signal region, and MC modeling validation (third column) in the 2J0T sideband region of the NN discriminating input variables for TLC events. Distributions of (a)–(c) $M^{\ell\nu b}$, (d)–(e) jet flavor separator, (f)–(h) M_{jj}, and (i)–(k) $Q \times \eta$ are shown.

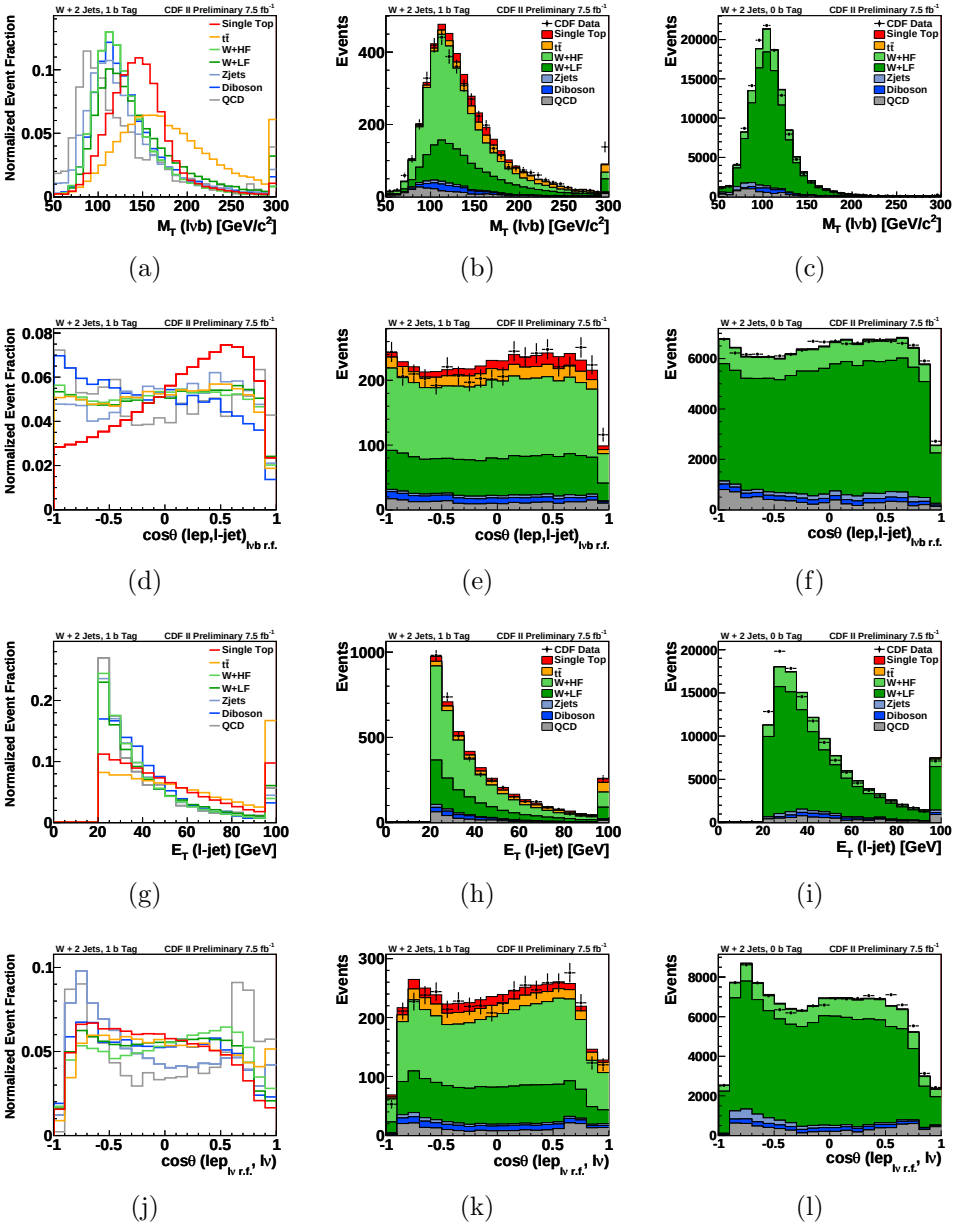


Figure 7.5: Shape comparison (first column) and MC modeling validation (second column) in the 2J1T signal region, and MC modeling validation (third column) in the 2J0T sideband region of the NN discriminating input variables for TLC events. Distributions of (a)–(c) $M_T^{\ell\nu b}$, (d)–(f) $\cos\theta^t_{\ell j}$, (g)–(i) E_T^{light}, and (j)–(l) $\cos\theta^W_{\ell W}$ are shown.

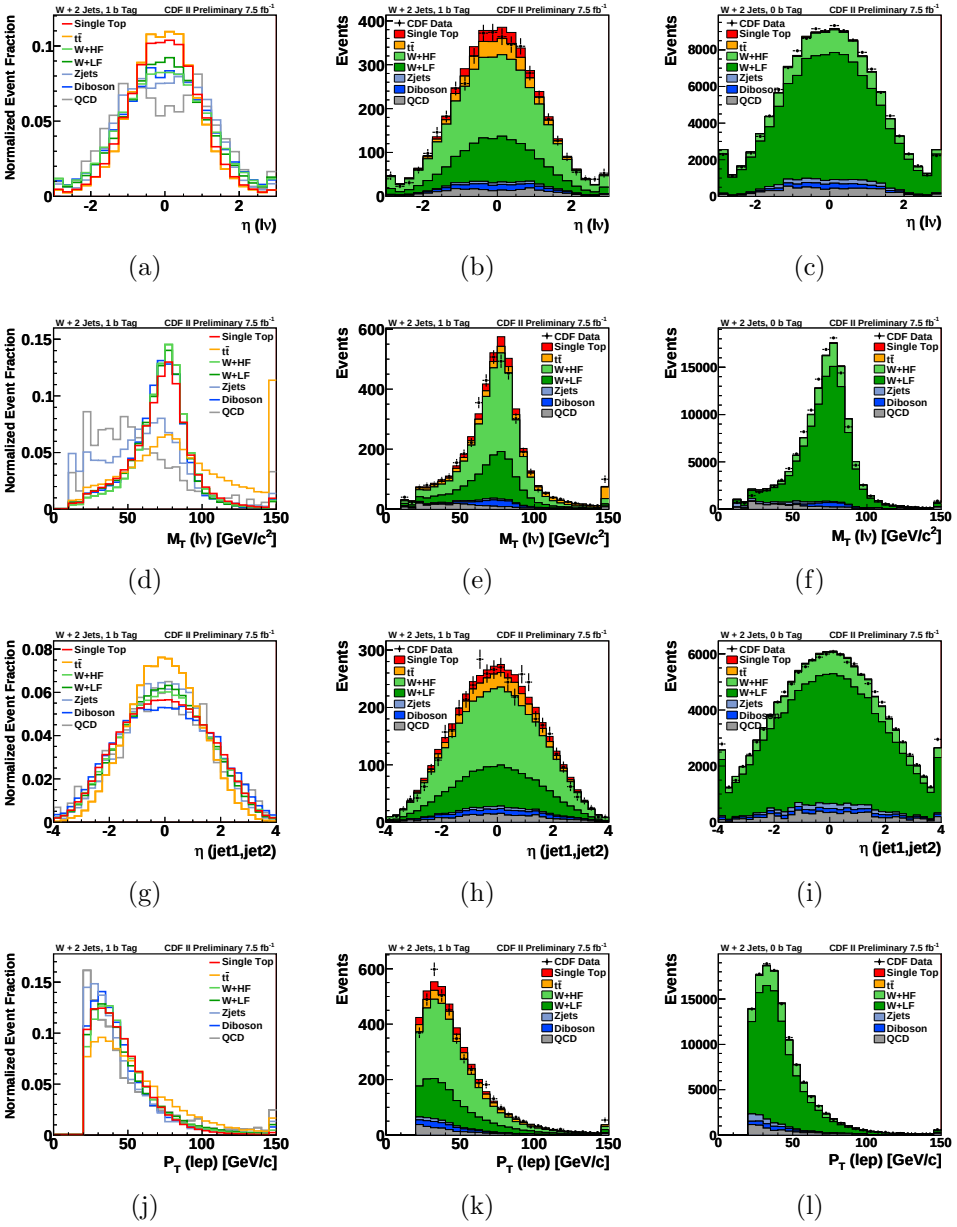


Figure 7.6: Shape comparison (first column) and MC modeling validation (second column) in the 2J1T signal region, and MC modeling validation (third column) in the 2J0T sideband region of the NN discriminating input variables for TLC events. Distributions of (a)–(c) η_W, (d)–(f) M_T^W, (g)–(i) $\sum \eta_j$, and (j)–(l) p_T^ℓ are shown.

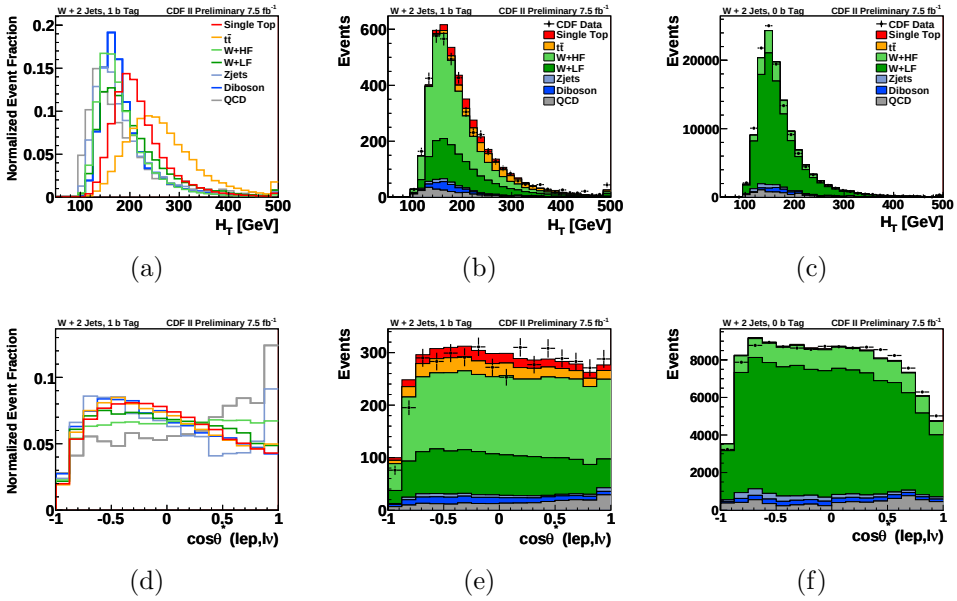
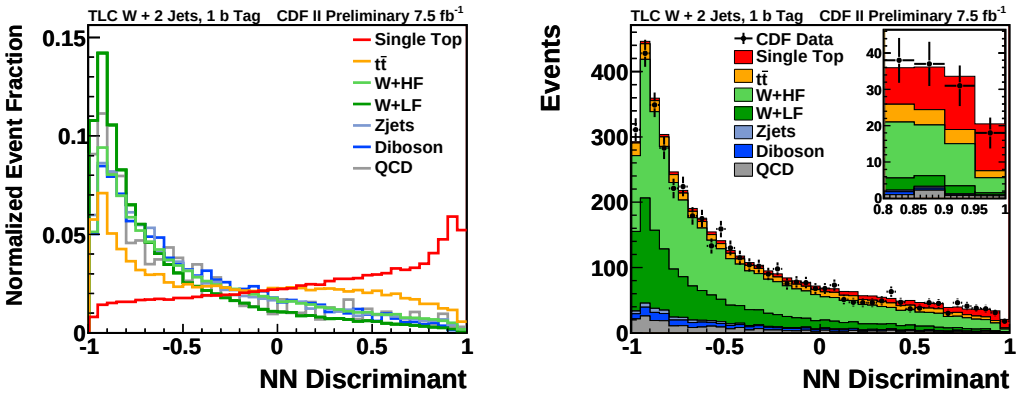


Figure 7.7: Shape comparison (first column) and MC modeling validation (second column) in the 2J1T signal region, and MC modeling validation (third column) in the 2J0T sideband region of the NN discriminating input variables for TLC events. Distributions of (a)–(c) H_{T} and (d)–(f) $\cos\theta^{W}_{\ell W}$ are shown.

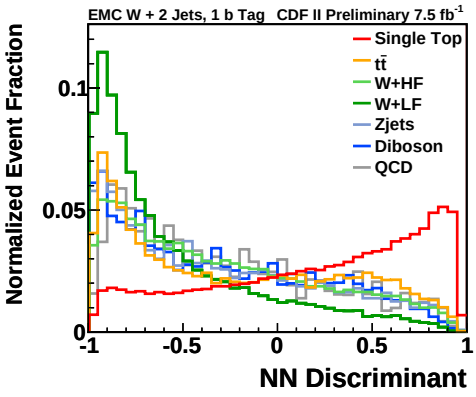
7.3 Training Results and Template Construction

In each data sample — distinguished by the number of tight jets, the number of b tags, and the type of lepton — a NN is trained with the input variables described in Section 7.2.2. The output discriminant is continuously distributed between –1.0 and +1.0, where –1.0 indicates background-like events and +1.0 indicates signal-like events. Templates are constructed with each NN for the various samples of the signal and background processes. Figures 7.8–7.11 shows the NN templates constructed (both the shape comparison and the stacking plots) in the corresponding signal regions.

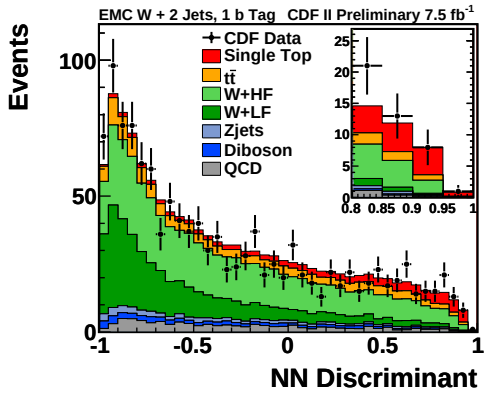


(a)

(b)



(c)



(d)

Figure 7.8: (a) (c) Shape comparison and (b) (d) stacking distribution of the NN discriminant in the 2J1T signal region for TLC and EMC events.

