A SEARCH FOR ANOMALOUS HEAVY-FLAVOR QUARK PRODUCTION IN ASSOCIATION WITH W BOSONS

Wade Cameron Fisher

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PHYSICS

January 2005

Abstract

The production of W bosons in association with jets at the Fermilab Tevatron provides an opportunity to test predictions for electroweak and QCD processes described by the standard model. Complicating this picture, evidence for anomalous production of heavy-flavor quarks (t, b, c) in association with W bosons was reported in Run I by the CDF collaboration. In this dissertation, I present an examination of the exclusive jet spectrum in the W+jets final state in which the heavy-flavor quark content has been enhanced by requiring at least one b-tagged jet in an event. For this measurement, W bosons have been selected in $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ decay channels. I present a measurement of the exclusive jet spectrum for events which contain one jet tagged with more than one b-tagging algorithm. I compare data on $e + \text{jets}$ (164.3 pb^{-1}) and $\mu + \text{jets}$ (145.3 pb^{-1}) channels, collected with the DØ detector during Run II of the Fermilab Tevatron $p\bar{p}$ collider, to expectations from the standard model. The results of the search are used to set upper limits on anomalous production of such events.

Acknowledgments

The path I have followed throughout my education in physics has been marked by the influence of exceptional people. In most instances, I've been presented with serendipitous opportunities that have provided guidance along this path. These opportunities arose in circumstances where someone saw something in me that I wasn't able to see on my own. Here, I would like to acknowledge the people who have had the greatest influence on my journey.

Being raised in rural North Dakota, the educational experiences available for myself were quite restricted when compared with those considered commonplace in more populous regions of the country. Although I aspired to a higher education in sciences, my lack of instruction in advanced science and mathematics loomed as an enormous liability for admission to colleges and universities. Fortune, however, smiled and placed me in friendship with a rare teacher with experience in college level calculus. Ron Parfit, who spent most of the day teaching the math courses necessary for meeting minimum graduation requirements, agreed to forfeit time in his schedule to provide one-on-one tutoring in calculus. Although I believe that Mr. Parfit enjoyed teaching me as much as I enjoyed learning, the sacrifice he made to teach one student during his spare time and without monetary compensation was as rare as his unique abilities.

Also during this time in high-school, I had the fortune of meeting a woman who has spent much of her life opening doors to educational experiences for students who were previously oblivious to such possibilities. The details of our acquaintance, although an exceptional story, belong somewhere with much more space for my reminisce. Within hours of meeting Harriet Howe, we had boiled down my upcoming summer plans to performing research in biophysics at Boston University or taking advanced math and science courses at the University of Hawaii at Hilo. I chose Boston. Aside from the wonderful educational experience, my interaction with Harriet opened my mind to a world full of possibilities. Finding the courage to step through a door into an unfamiliar world is always difficult. When it's your first of these doors, sometimes it takes a bit of a nudge to take that step. Harriet provided something more like a healthy shove, and for that I will always be grateful.

With my high-school experiences and, no doubt, a healthy dose of fortune, I was accepted to the University of Minnesota Institute of Technology Honors Program as a physics major. My first year or so constitute another story which belongs somewhere else, as I underwent procedures to remove an intracranial tumor, subsequent radiation therapy, and recovery period. In my second year, I took my first quantum physics course. This course, very well-taught by Prof. Priscilla Cushman, was my introduction to so many ideas and phenomena that I had never before even imagined. In fact, her enthusiasm and commanding grasp of the topic soon brought me to her office asking for a job where I could learn more about what she'd taught me. I was granted such an opportunity and found myself plunged into the ways and means of high energy particle physics. This experience, more than any other, solidified my desire to continue in particle physics. During my second, third, and fourth years at the University of Minnesota, I worked for Prof. Cushman and Prof. Roger Rusack and became very excited about the physics they pursued in their research in particle physics. The confidence I gained via their encouragement and my experiences in their laboratories has been an invaluable asset in my journey as a physicist.

While working for Prof. Cushman, I was given the opportunity to work at the CERN particle physics research center just outside Geneva, Switzerland. For a student at my point in education, this experience was a singularly wonderful introduction to the world of particle physics. While at CERN, I met and became friends with another very influential person in my career, Prof. Andris Skuja. Although I wasn't fully aware at the time, Andris possesses a unique insight to the landscape and structure of the field of high energy physics. The advice and support Andris has provided me is as dear to me as his friendship. I have learned much of what I know about how to be a physicist from Andris. I count my friendship with Andris among my most meaningful experiences.

Upon graduating from Minnesota, I chose to attend Princeton University. I joined the group working on the L3 experiment and flew to CERN to begin working. Here I met Dr. Jeremy Mans, who was a fellow graduate student at the time. Immediately, Jeremy became my friend and mentor. There are very few individuals as exceptionally talented as Dr. Mans, and I feel very fortunate to have been under his tutelage. His ability to understand, master, and communicate the arts of physics are second to none. In addition, Jeremy's friendship was a great benefit to our working relationship. I have appreciated our friendship every moment, as his patience and good humor with me were invaluable to my experience. I am so very glad to have been in the right place at the right time to work with Jeremy.

I chose to attend Princeton University because I wanted to work for Prof. Chris Tully. Many people had recommended him as a very skilled physicist with an well-developed outlook of the field. More convincing to me at the time was his youth and enthusiasm. Chris has been a very unassuming advisor, but his expertise shines through every effort at modesty. He has the extraordinary ability to understand physics with a speed that never ceases to amaze me. Discussions rarely last more than a few moments before he's able to grasp even the smallest details of the topic. As a friend, role model, and advisor, Chris has exceeded all of my expectations. I feel that I've been extremely fortunate to have an advisor of Chris's quality and compassion. My experience as a graduate student has been defined by his skills and ability to relate. I owe Chris a very large debt of gratitude for agreeing to take me on as a student.

Of course, none of what I've described here would have been possible without my parents, Bart and Kay. At every step in my life, they have made it clear that the sky is the limit. The importance they put on education and the pains they undertook to remove all obstacles from my path as a student are traits every child should be able to have. I will be eternally grateful for the encouragement and support (emotional and financial) they've given me. The ability to lead such a selfless life in favor of their children is still amazing to me. I endeavor to follow in my parents' footsteps when I am faced with the same decisions.

The list of people I have included above is by no means meant to be exhaustive. I could fill volumes with the descriptions of the numerous people who have helped shape my path as a physicist. For everyone not listed, please accept my thanks for helping me reach this goal.

Contents

List of Tables

List of Figures

Part I

Invitation

Foreword to Part I

The work described in this dissertation is a continuation of the ongoing search for new phenomena in elementary particle physics and was completed at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, IL. Fermilab is the home of the world's highest-energy particle accelerator, referred to as the Tevatron, which collides counter-rotating beams of protons and anti-protons. These beams are brought into collision at two points on the ring. The DØ experiment sits at one of these points and was used to collect the data used in this thesis.

I have separated this document into two parts: the first, which you are beginning now, attempts to provide a landscape view of the physics studied by scientists at Fermilab and the methods utilized in this study. The second part focuses on the data analysis to which this thesis is devoted and the motivations for such an endeavor. I like to think of the first five chapters of this document as an invitation to the reader, providing background for the physics discussed here. As alluded to above, the work described in this thesis is a continuation of efforts which have been in progress for many years. To immediately leap into the details of the analysis would, in numerous ways, take undue credit for the many lifetimes of work which have defined this field. Furthermore, it would provide a disadvantage to readers who are not intimately familiar with the ideas behind elementary particle physics or its experimental methods. To allay such misdirections, Chapter one provides a short discussion of the means and motivations of the particle physicist; Chapter two offers an introduction to the theoretical advances (and weaknesses) embodied in the Standard Model of particle physics; Chapter three describes the experimental apparatus used to initiate and collect the data used in this analysis; Chapter four outlines the methods used to translate the collected data into meaningful physics; and Chapter five introduces the concept of simulating the predictions of theories.

Chapter 1

Introduction

For millenia, the nature of the universe has fascinated those who live in it. The idea of basic building blocks of nature was devised early on by the Greek philosopher Empedocle (*cf.* 450 BC) who broke matter down into four "roots": water, earth, air, and fire[1]. These roots were bound by the forces of love and hate, with love creating unity and hate forcing division. The mixing of these basic elements by the forces gave rise to all matter. The science of particle physics follows in the footsteps of this philosophy, struggling to describe the fundamental components of matter and their interactions. A revolution in the understanding of particle interactions occurred in the 19th and 20th centuries as scientists rapidly progressed from Dalton's 1803 inference of atoms as the basic components of matter to the development of quantum mechanics and relativity in the early 19th century. Shortly after, technological advances drove a furious wave of advancement in the field culminating spectacularly in the 1983 discovery of W and Z bosons and the 1995 discovery of the top quark.

This chapter is intended to introduce the philosophy and basic practices of experimental particle physics. Furthermore, in an effort to reach the widest possible audience, it is written as a pedagogical discussion for those who are not familiar with the field this thesis addresses.

1.1 Fundamental Questions

The field of particle physics[2] has evolved over many years, but has remained constant in the desire to answer fundamental questions about the universe in which we live. The most basic methodology has been to identify the smallest indivisible components of matter. In this regard, we have been very successful in identifying a handful of particles believed to be basic. Further efforts have shown how all forms of matter observed thus far can be constructed from these basic constituents. With this evidence, one can address questions such as the evolution of the universe, asymmetries between matter and anti-matter, and the origin of mass. Answering these questions involves a long journey of smaller questions. Each step along this journey paradoxically demonstrates how little we understand, rather than simply advancing our knowledge. Indeed, each new discovery simultaneously enlightens us while unearthing numerous new questions. In this way, the identification and pursuit of

---

[1]Hindu scholars of the same epoch added sky to this list.
[2]This field is also commonly referred to as high-energy physics or elementary particle physics.

fundamental questions drives the evolution of the field. Furthermore, accumulating evidence allows for the formulation of predictions for questions that remain unanswered.

Physicists have identified several particles that they believe to be the most fundamental. Some of these particles are stable and form the normal matter we interact with every day while the others live for only fractions of a second before decaying to the stable ones. The matter that we interact with on a daily basis was created in the birth of the universe, the so-called Big Bang. The intense energy released manifested itself as enormous temperatures which facilitated particle collisions with energies large enough to produce all of the elementary particles and their bound states. As the universe cooled[3], most of these interactions became increasingly inefficient and eventually energetically impossible. Continued cooling has produced the universe we live in today. Thus, to reproduce the conditions of the early universe (and thereby study them) we require the means of achieving particle interactions with energies equal to those of the earliest times.

1.2   Tools and Goals

As with all vocations, the practitioners of particle physics have crafted very specialized tools and methods to facilitate the experimental observations imperative to the science. This toolbox is employed to achieve the primary goal of particle physicists: to produce particles not found in ordinary matter for the purposes of studying their properties. This task is clearly separated into two problems: particle production and particle detection. Each of these problems is addressed with a unique set of methods.

1.2.1   Particle Production

The most important method in the particle physicist's toolbox is the ability to create and annihilate particles. Colliding two particles can cause the annihilation or scattering of the particles. If the momenta of the colliding particles is large compared to their rest mass, the resulting annihilation or scattering can create final-state particles that are different from the original particles. In this way, physicists can carefully construct collisions which will result in particles of interest. A large fraction of the effort to obtain the necessary collisions goes into preparing uniform beams of particles with a precise energy. There are two main types of particle accelerators capable achieving this: linear accelerators and synchrotrons.

Linear accelerators are, as the name suggests, machines which accelerate electrically charged particles to a specific energy over a straight path. This acceleration occurs by using successive bursts of energy from a series of electric fields along the length of the accelerator to push particles to higher and higher energies. Particles are confined to the path of the accelerator using magnetic fields. Collisions between particles can be arranged by directing the particle beams of two opposite-facing linear accelerators at a single interaction point. The drawback of such a machine is that larger energies require longer accelerators, eventually becoming unrealizable due to cost and space limitations.

---

[3]The explosion of the Big Bang caused the infant universe to expand. This expansion increased average particle separation, and thus lowered average energy density, or, equivalently, temperature.

Synchrotrons are circular accelerators. Particles are accelerated in the same way as with linear accelerators, but they are forced to follow a circular path (a ring). Such accelerators allow for two counter-rotating beams in the accelerator at one time. These beams can be brought into collision at points along the ring. The advantage of this design is that the accelerating path is followed many times by each particle, thus reducing the length of accelerating fields required to achieve higher energies. Furthermore, the beams of particles are recycled and caused to collide many times, thus increasing the efficiency of the accelerator. The drawback is that the constant turning of the particles required to remain in the circular path causes them to lose energy in the form of synchrotron radiation. This loss increases with particle energy and eventually limits the maximum energy obtainable with the accelerator.

1.2.2 Particle Detection

After successfully initiating particle collisions which result in the production of new particles, the task remains to observe the behavior of the new particles created. The most useful information for physicists is the energy and momentum for each final state particle. This knowledge allows one to reconstruct the event and thus infer the physics involved. To achieve these observations, immense detectors are constructed around the collision point. These detectors contain two different types of subdetectors: particle tracking and calorimetry. These subdetectors measure the paths and energies of the particles produced in the collision, respectively.

It is the jobs of particle tracking detectors to precisely measure the trajectories of particles leaving the interaction region without significantly influencing their original momenta. This is achieved by using gases or thin layers of material sensitive to passing charged particles, which are arranged close to the collision point. When a charged particle passes through these materials, some atoms are ionized. This ionization is used to reconstruct the path of the particle. The introduction of a uniform solenoidal magnetic field will cause charged particles to follow a curved path, yielding a measurement of the particle's momentum.

Once the direction and momentum of particles are measured, the energy can be determined. This technique, known as calorimetry, is accomplished by bringing the particle to a halt in carefully chosen materials outside the tracking detectors. Interactions with these materials cause the particle to continuously lose energy until it stops. The energy loss initiates a calibrated response in the detector which is translated to an energy measurement.

