Measurement of the $W^{\pm} + b\overline{b}$ Cross-Section in 695 pb^{-1} of $p\overline{p}$ Collisions at CDF II

by

Mitchell P. Soderberg

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2006

Doctoral Committee:

Associate Professor David W. Gerdes, Chair
Professor Ratindranath Akhoury
Professor Daniel M. Burns Jr.
Professor Jianming Qian
Adjunct Professor Stephen Mrenna

DEDICATION

To mom and dad.

ACKNOWLEDGEMENTS

Much like a good book, the journey through graduate school brings you into contact with many characters who all help shape the story. I'd like to thank just a few of the people who have played important roles in my journey.

The Michigan CDF group is a collection of truly amazing scientists and I have been fortunate to work with them during my graduate school career. First and foremost, I owe many thanks to my advisor, Dave Gerdes, who gave me the opportunity to work in high-energy physics. Dave is a great teacher, who always seems to know the best way to proceed in an analysis, and I appreciate all the help he has given me. Dan Amidei is the exceptional leader of the CDF group, and I'm very grateful for the enthusiasm he brings to every meeting and for the countless pieces of good advice that he has given me. Stephen Miller was the incredible postdoc who helped create the analysis that I carried out, and I always could count on him for help when I got stuck. I thank the other talented faculty and postdocs who have been a part of the Michigan group during my time here, all of whom have helped me along the way: Myron Campbell, Ken Bloom, Claudio Ferretti, Monica Tecchio, Tom Wright, and Alexei Varganov. I'd like to give a shout out to my fellow students, and sometimes housemates, in the group: Kathy Copic, Clark Cully, Nate Goldschmidt, Bo Jayatilaka, Mike Kagan, Fred Niell, and Tom Schwarz. You have all been great friends to me and I thank you for making life fun, even though the work was hard, these past few years.

I met many great people while living out at Fermilab, and I'd like to thank a few of them who played a part in my work. Chris Neu, from the University of Pennsylvania, has been an invaluable contributer to this analysis and I thank him for all his useful insights and guidance. The past and present conveners of the Top/Top-Properties/Lepton+Jets groups all helped bring my analysis to its final form: Robin Erbacher, Kevin Lannon, Taka Murayama, Jason Nielsen, Anyes Taffard, and Evelyn Thomson. Eric James was a great mentor when I was working with TDC boards, and I thank him for all he taught me.

To my mom and dad, I want to thank you for your endless love and support while I've been away. To my brothers Mike and Matt, and my sister Gretchen, thanks for putting up with me during my many short trips back to St. Louis these past few years. To my Grandma Mullen, you're the best!

Finally, and most importantly, Kathy-Anne Brickman. Your love and friendship mean the world to me, and I can't thank you enough for all the support you give me.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Measurement of the $W^{\pm} + b\bar{b}$ Cross-Section in 695 pb^{-1} of $p\overline{p}$ Collisions at CDF II

by
Mitchell P. Soderberg

Chair: David Gerdes

$W^{\pm}$+ $b\bar{b}$ events contain the associated production of a $W^{\pm}$ boson, a pair of bottom quarks ($b\bar{b}$), and any number of additional partons. This process is of much importance at hadron collider experiments due to its role as a background source in searches for Standard Model Higgs boson and single top-quark production. In this thesis the results are presented for a measurement of the b-jet cross-section in $W^{\pm}$+$b\bar{b}$ events containing 1 or 2 jets in 695 pb^{-1} of $\sqrt{s}$ =1.96 TeV $p\overline{p}$ collisions at the CDF experiment. This is the first measurement of the cross-section of $W^{\pm}$+$b\bar{b}$ performed in any experiment. The cross-section is defined to be proportional to the number of b-jets from $W^{\pm}$+$b\bar{b}$ events with one or two jets, and a leptonically decaying $W^{\pm}$ with decay products passing kinematics cuts ($p_T(\ell^{\pm})$ ≥20.0 GeV, $|\eta(\ell^{\pm})|$ ≤1.1, $p_T(\nu)$ ≥25.0 GeV). The invariant mass distribution of jets identified as containing a long-lived hadron is fit with components for bottom, charm, and light-flavor to find the fraction due to true b-decays. Background b-jet sources are subtracted to isolate the contribution of $W^{\pm}$+ $b\bar{b}$ to the data. The cross-section is measured to be 0.90±0.20(stat.)±0.26(syst.)pb, which compares well with the leading order theoretical prediction of 0.74±0.18 pb.

CHAPTER 1

The Standard Model of Particle Physics and $W^{\pm}+b\bar{b}$

Last year marked the 100^{th} anniversary of Albert Einstein's *annus mirabilis*, 1905, during which he published three papers that revolutionized the world of physics and eventually changed the way those inside and outside the physics community view the universe. Two of these papers are responsible for elucidating the now commonly accepted ideas that light is composed of discrete packets of energy known as photons (which explained the photoelectric effect), and travels at 3×10^8 m/s for all inertial observers (which is part of the Special Theory of Relativity) . These were stunning revelations to physicists and non-physicists of the day, and it took many years for the implications of these ideas to be fully understood and verified experimentally. Ultimately, Einstein's ideas about light helped in the development of the theory of quantum mechanics, which explains the physics of microscopic scales in terms of the probabilistic interactions of non-relativistic particles with external potentials.

If the physics of the early part of the 20^{th} century was defined by struggles to understand the implications of quantum mechanics and Einstein's relativity, then the physics of the latter part of the century was defined by efforts to merge these two ideas into a fully relativistic theory of microscopic scales. This theory would need to explain the steady stream of new phenomena observed in a growing number of particle physics experiments.

By the end of the century particle physicists had discovered that there are heavier versions of the electron, known as the muon and the tau, all of which have their own ghostly partners known as neutrinos that are capable of passing through millions of miles of dense matter without interacting at all! Other particles known as quarks were discovered in the form of a veritable zoo of bound-states collectively known as hadrons, and the manner of their binding was the source of much theoretical and experimental intrigue. To top things off it was realized that these new particles could interact in far more interesting ways than just those of the familiar forces of gravity and electromagnetism, though it was becoming apparent that all interactions could be described on the microscopic level as being due to the exchange of force-mediating particles. New forces named Weak and Strong were discovered that were more powerful in magnitude than electromagnetism, though restricted in range to subatomic distance scales.

Perhaps most amazing is the fact that all of these discoveries are now understandable thanks to a single theory known as the Standard Model. The Standard Model is the culmination of a theoretical framework that grew out of, and sometimes ahead of, the exciting discoveries in particle physics during the 20^{th} century. This single theory provides the mathematical description of the fundamental interactions of elementary particles that weds the ideas of quantum mechanics and Einstein's relativity. The Standard Model has been incredibly successful at explaining the results of experiment, and also makes predictions that are actively being tested today.

This thesis is a study of one particular process that can occur when a highly-relativistic proton is made to collide with a similarly energetic antiproton. This process is the simultaneous production of a $W^{\pm}$ boson and a pair of bottom quarks, which will be referred to as $W^{\pm}+b\bar{b}$. By itself, the properties of $W^{\pm}+b\bar{b}$ production are an interesting test of the predictive powers of the Standard Model, since it involves both the Weak and Strong forces. In a perhaps more important sense, $W^{\pm}+b\bar{b}$ is interesting due to its final-state's similarity to that of the production of certain types of Higgs boson and also to processes where a single top quark is produced. It is for all of these reasons that this thesis has been carried out.

We begin the thesis by introducing the Standard Model in depth before moving on to the more specific case of $W^{\pm}+b\bar{b}$ production. Along the way the theoretical issues that will make this analysis challenging will be explained, while at the same time motivation will be given for why it is important.

1.1 The Standard Model

The Standard Model is the theoretical framework, compatible with both quantum mechanics and special relativity, used to describe the interactions of fundamental particles via the exchange of force mediating particles[1, 2, 3, 4, 5]. This theory successfully describes the electromagnetic, weak, and strong forces which are responsible for the vast majority of interactions (gravity is excluded) that are believed to be allowed between elementary particles.

1.1.1 Group Structure and Gauge Bosons

The Standard Model is a gauge field theory that is invariant under a set of transformations that form the group

$$G = SU(3) \times SU(2) \times U(1). \tag{1.1}$$

The particles of the Standard Model have internal invariance under transformations in each of the subgroups of G, with associated gauge bosons (spin-1 particles) for each transformation. The theory that describes the $SU(3)$ component of the Standard Model, which defines the structure of the strong interaction, is known as Quantum Chromodynamics

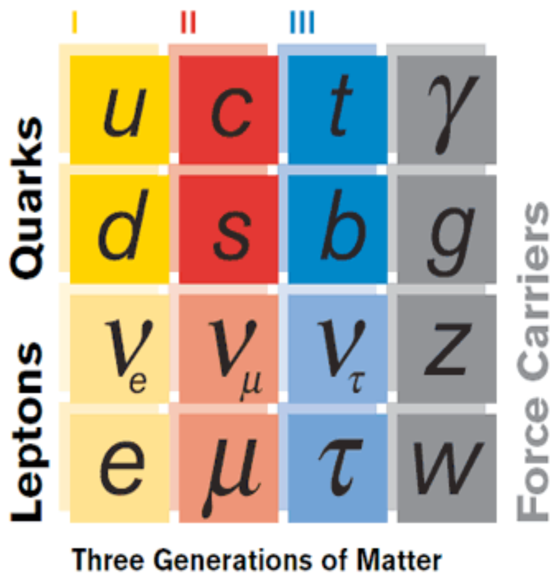


Figure 1.1: The particles that comprise the Standard Model are arranged in three generations, and interactions between them are felt by exchanging force carrying particles.

(QCD)[6, 7]. The theory that describes the combined $SU(2) \times U(1)$ component of the Standard Model, which defines the structure of the electromagnetic and weak interactions, is the Glashow-Salam-Weinberg (GSW) theory[8, 9, 10].

All particles have $U(1)$ invariance, which is related to the electromagnetic and weak interactions. The gauge boson associated with this invariance is called B^{μ}. All particles also have $SU(2)$ invariance, which is related to the weak interaction. The gauge bosons associated with the three generators of $SU(2)$ transformations are known as W_i^{μ} (i=1,2,3). Finally, all particles possess an $SU(3)$ invariance, which is related to the strong force. The gauge bosons associated with the eight generators of the $SU(3)$ transformations are labeled G_a^{μ} (a=1, 2, ..., 8). Table 1.1 summarizes the mass eigenstates of the gauge bosons responsible for each of the forces described by the Standard Model. As will be discussed later in Section 1.1.3, the gauge fields of the $U(1)$ and $SU(2)$ symmetries mix to form the physical gauge bosons listed in Table 1.1.

Boson	Mass (GeV/c^2)	Force
Photon (γ)	0	Electromagnetic
W boson ($W^{\pm}$)	80.4	Weak Charged-Current
Z boson (Z^0)	91.2	Weak Neutral-Current
Gluon (G_a)	0	Strong

Table 1.1: The gauge bosons of the Standard Model

1.1.2 Fermions

Fermions are spin-$\frac{1}{2}$ particles that are classified based on their interactions with the different forces of the Standard Model. There are three generations of fermions in the Standard Model that each have the same structure and quantum numbers, but differ significantly in the mass values of the constituent particles. Most of the observable matter in the universe is composed of particles from the first generation of fermions. Table 1.2 summarizes the fermions of the Standard Model and their properties.

Particle	Symbol	Charge ($\|e\|$)	Mass (MeV/c^2)	Interactions
Quarks				
up	u	+2/3	1.5-5	strong, weak, EM
down	d	-1/3	3-9	strong, weak, EM
charm	c	+2/3	1100-1400	strong, weak, EM
strange	s	-1/3	60-170	strong, weak, EM
top	t	+2/3	172000	strong, weak, EM
bottom	b	-1/3	4100-4400	strong, weak, EM
Leptons				
electron	e	-1	0.511	weak, EM
electron neutrino	ν_e	0	≈0	weak
muon	μ	-1	105.7	weak, EM
muon neutrino	ν_μ	0	≈0	weak
tau	τ	-1	1777.1	weak, EM
tau neutrino	ν_τ	0	≈0	weak

Table 1.2: Standard Model Quarks and Leptons. Charges are expressed in terms of the magnitude of the electron charge, $|e|$.

Leptons

Leptons are fermions that can interact only through the electroweak force. Each generation contains one lepton with nonzero electric charge and one neutral lepton known as a *neutrino*. It is experimentally observed that the electroweak interaction is a parity violating interaction, consequently left-handed and right-handed particles are represented differently in the theory. The $SU(2)$ doublets for the three generations of left-handed leptons are written as:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_{\mathrm{L}} \tag{1.2}$$

while the $SU(2)$ singlets for the three generations of right-handed leptons appear as:

$$(e_{\mathrm{R}}, \mu_{\mathrm{R}}, \tau_{\mathrm{R}}). \tag{1.3}$$

The neutrino sector of the Standard Model is currently the source of much excitement in the high-energy physics community. Right-handed neutrinos (and left-handed antineutrinos) are not present in the Standard Model due to their absence from experimental data.

Recent experimental results indicate that neutrinos oscillate in flavor and are not massless (m_{ν_e} <2 eV, m_{ν_μ} <190 keV, m_{ν_τ} <18.2 MeV)[11, 12, 13, 14]. In fact, there might be additional massive "sterile" neutrinos beyond the three flavors predicted by the Standard Model[15]. Massive neutrinos may indicate that there are right-handed neutrinos (and left-handed antineutrinos) that need to be included in the Standard Model. Massive neutrinos can be accommodated by the Standard Model if they are Majorana particles (not-conserving lepton number) instead of Dirac particles (conserving lepton number), and future experiments will attempt to determine which scenario describes the neutrino sector.

Quarks

Quarks are fermions that can interact with both the electroweak and strong forces. As with the leptons, right-handed and left-handed quarks are represented differently in the theory describing their electroweak interactions. The $SU(2)$ doublets for the three generations of left-handed quarks are written as:

$$\begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} c \\ s \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} t \\ b \end{pmatrix}_{\mathrm{L}} \tag{1.4}$$

while the $SU(2)$ singlets for the three generations of right-handed up-type and down-type quarks appear as:

$$(u_{\mathrm{R}}, c_{\mathrm{R}}, t_{\mathrm{R}}) \qquad (d_{\mathrm{R}}, s_{\mathrm{R}}, b_{\mathrm{R}}). \tag{1.5}$$

Unlike leptons, quarks can interact with each other via the strong force by exchanging gluons. The color charge of a quark is either red, green, or blue. Each quark, q, transforms as a triplet under $SU(3)$

$$q = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix}. \tag{1.6}$$

1.1.3 The Higgs Mechanism

Masses in the Standard Model arise through interactions with a scalar field, known as the Higgs field, that permeates all space. The Higgs field couples to the bosons and fermions without spoiling the gauge-invariance or renormalizability of the Standard Model. The Higgs field can be thought of as a viscous fluid that all particles have to constantly travel through, and the resulting drag on the particles due to this fluid is what causes mass. The interaction with the Higgs field mixes the $SU(2)$ and $U(1)$ gauge bosons and endows them with mass, in a process known as *spontaneous symmetry breaking*, giving rise to the more familiar mass eigenstates of equations 1.7-1.9 that are observed in experiments[16].

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp \imath W_\mu^2) \tag{1.7}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g^2 + g'^2}}(gW_\mu^3 - g'B_\mu) \tag{1.8}$$

$$A_\mu^0 = \frac{1}{\sqrt{g^2 + g'^2}}(g'W_\mu^3 + gB_\mu). \tag{1.9}$$

The g and g' variables in Eqs. 1.7-1.9 characterize the coupling strengths of the fields, and in fact can be identified with the charge of the electron, e:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}. \tag{1.10}$$

The masses of the $W^\pm$ and Z^0 bosons result from their interaction with the Higgs field, and can be written as:

$$m_W = \frac{v}{2}g \tag{1.11}$$

$$m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \tag{1.12}$$

where v is the vacuum expectation value of the Higgs field when the Higgs is represented as an $SU(2)$ doublet of complex scalar fields. [1] The photon (A^0 in Eq. 1.9) remains massless, which is a consequence of the conservation of electric charge.

Excitations of the Higgs field from its vacuum expectation value show up in the form of the Higgs boson, H. The Higgs boson couples to mass, meaning that it prefers to decay to the heaviest possible channels available to it. Figure 1.2 displays the predicted branching ratios of the Standard Model Higgs as a function of possible mass values[17]. It is worth noting, as it pertains to the main focus of this thesis, that the current expected value of the Higgs mass is $89\pm^{42}_{30}$ GeV/c^2 (<175 GeV/c^2 at 95% CL), indicating that a decay to a $b\bar{b}$ quark pair seems favored[18]. Direct searches for the Standard Model Higgs at the LEP experiments have excluded mass values below 114.4 GeV/c^2 at the 95% CL.

[1] $W^\pm$ will be used to generically refer to either a W^+ or W^- boson.

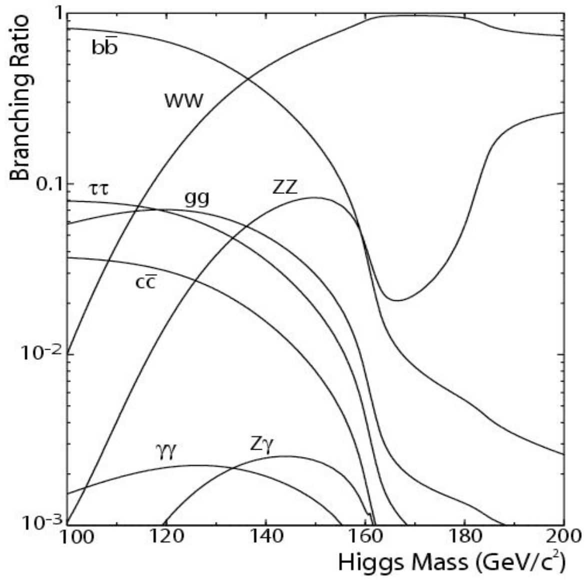


Figure 1.2: The branching ratio of the Standard Model Higgs in different decay channels as a function of possible mass values.

1.1.4 $W^{\pm}$ Decays

Identifying proton-antiproton collisions in which a $W^{\pm}$ boson is produced will be a crucial part of this analysis. The $W^{\pm}$ boson decays into fermion pairs that differ in electric charge by ± 1 when kinematically allowed (*i.e.* - decays to a top-bottom quark pair are highly suppressed due to the large mass of the top quark). Figure 1.3 displays the predominant decay modes of the $W^{\pm}$, which approximately 2/3 of the time is to a quark-antiquark pair and the remaining 1/3 of the time is to a charged-lepton/neutrino pair. This analysis will focus on the electron and muon channels which the experimental apparatus is well suited to study, as will be explained in the following chapters. The final cross-section measured will include the branching-ratio for a $W^{\pm}$ to decay leptonically.

The cross-section for $W^{\pm}$ production drops with each additional high transverse momentum (p_T) parton present in the final-state of the hard-scatter collision. [2] Inclusive ($\geq$0 final-state partons with $p_T \geq$15 GeV/c) $W^{\pm}(\rightarrow \ell^{\pm}\nu)$ production has a cross-section of $\approx$2

[2] The transverse component of momentum will be used extensively in the experimental identification of charged particles.

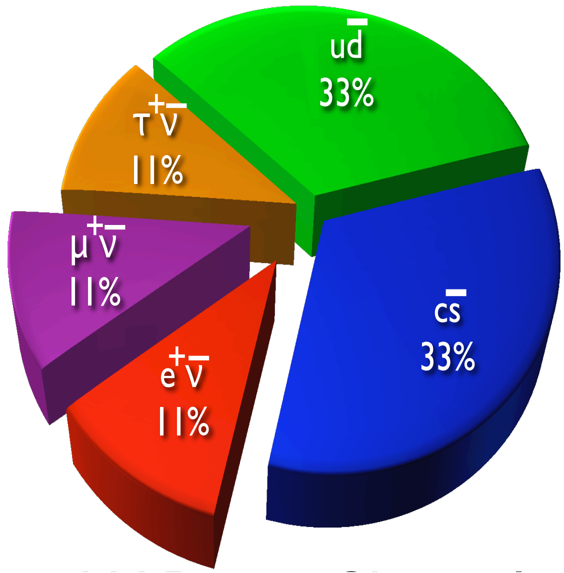


Figure 1.3: Decay channels of the W^+ boson.

nb, which drops to ≈.7 nb with the requirement of one or more final-state partons with $p_T \geq 15$ GeV/c, and to ≈0.2 nb with the requirement of at least two final-state partons with $p_T \geq 15$ GeV/c, and so on[19]. The difficulty of theoretical calculations increases dramatically as more partons are required in the final state.

1.2 Hadron Collider Cross-Sections

Protons have internal structure in the form of quarks and gluons, which makes calculations at a hadron collider like the Tevatron difficult. At sufficiently high energies the QCD coupling between proton constituents, or *partons*, is diminished to the point that collisions between protons and antiprotons can be treated as free particle interactions of one parton in the proton with another in the antiproton. In this high-energy limit the total cross-section for a process can be expressed in the parton model as a sum over all possible parton interactions, as represented in Eq. 1.13.

$$\sigma(p\bar{p} \to X) = \sum_{i,j} \int dx_i dx_j f_{i/p}(x_i, Q^2) f_{j/\bar{p}}(x_j, Q^2) \sigma_{ij}(Q^2, \beta_1, \beta_2, ...) \tag{1.13}$$

$\sigma_{ij}(Q^2, \beta_1, \beta_2, ...)$ is the parton level cross-section for the process ($ij \rightarrow X$) that is calculated using perturbation theory, and can depend on some number of relevant parameters (β_1, β_2, ...). Q^2 is the value of the factorization and renormalization scale (typically set equal to each other), chosen to characterize the momentum scales in the collision, which is introduced in the factorization and renormalization procedures[20]. The *parton distribution functions* (PDFs), $f_p(x_i, Q^2)$ and $f_{\overline{p}}(x_j, Q^2)$, characterize the probability for a constituent parton in the proton (antiproton) to carry a fraction, x_i (x_j), of the proton's (antiproton's) total momentum. The PDFs are determined from fits to several data sources, such as deep inelastic scattering (DIS) experiments at low Q^2[21]. The PDFs determined in these low Q^2 DIS experiments can be evolved to the higher Q^2 values that characterize collisions at the Tevatron using the DGLAP equations[22, 23, 24]. Figure 1.4 shows PDF functions for different Q^2 values from a particular fitting procedure.

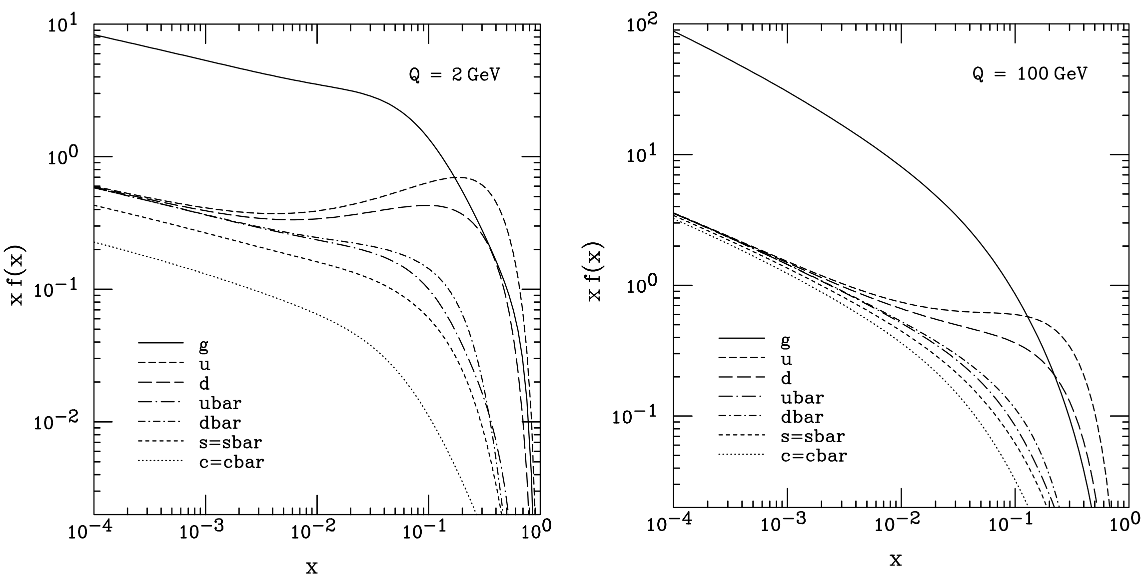


Figure 1.4: Examples of PDF functions at Q values of 2 GeV (Left) and 100 GeV (Right).

As will be discussed in the following chapters, the cross-section for a process can also be measured by counting the number of occurrences of the process in a given amount of data. Monte Carlo techniques used in this analysis will employ the full treatment of Eq.

1.13, which requires knowledge of the parton level cross-sections ($\sigma_{ij}(Q^2, X)$).

1.3 $W^{\pm}+b\bar{b}$

$W^{\pm}+b\bar{b}$ production is one possible process that can occur when quarks are made to collide at very high energies. As will be shown, this process is of much interest due to its challenging QCD/electroweak nature, and also because of its role as a significant background source in searches for the Standard Model Higgs and other new phenomena.

$W^{\pm}+b\bar{b}$ refers to the associated production of a $W^{\pm}$ boson, a pair of b-quarks ($b\bar{b}$) which are almost entirely produced by gluon splitting (*i.e.* - $g \rightarrow b\bar{b}$), and any number of additional observed and unobserved partons in a proton-antiproton collision. The leading-order (LO) Feynman diagram, which is a notation used in the calculation of probabilities, is depicted in Figure 1.5. This is an important and challenging process that must be accounted for since it is a leading source of background, and uncertainty, in Higgs searches, single-top searches, and the $t\bar{t}$ cross-section analysis. Figure 1.6 depicts Feynman diagrams for Standard Model Higgs and s-channel single-top processes, which have the same final-state as $W^{\pm}+b\bar{b}$.

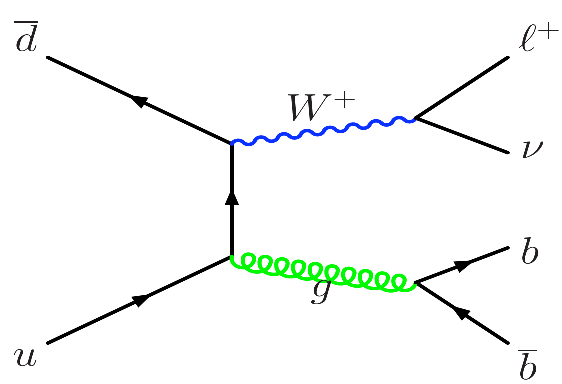
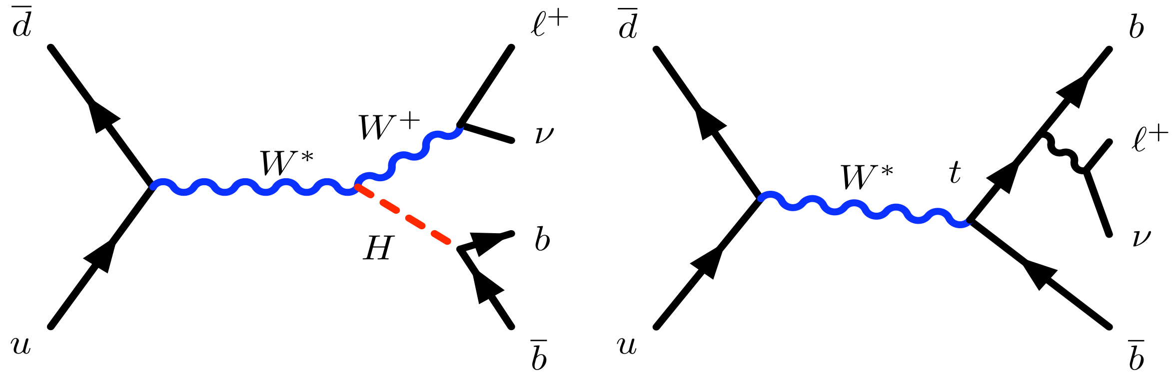


Figure 1.5: Leading order Feynman diagram for $W^+ + b\bar{b}$ production. A similar diagram can be drawn for the charge conjugate process.

Conclusively discovering these processes in a collider experiment will require under-

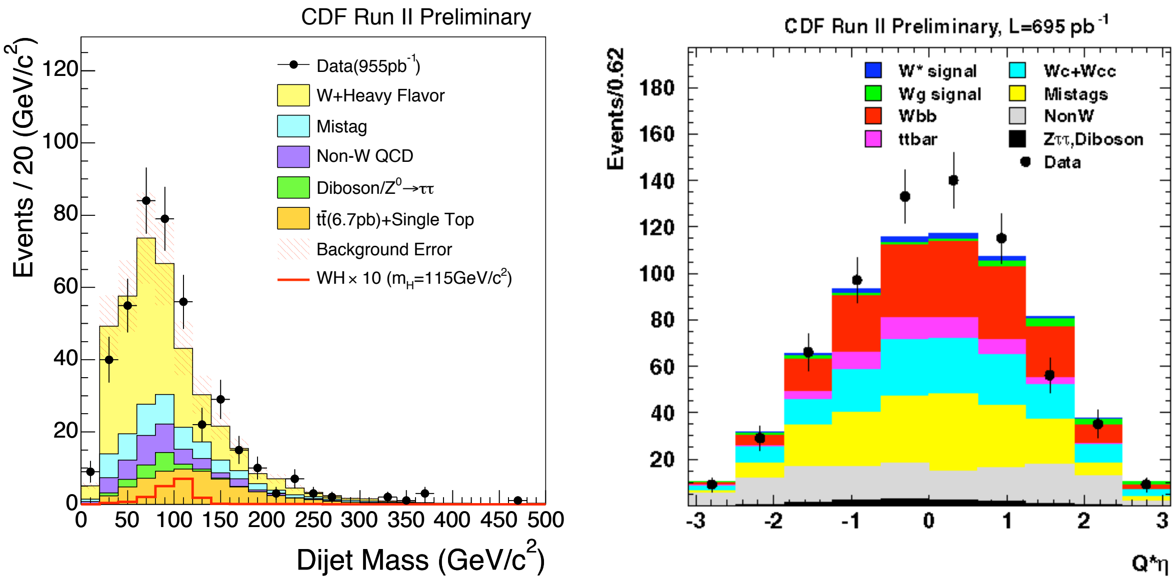


(a) ***Standard Model Higgs*** (b) ***s-channel Single-Top***

Figure 1.6: Feynman diagram for Standard Model Higgs and s-channel single-top quark production. Similar diagrams can be drawn for the charge conjugate processes.

standing the contributions that $W^{\pm}$+$b\bar{b}$ makes in the data. Figure 1.7(a) shows recent results from CDF searches for the Standard Model Higgs, and Figure 1.7(b) summarizes recent results for single-top production[25, 26] The left-hand figure is roughly the equivalent of the invariant mass of the parton pair in $W^{\pm}$+2 parton events (which should contain a peak around any Standard Model Higgs signal), while the right-hand figure is the charge of the lepton times the pseudorapidity of a parton in $W^{\pm}$+1,2,3 parton events. [3] In both cases the desired signal must contend with large amounts of background processes, of which $W^{\pm}$+$b\bar{b}$ is among the leading contributers. These results indicate the Standard Model Higgs process described by Figure 1.6(a) has a cross-section less than 3.9-1.3 pb (95% C.L. for masses ranging from 110 GeV/c^2 to 150 GeV/c^2), and the single-top production process for the combined s (Figure 1.6(b)) and t channels has a cross-section less than 3.4 pb (95% C.L.). Theoretical predictions for the total $W^{\pm}$+$b\bar{b}$ cross-section are approximately 10 times higher than these Higgs and single-top limits, which underscores why it is very important to have good estimates of the contribution of $W^{\pm}$+$b\bar{b}$ to the data.

[3] Pseudorapidity is a variable which will be used in place of the polar angle at hadron collider experiments.



(a) ***Standard Model Higgs Search*** (b) ***Single-Top Search***

Figure 1.7: Left: Search for SM Higgs peak in dijet mass spectrum of $W^{\pm}$+2 jet events. Right: Search for single-top quark production in lepton charge × jet rapidity distribution with $W^{\pm}$+jets data.

1.3.1 $W^{\pm}+b\bar{b}$ Theoretical Issues

$W^{\pm}+b\bar{b}$ is a challenging process to calculate in perturbation theory[27, 28]. Typically Monte Carlo techniques are relied upon to estimate the cross-sections and kinematics of $W^{\pm}+b\bar{b}$ production, though the available programs often only utilize the LO Feynman diagram in the perturbative expansion for this process. Calculations that incorporate the next-to-leading-order (NLO) diagrams for $W^{\pm}+b\bar{b}$ have only become available in the past few years[29]. It is easier to perform calculations for massless quarks, but the exact treatment of the b-quarks that come from the gluon splitting can have a significant impact. Divergences can appear in the calculations if the b-quarks are treated as massless and the $b\bar{b}$ pair from the gluon splitting are collinear or have very low momentum, making it difficult to compare with experimental results.

1.3.2 $W^{\pm}$+$b\bar{b}$ Methods at CDF

The CDF experiment, which will be discussed at length in the next chapter, is capable of identifying proton-antiproton collisions (*events*) in which a $W^{\pm}$ boson and some number of final-state partons were produced. Partons appear as collimated sprays of particles in the CDF detector, known as *jets*, while leptonic $W^{\pm}$ decays can be identified by the presence of a highly energetic charged lepton and a large amount of unaccounted for transverse energy due to the escaping neutrino.

Identifying $W^{\pm}$+$b\bar{b}$ production at CDF will be challenging for several reasons. The definition of a jet can have a significant impact on the measured results, so studies will have to be performed to see how well CDF jets represent the partons that created them. There will be many competing processes, or *backgrounds*, that can mimic the signature of $W^{\pm}$+$b\bar{b}$ so care must be taken to remove as many of these events as possible. Finally, $W^{\pm}$+$b\bar{b}$ production makes up only a small fraction of generic $W^{\pm}$+jet events, the vast majority of which are dominated by lighter flavor (u, d, s, c, g) partons in the final-state. Techniques for identifying b-quark production must be employed to isolate the desired $W^{\pm}$+$b\bar{b}$ events over the large background of generic $W^{\pm}$+light-flavor events.

Method II

The standard method of accounting for $W^{\pm}$+$b\bar{b}$ at CDF relies entirely on LO Monte Carlo predictions that take into account the mass of the b-quark. [4] Using Monte Carlo, the fraction of $W^{\pm}$+jet events due to $W^{\pm}$+ $b\bar{b}$ is determined, and this "heavy-flavor" fraction is multiplied by the observed number of $W^{\pm}$+jet events in the data to determine the number of $W^{\pm}$+$b\bar{b}$ events in the data. This procedure, which is also used to estimate the contributions of $W^{\pm} + c\bar{c}$ and $W^{\pm} + c$, is referred to as the "Method II" analysis

[4] Here, and in the following, "LO" will refer to calculations performed using only leading order matrix-element calculations interfaced with a parton shower event generator.

[30]. Figure 1.8 shows a summary of the results of the Method II analysis using the same data sample that will be used in this analysis. This distribution represents the sample composition of the $W^{\pm}+\geq 1$ jet event sample, for events in which at least one of the jets has been identified as coming from a heavy-flavor hadron decay. There is a large amount of $W^{\pm}+b\bar{b}$ present in the 1 and 2 jet bins.

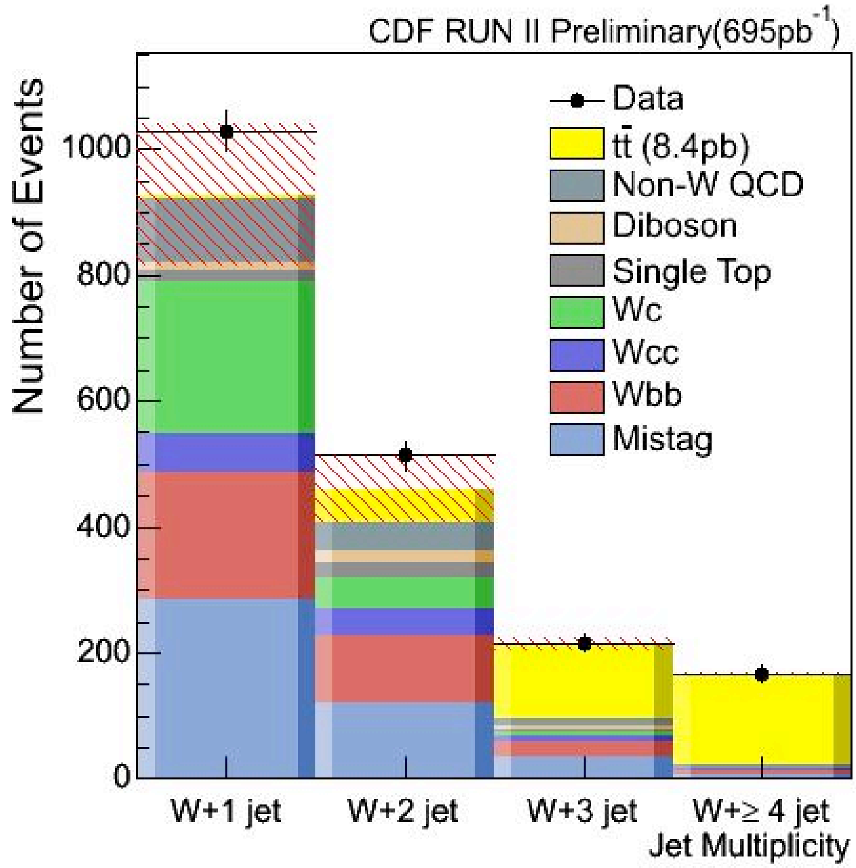


Figure 1.8: N-jet sample composition plot from the Method II analysis at CDF.

The Method II analysis suffers from large uncertainties due to its estimation of the $W^{\pm}+b\bar{b}$, $W^{\pm}+c\bar{c}$, and $W^{\pm}+c$ processes. The heavy-flavor fractions for these processes, derived from LO Monte Carlo, are scaled by a "K-factor" of 1.5±0.4 (derived from comparisons between data and the absolute predictions of Alpgen for dijet processes) that is supposed to mimic the impact that NLO effects would have on the calculated fractions. The large uncertainty on the K-factor causes large uncertainties on the calculated numbers of $W^{\pm}$+$b\bar{b}$, $W^{\pm}+c\bar{c}$, and $W^{\pm}+c$ events in the selected data.

The large uncertainties associated with the $W^{\pm}$+$b\bar{b}$, $W^{\pm}+c\bar{c}$, and $W^{\pm}+c$ processes in the Method II analysis are not of primary concern due to their diminished presence in the $W^{\pm}+\geq$3 jet bin where the $t\bar{t}$ signal of interest dominates. However, in the $W^{\pm}$+1 and $W^{\pm}$+2 jet bins, where likely Higgs and single-top signals show up, the problem of how to precisely account for $W^{\pm}$+$b\bar{b}$, $W^{\pm}+c\bar{c}$, and $W^{\pm}+c$ is more acute.