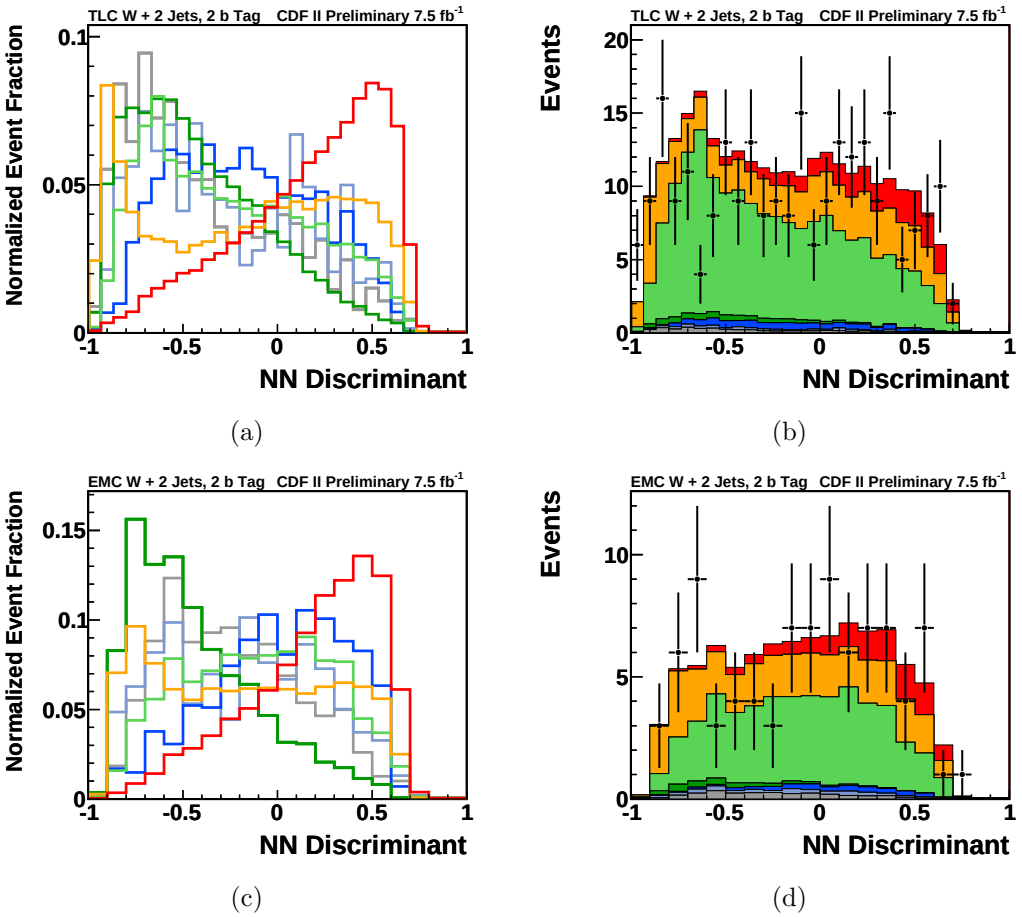


Figure 7.9: (a) (c) Shape comparison and (b) (d) stacking distribution of the NN discriminant in the 2J2T signal region for TLC and EMC events.

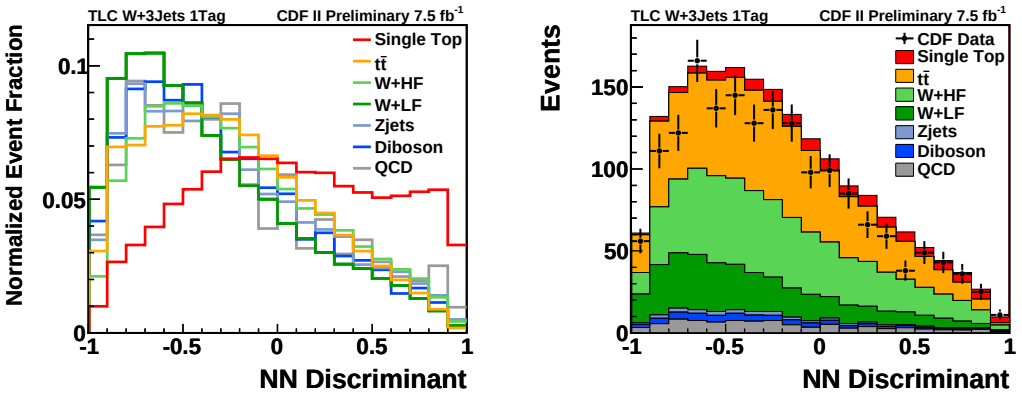


Figure 7.10: (a) (c) Shape comparison and (b) (d) stacking distribution of the NN discriminant in the 3J1T signal region for TLC and EMC events.

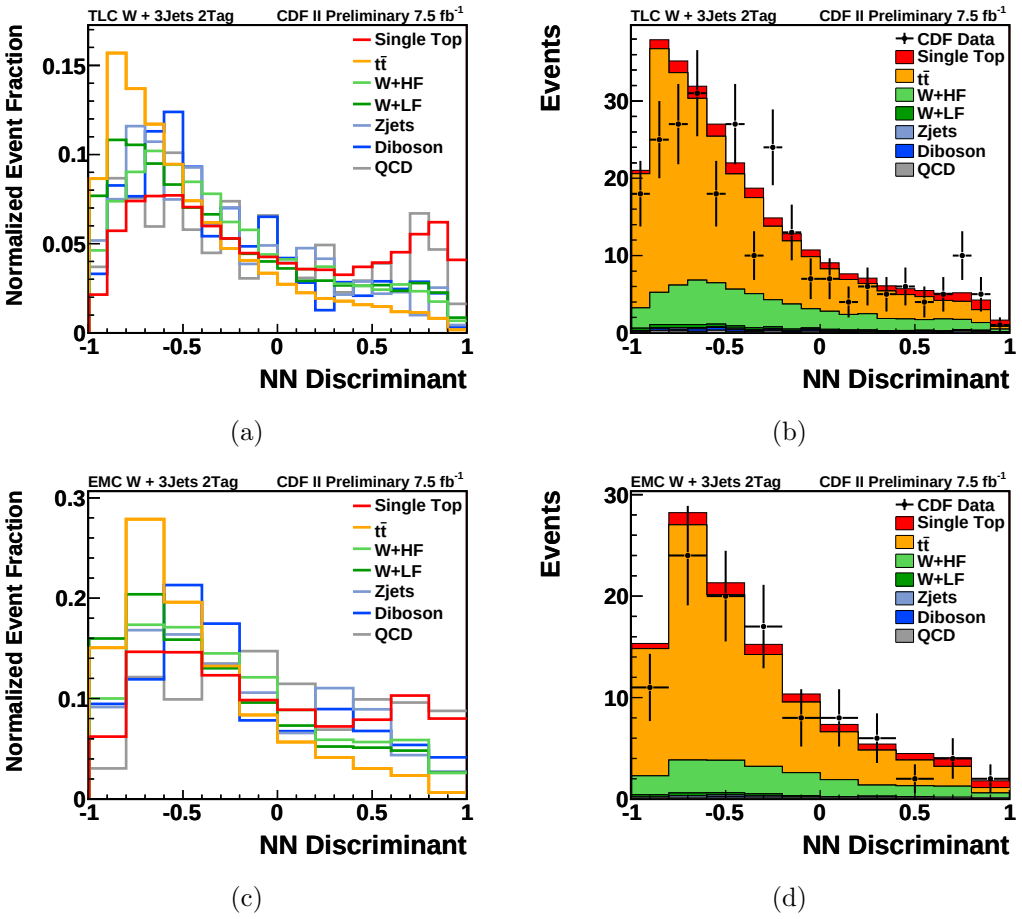


Figure 7.11: (a) (c) Shape comparison and (b) (d) stacking distribution of the NN discriminant in the 3J2T signal region for TLC and EMC events.

CHAPTER EIGHT

Systematic Uncertainties

The measurement of the single top quark cross section heavily relies on inputs from theoretical models, Monte Carlo simulations, and extrapolations from control samples in the data. The systematic uncertainties associated with each input create an uncertainty in the measurement. We include the effects of these uncertainties on the measurement of the single top cross section by varying the rate of predicted signal and background events and the shape of the discriminant histograms. In total, three categories of uncertainties are considered in this analysis: the uncertainty in the predicted rates of the signal and background processes (rate uncertainty), the uncertainty in the shape of the discriminant templates (shape uncertainty), and the uncertainty due to a limited number of events in each bin of the discriminant templates (statistical uncertainty). Both rate and shape uncertainties are considered as systematics uncertainties, which will be described in this chapter.

Sources of uncertainty may affect multiple signal and background processes. Uncertainties arising from the same source are considered to be fully correlated while uncertainties from different sources are considered to be uncorrelated. The sources of systematics uncertainties are listed in Table 8.1.

8.1 Rate Uncertainties

Rate uncertainties affect the expected production yields (number of events) for both signal and background processes. The sources of rate uncertainties in this analysis are described below. Some of the uncertainties also affect the shape of the discriminant templates.

Table 8.1: Sources of systematic uncertainties considered in this analysis. Some rate uncertainties are listed as a range because their impact differs for each of the categories indexed by the numbers of jets and b tags. The symbol X indicates that an uncertainty is applied as a shape variation. The last two sources of uncertainty are used only in the $|V_{tb}|$ measurement.

Source of Uncertainty	Rate	Shape	Processes affected
Luminosity	6%		all
Acceptance and efficiency	1–7%		all
Monte Carlo generator	3–7%		single top, $t\bar{t}$
Initial- and final-state radiation	0–6%	X	single top, $t\bar{t}$
Parton distribution functions	0–1%	X	single top, $t\bar{t}$
Jet energy scale	0–8%	X	all
Jet flavor separator		X	all
Mistag model		X	W + LF
Non-W model		X	Non-W
Factorization and renormalization		X	$Wb\bar{b}$
Jet η and ΔR distribution		X	W + LF
Non-W normalization	40%		Non-W
$Wb\bar{b}$ and $Wc\bar{c}$ normalization	30%		$Wb\bar{b}$, $Wc\bar{c}$
Wc normalization	30%		Wc
Mistag normalization	10–20%		W + LF
$t\bar{t}$ normalization	8%		$t\bar{t}$
Single top normalization	7%		single top
Top mass	2–12%	X	single top, $t\bar{t}$

8.1.1 Integrated Luminosity

The MC-based background predictions rely on the input luminosity, which is assigned a systematic uncertainty of ±6%. This uncertainty includes the systematic error of the luminosity measurement, which is dominated by the uncertainty on the inelastic $p\bar{p}$ cross section and the acceptance uncertainty of the CLC counter [146]. It also includes the acceptance uncertainty that arises from requiring the primary vertex position in z to be within ±60 cm around the origin.

8.1.2 Acceptance and Efficiency Scale Factors

Uncertainties on the lepton identification scale factor, the trigger efficiency, and the b-tagging scale factor are included for each sample that uses these factors (see Section 6.2).

8.1.3 Theoretical Cross Section

We normalized those MC-based background processes to their NLO (or higher-order) theoretical predictions. The associated theoretical uncertainties are applied in the analysis (see Table 6.1). The effects of the top quark mass on the theoretical predictions for the $t\bar{t}$ and single top processes are separated out. For the cross section measurement, we omit these effects along with the theoretical uncertainties on the single top quark cross section; this is because the single top quark cross section is the quantity being measured. For the extraction of the $|V_{tb}|$ value, these uncertainties are applied.

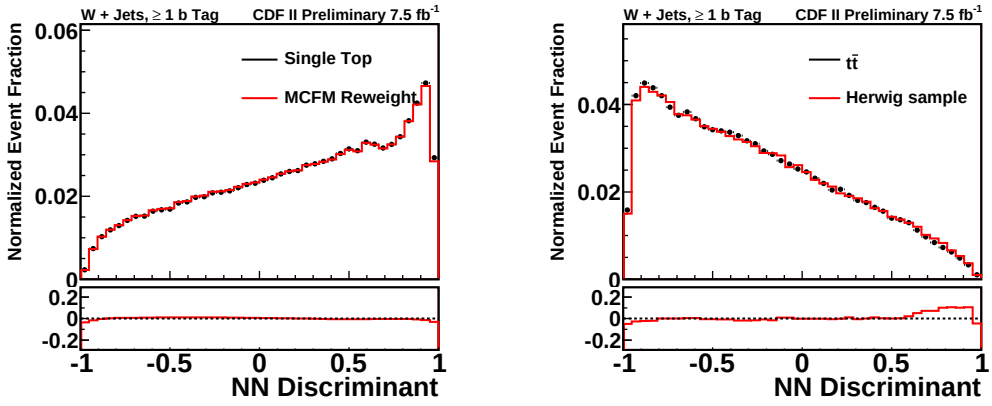
8.1.4 Monte Carlo Generator

Using a different MC generator for the signal or background simulation may result in a different measurement. We compare the kinematic distributions of the primary partons obtained from the s- and t-channel POWHEG samples with the theoretical differential cross section calculated by MCFM. The discrepancies between the MC simulation and the theoretical calculation are quantified by assigning weights to the simulated events.

For the t-channel, we derive the weights from comparisons of three consecutive two-dimensional kinematic distributions: p_T versus η of the top quark, p_T versus η of the separator b jet, and p_T versus η of the light jet. The comparisons are done with a MC sample without any selection, and a weight is calculated for each event. The rate uncertainty is taken as the acceptance difference between the selected samples

and the selected re-weighted samples. An alternative shape template is constructed with the weights to take into account the shape difference in the signal region. For the s-channel, a similar method is applied with three comparisons: p_T versus η of the separator b jet, p_T versus η of the top quark, and p_T versus η of the b jet from the top quark decay. For the Wt-channel, since we do not have the MCFM prediction of its kinematic distributions, the rate difference alone is calculated; it is obtained from the difference in the acceptance between the DR and DS methods. No shape uncertainty is applied for the Wt-channel in this analysis. The effect of the MC uncertainty for the single top process on the NN discriminant is shown in Figure 8.1(a).