In this way, the particles produced in each collision are observed and recorded. This process is repeated many, many times to observe different final states[4] and to collect many instances of the same physics process to increase the statistical certainty of the observation.

1.3 Predictions and Horizons

The ability to predict the results of particle physics experiments has steadily progressed to a level which currently allows accurate predictions in most instances. The predictive paradigm used today is referred to as the Standard Model (SM) of particle physics. Within the SM,

---

[4]Each collision typically only produces one interaction type of interest. Collecting a significant number of interesting events requires many billions of particle collisions.

there exist descriptions of the most basic particles and their properties. Also described are the means by which these particles interact with one another, which gives rise to the forces observed in nature. This means of predicting particle interactions has proven to be a very powerful tool for understanding matter on its most fundamental level.

However, the SM is an approximate description and is not believed to be the ultimate description of nature. In particular, the SM predicts the existence of a particle (or particles, depending on the exact formulation of the theory) needed to explain the origin of mass. This particle remains undetected, despite large efforts designed to observe such a particle. Furthermore, the SM does not explicitly exclude physics not described by the SM itself. These issues reinforce the understanding that the SM only describes a small portion of physics and that a more complete theory could someday be formulated. As such, physicists continue to test the SM's predictive ability in an effort to identify physics not described by the model. Such searches for new phenomena will continue to expand the horizons of particle physics, both the known and the unknown.

1.4   Notational Conventions

Before continuing on in this discussion, it is necessary to introduce the common use of *natural units*. In high energy physics, there are two fundamental constants: Planck's constant, h, and the speed of light (in vacuum), c. These constants appear ubiquitously in the notation of this field and it becomes convenient to adopt a system of units in which the values of these constants are equal to one

$$\hbar = \frac{h}{2\pi} = 1.055 \times 10^{-34} J\ sec \rightarrow \hbar = 1$$
$$c = 2.998 \times 10^{10} cm/s \rightarrow c = 1$$

In this way, an equation such as Einstein's famous energy-matter relationship $E = mc^2$ becomes simply $E = m$. It is then appropriate to convert units to reflect this change: energy, momentum, and mass are measured in GeV, distance and time in GeV^{-1}. This notational convention, although not quite accurate from a mathematical point of view, is commonplace and will thus be used throughout this dissertation.

Chapter 2

Theoretical Background

By the early 20th century, the theories of special relativity and quantum mechanics had arrived at the forefront of modern physics. However, quantum mechanics broke down at large velocities and relativity failed to make predictions on small distance scales. In 1928, the era of relativistic quantum mechanics was ushered in by Dirac when he merged the two theories in a relativistically covariant equation describing a quantum-mechanical spin-1/2 particle. Over the next three decades, a wealth of experimental evidence lead to many ad hoc applications of this approach but no coherent theory describing the observations. During the late 1960's and early 1970's Glashow, Salam, and Weinberg [1, 2, 3] developed a relativistic quantum field theory capable of describing the physical world at high energies and small distances based on the principle of local gauge invariance. This theory is known as the Standard Model (SM) of particle physics. In my description of the SM, I will follow the conventions of [4] and [5].

2.1 The Standard Model

In quantum field theories (QFTs), particles are described by quantized field excitations above the ground state, with different fields representing different particle types. The SM contains two kinds of elementary particles: matter particles and force carriers. The matter particles can be divided into two types of fermions (particles with intrinsic spin of 1/2): quarks and leptons. These fermions all have mass and are the constituents of the physical universe. The force carriers have integral spin and are called bosons. Interactions between the matter particles are mediated via boson exchange. These exchanges give rise to the three SM forces: electromagnetic, weak, and strong[1]. In total, six quarks, six leptons, and five bosons have been observed and are listed in Tables 2.1 and 2.2.

2.1.1 Local Gauge Invariance

The SM is a Lagrangian field theory and a discussion of which should begin with a description of its construction. To properly describe a field, our Lagrangian must be a function of a complex field, $\phi(x)$, and its derivatives, $\partial_\mu \phi(x)$

---

[1]The fourth fundamental force, gravity, is not included in the formulation of the SM.

Leptons			Quarks		
Particle Type	Symbol	Charge	Particle Type	Symbol	Charge
electron neutrino	ν_e	0	up quark	u	2/3
electron	e	+1	down quark	d	-1/3
muon neutrino	ν_μ	0	charm quark	c	2/3
muon	μ	+1	strange quark	s	-1/3
tau neutrino	ν_τ	0	top quark	t	2/3
tau	τ	+1	bottom quark	b	-1/3

Table 2.1: The spin-1/2 particles (fermions) of the Standard Model of particle physics.

Bosons			
Particle Type	Symbol	Charge	Force Mediated
photon	γ	0	Electromagnetic
W^+ boson	W^+	+1	Weak
W^- boson	W^-	-1	Weak
Z boson	Z	0	Weak
gluon	g	0	Strong

Table 2.2: The integral-spin particles (bosons) of the Standard Model of particle physics.

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.1}$$

All of the force carrier particles and the interactions they mediate can be shown to arise from symmetries in such a Lagrangian. In particular, requiring the SM Lagrangian to be symmetric under local gauge[2] transformations provides a mechanism for generating each boson in turn. An example of such a gauge transformation can be provided by the $U(1)$ group of local phase transformations $U(\beta(x)) = e^{i\,q\,\beta(x)}$ where $\beta(x)$ has an arbitrary dependence on space-time coordinates and q is some constant. If we assume a fermion field with no external potential, the Lagrangian takes the form

$$\mathcal{L} = \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi \tag{2.2}$$

where γ^μ are the Dirac matrices. Upon inspection of this Lagrangian, it is clear that it is not invariant under the $U(1)$ transformation of the field

$$\psi \rightarrow e^{iq\,\beta(x)}\psi \tag{2.3}$$

However, if we rewrite the Lagrangian as

---

[2]The term "gauge" can be taken as synonymous with phase.

$$\mathcal{L} = \overline{\psi} \left(i\gamma^{\mu}\mathcal{D}_{\mu} - m \right) \psi \tag{2.4}$$

where $\mathcal{D}_{\mu} = \partial_{\mu} + iA_{\mu}(x)$ and A_{μ} is a vector field (potential) which transforms under $U(1)$ as

$$A_{\mu} \to A_{\mu} - \partial_{\mu}\beta(x) \tag{2.5}$$

we explicitly preserve local gauge invariance and the Lagrangian takes on the new form of

$$\mathcal{L} = \overline{\psi} \left(i\gamma^{\mu}\partial_{\mu} - m \right) \psi - q\,\overline{\psi}\gamma^{\mu}\psi\,A_{\mu} \tag{2.6}$$

In this way, we have added to the Lagrangian a massless vector field which couples to the fermion field with strength q. This new particle can be interpreted as the photon. If we further require $SU(2)$ and $SU(3)$ gauge invariance of the fermion Lagrangian, we explicitly obtain three more massless vector bosons (carriers of weak force) and eight massless scalars (carriers of strong force), respectively.

2.1.2 Electroweak Unification

As mentioned above, the weak force carriers can be introduced into the SM by requiring $SU(2)$ local gauge invariance. However, this alone is unsatisfactory as it allows only one neutral vector boson, while Table 2.2 lists two. We can get around this by combining the (commutable) $SU(2)$ and $U(1)$ groups in one group: $SU(2) \times U(1)$. This done by requiring the left handed and right handed particle fields to transform differently

$$\chi_L \to e^{i\,\alpha^a \cdot T^a}\,e^{i\,\beta Y}\chi_L$$
$$\psi_R \to e^{i\,\beta Y}\psi_R \tag{2.7}$$

where $T^a = \sigma^a/2$ is the generator for the $SU(2)$ rotation group and σ^a are the three Pauli matrices. The phase Y is analogous to the phase $\beta(x)$ introduced in the last section, but I've changed its name to avoid confusion with the previous example. The χ_L are isospin doublets of left-handed fermions and the ψ_R are isosinglets of right-handed fermions[34]. We can create a Lagrangian invariant under these transformations by writing a covariant derivative of the form

$$\mathcal{D}_{\mu} = \partial_{\mu} + ig\,W_{\mu}^a\,T^a + \frac{i}{2}\,kB_{\mu} \tag{2.8}$$

for left-handed doublets and

$$\mathcal{D}_{\mu} = \partial_{\mu} + \frac{i}{2}\,kB_{\mu} \tag{2.9}$$

for right-handed singlets. This technique introduces three $SU(2)$ gauge bosons, W_{μ}^a, and one $U(1)$ gauge boson, B_{μ}, which couple to the fermion fields with strengths g and k, respectively.

---

[3]Handedness is the same as helicity and is determined by the overlap of the particle's spin and momentum vectors.

[4]It should be noted that in the SM, the neutrino is massless. Thus, there exist only left-handed neutrinos.

We have now obtained a Lagrangian that is invariant under local $SU(2) \times U(1)$ gauge transformations, and have produced four massless vector bosons. This feature exemplifies the beautiful interaction between theory and experiment in the field of particle physics. The physical model predicted three weak bosons, which were discovered in a celebrated dual victory of theory and experiment. However, these weak force carrier bosons were observed to have masses nearly 100 times larger than the proton mass. A very large problem, indeed, as adding a mass term of the form $\frac{1}{2}m^2 W^\mu W_\mu$ to the Lagrangian would destroy the local gauge invariance we worked so hard to obtain. Ignoring this issue and introducing mass terms "by-hand" would inject divergences into the theory, rendering it unrenormalizable and thus retaining no predictive power. This problem is solved in the Glashow-Weinberg-Salam (GWS) model of electroweak interactions by the method of spontaneous symmetry breaking.

2.1.3 Electroweak Symmetry Breaking

The GWS model addresses the problem of gauge boson masses by introducing a $SU(2)$ doublet of complex scalar (spin-0) fields to the SM Lagrangian

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \text{with} \quad \begin{matrix} \phi^+ = (\phi_1 + i\phi_2)/\sqrt{2} \\ \phi^0 = (\phi_3 + i\phi_4)/\sqrt{2} \end{matrix} \tag{2.10}$$

which interacts with itself via a potential $V(\phi) = \mu^2\, \overline{\phi}\phi + \lambda\left(\overline{\phi}\phi\right)^2$. To force non-zero boson masses, we consider the case of $\mu^2 < 0$ and $\lambda > 0$ giving $V(\phi)$ two minima at $\phi_0 = \pm\sqrt{-\mu^2/2\lambda} = \pm v$. We have the freedom to choose either solution as the ground state for ϕ and in this discussion we shall choose the positive solution. We can then rewrite the scalar field as an excitation relative to this non-zero vacuum expectation value

$$\phi(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.11}$$

which manifestly breaks the $SU(2) \times U(1)$ gauge symmetry and introduces a single real scalar field, $h(x)$, which is referred to as the Higgs field. Via interactions with the Higgs field, the SM Lagrangian obtains a new term which generates the gauge boson masses (which is sufficient to evaluate at the ground state of ϕ)

$$\mathcal{L}' = \frac{1}{2}\begin{pmatrix} 0 & v \end{pmatrix}\left(gW^a_\mu T^a + \frac{1}{2}kB_\mu\right)\left(gW^{b\mu} T^b + \frac{1}{2}kB^\mu\right)\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.12}$$

$$= \frac{v^2}{4}\left(g^2\left(W^1_\mu\right)^2 + g^2\left(W^2_\mu\right)^2 + \left(-gW^3_\mu + kB_\mu\right)^2\right) \tag{2.13}$$

(recall that T^a are the Pauli matrices). We therefore can write

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2 \right) \quad \text{with} \quad m_W = \frac{v\,g}{2} \tag{2.14}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g^2 + k^2}} \left(g W_\mu^3 - k B_\mu \right) \quad \text{with} \quad m_Z = \frac{v}{2}\sqrt{g^2 + k^2} \tag{2.15}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + k^2}} \left(k W_\mu^3 + g B_\mu \right) \quad \text{with} \quad m_A = 0 \tag{2.16}$$

where we can now interpret the forms of the weak gauge fields and the photon field. The three extra degrees of freedom obtained from $\phi(x)$ appear as longitudinal polarizations of the original W_μ^i fields, allowing them to become massive. Thus, the introduction of the Higgs field, although it spontaneously breaks $SU(2) \times U(1)$ symmetry upon the choice of a ground state for the Higgs potential, solves the problem of weak boson masses while keeping the theory renormalizable. Furthermore, fermion masses are obtained in a similar manner through their interactions with the Higgs field. Reinterpreting Eq. 2.13 with Higgs field excitations above the ground state of ϕ_0, we obtain self interaction terms and a mass term for the Higgs field

$$\mathcal{L}' = -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 \tag{2.17}$$

$$= -\frac{1}{2}m_h^2 h^2 - \sqrt{\frac{\lambda}{2}}m_h h^3 - \frac{1}{4}\lambda h^4 \tag{2.18}$$

where $m_h = \sqrt{2}\mu^2$. The only unsatisfactory feature of this new Higgs field is that its mass is not predicted by the theory, even though the $W^\pm$ and Z^0 masses can be determined by measurements of weak decays.