The goal of this analysis is to measure the cross-section for $W^{\pm}$+$b\bar{b}$ in a more data-driven approach, minimizing as much as possible the uncertainties associated with LO predictions. The measured cross-section will be constructed to be proportional to the number of b-jets from $W^{\pm}$+$b\bar{b}$ events containing one or two jets.

CHAPTER 2

Experimental Apparatus

Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois is currently the home of the world's most energetic man-made particle collisions[31]. These collisions between protons and antiprotons traveling at nearly the speed of light are used to study many interesting processes, such as $W^{\pm} + b\bar{b}$.

This chapter describes how the accelerators at Fermilab produce the protons and antiprotons necessary to create the collisions used to study the physics of interest. The components of the CDF detector that are used to examine the particles produced when a proton and antiproton collide are also described.

2.1 The Accelerator Chain at Fermilab

Preparing the beams of protons and antiprotons that collide at Fermilab at speeds nearing that of light is an intricate process. Several accelerating machines — the Cockcroft-Walton, Linac, Booster, Main Injector, and Tevatron — collaborate, as depicted in Figure 2.1, to gradually push the beams up to the incredible energies necessary to create collisions capable of producing massive particles not ordinarily found since the time of the early universe. The timing of the collisions is almost as important as their energies since the detector experiments must have the opportunity to analyze a given collision and then reset themselves in time for the next *event* (the common term for a single proton-antiproton collision). The accelerating machines used at Fermilab employ radio frequency (RF) cavities to accelerate charged particles and to induce a bunch structure in the beams which allows for the timing of collisions to be controlled. In this section the process of making the beams of protons and antiprotons to be collided is described. Much of the information about the accelerator chain at Fermilab can be found in the accelerator division's Rookie Books[32].

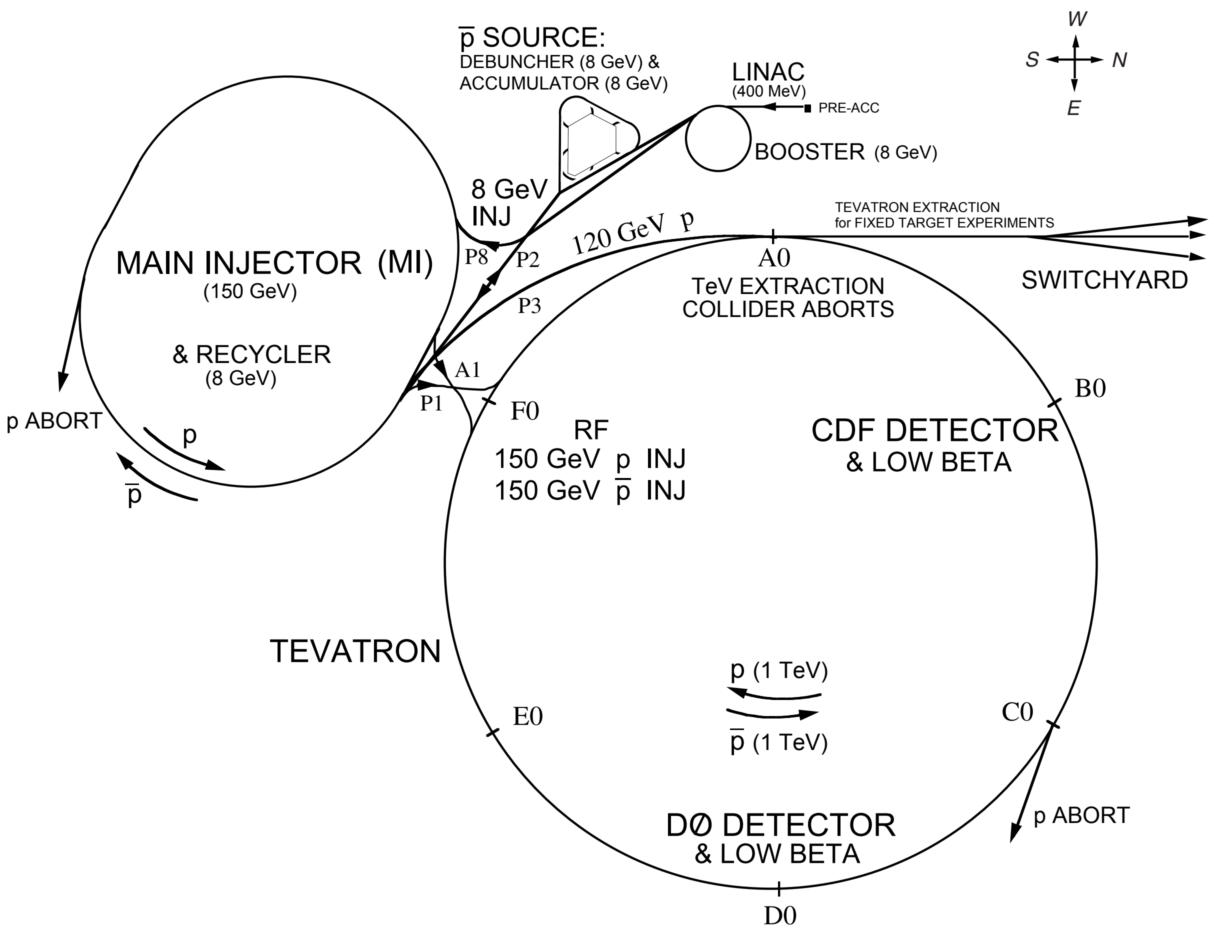


Figure 2.1: The accelerator chain at Fermi National Accelerator Laboratory. Protons and antiprotons are accelerated through various stages, ending up in the Tevatron circulating in opposite directions at 980 GeV. The protons and antiprotons collide at a center-of-mass energy of 1.96 TeV at two positions (BØ and DØ, where the CDF and DØ detectors are respectively located).

2.1.1 Cockcroft-Walton Preaccelerator

The initial source of protons at Fermilab is a simple bottle of molecular hydrogen gas. The gas is injected into a device known as a magnetron, which ionizes the hydrogen gas into a plasma of H^+ ions (*i.e.* - protons) and electrons. The H^+ ions collide with a cesium-coated cathode, located in the magnetron, where they sometimes pick up two electrons to form H^- ions. These negative H^- ions will eventually become part of the final proton beam or be used to help create the antiproton beam.

The magnetron is housed in a dome, electrically charged to -750 kV, which is part of the Cockcroft-Walton electrostatic generator. The H^- ions exit the dome and accelerate toward a grounded wall, acquiring 750 keV of energy before exiting the Cockcroft-Walton through a transport line that will take them to the next stage of acceleration. The magnetron and Cockcroft-Walton, collectively known as the *Preacc*, repeat this process every 66 ms (or 15 Hz) producing a pulse of H^- ions each cycle.

2.1.2 Linear Accelerator

The next accelerator the H^- ions arrive at from the 750 keV transport line is a two-stage linear accelerator known as the Linac which will push the ions up to an energy of 400 MeV. The first stage of the Linac, known as the Drift Tube Linac (DTL), contains five RF cavities operating at a frequency of 201.25 MHz. Each cavity consists of a series of electrostatically grounded drift tubes of increasing lengths separated by accelerating gaps, as shown in Figure 2.2. In the gap regions the ions are accelerated towards the next drift tube by the electric field set up in the cavity by the RF generator, while in the the grounded drift tubes the ions are shielded from the RF field and move at constant velocity. The increasing length of the drift tubes is necessary for the ions to always enter the gaps as the RF field is pointing in the accelerating direction. The DTL accelerates the H^- ions from 750 keV to 116 MeV.

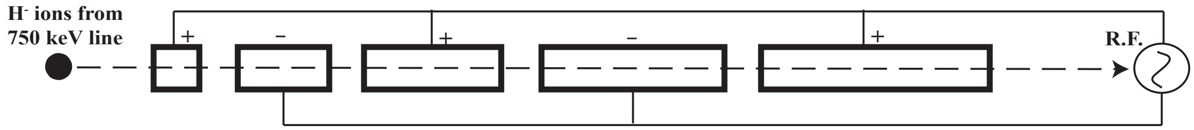
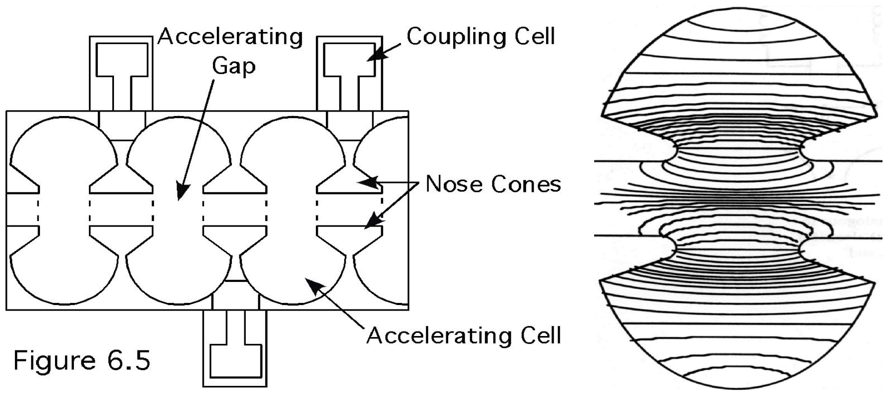


Figure 2.2: Example of an RF linear accelerator. H^- ions are accelerated across a potential difference set up by an RF generator, and then drift through tubes when the RF potential is pointing in the deaccelerating direction. The tubes are gradually increased in length to allow the ions, which are increasing in velocity as they travel through the Linac, to always arrive at the accelerating gap as the RF is pointing in the right direction.

It is already apparent that a Linac setup won't work for a circular accelerator, since the drift tubes would have to be increased in length every time a particle made a revolution and acquired more energy. It should also be noted that the operation of the Linac's RF cavities will create structure in the ion beam since only those ions entering the accelerating gaps in phase with the RF field will consistently increase in energy. These accelerated regions of the beam are referred to as *buckets*, and the collection of ions that occupy a bucket are referred to as a *bunch*.

The second stage of the Linac, known as the Side Coupled Linac (SCL), contains seven side-coupled modules operating at 805 MHz. Each of the modules in the SCL consists of four coupled sections which each contain 16 accelerating cells and 15 coupling cells. Each of the cells within a SCL module help carry the RF signal throughout the module. The shape of the accelerating cells is referred to as a *nose-cone*, and it is this shape which causes a concentrated electric field to be set up along the beam direction when the RF signal is applied, as shown in Figure 2.3. The accelerating cells are separated by coupling cells that the ions drift through while the fields in the neighboring accelerating cells, that are $\pi/2$ out of phase with one another, change direction.

As in the DTL, the modules of the SCL are gradually increased in length to compensate for the increasing speed of the ions. The beam pulses from the DTL travel eight accelerating cells apart in the SCL (due to the different RF frequencies), so up to two pulses can be in the same section, and up to eight pulses can be in the same module. The SCL accelerates the ions from 116 MeV to 400.0 MeV. The entire Linac has the same cycle time of 66 ms (15 Hz) as the Preacc, producing a beam of H^- ions with a 201.25 MHz bunch spacing.



(a) ***SCL Module Section*** **(b)** ***Nose-Cone Electric Field***

Figure 2.3: (a) Schematic of Side-Coupled Linac module section and (b) nose-cone field.

2.1.3 Booster

The H^- ions exit the Linac and are directed down (literally, about 15 feet below the Linac) a 400 MeV transport line into the next accelerator, the Booster. The Booster is the first *synchrotron* accelerator the ions encounter at Fermilab. Unlike a linear accelerator where an RF field of constant frequency is used to increase the energy of ions, the frequency of the RF in a synchrotron is gradually increased along with an increase in the magnitude of a steering magnetic field which bends the particles around in a circular orbit. The net effect is charged particles traveling in a circle of constant radius but with increasing energies every time they encounter the RF. The Booster is a synchrotron 150 m in diameter that will accelerate protons, that are left after stripping the H^- ions of their electrons, from 400 MeV to 8 GeV. The Booster can accelerate beam once every 66 ms (15 Hz), which determines the cycle time of the Preacc and Linac.

The ions exiting the Linac are actually directed at a beam dump site, where they will end up unless acted upon by a device known as a "chopper" which selects a section of the 400 MeV beam and guides it down into the transport line leading to the Booster. Upon entering the Booster, the H^- ions are merged into the path of any proton beam already circulating by passing both beams through two dipole magnets of opposite polarity, as shown in Figure 2.4. The combined beam of protons and H^- ions passes through a carbon foil which strips the H^- ions of their electrons, which are then removed via an "electron catcher". The resulting beam, which is now composed of mostly protons, is separated from any remaining H^- ions by two more dipole magnets of opposite polarity. The chopper is usually run for a length of time equal to several multiples of the period of a 400 MeV particle orbiting the Booster, meaning that protons (after being stripped of their electrons as described above) are layered in on top of protons already circulating. A *turn* is the amount of proton beam equal to one lap of a 400 MeV particle around the Booster ($\approx 2.2\mu s$), and the amount of turns (typically 5 or 6) the chopper layers into the Booster will be one of the main criteria for controlling the intensity of the bunches that collide in the Tevatron.

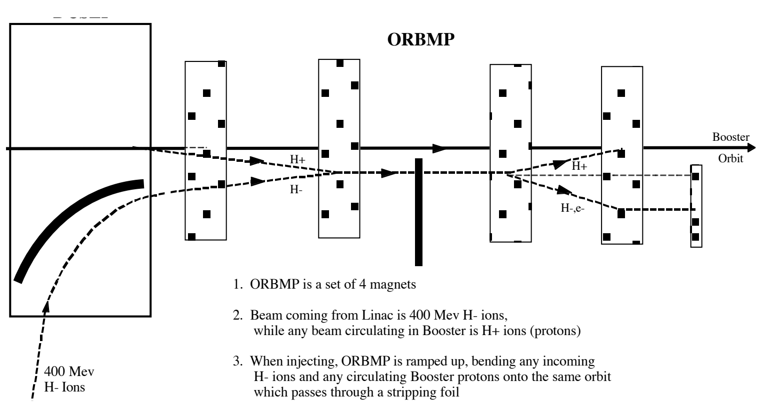


Figure 2.4: A series of dipole magnets of opposing polarities is used to combine the H^- ions coming from the Linac into any H^+ beam already circulating in the Booster.

The Booster uses 96 combined function dipole/quadrupole magnets to steer the proton beam, and 18 RF cavities to accelerate the beam. When the beam in the Booster is fully loaded and ready for acceleration the RF frequency is gradually increased from 37.8 MHz to 52.8 MHz, accompanied by an increase in the field strength supplied by the dipole/quadrupole magnets. It is noted that the initial RF frequency of the Booster (37.8 MHz) is different than the bunch structure of the beam arriving from the Linac (201.25 MHz), which is dealt with in a process known as *paraphasing*. In paraphasing, half of the RF cavities in the Booster are initially $\pi/2$ out of phase with the other half, causing no net increase in energy as a particle orbits (i.e.- the bunches will receive as many deaccelerating kicks as accelerating kicks during one revolution). After the desired number of turns are loaded into the Booster by the chopper, the proton beam is allowed to circulate with the RF cavities in their out of phase state, eventually causing the beam to lose its 201.25 MHz bunch spacing and form an almost DC beam. The RF cavities are then slowly brought back into phase over several hundred nanoseconds, which induces the 37.8 MHz bunch spacing structure to form, and then the acceleration process can begin.

2.1.4 Main Injector

The proton beam from the Booster travels down the MI-8 transport line to the next stage in acceleration, the Main Injector (MI). The MI is a synchrotron accelerator that is elliptical in shape and has a circumference of 10890 feet. 344 dipole and 208 quadrupole magnets steer beams of protons and antiprotons, while 18 RF cavities accelerate them. The RF frequency of the MI starts at 52.8 MHz and is gradually increased to 53.1 MHz, accompanied by an increase in the field strength of the steering magnets. It is noted that, unlike the transfer from the Linac to the Booster, the MI is matched in RF frequency to the Booster during transfer allowing for the bunch spacing to be preserved. This is known as a *bucket-to-bucket* transfer.

The Main Injector has several functions. It accelerates protons and antiprotons up to an energy of 150 GeV for injection into the Tevatron, it accelerates protons up to an energy of 120 GeV to be sent to the antiproton source, and it also receives 8 GeV antiprotons from the Accumulator (described below) and sends them to the Recycler (described below).

To initiate collisions for the experiments located on the Tevatron, the MI receives 7 proton bunches from the Booster and accelerates them up to 150 GeV. These 7 bunches are merged into one larger bunch during a process known as "coalescing", which is necessary to ensure the beam intensities required to have statistics adequate to see rare processes, and the coalesced bunch is sent to the Tevatron via the P1 transport line. This process is repeated 36 times to form the 36 proton bunches that will be used by the collider experiments.

To form the antiproton bunches necessary for the collider experiments, four sets of seven bunches are sent from the antiproton source (described below) to the MI and accelerated to 150 GeV. The sets of seven bunches are coalesced to form four larger bunches which are sent to the Tevatron via the A1 transport line, circulating in the opposite direction as the proton bunches. This process is repeated nine times to form the 36 antiproton bunches that will be used by the collider experiments.

2.1.5 Tevatron

The final stage of acceleration for protons and antiprotons is a synchrotron 2000 m in diameter known as the Tevatron. The Tevatron will accelerate the particles from 150 GeV to 980 GeV and ultimately collide the beams at two places at a center of mass energy of 1.96 TeV. 774 dipole and 216 quadrupole magnets steer the beams, while eight RF cavities accelerate them by increasing from a frequency of 53.103 MHz to one of 53.104 MHz.

Unlike every other accelerator at Fermilab, the magnets in the Tevatron are made of superconducting niobium/titanium that must be cryogenically cooled to 4 K to produce the superconducting fields. This choice was made due to the large magnetic fields, that typical resistive magnets would require excessively large currents to achieve (at a much greater cost), necessary for steering the highly relativistic protons and antiprotons.

The proton and antiproton beams orbit the Tevatron in opposite directions in a helix formation, which allows them to be in the same beampipe without unwanted collisions taking place. At two places on the ring known as BØ and DØ, home of the CDF and DØ experiments respectively, the beams are focused very tightly together by large quadrupole magnets located on either side of the experiments. This tight focusing brings the beams into an almost head-on orbit, allowing high-energy collisions to occur. After passing through the experiments the beams are separated by the large quadrupole magnet on the opposing side. Once injected into the Tevatron the beams can provide a sufficient collision rate for upwards of two days, though the quality of the beams is gradually reduced as more collisions occur. The length of time a given set of beams is allowed to collide in the the Tevatron is known as a *store*, and Fermilab is generally colliding beams 24 hours a day for nine months every year, punctuated by extended periods of maintenance and equipment upgrades.

2.1.6 Antiproton Production: Debuncher, Accumulator, and Recycler

The production and storage of antiprotons is a multistep process involving several accelerator components. The process begins with a batch of 120 GeV protons from the MI being directed towards a nickel target. Just before the protons are sent to the target the bunches undergo a 90° rotation in phase space, accomplished by manipulating the phases of RF cavities of the MI such that the net voltage the bunches travel through during a revolution is severely decreased. The bunches go from having a very small spread in energy and significant spread in time (with respect to the RF phase), to having a large spread in energy and small spread in time.

The collision of the rotated proton bunch on the nickel target produces many types of particles with large variation in momentum (due to the spread in energy of the rotated bunch), among which are antiprotons (typically only ≈15 antiprotons are collected for every million incident protons) with an average energy of 8 GeV. The benefit of rotating the proton bunch as described above is that the particles coming off the nickel target will be closely spaced in time, allowing for a more efficient capture.

The particles created in this collisions are directed onto a Lithium lens with a 0.65 MA current running through it, which has the effect of focusing the particles' momentum. The focused beam of particles is then sent through a dipole magnet that is tuned to deflect negatively charged 8 GeV particles away from the other particles and towards the Debuncher, as shown in Figure 2.5.

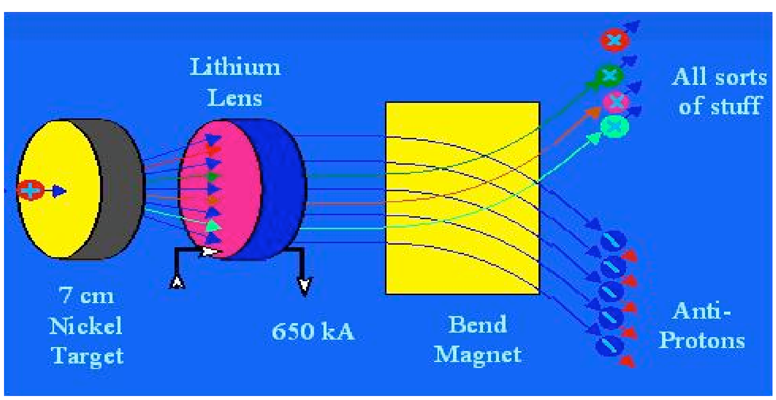


Figure 2.5: Schematic of antiproton production. Protons impact on a nickel target, producing numerous secondary particles that are focused and steered to isolate any antiprotons present.

The Debuncher is a triangular storage ring that is 90 m in radius on average. The 8

GeV antiprotons arriving from the nickel target have a significant spread in longitudinal momentum. By performing another 90° rotation in phase space, in the opposite direction as was done to the proton bunches sent to the nickel target from the MI, the momentum spread is reduced. The transverse spread of the beam is further reduced via stochastic cooling (the term ‘cooling’ is used generically at Fermilab to refer to any process in which the particles in a bunch are made to be more uniform in momentum/timing/etc...), which involves measuring the deviation of antiprotons from a desired orbit and then sending a signal to a “kicker”, located across the Debuncher, before the measured antiprotons arrive. The kicker applies an electrostatic force to the antitprotons when they arrive which corrects their orbit back to the desired one. After being stochastically cooled the antiprotons are transferred to the Accumulator, which is another triangular storage ring in the same tunnel as the Debuncher. This process occurs on a time scale of 2-4 seconds, just before the next bunch of antiprotons arrives from the target. When enough antiprotons have been gathered, or “stacked”, in the Accumulator they are sent to either the MI or to the Recycler.

The Recycler is a storage ring that sits in the same enclosure as the Main Injector. The original purpose of this ring was to take antiprotons from the Tevatron that are left over at the end of a store and maintain them at 8 GeV for use in subsequent stores. This function was never achieved, and instead the Recycler has been used in the stacking process to receive antiprotons from the Accumulator and maintain them, thus freeing up the Accumulator to receive more antiprotons from the Debuncher. About every four hours the stack of antiprotons in the Accumulator are transferred to the Recycler.

To keep the antiprotons cooled in the Recycler, *electron cooling* is used to reduce the longitudinal spread. Electron cooling involves circulating a beam of electrons, with a very small spread in energy, parallel to an uncooled antiproton beam. The beams will interact, with the cold electron beam receiving energy from the warm antiproton beam, effectively cooling the antiproton beam. Since August 2005, electron cooling has been used for every store, significantly helping to make more uniform antiproton beams.

Initially in the Run II period of data taking, only the Accumulator was used to transfer antiprotons to the Tevatron, which limited the resulting collision rates since the Accumulator could only hold a certain amount of antiprotons. In January of 2005, both the Accumulator and Recycler supplied the Tevatron with antiprotons in what were referred to as "mixed-mode" shots, which established the usefulness of storing antiprotons in the Recycler. By the end of 2005, only the Recycler was being used to provide the antiprotons to the Tevatron, almost doubling the number of antiprotons available for collisions.

2.1.7 Luminosity

Luminosity is the quantity used to describe the potential for collisions to occur when colliding beams of protons and antiprotons. It is dependent on the rate at which the bunches of protons and antiprotons come into contact, as well as on the structure of the bunches. Luminosity is most simply defined as:

$$L = \frac{fnN_pN_{\overline{p}}}{A} \tag{2.1}$$

for n bunches of protons and antiprotons orbiting with frequency f and composed of N_p protons and $N_{\overline{p}}$, respectively, with cross-sectional overlap area A. The numerator of Equation 2.1 gives the total rate of possible proton-antiproton interactions.

The proton and antiproton beams could have different cross-sectional areas, so the luminosity is typically modified to reflect this by assuming that the transverse spread is Gaussian with respective widths σ_p and $\sigma_{\overline{p}}$ at the interaction points, giving:

$$L = \frac{fnN_pN_{\overline{p}}}{2\pi\left(\sigma_p^2 + \sigma_{\overline{p}}^2\right)}H(\frac{\sigma_l}{\beta^*}) \tag{2.2}$$

Here, the luminosity is further modified by the multiplication of a form factor, $H(\frac{\sigma_l}{\beta^*})$, that depends on the longitudinal (along the beam direction) spread of the bunches and the value of β^*, which is a measure of the beam's transverse spread at the interaction point. Values for the relevant parameters of the luminosity calculation during the early periods of Run II are shown in Table 2.1, which indicate a luminosity value of approximately 10^{31} $\mathrm{cm^{-2}s^{-1}}$

Bunches ($p\times\overline{p}$)	36×36
Protons/bunch (N_p)	2.7×10^{11}
Antiprotons/bunch ($N_{\overline{p}}$)	3.0×10^{10}
β^* (cm)	35
Bunch length (m)	0.37
Bunch spacing (ns)	396

Table 2.1: Typical values for Luminosity parameters in the early stages of Run II.

2.2 The Collider Detector at Fermilab

We now have a sense of how the accelerators at Fermilab produce the 1.96 TeV collisions of protons and antiprotons in the Tevatron that occur millions of times per second. Creating the collisions is, of course, only half the battle in trying to understand them. In order to gain a complete understanding of what occurs when protons and antiprotons collide, a device is needed that is capable of fully reconstructing each collision that occurs in the Tevatron. This device needs to be able to identify many different types of particles that might be produced in a collision along with their kinematical properties, and also the device needs to be fast enough to record all this information in the window between consecutive collisions. The device should have regions of minimally absorbing detectors which give accurate measurements of the particle's charge and momentum without perturbing these quantities significantly, and regions of maximally aborbing detectors that turn the total energy of a particle into an observable signal that can be recorded. The *hermeticity*, a measure of the instrumented spatial coverage surrounding a given collision, of the device should be very near to 4π to ensure that the full kinematical space of possible collisions is recordable. Furthermore this apparatus must be able to separate the proverbial wheat from the chaff and identify potentially "interesting" events, which must be defined, from garden-variety collisions since the collision rate at Fermilab is far too great and contains too much information to physically allow for every single collision to be examined and stored.

The unique way in which different particles are absorbed (or not absorbed, in the case of muons and neutrinos), shown schematically in Figure 2.6, is used to to determine their identities. By combining multiple layers of detection to identify all the particles created in a collision, the proposed device described above can give a very good picture of the proton-antiproton collision that took place.

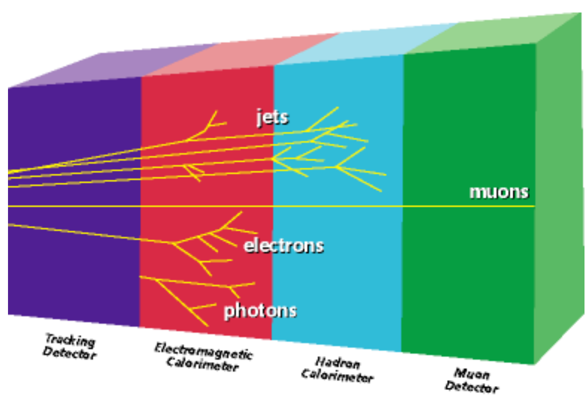


Figure 2.6: The unique ways in which particles of different types are absorbed, or not absorbed, by several layers of detection allows for their identification.

The manifestation of the described device used for this analysis comes in the form of the Collider Detector at Fermilab (CDF) experiment, which is used to capture as much information as possible about the high-energy collisions of protons and antiprotons in the Tevatron[33, 34]. It is centered around the BØ collision point of the Tevatron, and is instrumented with many detector subsystems that allow for near-full event reconstruction.

The detector is nearly forwards-backwards as well as azimuthally symmetric and, though not completely hermetic, provides excellent detection capability over most of the 4π region. It has components that allow for the identification of the charge, momentum, and energy of particles produced in collisions and a fast electronic triggering system that allows for the quick selection of interesting collisions to be recorded for physics analysis. Figure 2.7 shows a longitudinal view of one quadrant of CDF.

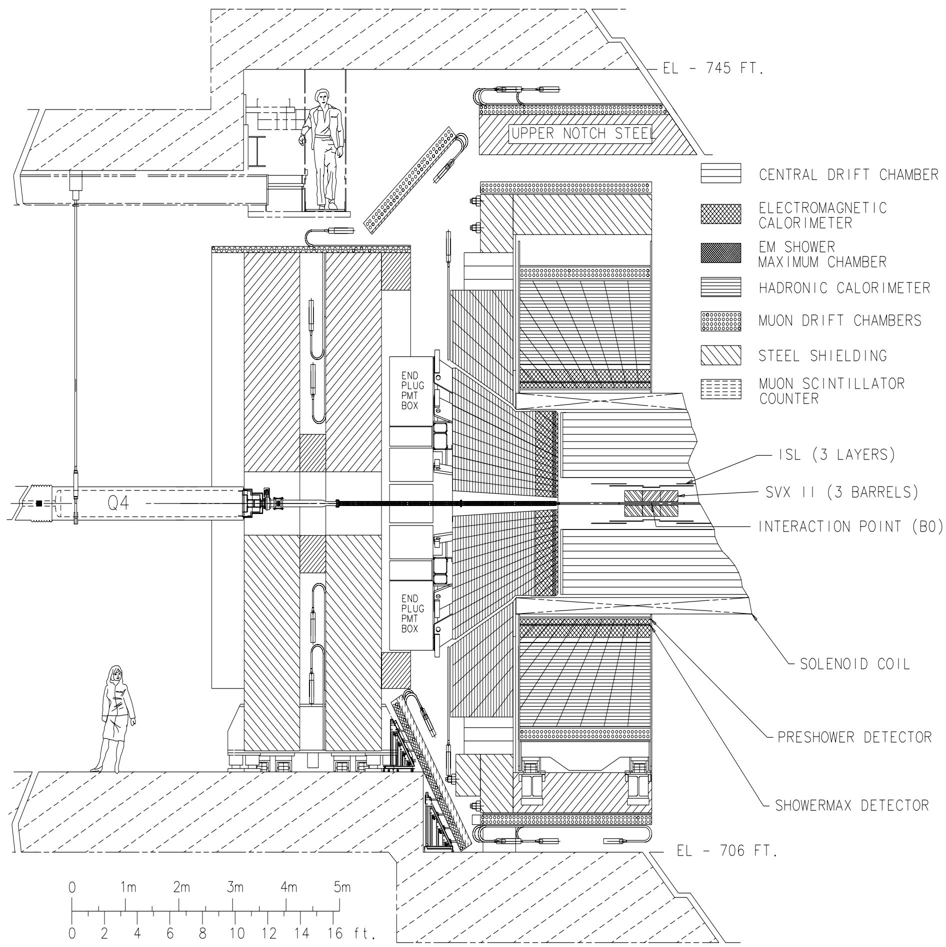


Figure 2.7: Longitudinal cross-section of one quadrant of the CDF detector.

2.2.1 Coordinates

The coordinate system used at CDF is right-handed with the origin located in the center of the detector, the positive $\hat{z}$-axis pointing along the direction of the proton beam, the positive $\hat{x}$-axis pointing radially outwards in the plane of the Tevatron, and the positive $\hat{y}$-axis pointing vertically upwards. Due to the mixture of spherical and cylindrical shapes of parts of the CDF detector, polar coordinates are more commonly used, with the polar angle (θ) measured from the positive $\hat{z}$-axis and the azimuthal angle (ϕ) measured from the positive $\hat{x}$-axis in the $x-y$ plane (also referred to as the "transverse" plane). The polar coordinates can be written in terms of the cartesian coordinates as follows:

$$r = \sqrt{x^2+y^2}, \tag{2.3}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right), \tag{2.4}$$

$$\theta = \tan^{-1}\left(\frac{r}{z}\right). \tag{2.5}$$

Another useful coordinate known as *rapidity* is defined as:

$$y \equiv \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \tag{2.6}$$

for a particle of energy E and with a momentum of p_z along the z-axis (rapidity is not to be confused with the spatial y-component, though unfortunately the symbol for both is the same). Rapidity transforms as $y \to y + \tanh^{-1}\beta$ for Lorentz boosts along the z-axis from one inertial frame to another with relative-velocity β along that axis. Consequently, differences in rapidity are Lorentz-invariant quantities. In the limit of highly relativistic particles ($p \gg m$) the rapidity can be approximated by the *pseudorapidity* (also called "eta", η), defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.7}$$

This quantity is used in two different manners at CDF: the event pseudorapidity (η) is defined with respect to the interaction point of a given collision, while the detector pseudorapidity (η_{det}) is defined with respect to the center of the detector. The multiplicity of particles produced in the $p\overline{p}$ collisions at Fermilab is nearly constant in η.

2.2.2 Charged-particle Tracking

To identify the charge and transverse-momentum of particles produced in collisions the innermost sections of the CDF detector are surrounded by a superconducting solenoid 5 m in length and 3.2 m in diameter that produces a 1.4 T magnetic field parallel to the z-axis. Particles with a non-zero electric charge will be deflected by this magnetic field and travel in helical orbits (*i.e.* - circular orbits in the $x-y$ plane and constant velocity along the z-axis). By measuring the radius of the helix in the $x-y$ plane the transverse momentum can be determined, and by noting the orientation of the helix rotation (clockwise vs. counterclockwise) the sign of the charge can be determined. CDF has two detector systems, the Silicon detector and the Central Outer Tracker (COT), located inside of the solenoid that locate charged particles as they move outward from the interaction vertex and allow for the reconstruction of the helical trajectories. Tracks with only COT information have a transverse momentum (p_T) resolution of $\delta p_T/p_T^2 = 0.0030/\text{GeV}/c$, while tracks with both COT and silicon information have resolution $\delta p_T/p_T^2 = 0.0012/\text{GeV}/c$. Figure 2.8 depicts the tracking volume at CDF.

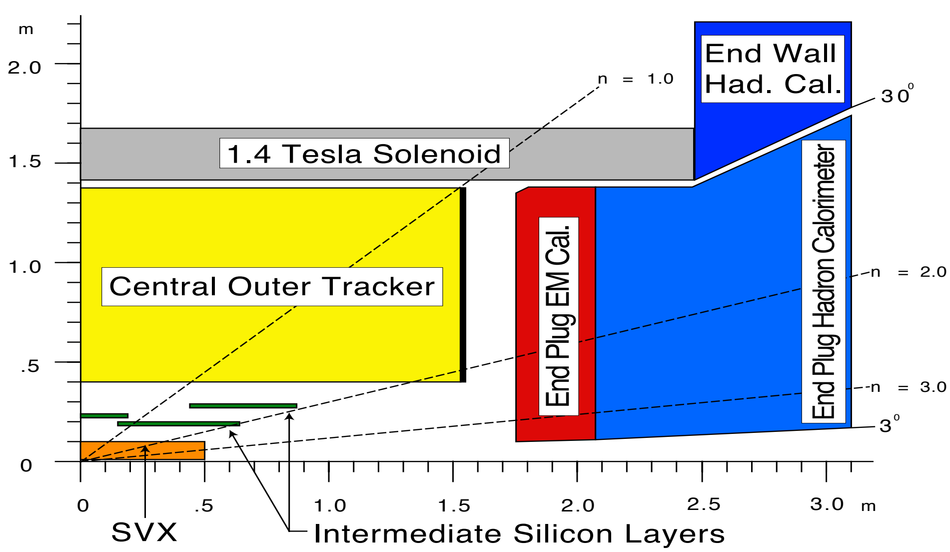


Figure 2.8: Charged-particle tracking systems at CDF.

Silicon Detectors (L00, SVX, ISL)

Located directly outside of the Tevatron beampipe are three concentric silicon microstrip detectors which enable high-precision reconstruction of charged-particle trajectories[35, 36, 37]. The excellent resolution of the silicon detectors is essential in identifying the decay of long-lived particles, such as the B mesons, that can travel several dozen millimeters away from the primary interaction point before decaying. The impact parameter resolution of tracks with silicon information is $\sigma_{d_0} \approx 60\times10^{-6}$m. Figure 2.9 shows the coverage provided by the three silicon detectors.

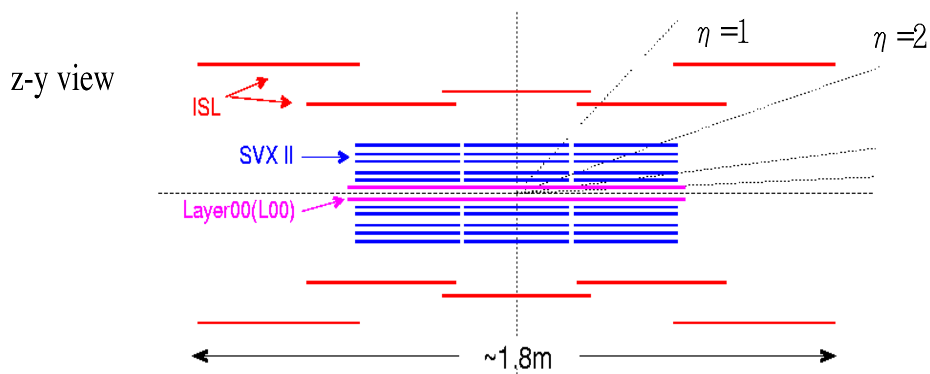


Figure 2.9: Close-up view of the coverage of the three silicon detectors.

Semiconductor materials, such as silicon, are characterized by the small gap in energy between their valence band and conduction band. Dopants are added to the silicon to introduce new energy levels that move the Fermi level of the material close to the valence band (for p-type doping) or conduction band (for n-type doping). When n-type and p-type silicon are brought into contact the majority carriers of both types diffuse across the boundary setting up a region that is depleted in major charge carriers and subject to an electric field that prevents any more diffusion across the junction. Applying a reverse bias voltage, as shown in Figure 2.10, to the silicon extends this depleted region throughout the n-type material (chosen due to the lower mobility of holes requiring more voltage to deplete). Charged-particles passing through the depleted region will produce electron-hole pairs that are pulled apart by the bias voltage, and then free to drift without recombining due to the depletion of charge carriers. Silicon detectors at CDF can also have n^+-type junctions on the opposite side of the n-bulk from the p-type junction, allowing for two different measurements to be made of the same ionizing particles.