For the $t\bar{t}$ process, differences in the acceptance rate and the discriminant shape (see Figure 8.1(b)) between the default PYTHIA sample and the alternative HERWIG sample are considered as the MC uncertainty.



(a) (b)

Figure 8.1: The shape comparison in the NN discriminant distributions due to the MC generator uncertainty for the (a) single top and (b) $t\bar{t}$ processes.

8.1.5 Parton Distribution Functions (PDF)

In this analysis, we use CTEQ5L for the $t\bar{t}$ PYTHIA sample and CTEQ6M for the single top POWHEG samples. The determination of PDFs is carried out by several groups, using various data from different experiments and with various values of the

strong coupling constant α_s. The PDF uncertainty is thus obtained by adding in quadrature the differences between the predictions of different PDF sets and the default PDF, which are estimated by reweighting events with weights associated with the different PDF sets.

The CTEQ61 PDF set contains a set of 20 orthogonal pairs of eigenvectors corresponding to the determined uncertainties. Each pair of eigenvectors is compared with the default PDF of the corresponding process. The differences are summed up in quadrature since the eigenvectors represent independent systematic sources. Besides the CTEQ PDFs used in this analysis, we also compare to the MRST72 [147] PDF from the MRST group. The difference between the MRST72 PDF and the default PDF is compared to the quadrature sum of the CTEQ61 eigenvectors, and the larger difference is taken as the uncertainty. In addition, the difference is evaluated between the MRST72 and MRST75 sets, which differ in their value of α_s. The final sum gives an event weight that is used to evaluate the rate and shape uncertainty due to the PDF uncertainty. Since the PDF uncertainty is very small, it is only calculated for the $t\bar{t}$ and single top processes.

8.1.6 Initial- and Final-State Radiation (IFSR)

The initial-state radiation is calculated in PYTHIA by the "backwards evolution" shower algorithm [148]. The effects of ISR are controlled by the DGLAP evolution equation, which has been studied extensively in electron/neutrino-nucleon deep inelastic scattering experiments. From the DGLAP equation, we know that ISR effects depend on the momentum scale of the interaction Q^2, the value of λ_{QCD}, splitting functions, and parton distribution functions. λ_{QCD} is a scale parameter introduced in the renormalization of perturbation theory; it is the energy scale at which perturbative quantum chromodynamics begins to break down [88]. Though the PYTHIA tunes we use include an estimation of ISR effects from the data, a sys-

tematic uncertainty is applied to account for the incomplete knowledge of the initial state radiation.

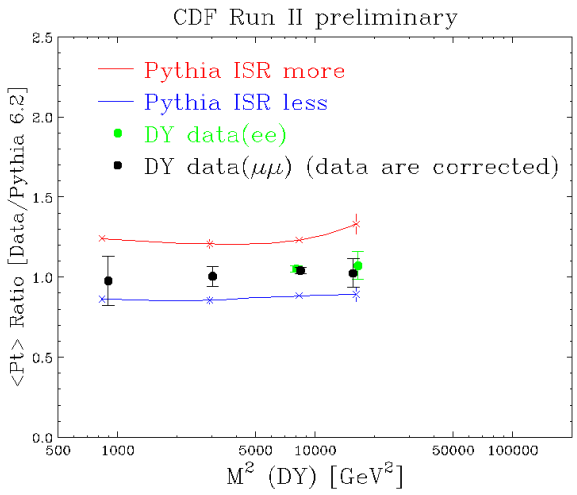


Figure 8.2: The ratio of the transverse momentum of the dilepton system in the data to the same quantity in the MC prediction, as a function of the invariant mass of the dilepton system. The systematic samples bracket the data points and their statistical uncertainty.

The ISR uncertainty was studied with about 200 pb^{-1} of collected data with the Tune AW at CDF [149]. It is evaluated by comparing Drell-Yan (DY) MC events with DY data events, where DY refers to events in which a Z/γ^* decays into leptons ($Z/\gamma^* \rightarrow e^+e^-, \mu^+\mu^-$). The DY sample has the advantage that it has no final-state radiation, which provides a unique sample to study ISR effects. Both the dilepton p_T and the number of jets are sensitive to ISR activity. Only the dilepton p_T is chosen in the study since it is not affected by the jet energy uncertainty. The dilepton p_T distribution as a function of the dilepton invariant mass is studied and the MC-simulated samples with more/less ISR are then tuned to generously bracket the data [150] as shown in Figure 8.2. The large variation is chosen to bracket the data in order to include the extrapolation uncertainty of the ISR prediction from the Z mass scale to higher scales of $t\bar{t}$ and single top quark events. The resulting ISR uncertainty is constructed by varying the parameters of PYTHIA. For more ISR and less ISR respectively, λ_{QCD} is doubled or divided in half and the initial transverse

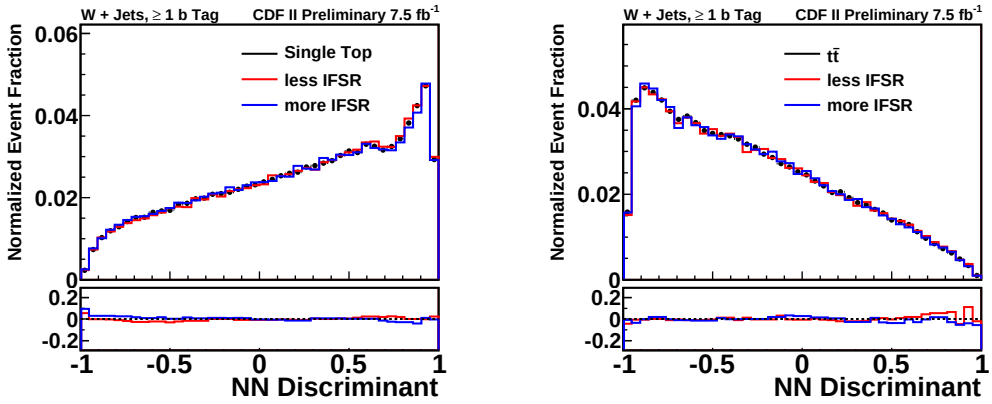
momentum scale and the hard scattering scale of the shower are multiplied by four or divided by four.

Unlike ISR, the final state radiation cannot be isolated at the Tevatron. This is because each event is initiated by a quark or gluon interaction at the Tevatron. Thus, ISR effects are impossible to remove. In PYTHIA, the model of gluon radiation from partons emitted from the hard-scattering interaction has been tuned to data collected at the Large Electron-Positron (LEP) experiment [148]. As its name indicates, LEP had leptons in the initial state, and thus it was possible to measure the FSR effect cleanly. Parameters analogous to those used for ISR are adjusted in PYTHIA for the FSR uncertainty.

The effects of variations in ISR and FSR are treated as 100% correlated, and we can combine them as the IFSR uncertainty. Because of the very conservative estimation of ISR uncertainty to bracket the Z/γ^* data, and because much more Z/γ^* data have been collected, this IFSR uncertainty is considered to be an overestimate. Since the Tune QW is very similar to the Tune AW, we use the same variation of parameters for the IFSR uncertainty of the Tune AW on the Tune QW samples to estimate the IFSR uncertainty of the Tune QW. A pair of systematic samples are generated: one with increased PYTHIA parameters for more ISR and FSR and one with decreased PYTHIA parameters for less ISR and FSR. These two samples are used to obtain $\pm 1\sigma$ differences from the central value for both rate and shape uncertainties. The ISR and FSR uncertainties are only applied to single top and $t\bar{t}$ production; their effects on the shape of the discriminant distribution are shown in Table 8.2 and Figure 8.3, respectively. For W + jets backgrounds, they are not evaluated since their rates are derived from a data-driven method and their shapes have factorization and renormalization scale uncertainties, which will be discussed later.

Table 8.2: The rate difference due to the MC, PDF, and IFSR uncertainties for s- and t-channel, combined single top production, and the $t\bar{t}$ process. The listed differences are calculated from the combination of all eight jet-tag-lepton categories.

Channel	MC (%)	PDF (%)	IFSR (%)
s-channel	3.53/3.53	0.80/–0.89	–3.06/4.46
t-channel	6.58/6.58	0.86/–0.89	–2.95/5.69
single top	4.87/4.87	0.77/–0.82	–2.94/5.05
$t\bar{t}$	–5.46/–5.46	0.65/–0.59	–4.68/–0.00



(a) (b)

Figure 8.3: The shape comparison in the NN discriminant distributions due to the IFSR for the (a) single top and (b) $t\bar{t}$ processes.

8.1.7 Jet Energy Scale (JES)

As discussed in Section 3.4, the measured raw jet energies are corrected with a multi-level correction that accounts for detector non-uniformity, multiple interactions, calorimeter non-linearity, underlying event, and fixed jet cone size. Each correction level introduces an uncertainty that is added in quadrature into the final jet energy scale as shown in Figure 8.4. All MC samples are reprocessed with the jet energy scale varied upwards and downwards as alternative systematic samples. Because the E_T of a jet changes when the JES is varied upward or downward, some events are re-categorized since the number of jets in the event will also differ. The background estimation is then re-evaluated to determine the rate uncertainty due

to JES (see Table 8.3). The kinematic properties of each event are also affected and carried to the NN discriminant templates. The JES uncertainty is the largest source of systematic uncertainty in this analysis even though the NN discriminant is trained with the JES-varied samples. The effects of the JES uncertainty on the NN discriminant shape are shown in Figure 8.5.

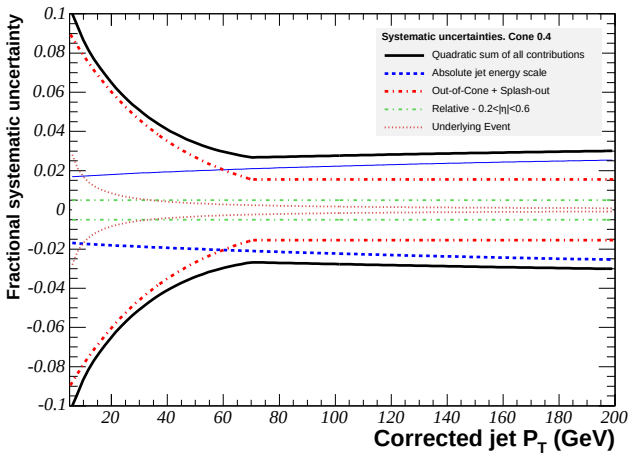


Figure 8.4: The systematic uncertainty on the jet energy scale as a function of the corrected transverse momentum of the jet.

Table 8.3: The rate difference due to the JES uncertainty in both signal and background processes. The listed differences are calculated from the combination of all eight jet-tag-lepton categories.

Channel	JES (%)
s-channel	0.14/–1.87
t-channel	0.28/–1.87
single top	0.21/–1.40
$t\bar{t}$	–0.15/7.04
W + HF	–0.54/7.68
W + LF	–0.39/0.21
Diboson	1.11/–5.26
Z + jets	2.79/–6.99

8.1.8 Heavy Flavor Fraction in W + Jets

As described in Section 6.5.1, it is found that the ALPGEN simulation underestimates the $Wb\bar{b}$ and $Wc\bar{c}$ predictions in the data. An additional K-factor is derived

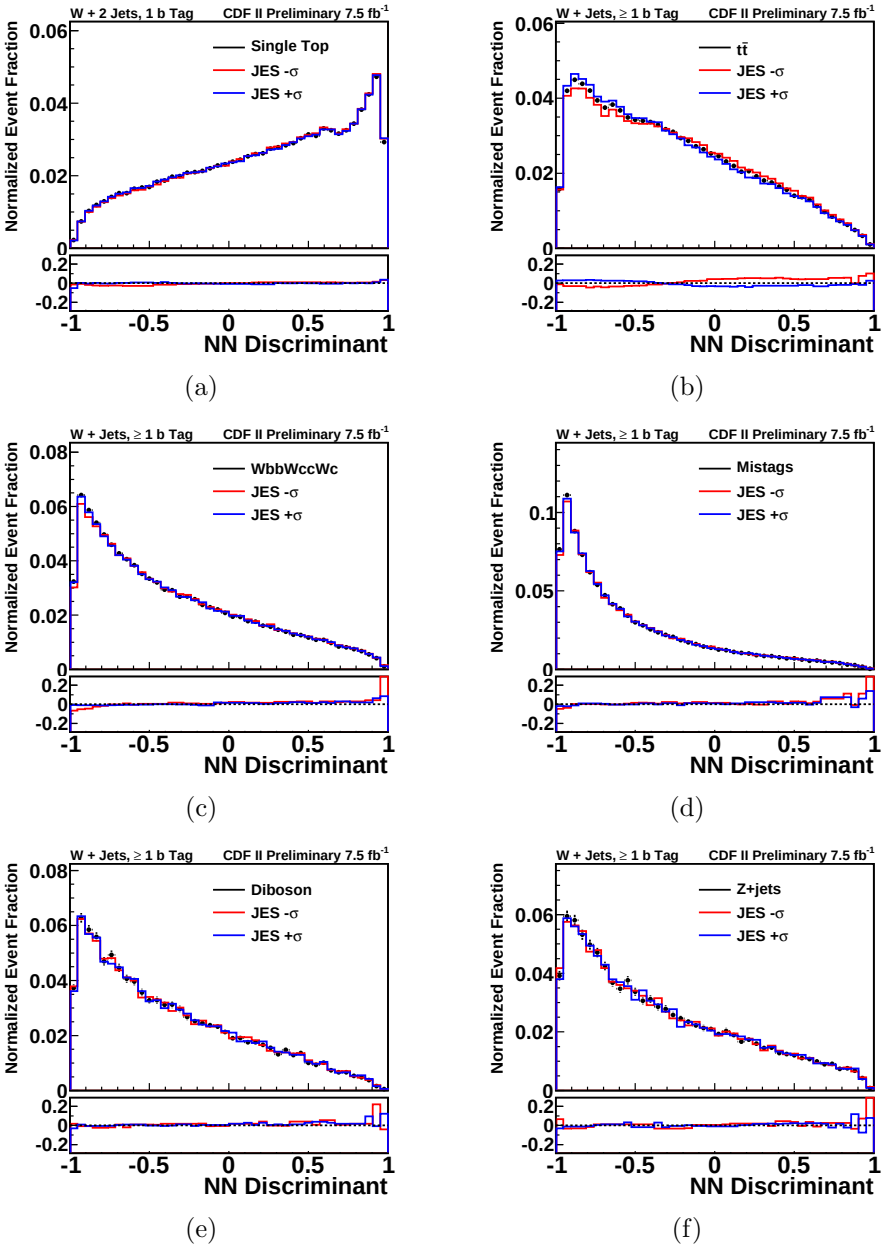


Figure 8.5: The shape comparison in the NN discriminant distributions due to the JES uncertainty for the (a) single top, (b) $t\bar{t}$, (c) W + HF, (d) W + LF, (e) diboson, and (f) Z + jets processes.

from the W + 1 jet sample and extrapolated to the W + 2 jets and W + 3 jets samples. An uncertainty of 30% is assigned to the K-factor that considers the spread of the measured heavy-flavor fraction using different variables in the fit and the extrapolation to the W + 2 and W + 3 jets samples. The Wcj fraction is obtained

from a direct measurement using lepton charge correlations. It is found that no rescaling is needed. A 30% uncertainty is assigned to the Wcj prediction — the same relative magnitude as the $Wb\bar{b}$ and $Wc\bar{c}$ uncertainty. Since these uncertainties come from different sources and are measured with different methods, they are treated as uncorrelated. The heavy flavor fraction uncertainty is one of the largest systematic uncertainties in this analysis. Its effect is further constrained with the likelihood fit of the background normalization to the data.