2.1.4 Quantum Chromodynamics

The third and final SM force remaining is the strong force. This force is incorporated into the SM Lagrangian via the requirement of $SU(3)$ gauge invariance. The fundamental representation of this group requires three quantum numbers, which are referred to as "colors" in the quark sector, thus motivating the name quantum chromodynamics (QCD). As $SU(2)$ gave rise to $2^2 - 1 = 3$ gauge bosons, QCD adds $3^2 - 1 = 8$ new fields (G^μ), referred to as gluons. Given color field transformations under $SU(3)$ given by

$$\psi(x) \rightarrow e^{i\alpha_a(x)\Gamma^a}\psi(x) \tag{2.19}$$

where Γ^a are the eight generators of the $SU(3)$ gauge group and α_a is allowed to vary over all space-time points (a summation over the subscript a is implied), we can create a Lagrangian invariant to $SU(3)$ phase transformations by writing a covariant derivative of the form

$$\mathcal{D}_\mu = \partial_\mu + ig_s\,G_\mu^a\Gamma^a \tag{2.20}$$

G^μ are the eight gluon fields and g_s is the strong coupling constant. However, due to the non-Abelian nature of the $SU(3)$ generators, we find that

$$\left(\overline{\psi}\gamma^{\mu}\Gamma_a\psi\right) \rightarrow \left(\overline{\psi}\gamma^{\mu}\Gamma_a\psi\right) + f_{abc}\alpha_b\overline{\psi}\gamma^{\mu}\Gamma_c\psi \tag{2.21}$$

where f_{abc} are the $SU(3)$ structure constants. This indicates that the fields G_{μ}^a must transform as

$$G_{\mu}^a \rightarrow G_{\mu}^a - \frac{1}{g_s}\partial_{\mu}\alpha^a - f_{abc}G_{\mu}^c\alpha_b \tag{2.22}$$

In QCD, quarks are defined as one of three color fields (chosen as red, green, and blue), defining three quark and three anti-quark quantum states. The gluons occupy eight color states: $r\overline{g}$, $r\overline{b}$, $g\overline{r}$, $g\overline{b}$, $b\overline{r}$, $b\overline{g}$, $\left(r\overline{r} - g\overline{g}\right)/\sqrt{2}$, and $\left(r\overline{r} + g\overline{g} - 2b\overline{b}\right)/\sqrt{6}$. These gluons are massless and are the carriers of the strong force. Quarks and anti-quarks are bound into color singlet combinations by gluons to form mesons and hadrons.

2.1.5 The Standard Model Lagrangian

At this point, we have nearly constructed the full SM Lagrangian in parts. We can collect all the parts and summarize as

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{H} \tag{2.23}$$

Up to this point, we've ignored the kinetic energy terms in the Lagrangians. We reintroduce them here for completeness. These components are given by

$$
\begin{aligned}
\mathcal{L}_{EW} = \quad & -\tfrac{1}{4}W_{\mu\nu}W^{\mu\nu} - \tfrac{1}{4}B_{\mu\nu}B^{\mu\nu} && W^{\pm}, Z^0, \gamma \text{ kinetic energies} \\[6pt]
& + \overline{\chi}_L\gamma^{\mu}\left(\partial_{\mu} + ig\,W_{\mu}^a\,T^a + \tfrac{i}{2}\,kB_{\mu}\right)\chi_L && \text{Left-handed quark and lepton} \\
& && \text{kinetic energies and EW interactions} \\[6pt]
& + \overline{\psi}_R\gamma^{\mu}\left(\partial_{\mu} + \tfrac{i}{2}\,kB_{\mu}\right)\psi_R && \text{Right-handed quark and lepton} \\
& && \text{kinetic energies and EW interactions}
\end{aligned}
\tag{2.24}
$$

where χ_L and ψ_R extend for all quark and lepton left-handed (LH) doublets and right-handed (RH) singlets, respectively, $W_{\mu\nu} = \partial_{\mu}W_{\nu} - \partial_{\nu}W_{\mu}^a - gW_{\mu} \times W_{\nu}$, and $B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}$;

$$
\begin{aligned}
\mathcal{L}_{QCD} = \quad & -\tfrac{1}{4}F_{\mu\nu}^{(a)} \cdot F^{(a)\mu\nu} && \text{Gluon kinetic energies} \\[6pt]
& + i\overline{q}^i\left(ig_s G_{\mu}^a\Gamma^a\right)q^i && \text{Gluon couplings to the quarks}
\end{aligned}
\tag{2.25}
$$

where $F_{\mu\nu}^{(a)} = \partial_{\mu}G_{\nu}^a - \partial_{\nu}^a - \alpha_b f_{abc}G_{\mu}^a G_{\nu}^c$, $a = 1, ..., 8$ runs over the gluon indices, and $i = 1, 2, 3$ runs over the three quark colors;

$$\mathcal{L}_H = \quad \left| \left(\partial_\mu + ig\, W_\mu^a\, T^a + \tfrac{i}{2}\, kB_\mu \right) \phi \right|^2 \qquad \text{Higgs kinetic energy}$$

Higgs kinetic energy
and EW interactions

$$-\mu^2 \bar{\phi}\phi + \lambda \left(\bar{\phi}\phi \right)^2 \qquad \text{Higgs potential energy} \tag{2.26}$$

Higgs potential energy

$$- \left(g_1 \bar{\chi}_L\, \phi\, \psi_R + g_2 \bar{\chi}_L\, \phi_c\, \psi_R + HC \right) \qquad \text{Lepton and quark Higgs couplings}$$

Lepton and quark Higgs couplings
and mass terms

where ϕ_c is given by

$$\phi_c = \begin{pmatrix} \bar{\phi}^0 \\ \phi^- \end{pmatrix} \qquad \text{with} \qquad \begin{array}{l} \bar{\phi}^0 = (\phi_3 - i\phi_4)/\sqrt{2} \\ \phi^- = (\phi_1 - i\phi_2)/\sqrt{2} \end{array} \tag{2.27}$$

and HC indicates the hermitian conjugate of the terms. The Higgs Lagrangian $\mathcal{L}_H$ can be written in the more intuitive form of (expanding about the minimum)

$$\mathcal{L}_H = \quad \tfrac{1}{2} \left(\partial h \right)^2 - \tfrac{1}{2} m_h^2 h^2 \qquad \text{Higgs kinetic energy and mass}$$

$$-\tfrac{1}{2} \frac{m_h^2}{v} h^3 - \tfrac{1}{8} \frac{m_h^2}{v^2} h^4 \qquad \text{Higgs self-interactions}$$

$$+\tfrac{1}{4} W_\mu^+ W^{-\mu} \left(2v\, h + h^2 \right) \qquad WWh \text{ and } WWhh \text{ vertices}$$

$$+\tfrac{1}{8} \frac{g^2+k^2}{\sin \theta_W{}^2} Z_\mu Z^\mu \left(2v\, h + h^2 \right) \qquad ZZh \text{ and } ZZhh \text{ vertices} \tag{2.28}$$

$$-m_W^2 W_\mu^+ W^{-\mu} + m_Z^2 Z_\mu Z^\mu \qquad W^\pm, Z^0 \text{ masses}$$

$$-\sum_f m_f \bar{\psi}^f \psi^f \qquad \text{Fermion masses}$$

This completes the full SM Lagrangian in its GWS formulation, providing for the kinetic energies of all SM particles, the masses (or lack thereof) of all SM particles, and the gauge-boson mediated fundamental forces.

2.2 Beyond the Standard Model

The SM has proven to be a sufficient theory describing fundamental particles and their interactions up to the energies presently achievable. Indeed, the SM has proven itself in an era of precision electroweak measurements, and the robustness of this theory is due largely to the fundamental simplicity it uses to approach a description of matter. Despite its successes, however, there remain a few glaring problems that remind us that there must be physics beyond the SM. This section outlines some of the problems faced by the SM and a few of the currently promising alternative formulations. By no means is it intended to be a comprehensive discussion of the topic. The goal is to motivate the reader's imagination for new ideas. I also hope to reinforce the idea that elementary particle physics is a journey and not a destination.

2.2.1 Problems of the Standard Model

Given sufficient time and energy, a well-educated person could fill several books on the topic of problems of the SM. However, I would like to draw attention to only a few problems which I feel exemplify the most important issues. First, there exist several important problems that the SM was not designed to address. These issues are not technically a failure of the SM, but simply describe the litany of physics that is not described by the SM.

- **Gravity:** The one fundamental force that the SM is silent about is gravity. Although the strength of the gravitational force is tiny at the very small scales probed by the SM, it remains the dominant force throughout the universe. Attempts to find a quantum mechanical description of gravity consistent with the SM have failed spectacularly. This omission provides the first suggestion that the SM is merely an effective field theory (EFT) which describes nature well at the chosen length scales, but ignores the additional degrees of freedom at smaller length scales (higher energies).

- **Unification:** The SM description of fermion-boson interactions is fixed by the gauge coupling strengths, which are different for each gauge boson. A more complete theory will create a single charge quantization and thus unify all three forces. Such a unification in the SM would occur at higher (unprobed) energies. Typically, these energies are at a very large scale ($\Lambda_{Pl} \sim 10^{18}$ GeV) known as the Planck Scale. However, calculations of this phenomenon indicate that without the discovery of new physics, the gauge couplings do not converge even at large energies as seen in Figure 2.1.

- **Dark Matter:** Measurements of the matter-energy density of the universe find that the density prescribed to the particles of the SM accounts for very little of the total density. Barring a change in the behavior of gravity over very large length scales, the most probable explanation is a new particle not directly predicted by the SM. This particle would necessarily be non-luminous to explain the inability to observe it, hence the name Dark Matter. Depending on the properties of such a particle, the SM could either require cosmetic changes or fail completely.

- **The Cosmological Constant Problem:** Vacuum energy calculations of the vacuum energy density (ρ_V) based on the predictions of the SM result in a very large value of $\rho_V \sim \Lambda_{Pl}^4$, if one believes that the SM holds up to the Planck Scale, and $\rho_\Lambda \sim \chi_{EW}^4 \simeq (100 \text{ GeV})^4$ otherwise. This is many orders of magnitude larger than recent cosmological observations, which suggest a cosmological constant of approximately $\rho_\Lambda = 10^{-39}$ GeV^4. While this problem could be resolved by fine-tuning – subtracting off the present cosmological constant by fiat – it seems odd that two numbers should cancel to over 50 orders of magnitude [6]. As the SM contains no mechanism (symmetry) which could account for a negative vacuum energy contribution, this evidence requires an explanation external to the SM.

- **Matter / Antimatter Asymmetry:** The world we live in is made purely of matter. Predictions of the very early universe suggest that there was a slight excess of matter over antimatter. SM matter-antimatter interactions could not have driven this resulting asymmetry without producing an enormous energy density beyond what is currently
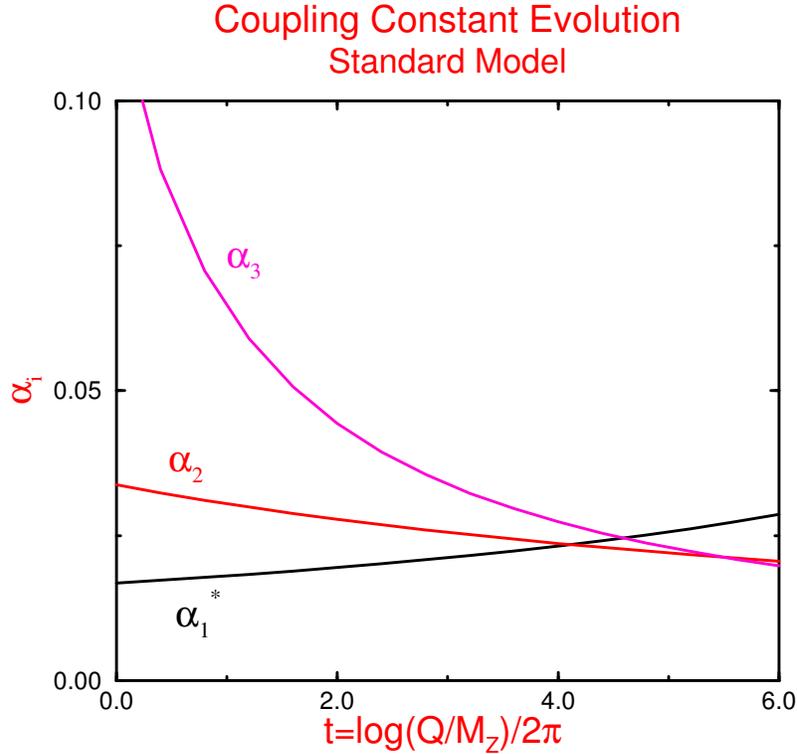
Figure 2.1: Evolution of the gauge coupling constants in the Standard Model from the experimentally measured values at the Z-pole. $\alpha_1^* \equiv 5/3\alpha_1$, since this is the relevant coupling in Grand Unified Theories [7].

observed. This puzzle suggests a primordial baryon asymmetry, which is explicitly ruled out by the SM. If baryon number truly is not a conserved quantity, then the SM must contain an underlying substructure.

In addition to the problems that are external to the SM, there exist a number of problems which arise in the formulation of the SM itself.

- **The Hierarchy Problem:** In the context of unification, which was introduced above, there is no explanation within the SM for why the Planck scale energies at which unification becomes apparent are so much larger than the other energy scales of the SM. The natural SM energy scales are the electroweak scale (~ 100 GeV) and the Higgs vacuum expectation value (~ 250 GeV), leaving an unoccupied "desert" of physics over 17 orders of magnitude in energy scale.

- **The Fine-Tuning Problem:** As given in Eq.2.18, the quartic self-interaction term generates a quadratically divergent contribution to the Higgs mass, arising from radiative quantum corrections. These divergences can remain finite if a large energy scale limit is introduced. However, the only other energy scale in the SM is the Planck scale, at which the corrections become enormous. In order for the Higgs mechanism to maintain unitarity[5] in longitudinal gauge boson scattering, the Higgs mass must remain smaller than ~ 900 GeV. Achieving this constraint requires cancellation by mass counter terms to one part in roughly 10^{16}, which, although not formally impossible, is regarded as an unacceptable fine-tuning of parameters [7].

- **Neutrino Mass:** In the SM, the neutrino is massless. However, present data on atmospheric and solar neutrinos as well as accelerator neutrino experiments indicate that neutrinos indeed have mass. As a consequence, evidence suggests that neutrinos exhibit a mixing behavior which is different than that observed in the quark sector. Simply adding neutrino mass terms to the SM Lagrangian causes undesirable behaviors of the theory, as well as requiring the inclusion of right-handed neutrinos. Aside from forcing a reformulation of the SM, this problem would introduce several independent parameters which are not predicted by the theory.