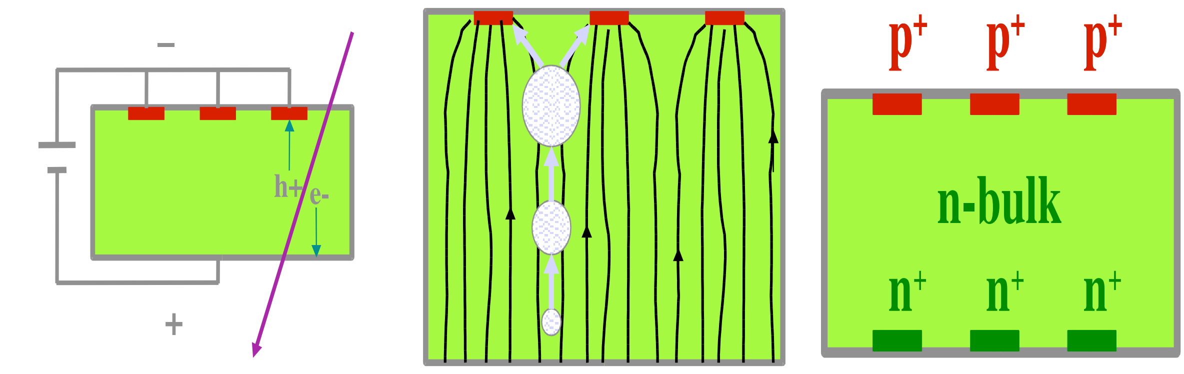


Figure 2.10: Basics of silicon detection. Left: Reverse bias voltage is applied across the $p-n$ junction to deplete the material of free charge carriers. Middle: Passing charged particles produce electron-hole pairs, which drift along electric field to collection planes. Right: Double-sided silicon detectors allow for independent $r-\phi$ and $r-z$ measurements.

Collecting the charges deposited by the ionizing particle and noting their arrival time can give a measurement of the particles charge and position. Electrons and holes created in the CDF silicon detectors typically drift for less than 30 ns before being collected by readout electronics, so the silicon detectors are able to resolve consecutive bunch crossings easily. The combined silicon detector provides 722,432 channels of electric readout, and provides coverage over an area of 6 m^2. Figure 2.11 shows a transverse view of the combined CDF silicon system, while Figures 2.12 and 2.13 respectively show drawings and photographs of the individual detectors. Information from the SVX detector is used to rapidly select events during data taking, while information from L00 and the ISL is added later to "refit" the tracks and improve their resolution.

The innermost of the silicon detectors, located directly on the beampipe, is a single-

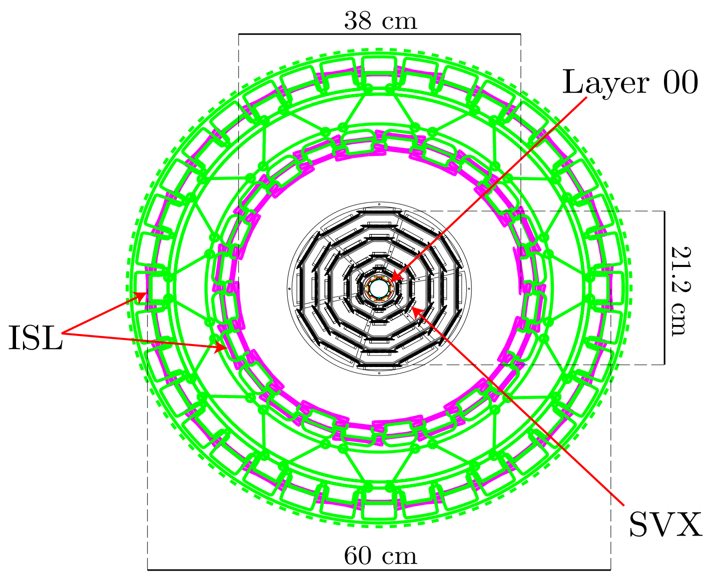


Figure 2.11: Transverse view of the three CDF silicon detectors.

sided single-layer radiation-hard silicon detector known as Layer 00 (L00) which allows for tracking in the $r - \phi$ plane. L00 extends from a radius of 1.15 cm out to 2.1 cm. The single layer of L00 is arranged in the form of two overlapping hexagons of slightly different size (one at r=1.25 cm, and one at r=1.5 cm), as shown in Figure 2.12. This detector was built to improve the resolution of the impact parameter, σ_{d0}, of tracks with hits in the silicon detectors.

Outside of L00 is the primary silicon detector system, known as the Silicon Vertex Detector (SVX). This detector has five double-sided layers allowing for tracking reconstruction in the $r - \phi$ and $r - z$ planes. It covers the radius from 2.5 cm out to 10.6 cm, and is 90 cm in length. The SVX is composed of 6 identical barrels which are each segmented into 30° wedges, giving the detector a 12-fold symmetry in ϕ. All five layers have one side that provides $r - \phi$ tracking information, while the other side provides 90° degree stereo information for three of the layers (L0, L1, and L3), and ±1.2° small-angle stereo information for the remaining two layers (L2 and L4).

The outermost silicon detector is the Intermediate Silicon Layer (ISL). It has one central layer ($|\eta| \leq 1.0$) that is 1.0 m in length that helps link tracks from the COT to the SVX. The ISL also has two forward layers ($1.0 \leq |\eta| \leq 2.0$) that allow for silicon tracking in the forward regions where the COT (described below) does not provide full coverage. Similar to L2 and L4 of the SVX, the ISL layers are double-sided with one side providing measurements in $r - \phi$ and the other providing 1.2° small-angle stereo information.

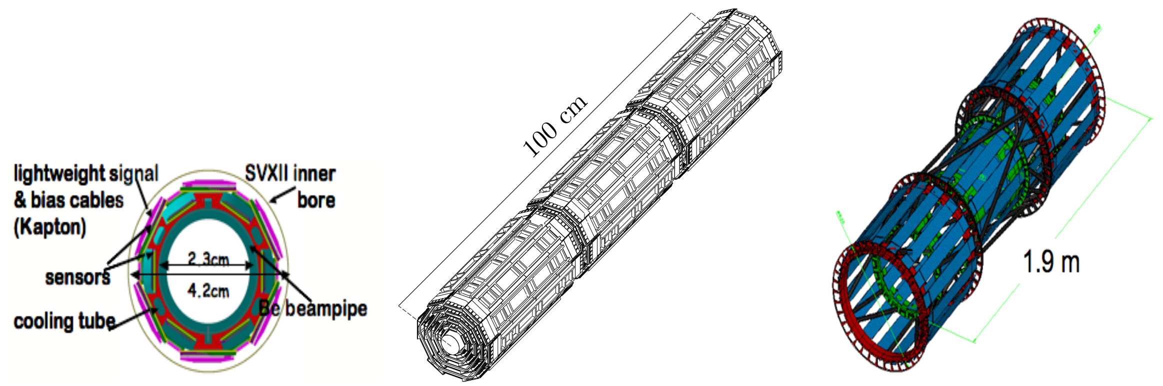


Figure 2.12: Schematics of the three CDF silicon detectors. Left: Transverse view of L00. Middle: Three of the six barrels of the SVX. Right: The ISL detector.

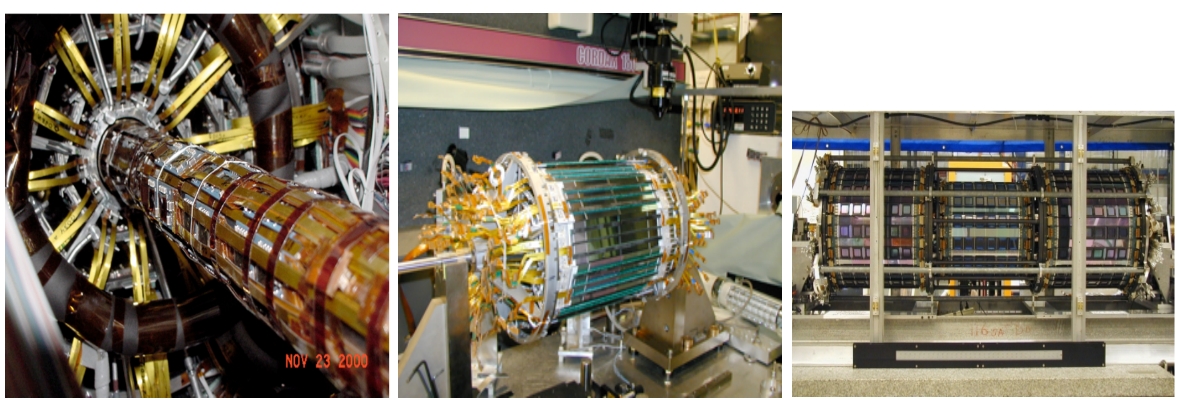


Figure 2.13: Photographs of the three CDF silicon detectors. Left: L00 being inserted into the SVX. Middle: One barrel of the SVX under construction. Right: The completed ISL waiting for installation.

Central Outer Tracker

The other charged particle tracking system at CDF, located directly outside of the ISL, is a 310 cm long cylindrical proportional drift-chamber known as the Central Outer Tracker (COT)[38]. A single proportional chamber could be envisioned as a wire of radius

r_1 kept at positive potential (anode), surrounded by a metal tube of radius r_2 kept at a negative potential (cathode), with the metal tube filled with gas. The electric field between these will point radially from the outer tube to the inner wire, increasing in magnitude as the inner wire is approached. A charged particle traversing the inner region will ionize the gas, creating an electron positive-ion pair at radius r_i. The liberated electron will accelerate towards the anode, gaining energy as it goes. If it gains enough energy to be above the ionization threshold of the gas, it will ionize more gas and the process will repeat until an *avalanche* of electrons and positive ions is created. A large number of electrons will arrive at the anode producing a negative pulse which can be used to estimate the distance of the initial ionization from the anode if the drift velocities of the gas and electric/magnetic field configurations of the chamber are known. Typically the avalanche consists of 10^5 electrons per initial ionized electron, independent of the the initial number of primary ions, which is where the name "proportional" drift-chamber originates.

The COT is a much more elaborate form of the simple drift-chamber described above. The COT provides radial coverage from r =43 cm to r=133 cm, full coverage out to $|\eta|$ =1.0, and partial coverage out to $|\eta|$ =2.0. The COT is filled with an approximately $500:500:173$ mixture of Argon, Ethane, and isopropyl alcohol gases, which are ionized by charged particles traversing the chamber. This gas mixture is chosen, among other reasons, for its high drift velocity ($\approx$50 μm/ns) which allows for consecutive bunch crossing to be measured with minimal overlap.

The COT's inner structure is composed of 2520 *cells*, which are each the same length as the COT but cover only a small fraction of the total chamber volume. Each cell consists of two cathodes made out of sheets of 350 angstrom thick gold on 6.4 μm thick Mylar, and 12 anodes (also called Sense wires) in the form of gold-plated tungsten wires 40 μm in diameter. The cathodes are kept at ground potential, while the anode wires are kept at a positive voltage of 3 kV, thus ensuring that any ionized electrons from the gas will be directed towards the anode, and ions are directed towards the cathode.

13 Potential wires and 4 Shaper wires, made out of the same gold-plated tungsten wires, are also present in each cell to allow for the field shaping to make drift velocities uniform. The Potential wires are kept at a voltage of 2 kV. Figure 2.15 depicts the wire configuration and drift trajectories found inside a COT cell.

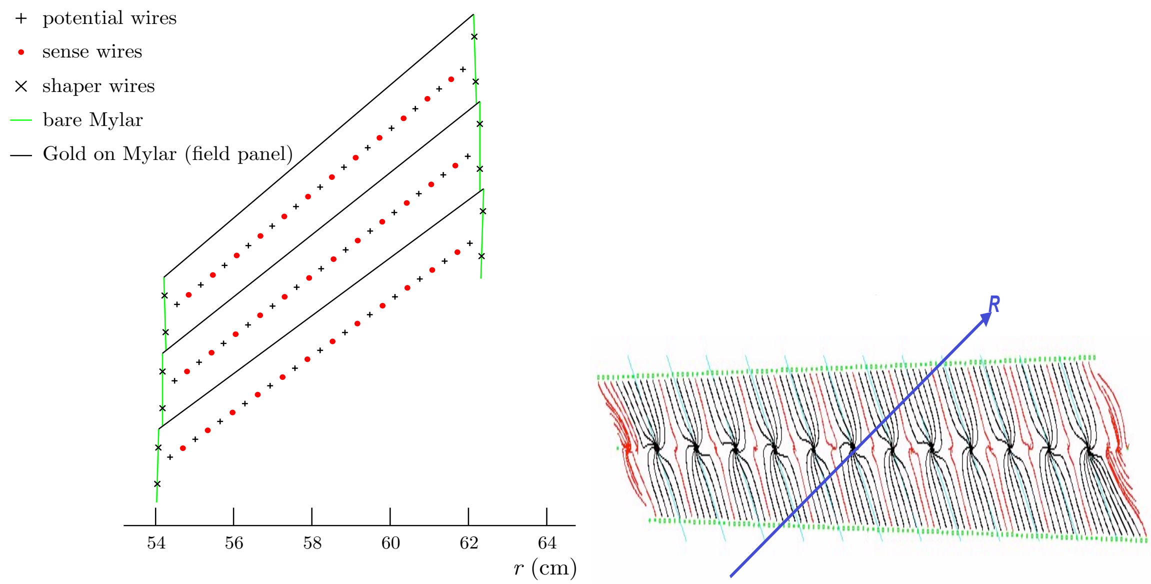


Figure 2.14: COT cells and drift trajectories.

The cells are tilted 35° from the radial direction to allow for a Lorentz force effect that ensures the ionized particles will drift in the azimuthal direction. The cells are arranged in eight "superlayers", shown in Figure 2.15, with each superlayer containing one cell in the radial direction. Four of these superlayers are axial allowing for $R-\phi$ measurements, and four of them are $\pm 3°$ stereo layers allowing for $R-z$ measurements. Table 2.2 summarizes the COT superlayer parameters.

SL	Stereo Angle (°)	# cells	<r> (cm)
1	+3	168	46.774
2	0	192	58.534
3	-3	240	70.295
4	0	288	82.055
5	+3	336	93.815
6	0	384	105.575
7	-3	432	117.335
8	0	480	129.096

Table 2.2: COT Superlayer Parameters.

2.2.3 Calorimetry

If tracking is the art of measuring a particle's trajectory without altering it, then the opposite can be found in calorimetry, which is the art of destroying the particle in order to measure it. Particles exiting the tracking region of CDF will encounter a series of sampling calorimeters, where they will gradually be absorbed through interactions with the dense (*e.g.*- iron or lead) calorimeter material[39, 40, 41]. The interactions will instigate electromagnetic or hadronic showers (depending on the particle's identity) which are sampled by layers of scintillating material interleaved with the dense calorimeter matter. Figure 2.16 depicts the basics of calorimeter design.

The combined calorimeter system at CDF covers the full 2π in azimuth and pseudo-rapidity $|\eta|$ <3.6. Each calorimeter system at CDF is segmented into $\eta - \phi$ *towers* that are arranged in a projective geometry to point back to the center of the detector. Each tower consists of alternating layers of scintillator, which sends signals via light guide to photomultiplier tubes (PMTs) at the top of the tower, and absorber. The PMT signal is integrated to measure the energy deposited in the tower in each collision. Figure 2.17 represents a calorimeter wedge, which is all towers for a given η section, at CDF.

To effectively measure the energies of all the varieties of particles created in collisions CDF has two types of calorimeters, known as the *Electromagnetic* (EM) and *Hadronic* calorimeters (HAD). Both have the same basic design principle, which is that of a series

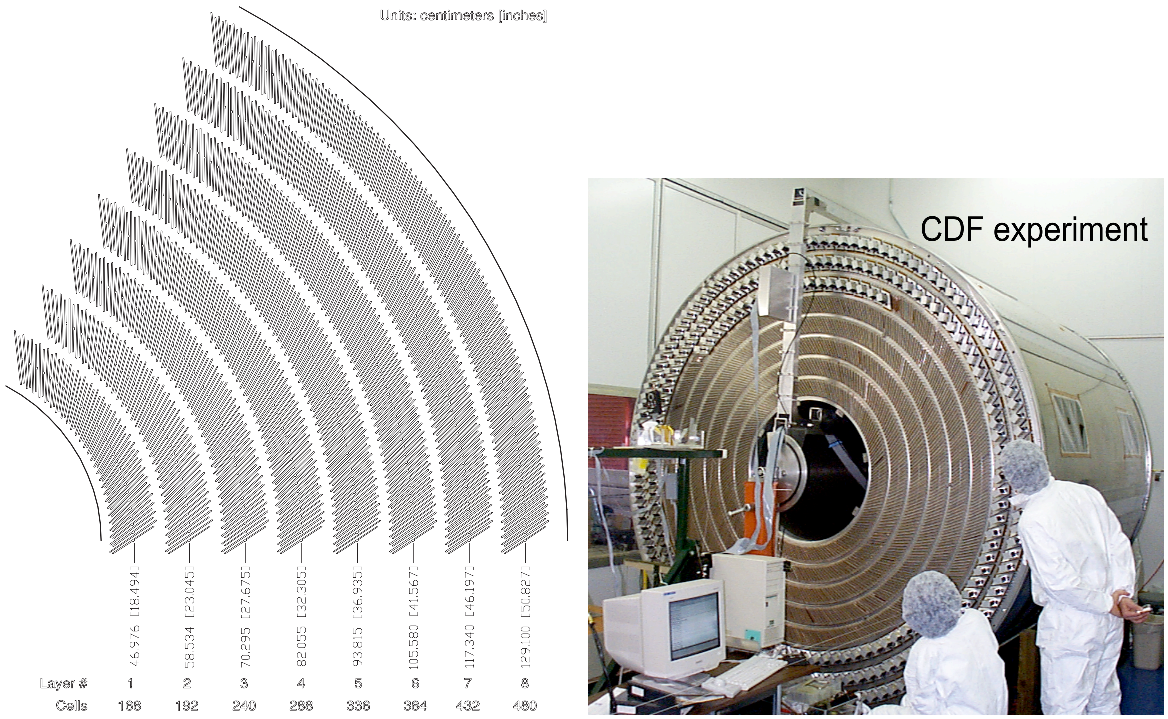


Figure 2.15: Left: COT endplate showing superlayer configuartion. Right: Stringing of the 6^{th} superlayer of the COT chamber.

of interleaved absorbing regions and detecting regions, but they differ in the absorber material which is typically a metal of large atomic number chosen for its ability to capture certain types of particles.

The electromagnetic calorimeter samples the electromagnetic showers of high-energy electrons and photons, which respectively are dominated by bremsstrahlung and e^+e^- pair production. The basic occurrence inside of the electromagnetic showers is an electron radiating a photon, which converts into an electron-positron pair, which both radiate photons, developing a cascade. This process repeats itself until the particles no longer have enough energy to continue the radiation. The shower is characterized by a particle multiplicity that increases logarithmically until a maximum (known as “shower max”) is reached, at which point the lateral extent of the shower is also reached. X_0, known as the radiation length in the EM calorimeters, is the mean distance for an electron to lose

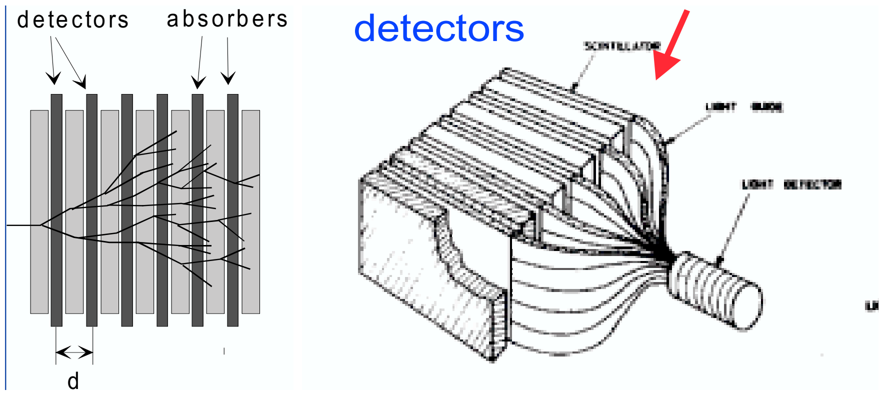


Figure 2.16: Left: Schematic of a typical calorimeter stack, with alternating layers of scintillator and absorber. Right: Wavelength shifting fibers carry the scintillator signals to phototubes that collect them.

all but $1/e$ of its energy (by bremsstrahlung). For the lead that composes the CDF EM calorimeters, the value of X_0 is 0.56 cm, and the calorimeters are many multiples of this length.

Hadronic showers are much more complex than electromagnetic showers due to the much wider range of contributing processes. Inelastic collisions of hadrons with nucleii leads to fission of the nucleus, producing nucleii fragments and many secondary particles, which undergo further inelastic collisions and produce more fragments and secondaries, developing a cascade. Much of the energy of a hadronic shower is lost due to the breakup of nuclei, nuclear excitation, and escaping neutrons. Because there are large fluctuations in the lost energy, the resolution of hadronic energy measurements is much poorer than that of electromagnetic energies. Hadronic showers are much longer and wider than EM showers. λ_I, known as the nuclear absorption length, is the mean free path of a particle before undergoing an inelastic nuclear collision. For the iron that composes much of the CDF hadronic calorimeters, the value of λ_I is 16.7 cm, so the calorimeters only contain a few multiples of this length. Table 2.3 summarizes the various electromagnetic and hadronic calorimeters at CDF, and their design parameters.

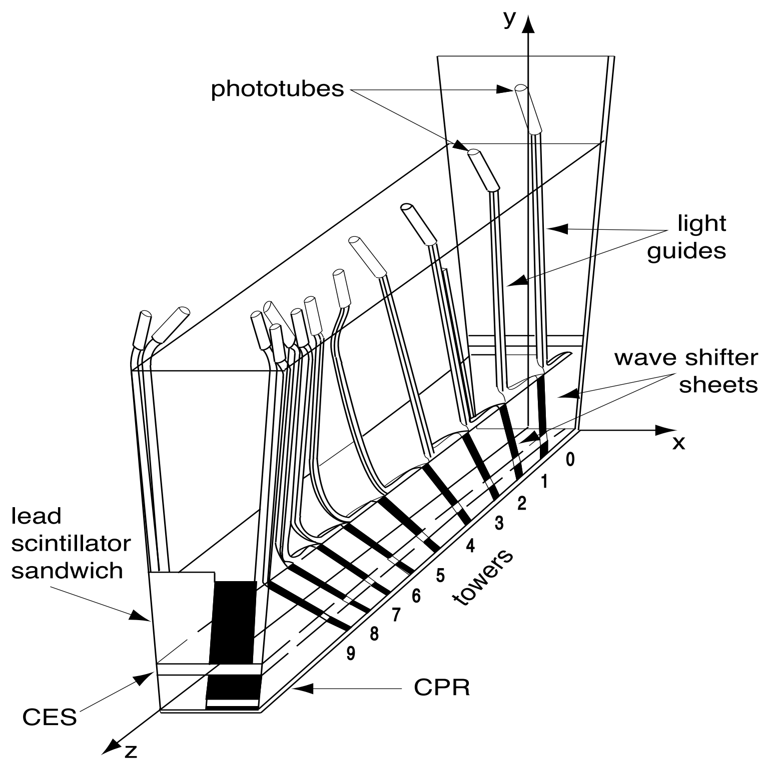


Figure 2.17: Calorimeter wedge at CDF. Light from various scintillator layers throughout the calorimeter is collected by phototubes located on the outside of the detector.

System	$\|\eta\|$ Coverage	Thickness	Energy Resolution
Electromagnetic			
CEM	0.0→1.1	$19X_0$	$13.7\%/\sqrt{E_T}\oplus 2\%$
PEM	1.1→3.6	$21X_0$	$16\%/\sqrt{E_T}\oplus 1\%$
Hadronic			
CHA	0.0→0.9	$4.5\lambda_I$	$50\%/\sqrt{E_T}\oplus 2\%$
WHA	0.7→1.2	$4.5\lambda_I$	$75\%/\sqrt{E_T}\oplus 4\%$
PHA	1.2→3.6	$7\lambda_I$	$80\%/\sqrt{E_T}\oplus 5\%$

Table 2.3: Summary of the CDF calorimeters. Calorimeter energy resolution is written as σ/E = A/$\sqrt{E}\oplus$ B, where B is an energy independent term controlled by monitoring relative light yields among scintillator tiles.

Central Calorimeters

The Central Electromagnetic Calorimeter (CEM) covers the pseudorapidity range out to $|\eta|$ <1.1. The towers of this calorimeter are composed of alternating layers of lead absorber and polystyrene scintillator. The Central Hadronic Calorimeter (CHA), which sits right outside of the CEM, covers the pseudorapidity range out to $|\eta|$ <0.9 and is composed of alternating layers of iron absorber and scintillator. Both the CEM and CHA are segmented into towers 15° in ϕ and 0.1 in η.

Interspersed with the CEM are two proportional chambers that are used to monitor the profile of electromagnetic showers. The central pre-radiator (CPR) is a proportional chamber that sits directly outside of the solenoid and before the first layer of the CEM. The shower-maximum detector (CES) is a proportional chamber that sits $6X_0$ inside of the CEM, where CDF electromagnetic showers reach their maximum lateral extent. The CPR and CES can be used to improve the efficiency of electron and photon identification and reduce any background signals.

End-Wall and Plug Calorimeters

Energy information for particles produced in the forward regions of the detector is provided by a similar arrangement of electromagnetic and hadronic calorimeters. The plug electromagnetic (PEM) and plug hadronic (PHA) cover pseudorapidity ranges out to $|\eta| <3.6$, while the end-wall hadronic calorimeter (WHA) covers the η range $0.7< |\eta| <1.3$. The WHA provides some overlap coverage where there are gaps in the CHA and PHA coverage. As in the central region, there is a proportional chamber in the plug region known as the plug shower-maximum detector (PES) that provides further electron and photon identification.

2.2.4 Muon Detection

Calorimetry fails to slow down muons, which are two hundred times heavier than the electron. Muons feel the same magnitude of force as electrons in the calorimeter, but their mass causes them to be much less affected. Muons will penetrate through the CDF calorimeters without losing much of their energy. To help identify muons, CDF has several systems located in the outermost regions of the detector[42]. By matching charged particle tracks in the central tracking chambers to track "stubs" (track segments formed in the muon chambers) in the muon systems, muon candidates are identified. The muon systems sit outside of a layer of steel shielding that absorbs any charged hadrons (pions, kaons, etc...) that are able to penetrate completely through the hadronic calorimeters and could fake a muon signal.

Due to spatial constraints inside the collision hall there is not complete $\eta - \phi$ coverage for muon detection, but the subsystems manage to cover much of the central region, as shown in Figure 2.18.

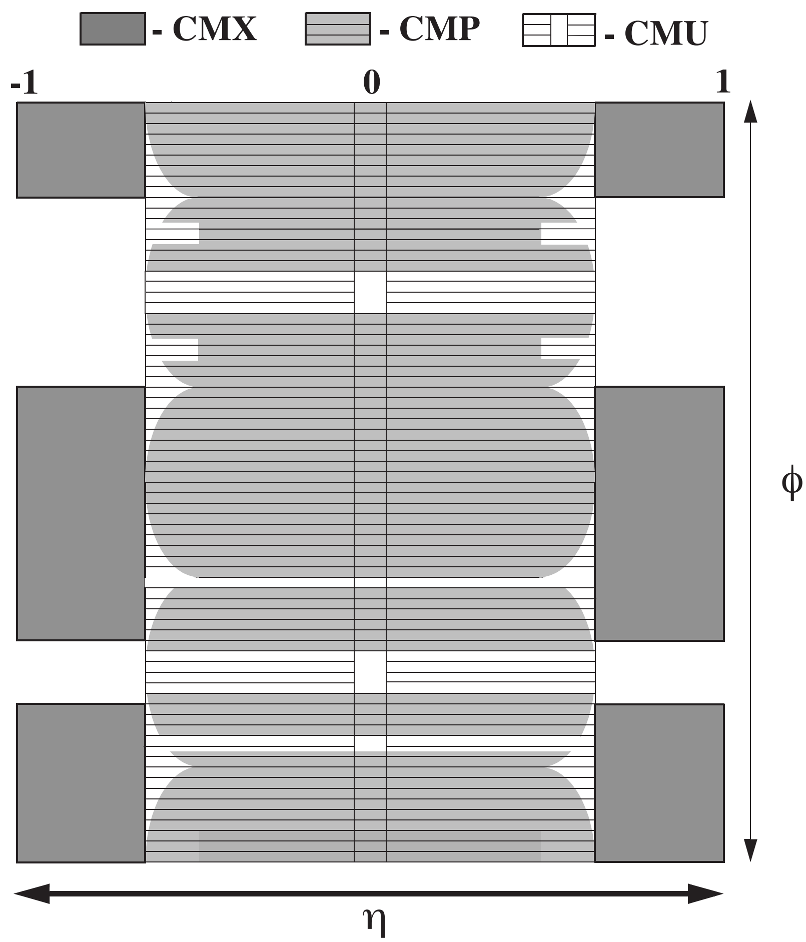


Figure 2.18: Muon $\eta - \phi$ Coverage provided by the CMU/CMP/CMX detectors.

The muon detectors are composed of proportional chambers filled with a 50%-50%

argon/ethane gas mixture (with <1% isopropyl alcohol). The central muon detector (CMU) consists of 2304 rectangular cells, arranged in 144 modules of 16 cells. The CMU provides a full 2π coverage in azimuth and eta coverage for $|\eta|$ <0.6. Figure 2.19 shows a typical CMU module.

Outside of the CMU, beyond 60 cm of additional steel shielding, is the central muon upgrade (CMP) detector. The CMP contains four offset layers of rectangular drift tube cells, as shown in Figure 2.19, and 1076 total cells. The CMP is covered by a layer of scintillator tiles known as the central scintillator upgrade (CSP). The information from the CSP can be used along with the CMU and CMP to improve muon identification.

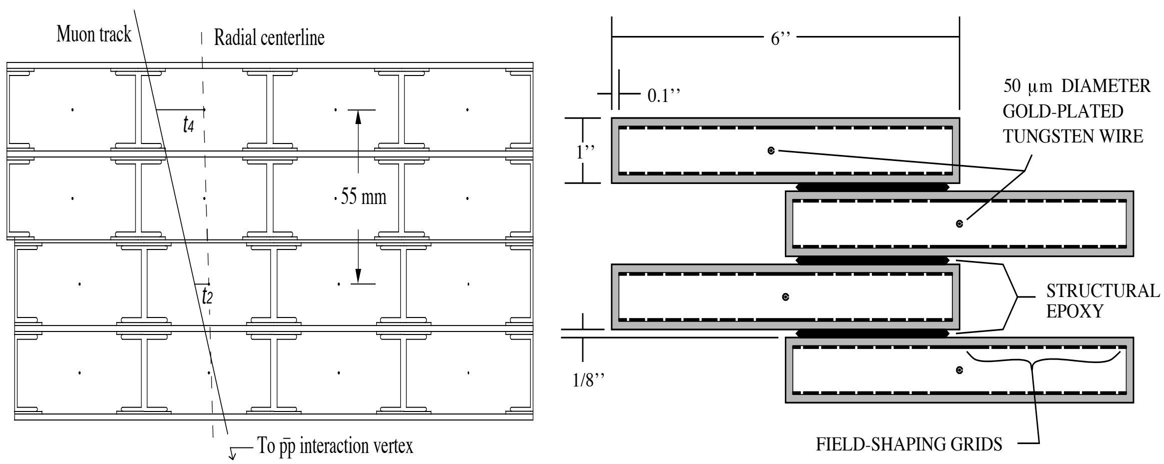


Figure 2.19: Left: CMU Module Right: CMP Module

The final muon system considered, known as the central muon extension (CMX), is arranged in a conical geometry and covers 0.6< $|\eta|$ <1.0. Figure 2.20 shows a picture of the CMX as well as the cells that compose it. The CMX is four cells thick at the small end of the cone geometry, and eight cells thick at the large end of the cone geometry, with a total of 2208 cells. Like the CMP, the CMX is covered by a layer of scintillator known as the central muon scintillator extension (CSX) that is used to improve muon identification.

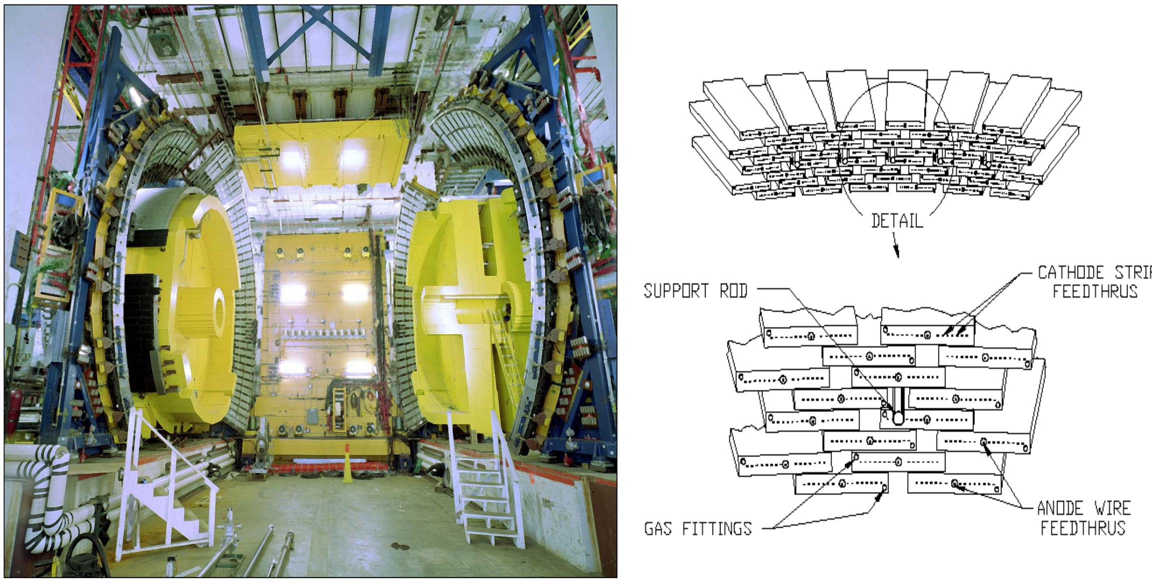


Figure 2.20: Left: CMX detector, with main CDF detector "rolled out" of collision hall. Right: CMX Cells.

2.2.5 Trigger and Data Acquisition

The total inelastic cross-section, which gives the rate for any type of interaction to occur, for proton-antiproton collisions at a center-of-mass energy of 1.96 TeV is ≈50 mb. [1] With this cross-section, and typical instantaneous luminosities of 1.0×10^{31} cm^{-2}s^{-1}, 500,000 proton-antiproton collisions can be expected at CDF every second. Each proton-antiproton collision that takes place in the CDF detector produces approximately 150 kB of recorded information that can be used to analyze the data, which translates into ≈75 GB of storage space needed every second if every collision were to be recorded!

A crucial requirement for a hadron collider experiment is the ability to collect and analyze quickly the information produced in a collision and decide whether or not the event is interesting enough to record to storage media. As discussed, with over two million bunch crossing per second at the CDF detector, each of which has the potential to produce collisions, it is impossible to record every collision to tape. The vast majority

[1] 1 barn = 10^{-24} cm^2

of collisions that take place at CDF are due to diffractive (glancing) collisions between proton and antiproton that are not relevant to the physics of $W^{\pm}+b\bar{b}$ production that is of interest. *Triggering* is used to monitor collisions for the presence of more interesting phenomena.

CDF employs a three-level "deadtimeless" triggering system to rapidly filter the generic proton-antiproton collisions down to the interesting events that will be analyzed, as shown in Figure 2.21. At each successive level of triggering, events are examined in greater detail, requiring longer processing times, and a decision is made whether to pass the event on to the next level or not. The overall rate for accepted events is reduced in this fashion until it is low enough to record the events to storage media, which occurs at $\approx$100 Hz.

In this analysis, the triggers used all are trying to identify events in which charged leptons with high ($\geq$18 GeV/c) transverse momentum are produced. Events passing all three levels of triggering due to either an electron in the CEM detector, a muon in the CMUP detector, or a muon in the CMX detector will be used in this analysis.

Level 1 trigger

The Level 1 trigger (L1) is a hardware trigger that works synchronously with the 396 ns master clock cycle of the Tevatron to quickly examine the information from every bunch crossing. There is a 14 clock cycle deep pipeline that stores the event information while the L1 decision is being made. After every clock cycle a new event enters the pipeline, and if an event has not received a L1 accept by the time it reaches the end of the pipeline it will be discarded. The time for a L1 accept decision to be made is 5.544 μs, which (along with the Tevatron clock cycle) determines the length of the pipeline.

The physics objects that are examined at L1 include charged-particle tracks in the COT, calorimeter activity, and muon chamber hits. L1 tracks are found by the eXtremely

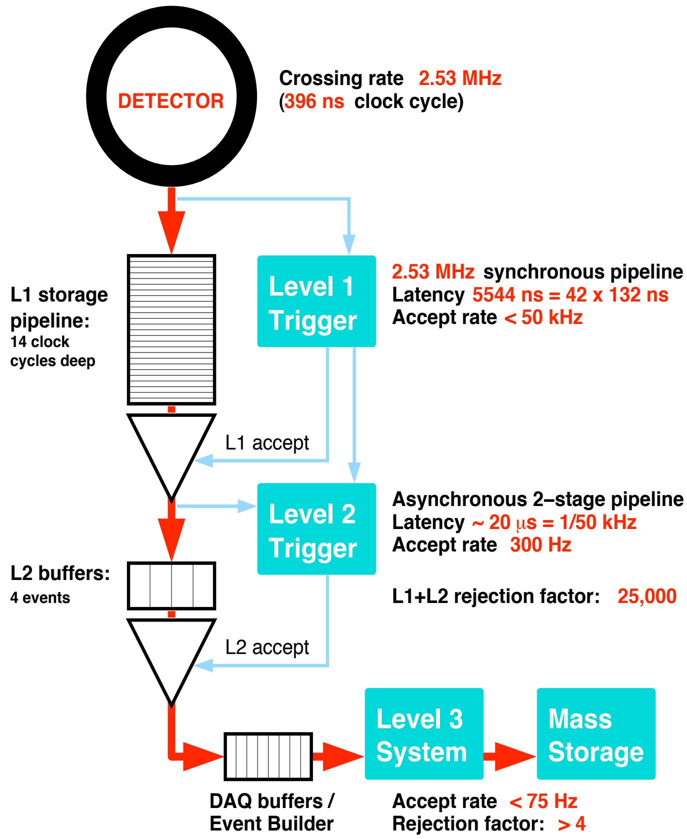


Figure 2.21: Run II Trigger System.