8.1.9 Mistag Estimate

The estimation of the number of incorrectly b-tagged events in the data is derived from the mistag probability as described in Section 6.5.2. The systematic uncertainty of the mistag event estimation includes the extrapolation from the negative tags to the positive tags in the data by estimating the asymmetry between positive and negative light-flavor tags. It also takes into account the differences in the negative tag rates of different data samples used to construct the mistag matrix. For events with two b tags, the uncertainty is doubled.

8.1.10 Non-W Multijet Estimate

The non-W prediction is estimated from the $\not\!\!E_T$ fit to data. The fit is dependent on the number of bins chosen for the $\not\!\!E_T$ distribution and on the different models used for the non-W templates. A conservative uncertainty of $\pm 40\%$ is assigned on all the non-W predictions.

8.2 Shape-Only Uncertainties

Some sources of uncertainty discussed above that affect the rate of a process will also affect the shape of the signal or background template. The sources of shape uncertainties that do not have an associated rate variation already described are enumerated below.

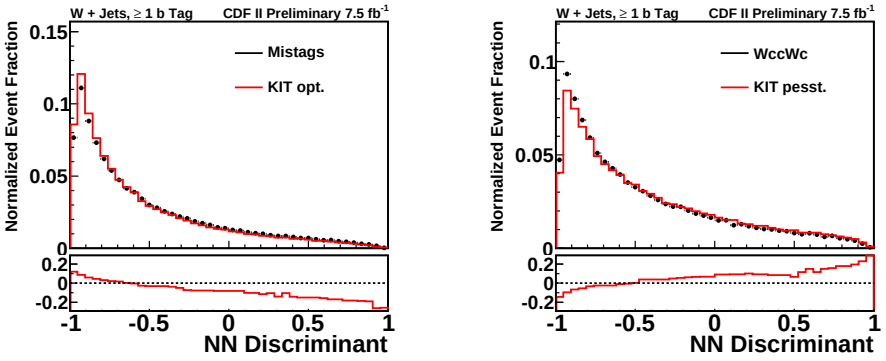
8.2.1 Jet Flavor Separator Modeling

Since we do not select events based on the output of the jet flavor separator, only a shape uncertainty on the NN distribution is associated with this separator. As described in Section 7.2.1, it was found that the b jet properties are well described by the MC simulation when compared with the data, while light-flavor jets require an additional correction. Nevertheless, it is impossible to test the shape of the jet flavor separator distribution of c jets since we lack a pure sample of charm jets in the data. As a consequence, two scenarios are considered for the systematic uncertainty on the jet flavor separator. They are called the "optimistic" and "pessimistic" scenarios based on their impact on the b-enriched signal region [141, 151].

For the optimistic scenario, the uncorrected light-flavor Monte Carlo prediction is taken as a one-sided systematic uncertainty. In this case, light-flavor jets are further separated from b jets. In the pessimistic scenario, the correction function of the light flavor jet is applied for the c-jets, conservatively assuming that the c jets are mis-modeled as badly as the light jets to cover any effects from the lack of understanding of the charm jet data. The resulting jet flavor distribution of c-jets is shifted to the more signal-like region while the distribution of the simulated light jets is corrected as usual. The shifts in the jet flavor distribution of both scenarios are propagated through to the NN discriminant distributions of the corresponding processes as shown in Figure 8.6.

8.2.2 Mistag Model

The predicted rate of mistag events is obtained from the mistag matrix of the data, while the mistag NN shape is modeled by W + LF ALPGEN samples. Since an understanding of the mistag events in the data remains inconclusive, we assign a systematic uncertainty for the shape of the mistag events by constructing an alternative mistag template using 0-tagged data weighted by the mistag probability.



(a) (b)

Figure 8.6: The shape comparison in the NN discriminant distributions due to the jet flavor uncertainty for the (a) optimistic scenario of the W + LF process and the (b) pessimistic scenario of the $Wc\bar{c}$ and Wc processes.

Because the 0-tagged data contain mostly W + LF events, this uncertainty should cover the "real" shape of the W + LF process.

8.2.3 Factorization and Renormalization Scales

The generation of the W + jets samples in ALPGEN requires the factorization and renormalization scale as inputs. The factorization and renormalization scale Q is set to be

$$\sqrt{M_W^2 + \sum_{\text{partons}} m_T^2}, \tag{8.1}$$

where $m_T = \sqrt{m^2 + p_T^2/c^2}$ is the transverse mass of the generated parton. The masses of the light partons (u, d, s, and g) are set to zero; m_b is set to 4.7 GeV/c^2 and m_c is set to 1.5 GeV/c^2. The sum is over all final-state partons excluding the W boson decay products. In addition, the scale of the strong coupling constant α_s is set to the transverse momentum of the vertex in ALPGEN. Since these scale values are artificial inputs, an uncertainty is assigned to cover a variety of different choices. The scales are doubled and halved to create samples that cover the scale uncertainty. The NN discriminant templates of these scale-varied samples are constructed. Their comparison with the central templates is shown in Figure 8.7.

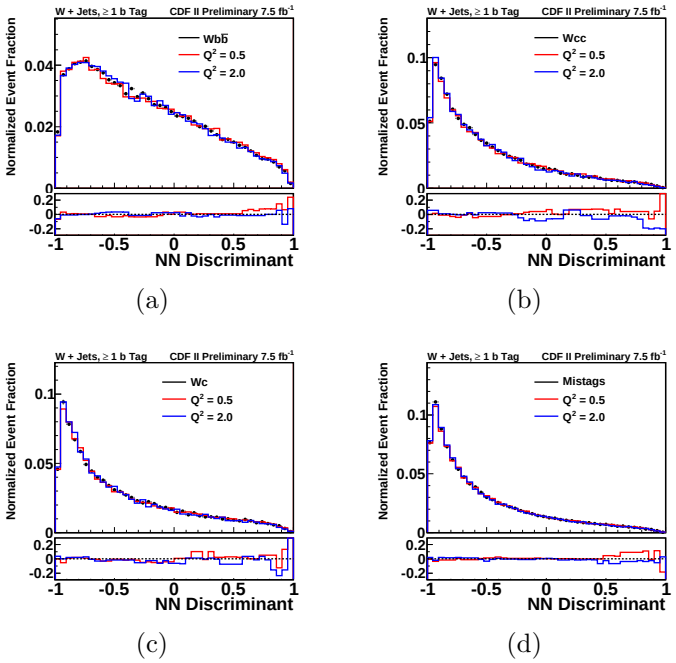


Figure 8.7: The shape comparison in the NN discriminant distributions due to the factorization and renormalization scale variations for the (a) $Wb\bar{b}$, (b) $Wc\bar{c}$, (c) Wc, and (d) W + LF processes.

8.2.4 Non-W Flavor Composition

As described in detail in Section 7.2.1, the flavor composition of jets in non-W events is estimated by fitting different jet flavor templates in the low-$\not\!\!E_T$ control samples. An uncertainty on the non-W flavor composition is needed due to the extrapolation to the high-$\not\!\!E_T$ signal region and due to the limited statistical precision of the fits. While the central prediction for the flavor composition is 45% b jets, 40% c jets, and 15% light-flavor jets, an alternative flavor composition of 60% b jets, 30% c jets, and 10% light-flavor jets is used to construct the alternative non-W templates. The alternative flavor composition is considered the "worse-case" variation since the resulting non-W sample is more signal-like. This uncertainty has no effect on the predicted event yields, but the variation on the jet flavor composition is propagated to the NN discriminant template of the non-W models.

8.2.5 *Mismodeling of Jet Angular Distribution*

From the validation checks in the 0-tagged sideband region, mismodelings were found in the jet E_T and jet η distributions and the $\Delta R_{j_1 j_2}$ distribution. Relative to the data, the rate of jets with low E_T and high $|\eta|$ are underestimated by the prediction (see Figure 8.8(a)). The mismodeling of the jets in the high $|\eta|$ region has a significant impact on the t-channel signal, since the variable $Q \times \eta$ is one of the most significant discriminating variables for jets at large η. Inaccurate modeling in that region might create a false signal in this analysis. The modeling of the $\Delta R_{j_1 j_2}$ distribution, where $\Delta R_{j_1 j_2} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the angular separation between two jets in the $\eta - \phi$ plane, is also of considerable importance. The discriminant variable M_{jj}, which is useful to separate the top quark process from background processes, is highly correlated with $\Delta R_{j_1 j_2}$. Therefore, the mismodeling of the $\Delta R_{j_1 j_2}$ distribution (see Figure 8.8(b)) could potentially affect the NN discriminant.

The mismodeling of the jet angular distributions mentioned above was discovered in the previous CDF single top analysis [37]. Various explanations had been proposed and intensive studies of this problem had been carried out. With more collected data in the 0-tagged sideband data, the mismodeling is enhanced in the η and E_T distributions of the first- and second-most energetic jets as well as in the $\Delta\rho$ and ΔR of these two jets. We believe the mismodeling is mainly due to W + LF events from ALPGEN. Thus we assign a systematic uncertainty by reweighting the W + LF samples to 0-tagged data using three consecutive reweighting functions: a 2D reweighting of η^{jet1} versus η^{jet2}, a 2D reweighting of E_t^{jet1} versus E_t^{jet2}, and a 1D reweighting function of the $\Delta\phi(j_1, j_2)$ distribution. Since $\Delta R(j_1, j_2)$ is highly correlated to these variables, this reweighting procedure also covers the discrepancy in that distribution. Figure 8.8(c) and 8.8(d) show the distribution after reweighting the W + LF samples. No rate uncertainty is applied and the alternative template from the reweighted W + LF sample is taken as a one-sided systematic uncertainty.

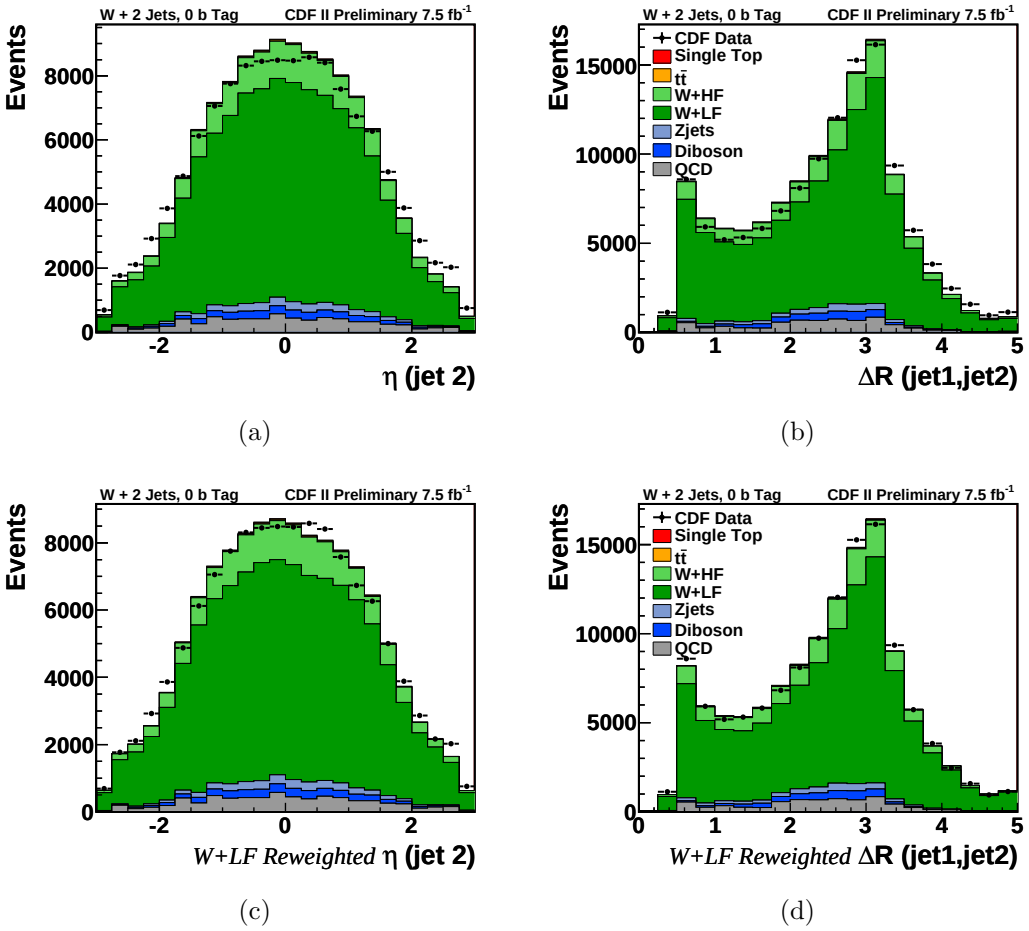


Figure 8.8: The (a) η_{j2} and (b) $\Delta R_{j_1 j_2}$ distributions in the zero-tag region, respectively. The (c) η_{j2} and (d) $\Delta R_{j_1 j_2}$ distributions in the zero-tag region after the W + LF samples are reweighted to data.