2.2.2 Potential Future Paths

In this section, I would like to introduce two potential alternatives to the SM of particle physics. Here, I limit myself to formulations which provide observables that can be tested in the foreseeable future. This choice, I regret, leaves out promising advances in high energy theory (such as String theory) and in cosmology theory. Although it is prediction which drives experimental innovation, in this context it seems best to focus on those theories which are experimental contemporaries of the SM.

- **New Minimal Standard Models:** There have been many attempts to "fix" the existing problems of the SM by writing new SM-like Lagrangians. These modified Lagrangian terms can be assembled to form a complete SM-like Lagrangian, incorporating terms for Dark Energy, Dark Matter, neutrino masses, and gravity [8]. Although these models are narrowly confined by experimental measurements, they should be investigated as they predict observables that can be probed at today's achievable energy levels. In particular, such a model can be formulated to include neutrino masses in two ways: with Dirac-type or Majorana neutrinos. Each of these choices has drawbacks. Choosing Dirac neutrinos does not provide a solution for universal baryon asymmetry. Choosing Majorana neutrinos creates a high-energy CP-violating parameter which can account for baryon asymmetry, but forces the possibility of neutrinoless double beta decay. Measurements of the parameters of a New Minimal Standard Model would allow a full evaluation of the phase space for direct extensions to the SM.

- **Supersymmetry:** Supersymmetry (SUSY) is a model which introduces a symmetry relating particles of different spin. Particles are combined in a *superfield* containing

---

[5]Unitarity is basically the conservation of probability. This constraint is included in the construction of any quantum field theory, including the SM.

two fields differing by one-half unit of spin. Thus each fermion of the SM is given a bosonic superpartner (forming a chiral superfield) and each boson of the SM is given a fermionic superpartner (forming a vector superfield). The addition of this symmetry provides solutions to many problems of the SM such as the Fine-Tuning problem, the Cosmological Constant problem, and unification (see Figure 2.2). However, particles in a superfield have the same masses and quantum numbers aside from the $1/2$ unit of spin. This is a problem as no scalar particles with the (small) masses of the SM leptons have been observed, directly or indirectly. Thus, SUSY must be a broken symmetry, and the mechanism for this breaking is not well described. The currently accepted means of breaking this symmetry requires the superpartners to the SM particles to have masses less than roughly 1 TeV. Furthmore, the simplest anomaly-free SUSY model requires two $SU(2)$ Higgs-type doublets of complex scalar fields, which predicts not one but five scalar Higgs fields. This, in itself, is not fundamentally a problem, but it introduces a second, unpredicted parameter to the Higgs sector describing the ratio of the vacuum expectation values for the two Higgs doublets [7]. Despite its deficiencies, SUSY provides a badly needed substructure for a SM-like gauge field theory. The predictions of this model provide an excitingly rich array of new physics which can be tested by today's experimental technology.
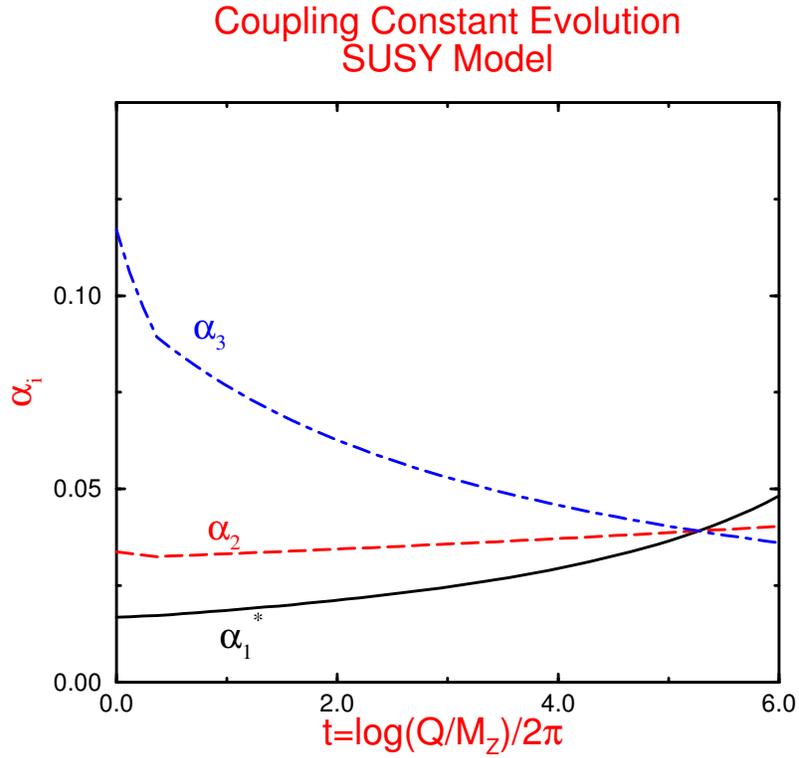
Figure 2.2: Evolution of the coupling constants in a low energy SUSY model from the experimentally measured values at the Z-pole. The SUSY thresholds are taken to be at 1 TeV. $\alpha_1^* \equiv 5/3\alpha_1$, since this is the relevant coupling in Grand Unified Theories [7].

Chapter 3

Fermilab and The DØ Detector

The data analyzed in this thesis were produced via the orchestrated interaction of two primary experimental instruments: the Fermilab Tevatron and the DØ detector. The data were recorded during Run II of the Tevatron in the years 2003-2004. This experimental procedure consists of the Tevatron preparing high-energy beams of protons and anti-protons which are brought into collision. These collisions occur at the center of two particle detectors: the collider detector at Fermilab (CDF) and the DØ detector. These detectors measure the final states of the particles that are produced in the interactions initiated in the colliding beams.

This chapter provides a basic description of the preparation of proton and anti-proton beams by the Tevatron accelerator complex and the means by which the DØ detector observes and records the collisions of these beams. This chapter also includes a discussion of proton anti-proton collisions and the resulting physics.

3.1   The Fermilab Accelerator Complex

The Fermilab Tevatron delivers proton and anti-proton ($p\bar{p}$) beams, each with energies of 980 GeV. In collider mode, these beams are brought into collision with a center-of-momentum energy of $\sqrt{s} = 1.96$ TeV. Each beam type starts at low energy and is brought to full energy through two different series of acceleration steps. Figure 3.1 gives a schematic description of the Fermilab accelerator complex. In this section, we will briefly describe the acceleration stages necessary to prepare the final beams. A very complete discussion of the Run II Tevatron acceleration complex operation can be found in [9].

The proton beam is accelerated in five distinct stages

1. The first stage is known as the preaccelerator. The proton beam originates as hydrogen gas which is ionized via a magnetron source to H^- ions. These ions are magnetically selected and accelerated to 750 keV using an electrostatic Crockroft-Walton accelerator.

2. The proton linac (a linear accelerator) accepts the 750 keV H^- ions and uses radio-frequency (RF) fields to accelerate the ions to an energy of 400 MeV over about 150 m.

3. Next in line is the Booster, which is a synchrotron ring with a radius of 75 m. Before injection to the Booster, the 400 MeV H^- ions are passed through a carbon foil that
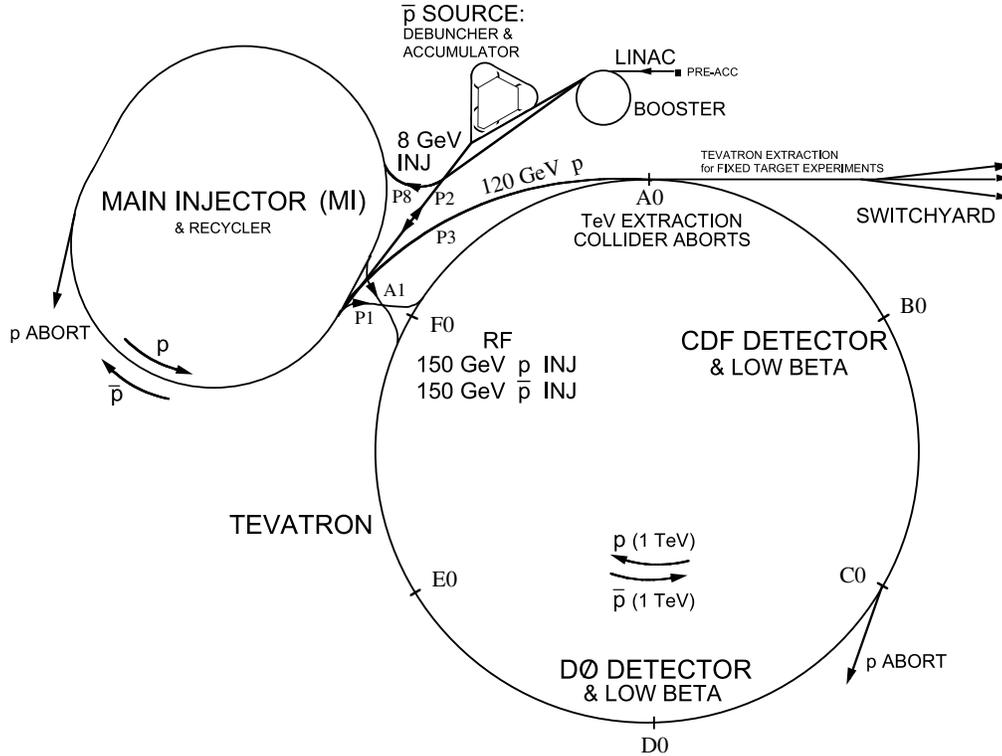
Figure 3.1: The accelerator complex at the Fermi National Accelerator Laboratory.

strips off the electrons, leaving bare protons. The protons are injected into the Booster and constrained to a circular path using dipole magnets. Magnets of higher-order poles are used to maintain a focused beam. The Booster uses RF cavities to accelerate the protons over the course of about 20,000 revolutions[1] to 8 GeV. During the acceleration process, the protons are grouped into a pulse train containing five to seven bunches.

4. The Booster injects its 8 GeV proton beam to the Main Injector, which is a circular synchrotron with a radius of about 500 m. The Main Injector coalesces the protons from the Booster into one bunch and accelerates them to either 120 GeV or 150 GeV, depending on their target location. The 150 GeV proton bunches are injected to the Tevatron, while the 120 GeV bunches are delivered to the anti-proton facility.

The anti-proton beam is the largest limiting factor in the operation of the Tevatron.

---

[1]Over the course of the acceleration, the RF frequency and magnetic field strengths are increased synchronously to maintain a circular orbit. Hence the name synchrotron.

Although preparing a proton beam for the Tevatron is (comparatively) quick and easy, obtaining enough anti-protons for collider operation takes more time and care.

1. In the $\bar{p}$ stacking mode, the Main Injector delivers bunches of 120 GeV protons to the anti-proton source.

2. The proton beam from the Main Injector is directed to a nickel target. The proton-Ni collisions create many secondary particles, including anti-protons. These secondaries are collected via a lithium lens and 8 GeV anti-protons are selected using a magnetic dipole spectrometer. The efficiency of this process is about 15 anti-protons produced for every million protons on target.

3. The 8 GeV anti-protons are transferred next to the Debuncher, which is a triangular storage ring of about 520 m in circumference. Here the bunch structure from the Main Injector is removed and the transverse momentum profile of the anti-protons is reduced. Using stochastic cooling[2], the anti-protons are induced into an ideal orbit.

4. Next, the anti-protons are transferred to the anti-proton Accumulator, which is housed along with the Debuncher. Here, the anti-protons are bunched and stored until a sufficient number has been prepared. The anti-protons are maintained at an energy of 8 GeV using RF cavities in both the Accumulator and Debuncher.

5. Once the anti-protons have achieved the same timing structure as the protons in the Main Injector, they are transferred over to the Main Injector and accelerated to 150 GeV.

When the accelerator complex is ready, the Main Injector delivers 36 bunches of protons (about 5×10^{12} protons per bunch) to the Tevatron. The Tevatron is a circular synchrotron with eight accelerating RF cavities and is about 1 km in radius. The Tevatron is the only Fermilab accelerator with superconducting magnets, producing fields of up to 4 Tesla. After proton injection, anti-protons are transferred from the Main Injector four bunches at a time (up to 36 bunches). Then, the Tevatron accelerates the proton and anti-proton bunches in two oppositely rotating beams up to a final energy of 980 GeV. The high-energy beams are then squeezed to a high density using focussing magnets and brought into collisions with a center-of-momentum energy of 1.96 TeV. The Tevatron operates with a 36×36 $p\bar{p}$ bunch structure that creates a 396 ns bunch crossing. These collisions occur at two points on the Tevatron ring, referred to as interaction regions. One of these interaction regions is occupied by the DØ Detector.

3.2 Cross Sections, Coordinates, and Collisions

This section briefly describes the physics of colliding protons and anti-protons. We also include a discussion of the rate of interactions between two colliding beams and the coordinate

---

[2]Stochastic cooling is a beam-cooling method in which the beam orbit is measured at one point on the ring and altered in a downstream part of the ring. By sending the correction information across the center of the ring, it can arrive before the particles, which take a longer path along the ring.

systems which are used to measure the final states of such collisions. Also included here is a phenomenological discussion of the basic interactions of elementary particles with bulk matter.

3.2.1 Cross Sections and Luminosities

The simplest description of two colliding particles (either one stationary and one in motion, or both in motion) is in the center-of-momentum frame (in both classical and relativistic descriptions). This frame of reference occurs when there is zero total momentum between the two particles, and is almost always moving with respect to the laboratory frame, the frame of a stationary observer. In the scenario that the two particles are produced with identical (but perfectly opposite) momenta in the laboratory frame, the two frames are equivalent. It is this scenario that is created at the Tevatron. In addition to simplifying the description of the collision, this preparation maximizes the total collision energy in the laboratory frame. Incidentally, the center-of-momentum frame is also historically referred to as the center-of-mass frame, but we'll use the more intuitive center-of-momentum terminology.