Fast Tracker (XFT), which forms track segments in the axial superlayers of the COT and tries to link them together by fitting a curve through them, as shown in Figure 2.22. XFT tracks used in this analysis must contain segments, each of which must be composed of 11 or more hits, in at least 3 axial superlayers.

For the L1 electron triggers used in this analysis, calorimeters are checked for the presence of at least one tower with ≥8 GeV of energy and a Had/EM energy ratio of ≤0.125. The XFT track must have $p_T \geq 8$ GeV/c and be matched to the trigger towers.

For the L1 muon triggers used in this analysis there must be an XFT track matched to a stub found in either the CMX or CMUP muon detectors. The extrapolation of the XFT track to the muon detectors takes into account the effect of multiple scattering. The

CMUP trigger requires a stub in the combined CMU and CMP detectors with $p_T \geq 6$ GeV/c matched to an XFT track with $p_T \geq 4$ GeV/c. The CMX trigger requires a stub in the CMX detector with $p_T \geq 6$ GeV/c matched to an XFT track with $p_T \geq 8$ GeV/c.

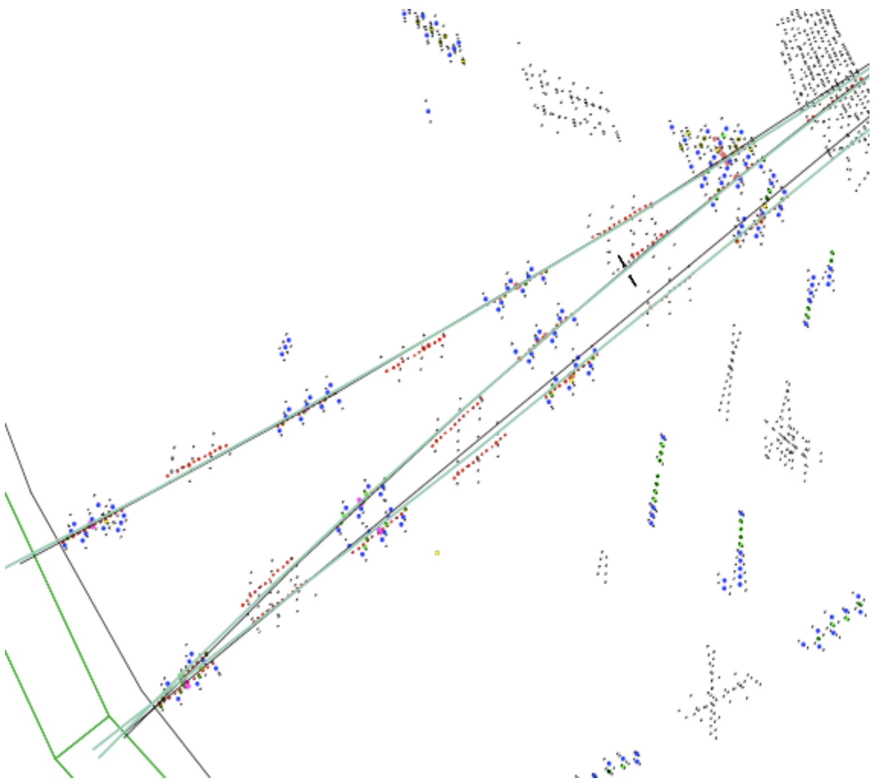


Figure 2.22: The XFT system finds track segments based on wire hits in the different COT superlayers, and then links these tracks together with a smooth curve.

Level 2 trigger

The Level 2 trigger (L2) is a combination hardware and software trigger that works asynchronously with the master clock cycle. Events accepted at L1 are placed into one of four L2 buffers, where they are stored while more sophisticated algorithms are used to examine the physics objects of interest. Events that are stored in the L2 buffers cannot be overwritten while they are being processed, so if any L1 accepts occur while all four buffers are full then the L1 event is lost and “deadtime” is said to occur. To minimize deadtime the L2 decision time is designed to be less than 80% of the average time between

L1 accepts, which amounts to a L2 latency of $\approx 20\mu$s.

For the electron triggers used in this analysis, the trigger calorimeter tower is checked to ensure that it has ≥ 16 GeV of electromagnetic energy, as well as a Had/EM energy ratio of ≤ 0.125. For the muon triggers used in this analysis, the thresholds on the p_T requirement for the XFT track are raised to 8 GeV/c for CMUP triggers, and CMX triggers are automatically accepted if they pass the L1 requirements.

Level 3 trigger

The Level 3 trigger (L3) is a software trigger that provides the final decision on whether or not to keep an event[43]. All information from events accepted at L2 are collected by hardware known as the "event builder", where they are placed in the final CDF data format, and then fed into one of several hundred PC computers where final selection criteria are applied. The output rate of the L3 trigger is 75 Hz, which is set by the ability to write the accepted event to storage media.

The L3 trigger requirement for the CEM, CMUP, and CMX triggers in this analysis requires that the p_T of the XFT track be ≥ 18 GeV/c.

CHAPTER 3

Event Selection and Data Samples

In this chapter the details of the analysis data sample are presented, and event selection criteria are explained.

3.1 Data Samples

This analysis uses all data taken in Run II through the fall of 2005, which are stored in datasets labeled `bhel0d`, `bhel0h`, `bhmu0d`, and `bhmu0h`. The data has been processed using the 5.3.3_nt version of the CDF reconstruction software. The CEM, CMUP, and CMX triggers are relied upon to select high-p_T electrons and muons. The integrated luminosity for this data sample is $695.5 \pm 41.7 \text{pb}^{-1}$ for the CEM and CMUP triggers, and $682.1 \pm 40.9 \text{pb}^{-1}$ for the CMX trigger.

3.2 Selection Details

In this analysis, events containing the simultaneous production of a $W^{\pm}$ boson and jets (which are the signature of quark and gluon production in a collision, as will be discussed later), at least one of which is identified as containing a heavy-flavor hadron decay, are isolated. This analysis will focus on the $W^{\pm}$ decays in the electron and muon channels, which represent 2/9 of all $W^{\pm}$ decays, since CDF has excellent capability to identify these charged leptons. Neutrino production will be indirectly identified by the presence of a large imbalance in transverse missing energy, since neutrinos escape the detector without interacting.

3.2.1 Charged Lepton Identification

Starting from the basic detector objects that cause a trigger to fire, tracks matched to EM clusters for electrons and tracks matched to stubs for muons, quality cuts are applied to ensure that an event has a valid lepton candidate present. In this section the definitions and values of the lepton quality cuts are provided. Unless otherwise noted, the selection criteria applies to both electrons and muons. The tracks responsible for firing the CEM, CMUP, and CMX triggers are "refitted" offline after including any information from the silicon detectors, possibly altering the kinematic quantities initially found for the track.

- Isolation - The ratio of the calorimeter energy in a cone of radius 0.4 ($\Delta\mathrm{R} = \sqrt{\Delta\phi^2 + \Delta\eta^2} \leq 0.4$) around the lepton track that is not associated with the track, divided by the energy that is associated with the track, as defined in Eq. 3.1. The isolation of the lepton candidate must be $\leq$0.1 to pass selection.

$$\mathrm{I}_{\mathrm{lepton}} = \frac{\sum E_T^{\Delta R \leq 0.4} - E_T^{\mathrm{lepton}}}{E_T^{\mathrm{lepton}}} \tag{3.1}$$

- E_T - Electrons must have an EM calorimeter cluster $\geq$20.0 GeV matched to the lepton track.

- p_T - Electrons (muons) must have a charged track with transverse momentum $\geq$10.0 (20.0) GeV/c. If the E_T of an electron is $\geq$100 GeV, then the p_T threshold is raised to 50 GeV/c to minimize high energy hadrons faking electrons

- E/p - For electrons with E_T $\leq$100 GeV, the ratio of the electron's energy to its momentum should be $<$2. This ensures that the matching track and EM energy cluster from the CEM trigger are roughly equal.

- $E_{\mathrm{HAD}}/E_{\mathrm{EM}}$ - The ratio of the hadronic to electromagnetic energies in the tower containing an electron candidate must be less than 0.055+0.00045·E to reduce the rate of jets faking electrons. Electrons should have a very high percentage of their energy contained in the EM calorimeter, though the linear energy term is included since higher-energy electrons deposit more energy in the hadronic calorimeter.

- Axial/Stereo Segments - The XFT track for both electrons and muons must contain at least 3 axial segments and at least 2 stereo segments in the COT, with each segment having at least 5 hits.

- -3.0 cm < $Q \cdot \Delta x$ < 1.5 cm - Check that the difference in x-coordinate between the extrapolated COT track and matching CES hits is small. The track extrapolation is charge dependent, necessitating different cut values.

- $|\Delta z|$ <3.0 cm - For electrons, the difference in z-coordinate between the extrapolated COT track and the matching CES hits should be small.

- $|z_0| < 60$ cm - The z-intercept of electron and muon tracks must be within 60 cm of the center of the detector to guarantee full coverage by the COT detector.

- E_{EM} < Max.(2.0, 2.0+0.0115·(p-100.0)) GeV - Check that muon candidates leave little energy in the electromagnetic calorimeters. The linear momentum dependent term is added due to the increased EM deposition of high momentum muons.

- E_{HAD} < Max.(6.0, 6.0+0.0115·(p-100.0)) GeV - Check that muon candidates leave little energy in the hadronic calorimeters. The linear momentum dependent term is added due to the increased HAD deposition of high momentum muons. The thicker hadronic calorimeters should contain more of the muon's energy than the EM calorimeters, hence the different cut compared to E_{EM}.

- $|d_0|$ and Muon Track Silicon Hits - The impact parameter of a track is the closest approach the track makes to the beam axis. Muon tracks with no hits in the silicon detectors are required to have impact parameters $|d_0|$ <0.2 cm, while tracks with silicon hits must pass $|d_0|$ <0.02 cm.
- ρ >140 cm - CMX muons are required to have tracks exiting the COT at a minimum radius of 140 cm, to ensure they can have enough segments to form a high quality track.

Muons are required to pass several additional requirements on the distances of the extrapolated COT track to the various CMU, CMP, and CMX muon chambers. Table 3.1 summarizes those so-called fiducial requirements, as well as all other requirements used in this analysis.

Leptons that pass all the requirements described above and in Table 3.1 are referred to as "tight" leptons. Leptons that pass all requirements except for the isolation requirement are referred to as "loose" leptons. This analysis will require that a single tight lepton, and no additional tight or loose leptons, is present in the events considered.

3.2.2 Jet Identification

Quarks and gluons produced in a $p\bar{p}$ collision will immediately *hadronize* into a shower of neutral and charged particles known as a *jet*. Jets are detected as a collimated grouping of charged particle tracks pointing to deposits of energy in both the electromagnetic and hadronic calorimeters. Jets in this analysis are identified using the JetClu algorithm, which groups energy deposits in the calorimeter that fall within a cone of radius 0.4 in the $\eta - \phi$ plane[44]. The algorithm works as follows:

1. Identify 'seed' towers in the calorimeter, which are defined as towers with energy $\geq$3.0 GeV

2. Draw a cone of radius $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the seed tower.

3. Calculate the E_T,η, and ϕ of the jet by applying the following definitions to the towers within the cone:

$$E_{T_{\text{Jet}}} = \sum_{i \in \text{cone}} E_{T_i}, \tag{3.2}$$

$$\eta_{\text{Jet}} = \frac{1}{E_{T_{\text{Jet}}}} \cdot \sum_{i \in \text{cone}} E_{T_i} \cdot \eta_i, \tag{3.3}$$

$$\phi_{\text{Jet}} = \frac{1}{E_{T_{\text{Jet}}}} \cdot \sum_{i \in \text{cone}} E_{T_i} \cdot \phi_i. \tag{3.4}$$

4. If the calculated η and ϕ of the jet do not coincide with the seed tower's η and ϕ, iterate the algorithm starting from the calculated jet η and ϕ until a stable cone center is found.

The energy of the jets obtained using the JetClu algorithm will in general not be equal to the energy of the hadrons that were absorbed by the calorimeters. Several levels of correction are applied to the raw jets from JetClu in order to make their energies more reflective of the true physics objects that gave rise to them[45].

The first correction applied, referred to as L1 or the "relative" correction, attempts to make the response of the CDF calorimeters uniform in eta. The raw energies reported by the CDF calorimeters will exhibit a dependence on eta. Conservation of momentum requires that the transverse energy of the two jets in a 2→2 (dijet) process should be equal. Dijet events that have one jet in the $0.2 \leq |\eta| \leq 0.6$ region and one jet outside of this region are used to derive a correction function. The energy of the jet outside of this central region is scaled until it equals the central jet, providing the η-dependent correction factor. The $0.2 \leq |\eta| \leq 0.6$ region does not include any cracks and is a well understood part of the calorimeter, which is why its energy measurements are relied upon.

The next level of correction applied, referred to as L4 or the "multiple interaction" correction, corrects for the presence of additional proton-antiproton interactions, aside from the one of interest, in a given bunch-crossing at CDF. Any deposited energy from these additional interactions can increase the energy of measured jets from the proton-antiproton collision of interest. A correction function, parameterized by the number of primary vertices observed in the event, is determined using minimum-bias data (*i.e.* - bunch crossings in the detector where no hard scatter collision was observed). By examining how much energy these minimum-bias events contribute to the calorimeters, an average correction function is constructed that takes into account the expected number of minimum-bias events at different luminosities. This function is used to calculate an average energy due to additional interactions, which is then subtracted from the calorimeter towers in an event.

The final correction used in this analysis, referred to as L5 or the "absolute" correction, accounts for any calorimeter response non-linearity or energy loss due to uninstrumented regions of the detector. Particles of a given energy might not be measured as having that energy by the calorimeter towers due to nonlinear response issues. Jet energies measured in MC are corrected to be equal in energy to the sum p_T of stable particles that fall within ΔR$\leq$0.4 of the jet direction. After applying the L5 correction, as well as the preceding levels, the energy of a jet should be equivalent to that of the stable particles that were present in the cone at the end of hadronization.

There are two additional levels of correction available at CDF that are not used in this analysis. The "underlying event" correction (L6) corrects for any energy in the calorimeters due to the spectator partons in the proton and antiproton that undergo a hard collision. The "out-of-cone" correction (L7) attempts to add energy to jets due to any energy that falls outside of the defined ΔR cone. Applying all corrections through

L7 should result in obtaining all the parton energies involved in the proton-antiproton collisions.

The L6 and L7 corrections are not used in this analysis because they introduce unwanted model dependencies and they suffer from large uncertainties. Figure 3.1 shows the systematic uncertainty on the measured jet energy after the different levels of correction, as a function of the jet's corrected transverse momentum. Lower momentum jets ($p_T \leq$50.0 GeV) suffer from much out-of-cone energy loss due to the magnetic field inside the COT bending low momentum particles out of the jet. Also, calorimeter response is worse for lower momentum particles, as can be understood from Table 2.3. The data considered in this analysis are dominated by low E_T jets, which is why the L6 and L7 corrections are not used.

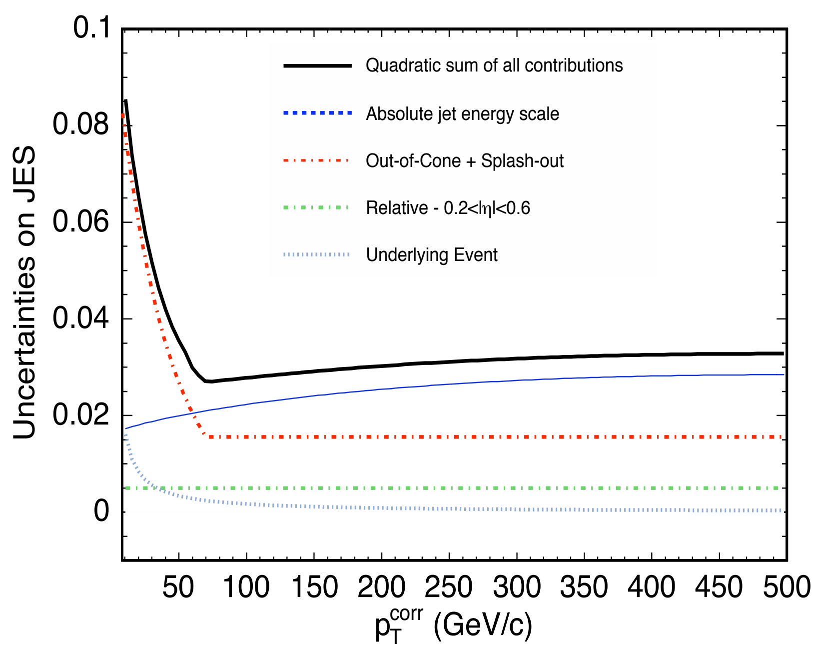


Figure 3.1: Relative Jet Energy Systematic uncertainty as a function of the corrected transverse momentum of jets.

Jets used in this analysis must have L5-corrected $E_T \geq$20.0 GeV, and must be in the central region of the detector ($|\eta_{\text{detector}}| \leq$2.0) where calorimeter response is well understood. Jets that pass this criteria are referred to as "tight" jets. One or two tight jets will be required for events to be considered in this analysis. Jets that aren't tight, but have E_T ≥12.0 GeV and $|\eta_{\text{detector}}|$ ≤2.4 are referred to as "loose" and are used to correct the missing transverse energy as described in the next section.

3.2.3 Missing Transverse Energy

Neutrinos produced in the leptonic decay of a $W^\pm$ boson will traverse the entire detector without interacting. Their presence is inferred based on any imbalance of transverse energy in collisions between protons and antiprotons, which initially have zero transverse energy. This imbalance of transverse energy is referred to as "missing energy" ($\not\!\!E_T$), and it is used to select event samples enriched in leptonic $W^\pm$ decays.

The initial values for the x andy y components of the $\not\!\!E_T$ are calculated as the negative of the vector sum of the raw transverse energies of the calorimeter towers.

$$\not\!\!E_{T_x}^{\text{raw}} = -\sum_{\text{towers}} E_T^{\text{towers}} \cdot \cos(\phi_{\text{tower}}) \tag{3.5}$$

$$\not\!\!E_{T_y}^{\text{raw}} = -\sum_{\text{towers}} E_T^{\text{towers}} \cdot \sin(\phi_{\text{tower}}) \tag{3.6}$$

Events containing muons must have additional corrections applied to the raw missing energy components. Muons are minimum-ionizing particles that leave only a small portion of their energy in the calorimeters, which would give the indication of a large amount of missing energy even if no neutrino was present in the collision. The raw missing energy components are corrected for the presence of muons by subtracting any calorimeter energy associated with the muon and adding the track momentum.

$$\not\!\!E_{T_x}^{\text{muon-corr}} = \not\!\!E_{T_x}^{\text{raw}} - p_T^{\text{muon}} \cdot \cos(\phi_{\text{muon}}) \tag{3.7}$$

$$\not\!\!E_{T_y}^{\text{muon-corr}} = \not\!\!E_{T_y}^{\text{raw}} - p_T^{\text{muon}} \cdot \sin(\phi_{\text{muon}}) \tag{3.8}$$

Finally, any corrections applied to the energy of tight and loose jets must be accounted for by replacing raw jet energies with corrected jet energies.

$$\not{E}_{T_x}^{\mathrm{L5-corr}} = \not{E}_{T_x}^{\mathrm{muon-corr}} + \sum_{\mathrm{jets}} E_{T_{\mathrm{raw}}}^{\mathrm{jet}} \cdot \cos(\phi_{\mathrm{jet}}) - \sum_{\mathrm{jets}} E_{T_{\mathrm{L5}}}^{\mathrm{jet}} \cdot \cos(\phi_{\mathrm{jet}}) \quad (3.9)$$

$$\not{E}_{T_y}^{\mathrm{L5-corr}} = \not{E}_{T_y}^{\mathrm{muon-corr}} + \sum_{\mathrm{jets}} E_{T_{\mathrm{raw}}}^{\mathrm{jet}} \cdot \sin(\phi_{\mathrm{jet}}) - \sum_{\mathrm{jets}} E_{T_{\mathrm{L5}}}^{\mathrm{jet}} \cdot \sin(\phi_{\mathrm{jet}}) \quad (3.10)$$

The final $\not{E}_T$ magnitude and azimuthal angle are calculated from the x and y components.

$$\not{E}_T = \sqrt{(\not{E}_{T_x}^{\mathrm{L5-corr}})^2 + (\not{E}_{T_y}^{\mathrm{L5-corr}})^2} \quad (3.11)$$

$$\phi_{\not{E}_T} = \tan^{-1}(\not{E}_{T_y}^{\mathrm{L5-corr}} / \not{E}_{T_x}^{\mathrm{L5-corr}}) \quad (3.12)$$

In this analysis, the L5 corrected $\not{E}_T$ must be $\geq$25.0 GeV to pass selection. This cut is chosen to enrich the sample in events where $W^{\pm}$ bosons have been produced, and to reject possible background sources that have small amounts of $\not{E}_T$ due to mismeasured jet energies. A complete description of background sources will be given later.

3.2.4 Event Quality Cuts

After selecting events that have at least one good lepton candidate and were taken during a store that had all detector components functioning properly (as indicated by the CDF "GoodRun" list), several further rounds of cuts are applied to ensure that any potential background processes are removed as much as possible from the selected data sample.

Approximately 5% of top-quark pair production events have final states containing 2 charged leptons, 2 neutrinos, and 2 b-quarks. If one of the charged leptons is lost or misidentified as a jet, then these "dilepton" events can end up passing all the selection criteria. A dilepton veto is applied that looks for events with additional lepton candidates, beyond the good isolated candidate that fires the trigger and passes all lepton selection cuts.

Events in which a Z boson was produced and decayed to two leptons can enter our signal region if one of the leptons is lost or fakes a jet. A Z-boson veto is applied by trying to pair the candidate lepton with a second object (in the form of a jet, a loose lepton, or a charged track) such that the invariant mass of the pair falls within a Z window of $76.0 \leq m_{pair} \leq 106.0$ GeV. If such a pair is found, the event is flagged as a possible Z event and removed from the sample.

High energy photons interacting with detector material can convert into an electron-positron pair. These photons are referred to as "conversions", and they are identified by a pair of oppositely charged tracks traveling in the same direction with an invariant mass very near zero. Events in which the trigger electron is identified as being part of a conversion pair are rejected.

Finally, events that could potentially be due to cosmic ray showers are identified and removed from the muon data sample. These events are characterized by hits in the muon chamber matching to tracks in the COT that occur at a time significantly different from the timing of the nearest bunch crossing.

For completeness, Table 3.1 lists all selection cuts used when running over MC and data.

3.3 CDF Secondary Vertex Tagging

Identifying heavy-flavor quark production is obviously of much importance in the study of $W^{\pm}$+$b\bar{b}$. If quarks are produced in a collision at CDF they will almost instantly form

Selection	Cut		
-	Electron	Muon	
		CMUP	CMX
Isolation	<0.1	<0.1	
Region	=0	-	
Fiducial	=1	-	
E_T	≥20.0 GeV	-	
p_T	≥10.0 GeV	≥20.0 GeV	
p_T && E_T	!(<50.0 && ≥100.0)	-	
Had/EM	≤(0.055 + 0.00045· E)	-	
Had	-	≤max(6.0,6.0+0.0280·(p-100.0))	
EM	-	≤max(2.0,2.0+0.0115·(p-100.0))	
E/P && E_T	!(≥2.0 && <100.0)	-	
Δx · Charge	≥-3.0 && ≤1.5	-	
$\lvert\Delta z\rvert$	≤3.0 cm	-	
χ^2	StripChi2 ≤10.0	TrkRedChi2CT ≤2.75 (run<190697), ≤2.3 (run≥190697)	
Track L_{shr}	≤0.2	-	
Track Axial Seg.	≥3	≥3	
Track Stereo Seg.	≥2	≥2	
$\lvert z_0\rvert$	<60.0 cm	<60.0 cm	
Conversion	false	-	
D0 && Track Si Hits	-	!(≥0.2 && =0) && !(≥0.02 && ≠0)	
CMU $\lvert\Delta x\rvert$		<3.0	
CMU Fiducial x		<0.0	
CMU Fiducial z		<0.0	
CMP $\lvert\Delta x\rvert$		<5.0	
CMP Fiducial x		<0.0	
CMP Fiducial z		<-3.0	
Bluebeam		false (run<15449)	-
CMX $\lvert\Delta x\rvert$			< 6.0
CMX Fiducial x			<0.0
CMX Fiducial z			<-3.0
CMX ρ			<140.0 cm
Arches			
Keystone			
Miniskirt			
Jet Correction Level	Level 5 (jetCorr05b)		
Tight (Loose) Jet E_T	≥20.0 (12.0) GeV		
Tight (Loose) Jet $\lvert\eta_{\mathrm{detector}}\rvert$	≤2.0 (2.4)		
GoodRun (w/wo Silicon)	true		
$\lvert z_0\rvert \leq 60$ cm	≤60.0 cm		
≥1 Lepton Candidate	true		
$\not\!\!E_T$	≥25.0 GeV		
Lepton Isol. ≤0.1	true		
DiLepton	false		
Z-veto	false		
$\lvert z_0 - \mathrm{lepton}z_0\rvert$	≤5.0 cm		
QCD veto	Not Used		
GoodRun w/ Silicon	true		
Monte Carlo Only			
HEPG Lepton p_T	≥20.0 GeV		
HEPG Lepton $\lvert\eta\rvert$	≤1.1		
HEPG Neutrino p_T	≥25.0 GeV		

Table 3.1: Summary of cuts used in this analysis to select events.

hadrons (except for the top quark, whose lifetime is so short that it decays before it can hadronize). Hadrons that contain b-quarks (*e.g.* - B^0, $B^{\pm}$, B_s^0 mesons and Λ_B baryons) can have significantly long lifetimes (several picoseconds). The long lifetimes of these B hadrons, combined with the significant boost they possess since they carry much of the initial b-quark's momentum, allow them to travel several millimeters away from a proton-antiproton collision before they decay.

To identify jets that potentially originate from heavy-flavor quark production, a search is performed for tracks within the jet that point back to a vertex that is displaced from the primary proton-antiproton interaction vertex. The impact parameter of charged-particle tracks is used to determine whether or not the track points to the primary vertex. Such a confluence of tracks would indicate that a heavy-flavor hadron was produced in a collision and traveled some distance before decaying. Identifying events with jets that are likely due to b-quark production would vastly improve the signal to background ratio in the data sample that is used to study $W^{\pm}$+$b\bar{b}$.

The algorithm used to identify displaced vertices is known as Secondary Vertex Tagging, or "SECVTX"[46]. The algorithm makes two passes at finding displaced tracks within a jet (*i.e.* - within ΔR$\leq$0.4 of an identified jet) and associating them to a vertex. If a jet is found to contain a good displaced vertex it is said to be "tagged". The details of the algorithm are provided here:

1. Pass 1 track selection. Make list of all tracks with following criteria:
    - pass 1 track cuts: p_T >0.5 GeV/c, impact parameter significance $(|\frac{d_0}{\sigma_{d_0}}|)$ >2.5
2. K^0 track removal from list of pass 1 tracks.
3. Sort tracks in list based on quality and impact parameter significance
4. Apply Pass 1 Vertex Finding
    - Pair tracks in the list, starting from highest ranked tracks, to form seed vertices.
    - Attach additional tracks to seed vertex if their impact parameter significance with respect to the seed vertex is smaller than 3.0.
    - If at least one additional track is attached to seed vertex, the vertex is retained and all contributing tracks are constrained to a common vertex.

- If a track contributes to χ^2 more than 50, discard track and repeat procedure until all tracks in vertex contribute less than 50 to χ^2.

5. Pass 1 Vertex Cuts (applied to Pass 1 vertex)
    - Number of pass 1 tracks in the vertex: $\geq$3
    - $c\tau$ upper cut to remove poorly reconstructed tracks and possible long-lived/material-interaction light-flavor hadron tracks.
    - Vertex decay length significance ($|\frac{L}{\sigma_L}|$) >3
6. Pass 2 track selection (if no pass 1 vertices are found)
    - pass 2 track cuts: p_T >1.0 GeV/c, impact parameter significance ($|\frac{d_0}{\sigma_{d_0}}|$) >3
7. Pass 2 Vertex Finding
    - No seed vertexing is performed. All tracks are constrained to a common vertex. As in Pass 1 vertexing, worst contributing tracks are discarded.
8. Pass 2 Vertex Cuts
    - Number of pass 2 tracks in the vertex: $\geq$2
    - $c\tau$ upper cut
    - Decay length significance ($|\frac{L}{\sigma_L}|$) >3

The first pass over the tracks allows lower quality tracks, but requires more of them to make a vertex. The second pass requires higher quality tracks, and allows less of them to form a vertex. The algorithm only looks for one vertex inside a jet, and does not attempt to find additional vertices in the jet after one has been found. Figure 3.2 depicts a transverse view of the SECVTX algorithm at CDF, and Figure 3.3 shows an event containing two SECVTX tagged jets. Both figures refer to the transverse decay length variable, L_{xy}, which is the transverse distance between the SECVTX vertex and the primary proton-antiproton vertex. If a vertex found by the SECVTX algorithm appears to originate on the wrong side of the primary vertex (*i.e.* - the jet direction points towards the primary vertex instead of away from it), then the jet is said to be "negatively" tagged. Negatively tagged jets are unphysical and are typically due to the resolution of the CDF tracking volume.

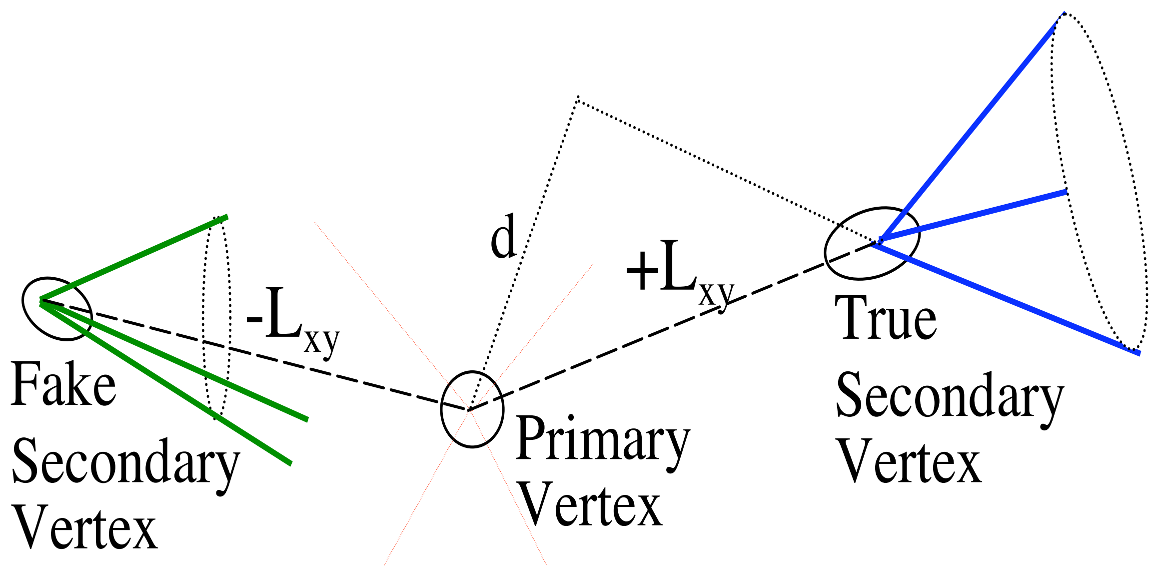


Figure 3.2: Transverse view of SECVTX tagging at CDF. Tracks inside of jets are clustered to form possible secondary vertices.

3.3.1 SECVTX Tagging Efficiency

MC samples are used to estimate the efficiency of the SECVTX algorithm to tag a jet that originates from a b-quark (*i.e.* - a "b-jet"). Figure 3.4 shows the b-jet tagging efficiency, as a function of the jet's transverse energy and also as a function of the jet's pseudorapidity, obtained using $t\bar{t}$ MC where SECVTX tagged reconstructed jets have been matched (via a $\Delta R \leq 0.4$ cone) to b-quarks from the matrix element. The average b-jet tagging efficiency determined in MC is multiplied by a "scale-factor" (SF) of 0.89 ± 0.07 that addresses differences in tagging efficiency found between dijet data and MC[47].

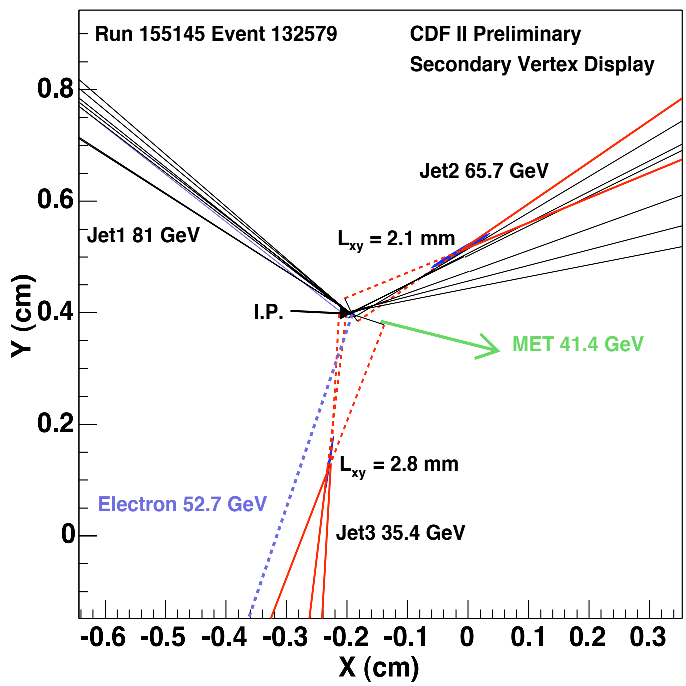


Figure 3.3: An example of a $W^{\pm}+b\bar{b}$ candidate event, containing three jets, two of which have been positively tagged by the SECVTX algorithm. This is a transverse view of the tracks in the event, and the tracks that are used to form the tags are highlighted in red.

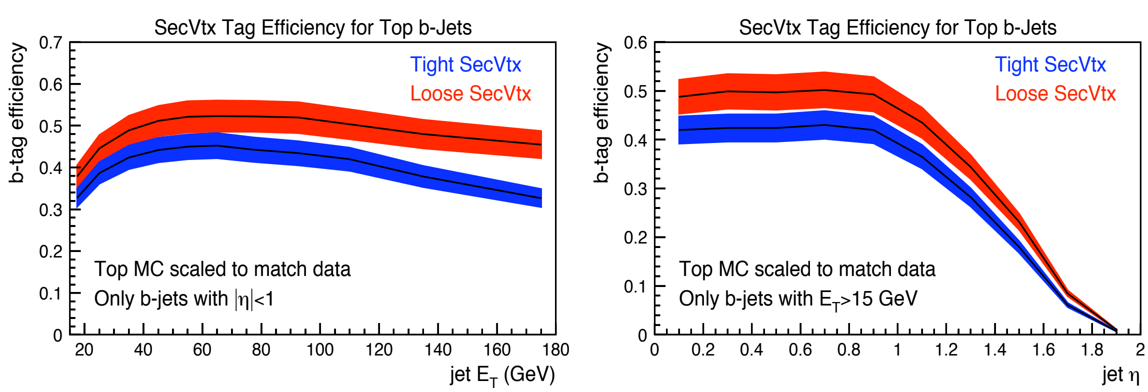


Figure 3.4: SECVTX Tagging Efficiency as a function of jet E_T (left) and η (right).

3.3.2 SECVTX Tagging Issues

Though the main motivation for the SECVTX algorithm is to identify jets containing the decay of heavy-flavor hadrons, sometimes the algorithm will tag jets that have no real heavy-flavor content. These tags of light-flavor jets are referred to as “mistags”. Mistags are predominantly caused by the limited resolution of the CDF tracking detectors, but also can be due to long-lived light particle decays (K_s and Λ), and interactions with the beampipe/detector material.

We want to know what portion of the positively tagged jets in our selected data sample are due to mistags. The rate of positive mistags is estimated using a sample of events containing negatively tagged jets. It is assumed that mistags due to detector resolution will have a symmetric pseudo-lifetime ($c\tau$) distribution, where positive (negative) mistags are assigned positive (negative) $c\tau$ values. [1] The $c\tau$ for negative and positive tagged jets from a sample of 50 GeV dijet data is fit with components for bottom, charm, and light flavors, as shown in Figure 3.5[48]. The results indicate that the rate of positive and negative light-flavor tags is not exactly equal as assumed, and a mistag asymmetry of 1.36±0.23 is measured.

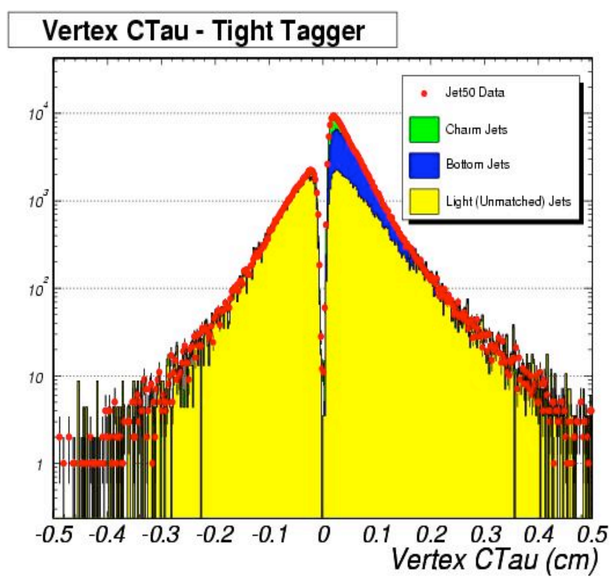


Figure 3.5: $c\tau$ distribution for Jet50 data containing negative and positive SECVTX tagged jets, fitted with templates for bottom, charm, and light-flavor.

A parameterization of negatively tagged jets is used to estimate the rate of positive

[1] $c\tau \equiv L_{xy} \times (\frac{M}{P_T})_{vtx}$, where M and P_T are the invariant mass and combined transverse momentum of the tracks in the vertex.

mistags as a function of several kinematic variables. The parameterization of negative tags is scaled by the positive/negative mistag asymmetry of 1.36±0.23. Figure 3.6 shows the mistag rate as a function of jet transverse energy and also as a function of pseudorapidity, using dijet data that is dominated by light-flavor jets. Though it can be seen that the mistag rate is small (≈2%) compared to the b-tag rate of Figure 3.4 (≈40%), the much larger cross-section for light-flavor production (compared to b-quark or c-quark production) makes this a non-negligible source of tags in our selected data sample.