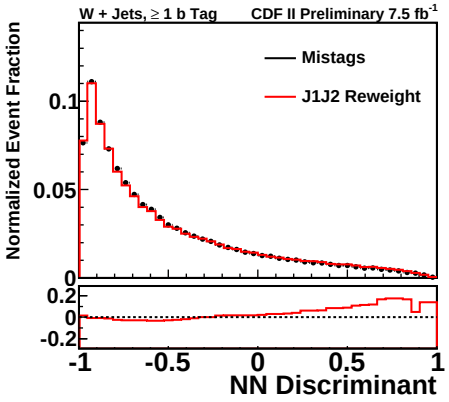


Figure 8.9: The shape comparison in the NN discriminant distribution of the W + LF sample before and after reweighting to the zero-tag sideband data.

CHAPTER NINE

Measurement of the Single Top Cross Section

To measure the single top quark cross section, a correct interpretation of the NN discriminant templates with all uncertainties is crucial. We use CDF's MCLimit package to calculate the cross section using a Bayesian approach. The MCLimit package can be obtained from Ref. [152] and is described in Ref. [153–155]. We will briefly describe our implementation of this package in the following section.

9.1 Statistical Method

9.1.1 Bayesian Approach

Consider a sample set S that has two subsets A and B. We can define the conditional probability $P(A|B)$ (the probability P of A, given B) as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{9.1}$$

Based on the fact that $A \cap B$ and $B \cap A$ are the same, we can obtain *Bayes' theorem*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{9.2}$$

In *frequentist statistics*, the probability $P(A)$ is interpreted as the limited frequency of occurrence of A. In *Bayesian statistics*, the probability $P(A)$ is a subjective probability, which can be interpreted as the credibility level that the hypothesis A is true. In the case of high energy physics, the "theory" represents a hypothesis A while B is the outcome of the experiments from the "data". Bayes' theorem becomes

$$P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory})P(\text{theory}). \tag{9.3}$$

$P(\text{theory})$ is the *prior* probability for the theory, reflecting our confidence of the theory based on a priori knowledge. $P(\text{data}|\text{theory})$ is the probability of having the

data actually obtained given the theory; this is called the *likelihood*. Since we always consider the data to be real, $P(\text{data})$ ($P(B)$ in Eq. 9.2) is a constant, serving as a normalization term. Equation 9.3 tells how the probability for the theory is modified with the given data; this is called the *posterior* probability $P(\text{theory}|\text{data})$.

9.1.2 Likelihood Function

From the NNs trained in Chapter 7, we construct template histograms for each physics process (signal and background). This is done for each of the orthogonal categories, which are indexed by the number of jets, the number of b tags, and whether the charged lepton is a TLC or an EMC lepton. In particle physics, the number of events in the data is normally described by a Poisson distribution. Thus, the likelihood function is the product of Poisson probabilities for each bin in each histogram of the neural network distribution for each category. The Poisson probabilities are functions of the number of observed data events in each bin d_i and the predictions of each physics process in each bin μ_i, where i ranges from 1 to the number of total bins n_{bins}. The likelihood function is given by

$$L = \prod_{i=1}^{n_{\text{bins}}} \frac{\mu_i^{d_i} e^{-\mu_i}}{d_i!}. \tag{9.4}$$

The expected number of events in each bin includes both the signal and background contributions:

$$\mu_i = \sum_{k=1}^{n_{\text{bkg}}} b_{ik} + \sum_{k=1}^{n_{\text{sig}}} s_{ik} \tag{9.5}$$

where b_{ik} is the predicted number of events in bin i for background process k; n_{bkg} is the number of background contributions considered in this analysis. The s_{ik} are the predicted signal yields in each bin, and n_{sig} is the number of signal processes. In this analysis, we consider the sum of three single top processes as the signal contribution: the s-, t-, and Wt-channel.

Each source of systematic uncertainty described in Chapter 8 varies the b_{ik} and s_{ik} predictions for each systematic source. Hence, the b_{ik} and s_{ik} predictions de-

pend on a set of the uncertainty nuisance parameters θ_m, where $m = 1...n_{\text{nuis}}$, one for each independent source of systematic uncertainty. n_{nuis} is the number of systematic uncertainties considered for the corresponding signal or background process. Each nuisance parameter is assigned a Gaussian prior that is centered on zero and has unit width. Therefore, zero represents the central prediction and ± 1 represents the $\pm 1\sigma$ variation of the systematic uncertainty. The list of nuisance parameters is shared among the signal and background processes. Systematic variations that come from the same source of uncertainty are considered to be correlated. In addition to the systematic uncertainties (both shape and rate), bin-by-bin statistical uncertainties are also taken into account. Bin-by-bin uncertainties come from the limited number of MC events (or data events from a control sample) used to predict the corresponding signal or background in each bin. The bin-by-bin uncertainties are taken to be independent of (uncorrelated to) each other and the systematic uncertainties.

The procedure for applying uncertainties to the signal and background predictions in each bin is performed step by step. First the shape uncertainties are applied, then the bin-by-bin uncertainties, and finally the rate uncertainties. We use $t_{ik}^{0^{th}}$ to represent the central (unvaried) prediction in bin i of the template histogram k (either signal $s_{ik}^{0^{th}}$ or background $b_{ik}^{0^{th}}$). The superscript index labels the step of the procedure (i.e., the zeroth step for the case of $t^{0^{th}}$).

For the shape uncertainty, in order to remove statistical fluctuations in the systematically shifted histogram, a median smoothing process is applied to each histogram. In every five consecutive bins, the ratios between the varied prediction and the central prediction are calculated and the median of these ratios is found. Then, each bin is recalculated by multiplying the central prediction by the median ratio. This efficiently filters out extreme outliers and high frequency noise and results in a smoother shape-varied histogram. The contribution to the prediction of each bin from a given shape uncertainty is linearly added to the central prediction. This

contribution is obtained from the difference between the central prediction $t_{ik}^{0^{th}}$ and the prediction corresponding to a $+1\sigma$ variation $\kappa_{t,ik}^{m+}$ (-1σ variation $\kappa_{t,ik}^{m-}$) if $\theta_m > 0$ ($\theta_m < 0$):

$$t_{ik}^{1^{st}} = t_{ik}^{0^{th}} + \sum_{m=1}^{n_{\text{nuis}}} \begin{cases} (\kappa_{t,ik}^{m+} - t_{ik}^{0^{th}})\theta_m & : \quad \theta_m \geq 0 \\ (t_{ik}^{0^{th}} - \kappa_{t,ik}^{m-})\theta_m & : \quad \theta_m < 0 \end{cases} . \tag{9.6}$$

It is non-physical to have a negative prediction for a physics process. This procedure should not produce a negative prediction in any bin for any source of background or signal. The second step of the procedure therefore produces a prediction of

$$t_{ik}^{2^{nd}} = \max(0, t_{ik}^{1^{st}}). \tag{9.7}$$

The bin-by-bin uncertainties are next linearly interpolated. The number of events in each bin follows a Poisson distribution, thus the bin-by-bin uncertainty on $t_{ik}^{0^{th}}$ is $\delta_{t,ik}^{0^{th}}$, which is approximated as $\sqrt{t_{ik}^{0^{th}}}$ since we do not know the true mean of that bin. Systematically varied histograms also have a statistical uncertainty in each bin. The bin-by-bin uncertainty on the bin content of a systematically varied histogram $t_{ik}^{m\pm}$ is $\delta_{t,ik}^{m\pm}$. Because some of the shape-varied templates are filled with independent systematic samples, directly summing up $\delta_{t,ik}^{0^{th}}$ and $\delta_{b,ik}^{m\pm}$ could overestimate the bin-by-bin uncertainties. Instead, the bin-by-bin uncertainty is calculated as

$$\delta_{t,ik}^{1^{st}} = \delta_{t,ik}^{0^{th}} + \sum_{m=1}^{n_{\text{nuis}}} \begin{cases} (\delta_{t,ik}^{m+} - \delta_{t,ik}^{0^{th}})\theta_m & : \quad \theta_m \geq 0 \\ (\delta_{t,ik}^{0^{th}} - \delta_{t,ik}^{m-})\theta_m & : \quad \theta_m < 0 \end{cases} . \tag{9.8}$$

A nuisance parameter $\eta_{t,ik}$ is given for the bin-by-bin uncertainty such that $t_{ik}^{3^{rd}}$ is given by

$$t_{ik}^{3^{rd}} = t_{ik}^{2^{nd}} + \delta_{t,ik}^{1^{st}} \eta_{t,ik}, \tag{9.9}$$

where $\eta_{t,ik}$ is drawn from a Gaussian centered on zero with unit width.

Finally, rate uncertainties are applied. The systematic uncertainties on the predicted rates are often parameterized by a symmetric, multiplicative scale factor,

such as the luminosity and acceptance uncertainties. The fractional uncertainty on $t_{ik}^{0^{th}}$ due to a nuisance parameter m is given by $\rho_{t,ik}^{m+}$ for a $+1\sigma$ variation and given by $\rho_{t,ik}^{m-}$ for a -1σ variation. The bin content including rate uncertainties has the form

$$t_{ik}^{\text{varied}} = t_{ik}^{\text{central}} \prod_{m=1}^{n_{\text{nuis}}} \left(1 + \rho_{b,ik}^{m\pm}\theta_m\right) \tag{9.10}$$

However, some uncertainties (like that for the jet energy scale) can have an asymmetric impact on one process, while having a symmetric impact on another process. From Ref. [156], the above method is indeed wrong for asymmetric uncertainties. Model 2 of [156] is used to parameterize the asymmetric uncertainties with a quadratic function on a smoothly varying nuisance parameter. With the shape and bin-by-bin uncertainties calculated, the final bin prediction of $t_{ik}^{0^{th}}$ is then given by

$$t_{ik}^{4^{th}} = t_{ik}^{3^{rd}} \prod_{m=1}^{n_{\text{nuis}}} \left(1 + \frac{\rho_{t,ik}^{m+} + \rho_{t,ik}^{m-}}{2}\theta_m^2 + \frac{\rho_{t,ik}^{m+} - \rho_{t,ik}^{m-}}{2}\theta_m\right). \tag{9.11}$$

For a symmetric impact, $\rho_{t,ik}^{m+}$ equals $\rho_{t,ik}^{m-}$; Eq. 9.11 falls back to Eq. 9.10.

For some uncertainties, the systematic variation beyond $\pm1\sigma$ leads to a non-physical prediction in a bin. This is the case for those uncertainties obtained from reweighting a MC prediction to match the data or a higher-order prediction. An example is the jet angular mismodeled systematic uncertainty and the MC systematic uncertainty. Thus we apply an additional Heaviside step function to the Gaussian prior of that particular nuisance parameter, which is called a "truncated" Gaussian prior.

The above chain of application of nuisance parameters is applied to both the signal and background predictions. For the signal predictions from the fourth step, a scale factor β_k is applied on the Standard Model cross section prediction:

$$s_{ik} = s_{ik}^{4^{th}}\beta_k. \tag{9.12}$$

The index k for the signal runs over the three signal processes. Then, β_s scales the s-channel signal, β_t scales the t-channel signal, and β_{Wt} scales the Wt-channel signal.

Therefore, the likelihood is a function of the observed data $\mathbf{D} = \{d_i\}$, the signal scale factors $\boldsymbol{\beta} = \{\beta_s, \beta_t, \beta_{wt}\}$, the nuisance parameters $\boldsymbol{\theta} = \{\theta_m\}$ and $\boldsymbol{\eta} = \{\eta_{s,ik}, \eta_{b,ik}\}$, the central values of the signal and background predictions $\boldsymbol{s} = \{s^0_{ik}\}$ and $\boldsymbol{b} = \{b^0_{ik}\}$, and the rate, shape, and bin-by-bin uncertainties $\boldsymbol{\rho} = \{\rho^{m\pm}_{b,ik}, \rho^{m\pm}_{s,ik}\}$, $\boldsymbol{\kappa} = \{\kappa^{m\pm}_{b,ik}, \kappa^{m\pm}_{s,ik}\}$, $\boldsymbol{\delta} = \{\delta^0_{b,ik}, \delta^{m\pm}_{b,ik}, \delta^0_{s,ik}, \delta^{m\pm}_{s,ik}\}$:

$$L = L(\mathbf{D}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\rho}, \boldsymbol{\kappa}, \boldsymbol{\delta}). \tag{9.13}$$

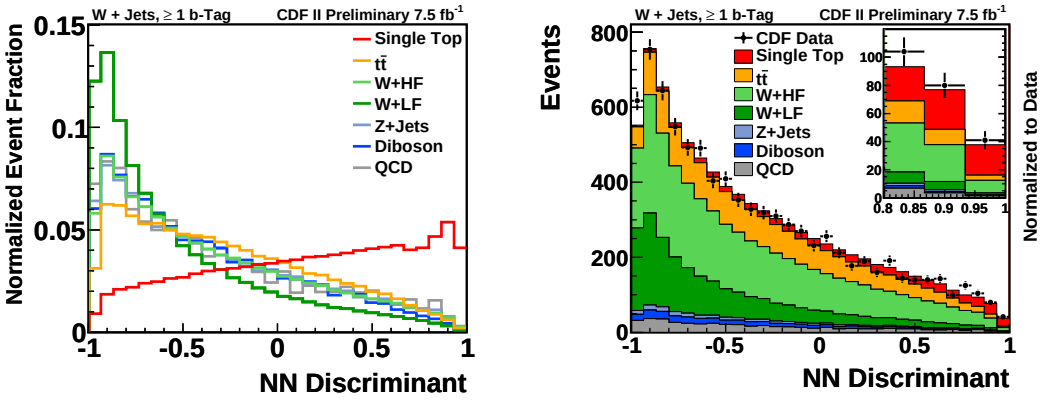
9.2 Cross Section Measurement

9.2.1 Profile and Marginal Likelihood

For the measurement of the cross section, we aim to determine the maximum of the reduced likelihood function with respect to the signal scale $\boldsymbol{\beta}$ from Eq. 9.13. This can be achieved by two approaches: profile and marginal likelihood. The method of profiling the likelihood is to determine all the nuisance parameters that maximize the likelihood for each fixed $\boldsymbol{\beta}$. The method of marginalization is to integrate all the nuisance parameters for the posterior. The cross section can be extracted from the maximum of the posterior. The profiling procedure is done by fitting the nuisance parameters to the data for the minimum of the negative log-likelihood function using the MINUIT package [157]. It is limited in its ability to correctly estimate the uncertainty [45]. Instead, marginalization is used in order to incorporate the effects of systematic uncertainties.