The rate of interactions between two identical, colliding beams of particles is classically described as

$$\mathcal{R} = \rho \, v \, \sigma \tag{3.1}$$

where ρ is the density of the (combined) beams, v is the relative velocity of the beams, and σ is the classical cross section for the process (nominally the cross-sectional area of the particles). In this classical description, the cross section is fixed and the rate of interactions is linear with both velocity and beam density. This equation can be equivalently written in the form

$$\mathcal{R} = L \, \sigma \tag{3.2}$$

by introducing the quantity, L, or the instantaneous luminosity. The instantaneous luminosity is a measure of the particle flux of the colliding beams and (again, for identical beams) is proportional to the square of the number of beam particles passing through a unit area per unit time.

In a SM description of collisions, the classical description of the cross section manifestly failes and thus must take on a new definition. This failure occurs because SM particles are point particles and occupy a single space-time point. The overlap of the spatial extent of two point particles is a delta function and the interaction probability is vanishingly small. Furthermore, the quantum-mechanical description of such interactions provides for many different possible interaction types. Thus, the outcome of each interaction isn't fixed, but is selected from a large list of outcomes with different probabilities based on the properties of the incident particles. Finally, the SM description of the cross section requires it to change as a function of the center-of-momentum energy in order to conserve probability. Hence, when colliding relativistic beams of SM particles, we are lead to a new description of the cross section. In this description, one specifies a specific final state or ensemble of final states and an energy-dependent cross section is calculated using the rules defined by the SM. This cross section is no longer related to the size of the particles, but represents an interaction

probability per unit flux. Historically, cross sections were measured in units of barns, where 1 barn $= 10^{-24}$ cm^2, as this was roughly the cross-sectional area of a small atomic nucleus. This notation is still used, although cross sections of interest are typically much smaller and are usually on the order of pico-barns (pb), or 10^{-36} cm^2.

In the evaluation of a dataset, the quantity of interest is the number of events recorded. This number is dependent on the duration of time collisions are observed. Thus, a new quantity is introduced, the integrated luminosity $\mathcal{L}$

$$\mathcal{L} = \int L \, dt \tag{3.3}$$

which is the time integral of the instantaneous luminosity during the experimental exposure period. This quantity is frequently referred to as simply the luminosity and we will adopt this convention. To simplify calculations of event numbers, the luminosity is generally measured in units of inverse barns, with common luminosities on the order of inverse pico-barns (pb^{-1}) and inverse femto-barns (1fb^{-1} = 1000pb^{-1}).

3.2.2 Coordinate Systems

The DØ detector uses a right-handed coordinate system. In this system, the $+x$ axis is defined by a vector pointing radially outward from the center of the Tevatron ring (east). The $+y$ axis points vertically and the $+z$ direction is tangent to the direction of proton travel at the center of the detector (south), thus completing the right-handed coordinate system.

DØ has roughly cylindrical symmetry and particle collisions exhibit spherical symmetry (in their rest frame), motivating the choice of a combination of spherical and cylindrical coordinates (θ, ϕ, z). The polar angle θ is defined from the $+z$ axis and the azimuthal angle ϕ is defined with $\phi = 0$ at the $+x$ axis and $\phi = \pi/2$ at the $+y$ axis. As the variable θ is not Lorentz invariant (as ϕ is), it becomes difficult to work with at the large Lorentz boosts common at the Tevatron. Thus, a Lorentz invariant variable rapidity, y, is introduced

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{3.4}$$

where p_z is the particle momentum along the z axis. In the zero-mass limit, this variable transforms to the new variable pseudorapidity, η,

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \tag{3.5}$$

As the position of the beam collision isn't constrained to occur exactly at $z = 0$, it becomes useful to define η in two ways: physics-η and detector-η. Physics-η is defined with respect to the primary interaction and detector-η is defined with respect to the center of the detector $(x, y, z = 0)$. Solid angles are measured using the two Lorentz invariant angles (η, ϕ)

$$\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} \tag{3.6}$$

3.2.3 $p\bar{p}$ Collisions

Most of the $p\bar{p}$ collisions initiated at the Tevatron result in a very small momentum exchange between the two hadrons. These interactions do not produce physics useful for probing SM physics. Occasionally, a larger momentum transfer occurs and the original proton and anti-proton are broken apart. A *parton* (a quark or gluon constituent) of the proton exchanges a force carrier boson with a parton in the anti-proton to create a *hard-scatter* reaction. The fragments of the proton and anti-proton recieve little transverse momentum in the collision and continue along nearly parallel to the beam-line, while the ejected partons enter the detector [10].

Such hard-scatter interactions of the protons can result in the production of any of the SM particles. However, many of the SM particles are unstable and thus decay rapidly to lighter particles. Common examples of this are W and Z bosons, as well as the top quark. Generally, only electrons, muons, neutrinos, photons, and a few bound states of light quarks (u,d,s) live long enough to reach the detector.

The parton structure of the proton and anti-proton (hadrons) leads to a careful consideration for the physicist. The hadrons are accelerated to a uniform 980 GeV by the Tevatron, but this energy is shared by the constituent partons, which are in continuous relative motion within the hadrons. Thus, the collision of two partons may not occur at 1.96 TeV and may not have zero net momentum along the z-axis. However, to a good approximation the net transverse momentum of the event is zero, which leads to the useful variables of transverse momentum ($p_T = p\, sin(\theta)$) and transverse energy ($E_T = E\, sin(\theta)$) for each particle.

3.2.4 Particle Interactions with Matter

The primary means of measuring the properties of particles is to look for energy deposited as those particles pass through matter. Different particles lose energy in different ways and the energy loss of particles can depend highly on its initial energy. In this section, we'll briefly review the means by which particles are detected. This discussion is largely derived from the detailed treatment in [11].

Ionization and Excitation

As a charged particle passes through matter, its charge will interact electromagnetically with the electrons of the medium's atoms. These interactions result in a net transfer of momentum from the incident particle to the medium. This momentum transfer represents energy loss for the particle and serves to either ionize or excite the atom it interacts with. Such interactions are mediated by photons and the induced energy loss can be phenomenologically described by the Bethe-Bloch equation [11]. This interaction depends on the momentum of the indicent particle and the charge and mass densities of the matter through which it passes.

In the event that the momentum transfer is large enough to eject an electron from its atomic orbit, ionization occurs. The resulting free electron and cation (referred to as a *hole*) can be collected via an electric field, thus signaling the passage of a charged particle. This technique is capitalized on in solid state semiconductor (silicon) detectors.

When the momentum transfer is insufficient to ionize an atom, an atomic electron can absorb this energy and be promoted to a higher-energy orbital above its ground state. The

relaxation this excitation results in photon emission. This photon emission (referred to as scintillation light) is used in scintillator detectors to indicate the passage of a charged particle.

Coulomb Scattering

In addition to interactions with atomic electrons, charged particles can interact with the nuclei of the atoms in a material. This Coulomb scattering results in a deflection of the path of the particle, with almost no energy loss. The scattering angle falls with increasing particle momentum and radiation length of the material in question, but multiple Coulomb scattering can result in a significant alteration in the original trajectory of a particle. Coulomb scattering represents an undesireable matter interaction from the point of view of the physicist as it changes the trajectory of a particle in an unpredictable way.

Particle Cascades

As high-energy particles (we now include neutral particles such as photons and π^0 mesons) pass through matter, interactions can result in the production of secondary particles with lower energies than the original particle. Subsequent interactions create cascades of secondaries which lose their energy via ionization and excitation in a *particle shower*. In the presence of sufficient amounts of material, such particle showers result in the full energy depletion of the incident particle, and thus a means of measuring the particle's energy. This process of absorbing a particles energy for measurement is referred to as calorimetery. The mechanism of shower formation varies for different particles, so we will describe them separately.

Electrons with energies above a few hundred MeV dominantly lose energy via photon emission, or bremsstrahlung. The energy loss is exponential with increasing distance traveled in a medium. The characteristic distance of this energy loss is referred to as the *radiation length*, X_0, given approximately as

$$X_0 = \frac{716.4\ A}{\rho Z(Z-1)\ln\left(287/\sqrt{Z}\right)}\ \text{cm} \tag{3.7}$$

where A is atomic mass of the material, Z is the atomic number, and ρ is the mass density. The energy loss is inversely proportional to the square of the particle's mass, and thus bremsstrahlung losses for muons and hadrons are typically neglected at the energies of the Tevatron. When encountering the electromagnetic field of a nucleus, high-energy photons can convert to electron-positron pairs. The energy loss for this process can be parameterized in terms of the radiation length of the material the photon is traversing, $\frac{7}{9}X_0$. The electron-positron pair will then each lose energy via bremsstrahlung, creating more photons, thus propagating an electromagnetic shower.

Although hadrons can lose energy via ionization (provided they are charged), they also interact with the nuclei of matter via the strong force. These strong interactions are generally inelastic, producing secondary quarks and gluons which hadronize, producing a hadronic shower. The characteristic length for hadronic showering is the nuclear interaction length, λ_I, given approximately by

$$\lambda_I = \frac{35 \, A^{1/3}}{\rho} \; \text{cm} \tag{3.8}$$

Hadronic shower formation can give rise to particles which decay or interact electromagnetically, producing photons and electrons in the hadronic shower. As the nuclear interaction length is generally much larger than the radiation length for a given material, electromagnetic and hadronic showers have much different signatures in a calorimeter.

3.3 The DØ Detector

DØ is a multi-purpose detector designed to identify the elementary particles that are produced in the $p\bar{p}$ collisions at the Tevatron [12, 13]. The detector, shown in Figure 3.2, was built in the late 1980's and early 1990's, operated from 1992 to 1996 in Run I of the Tevatron, and ultimately upgraded to handle the increased luminosities in Run II of the Tevatron [14]. The detector weighs approximately 5,600 tons, and its dimensions measure roughly $13 \times 11 \times 17$ m^3. The primary utility of the detector is to measure the properties of electrons, muons, and stable hadrons and mesons. To perform these measurements, the DØ detector uses three main detection components: tracking, calorimetry, and muon identification, all symmetric about the Tevatron beam line. At the center of the detector are the tracking detectors, which are designed to measure the trajectories of particles while minimizing scattering and energy loss. These requirements are met by using the least amount of material possible to obtain an accurate measurement of the ionization induced by the passage of charged particles. Outside the tracking detectors is the calorimeter, whose job it is to make a measurement of each particle's energy. This is done by using dense materials sufficient to absorb the full energy of most incident particles, while making a measurement of the energy deposition. Muons, unlike electrons and strongly-interacting particles, do not interact appreciably with the calorimeter and must be detected in the muon system outside the calorimeter. The muon system uses tracking detectors designed to measure the trajectory and charge of the muon. Neutrinos interact so rarely with matter that their presence must be inferred via a net imbalance in transverse momentum. This section will briefly describe the DØ subdectectors and their operation. More complete descriptions can be obtained in [13] and [14].

3.3.1 Inner Tracking Detectors

The tracking detectors are constructed directly outside of the Tevatron beam line and consist of two subsystems: the Silicon Microstrip Tracker (SMT) and the Central Fiber Tracker (CFT). These tracking detectors are surrounded by a superconducting solenoid magnet, providing a 2 Tesla field parallel to the beam line. The tracking detectors perform charged particle detection up to $|\eta| = 3.0$. The solenoidal field causes electrically charged particles to follow a curved path, with the curvature inversely proportional to the particle's momentum. This allows for precise measurements of momentum and a determination of the particle's charge. Futhermore, the tracking detectors provide a means to measure the hard-scatter vertex and any secondary vertices produced by the decay of short-lived particles.
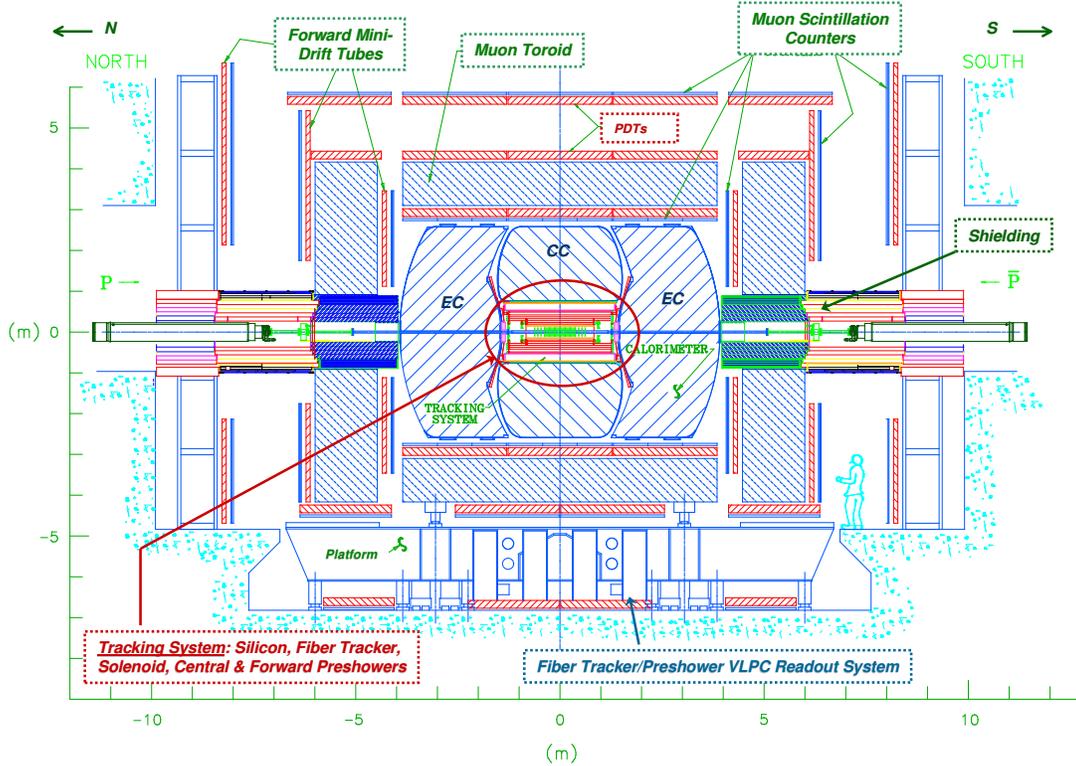
Figure 3.2: Schematic view of the DØ detector, illustrating the layout of the three major subdetector components: the central tracking, the calorimeter, and the muon system. Also shown is the support structure and the Tevatron beam line.