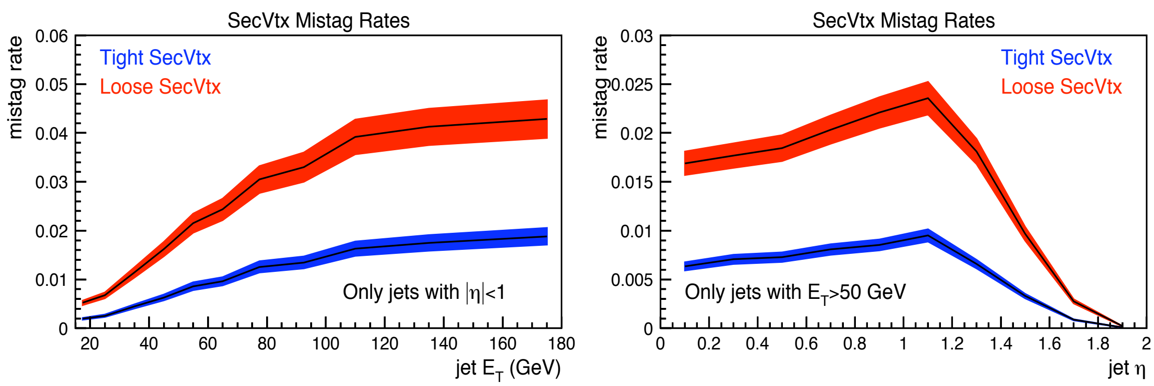


Figure 3.6: SECVTX Mistag rates as a function of jet E_T (left) and η (right).

3.4 Event Yields

All the selection criteria discussed in this chapter are applied to the combined **bhel0d**, **bhmu0d**, **bhel0h**, and **bhmu0h** data samples. Table 3.2 summarizes the number of pretag (*i.e.* - After all selection except SECVTX tagging), ≥1 SECVTX tag, and ≥2 SECVTX tag events after selection. In the 1 and 2 jet bins there are 1267 tagged jets which will be used in fits of SECVTX variables to determine the b-fraction of the sample.

Figure 3.7 shows plots for pretag events selected using the CEM trigger that pass all cuts and have one or two jets.

-	1 Jet	2 Jet
Pretag		
CEM	33306	5466
CMUP	16024	2592
CMX	8501	1282
$\geq$1 Tag		
CEM	440	239
CMUP	234	136
CMX	138	53
$\geq$2 Tags		
CEM	-	14
CMUP	-	12
CMX	-	1

Table 3.2: Number of pretag, $\geq$1 tag, and $\geq$2 tag events in the 1 and 2 jet bins of the `0d` and `0h` data samples, after all selection has been applied.

3.5 Monte Carlo Samples

The selected data sample will contain contributions from many other processes besides $W^{\pm}$+$b\bar{b}$. We would like to know the efficiency of our selection requirements at accepting $W^{\pm}$+$b\bar{b}$ and rejecting these other processes. Since there is no completely pure data sample of $W^{\pm}$+$b\bar{b}$ available to study, Monte Carlo techniques are commonly used to generate samples of events for $W^{\pm}$+$b\bar{b}$ and many processes of interest. These samples can be used to construct kinematic distributions and to check the efficiency of selection cuts.

The Alpgen matrix-element generator program is used in this analysis to generate MC samples for $W^{\pm}$+$b\bar{b}$[49]. This program uses the LO matrix-elements of processes like $W^{\pm}$+$b\bar{b}$ to generate a phase-space grid covering the range of possible kinematics of initial and final-state particles, from which event samples are generated. These event samples are fed into parton shower evolution programs such as Herwig and Pythia where the final-state matrix-element particles are hadronized and evolved into the stable particles that are observed by CDF[50, 51]. Table 3.3 lists the generation parameters used to make the Alpgen $W^{\pm}$+$b\bar{b}$ samples for this analysis. Finally, the event samples are run through a full GEANT simulation of the CDF detector to allow the MC to have the same information

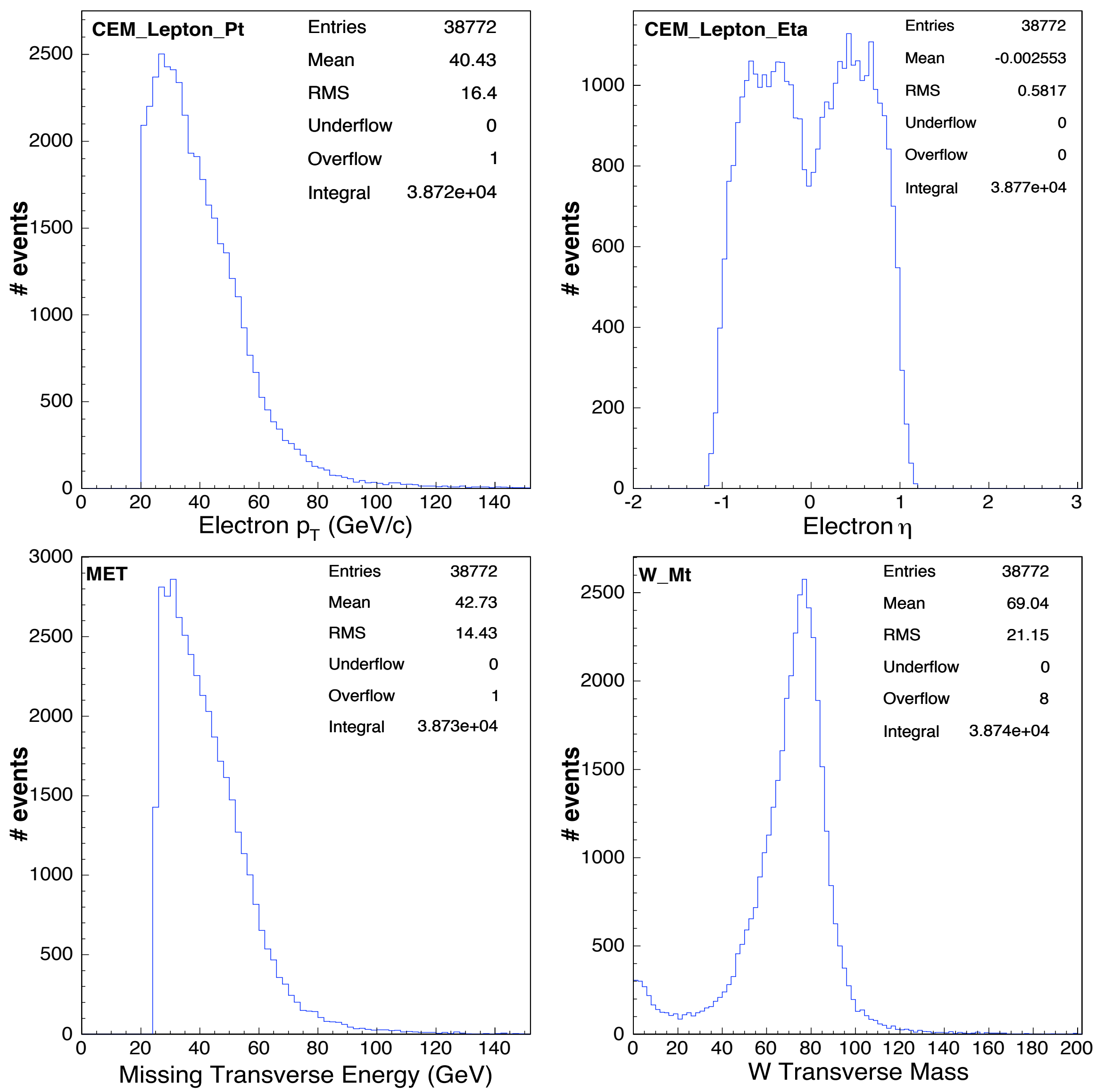


Figure 3.7: Selected plots from pretag CEM Events - Upper Left: Electron p_T. Upper Right: Electron pseudorapidity. Lower Left: Missing transverse energy. Lower Right: W transverse mass.

and format as the data[52].

Alpgen is able to produce LO event samples for $W^{\pm}+b\bar{b}$ plus up to two additional partons. In the rest of this document the samples will be referred to as $W^{\pm}+b\bar{b}$+0p, $W^{\pm}+b\bar{b}$+1p, and $W^{\pm}+b\bar{b}$+2p to indicate how many additional partons, beyond the $b\bar{b}$ pair, are present in the matrix-element. Each of the $W^{\pm}+b\bar{b}$+Np MC samples contributes to the overall $W^{\pm}+b\bar{b}$ phase-space, so it is desirable to use information from all of them

Parameter	Alpgen v2.1
Lepton: η_{max}	5.0
Lepton: p_T min.	1.0 GeV/c
ΔR(lepton,partons)	0.0
Neutrino: p_T min.	1.0 GeV/c
Light Partons: η_{max}	3.0
Light Partons: p_T min.	15.0 GeV/c
Light Partons: ΔR	0.4
Heavy Flavor: η_{max}	3.0
Heavy Flavor: p_T min.	8.0 GeV/c
Heavy Flavor: ΔR	0.0
Heavy Flavor: b-mass	4.7
Q^2	$M_W^2 + \sum m_T^2$

Table 3.3: Alpgen generator parameters used to make the $W^{\pm}$+$b\bar{b}$ samples for this analysis. The Q^2 value uses the m_T^2 $(= m^2 + p_T^2)$ values of all final-state partons.

in the analysis. The samples used in this analysis are listed in Table 3.4. The main samples used are those generated with the Q^2 value listed in Table 3.3, while the samples generated with other Q^2 values will be used to estimate systematic effects.

Alpgen allows the b-quarks in $W^{\pm}$+$b\bar{b}$ to be treated differently than any light-flavor quarks present in the matrix-element. The η and ϕ requirements applied to b-quarks are chosen to be looser than those applied to light-flavor partons to allow for the full range of b-quark kinematics to be covered by the matrix-element generator. This is because the parton shower programs are not as efficient at adding in heavy-flavor particles as the matrix-element generator. Conversely, the requirements applied to light-flavor partons by Alpgen are much tighter than those applied by Herwig or Pythia. The parton shower programs are more adept than Alpgen at adding additional low-momentum or collinear light-flavor radiation to the matrix-element calculation, so it is better to not rely too much on Alpgen to handle those regimes.

Finally, it is noted that the b-quarks in Alpgen can even be allowed to "fail" some set of requirements as long as other criteria are met. For instance, one of the b-quarks in $W^{\pm}$+$b\bar{b}$ is allowed to fail the p_T $\geq$8.0 GeV/c and $|\eta|$ $\leq$3.0 requirements as long as the

Process	Fileset	Alpgen Version	Generator $\sigma \times$ BR (pb) Before MLM	After MLM	Q^2	N_{gen}
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	mtop0x	2.1	3.438±0.102	3.168±0.105	$M_W^2 + \sum m_T^2$	371914
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p	mtop1x	2.1	1.697±0.017	1.035±0.023		376566
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p	mtop2x	2.1	0.678±0.014	0.387±0.017		376136
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	mtop3x	2.1	2.786±0.081	2.562±0.083	$4M_W^2$	377204
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p	mtop4x	2.1	1.380±0.015	0.839±0.018		370246
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p	mtop5x	2.1	0.569±0.012	0.329±0.017		376119
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	mtop6x	2.1	3.561±0.119	3.278±0.120	M_W^2	129283
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p	mtop7x	2.1	1.855±0.033	1.134±0.032		129292
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p	mtop8x	2.1	0.792±0.025	0.462±0.027		122144
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	mtop9x	2.1	4.648±0.158	4.319±0.165	$\frac{1}{4}M_W^2$	129164
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p	mtop10x	2.1	2.567±0.030	1.615±0.055		129290
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p	mtop11x	2.1	1.141±0.028	0.676±0.038		127780
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	mtop0y	2.1	3.443±0.097	3.170±0.103	$M_W^2 + \sum m_T^2$	376928
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p	mtop1y	2.1	1.697±0.017	1.035±0.019		376462
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p	mtop2y	2.1	0.678±0.014	0.387±0.018		374353
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	mtop3y	2.1	2.784±0.086	2.558±0.092	$4M_W^2$	108751
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p	mtop4y	2.1	1.381±0.016	0.840±0.017		113270
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p	mtop5y	2.1	0.569±0.012	0.330±0.015		127998
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	mtop6y	2.1	3.555±0.122	3.268±0.125	M_W^2	129034
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p	mtop7y	2.1	1.855±0.030	1.134±0.039		129286
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p	mtop8y	2.1	0.790±0.016	0.462±0.023		126927
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	mtop9y	2.1	4.645±0.151	4.315±0.157	$\frac{1}{4}M_W^2$	129313
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p	mtop10y	2.1	2.567±0.030	1.615±0.055		129290
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p	mtop11y	2.1	1.141±0.028	0.676±0.038		127407

Table 3.4: Alpgen samples used to calculate Acceptance numbers and also to study of the effect of Q^2 on $W^{\pm}$+$b\bar{b}$ cross-section. There is a noticeable trend for decreasing event cross-sections as Q^2 is increased that can be attributed to the diminishing value of α_s as Q^2 grows. The "Before MLM" and "After MLM" columns contain the generator-level cross-sections before and after a parton-jet matching scheme, referred to as "MLM matching", are applied. This will be discussed further in the next chapter.

other b-quark passes these cuts. In the case where both of the $b\bar{b}$ pair fail the η and ϕ requirements, the combined $b\bar{b}$ pair can be treated as a single object and checked to see if it passes the cuts as long as the pair fall within a ΔR cone of 0.4. The idea of this behavior is to allow for all possible kinematic configurations of $W^{\pm}$+$b\bar{b}$ to be present in the MC. Figure 3.8 shows the p_T and η behavior of the b-quarks from $W^{\pm}$+$b\bar{b}$+0p MC.

3.5.1 Monte Carlo Selection

A few selection requirements applied to MC samples are not applicable to data samples. The full "HEPG" information, including the identity and kinematics of all particles in an event, is available in MC samples. There will be additional requirements placed on

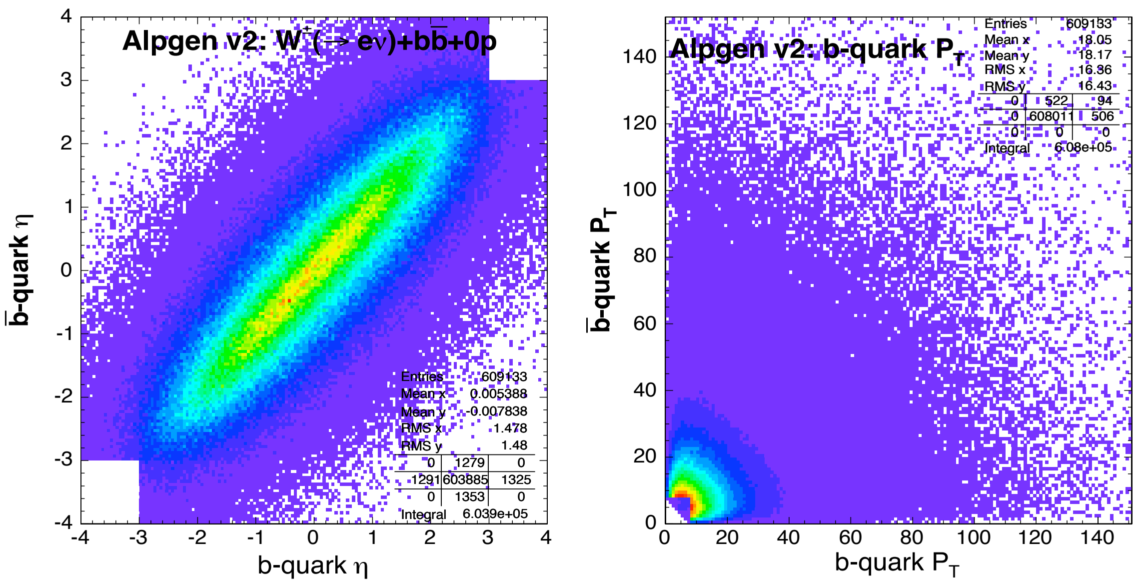


Figure 3.8: η and p_T of b-quark vs. that of the $\bar{b}$-quark in $W^{\pm}$+$b\bar{b}$+0p Alpgen v2.1 MC. One of the b-quarks in these events is allowed to fail the p_T $\geq$8.0 GeV and $|\eta|$ $\leq$3.0 generation cuts as long as the other passes, or both can fail as long as the b-quarks are within a ΔR cone of 0.4 of each other and the combined pair passes the cuts . These plots show that behavior.

the p_T and η of neutrinos and electrons/muons from $W^{\pm}$ decays. It will be made clear in the following chapters why this requirement helps remove some dependence on MC in the final result. Table 3.1 lists the values used when cutting on the HEPG information.

CHAPTER 4

Method

The goal of this analysis is to measure the b-jet cross-section for $W^{\pm}+b\bar{b}$ production in a manner that is minimally dependent on theoretical input. The information provided by the SECVTX algorithm will be used to isolate $W^{\pm}$+jet candidate events that contain real b-quark production. The amount of $W^{\pm}+b\bar{b}$ in this pool of events will be determined by accounting for all background sources and subtracting their contributions. The $W^{\pm}+b\bar{b}$ signal found in the selected data will be corrected for any possible inefficiencies associated with the selection process, giving a measured cross-section that is defined independent of CDF effects and that can be compared against theoretical predictions.

The invariant mass of all charged-particle tracks in SECVTX tags, referred to as the SECVTX mass, has been found to be an excellent discriminant between jet flavors. Jets originating from b-quark production will be shown to have a much larger, on average, SECVTX mass than jets originating from charm or light-flavor quark production. The SECVTX mass will be used, along with background subtraction, to isolate $W^{\pm}+b\bar{b}$ in a manner that is essentially independent of $W^{\pm}+b\bar{b}$ theoretical predictions.

4.1 Cross-Section Definition

The $W^{\pm}$+$b\bar{b}$ b-jet cross-section times branching ratio of $W^{\pm}$ bosons to leptons is generically written as:

$$\sigma_{W^{\pm}+b\bar{b}} \times \mathrm{BR}(W^{\pm} \rightarrow \ell^{\pm}\nu) = \frac{n_{W^{\pm}+b\bar{b}\ b-\mathrm{jets}}}{\mathcal{L}} \tag{4.1}$$

where $n_{W^{\pm}+b\bar{b}\ b-\mathrm{jets}}$ is the number of b-jets in the luminosity, $\mathcal{L}$, from $W^{\pm}$+$b\bar{b}$ events with the following properties:

1. 1 or 2 jets, where jets are defined as:
   - JetClu (cone 0.4) clusters of final-state hadrons.
   - Jet E_T $\geq$20.0 GeV
   - Jet $|\eta|$ $\leq$2.0
2. decay products of the $W^{\pm}$ passing:
   - $p_T(\ell^{\pm})\geq$20.0 GeV/c
   - $p_T(\nu)\geq$25.0 GeV/c
   - $|\eta(\ell^{\pm})|$ $\leq$1.1.

The cross-section is defined in a restricted phase space, defined by the $W^{\pm}$ decay product cuts and jet properties listed above. This definition is chosen in order to keep the final cross-section as close as possible in definition to what is actually measured and reduce model-dependent assumptions. Since the trigger leptons by definition have p_T $\geq$18 GeV/c and are in the central part of the detector, allowing the decay products of the $W^{\pm}$ boson to have unrestricted kinematics in the cross-section definition would require the use of MC to estimate what fraction of $W^{\pm}$+$b\bar{b}$ events have $W^{\pm}$ decay products outside of the measured regime.

The inclusion of hadron jets in the cross-section is also significant. The jets defined in the cross-section are clusters of stable hadrons. L5-corrected jets are supposed to be equal in energy, on average, to the cluster of stable hadrons that deposited in the calorimeter to form the jets. A check will have to be made to see how well L5 jets truly agree with the hadron-level jets, and corrections will have to be applied if there is any difference. By defining the cross-section in terms of hadron jets, the definition is truly independent of CDF-specific effects. Jets will be described as "calorimeter" or "hadron" depending on whether they come from, respectively, clusters of calorimeter towers or final-state hadrons in MC.

4.1.1 Cross-Section Details

Eq. 4.1 is written explicitly as:

$$\sigma_{W^\pm+b\bar{b}} \times \mathrm{BR}(W^\pm \to \ell^\pm \nu) = \frac{n_{b-\mathrm{jets}}(\mathrm{fit}) - n_{b-\mathrm{jets}}(\mathrm{back})}{\epsilon \cdot \mathcal{A} \cdot \mathcal{L}} \tag{4.2}$$

In Eq. 4.2, $n_{b-\mathrm{jets}}(\mathrm{fit})$ is the number of tagged calorimeter b-jets returned by a fit of the SECVTX mass of tagged jets in data events passing all selection and having 1 or 2 L5-corrected calorimeter jets. $n_{b-\mathrm{jets}}(\mathrm{back})$ is the number of tagged L5 calorimeter b-jets estimated in the selected data that are not from $W^\pm$+$b\bar{b}$. The denominator consists of an efficiency, ϵ, and acceptance, $\mathcal{A}$, term that will correct the number of tagged L5 calorimeter b-jets from $W^\pm$+$b\bar{b}$ in the selected data back to the true number of hadron b-jets produced (in $W^\pm$+$b\bar{b}$ events passing the initial criteria listed above) in the luminosity, $\mathcal{L}$.

The denominator of 4.2 is expanded into a sum over the different trigger types used:

$$\epsilon \cdot \mathcal{A} \cdot \mathcal{L} \equiv \epsilon_{\mathrm{z0}} \cdot \epsilon_{b-\mathrm{tag}} \cdot \sum_{\mathrm{trig}} \left(\epsilon \cdot f_{\mathrm{lepton\ ID}} \cdot \mathcal{A} \cdot \mathcal{L}\right)_{\mathrm{trig}} \tag{4.3}$$

where ϵ_{z0} is the efficiency of the requirement that the primary lepton vertex be within 60 cm of the center of the detector, and $\epsilon_{b-\mathrm{tag}}$ is the efficiency to tag a b-jet with the default

SECVTX tagger[53]. In the sum over triggers (CEM, CMUP, CMX), ϵ is the trigger efficiency for the given channel, which is a measure of how well the triggers identify all charged leptons produced at CDF[54, 55]. $f_{\text{lepton ID}}$ is the lepton ID scale-factor, which accounts for any differences in the efficiency to select a charged lepton in MC versus data, since MC will be used to estimate the efficiency of the lepton selection requirements of the previous chapter[56].

It can immediately be seen that the number of b-jets from $W^{\pm}$+$b\bar{b}$ in the Pretag sample (*i.e.* - passing all selection, but no tagged jet requirement) is given by:

$$n_{W^{\pm}+b\bar{b}\ b-\text{jets}}(\text{Pretag}) = \frac{n_{b-\text{jets}}(\text{fit}) - n_{b-\text{jets}}(\text{back})}{\epsilon_{b-\text{tag}}} \tag{4.4}$$

$W^{\pm}$+$b\bar{b}$ events in the selected sample can potentially have two tagged b-jets, so there is an assumption in Eq. 4.4 that the (per-jet) efficiency to tag a b-jet is independent of the number of b-jets in an event. This assumption probably holds in events with only 2 b-jets, but potentially breaks down once the number of b-jets becomes so great that tracks from different jets start to overlap. There is also an assumption that the efficiency to tag a b-jet is independent of whether the jet contains more than one b-hadron. These assumptions will be addressed and systematic uncertainties will be assigned if appropriate.

4.1.2 Acceptance Definition

The Acceptance term, $\mathcal{A}$, in Eq. 4.3, is used to extrapolate the measured number of $W^{\pm}+b\bar{b}$ b-jets back to some pre-selection level. This necessitates the use of MC since there is no 100% pure sample of $W^{\pm}+b\bar{b}$ data to study. As mentioned, events used in the Acceptance calculation are restricted to those in which the decay products of the $W^{\pm}$ boson pass the specified cuts at the HEPG level. The reason for this approach is a desire to reduce the dependence of the final result on theoretical input, and allowing the $W^{\pm}$ decay products to have unrestricted kinematics at the HEPG level only increases the amount theory is relied upon to extrapolate to regions of phase-space outside of the selected data.

The other method of restricting the scope of the acceptance comes by only correcting back to events that have one or two hadron jets (passing the cuts listed above), including at least one identified as a b-jet, as opposed to correcting back to the total number of events (regardless of their jet multiplicity). The definition of the Acceptance is also motivated by wanting to test the reconstructed L5 calorimeter jets against the actual hadron jets that were produced in collisions.

The $W^{\pm}+b\bar{b}$ acceptance (for each trigger channel) is written as a weighted sum over several $W^{\pm}+b\bar{b}$+parton samples:

$$\mathcal{A}_{\mathrm{trig}} \equiv \left(\sum_{\mathrm{i}=0}^{2} \mathrm{w_i} \cdot \mathcal{A}_{W^{\pm}+b\bar{b}+\mathrm{i\ partons}} \right)_{\mathrm{trig}} \tag{4.5}$$

In Eq. 4.5 $\mathrm{w_i}$ is a weighting factor calculated as the generator cross-section of the $W^{\pm}+b\bar{b}+\mathrm{i}$ parton sample divided by the sum of the generator cross-sections for all $W^{\pm}+b\bar{b}$ samples. The $\mathrm{w_i}$ factors will allow each $W^{\pm}+b\bar{b}+\mathrm{N}$ parton MC sample to contribute appropriately to the overall acceptance, which should give the best description of the true $W^{\pm}+b\bar{b}$ physics.

Alpgen v2.1 has a built-in parton-jet matching scheme, referred to as "MLM-matching", which ensures that quantities from the different $W^{\pm}+b\bar{b}$+Np samples can be added together without over/under-representing any regions of phase-space[57]. The MLM-matching checks to see that all partons from the matrix-elements are matched (via a ΔR cone) to a unique cluster of final-state hadrons from Herwig. If there is a parton that is not matched to a cluster of hadrons, or there is an additional cluster of hadrons left over after all partons have been matched up (due to Herwig adding in excessively energetic radiation, for example), Alpgen throws the event out. Exclusive matching means that each parton is uniquely matched to a cluster of hadrons with no hadron clusters left over.

Inclusive matching allows for additional clusters of hadrons to be left after all partons have been matched, to allow for some modeling of higher-order phase-space for which there is no MC. The various $W^{\pm}$+$b\bar{b}$+0,1 parton samples have had exclusive MLM matching applied during generation, while the $W^{\pm}$+$b\bar{b}$+2 parton sample has been inclusively matched. Figure 4.1 depicts the idea behind MLM-matching. The weights, w_i, of Eq. 4.5 will be taken from the contribution of the generator-level cross-section of each $W^{\pm}$+$b\bar{b}$+i parton sample to the combined generator-level cross-section.

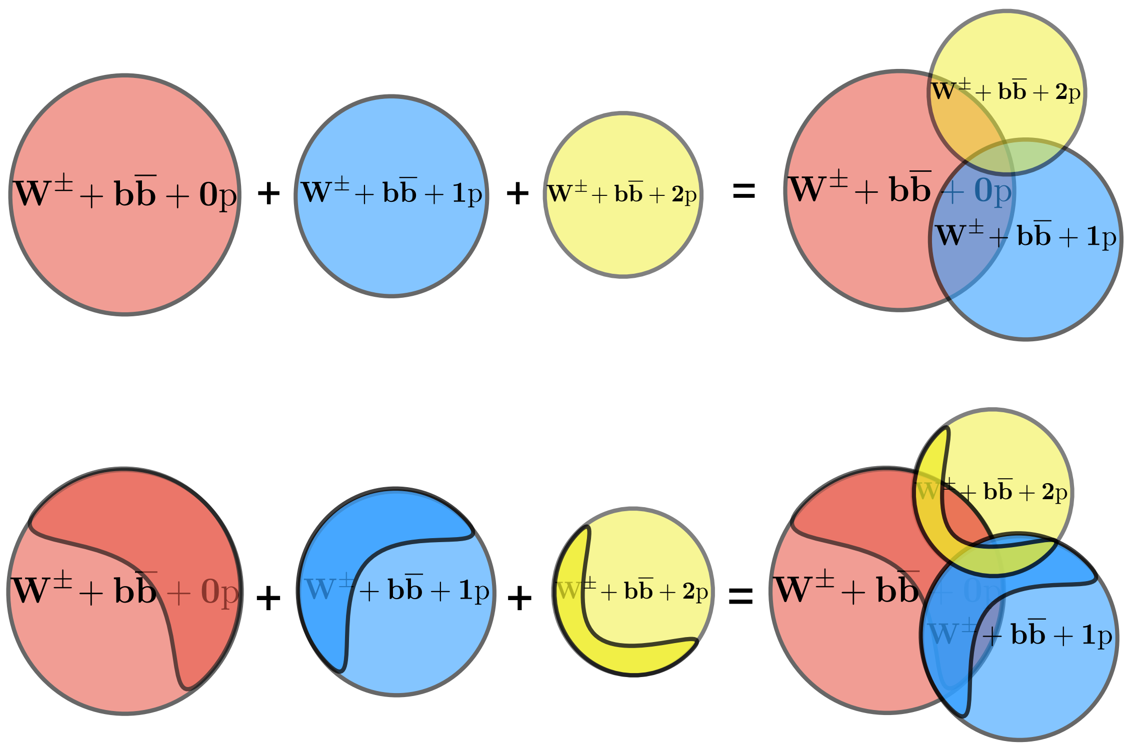


Figure 4.1: These diagrams represent the "double-counting" problem that can occur if one were to include all events from all $W^{\pm}$+$b\bar{b}$ samples into one result. The circles represent the phase-space of $W^{\pm}$+$b\bar{b}$ events from the different multiplicity MC samples. The goal is to merge these samples into one inclusive set that best represents all of $W^{\pm}$+$b\bar{b}$, but certain regions of phase space could be over-represented and this would carry through to the acceptance (ratio of shaded region to total area of each circle) of each sample. MLM-matching ensures that the different $W^{\pm}$+$b\bar{b}$+Np samples are appropriately generated and can be added together.

The acceptance for a given $W^{\pm}$+$b\bar{b}$ sample is factorized into two terms:

$$\mathcal{A}_{W^{\pm}+b\bar{b}+\mathrm{i\ partons}} \equiv \mathcal{A}_{\mathrm{jet}} \cdot \mathcal{A}_{\mathrm{selection}} \tag{4.6}$$

where $\mathcal{A}_{\text{jet}}$ will allow hadron-level jets to be connected with calorimeter jets, and $\mathcal{A}_{\text{selection}}$ will correct for the geometric/kinematic/veto cuts applied. $\mathcal{A}_{\text{jet}}$, which is calculated directly after applying the $|z_0|$ cut, is written as:

$$\mathcal{A}_{\text{jet}} \equiv \frac{n_{b-jets}(1+2 \text{ L5 Cal. jets})}{n_{b-jets}(1+2 \text{ Had. jets})} \tag{4.7}$$

while $\mathcal{A}_{\text{selection}}$ is written:

$$\mathcal{A}_{\text{selection}} \equiv \frac{n_{b-jets}(\text{after selection; } 1+2 \text{ L5 Cal. jets})}{n_{b-jets}(1+2 \text{ L5 Cal. jets})} \tag{4.8}$$

In Eq. 4.7 the number of calorimeter b-jets in events with 1 or 2 L5 calorimeter jets is compared to the number of hadron b-jets in events with 1 or 2 hadron jets. This term allows the performance of L5 jet corrections to be checked, since these are supposed to be equivalent to hadron jets (on average) in energy.

In Eq. 4.8 the number of L5 calorimeter b-jets (in events with 1 or 2 L5 calorimeter jets) is compared before and after applying selection cuts. The selection cuts are applied to improve the signal to background ratio of the final data, but it is unavoidable to remove some signal with these cuts. The value of Eq. 4.8 will be used to correct for the fraction of $W^{\pm}$+$b\bar{b}$ signal removed during the selection process.

4.2 MC SECVTX Mass Templates

Templates are constructed of the SECVTX mass of different quark flavors (bottom, charm, and light) using tagged jets in various MC samples. After applying all event selection bottom and charm hadrons in the event shower are matched to tagged jets in the 1 and 2 calorimeter jet bins by requiring them to be within a cone of ΔR$\leq$0.4. If a jet is matched to both a bottom and charm hadron, then it is included in the shape for b-jets. For light-flavor taken from the MC, any tagged jet not matched to either a b or c

hadron is called a light-flavor jet. Negatively tagged jets from the 1 and 2 jet bins of the data are used as an alternative model of light-flavor tags.

Templates are made in a wide variety of MC samples, listed in Table 4.1, and then combined into one final template for each quark flavor. After applying the selection described in Chapter 3, but before merging all MC samples together into a final template, the templates from each MC sample are normalized to the expected shape in a given luminosity for each process (accomplished based on the generator-level cross-section and number of events available for each sample). The normalized templates for each process and each flavor are combined, and the final templates are displayed in Figure 4.2.

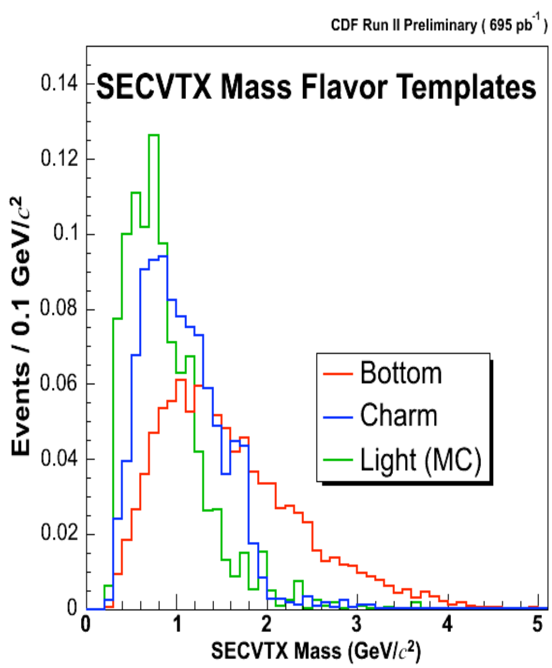


Figure 4.2: Templates of SECVTX mass for various quark flavors. We use the light-flavor model taken from the MC as our default, and use the model of negatively-tagged jets as a systematic check.

4.3 SECVTX Mass Fitting Procedure

A binned negative log-likelihood fit is used to determine the fraction of tagged calorimeter jets in the selected data sample that are due to b-jets. The likelihood function for

Process	Fileset	Cross-Section (pb)
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+0p	mtop0x	3.168
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+0p	mtop1x	1.035
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+0p	mtop2x	0.387
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+0p	mtop0y	3.170
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+1p	mtop1y	1.035
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+2p	mtop2y	0.387
$W^{\pm}(\rightarrow e\nu) + c\bar{c}$+0p	ltop0c	2.453
$W^{\pm}(\rightarrow e\nu) + c\bar{c}$+1p	ltop1c	1.221
$W^{\pm}(\rightarrow e\nu) + c\bar{c}$+2p	ltop2c	0.550
$W^{\pm}(\rightarrow \mu\nu) + c\bar{c}$+0p	ltop3c	2.509
$W^{\pm}(\rightarrow \mu\nu) + c\bar{c}$+1p	ltop4c	1.221
$W^{\pm}(\rightarrow \mu\nu) + c\bar{c}$+2p	ltop5c	0.550
$W^{\pm}(\rightarrow e\nu) + c$+0p	ltop0a	10.767
$W^{\pm}(\rightarrow e\nu) + c$+1p	ltop1a	4.447
$W^{\pm}(\rightarrow e\nu) + c$+2p	ltop2a	1.429
$W^{\pm}(\rightarrow e\nu) + c$+3p	ltop3a	0.436
$W^{\pm}(\rightarrow \mu\nu) + c$+0p	ltop4a	10.748
$W^{\pm}(\rightarrow \mu\nu) + c$+1p	ltop5a	4.447
$W^{\pm}(\rightarrow \mu\nu) + c$+2p	ltop6a	1.429
$W^{\pm}(\rightarrow \mu\nu) + c$+3p	ltop7a	0.475
$W^{\pm}(\rightarrow e\nu)$+1p	ltop1n	436.8
$W^{\pm}(\rightarrow e\nu)$+2p	ltop2n	99.0
$W^{\pm}(\rightarrow e\nu)$+3p	ltop3n	22.7
$W^{\pm}(\rightarrow e\nu)$+4p	ltop4n	5.2
$W^{\pm}(\rightarrow \mu\nu)$+1p	ltop1m	436.8
$W^{\pm}(\rightarrow \mu\nu)$+2p	ltop2m	99.0
$W^{\pm}(\rightarrow \mu\nu)$+3p	ltop3m	22.7
$W^{\pm}(\rightarrow \mu\nu)$+4p	ltop4m	5.2
$t\bar{t}$	ttopkl	6.7
Single-Top (s-chan.)	mtopya	0.88
Single-Top (t-chan.)	mtopta + ua_match	1.98
$W^{\pm}W^{\pm}$	wtop1w	12.4

Table 4.1: MC samples used in construction of SECVTX mass templates.

this analysis, defined in Eq. 4.9, is a function of three parameters representing the three flavor fractions.

$$\mathscr{L}(f_b, f_c, f_l) = \prod_{i=1}^{N_{bins}} P_i(n_i; f_b, f_c, f_l). \tag{4.9}$$

$P_i(n_i; f_b, f_c, f_l)$ is the probability of observing n_i tagged jets in the i^{th} bin of the SECVTX mass distribution given the fractions of tagged jets coming from bottom (f_b), charm (f_c), and light (f_l) jets. The data, n_i, are assumed to be Poisson distributed in each bin:

$$P_i(n_i; f_b, f_c, f_l) = e^{-\mu_i} \cdot \frac{\mu_i^{n_i}}{n_i!} \tag{4.10}$$

where μ_i is the expected number of tagged jets in each bin and is given by:

$$\mu_i \equiv (f_b \cdot P_{b_i} + f_c \cdot P_{c_i} + f_l \cdot P_{l_i}) \cdot N_{\text{total}} \tag{4.11}$$

In Eq. 4.11 the P_{x_i} variables are the probability of a jet of flavor x in bin i, and N_{total} is the total number of tagged jets (1267) observed in the selected data sample. The P_{x_i} values are taken from the SECVTX mass templates for each flavor that were discussed in Section 4.2.

The MINUIT package is used to find the values of f_b, f_c, and f_l that minimize the negative logarithm of the likelihood function[58]. All three of these parameters are allowed to float during the minimization procedure. Figure 4.3 shows the fit to the data when using a MC derived shape for the light-flavor template, and also shows the negative logarithm of the likelihood function. A b-fraction of $32.81\pm^{3.41}_{3.24}\%$ is measured, corresponding to $415.7\pm^{43.2}_{41.1}$ tagged b-jets. The χ^2/DOF of the fit is close to one, indicating that the templates are doing a good job of describing the data.

Figure 4.4 shows the $\pm$1,2,3,4,5σ confidence bands for the three flavor-fractions plotted against one another. The b-fraction appears largely uncorrelated with the light-flavor fraction, but somewhat correlated with the c-fraction. The c-fraction and light-flavor fraction appear strongly anticorrelated, which is perhaps not surprising given the similarity of their SECVTX mass shapes.