In this analysis, we employed both methods in the cross section measurement. First, the profile likelihood is used to improve the sensitivity. With a priori estimations of the background rate and shape with large uncertainties, the marginalization procedure has weak discrimination power for the single top signal. Fitting for

the background, using correlated fits with several histograms of differently selected data, allows us to constrain the background better than our a priori estimations. Figure 9.1(b) shows the combined NN discriminant distribution by fitting to the data.



(a) (b)

Figure 9.1: The (a) shape comparison and the (b) stacking plot of the combined NN discriminant distribution.

After fitting the NN discriminant distribution to the data in each category, the marginalization is performed with integration over all the nuisance parameters. The number of dimensions is proportional to the number of uncertainties, physics processes, categories, and bins. This results in integrations in a high number of dimensions. With a standard Monte Carlo method, this procedure is very CPU time consuming. A markov chain monte carlo method [2] is exploited in the MCLimit package to speed up this process.

9.2.2 Measurement of Cross Section

Assuming the SM ratio between the s-, t-, and Wt-channel production rates: $\beta_s = \beta_t = \beta_{wt} \equiv \beta$, the total cross section for single top quark production σ_{st} can be measured from the likelihood function in Eq. 9.13. We use a Bayesian marginalization technique [2] to incorporate the effects of systematic uncertainties as described in Ref. [36] and obtain the posterior density function $L'(\beta)$. This is achieved by

integrating out the nuisance parameters:

$$L'(\beta) = \int L(\mathbf{D}|\beta,\ \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\rho}, \boldsymbol{\kappa}, \boldsymbol{\delta})\ \pi(\boldsymbol{\theta})\ \pi(\boldsymbol{\eta})\ d\boldsymbol{\theta}\ d\boldsymbol{\eta}, \tag{9.14}$$

where the π functions are the Bayesian priors assigned to each nuisance parameter, which are Gaussian functions centered on zero with unit width.

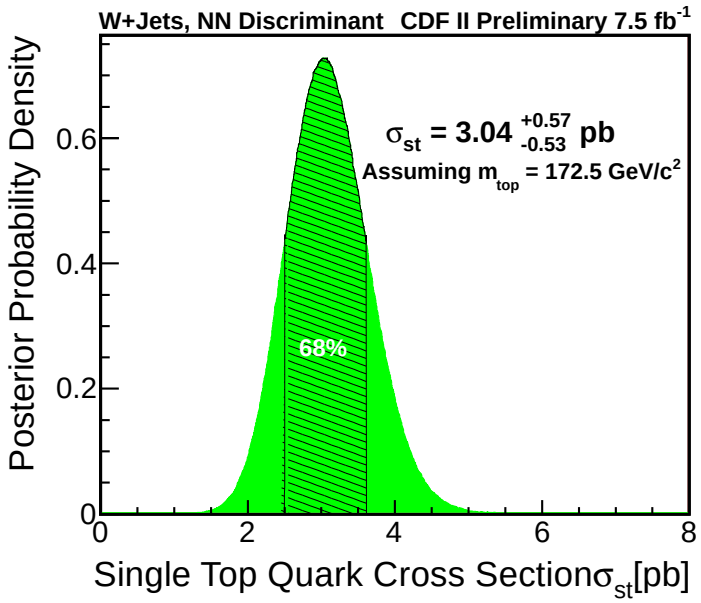


Figure 9.2. The posterior probability density for the cross section measurement.

The obtained posterior distribution is shown in Figure 9.2. The measured cross section $\sigma_{st}^{\text{measured}}$ is obtained from the value β that maximizes the posterior function $L'(\beta^{\max})$:

$$\sigma_{st}^{\text{measured}} = \sigma_{st}^{\text{SM}}\beta^{\max}. \tag{9.15}$$

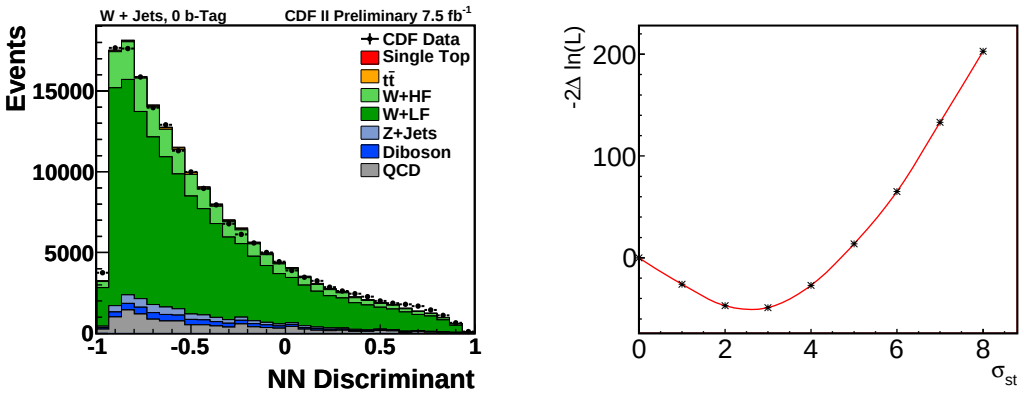
The uncertainty in $\beta^{\max}$ corresponds to the shortest interval $[\beta_{\text{low}}, \beta_{\text{high}}]$ containing 68% of the integral of the posterior:

$$0.68 = \frac{\int_{\beta_{\text{low}}}^{\beta_{\text{high}}} L'(\beta)\pi(\beta)d\beta}{\int_0^{\infty} L'(\beta)\pi(\beta)d\beta}. \tag{9.16}$$

The $\pi(\beta)$ is the prior on β, assumed to be positive and uniform. From the posterior distribution, we obtain a single top quark cross section measurement of $\sigma_{st} = 3.04^{+0.57}_{-0.53}$ pb, assuming a top quark mass of 172.5 GeV/c^2.

9.2.3 *Validation of the Measurement*

To validate the NN discriminant trained for single top quark production, the same NNs are applied on the zero-tag sideband data. This allows us to check for mismodeling effects of the NN discriminant. However, the output value of the jet flavor separator is not available in zero-tag data. The choice of a replacement value of the jet flavor separator turned out to be marginal for the validation. We set the jet flavor separator value to 0 for all the jets in the event. As shown in Figure 9.3(a), the combined NN output distribution shows good agreement between the expected background processes and the observed zero-tag data.



(a) (b)

Figure 9.3: (a) The expected and observed output of the NN applied in zero-tag sideband data. (b) The likelihood fit for the measured cross section in the data.

The posterior function is evaluated by the marginalization method and the cross section is extracted from the maximum of the posterior. This extracted cross section should be consistent with the result from the profile likelihood method, despite differences in their treatment of the uncertainty. Figure 9.3(b) shows that the minimum of the negative log-likelihood lies around the cross section extracted from the marginalization methods.

9.2.4 *Extraction of* $|V_{tb}|$

To extract $|V_{tb}|$, we take advantage of the fact that the production cross section σ_{st} is directly proportional to $|V_{tb}|^2$. Under the assumption that the top quark decays to Wb 100% of the time ($|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$) and that new physics contributions only affect $|V_{tb}|$, the parameter $\beta = \sigma_{st}^{\text{measured}}/\sigma_{st}^{\text{SM}}$ can be identified as $|V_{tb}|^2$ in the SM. The theoretical uncertainty on the single top and $t\bar{t}$ cross sections is included to obtain another posterior using the exact same method as the cross section posterior. We extract $|V_{tb}| = 0.96 \pm 0.09$ (stat. + syst.) ± 0.05 (theory) from the resulting posterior using the relation

$$|V_{tb}|^2_{\text{measured}} = \frac{\sigma_{st}^{\text{measured}}}{\sigma_{st}^{\text{SM}}} |V_{tb}|^2_{\text{SM}}, \tag{9.17}$$

where $|V_{tb}|^2_{\text{SM}} \approx 1$ and $\sigma_{st}^{\text{SM}} = 3.37 \pm 0.34$ [19].

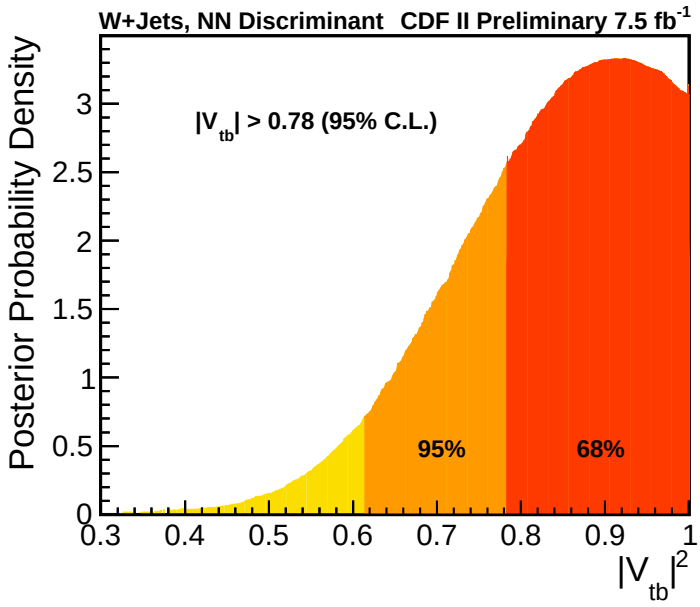


Figure 9.4. The posterior probability density as a function of $|V_{tb}|^2$.

The 95% credibility level limit on $|V_{tb}|$ is drawn from the same posterior by requiring $0 \leq |V_{tb}| \leq 1$. The point of $|V_{tb}|$ at which 95% of the likelihood curve lies

to the right of the point is identified as shown in Figure 9.4. From that point, we obtain a 95% credibility level lower limit of $|V_{tb}| > 0.78$.

9.3 Two-Dimensional Fit Result

The measurement of the total single top cross section σ_{st} was obtained by constructing a one-dimensional Bayesian posterior. We can extend this calculation by forming the posterior in a two-dimensional (2D) plane of one signal cross section versus another signal cross section. In this way, we can extract the cross section in each dimension separately. As before, we assume a uniform prior for each signal cross section to be measured and integrate over the nuisance parameters in the exact same way as the one-dimensional measurement. Since we consider three signals (s-, t-, and Wt-channels), two of the three signals are grouped together in the 2D fit.

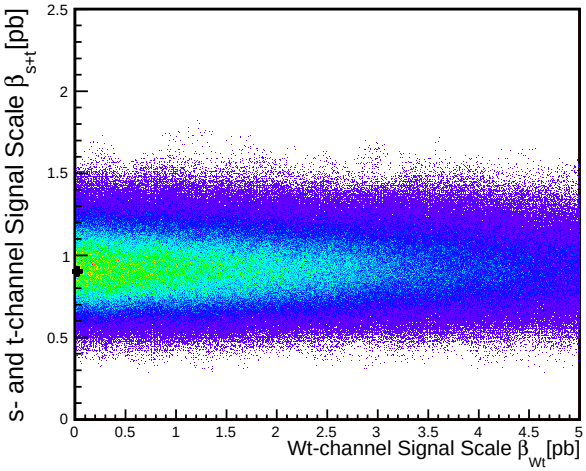


Figure 9.5: Two-dimensional fit result for σ_{s+t} versus σ_{Wt}. The black point is the best fit value. The colored region is the distribution of the 2D posterior, with green indicating a region where the posterior probability is higher and blue indicating where it is lower.

We first fit the Wt-channel cross section against the sum of the s- and t-channel cross sections. As shown in Figure 9.5, the Wt-channel contribution in the data is almost zero, with a large uncertainty. The combined s- and t-channel contribution is around 0.9 in the 2D fit. We further study the Wt-channel contribution in a

s-channel versus t-channel 2D fit by treating the Wt-channel as background, scaling the Wt-channel signal to zero, and completely ignoring the Wt-channel. In each scenario, the 2D fit result is almost identical (with less than 0.1% variation in the best fit point).