Silicon Microstrip Tracker

The innermost tracking system at DØ is the SMT [15], which is the closest detector component to the beryllium beam pipe[3] of the Tevatron. The SMT provides high-resolution measurements of the paths of charged particles leaving the interaction region. The large z distribution of the $p\bar{p}$ interaction region ($\sigma_z \simeq 26$ cm) provides a challenge for designing a detector in which tracks are predominantly perpendicular to detector surfaces. This challenge motivates a detector geometry consisting of six barrels and sixteen disks of silicon wafers, creating a tracking coverage out to $|\eta| = 3.0$. A schematic of the SMT geometry is shown in

---

[3]The Tevatron beam pipe is nearly all made of steel. However, the sections around the collision points are made of beryllium. Beryllium has a much smaller density and Z value than iron, reducing the probability that particles will interact with the beam pipe.

Figure 3.3.
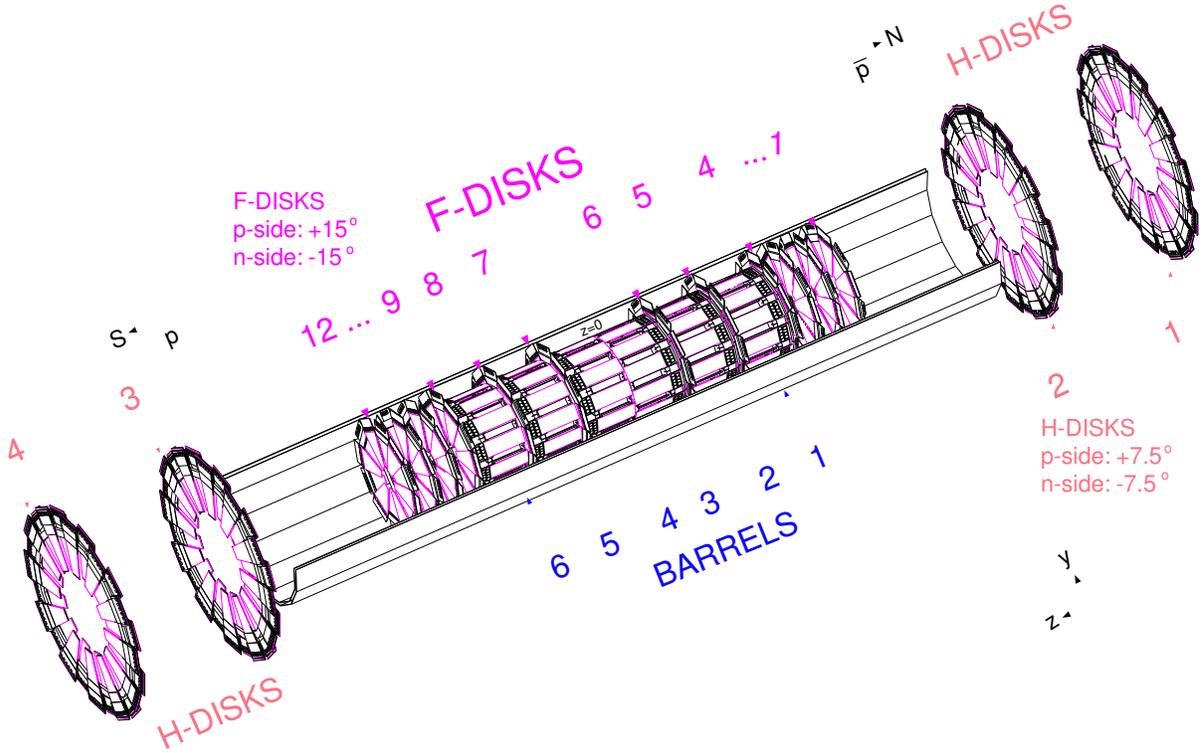
Figure 3.3: Schematic view of the SMT detector, illustrating the geometry of the barrel, F-disks, and H-disks.

The six barrel segments are 12 cm long and made up of four concentric layers of silicon wafers, allowing for $r - \phi$ measurements of central tracks. Each layer of silicon is slightly overlapped to ensure full ϕ coverage. The six barrels provide coverage of the $|\eta| < 1.1$ region. Along the axis of the barrels are twelve 8 mm-thick disks, referred to as F-disks. The disks are made of twelve overlapping, double-sided silicon wedges, creating an annulus with central radius 2.6 cm and outer radius 10.5 cm. Two larger disks, referred to as H-disks, are placed on either end of the detector. These H-disks are made of 16 overlapping, single-sided silicon wedges, each forming an annulus with inner radius 9.5 cm and outer radius 26 cm. The F-, and H-disks together provide $r - z$ and $r - \phi$ tracking coverage out to $|\eta| = 3.0$.

The SMT barrels and disk wedges are made of 300 μm-thick silicon wafers consisting of a n-type/p-type silicon interface (p-n junction). Interspersed on both sides of the silicon are thin conducting readout strips with a pitch varying from 50 μm to 153.5 μm. As charged

particles pass through the silicon wafer, ionization produces electron-hole pairs. An applied bias voltage pulls these pairs (in opposite directions) to the readout strips, and the collected charge is stored in a capacitor array until the information is ready to be processed.

Central Fiber Tracker

The Central Fiber Tracker (CFT) lies immediately outside the SMT and provides tracking coverage up to $|\eta| < 2.0$ [16]. The combined measurements of the SMT and CFT allow for improved tracking quality not achievable by either detector alone. The CFT consists of eight carbon fiber cylinders holding layers of scintillating fibers. Each cylinder supports a doublet layer of fibers oriented parallel to the beam line (*axial* fibers). The odd numbered cylinders (counting from the inside to outside) hold an additional doublet offset at alternating angles of $\pm 3^o$ (*stereo* fibers). The axial fibers provide ϕ measurements at a fixed radius and, when combined with the stereo fibers, can provide a measurement of z. Each fiber consists of a 775 μm polystyrene core that is doped with fluorescing molecules with peak emission at 535 nm. Surrounding the core are two 15 μm layers of cladding (acrylic and fluro-acrylic), increasing the light-collection efficiency. In total, the CFT contains 71,680 fibers. A quarter-view schematic of the CFT is shown in Figure 3.4.

As charged particles pass through the fibers, scintillation light travels their length in both directions. The fibers, which range in length from 166 cm for the innermost cylinder to 257 cm for the outermost cylinders, have an aluminum mirror coating at one end to reflect photons back into the fiber. The other end is joined to clear fibers which guide the scintillation photons to a solid-state silicon device called a Visible Light Photon Counter (VLPC). Photons incident upon the surface of the VLPC are converted to electron-hole pairs, which are subsequently collected via a 6 V bias voltage. The VLPC's are grouped together in "cassettes" of 1024 VLPC's which are kept in liquid helium dewars to reduce electronic noise, providing single-photon resolution.

3.3.2   Central and Forward Preshower Detectors

As noted above, the SMT and CFT are nested within a superconducting solenoid. This solenoid is constructed of very dense material and is uninstrumented. This solenoid material can interact with particles, causing early showering before the calorimeter that impacts the measurement of the particle's energy. To mitigate this problem, the region between the solenoid and the central calorimeter cryostat has been instrumented with a preshower detector. For uniformity, a complimentary preshower detector is constructed in the forward regions. These preshower detectors are designed to improve calorimetry measurements and are also sensitive enough to aid in tracking measurements.

The Central Preshower Detector (CPS) is a cylindrical detector consisting of three layers of scintillating strips which cover the region of $|\eta| < 1.2$ [17]. The scintillating strips have a triangular cross section with a 7 mm base and a 1 mm hole containing a wavelength shifting fiber. The innermost layer of strips is arranged axially, while the two outer layers are arranged at stereo angles of $\pm 23^o$. The geometry of the CPS and the orientation of the scintillating strips are shown in Figure 3.5. The Forward Preshower Detector (FPS) is very similar to the CPS in its construction, consisting of two layers of stereo scintillation
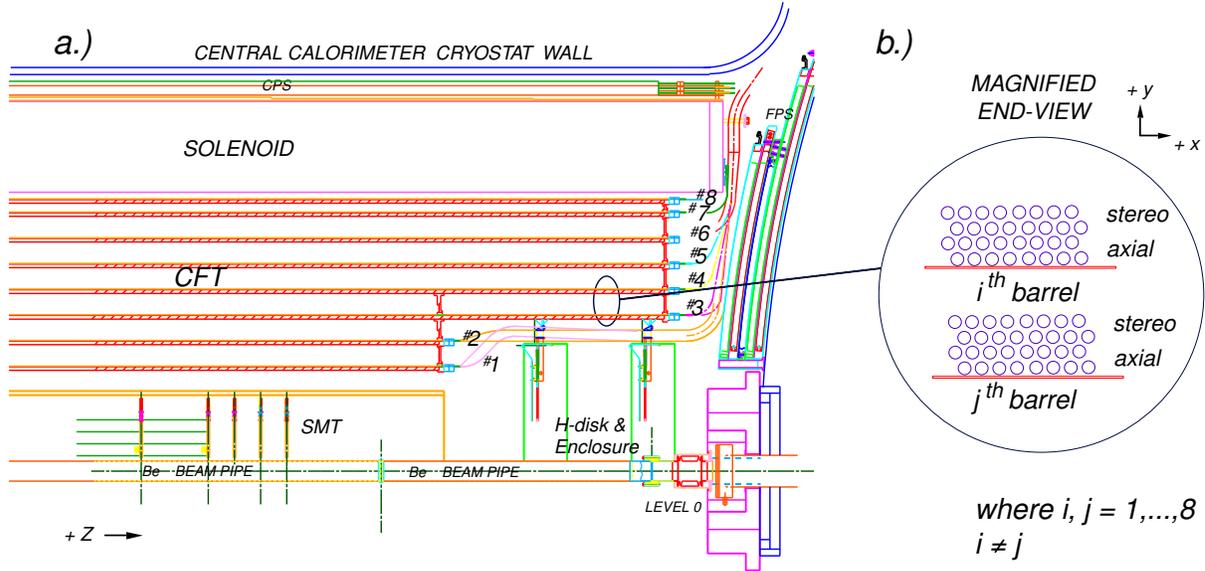
Figure 3.4: a) A quarter r-z view of the CFT detector, showing the nested eight barrel design. b) A magnefied r-ϕ view of the two ribbon doublet layer configuration for two different barrels. layering.

strips and no axial layer [18]. The FPS is mounted on the inner faces of the end calorimeter cryostats and is shown schematically in Figure 3.6. The preshower detectors are read out in the same manner as the CPS, with the scintillation light being collected via a clear fiber and transmitted to VLPC's for charge conversion.

3.3.3 The Calorimeter

The DØ calorimeter lies outside the solenoid and measures the energies of electromagnetic particles (electrons, photons) and hadrons. This measurement is made by inducing interactions with incident particles via the material of the calorimeter, creating showers of secondary particles which lose energy through ionization in the calorimeter's active medium. A measurement of a particle's total energy is made when the showering process is fully contained. The calorimeter is a compensating[4], sampling calorimeter in which liquid argon is used as the active medium and depleted uranium (as well as copper and steel) is used as an absorber material [12]. As it completely surrounds the inner detectors, the calorimeter has a modular

---

[4]The term *compensating* refers to the fact that the ratio of the electromagnetic to hadronic responses is nearly one. This is not true for all calorimeters and depends on the material used as an absorber
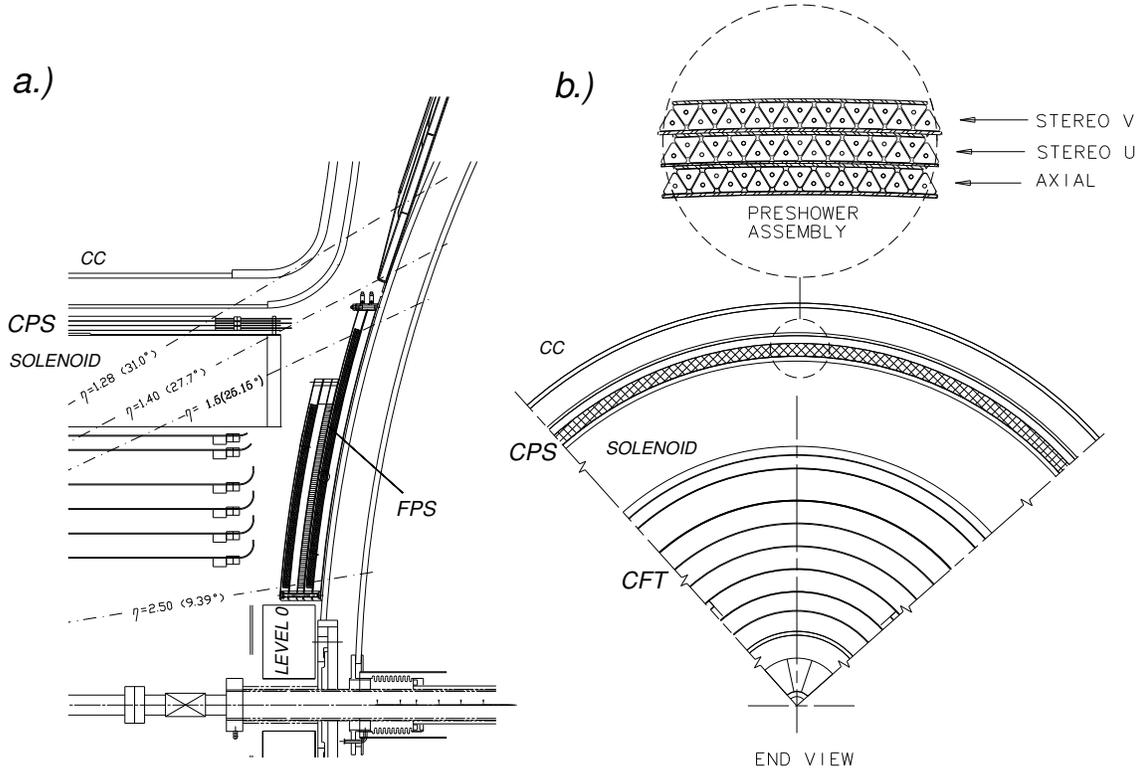
Figure 3.5: a) An r-z semi-quarter view of the CPS detector. b) A cross-sectional r-ϕ view of the CFT and CPS detectors. The inset shows a magnified view of the dove-tailed scintillating strips of the CPS.

design to provide access to the inner regions. This design consists of three cryostats, which are vessels containing the calorimeter and the cryogenics required to maintain the liquid argon at a constant temperature, and is shown in Figure 3.7.