As a check of the robustness of this approach of modeling the data, we apply the simple three-flavor model to other kinematic distributions. Figures 4.5-4.7 show distributions for tagged jets, broken down by flavor contribution. The three flavors in each plot are normalized to the same fractions as were found by the SECVTX mass likelihood fit used in the cross-section calculation. Good agreement is observed for all variables, giving us confidence in this method.

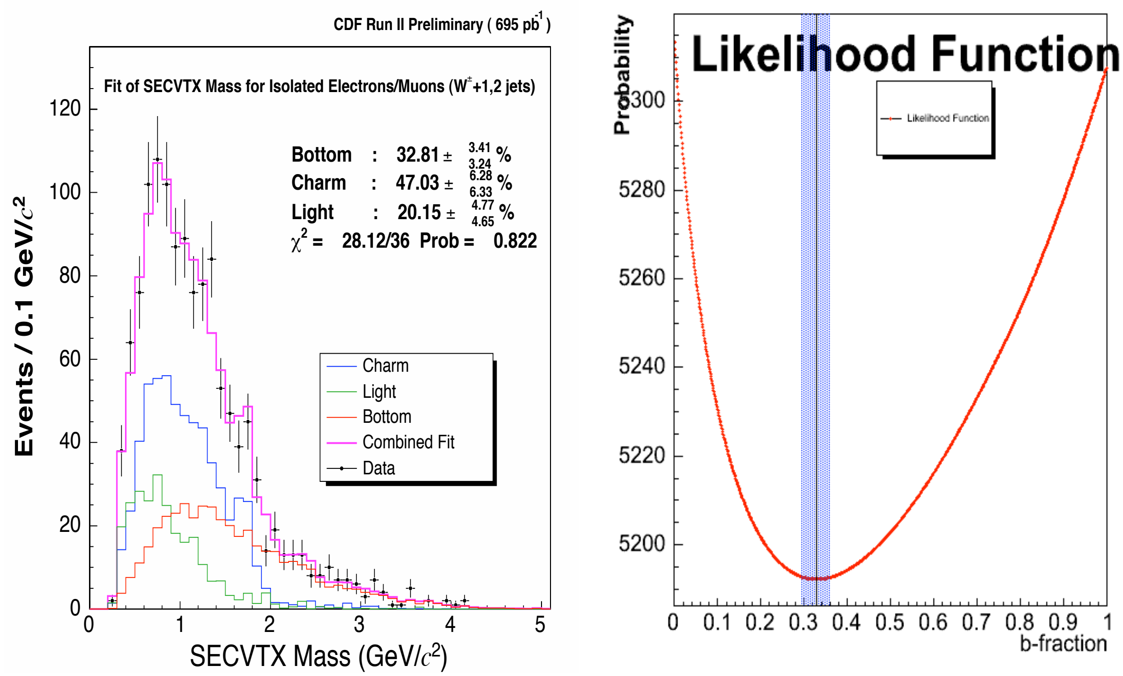


Figure 4.3: Left: SECVTX Mass fit when using MC to derive the shape for tagged light-flavor jets. Right: The negative logarithm of the likelihood function returned by MINUIT, as well as a band covering the $\pm\sigma$ uncertainties on the fitted b-fraction.

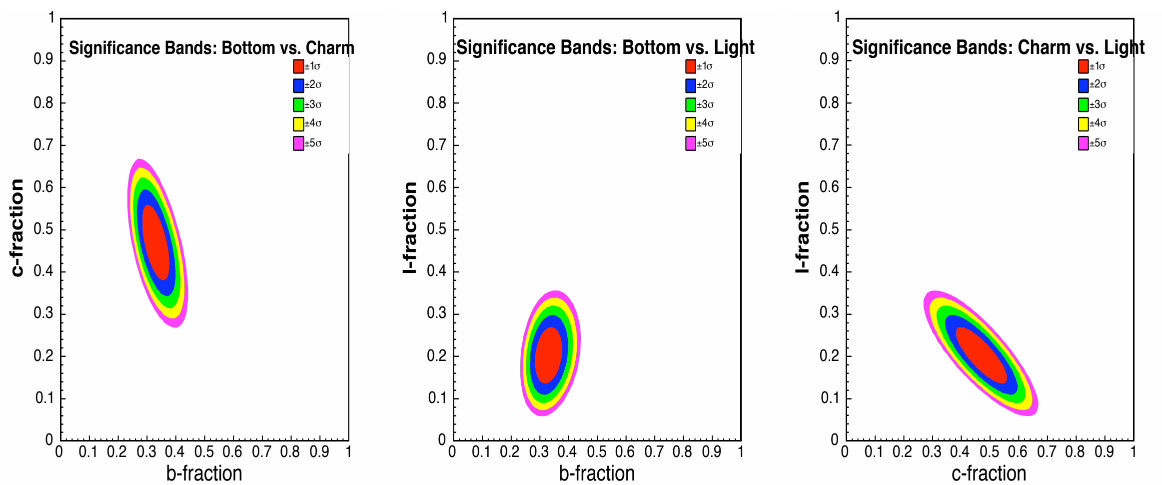
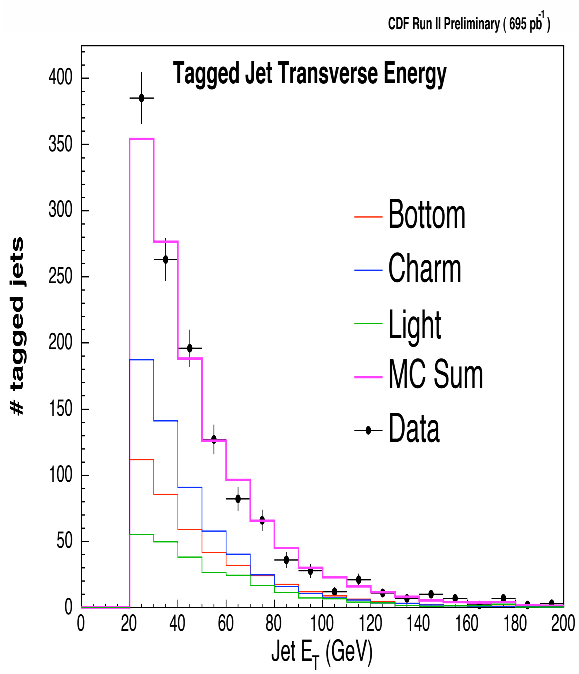


Figure 4.4: Confidence Bands for fitted b, c, and l fractions plotted against each other when using MC to derive the shape for tagged light-flavor jets.



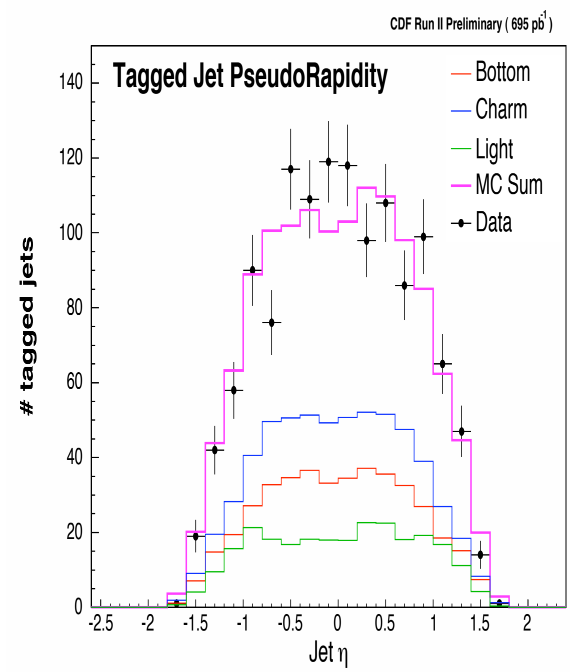
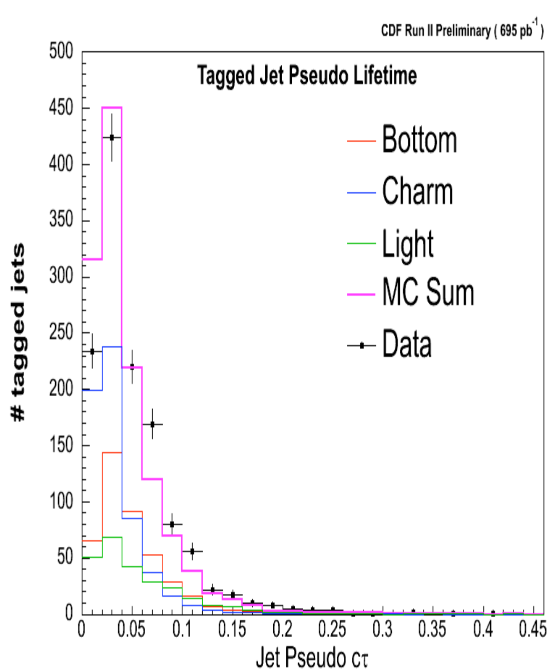


Figure 4.5: Left: E_T distribution for tagged jets passing all selection. Right: η distribution for tagged jets passing all selection.



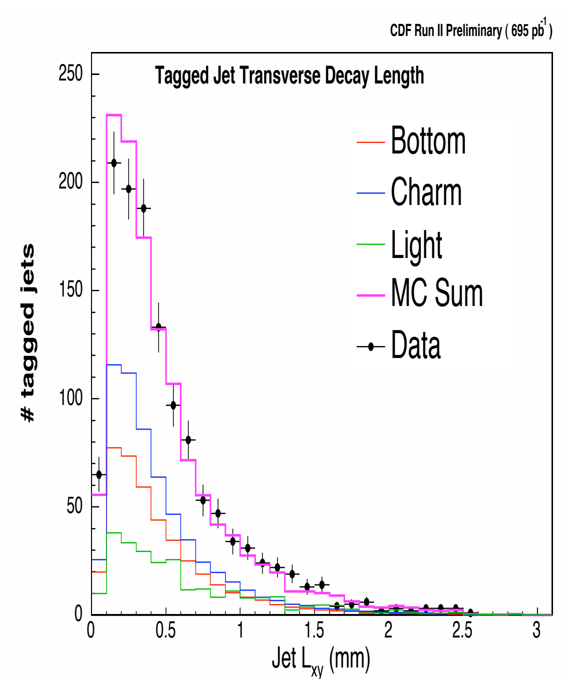


Figure 4.6: Left: Pseudo $c\tau$ distribution for tagged jets passing all selection. Right: L_{xy} distribution for tagged jets passing all selection.

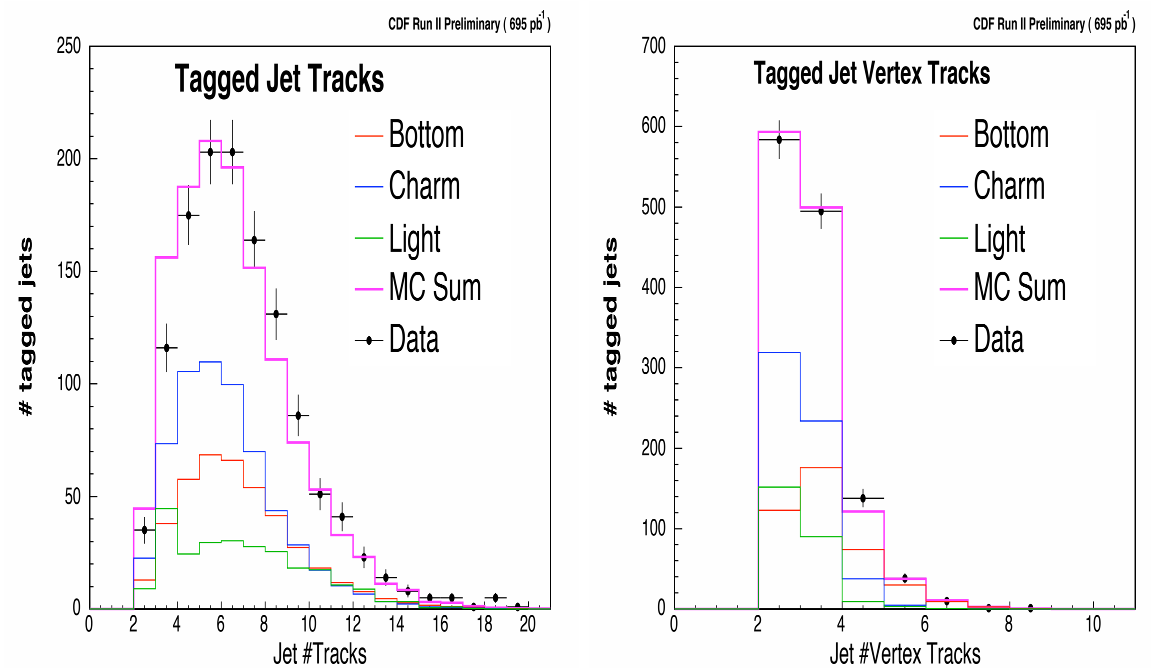


Figure 4.7: Left: Track multiplicity distribution for tagged jets passing all selection. Right: SECVTX Track multiplicity distribution for tagged jets passing all selection.

4.4 $\epsilon_{b-\text{tag}}$

In the cross-section calculation, tagged b-jets from $W^{\pm}$+$b\bar{b}$ are isolated in the data (background subtraction will be described in the next chapter), and then a luminosity-weighted acceptance is applied to correct back to the desired hadron level. Before doing this, the efficiency to tag a b-jet from $W^{\pm}$+$b\bar{b}$ must be accounted for, in order to estimate the number of pretag b-jets from $W^{\pm}$+$b\bar{b}$ that the acceptance operates on. MC is used to estimate the efficiency to tag a b-jet from $W^{\pm}$+$b\bar{b}$. The number of tagged b-jets in the 1+2 jet bins of the MC, after applying all selection, is divided by the total pretag number of b-jets in the 1+2 jet bins. The efficiencies from the various $W^{\pm}$+$b\bar{b}$+N parton samples are added together in a weighted sum (using the same weights as we do for the acceptance calculation), as shown in Eq. 4.12.

$$\epsilon_{b-\text{tag}} = SF \cdot \sum_{i=0}^{2} \left(w_i \cdot \frac{n_{tagged\ b-jets}}{n_{pretag\ b-jets}} \right)_{W^{\pm}+b\bar{b}+i} \tag{4.12}$$

The b-tag scale-factor is applied in Eq. 4.12, as mentioned previously, to account for any MC/data differences in tagging efficiencies. Table 4.2 summarizes the calculation of $\epsilon_{b-\text{tag}}$. In both the electron and muon channels, the b-tag efficiency comes out similarly. The value found is nearly identical to that found in the $t\bar{t}$ cross-section analysis, which makes sense since naively one would expect the efficiency to tag a b-jet to be independent of what process that jet comes from. This is also reassuring because it indicates that the effect of jets containing two b-hadrons, which is likely for $W^{\pm}$+$b\bar{b}$ but not for $t\bar{t}$, is small. The uncertainty quoted on the final b-tag efficiency is due to the uncertainty on the scale factor.

Sample	$n_{b-\text{jet}}$(tagged)	$n_{b-\text{jet}}$(pretag)	$\epsilon_{b-\text{tag}}$	Total
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+0p	12483	29573	0.422	0.37±0.029
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+1p	12433	31161	0.399	
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+2p	5936	15836	0.375	
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+0p	7526	17810	0.423	0.37±0.029
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+1p	7244	18403	0.394	
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+2p	3504	9290	0.377	

Table 4.2: The numbers used to calculate $\epsilon_{b-\text{tag}}$. The efficiencies from the individual $W^{\pm}$+$b\bar{b}$+N partons are summed, after weighting and applying the scale factor, to calculate the final efficiency in the last column.

4.5 Hadron Level Jets

Hadron-level jets are formed by running standard JETCLU (the exact same clustering algorithm used for calorimeter jets), with a cone-size of 0.4, over physics towers formed from final-state particles from Herwig or Pythia. The decay products of the $W^{\pm}$ (*i.e.* - $e^{\pm}$, $\mu^{\pm}$, ν) in $W^{\pm}$+$b\bar{b}$ are excluded from these physics towers to avoid their presence in the resulting jets.

As was mentioned previously, the motivation for using these hadron-level jets is to clearly define a cross-section so that any theorist can compare his/her models against the result. L5-corrected calorimeter jets are supposed to be equivalent in energy, on average, to the energy of the underlying hadrons that deposited into the CDF calorimeters. By incorporating a term ($\mathcal{A}_{\rm jet}$) in the acceptance that directly compares L5-corrected calorimeter jets and hadron-level jets, the claim about L5 jets equivalence to hadron jets can easily be shown to be true (in which case this part of the Acceptance should be $\approx$1), and if the claim is not exactly true than the term corrects for it and removes any CDF effects.

The first item checked is how the E_T of light-flavor calorimeter jets (*i.e.* - Jets that are explicitly not matched to b or c hadrons), that have been corrected to L4 or L5, compare to the E_T of any hadron jet they are matched to (via a ΔR$\leq$0.4 cone matching). The L4 comparisons are shown since they help illustrate the effect of the L5 corrections. Figure 4.8 shows this comparison for L4 (top, left) and L5 (top, right) light-flavor jets in a $W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b} + 1p$ MC sample. These same comparisons are also made for jets that are explicitly matched to b hadrons, as show in Figure 4.8 for L4 (bottom, left) and L5 (bottom, right).

In Figure 4.8 it is obvious that L5 is doing a much better job than L4 at matching the E_T of the light-flavor and bottom hadron jets. For light-flavor L5 jets, the fractional difference in energy between the matching calorimeter and hadron jets is roughly centered around 0.0 and flat over a wide range of calorimeter jet E_T. This is perhaps not unexpected, since the L5 corrections are built by tuning calorimeter jet energies to matching hadron jet energies in MC, but it gives confidence in the implementation of the hadron jet clustering.

For L5 b-jets the agreement in energy with the matching hadron jets is certainly better than it is at L4, but the fractional difference in energy with the hadrons jets is not centered around 0.0, especially at low E_T values. It is noted that there is approximately a 10% offset in E_T between the L5 b-jets and the hadron b-jets. Though this 10% offset is not necessarily flat over the whole range of L5 b-jet E_T, the practice is adopted of applying it as a b-specific correction to L5 b-jets. The CDF jet corrections are constructed using samples dominated by light-flavor jets, and since b-quarks have different fragmentation properties and more neutrinos in their final state, it is not surprising that the corrected b-jets don't adequately describe the underlying hadron level.

Figure 4.9 shows the effect of adding an additional 10% correction to all L5 b-jets while selecting events (right), as opposed to not applying any b-specific correction (left). The agreement between hadron b-jets and L5 b-jets after implementing the 10% correction, though still not as nice as that of the light-flavor, is much improved.

It seems that correcting calorimeter jets to L5 does a good job at modeling the underlying jet physics in $W^{\pm}$+$b\bar{b}$ events. To further probe that claim the agreement in E_T between hadron and calorimeter light-flavor jets and b-jets as a function of pseudorapidity is tested. Figure 4.10 shows these comparisons for light-flavor jets (left) and b-jets (right), for both L4 and L5. Both L4 and L5 seem to have no obvious dependence in η as far as matching the energy of the hadron level. Clearly L4 still is offset from the true hadron energy, for both light-flavor jets and b-jets, with most of this offset driven by the low E_T calorimeter jets, as shown in Figure 4.11.

As a final test of the effectiveness of L5 corrections, comparisons are made of the $\not\!\!E_T$ distributions for both L4 and L5 jets to that of the neutrino p_T from the HEPG information. The standard selection (including the 10% b-specific correction) is applied, with the exception of the neutrino p_T and $\not\!\!E_T$ cut, to a $W^{\pm}$+$b\bar{b}$+1p MC sample. Figure 4.12 shows the comparison, and confirms that L5 does a superior job at modeling the underlying neutrino physics of the MC.

4.6 Acceptance Calculation

The updated $W^{\pm}$+$b\bar{b}$ cross-section definition described in Section 4.1 relies upon an acceptance that is factorized into two parts, $\mathcal{A}_{\text{jet}}$ and $\mathcal{A}_{\text{selection}}$. Table 4.3 summarizes the acceptance values, calculated using Eq. 4.5 and Eq. 4.6 , for the main $W^{\pm}$+$b\bar{b}$ MC. The jet acceptances are all very close to unity, which is a sign that L5 jets and hadron jets are very similar.

Process	Trigger	$\mathcal{A}_{\text{jet}}$	$\mathcal{A}_{\text{selection}}$	$\mathcal{A} = \mathcal{A}_{\text{jet}} \cdot \mathcal{A}_{\text{sel.}}$	$\mathcal{A}_{\text{trig}}$
$W^{\pm}(\to e^{\pm}\nu) + b\bar{b}$+0p	CEM	0.952±0.006	0.538±0.004	0.512±0.005	0.505±0.007
$W^{\pm}(\to e^{\pm}\nu) + b\bar{b}$+1p	CEM	0.951±0.005	0.517±0.004	0.492±0.005	
$W^{\pm}(\to e^{\pm}\nu) + b\bar{b}$+2p	CEM	0.971±0.008	0.485±0.005	0.471±0.006	
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+0p	CMUP	0.953±0.006	0.320±0.003	0.305±0.003	0.301±0.004
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+1p	CMUP	0.955±0.006	0.307±0.003	0.293±0.003	
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+2p	CMUP	0.976±0.008	0.286±0.003	0.279±0.004	
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+0p	CMX	0.953±0.006	0.134±0.002	0.128±0.002	0.127±0.002
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+1p	CMX	0.954±0.006	0.130±0.002	0.124±0.002	
$W^{\pm}(\to \mu^{\pm}\nu) + b\bar{b}$+2p	CMX	0.976±0.008	0.126±0.002	0.123±0.002	

Table 4.3: Acceptance values for main $W^{\pm}$+$b\bar{b}$ MC samples. The last column contains the final acceptance, derived from a weighted sum over the individual $W^{\pm}$+$b\bar{b}$+N parton samples. The weights are taken from Table 7.5.

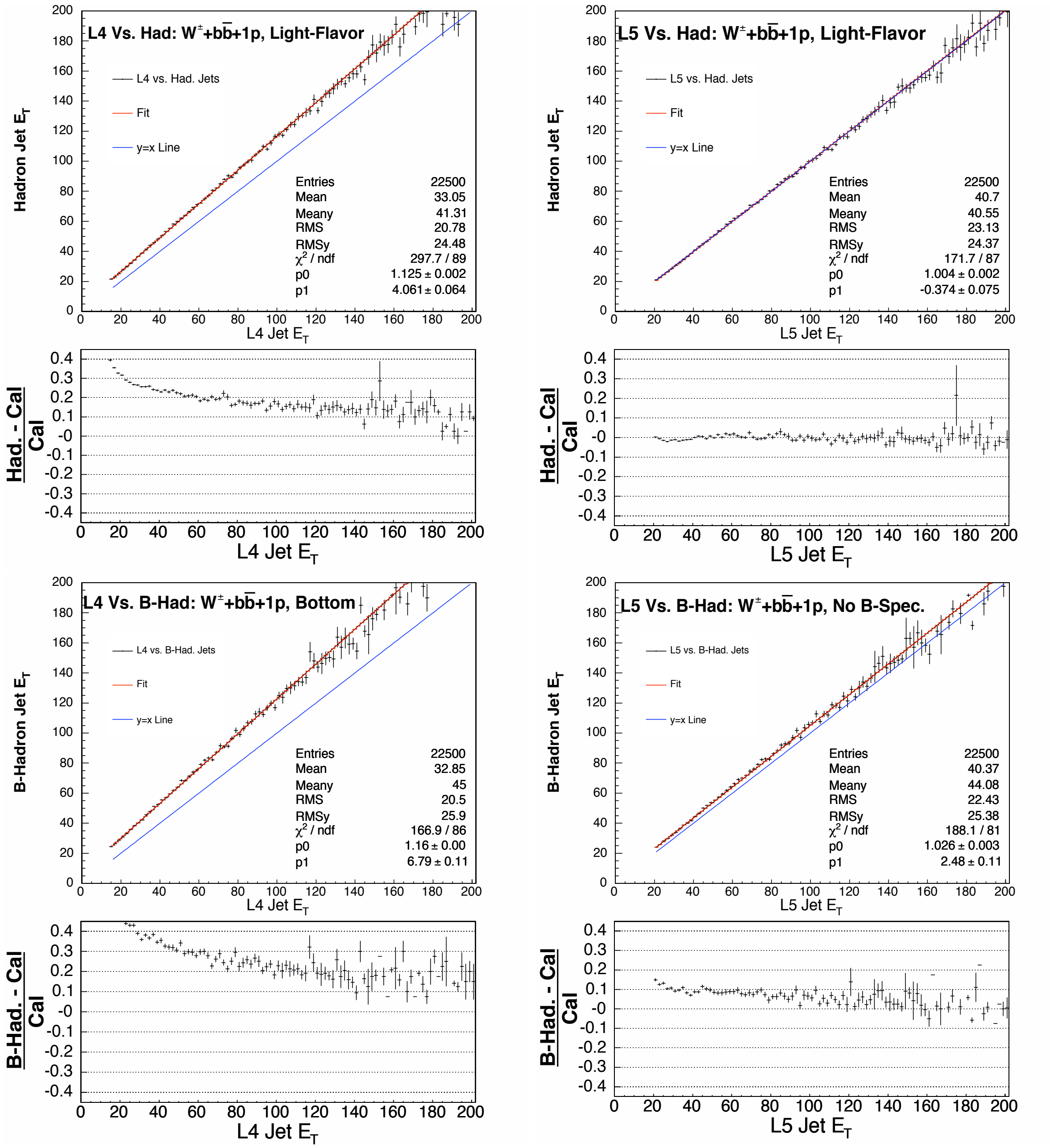


Figure 4.8: **Top**: Comparison of L4 (Left) and L5 (Right) corrected light-flavor (*i.e.* - jets that are specifically checked to not match any b or c hadrons) Calorimeter jet E_T with matching Hadron-level jet E_T. **Bottom**: Same comparisons, but for jets matched to a single b hadron.

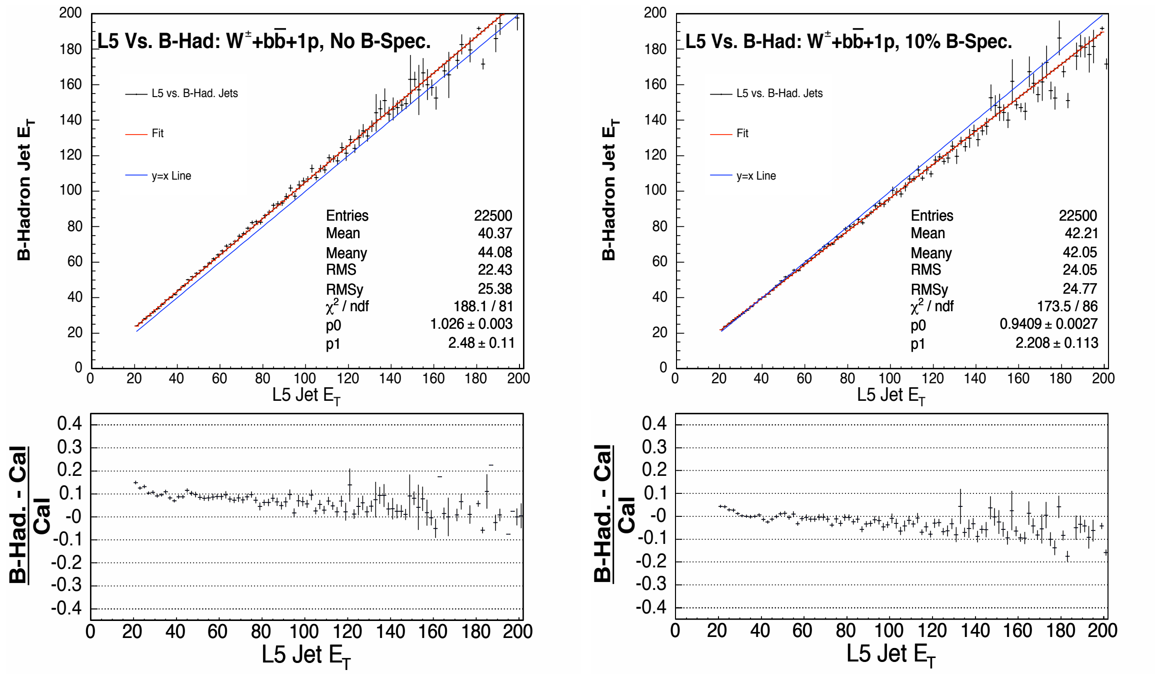


Figure 4.9: Comparison of L5 corrected Calorimeter b-jet E_T with no 10% b-specific correction (Left) and L5 corrected Calorimeter b-jet E_T with 10% b-specific correction (Right), with matching Hadron-level jet E_T.

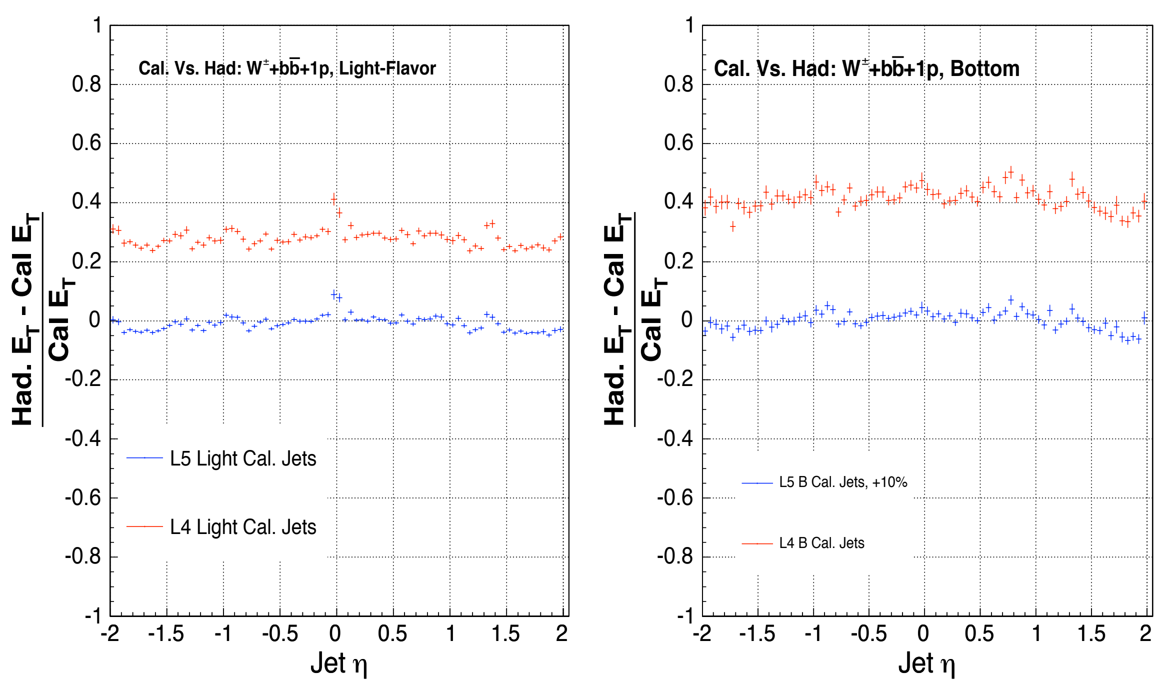


Figure 4.10: Comparison of L4 and L5 corrected Cal. jet fractional energy difference as a function of η with matching hadron jet, for light-flavor (Left) and bottom (Right).

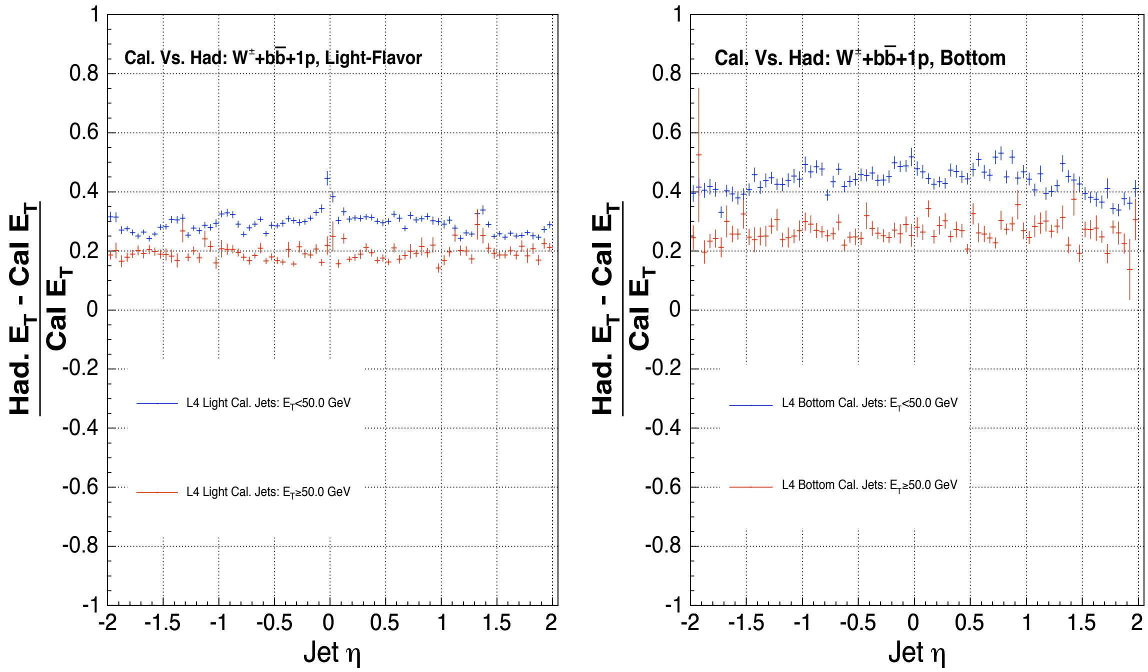


Figure 4.11: Comparison of L4 corrected Cal. jet and matching hadron jet fractional energy difference, as a function of η. Here, the comparison is split up into jets with E_T <50.0 GeV and E_T ≥50.0 GeV to show that most of the disagreement is driven by the low E_T jets.

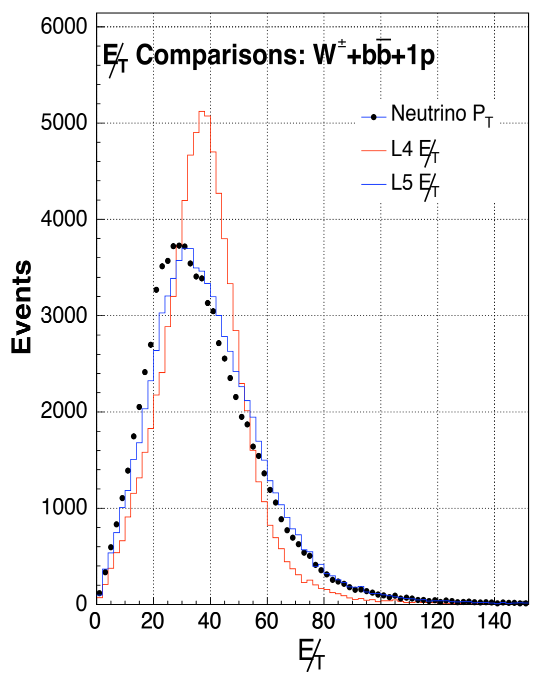


Figure 4.12: Comparison of neutrino p_T, L4 and L5 $\not{E}_T$, in $W^{\pm}$+$b\bar{b}$+1 parton MC. All selection criteria are applied except for those normally placed on the neutrino p_T and the reconstructed $\not{E}_T$. L5 corrections do a much better job than L4 of describing the true neutrino p_T distribution.

CHAPTER 5

Backgrounds

5.1 Background b-jet Sources

Not all of the b-jets estimated in our SECVTX mass fit are due to $W^{\pm}$+$b\bar{b}$. There are two main sources of background b-jets in the 1+2 jet bins of the data:

1. n_{QCD}: tagged b-jets from QCD, which we will calculate using data.
2. n_{MC}: tagged b-jets from processes that we must use MC to calculate.
   - $t\bar{t}$
   - Single-Top (s-channel and t-channel)
   - DiBoson - WZ/WW/ZZ
   - $W^{\pm}(\rightarrow \tau\nu) + b\bar{b}$+N partons
   - $Z \rightarrow \tau\tau$
   - $Z + b\bar{b}$

5.1.1 n_{QCD}

QCD background refers to multijet events, containing no real (*i.e.* - on-shell) $W^{\pm}$ boson content, that pass selection due to a combination of mismeasured jet energies, jets faking electrons, or semileptonic decays of b-quarks producing real leptons and missing energy. Figure 5.1 gives two examples of possible QCD background diagrams. The defining feature of this background, which will be used to estimate its contribution to the selected data, is that the lepton (whether it is a real lepton or a fake) should be very near jet activity.

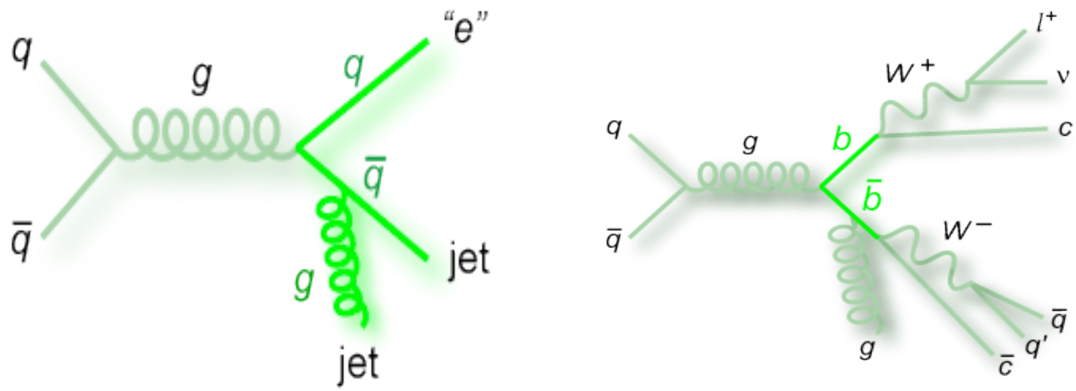
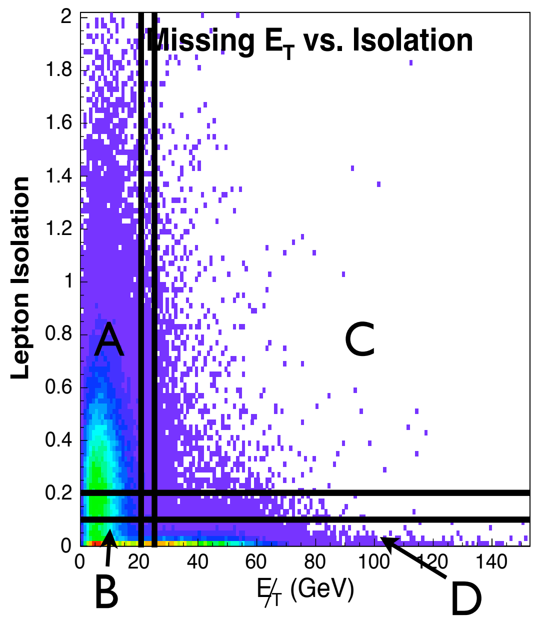


Figure 5.1: Examples of QCD background processes. Left: Jet fakes an electron Right: Semileptonic decay of b-quark produces real lepton and missing energy.