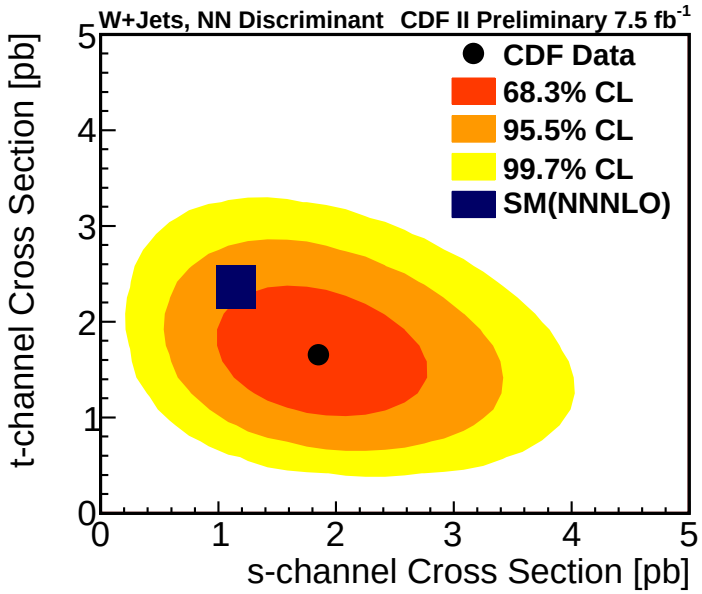


Figure 9.6: The two-dimensional fit result for σ_s versus σ_t. The black point is the best fit value. The 68.3%, 95.5%, and 99.7% credibility regions are shown as the shaded areas. The SM prediction from Ref. [19] is drawn with its theoretical uncertainty.

Here we combined the Wt-channel with the t-channel due to its similar final-state signature with the t-channel. The resulting 2D fit of the s- and t-channels is shown in Figure 9.6. The best-fit cross section values for σ_s and σ_t are the ones for which the posterior is maximized; this corresponds to $\sigma_s = 1.81^{+0.63}_{-0.58}$ pb and $\sigma_t = 1.49^{+0.47}_{-0.42}$ pb. The uncertainties on the measurements of σ_s and σ_t are correlated with each other because the s-channel and t-channel signals both populate the signal-like bins of our NN discriminant distributions. Regions of 68.3%, 95.5%, and 99.7% credibility are derived from the distribution of the posterior by integrating the contour around the best-fit point. We compare these with the NNNLO theoretical predictions of $\sigma_{t+Wt} = 2.32 \pm 0.27$ pb and $\sigma_s = 1.05 \pm 0.07$ pb [19], which

corresponds to the rectangular point in Figure 9.6. This 2D fit result is not in good agreement with the SM prediction; the difference is about one standard deviation of significance. Comparing with the result from the previous analyses [37] (about a 2σ deviation), we can see that the best-fit point moves toward the SM prediction.

CHAPTER TEN

Conclusions

We have presented a measurement of the single top quark cross section in the lepton plus jets final state. This measurement uses $p\bar{p}$ collision data corresponding to an integrated luminosity of 7.5 fb^{-1} collected by the Collider Detector at Fermilab. A neural network multivariate method is exploited to discriminate the single top quark signal from the comparatively large backgrounds. We measure a single top production cross section of $3.04^{+0.57}_{-0.53}$ (stat. + syst.) pb assuming $m_{\rm top} = 172.5$ GeV/c^2. In addition, we extract the CKM matrix element value $|V_{tb}| = 0.96 \pm 0.09$ (stat. + syst.) ± 0.05 (theory) and set a lower limit of $|V_{tb}| > 0.78$ at the 95% credibility level.

This analysis is based on a previous analysis described in Ref. [37] which used a neural network technique. A new Monte Carlo generator POWHEG was newly implemented at CDF to model the single top quark signal at NLO accuracy. The Wt-channel of single top production was included in this single top analysis for the very first time at the Tevatron. The selection acceptance was increased by adding a new lepton category ISOTRK. We further developed the neural network technique in this analysis. Even though many new input variables were found to be highly correlated to the variables used in the previous analysis and were omitted in the final version of the NN training, we successfully implemented the neural network training with systematically-mixed samples to constrain the uncertainty of the measured cross section.

With the shutdown of the Tevatron and the successful performance of the LHC, it is expected that an abundance of new analysis results from the LHC will supersede measurements from the Tevatron. However, the result of this analysis

is still important as it validates the SM at the 1.96 TeV energy scale. Another promising outcome of this single top analysis is the search for s-channel production, which is work in progress. Due to the small production rate of s-channel single top production at the LHC, the s-channel single top quark search will be a long standing legacy for the Tevatron.

BIBLIOGRAPHY

[1] Excerpt from http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[2] K. Nakamura *et al.* (Particle Data Group), "Review of particle physics," J. Phys. G: Nucl. Part. Phys. **37**, 075021 (2010).

[3] S. L. Glashow, "Partial Symmetries of Weak Interactions," Nucl. Phys. **22**, 579 (1961).

[4] A. Salam and J. C. Ward, "Electromagnetic and weak interactions," Phys. Lett. **13**, 168 (1964).

[5] S. Weinberg, "A Model of Leptons," Phys. Rev. Lett. **19**, 1264 (1967).

[6] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, "The standard model of particle physics," Rev. Mod. Phys. **71**, S96 (1999).

[7] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," Phys. Rev. Lett. **13**, 321 (1964).

[8] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," Phys. Rev. Lett. **13**, 508 (1964).

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," Phys. Rev. Lett. **13**, 585 (1964).

[10] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons," Phys. Rev. **145**, 1156 (1966).

[11] Excerpt from http://en.wikipedia.org/wiki/File:Elementary_particle_interactions.svg.

[12] S. Chatrchyan *et al.* (CMS Collaboration), "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **716**, 30 (2012).

[13] G. Aad *et al.* (ATLAS Collaboration), "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC," Phys. Lett. B **716**, 1 (2012).

[14] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," Phys. Rev. Lett. **10**, 531 (1963).

[15] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," Prog. Theor. Phys. **49**, 652 (1973).

[16] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, (Addison-Wesley, Reading, 1995).

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

[18] S. Abachi *et al.* (D0 Collaboration), "Observation of the Top Quark," Phys. Rev. Lett. **74**, 2632 (1995).

[19] N. Kidonakis, "Higher-order corrections to top-antitop pair and single top quark production," arXiv:0909.0037.

[20] E. Thomson, "Combination of CDF Top Quark Pair Production Cross Section Measurements with up to 4.6 fb^{-1}," CDF Public Note 9913, 2009 (unpublished).

[21] The CDF Collaboration, "The CDF Measurement of the Top Quark Charge using the Top Decay Products in Lepton + Jet Channel," CDF Public Note 10460, 2011 (unpublished).

[22] M. Datta, D. Glenzinski, and K. Walsh, "Measurement of W Boson Polarization Fractions in Top Quark Decay to Lepton + Jets Events using a Matrix Element Analysis Technique with 8.7 fb^{-1} of Data," CDF Public Note 10855, 2012 (unpublished).

[23] The Tevatron Electroweak Working Group for the CDF and D0 Collaborations, "Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb^{-1} of data," arXiv:1107.5255.

[24] D. Amidei *et al.*, "Study of the Top Quark Production Asymmetry and Its Mass and Rapidity Dependence in the Full Run II Tevatron Dataset," CDF Public Note 10807, 2012 (unpublished).

[25] V. M. Abazov *et al.* (D0 Collaboration), "Forward-backward asymmetry in top quark-antiquark production," Phys. Rev. D **84**, 112005 (2011).

[26] http://www-cdf.fnal.gov/physics/new/top/top.html.

[27] http://www-d0.fnal.gov/Run2Physics/top/.

[28] J. Alwall *et al.*, "Is $V_{tb} \simeq 1$?," Eur. Phys. J. C **49**, 791 (2007).

[29] T. M. Tait and C.-P. Yuan, "Single top quark production as a window to physics beyond the standard model," Phys. Rev. D **63**, 014018 (2000).

[30] A. Belyaev, "Single Top Quark in the SM and Beyond," arXiv:hep-ph/0007058.

[31] S. Bar-Shalom, D. Atwood, and A. Soni, "CP violation in single top production and decay via $p\bar{p} \to t\bar{b} + X \to W^+ b\bar{b} + X$ within the MSSM: A possible application for measuring $\arg(A_t)$ at hadron colliders," Phys. Rev. D **57**, 1495 (1998).

[32] D. Atwood, S. Bar-Shalom, G. Eilam, and A. Soni, "CP Violation in Top Physics," Phys. Rept. **347**, 1 (2001).

[33] G. Mahlon and S. Parke, "Improved spin basis for angular correlation studies in single top quark production at the Fermilab Tevatron," Phys. Rev. D **55**, 7249 (1997).

[34] G. Mahlon and S. Parke, "Single Top Quark Production at the LHC: Understanding Spin," Phys. Lett. B **476**, 323 (2000).

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

[36] T. Aaltonen *et al.* (CDF Collaboration), "Observation of Electroweak Single Top-Quark Production," Phys. Rev. Lett. **103**, 092002 (2009).

[37] T. Aaltonen *et al.* (CDF Collaboration), "Observation of single top quark production and measurement of $|V_{tb}|$ with CDF," Phys. Rev. D **82**, 112005 (2010).

[38] V. Abazov *et al.* (D0 Collaboration), "Observation of Single Top-Quark Production," Phys. Rev. Lett. **103**, 092001 (2009).

[39] The Tevatron Electroweak Working Group (CDF and D0 Collaboration), "Combination of CDF and D0 Measurements of the Single Top Production Cross Section," arXiv:0908.2171.

[40] S. Budd, J.-E. Jung, T. Junk, and S.-B. Kim, "Multivariate Likelihood Function Measurement of Single-Top-Quark Production with 3.2 fb^{-1} of Data," CDF Public Note 9699, 2009 (unpublished).

[41] S. R. Budd, "Search for Single Top Quark Production with the CDF Run II Detector Using a Multivariate Likelihood Method," Ph.D. thesis, University of Illinois at Urbana-Champaign, [Fermilab Report No. FERMILAB-THESIS-2008-41, 2008].

[42] F. Canelli, B. C. Laraña, P. Dong, B. Stelzer, and R. Wallny, "Measurement of Single Top Quark Production in 3.2 fb^{-1} of CDF II Data Using a Matrix Element Technique," CDF Public Note 9711, 2009 (unpublished).

[43] P. J. Dong, "Measurement of Electroweak Single Top Quark Production in Proton-Antiproton Collisions at 1.96 TeV," Ph.D. thesis, University of California at Los Angeles, [Fermilab Report No. FERMILAB-THESIS-2008-12, 2008].

[44] D. Hirschbühl *et al.*, "Search for Electroweak Single Top-Quark Production Using Neural Networks with 3.2 fb^{-1} of CDF II Data," CDF Public Note 9716, 2009 (unpublished).

[45] J. Lück, "Observation of Electroweak Single Top-Quark Production with the CDF II Experiment," Ph.D. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-THESIS-2009-33, 2009].

[46] F. Canelli *et al.*, "Search for Single Top Quark Production using Boosted Decision Trees in 3.2 fb^{-1} of CDF Data," CDF Public Note 9701, 2009 (unpublished).

[47] A. Apresyan, D. Bortoletto, F. Margaroli, and K. Potamianos, "First Measurement of the Single Top Production Cross Section in the $\not\!\!E_T$ Plus Jets Sample," CDF Public Note 9650, 2009 (unpublished).

[48] Linac Rookie Book, http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_RB_v2_3.pdf, 2012.

[49] L. Alvarez, Phys. Rev. **70**, 799 A (1946).

[50] Booster Rookie Book, http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V4.1.pdf, 2009.

[51] Main Injector Rookie Book, http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.1.pdf, 2010.

[52] Antiproton Source Rookie Book, http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_v2.2.pdf, 2011.

[53] Recycler Rookie Book, http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_RB_v1.42.pdf, 2007.

[54] Electron Cooling Rookie Book, http://www-ecool.fnal.gov/ecool_rookie_book.html, 2007.

[55] R. Blair *et al.* (CDF Collaboration), "The CDF II Detector: Technical Design Report," [Fermilab Report No. FERMILAB-DESIGN-1996-01, 1996].

[56] http://en.wikipedia.org/wiki/Depletion_region.

[57] C. S. Hill (CDF Collaboration), "Operational experience and performance of the CDFII silicon detector," Nucl. Instrum. Methods Phys. Res., Sect. A **530**, 1 (2004).

[58] A. Sill (CDF Collaboration), "CDF Run II silicon tracking projects," Nucl. Instrum. Methods Phys. Res., Sect. A **447**, 1 (2000).

[59] A. Affolder *et al.* (CDF Collaboration), "Intermediate silicon layers detector for the CDF experiment," Nucl. Instrum. Methods Phys. Res., Sect. A **453**, 84 (2000).

[60] A. A. Affolder *et al.* (CDF Collaboration), "CDF Central Outer Tracker," Nucl. Instrum. Methods Phys. Res., Sect. A **526**, 249 (2004).

[61] D. Acosta *et al.* (CDF Collaboration), "A Time-of-Flight detector in CDF-II," Nucl. Instrum. Methods Phys. Res., Sect. A **518**, 605 (2004).

[62] L. Balka *et al.* (CDF Collaboration), "The CDF Central Electromagnetic Calorimeter," Nucl. Instrum. Methods Phys. Res., Sect. A **267**, 272 (1988).

[63] S. Bertolucci *et al.* (CDF Collaboration), "The CDF Central and Endwall Hadron Calorimeter," Nucl. Instrum. Methods Phys. Res., Sect. A **267**, 301 (1988).

[64] M. Albrow *et al.* (CDF Collaboration), "The CDF plug upgrade electromagnetic calorimeter: test beam results," Nucl. Instrum. Methods Phys. Res., Sect. A **480**, 524 (2002).

[65] C. M. Ginsburg (CDF Collaboration), "CDF Run 2 muon system," Eur. Phys. J. C **33**, s1002 (2004).

[66] G. Ascoli *et al.*, "CDF Central Muon Detector," Nucl. Instrum. Methods Phys. Res., Sect. A **268**, 33 (1988).

[67] T. Dorigo (CDF Collaboration), "The muon system upgrade for the CDF II experiment," Nucl. Instrum. Methods Phys. Res., Sect. A **461**, 560 (2001).

[68] D. Acosta *et al.* (CDF Collaboration), "The CDF Cherenkov luminosity monitor," Nucl. Instrum. Methods Phys. Res., Sect. A **461**, 540 (2001).