The calorimeter is comprised of three distinct modules: the Central Calorimeter (CC) covering the region $|\eta| < 1.2$ and two End Calorimeters (EC North and EC South) that extend coverage to $|\eta| \simeq 4.5$. The calorimeter modules themselves are further segmented into three sections. In order of increasing radius, these are the electromagnetic (EM), fine hadronic (FH), and coarse hadronic (CH) sections.

The EM sections consist of four layers of depleted uranium absorber plates, each 3-4 mm thick. The FH section contain three (CC) or four (EC) layers of 6 mm-thick uranium-niobium (2%) alloy absorber plates. The outer CH section has one 46.5 mm-thick absorber plate made of copper (CC) or steel (EC). This layered structure is summarized in Table 3.1
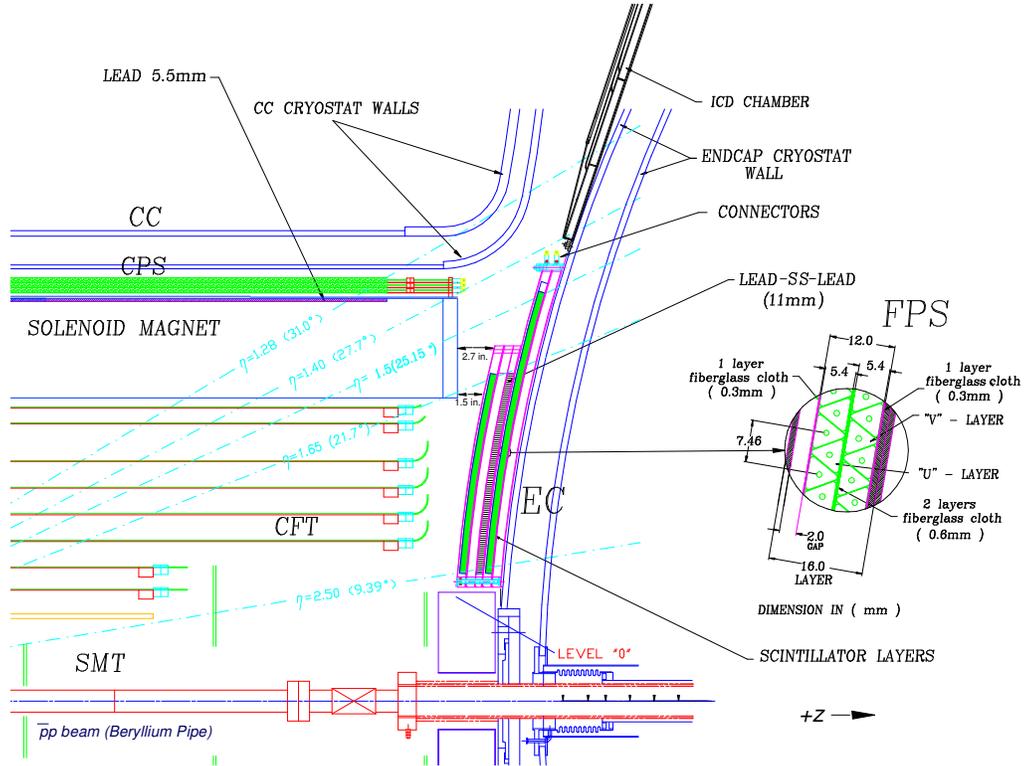
Figure 3.6: One quarter r-z view of the FPS detector. The inset shows details of the FPS scintillator layers.

in terms of the radiation and nuclear interaction lengths of each layer.

Each calorimeter layer is segmented into a set of *readout cells*. These cells are $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in size, except in the third EM layer where the segmentation doubles. These readout cells are grouped radially to form a $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ readout geometry referred to as a *tower*, shown in Figure 3.8. The readout cells consist of a group of adjacent *unit cells* immersed in the liquid argon of the calorimter. Each unit cell is a copper pad insulated with G10 plastic covered in a resistive epoxy coating. The resistive coating is held at a high voltage (~ 2.5 kV). The showering particles in the calorimeter ionize the liquid argon and the liberated electrons are drawn to the resistive coat. Via capacitive coupling, an image charge is induced on the copper pad. Readout electronics sample the charge on the pad, converting it to an analog signal proportional to the ionization energy recorded.
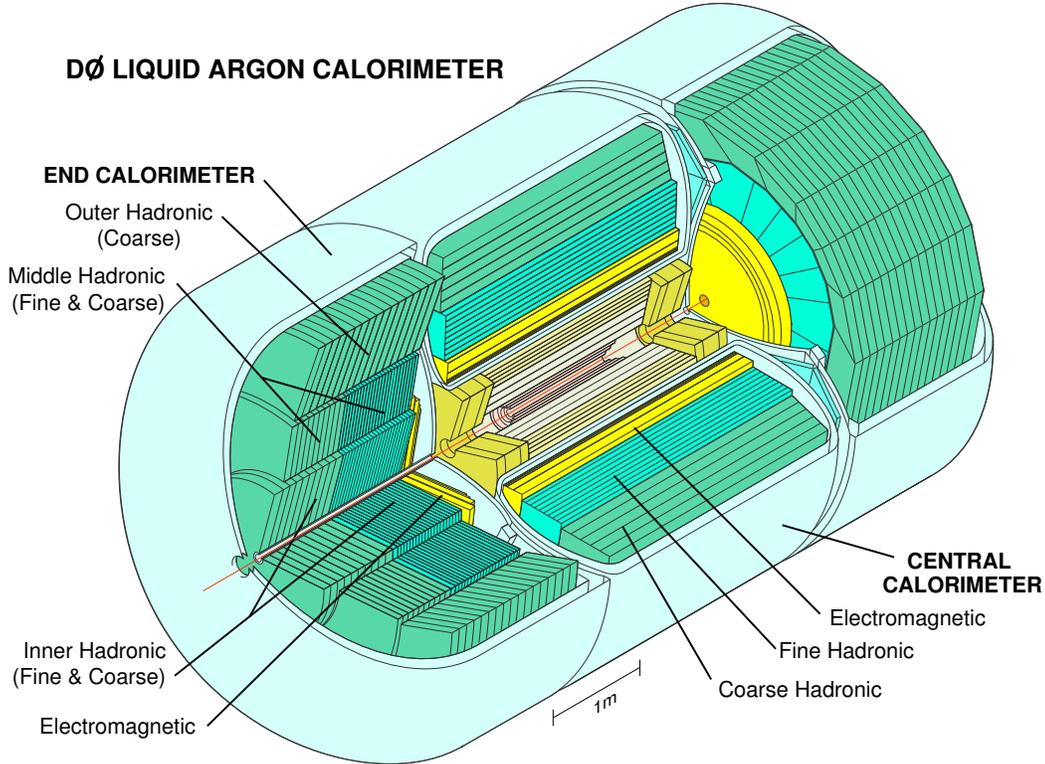
Figure 3.7: Three-dimensional cutaway view of the DØ calorimeter, showing the orientation of the three cryostats. Also shown is the segmentation of the calorimter layers.

3.3.4   Intercryostat and Massless Gap Detectors

As evident in Figure 3.8, there is an uninstrumented region between the CC and EC covering the region $1.1 \simeq |\eta| \simeq 1.4$. The material in this region (cryostat walls, support structures, cabling...) can participate in shower evolution, and thus can impact jet measurements. To augment the shower sampling in this region, scintillator detectors have been mounted on the EC cryostat walls facing the gap. Each intercryostat detector (ICD) consists of 384 scintillator tiles of the same size as the calorimeter cells, $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Separate single-cell structures, called massless gaps, are installed in the gap region to make further measurements of shower formation [12]

	EM	FH	CH
CC Depth	2.0,2.0,7.0,10 X_0	1.3,1.0,0.9 λ_I	3.2 λ_I
EC Depth	0.3,2.6,7.9,9.3 X_0	1.2,1.2,1.2 λ_I	3.6 λ_I

Table 3.1: Calorimter layer depths in terms of the radiation and nuclear interaction lengths of each layer.

3.3.5   The Muon System

Located outside the calorimeter, the DØ muon detector system is physically the largest sub-detector and is designed to detect the passage of muons while making a measurement of their momenta [14]. As mentioned in Section 3.2.4, the large muon mass ($\sim 200\ m_{electron}$) causes muons to lose little energy via Brehmsstrahlung. Energy loss for muons occurs primarily via ionization and excitation, which are low-energy loss processes. Thus, muons with energies above ~ 3 GeV exit the calorimeter and enter the muon system. The muon system consists of three primary components

- **W**ide **A**ngle **Mu**on **S**pectrometer (WAMUS) covering $|\eta| < 1$

- **F**orward **A**ngle **Mu**on **S**pectrometer (FAMUS) covering $1 < |\eta| < 2$

- A 1.8 Tesla iron toriodal magnet

The WAMUS consists of two types of detector components: proportional drift tubes (PDT's) and scintillator tiles. These components are arranged in three layers, referred to as A-,B-, and C-layers. The A-layer is located inside the toroid and the B- and C-layers are outside the toroid. The FAMUS has a similar structure using mini drift tubes (MDT's) and scintillator pixels. The geometry of the muon system can be seen in Figure 3.9.

The muon drift tubes are filled with a gas mixture (80% argon, 10% CH_4, 10% CF_4) which is easily ionized by the passage of charged particles. Each tube contains a gold anode wire held at high voltage (relative to cathode pads on the top and bottom of the tube). The ionization is collected at the wire and converted to a signal via readout electronics, allowing for good position measurements but poor timing measurements (~ 500 ns resolution). The scintillators provide additional spatial information and ~ 10 ns resolution time measurements, allowing for cosmic ray rejection. The iron toroid serves two purposes. First, it acts as an extra layer of dense shielding, effectively containing any hadronic showers which are not contained in the calorimeter. And second, its magnetic field provides a measurement of the muon's momentum by comparing the position of hits in the inner layer to the outer layers. Whenever possible, the high-resolution tracks of the inner tracking detectors are used for making muon momentum measurements.

3.3.6   Luminosity Monitor

The Luminosity Monitor (LM) is the subdetector responsible for measuring the instantaneous luminosity being delivered to the DØ experiment. As the instantaneous luminosity drops
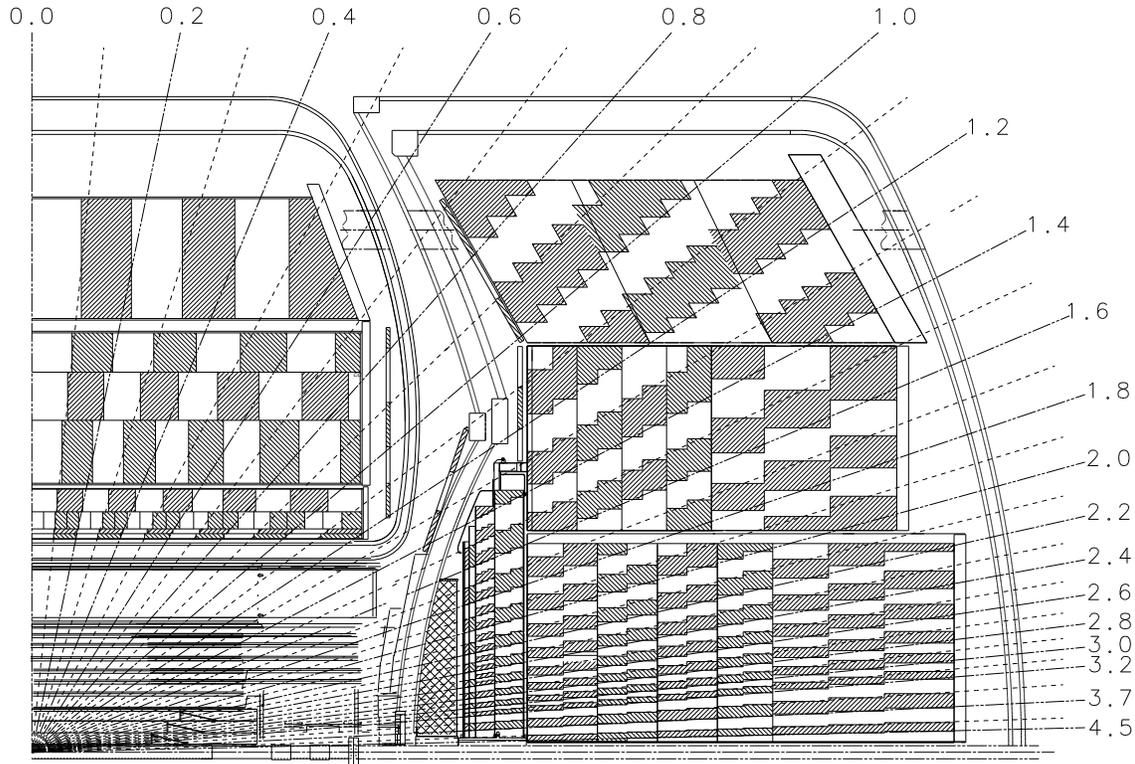
Figure 3.8: A one quarter r-z view of the calorimeter. Lines extending from the center of the detector denote the η coverage of projected readout towers.

steadily during beam collisions, an accurate measurement of the instantaneous luminosity allows for optimization of data taking rates and a reliable normalization measurement for specific event rates. The LM is constructed of two hodoscopes of plastic scintillation pixels mounted on the front faces of the EC calorimeters, as shown in Fig. 3.6 and labeled as "Level 0". The LM spans the region $2.7 < |\eta| < 4.4$ and measures the inclusive rate of inelastic $p\bar{p}$ scattering by detecting charged particles from the interaction region [45].