To estimate the number of tagged QCD events (N_{QCD}) in the data sample the $\not\!\!E_T$ vs. isolation method, which operates on the assumption that the $\not\!\!E_T$ and lepton isolation in an event are uncorrelated, is used[59]. The usual selection is applied, with the exception of the $\not\!\!E_T$ ≥25.0 GeV cut and the lepton isolation≤0.1 cut, which are both omitted. In events with non-isolated (*i.e.* - Isol>0.1) leptons any jets within ΔR≤0.4 of the lepton are not counted, but corrections to their energy are factored into the $\not\!\!E_T$ calculation. For electron and muon (CMUP or CMX) events in the data that have 1 or 2 L5 corrected jets, including at least one tagged jet, the $\not\!\!E_T$ of the event is plotted against the isolation of the lepton as in Figure 5.2.

The $\not\!\!E_T$ vs. isolation plane is divided into 4 sectors, as shown in Figure 5.2. Events from the signal region will be located in Region "D", with large $\not\!\!E_T$ and small isolation. The $\not\!\!E_T$ distribution is plotted for events with isolation≥0.2, which should be highly enriched in QCD, and also for events with isolation≤0.1, which should be enriched in QCD in the low $\not\!\!E_T$ region. To estimate the number of tagged QCD events in the signal region, the $\not\!\!E_T$ shape from the isolation≥0.2 region is scaled such that the integral below 20.0 GeV



Region	$\not\!E_T$ (GeV)	Isolation
A	≤20.0	≥0.2
B	≤20.0	≤0.1
C	≥25.0	≥0.2
D	≥25.0	≤0.1

Figure 5.2: Example of the $\not\!E_T$ vs. Isolation plane used to estimate the number of tagged QCD events in our signal (Region "D").

is equal to the integral below 20.0 GeV in the $\not\!E_T$ shape from the isolation≤0.1 region. The resulting $\not\!E_T$ shape for the isolation≥0.2 region is then used to estimate the QCD background by integrating the shape above the $\not\!E_T$ cut of 25.0 GeV. Figure 5.3 shows this procedure at work for 1-jet and 2-jet electron and muon events in the data. The data shapes in Figure 5.3 have been corrected for $t\bar{t}$, diboson, single-top, etc... contamination (since the $\not\!E_T$ vs. Isol. method assumes only QCD events populate the space outside of the signal region). Table 5.1 summarizes the estimates of tagged QCD events in the signal region. A 25% systematic uncertainty is assigned to these QCD event estimates based on differences obtained by varying the boundaries of the $\not\!E_T$ vs. isolation regions.

Lepton	1-jet	2-jet
Electron	91.0±6.1	42.6±4.1
Muon	16.1±1.5	10.7±1.4

Table 5.1: Estimates of tagged QCD events in the signal region. Uncertainties shown are statistical only, and don't include any systematic contributions.

The $\not\!E_T$ vs. isolation method gives an estimate of the number of tagged QCD events

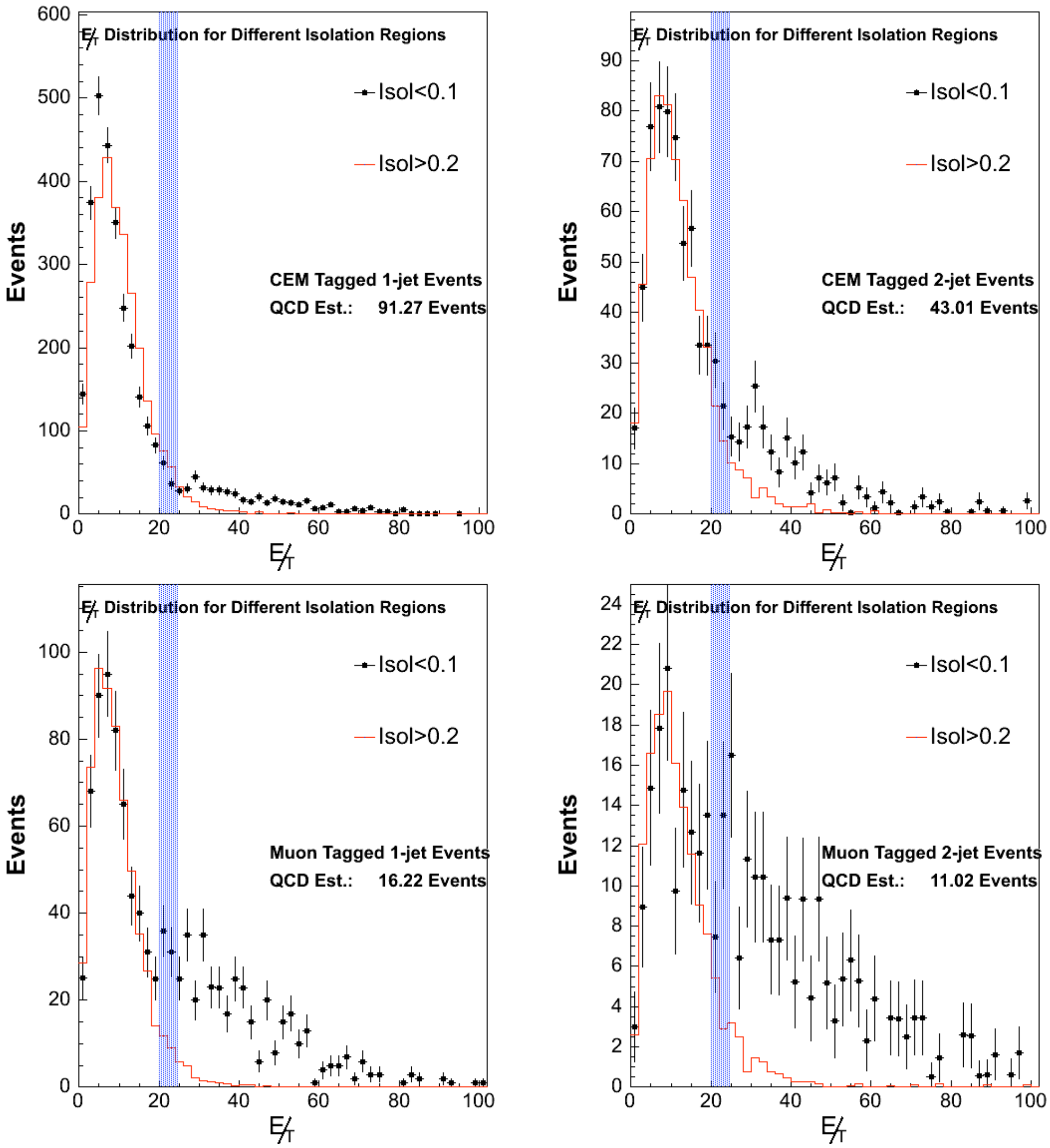


Figure 5.3: $\not\!\!E_T$ vs. isolation method at work. The $\not\!\!E_T$ shape from the isolation≥0.2 region is scaled to have the same integral below 20.0 GeV as that of the isolation≤0.1 region. The integral of the isolation≥0.2 $\not\!\!E_T$ shape above 25.0 GeV is taken to be the QCD background. The shaded band represents the portion of the $\not\!\!E_T$ spectrum that is ignored in the calculation of the QCD background.

in the signal region, but the portion of this background that is due to real b-decays must still be determined. To estimate the number of tagged b-jets in the 1+2 jet bins coming from QCD, the expected number of QCD events in the signal region, N_{QCD} (which is estimated using the $\not\!\!E_T$ vs. Isol. method, described above), is multiplied by the b-fraction of non-isolated lepton events, f_b, as shown in Eq. 5.1. To estimate the b-fraction of tagged jets coming from QCD the SECVTX mass distribution of tagged events with non-isolated leptons and low $\not\!\!E_T$ (*i.e.* - Region "A") is fit. These low $\not\!\!E_T$, non-isolated lepton events are a model of QCD events, and it is assumed that whatever b-fraction is found for them also applies to the isolated lepton QCD events in the signal region. Electrons and muons are treated separately since they have different b-fractions in the non-isolated samples.

$$n_{\mathrm{QCD}} = (f_b \cdot N_{\mathrm{QCD}})_{\mathrm{electrons}} + (f_b \cdot N_{\mathrm{QCD}})_{\mathrm{muons}} \tag{5.1}$$

The same flavor templates as were used in the fit of the signal data are used in this fit, and once again all three flavors are allowed to float in the likelihood fit. Figure 5.4 shows the SECVTX mass fit in events with lepton isolation≥0.2 and $\not\!\!E_T$ ≤20.0 GeV.

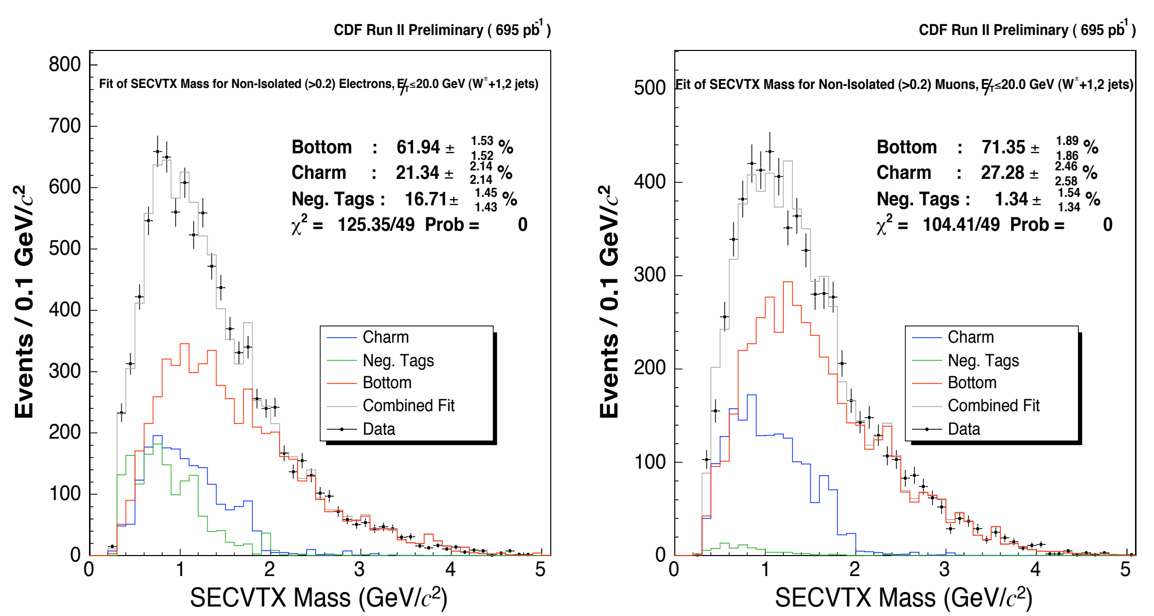


Figure 5.4: SECVTX mass fits of tagged jets in events with lepton isolation≥0.2 and $\not\!\!E_T$ ≤20.0 GeV. These shapes are taken to be a model of the SECVTX mass of QCD events, and the resulting b-fraction is applied to the QCD event estimates listed in Table 5.1. The electron QCD b-fraction is 0.62, while the muon QCD b-fraction is 0.71. The χ^2 values of the fits only take into account the statistics of the data points.

Using Eq. 5.1 the resulting b-fractions shown in Figure 5.4 are applied to the tagged

QCD event estimate of Table 5.1. The b-fractions obtained in Figure 5.4, which are the model of the b-fraction of QCD, are assigned systematic uncertainties based on differences with respect to b-fractions obtained in other regions of the $\not\!\!E_T$ versus isolation plane. The data in the $\not\!\!E_T$ $\geq$25.0 GeV, isolation$\geq$0.2 region (electron b-fraction = 0.625, muon b-fraction = 0.870) and also the data in the $\not\!\!E_T$ $\leq$20.0 GeV,isolation$\leq$0.1 region (electron b-fraction = 0.500, muon b-fraction = 0.564) is fit and half the spread from the b-fractions of Figure 5.4 is taken as the systematic on the QCD b-fraction. This gives an electron QCD b-fraction of 0.62$\pm$0.06 and a muon b-fraction of 0.71$\pm$0.08 . Applying these b-fractions to the tagged QCD events estimates gives 101.9$\pm$28.8 tagged b-jets from QCD in the 1+2 jet bins. The uncertainty quoted, which is treated as a systematic on the $W^{\pm}$+$b\bar{b}$ cross-section, is due to the 25% uncertainty of the event estimates of Table 5.1 (treated as correlated across electron and muon channels) and also due to the uncertainty of the b-fraction obtained in the non-isolated SECVTX mass fits (also treated as correlated across electrons and muons).

5.1.2 n_{MC}

To estimate the number of tagged b-jets in the 1+2 jet bins coming from the remaining backgrounds mentioned above, MC for each process is relied on. The standard selection is applied to these samples, and events in the 1 or 2 jet bins of the pretag data that have at least one b-jet (a jet matched to a b-hadron via a ΔR cone of 0.4) are counted. For these events the number of tagged b-jets appearing in the final sample is also counted. Eq. 5.2 represents the calculation of the number of background b-jets for a given process:

$$n_{\mathrm{MC}} = \sigma_{MC} \cdot \mathcal{L} \cdot \sum_{i=1}^{2} \Big(f_{1b} \cdot \epsilon_{tag1b} \cdot \mathrm{SF} + f_{2b} \cdot (\epsilon_{tag2b} \cdot \mathrm{SF} + 2 \cdot \epsilon_{2tag2b} \cdot \mathrm{SF}^2) \Big)_i \tag{5.2}$$

In Eq. 5.2 the following terms are used:

- σ_{MC}: Cross-section of process.

- $\mathcal{L}$: Luminosity to scale to.
- f_{1b}: Fraction of initial events that pass selection and have one identified b-jet
- f_{2b}: Fraction of initial events that pass selection and have two identified b-jets
- ϵ_{tag1b}: Fraction of 1b events in which the b-jet is tagged.
- ϵ_{tag2b}: Fraction of 2b events in which one of the b-jets is tagged.
- ϵ_{2tag2b}: Fraction of 2b events in which both of the b-jets are tagged.
- SF: b-tag scale factor, 0.89±0.07.

In Eq. 5.2 a sum is performed over the 1 and 2 jet bins of each MC process to get the number of tagged b-jets the process will contribute in the given luminosity. Clearly, for the 1-jet bin the contributions of the 2b category will be zero. This is essentially the procedure followed to estimate the W+HF process in the Method II analysis, but here the procedure is adding up b-jets instead of events (hence the factor of "2" in front of the ϵ_{2tag2b} term) since a b-jet cross-section is being calculated.

This procedure is applied to the MC samples listed in Table 5.2, which yields the results in Table 5.3. To obtain the numbers in Table 5.2 the calculations of Eq. 5.2 are separated out by trigger type (CEM, CMUP, CMX) to account for any differences in heavy-flavor fraction, tagging efficiencies, and luminosities.

Systematic uncertainty on these background estimates arises from the uncertainty of the b-tag scale-factor, the 6% uncertainty of the luminosity, and the inherent uncertainty of the calculation (due to MC statistics and cross-section uncertainties). Table 5.4 summarizes the various sources of uncertainty in the calculation of the number of background b-jets. The scale-factor uncertainty is treated as being correlated across trigger channels and jet bins, and the luminosity uncertainty is treated similarly.

Process	Cross-Section (pb)	Fileset	Events
$W^{\pm}(\to \tau\nu) + b\bar{b}$+0p	2.715±0.071	atop0t	250966
$W^{\pm}(\to \tau\nu) + b\bar{b}$+2p	0.283±0.012	atop2t	244986
$W^{\pm}W^{\mp}$	12.4±0.25	wtop1w	419728
$W^{\pm}Z$	3.96±0.06	wtop1z	409647
ZZ	1.58±0.02	ztopcz	412866
Single-top (s-channel)	0.88±0.05	mtopya	195928
Single-top (t-channel)	1.98±0.08	mtopta + ua_match	144525
$Z(\to \ell^+\ell^-) + b\bar{b}$+0p	0.5376±0.0008	ztop0b	148608
$Z(\to \tau\tau)$	252.0±9.0	ztop1i	925632
$t\bar{t}$	6.7±0.7	ttopkl	1846011

Table 5.2: Summary of MC sources and information needed to calculate background b-jet contributions from each. The cross-sections of $W^{\pm}(\to \tau\nu) + b\bar{b}$ are assumed to be the same as the electron/muon MC available for that process, since the actual cross-section for that MC is missing. $Z + b\bar{b}$ MC is only available for the electron channel, so the results obtained in that sample are doubled to account for any contributions from the muon channel.

Process	1-jet	2-jet	1+2 jets
$W^{\pm}(\to \tau\nu) + b\bar{b}$+0p	3.01 ± 0.21	2.05 ± 0.19	5.05 ± 0.28
$W^{\pm}(\to \tau\nu) + b\bar{b}$+2p	0.25 ± 0.02	0.54 ± 0.03	0.79 ± 0.04
$W^{\pm}W^{\mp}$	0.22 ± 0.09	0.51 ± 0.14	0.74 ± 0.17
$W^{\pm}Z$	1.79 ± 0.14	5.09 ± 0.28	6.88 ± 0.32
ZZ	0.04 ± 0.01	0.20 ± 0.04	0.23 ± 0.04
Single-top (s-channel)	3.57 ± 0.16	15.11 ± 0.60	18.69 ± 0.62
Single-top (t-channel)	13.26 ± 0.45	14.75 ± 0.49	28.01 ± 0.67
$Z(\to \ell^+\ell^-) + b\bar{b}$+0p	0.88 ± 0.09	1.55 ± 0.14	2.43 ± 0.16
$Z(\to \tau\tau)$	1.34 ± 0.68	1.34 ± 0.77	2.68 ± 1.03
$t\bar{t}$	5.86 ± 0.43	51.37 ± 3.53	57.22 ± 3.55
Total	30.22 ± 0.98	92.50 ± 3.71	122.72 ± 3.84

Table 5.3: Summary of tagged b-jets from MC sources. Uncertainties shown are those due to MC statistics and generator-level cross-section.

Source	1-jet	2-jet	1+2 jets
MC Stat./σ_{MC}	0.98	3.71	3.84
Scale Factor	2.38	9.22	11.59
Lum.	1.81	5.55	7.36
Total	30.22 ± 3.15	92.50±11.38	122.72±14.26

Table 5.4: Summary of uncertainties on the estimate of the number of tagged b-jets from various MC sources. The dominant uncertainty is due to the b-tag scale factor, followed by the 6% uncertainty of the luminosity.

CHAPTER 6

Systematic Uncertainties

Several sources of systematic uncertainty are considered for the various pieces of the cross-section. In this section the systematic uncertainties for the SECVTX mass fit of the data are described, as are those for the acceptance calculation. The sources of systematic uncertainty on the number of background *b*-jets from QCD and from MC have already been described in the preceding chapter.

6.1 SECVTX Mass Fit Systematics

Any effect that could cause the fit of the SECVTX mass to return an answer that doesn't truly reflect the real *b*-fraction of the data needs to be investigated.

6.1.1 Light-Flavor Model

Our default model for tagged light-flavor jets comes from MC. As an alternative model for light-flavor, negatively tagged jets from the data are used. A negative log-likelihood fit to the data is again performed, resulting in a *b*-fraction of $29.4\pm^{3.3}_{3.1}\%$, corresponding to $373.1\pm^{41.3}_{39.3}$ tagged *b*-jets. Half the difference in tagged *b*-jets between the two models of light-flavor, or 5.1%, is taken as the systematic on the light-flavor model. Figure 6.1 shows the SECVTX mass fit using negatively tagged jets as a model of light-flavor. Applying this same technique to our fits of non-isolated lepton data, in the QCD background calculation, produces a 1.0% systematic shift in the number of background b-jets from QCD.

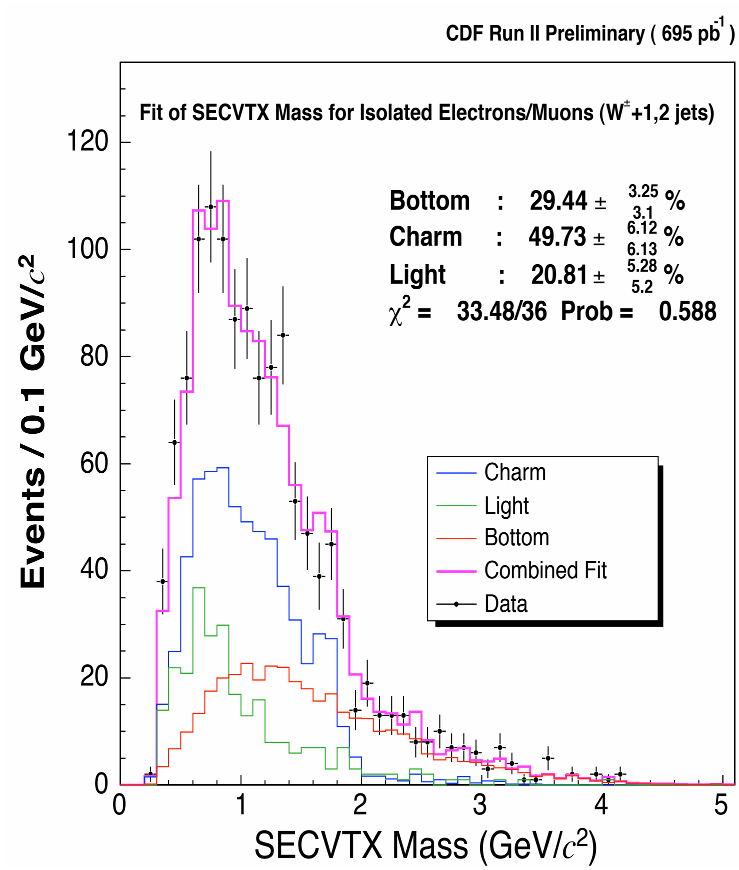


Figure 6.1: SECVTX Mass fit when using negatively tagged jets as the shape for tagged light-flavor jets.

6.1.2 MC/Data SECVTX Mass Comparison

In order to determine if our templates of SECVTX mass for b-jets in the MC resembles the distribution of b-jets in the data, we compare to an alternative data sample with very pure b-content. Double-tagged dijet events, with one identified semileptonic decay, are the data sample that can be used for this study. In these events, the two tagged jets are required to be roughly on opposite sides of the detector. The “away-jets” in these events, which are the jets opposite to the identified semileptonic decay jet, are found to have a very pure b-fraction (≈83%).

The default templates for b-jets are smeared based on differences observed in a study of these double-tagged dijet events with one identified semileptonic decay. Any differences in the SECVTX mass shapes of away-jets in the dijet data and dijet MC could be attributed to differences in modeling b-decays in the MC versus the data. The data shape in this dijet study is corrected for the small (≈17%) non-b component before the comparison with the MC b-shape is made. The default SECVTX mass templates for our $W^{\pm}$+jets analysis is convolved with the ratio of the data and MC b-shapes from the dijet study, as shown in Figure 6.2, and the SECVTX mass fit is redone with the smeared template.

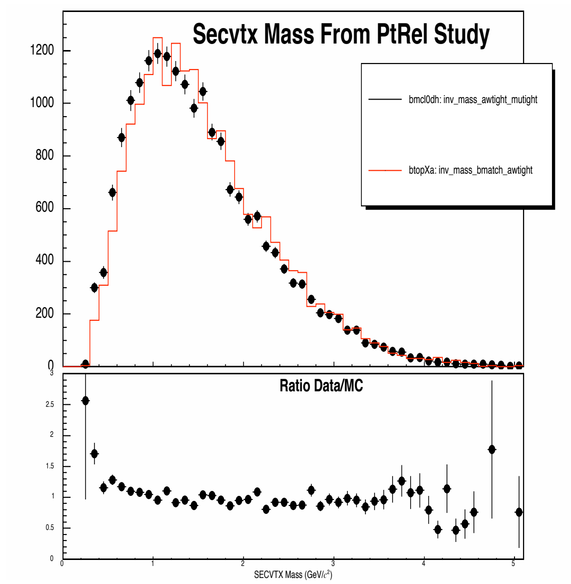


Figure 6.2: "Away"-jet SECVTX Mass shapes in double-tagged dijet data/MC with one identified semileptonic decay, and tags on opposite sides of the detector. The ratio of Data/MC shown here is used to smear the templates of SECVTX mass derived from $W^{\pm}$+jets MC, and the resulting templates are used in our fit to estimate a systematic uncertainty.

The systematic uncertainty for b-jet modeling is taken to be half the difference in b-fractions obtained when smearing with the dijet study shapes that have been corrected for 0.5× and 2× the estimated non-b contribution. This amounts to a 6.4% uncertainty on the b-fraction of the selected data sample.

Charm %	Light %	b-fraction
11%	6%	$0.349\pm^{0.037}_{0.035}$
16.5%	9%	$0.325\pm^{0.034}_{0.032}$
5.5%	3%	$0.373\pm^{0.039}_{0.037}$

Table 6.1: b-fraction obtained in the SECVTX mass fit when smearing with dijet data/MC shapes that have been corrected for the given charm/light components. The middle and bottom rows correspond to 2× and 0.5× the estimated non-b contribution to the dijet data.

6.1.3 Single/Double HF Jet Templates

It is possible that some jets in the SECVTX mass templates are actually matched to multiple HF hadrons. If the SECVTX mass shapes of these jets is significantly different than jets that are matched to only one HF hadron, then it would need to be determined if the templates had the correct proportions of both types of jet. In the MC templates of bottom and charm jets, the bottom jets are 89.2% single-b vs. 10.8% double-b, while the charm jets are 82.8% single-c vs. 17.2% double-c.

To estimate the systematic uncertainty due to possible differences between single and double HF jets, the proportions of single to double HF jets are fluctuated in the bottom and charm templates. Eq. 6.1 represents the sum of the number of single (N_1) and double (N_2) HF jets. The default templates have $x_1 = x_2 =1$.

$$x_1 \cdot \frac{N_1}{N_1 + N_2} + x_2 \cdot \frac{N_2}{N_1 + N_2} = 1 \tag{6.1}$$

The integral of the templates is kept constant while changing the relative proportions of single to double HF jets. The proportion of single HF jets is increased by changing x_1 up to 1.1 and at the same time the proportion of double HF jets must be decreased to $x_2 = 1 + \frac{N_1}{N_2}(1 - x_1)$. Similar changes are made to decrease (increase) the number of single (double) HF jets. Table 6.2 summarizes the b-fractions that result when changing the proportions of single and double HF jets.

The b-fractions appear relatively stable as the proportions of single and double HF

Single b (%)	Double b (%)	Single c (%)	Double c (%)	b-fraction (%)
89.2	10.8	82.8	17.2	32.8±3.4
98.1	1.9	82.8	17.2	33.9±3.5
80.2	19.8	82.8	17.2	31.7±3.3
89.2	10.8	91.1	8.9	33.6±3.3
89.2	10.8	74.6	25.5	32.3±3.5

Table 6.2: b-fraction obtained after varying proportions of single and double HF jets in the default templates. The top row represents the default templates.

jets are varied in the default templates. Half of the maximum deviation from the default templates is quoted as the systematic. This corresponds to a difference of 1.1% in the b-fraction and a change of 3.4% in the number of tagged b-jets returned by the fit. Varying the single and double HF jet proportions in the QCD fit has a very small effect on the resulting b-fraction, so no systematic is assigned to that estimate.

6.2 Acceptance Systematics

Several sources of systematic uncertainty have been considered on the calculated acceptance numbers. Table 6.3 summarizes the considered systematics, which are explained in more detail in the subsequent sections.

Trigger	Acceptance	Jet Energy	Collision Q^2	PDF	Shower	Total Sys.
CEM	0.505	5.7%	1.0%	0.5%	3.8%	0.035 (6.9%)
CMUP	0.301	6.0%	2.3%	1.2%	3.8%	0.023 (7.6%)
CMX	0.127	6.3%	2.4%	2.2%	3.8%	0.010 (8.1%)

Table 6.3: Acceptance and systematic uncertainties for each trigger channel.

6.2.1 Jet Energy Scale

In selecting events from the MC and data the CDF Jet Energy Correction package (`jetCorr05b`) is used to apply a correction factor to the raw transverse energy of the jets[45]. Events are classified based on the number of jets with L5 corrected transverse energy greater than 20 GeV, and detector pseudorapidity less than 2.0. Since the acceptance definition is directly dependent on how many events pass all selection cuts and end up in the 1 or 2 jet bins, then clearly any uncertainty in the jet energy correction factor will propagate into uncertainty in the acceptance.

The uncertainty in the acceptance due to jet energy corrections is estimated using the systematic uncertainty function provided by the CDF Jet Energy and Resolution group[45]. This function can return the correction factor plus its $\pm\sigma$ uncertainties, which are the quadrature sum of the uncertainties on all the jet correction levels up through L7. Selection code is applied three times to a given MC sample; once with the default jet correction factors applied, once with the default correction factor plus its one sigma uncertainty, and once with the default factor minus its one sigma uncertainty. Figure 6.3 shows the acceptance for each sample, and a bounding box which represents the spread of the $\pm\sigma$ variation in jet energy correction factor.

The cross-section weighted acceptance is calculated for each of the three scenarios, using the acceptance values in Table 6.4, and the largest deviation from the central acceptance value for each trigger is taken to be the systematic uncertainty due to jet energy corrections. Table 6.5 summarizes the final systematic uncertainty due to jet energy corrections for each trigger.

Process	Trigger	$\mathcal{A}_{\text{jet}}$		$\mathcal{A}_{\text{selection}}$	
		$+\sigma$ JES	$-\sigma$ JES	$+\sigma$ JES	$-\sigma$ JES
$W^\pm(\to e^\pm\nu)+b\bar{b}$+0p	CEM	1.026±0.006	0.880±0.005	0.535±0.004	0.539 ±0.004
$W^\pm(\to e^\pm\nu)+b\bar{b}$+1p	CEM	0.985±0.006	0.915±0.005	0.519±0.004	0.519±0.004
$W^\pm(\to e^\pm\nu)+b\bar{b}$+2p	CEM	0.909±0.007	1.027±0.008	0.484±0.005	0.488±0.005
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+0p	CMUP	1.028±0.006	0.881±0.005	0.318±0.003	0.321 ±0.003
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+1p	CMUP	0.987±0.006	0.917±0.005	0.307±0.003	0.306 ±0.003
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+2p	CMUP	0.914±0.007	1.033±0.008	0.285±0.004	0.287 ±0.003
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+0p	CMX	1.028±0.006	0.881±0.005	0.134±0.002	0.134 ±0.002
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+1p	CMX	0.987±0.006	0.917±0.005	0.130±0.002	0.129 ±0.002
$W^\pm(\to \mu^\pm\nu)+b\bar{b}$+2p	CMX	0.915±0.007	1.034±0.008	0.123±0.002	0.125 ±0.002

Table 6.4: Acceptance values for main $W^\pm+b\bar{b}$ MC samples when using $\pm\sigma$ Jet Energy Systematic corrections.

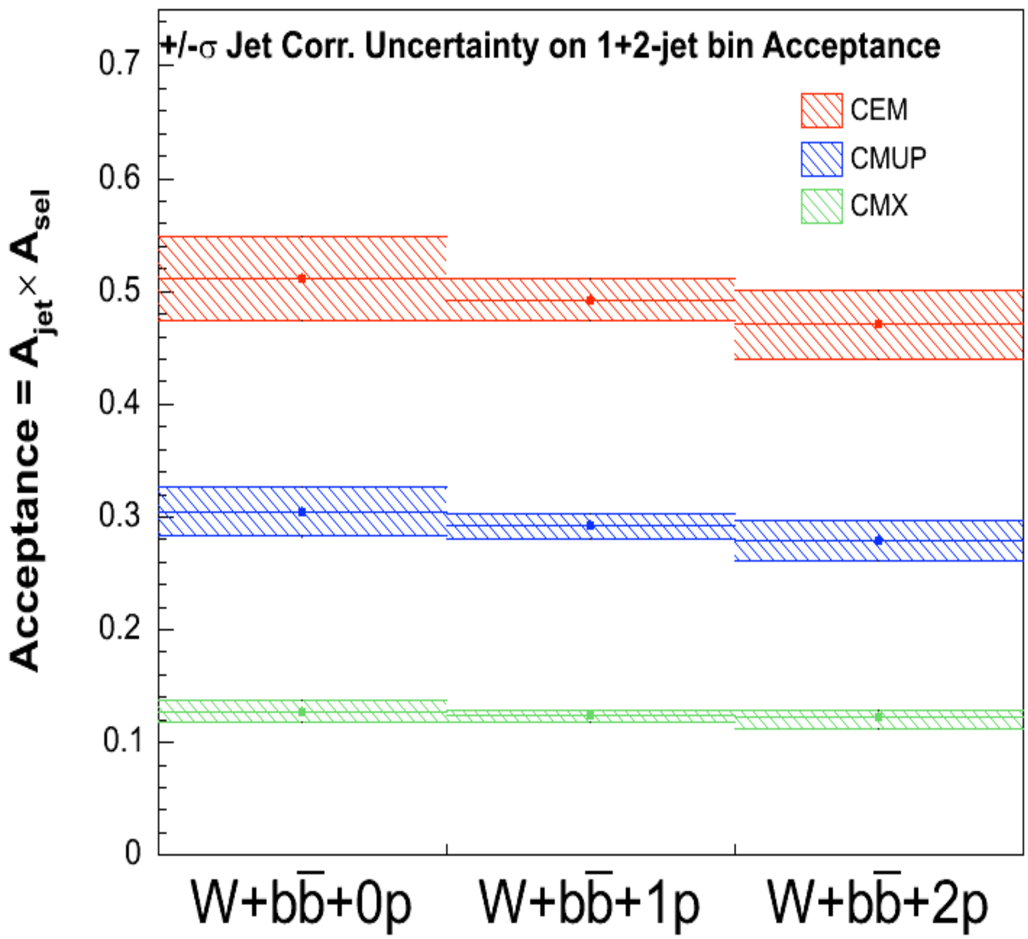


Figure 6.3: Jet Energy Uncertainty for each trigger channel and $W^{\pm}$+$b\bar{b}$+N parton sample.

Jet. Corr Variation	CEM	CMUP	CMX
$+\sigma$	0.534	0.318	0.134
default	0.505	0.301	0.127
$-\sigma$	0.476	0.283	0.119
Systematic	5.7%	6.0%	6.3%

Table 6.5: Cross-section weighted acceptances after varying jet-energy corrections. The systematic is the larger of the $\pm\sigma$ deviations from the default acceptance.

6.2.2 Q^2

All of the $W^{\pm}$+$b\bar{b}$ MC used in the calculation of the acceptance has been generated at a Q^2 value of $M_W^2 + \sum m_T^2$, where $m_T^2 = m^2 + p_T^2$ for all final-state partons (excluding any W decay products). The value of Q^2 determines the strength of the QCD coupling constant,

α_s. The larger the value of Q^2 the weaker the coupling of particles subject to the strong force will be, and so it is perhaps expected that the cross-sections for $W^{\pm}$+$b\bar{b}$+N partons will diminish with increasing Q^2 since it becomes increasingly harder to emit the gluons needed to form these final states. Alpgen allows the functional form of the Q^2 value to be modified, and also allows a multiplicative factor to be factored into the Q^2 value.

To estimate this effect several small samples were generated in Alpgen v2.1 with different values for the Q^2 functional form ($Q^2 \propto M_W^2$), and multiplicative factor, as listed in Table 3.4. The acceptance for each Q^2 subsample, parton multiplicity, and trigger type is plotted in Figure 6.4.

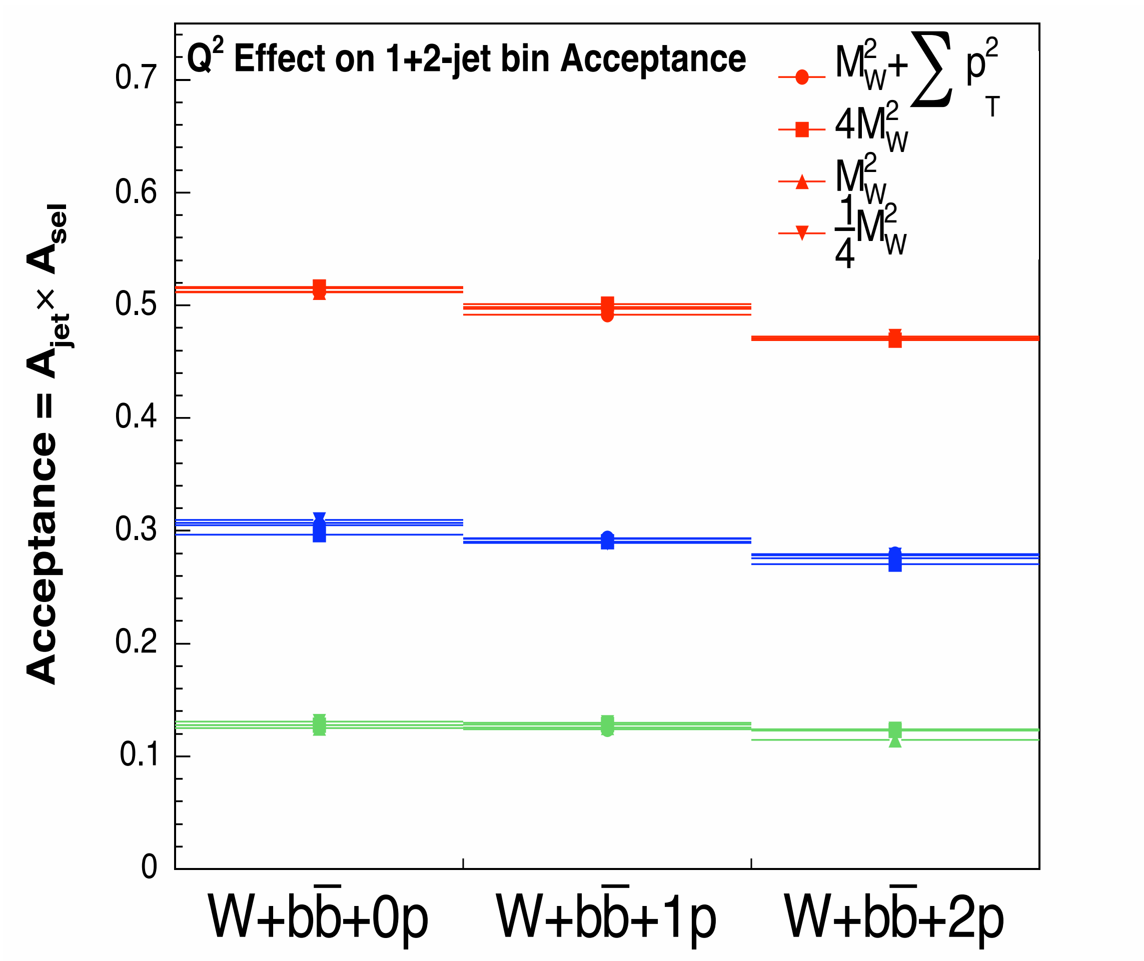


Figure 6.4: Q^2 effect on acceptance. The red histograms are for CEM acceptance, while the blue are for CMUP, and the green are for CMX.