[69] A. Abulencia *et al.* (CDF Collaboration), "eXtremely Fast Tracker trigger upgrade at CDF," Nucl. Instrum. Methods Phys. Res., Sect. A **598**, 328 (2009).

[70] B. Ashmanskas *et al.* (CDF Collaboration), "The CDF Silicon Vertex Trigger," Nucl. Instrum. Methods Phys. Res., Sect. A **518**, 532 (2004).

[71] J. Adelman *et al.* (CDF Collaboration), "The Silicon Vertex Trigger upgrade at CDF," Nucl. Instrum. Methods Phys. Res., Sect. A **572**, 361 (2007).

[72] R. Brun and F. Rademakers, "ROOT — An object oriented data analysis framework," Nucl. Instrum. Methods Phys. Res., Sect. A **389**, 81 (1997).

[73] E. J. Thomson *et al.*, "TopNtuple: A Public Analysis Module for the Top Group for CDF Run II," CDF Internal Note 6737, 2003 (unpublished).

[74] F. Snider (CDF Collaboration), "Tracking at CDF: Algorithms and experience from Run I and Run II," Nucl. Instrum. Methods Phys. Res., Sect. A **566**, 133 (2006).

[75] C. Hays, P. Tamburello, A. Kotwal, P. Wittich, and R. Snider, "The COT Pattern Recognition Algorithm and Offline Code," CDF Internal Note 6992, 2004 (unpublished).

[76] Y. Huang, C. Hays, and A. Kotwal, "Inside-Out Tracking," CDF Internal Note 6707, 2003 (unpublished).

[77] J. Goldstein, S. Harper, B. Heinemann, G. Manca, and P. Renton, "Reconstructing the Plug Electron Energy in 5.3.3," CDF Internal Note 7687, 2005 (unpublished).

[78] K. Burkett, J. ao Guimarães Da Costa, and S. Rappoccio, "First Look At SecVtx Using Event Primary Vertex Finder PrimeVtx," CDF Internal Note 6417, 2003 (unpublished).

[79] K. R. Bland, "Search for the Standard Model Higgs Boson in the Diphoton Final State in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV Using the CDF II Detector," Ph.D. thesis, Baylor University, [Fermilab Report No. FERMILAB-THESIS-2012-11, 2012].

[80] G. C. Blazey *et al.*, "Run II Jet Physics," arXiv:hep-ex/0005012.

[81] A. Bhatti *et al.* (CDF Collaboration), "Determination of the jet energy scale at the Collider Detector at Fermilab," Nucl. Instrum. Methods Phys. Res., Sect. A **566**, 375 (2006).

[82] D. Acosta *et al.* (CDF Collaboration), "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b-tagging," Phys. Rev. D **71**, 052003 (2005).

[83] J. Freeman, "Summer 2009 SecVtx Scale Factors Calculated Using the Electron Method through Period 22," CDF Internal Note 9848, 2009 (unpublished).

[84] S. Weinzierl, "Introduction to Monte Carlo methods," arXiv:hep-ph/0006269.

[85] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," J. High Energy Phys. **0605**, 026 (2006).

[86] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," J. High Energy Phys. **0307**, 001 (2003).

[87] F. Caravaglios and M. Moretti, "An algorithm to compute Born scattering amplitudes without Feynman graphs," Phys. Lett. B **358**, 332 (1995).

[88] J. Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation," J. High Energy Phys. **0709**, 028 (2007).

[89] T. Stelzer and W. Long, "Automatic Generation of Tree Level Helicity Amplitudes," Comput. Phys. Commun. **81**, 357 (1994).

[90] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms," J. High Energy Phys. **0411**, 040 (2004).

[91] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method," J. High Energy Phys. **0711**, 070 (2007).

[92] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," J. High Energy Phys. **0206**, 029 (2002).

[93] G. Corcella *et al.*, "HERWIG 6.5 Release Note," arXiv:hep-ph/0210213.

[94] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," Phys. Rept. **97**, 31 (1983).

[95] Sjöstrand, T., "Jet fragmentation of multiparton configurations in a string framework," Nucl. Phys. B **248**, 469 (1984).

[96] R. Brun and C. Carminati, "GEANT Detector Description and Simulation Tool," CERN Program Library Writeup W5013.

[97] J. Boudreau, C. Green, P. Murat, M. Paterno, and L. Sexton-Kennedy, "The CDF Geometry System for Run II Offline Software. A Guide for Subdetector Package Writers," CDF Public Note 5368, 2000 (unpublished).

[98] S. Carron, M. Kruse, and M. Gold, "Parametric Model for the Charge Deposition of the CDF Run II Silicon Detectors," CDF Internal Note 7598, 2005 (unpublished).

[99] P. Azzi *et al.*, "Charge Deposition Model in Silicon," CDF Internal Note 5080, 1999 (unpublished).

[100] R. Veenhof, "Garfield, a drift-chamber simulation program," CERN Program Library Writeup W5050.

[101] G. Grindhammer, M. Rudowicz, and S. Peters, "The fast simulation of electromagnetic and hadronic showers," Nucl. Instrum. Methods Phys. Res., Sect. A **290**, 469 (1990).

[102] S. Y. Jun *et al.*, "GFLASH Tuning in the CDF Calorimeter for the 2003 Winter Release," CDF Internal Note 7060, 2004 (unpublished).

[103] E. Gerchtein and M. Paulini, "CDF Detector Simulation Framework and Performance," arXiv:physics/0306031.

[104] R. Placakyte, "Parton Distribution Functions," arXiv:1111.5452.

[105] S. Moch, J. Vermaseren, and A. Vogt, "The Three-Loop Splitting Functions in QCD: The Non-Singlet Case," Nucl. Phys. B **688**, 101 (2004).

[106] H. L. Lai *et al.* (CTEQ Collaboration), "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions," Eur. Phys. J. C **12**, 375 (2000).

[107] D. Stump *et al.*, "Inclusive Jet Production, Parton Distributions, and the Search for New Physics," J. High Energy Phys. **0310**, 046 (2003).

[108] T. Affolder *et al.* (CDF Collaboration), "Charged jet evolution and the underlying event in $p\bar{p}$ collisions at 1.8 TeV," Phys. Rev. D **65**, 092002 (2002).

[109] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," Eur. Phys. J. C **71**, 1547 (2011).

[110] R. Field, "Studying the Underlying Event at CDF and the LHC," arXiv:1003.4220.

[111] Z. Sullivan, "Understanding single-top-quark production and jets at hadron colliders," Phys. Rev. D **70**, 114012 (2004).

[112] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and $e + e-$ Annihilation by Perturbation Theory in Quantum Chromodynamics.," Sov. Phys. JETP **46**, 641 (1977).

[113] V. Gribov and L. Lipatov, "Deep inelastic ep scattering in perturbation theory," Sov. J. Nucl. Phys. **15**, 438 (1972).

[114] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," Nucl. Phys. B **126**, 298 (1977).

[115] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, "Next-to-Leading-Order Predictions for t-channel Single-Top Production at Hadron Colliders," Phys. Rev. Lett. **102**, 182003 (2009).

[116] R. Frederix, E. Re, and P. Torrielli, "Single-Top t-channel Hadroproduction in the Four-Flavour Scheme with POWHEG and aMC@NLO," arXiv:1207.5391.

[117] J. M. Campbell, R. K. Ellis, and C. Williams, "MCFM v6.3 A Monte Carlo for FeMtobarn Processes at Hadron Colliders Users Guide," http://mcfm.fnal.gov/mcfm.pdf.

[118] J. Butterworth, J. Couchman, B. Cox, and B. Waugh, "KtJet: A C++ implementation of the $K_\perp$ clustering algorithm," Comput. Phys. Commun. **153**, 85 (2003).

[119] A. Belyaev, E. Boos, and L. Dudko, "Single top quark at future hadron colliders: Complete signal and background study," Phys. Rev. D **59**, 075001 (1999).

[120] T. M. Tait, "tW^- mode of single top production," Phys. Rev. D **61**, 034001 (2000).

[121] J. M. Campbell and F. Tramontano, "Next-to-leading order corrections to Wt production and decay," Nucl. Phys. B **726**, 109 (2005).

[122] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber, "Single-top hadroproduction in association with a W boson," J. High Energy Phys. **0807**, 029 (2008).

[123] A. Buzatu, "Search for the Standard Model Higgs boson produced in association with a W Boson in the isolated-track charged-lepton channel using the Collider Detector at Fermilab," Ph.D. thesis, McGill University, [Fermilab Report No. FERMILAB-THESIS-2011-24, 2011].

[124] A. Buzatu, A. Warburton, N. Krumnack, and W.-M. Yao, "A Novel in situ Trigger Combination Method," arXiv:1206.4813.

[125] J. M. Slaunwhite, "Search for the Higgs Boson Produced in Association with a W Boson at CDF Run II," Ph.D. thesis, The Ohio State University, [Fermilab Report No. FERMILAB-THESIS-2009-01, 2009].

[126] A. Papaikonomou, "Search for Single Top-Quark Production via Flavor-Changing Neutral Currents with the CDF II Experiment," Ph.D. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-THESIS-2009-21, 2009].

[127] T. Peiffer, "Search for the Higgs Boson in the WH Channel with the CDF II Experiment," M.S. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-THESIS-2008-01, 2008].

[128] J. Adelman *et al.*, "Method II For You," CDF Internal Note 9185, 2008 (unpublished).

[129] W. Sakumoto, "Event $|Z_{vtx}| < 60$ cm Cut Acceptance for Run II," CDF Internal Note 8318, 2006 (unpublished).

[130] V. Boisvert, "Trigger Efficiencies for the High E_T Central Electrons in the Gen6 Data," CDF Internal Note 7939, 2005 (unpublished).

[131] G. Chiarelli and F. Sforza, "MET_PEM Trigger Efficiency up to Summer 2008," CDF Internal Note 9493, 2008 (unpublished).

[132] S. Rolli and L. Lovas, "High-P_T Muons ID, Reconstruction and Trigger Efficiencies and Scale Factors for Period 9 through Period 12 Data," CDF Internal Note 9085, 2007 (unpublished).

[133] A. Buzatu, N. Krumnack, and A. Warburton, "MET-based Trigger Turnon Parametrizations for WH Search," CDF Internal Note 9623, 2008 (unpublished).

[134] R. J. Tesarek and S. Forrester, "Optimizing CDF Code Performance," CDF Internal Note 9418, 2008 (unpublished).

[135] B. Cooper and A. Messina, "Estimation of the Background to $W^{\pm} \to e^{\pm}\nu + n$ Jet Events," CDF Public Note 7760, 2005 (unpublished).

[136] T. Aaltonen *et al.* (CDF Collaboration), "Search for Heavy Top-like Quarks Using Lepton Plus Jets Events in 1.96-TeV $p\bar{p}$ Collisions," Phys. Rev. Lett. **100**, 161803 (2008).

[137] M. Feindt and U. Kerzel, "The NeuroBayes neural network package," Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).

[138] M. Feindt, "A Neural Bayesian Estimator for Conditional Probability Densities," arXiv:physics/0402093.

[139] M. Feindt, T. Muller, S. Richter, and W. Wagner, "A Neural Network b Tagger for Single-Top Analyses," CDF Internal Note 7816, 2005 (unpublished).

[140] Excerpt from http://www-cdf.fnal.gov/physics/new/top/2007/singletop/KANNbtag/.

[141] S. Richter, "Search for Electroweak Single Top-Quark Production with the CDF II Experiment," Ph.D. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-THESIS-2007-35, 2007].

[142] T. Chwalek *et al.*, "Update of the Neural Network b Tagger for Single-Top Analyses," CDF Internal Note 8903, 2007 (unpublished).

[143] D. Acosta *et al.* (CDF Collaboration), "Search for electroweak single-top-quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **71**, 012005 (2005).

[144] T. Walter, "Search for Electroweak Single Top Quark Production with CDF in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," Ph.D. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-THESIS-2005-28, 2005].

[145] Z. Wu *et al.*, "Single Top Quark in Lepton + Jets Analysis at CDF II using 7.5 fb^{-1} Data," CDF Internal Note 10703, 2012 (unpublished).

[146] D. Acosta *et al.*, "The performance of the CDF luminosity monitor," Nucl. Instrum. Methods Phys. Res., Sect. A **494**, 57 (2002).

[147] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton distributions: a new global analysis," Eur. Phys. J. C **4**, 463 (1998).

[148] T. Sjöstrand *et al.*, "High-energy-physics event generation with PYTHIA 6.1," Comput. Phys. Commun. **135**, 238 (2001).

[149] Y.-K. Kim and U.-K. Yang, "Initial State Gluon Radiation Studies on Drell-Yan Data for Top-Pair Production in Hadron Collider," CDF Internal Note 6804, 2003 (unpublished).

[150] A. Abulencia *et al.* (CDF Collaboration), "Top quark mass measurement using the template method in the lepton + jets channel at CDF II," Phys. Rev. D **73**, 032003 (2006).

[151] M. Renz, "*b*-jet and *c*-jet Identification with Neural Networks as well as Combination of Multivariate Analyses for the Search for Single Top-Quark Production," M.S. thesis, Universität Karlsruhe, [Fermilab Report No. FERMILAB-MASTERS-2008-06, 2008].

[152] T. Junk, MClimit Software, http://www-cdf.fnal.gov/~trj/mclimit/production/mclimit.html.

[153] J. Heinrich, "Bayesian Limit Software: Multi-Channel with Correlated Backgrounds and Efficiencies," CDF Public Note 7587, 2005 (unpublished).

[154] T. Junk, "Builiding a More General χ^2," CDF Public Note 7904, 2005 (unpublished).

[155] T. Junk, "Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Unvertainties," CDF Public Note 8128, 2006 (unpublished).

[156] R. Barlow, "Asymmetric Errors," arXiv:physics/0401042.

[157] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," Comput. Phys. Commun. **10**, 343 (1975).