Process	Q^2	Trigger	$\mathcal{A}_{\text{jet}}$	$\mathcal{A}_{\text{selection}}$	$\mathcal{A} = \mathcal{A}_{\text{jet}} \cdot \mathcal{A}_{\text{sel.}}$	$\mathcal{A}_{\text{trig}}$
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	$4M_W^2$	CEM	0.952±0.006	0.542±0.004	0.516±0.005	0.510±0.007
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p			0.956±0.005	0.524±0.004	0.501±0.005	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p			0.981±0.008	0.478±0.005	0.469±0.006	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	M_W^2	CEM	0.951±0.010	0.538±0.007	0.512±0.009	0.506±0.010
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p			0.955±0.010	0.522±0.006	0.499±0.008	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p			0.967±0.014	0.485±0.009	0.469±0.011	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	$\frac{1}{4}M_W^2$	CEM	0.951±0.010	0.542±0.007	0.515±0.009	0.508±0.010
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p			0.961±0.010	0.517±0.006	0.497±0.008	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p			0.986±0.014	0.479±0.008	0.472±0.010	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	$4M_W^2$	CMUP	0.948±0.011	0.313±0.005	0.297±0.006	0.294±0.006
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.961±0.010	0.302±0.005	0.290±0.006	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.959±0.013	0.282±0.006	0.270±0.007	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	M_W^2	CMUP	0.954±0.010	0.322±0.005	0.307±0.006	0.302±0.006
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.958±0.010	0.306±0.005	0.293±0.006	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.981±0.014	0.291±0.006	0.286±0.007	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	$\frac{1}{4}M_W^2$	CMUP	0.955±0.010	0.324±0.005	0.309±0.006	0.302±0.007
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.954±0.010	0.303±0.005	0.289±0.006	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.990±0.014	0.281±0.006	0.278±0.007	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	$4M_W^2$	CMX	0.947±0.011	0.135±0.003	0.128±0.003	0.128±0.003
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.962±0.010	0.133±0.003	0.128±0.003	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.960±0.014	0.128±0.004	0.123±0.004	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	M_W^2	CMX	0.954±0.010	0.131±0.003	0.125±0.003	0.125±0.003
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.960±0.010	0.131±0.003	0.126±0.003	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.981±0.014	0.117±0.004	0.115±0.004	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	$\frac{1}{4}M_W^2$	CMX	0.955±0.010	0.137±0.003	0.131±0.003	0.130±0.003
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+1p			0.952±0.010	0.136±0.003	0.130±0.003	
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+2p			0.992±0.014	0.125±0.004	0.124±0.004	

Table 6.6: Acceptance values for the Q^2 systematic $W^{\pm}$+$b\bar{b}$ MC samples.

The acceptances in the different Q^2 samples are listed in Table 6.6. Since the cross-sections for the various $W^{\pm}$+$b\bar{b}$+Np samples changes significantly with the value of Q^2, the weights that must be applied (as defined in Eq. 4.5) when combining acceptances also change. These differences are taken into account and the correct cross-section weighted acceptance for each trigger channel and Q^2 value is calculated. Whichever of the Q^2 samples has the largest difference between the weighted acceptance and the default sample acceptance will be quoted as the systematic due to the Q^2 value. For the CEM trigger this amounts to a 1% systematic, while for the muon triggers it is closer to 2%.

6.2.3 Parton Distribution Function

A reweighting technique is used to estimate the uncertainty in the acceptance caused by the choice of PDF used in the generated events. To accomplish this 46 PDF sets, listed in Table 6.7, are used to investigate this systematic uncertainty in the $W^{\pm}$+$b\bar{b}$ acceptance.

index	PDF	Note
0	CTEQ5L	Default in Gen. 5 MC, α_s=0.118
1	MRST72	
2	MRST75	
3	CTEQ6L	LO, α_s=0.118
4	CTEQ6L1	LO, α_s=0.130
5	CTEQ6M	NLO
5+i	CTEQ6M	i =1, ±PDF eigenvectors
.		
.		
.		
45	CTEQ6M	i =40, ±PDF eigenvectors

Table 6.7: Sets used in the investigation of the systematic uncertainty due to PDF issues. The first six entries correspond to various permutations of LO/NLO, α_s value, source (CTEQ/MRST), while the last 40 entries are the $\pm\sigma$ variations of the CTEQ6M set.

The uncertainty in the acceptance due to PDFs receives contributions from several sources: LO vs. NLO calculations, α_s used, MRST vs CTEQ (the two main collaborations responsible for producing PDFs), and inherent uncertainty of the PDF used in generation[60, 61]. Given a PDF set B, a weight is calculated relative to a different PDF set, A, by dividing the product of the value of the parton distribution functions, $f(x_1, x_2, Q)$, of A and B for the colliding partons of the proton and antiproton in an individual event:

$$w^{(A)}(x_1, x_2, Q) \equiv \frac{f_p^{(A)}(x_1, Q) \cdot f_{\overline{p}}^{(A)}(x_2, Q)}{f_p^{(B)}(x_1, Q) \cdot f_{\overline{p}}^{(B)}(x_2, Q)} \tag{6.2}$$

The sum of all weights for the L5 b-jets in the MC sample passing selection divided by the sum of all weights for the initial hadron b-jets in the MC (after the OBSV cut, and HEPG cuts) gives the acceptance that should be expected if the samples had been generated with set A instead of set B, as shown in Eq. 6.3:

$$\mathcal{A}^{(A)} = \frac{\sum_{i=1}^{n_{b-\text{jet}}(L5)} w_i^{(A)}(x_1, x_2, Q)}{\sum_{i=1}^{n_{b-\text{jet}}(Had.)} w_i^{(A)}(x_1, x_2, Q)} \tag{6.3}$$

To estimate the uncertainty due to the various contributions (LO vs. NLO, α_s, CTEQ vs. MRST, ±eigenset) Eq. 6.3 is used to estimate the acceptance for a given PDF set relative to another set. For instance, to investigate the differences caused by using a LO PDF versus a NLO PDF CTEQ6L is used as set B and CTEQ6M as set A. To investigate potential differences between CTEQ and MRST, CTEQ5L is used as B and MRST72 as A. To investigate differences caused by the value of α_s used in the PDF, CTEQ6L (MRST72) is used as B and CTEQ6L1 (MRST75) as A. The quadrature sum of these two is taken as the acceptance uncertainty coming from the value of α_s in the PDF.

The 40 positive/negative CTEQ6M eigenvectors are used to estimate the uncertainty inherent in the PDFs themselves. Each positive/negative pair (referred to as an *eigenset*) represent the positive/negative variation in one parameter of the CTEQ6M PDF. For each eigenset, the value of Eq. 6.3 is calculated assuming that the default CTEQ6M PDF (index 5 in Table 6.7) is B. For each eigenset, the positive (negative) variation eigenvector's acceptance uncertainty is added in quadrature with the uncertainties coming from all the other positive (negative) eigenvectors. One exception is for the situation where both members of the eigenset give an acceptance uncertainty of the same sign, in which case the mean value of the quadrature sum of these two is added in quadrature with the like-signed part of the total uncertainty. Figure 6.5 shows this procedure for all $W^{\pm}+b\bar{b}$ samples.

Table 6.8 summarizes the various PDF uncertainties in the individual $W^{\pm}+b\bar{b}$ MC

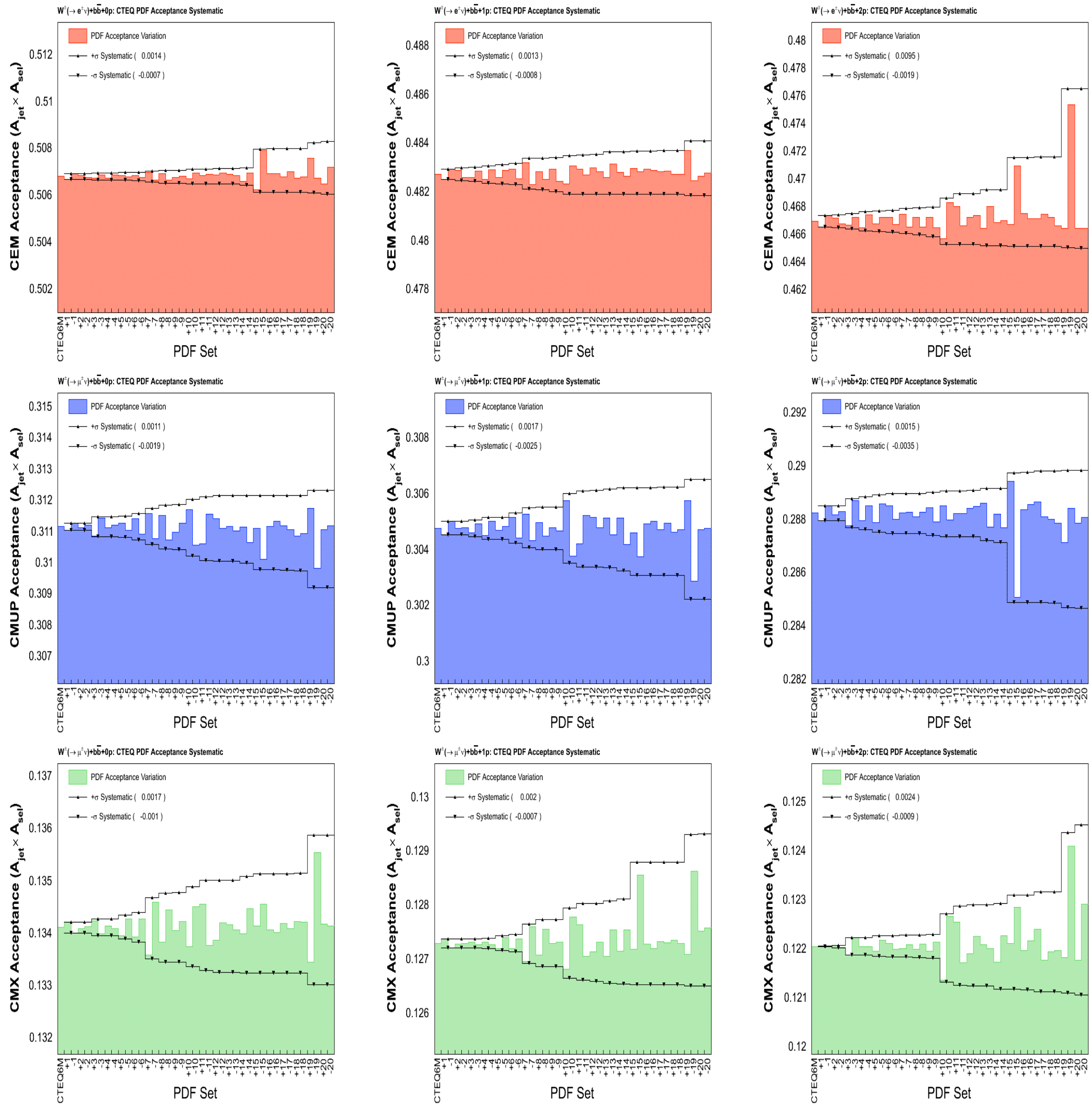


Figure 6.5: PDF reweighting for CTEQ6M eigensets on all $W^{\pm}{+}b\bar{b}$ MC.

samples. To calculate the total systematic due to PDFs, the quadrature sum of the individual PDF uncertainties of each sample (MRST vs. CTEQ, α_s, etc...) is calculated first. For each trigger type the cross-section weighted sum over the different samples is performed to calculate the final PDF uncertainty, as represented in Eq. 6.4. For the final PDF acceptance uncertainty the larger of the uncertainties in Table 6.9 will be taken rather than quote an asymmetric systematic.

Sample	Trig.	LO vs. NLO	±eigen	α_s	MRST vs. CTEQ	Tot.
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+0p	CEM	0.0001	0.0015 0.0008	0.0010	0.0003	0.0018 0.0013
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+1p	CEM	0.0012	0.0014 0.0009	0.0003	0.0011	0.0022 0.0019
$W^{\pm}(\rightarrow e\nu) + b\bar{b}$+2p	CEM	0.0003	0.0096 0.0020	0.0002	0.0036	0.0102 0.0041
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+0p	CMUP	0.0016	0.0012 0.0020	0.0002	0.0026	0.0033 0.0036
	CMX	0.0008	0.0018 0.0011	0.0005	0.0016	0.0025 0.0021
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+1p	CMUP	0.0008	0.0018 0.0025	0.0013	0.0023	0.0033 0.0037
	CMX	0.0015	0.0020 0.0008	0.0005	0.0021	0.0033 0.0027
$W^{\pm}(\rightarrow \mu\nu) + b\bar{b}$+2p	CMUP	0.0015	0.0016 0.0036	0.0014	0.0015	0.0030 0.0044
	CMX	0.0026	0.0025 0.0010	0.0009	0.0018	0.0041 0.0035

Table 6.8: Summary of PDF Uncertainties. For each sample, we quote a positive and negative total systematic due to the asymmetry possible in the ±eigenset uncertainty. For our final result we take the larger of the total positive and negative systematic for a sample to be the PDF systematic for that sample.

$$\delta(\text{PDF})_{\text{trig.}} = \sum_{\text{i}=0}^{2} \left(w_{\text{i}} \cdot \sqrt{\delta(\text{CTEQ/MRST})^2 + \delta(\alpha_s)^2 + \delta(\text{LO/NLO})^2 + \delta(\pm\text{eigen})^2} \right)_{\text{trig.}} \tag{6.4}$$

-	CEM	CMUP	CMX
+Systematic	0.0023 (0.5%)	0.0033 (1.1%)	0.0028 (2.2%)
−Systematic	0.0016 (0.3%)	0.0037 (1.2%)	0.0023 (1.8%)

Table 6.9: Final PDF Uncertainties obtained by applying Eq. 6.4 to the numbers in Table 6.8.

6.2.4 Shower

Samples of Alpgen+Pythia $W^{\pm}$+$b\bar{b}$ MC were generated to check the dependence of the acceptance on the shower evolution program. Any difference with the Alpgen+Herwig derived acceptance is treated as a systematic effect. A 3.8% systematic shift was found in the acceptance values. Table 6.10 summarizes the acceptance values calculated in the

Alpgen+Pythia samples. These samples were only generated in the electron channel, so the resulting systematic of 3.8% is taken for the muon channel as well, and treated as correlated across lepton channels.

Process	$\mathcal{A}_{\text{jet}}$	$\mathcal{A}_{\text{selection}}$	$\mathcal{A} = \mathcal{A}_{\text{jet}} \cdot \mathcal{A}_{\text{sel}}$	$\mathcal{A}_{\text{trig.}}$
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	0.957±0.010	0.514±0.006	0.492±0.008	0.486±0.009
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+1p	0.957±0.010	0.488±0.006	0.467±0.008	
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+2p	1.014±0.015	0.452±0.009	0.458±0.011	

Table 6.10: Acceptance values calculated in Alpgen+Pythia $W^{\pm}$+$b\bar{b}$ samples.

CHAPTER 7

Results

With all of the relevant pieces calculated, the $W^{\pm}$+$b\bar{b}$ b-jet cross-section can now be calculated and comparisons to theory can be made. In this section all of the numbers are collected and the final cross-section is reported.

7.1 Measured Value

Table 7.1 summarizes the numbers used in the calculation of the luminosity-weighted denominator of the cross-section. The uncertainties quoted on the acceptance numbers are the quadratic sum of all the various categories of systematics considered. In the final calculation of the "Denom." value of Table 7.1 the different categories of systematic uncertainty (PDF, Jet-Energy, etc...) on the acceptances are treated as being correlated across trigger channels. The efficiency term for tagging a b-jet, $\epsilon_{b-\mathrm{tag}}$, is not included in this calculation.

-	CEM	CMUP	CMX
ϵ_{common} (z_0)	0.9555±0.0004±0.003		
$\epsilon_{trig.}$	0.9755±0.0055	0.9157±0.0031	0.9623±0.0028
f_{id}	0.9810±0.0030	$0.9472\pm^{0.0034}_{0.0043}$	$1.0014\pm^{0.0039}_{0.0178}$
$\mathcal{A}$	0.505±0.029	0.301±0.020	0.127±0.009
$\mathcal{L}$	695.5±41.7 pb^{-1}	695.5±41.7 pb^{-1}	682.1±40.9 pb^{-1}
Denom.	**574.4±40.3(sys.)±34.5(lum.) pb^{-1}**		

Table 7.1: Final numbers needed for the calculation of the $W^{\pm}$+$b\bar{b}$ cross-section denominator.

Finally, the numbers in Table 7.2 are used to calculate the $W^{\pm}$+$b\bar{b}$ cross-section and its uncertainties. Using Eq. 4.2, the cross-section is calculated to be

$$\textbf{Measured}: \ \sigma_{W^{\pm}b\bar{b}} \times \mathrm{BR}(W^{\pm} \rightarrow \ell^{\pm}\nu) = 0.90 \pm 0.20(\mathrm{stat.}) \pm 0.26(\mathrm{syst.})\mathrm{pb} \qquad (7.1)$$

where the first uncertainty is statistical and the second is systematic. Table 7.3 summarizes all of the sources of systematic uncertainty in the analysis and their impact on the final cross-section number.

Variable	Value	Stat.	Sys.
n_{fit}	415.7	43.2	48.9
n_{MC}	122.7	-	14.3
n_{QCD}	102.6	-	28.8
$\epsilon_{b-\mathrm{tag}}$	0.37	-	0.029
$\epsilon \cdot \mathcal{A} \cdot \mathcal{L}$ (pb^{-1})	574.4	-	53.1
$\sigma_{W^{\pm}b\bar{b}} \times \mathrm{BR}(W^{\pm} \rightarrow \ell^{\pm}\nu)$ (pb)	0.90	0.20	0.26

Table 7.2: Final numbers needed for the calculation of the $W^{\pm}$+$b\bar{b}$ cross-section.

Variable	Source	$\Delta\sigma_{W^{\pm}b\bar{b}}(\%)$
n_{fit}	Dijet Mass Study	13.9
	Light-Flavor Model	11.1
	Double HF	7.4
	MC Statistics	2.2
n_{QCD}	MET vs. Isol.	13.5
	QCD Model	5.3
	Dijet Mass Study	3.7
	Light-Flavor Model	0.5
n_{MC}	b-tag Scale Factor	6.0
	Lum.	3.9
	MC Statistics	2.0
$\epsilon_{b-\mathrm{tag}}$	b-tag Scale Factor	8.7
Denominator	Luminosity	6.0
	Jet Energy	5.9
	Herwig/Pythia	3.8
	PDF	3.5
	Collision Q^2	1.6
	Trig. Efficiencies	0.33
	Lepton ID	0.33
	z_0 Efficiency	0.31
Total		28.8

Table 7.3: Final systematic uncertainties in the calculation of the $W^{\pm}$+$b\bar{b}$ cross-section.

7.2 Theoretical Predictions

The cross-section defined in Section 4.1 is that for the number of b-jets in $W^{\pm}+b\bar{b}$ events where there are 1 or 2 hadron jets with ($E_T \geq$20.0 GeV, $|\eta| \leq$2.0), and the $W^{\pm}$ decays leptonically to a charged lepton with ($p_T \geq$20.0 GeV/c, $|\eta| \leq$2.0) and a neutrino with $p_T \geq$25.0 GeV/c. We will use Alpgen (with MLM matching applied) interfaced with the Herwig parton shower program to calculate a theoretical prediction. To estimate the theoretical value of this cross-section, the Alpgen MC samples generated with much looser requirements are pared down until they are left with just the kinematic region chosen in the definition.

For all of the samples listed in Table 3.4 the number of hadron b-jets, in events that have 1 or 2 hadron jets and also have the $W^{\pm}$ decay products passing the chosen cuts at the HEPG level, is divided by the number of generated events. Multiplying the generator level event cross-section by this fraction gives the desired b-jet cross-section for each sample, as shown in Table 7.4. These b-jet cross-sections are summed over the various $W^{\pm}+b\bar{b}+N$parton samples for each Q^2 value to arrive at the theoretical value of

$$\textbf{Theoretical}: \ \sigma_{W^{\pm}+b\bar{b}} \times \mathrm{BR}(W^{\pm} \to \ell^{\pm}\nu) = 0.74 \pm 0.18(\mathrm{syst.})\mathrm{pb} \qquad (7.2)$$

Table 7.5 lists the theoretical b-jet cross-section for each of the $W^{\pm}+b\bar{b}+N$parton samples used to arrive at the final value of Eq. 7.2.

The samples generated at $Q^2 = M_W^2 + \sum m_T^2$ are used as the central value for the theoretical prediction since these are also the samples used to calculate the central value of the acceptance. The uncertainty quoted on the theory value is due to the spread in b-jet cross-sections of the different Q^2 samples.

Process	Fileset	Q^2	N_{gen}	$n_{b-\text{jets}}$	$\sigma' \times$ BR (pb)	w_i
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop0x`	$M_W^2 + \sum m_T^2$	371914	60897	0.519±0.017	0.701±0.008
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop1x`		376566	66925	0.184±0.004	0.249±0.007
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop2x`		376136	35503	0.037±0.002	0.050±0.003
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop3x`	$4M_W^2$	377204	63549	0.432±0.014	0.700±0.008
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop4x`		370246	67752	0.154±0.003	0.250±0.007
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop5x`		376119	35735	0.031±0.002	0.050±0.003
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop6x`	M_W^2	129283	21920	0.556±0.021	0.690±0.010
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop7x`		129292	23539	0.207±0.006	0.257±0.009
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop8x`		122144	11280	0.043±0.003	0.053±0.004
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop9x`	$\frac{1}{4}M_W^2$	129164	22203	0.742±0.029	0.680±0.011
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop10x`		129290	23449	0.293±0.010	0.269±0.010
$W^{\pm}(\rightarrow e^{\pm}\nu) + b\bar{b}$+0p	`mtop11x`		127780	10545	0.056±0.003	0.051±0.003
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop0y`	$M_W^2 + \sum m_T^2$	376928	61742	0.519±0.017	0.704±0.008
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop1y`		376462	66309	0.182±0.003	0.247±0.007
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop2y`		374353	35146	0.036±0.002	0.049±0.004
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop3y`	$4M_W^2$	108751	18535	0.436±0.016	0.704±0.009
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop4y`		113270	20513	0.152±0.003	0.246±0.007
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop5y`		127998	11992	0.031±0.001	0.050±0.002
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop6y`	M_W^2	129034	22138	0.561±0.022	0.692±0.010
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop7y`		129286	23542	0.207±0.007	0.255±0.010
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop8y`		126927	11714	0.043±0.002	0.053±0.003
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop9y`	$\frac{1}{4}M_W^2$	129313	22560	0.753±0.028	0.679±0.015
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop10y`		129290	23446	0.293±0.010	0.264±0.010
$W^{\pm}(\rightarrow \mu^{\pm}\nu) + b\bar{b}$+0p	`mtop11y`		127407	11846	0.063±0.004	0.057±0.004

Table 7.4: Theory b-jet cross-section numbers, derived by scaling down the event cross-sections of Table 3.4 by $n_{b-\text{jets}}/N_{\text{gen.}}$. The last column lists the weights, w_i, to be used in the Acceptance calculation.

Q^2	Electrons: $\sigma'_{W^{\pm}+b\bar{b}}$(pb)	Muons: $\sigma'_{W^{\pm}+b\bar{b}}$(pb)
$M_W^2 + \sum m_T^2$	0.740±0.018	0.737±0.017
$4\mathrm{M}_W^2$	0.617±0.015	0.619±0.016
M_W^2	0.806±0.022	0.811±0.023
$\frac{1}{4}\mathrm{M}_W^2$	1.091±0.031	1.109±0.030
Final	0.740±0.176 pb	0.737±0.186 pb

Table 7.5: Final b-jet cross-sections obtained by summing the values in Table 7.4 over all multiplicitites for the various Q^2 samples. The final uncertainty quoted is half of the maximum deviation from the default samples.

CHAPTER 8

Conclusions

The measured cross-section of 0.90±0.20±0.26 pb compares well with the theoretical prediction of 0.74±0.18 pb. The cross-section is defined to be proportional to the number of b-jets in $W^{\pm}+b\bar{b}$ events where there are 1 or 2 hadron jets with ($E_T \geq$20.0 GeV, $|\eta| \leq$2.0), and the $W^{\pm}$ decays leptonically to a charged lepton with ($p_T \geq$20.0 GeV/c, $|\eta| \leq$2.0) and a neutrino with $p_T \geq$25.0 GeV/c.

The leading sources of uncertainty on the measured result are those from MC b-jet SECVTX mass modeling study, and those related to modeling the QCD component of the background. Both of these have potential remedies that may reduce the systematic uncertainty on the cross-section. For the MC SECVTX mass modeling systematic, CDF is switching to a more updated b-hadron fragmentation model (EVTGEN) which should give better agreement with the data[62]. For the estimation of the number of tagged b-jets from QCD, reducing the uncertainty of the b-fraction of tagged jets from QCD could potentially be achieved with more sophisticated methods of estimating the value of this quantity in the signal region[19]. Parameterizing the "light" shapes (either from negative tags or from light-flavor MC) with a smooth function could also help reduce some of the systematic uncertainty.

Other potential improvements in the experimental method could come from optimizing selection cuts to further discriminate against QCD while keeping as much of the $W^{\pm}+b\bar{b}$ signal as possible (by either raising the $\not{E}_T$ cut or adding in additional cuts). CDF is in the process of making a whole new slate of MC samples, that will utilize the MLM matching routine to a much greater degree. This should help improve the modeling of the data and hopefully the reliability of this method.

NLO comparisons using programs such as MCFM would be very useful for the study of $W^{\pm}+b\bar{b}$[63]. Other NLO calculations of $W^{\pm}+b\bar{b}$ that take into account the mass of the b-quark, which MCFM does not do, also are starting to become available[64]. CDF jets would need to be corrected back to the parton level to facilitate comparisons with these NLO predictions, which are currently only available at the parton level. Correcting CDF jets back to the parton level introduces much uncertainty and model dependence into the measured result, so the best method of comparing with these NLO programs is still under investigation.

The theoretical value of 0.74±0.18 pb suffers from large uncertainty due to the choice of Q^2 used in the generation of our MC samples. NLO calculations of $W^{\pm}+b\bar{b}$ should not be as sensitive to Q^2 effects as LO calculations, which should allow for more solid conclusions to be drawn from the analysis. Previous studies have shown that the heavy-flavor predictions of Method II have reduced dependence on scale choices at NLO[65].

The measurement described in this thesis is the first ever of the $W^{\pm}+b\bar{b}$ cross-section at a hadron collider experiment, and the results are very encouraging. The simple three-component fit of the SECVTX mass of tagged jets allows for the true b-content of the tag sample to be determined, and background subtraction leaves just the $W^{\pm}+b\bar{b}$ contribution to the data. The methods used in this analysis are robust because they do not require the use of any $W^{\pm}+b\bar{b}$ MC, until the acceptance calculation, and also the large uncertainties in the LO predictions of processes such as $W^{\pm}+c\bar{c}$ and $W^{\pm}+c$ do not have any effect on our results as they do for the Method II analysis. CDF can expect to receive almost 6$\times$ the data used in this analysis by the end of Run II, allowing for the statistical precision of this technique to be significantly improved. The techniques developed in this analysis, which try to minimize the reliance on LO W+heavy flavor MC, should continue to be improved and used as more data is collected.

Refinements to this method can bring the uncertainties down to the level where this result can be used to better constrain $W^{\pm}+b\bar{b}$ in Higgs and single-top searches. Discovering the Higgs boson would be an incredible triumph for particle physicists and the Standard Model. If the Higgs turns out to have a relatively light mass, as current limits indicate, then understanding the role $W^{\pm}+b\bar{b}$ plays in the data will be a crucial part of this discovery.

BIBLIOGRAPHY

[1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books, Cambridge, MA, 1995. 3

[2] G. Kane, *Modern Elementary Particle Physics*, Perseus Books, Cambridge, MA, 1993. 3

[3] W. Greiner and A. Schäefer, *Quantum Chromodynamics*, Springer-Verlag, Berlin, 1 ed., 1995. 3

[4] W. Greiner and B. Müller, *Gauge Theory of Weak Interactions*, Springer-Verlag, Berlin, 2 ed., 1996. 3

[5] V. D. Barger and R. J. N. Phillips, *Collider Physics*, Perseus Publishing, Cambridge, MA, 1996. 3

[6] H. D. Politzer, *Asymptotic freedom: An approach to strong interactions*, Phys. Rept., 14 (1974), pp. 129–180. 4

[7] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. 1*, Phys. Rev., D8 (1973), pp. 3633–3652. 4

[8] S. Weinberg, *A model of leptons*, Phys. Rev. Lett., 19 (1967), pp. 1264–1266. 4

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

[10] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys., 22 (1961). 4

[11] W.-M. Yao et al., *Review of Particle Physics*, Journal of Physics G, 33 (2006), pp. 1+. 6

[12] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett., 81 (1998), p. 15. 6

[13] T. S.-K. Collaboration, *Constraints on neutrino oscillations using 1258 days of super-kamiokande solar neutrino data*, Physical Review Letters, 86 (2001), p. 5656. 6

[14] *Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam*, (2006). 6

[15] C. Athanassopoulos et al., *Evidence for neutrino oscillations from muon decay at rest*, Phys. Rev. C, 54 (1996), pp. 2685–2708. 6

[16] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett., 13 (1964), pp. 508–509. 7

[17] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for higgs boson decays in the standard model and its supersymmetric extension*, Computer Physics Communications, 108 (1998), p. 56. 8

[18] T. L. C. A. collaboration, D. Collaboration, L. Collaboration, O. Collaboration, and the LEP Electroweak Working Group, *A combination of preliminary electroweak measurements and constraints on the standard model*, tech. report, hep-ex/0511027, 2005. 8

[19] B. Cooper, A. Messina, and D. Waters, *Measurement of the w($\rightarrow e\nu$)+ $\geq$n jet differential cross-section using 320pb^{-1} of CDF Run II data*, Tech. Rep. 8046, The CDF II Collaboration, (2006). 10, 121

[20] G. 't Hooft and M. J. G. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys., B44 (1972), pp. 189–213. 11

[21] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP, 0207 (2002), p. 012. 11

[22] Y. L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e+ e- annihilation by perturbation theory in quantum chromodynamics. (in russian)*, Sov. Phys. JETP, 46 (1977), pp. 641–653. 11

[23] V. N. Gribov and L. N. Lipatov, *e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys., 15 (1972), pp. 675–684. 11

[24] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys., B126 (1977), p. 298. 11

[25] The CDF Collaboration, *Single-top-quark search with 0.7 fb^{-1}*, Tech. Rep. 8185, The CDF II Collaboration, (2006). 13

[26] ——, *Search for standard model higgs boson production in association with w$^{\pm}$ boson at CDF with 1 fb^{-1}*, Tech. Rep. 8390, The CDF II Collaboration, (2006). 13

[27] M. L. Mangano, *Production of w plus heavy quark pairs in hadronic collisions*, Nucl. Phys., B405 (1993), pp. 536–554. 14

[28] M. L. Mangano, M. Moretti, and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: W b anti-b + (n)jets as a case study*, Nucl. Phys., B632 (2002), pp. 343–362. 14

[29] R. K. Ellis and S. Veseli, *Strong radiative corrections to w b bbar production in p pbar collisions*, Physical Review D, 60 (1999), p. 011501. 14

[30] *Measurement of the t anti-t production cross section in p anti-p collisions at s**(1/2) = 1.96-tev*, (2006). 15

[31] D. McGinnis et al., *The tevatron run II handbook*, tech. report, http://www-ad.fnal.gov/runII/index.html, Fermilab Accelerator Division, 2001. 18

[32] Fermilab, *Operations rookie books*, tech. report, http://www-bdnew.fnal.gov/operations, 2006. 19

[33] R. Blair et al., *The CDF-II detector: Technical design report*. FERMILAB-PUB-96-390-E. 31

[34] D. Acosta et al., *Measurement of the J/ψ meson and $b-$hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ gev*, Phys. Rev., D71 (2005), p. 032001. 31

[35] C. S. Hill, *Operational experience and performance of the CDFII silicon detector*, Nucl. Instrum. Meth., A530 (2004), pp. 1–6. 35

[36] A. Sill, *CDF run II silicon tracking projects*, Nucl. Instrum. Meth., A447 (2000), pp. 1–8. 35

[37] A. Affolder et al., *Intermediate silicon layers detector for the CDF experiment*, Nucl. Instrum. Meth., A453 (2000), pp. 84–88. 35

[38] ——, *CDF central outer tracker*, Nucl. Instrum. Meth., A526 (2004), pp. 249–299. 38

[39] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nucl. Instrum. Meth., A267 (1988), p. 272. 41

[40] S. Bertolucci et al., *The CDF central and endwall hadron calorimeter*, Nucl. Instrum. Meth., A267 (1988), p. 301. 41

[41] S. Kuhlmann et al., *The CDF calorimeter upgrade for Run IIB*, Nucl. Instrum. Meth., A518 (2004), pp. 39–41. 41

[42] A. Artikov et al., *Design and construction of new central and forward muon counters for CDF II*, Nucl. Instrum. Meth., A538 (2005), pp. 358–371. 45

[43] G. Gomez-Ceballos et al., *Event builder and level 3 at the CDF experiment*, Nucl. Instrum. Meth., A518 (2004), pp. 522–524. 52

[44] F. Abe et al., *The topology of three jet events in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ tev*, Phys. Rev., D45 (1992), pp. 1448–1458. 56

[45] A. Bhatti et al., *Determination of the jet energy scale at the collider detector at fermilab*, Nucl. Instrum. Meth. (To be published), (2005). 57, 107, 108

[46] C. Neu, *CDF b-tagging: Measuring efficiency and false positive rate*, Tech. Rep. 8102, The CDF II Collaboration, (2006). 64

[47] F. Garberson et al., *Combination of the secvtx b-tagging scale factors for winter 2006 analyses*, Tech. Rep. 8025, The CDF II Collaboration, (2006). 66

[48] J. Guimaraes da Costa et al., *Measurement of the SECVTX mistag asymmetry in 5.3.3*, Tech. Rep. 7585, The CDF II Collaboration, (2005). 68

[49] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP, 0307 (2003), p. 001. 70

[50] G. Corcella et al., *HERWIG 6.5: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP, 0101 (2001), p. 010. 70

[51] T. Sjostrand et al., *High-energy-physics event generation with PYTHIA 6.1*, Computer Physics Communications, 135 (2001), p. 238. 70

[52] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth., A506 (2003), pp. 250–303. 71

[53] W. K. Sakumoto, *Event $|z_{vtx}| < 60$ cm cut acceptance for Run II (to Sep05)*, Tech. Rep. 7935, The CDF II Collaboration, (2005). 78

[54] V. Boisvert, *Trigger efficiencies for the high e_T central electrons in the gen6 data*, Tech. Rep. 7939, The CDF II Collaboration, (2005). 78

[55] U. Grundler et al., *High-pt muons recommended cuts & efficiencies for winter 2006*, Tech. Rep. 7956, The CDF II Collaboration, (2005). 78

[56] T. Spreitzer et al., *Electron identification in offline release 6.1.2*, Tech. Rep. 7950, The CDF II Collaboration, (2005). 78

[57] M. L. Mangano, *Merging multijet matrix elements and shower evolution in hadronic collisions*, tech. report, http://mlm.web.cern.ch/mlm/talks/lund-alpgen.pdf, 2004. 79

[58] F. James and M. Roos, *'MINUIT' a system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun., 10 (1975), pp. 343–367. 84

[59] K. Bloom, *Estimating QCD backgrounds to W production*, Tech. Rep. 6559, The CDF II Collaboration, (2003). 96

[60] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, *MRST partons and uncertainties*, 2003. 112

[61] H. L. Lai et al., *Global QCD analysis and the CTEQ parton distributions*, Phys. Rev., D51 (1995), pp. 4763–4782. 112

[62] D. J. Lange, *The evtgen particle decay simulation package*, Nucl. Instrum. Meth., A462 (2001), pp. 152–155. 121

[63] J. M. Campbell, *W/Z + b,bbar/jets at NLO using the monte carlo MCFM*, `hep-ph/0105226`, (2001). 122

[64] F. F. Cordero, L. Reina, and D. Wackeroth, *NLO QCD corrections to W boson production with a massive b-quark jet pair at the Fermilab Tevatron p-pbar collider*, 2006. 122

[65] J. Campbell and J. Huston, *Method 2 at NLO*, `hep-ph/0405276`, (2004). 122

ABSTRACT

Measurement of the $W^{\pm} + b\bar{b}$ Cross-Section in 695 pb^{-1} of $p\bar{p}$ Collisions at CDF II

by
Mitchell P. Soderberg

Chair: David Gerdes

$W^{\pm}$+ $b\bar{b}$ events contain the associated production of a $W^{\pm}$ boson, a pair of bottom quarks ($b\bar{b}$), and any number of additional partons. This process is of much importance at hadron collider experiments due to its role as a background source in searches for Standard Model Higgs boson and single top-quark production. In this thesis the results are presented for a measurement of the b-jet cross-section in $W^{\pm}$+$b\bar{b}$ events containing 1 or 2 jets in 695 pb^{-1} of $\sqrt{s}$ =1.96 TeV $p\bar{p}$ collisions at the CDF experiment. This is the first measurement of the cross-section of $W^{\pm}$+$b\bar{b}$ performed in any experiment. The cross-section is defined to be proportional to the number of b-jets from $W^{\pm}$+$b\bar{b}$ events with one or two jets, and a leptonically decaying $W^{\pm}$ with decay products passing kinematics cuts ($p_T(\ell^{\pm})$ $\geq$20.0 GeV, $|\eta(\ell^{\pm})|$ $\leq$1.1, $p_T(\nu)$ $\geq$25.0 GeV). The invariant mass distribution of jets identified as containing a long-lived hadron is fit with components for bottom, charm, and light-flavor to find the fraction due to true b-decays. Background b-jet sources are subtracted to isolate the contribution of $W^{\pm}$+ $b\bar{b}$ to the data. The cross-section is measured to be 0.90±0.20(stat.)±0.26(syst.)pb, which compares well with the leading order theoretical prediction of 0.74±0.18 pb.