3.3.7   Trigger and Data Acquisition Systems

With a Tevatron beam-crossing time of 396 ns, there are roughly 2.5 million possible events every second. Most of these events are due to low-p_T, non-diffractive $p\bar{p}$ scattering. These type of events have been studied extensively in the past and are not considered a physics priority at DØ. The task remains to identify the interesting events and record them. Iden-

tification of these events is performed using a technique known as *triggering*, which proceeds by matching event properties to a predefined set of patterns which are characteristic of the physics processes of interest. However, physical constraints limit the rate at which events can be triggered and recorded. First, the frequency at which the detector can be read out sets an upper limit on the event examination rate at about 10 kHz. Second, the maximum event processing and storage rate sets an upper limit on the rate of events which are ultimately recorded at about 100 Hz. The DØ detector utilizes a trigger structure comprised of three distinct stages, intuitively referred to as Level 1 (L1), Level 2 (L2), and Level 3 (L3) triggers. Each trigger level is increasingly more refined than the previous, creating a filtering system which maximizes the efficiency for identifying interesting physics events while satisfying the event rate constraint. The structure of this data acquisition path is shown in Figure 3.10.

The Level 1 trigger, shown in Figure 3.11, consists of algorithms implemented in the firmware of Field Programmable Gate Arrays (FPGA's). Condensed information from the calorimeter, preshower, CFT, and muon detectors is processed in parallel to make a preliminary triggering decision about each event. The latency for the L1 trigger is approximately 4.2 μs, allowing for a small deadtime compared to the maximum readout rate of the detector of ~ 10 kHz. The output of L1 is used to limit the rate for accepted events to ~ 1.5 kHz.

If the Level 1 trigger issues an accept, the Level 2 trigger queues the event for processing. The L2 trigger combines a hardware trigger scheme (as in L1) with a software trigger scheme. Different pieces of information from the subdetectors are correlated to construct basic physics objects (electrons, muons, tracks, jets) and this information is combined to make a global L2 trigger decision, further reducing the event rate to ~ 800 Hz.

When the L2 trigger system issues an accept, the event is passed to the L3/Data Acquisition (DAQ) system. At this point, the full detector information is collected from the subdetector read out crates (ROC's). As shown in Figure 3.12, this event information is then routed to one of ~ 125 Linux PC's in the L3 farm. Each PC processes the data with an identical copy of a filtering software package, reconstructing refined physics objects and applying sophisticated algorithms to arrive at a final trigger decision. Events which recieve a L3 accept are sent to a collection machine and are written to tape for future analysis.
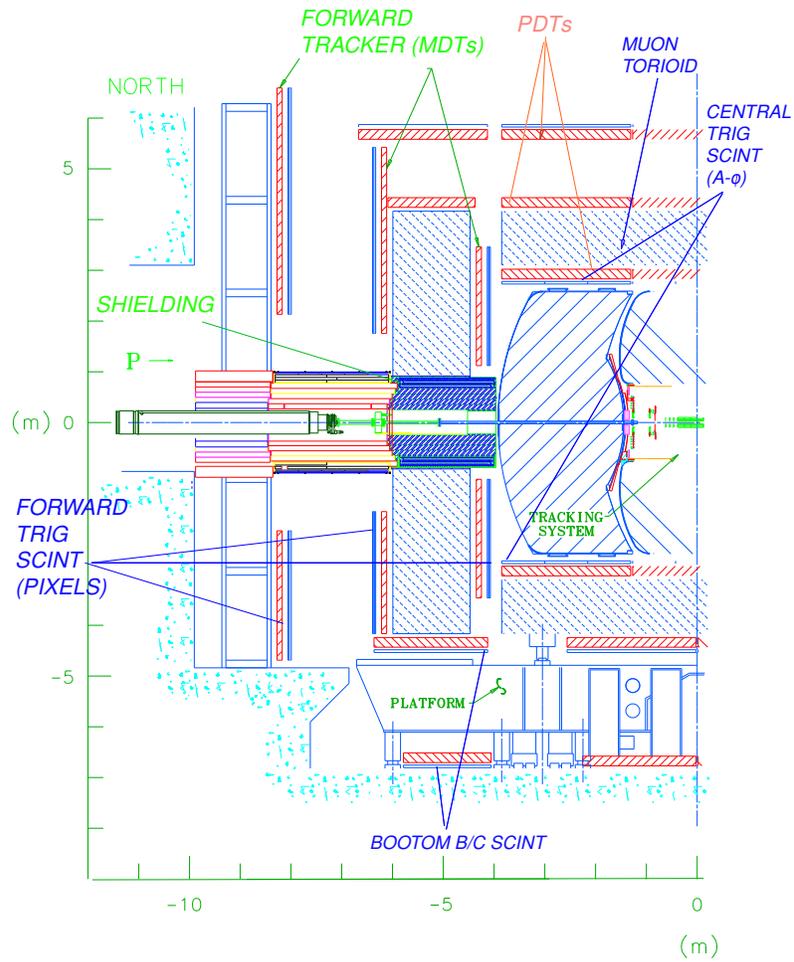
Figure 3.9: A one-half r-z view of the DØ Muon System. Components of both the Forward and Wide Angle systems are shown.

2.5 MHz Bunch Crossings

L1
4.2 μs

L1: Hardware Filter

1-3 kHz
128 bits

L2
100 μs

L2: Hardware
+ Software Filter

600-1000 Hz
128 bits

L3
100 ms

L3: Software Filter

40-80 Hz to Tape Storage
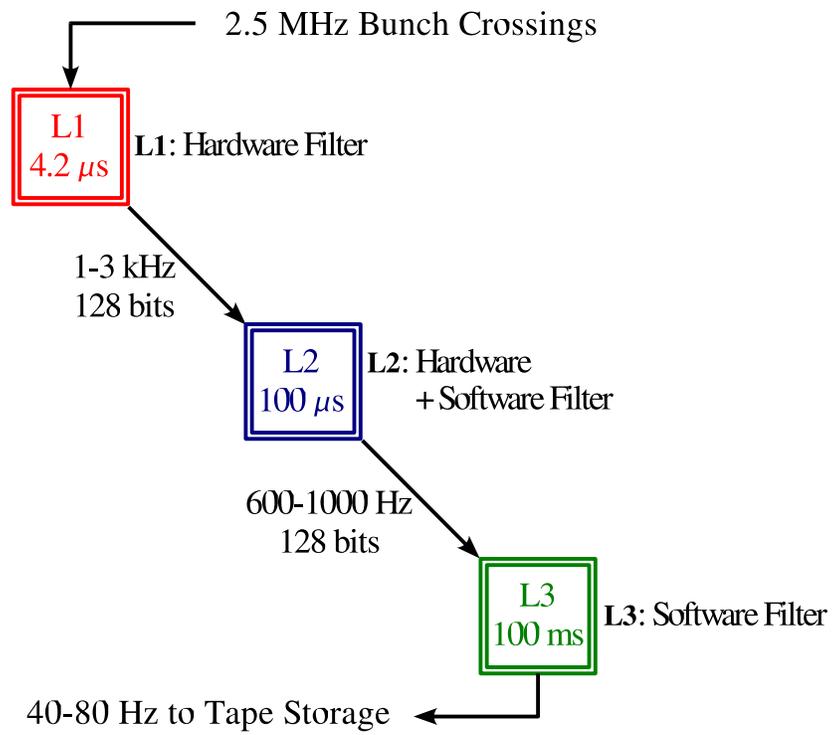
Figure 3.10: The DØ trigger layout and typical trigger rates.
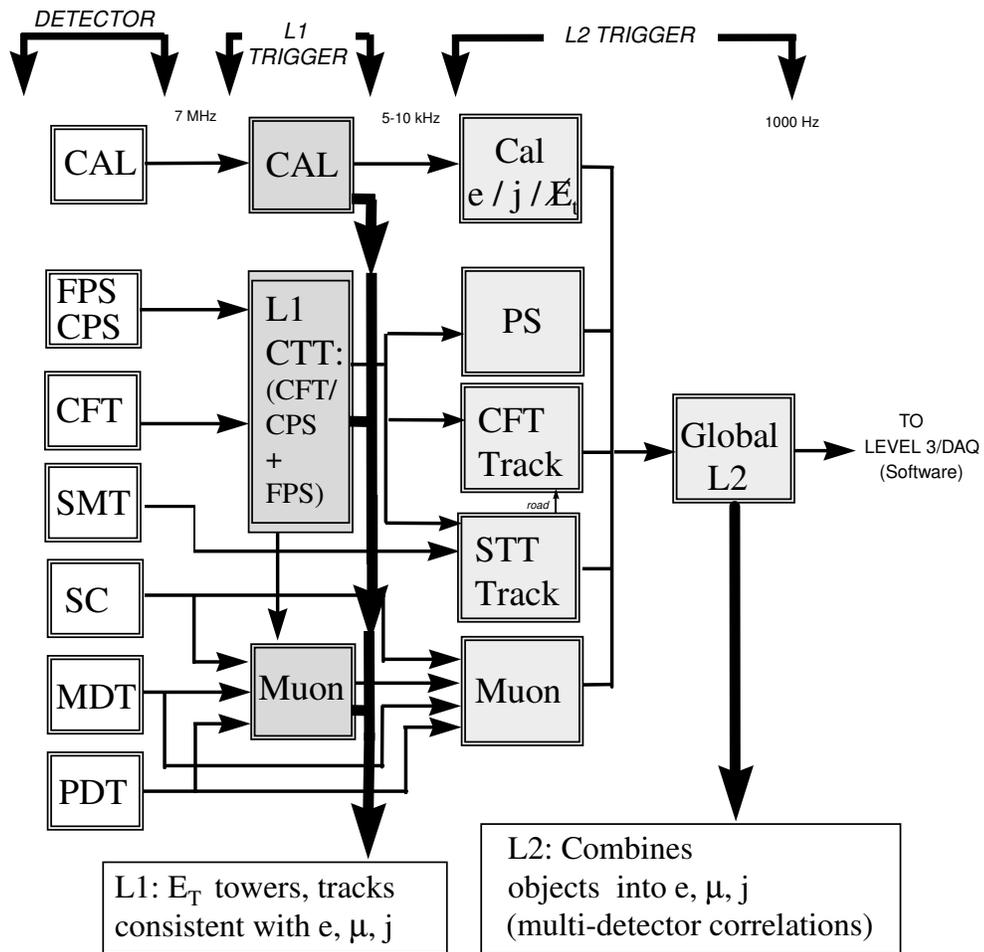
Figure 3.11: The Level 1 and Level 2 trigger data flow paths.
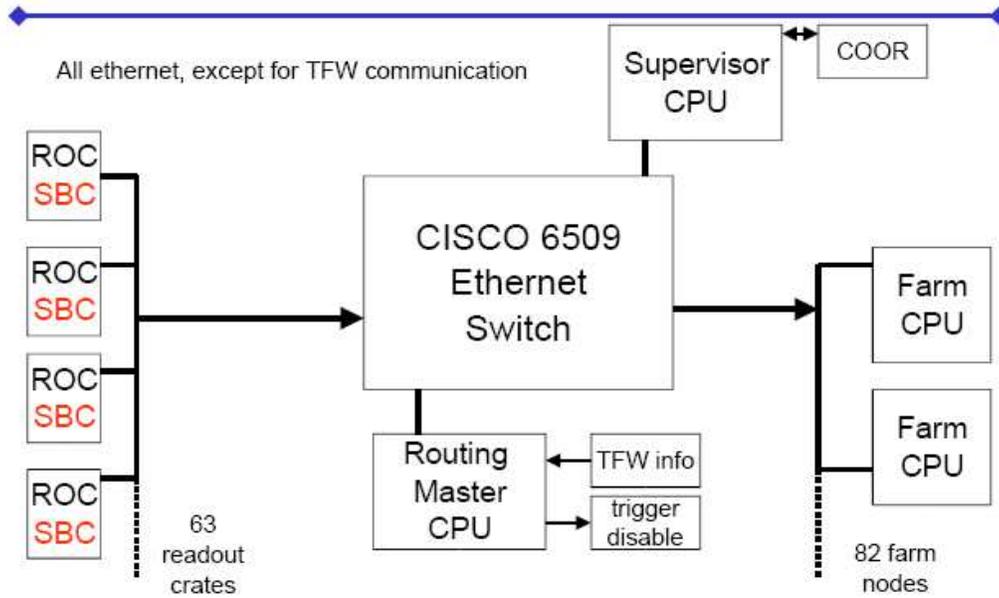
Figure 3.12: The L3 trigger and DAQ system architechture.

Chapter 4

Event Reconstruction

If an event satisfies the prerequisites for one of the physics triggers, the information measured by all the detector components is digitized and stored on disk to be analyzed in the future by researchers. During analysis, it is most convenient and intuitive for events to be recast in terms of the particles produced in the interaction. To this end, sophisticated algorithms are employed to identify quarks and leptons. Event reconstruction proceeds by using measurements made in the subdetectors to identify the particles produced in the hard-scatter interaction. The stable particles which reach the detector are electrons, muons, and hadrons. As mentioned previously, quarks experience a phenomenon called hadronization and manifest themselves in the detector as hadronic jets. The reconstruction of these physics objects from raw detector data can be divided into three primary stages

- *Hit finding*, wherein the digitized data is converted into "hits" defined by energy deposits in detector components (SMT silicon strips, CFT fibers, calorimeter cells, and muon scintillators and chambers). Each hit corresponds to an energy value, a spatial location, and their associated uncertainties.

- *Tracking and clustering*, where adjacent hits are combined into clusters consistent with the passage of a particle. Hits in the tracking detectors can be grouped to form trajectories referred to as tracks.

- *Particle identification*, during which tracks and clusters are combined to form candidate signatures of physics objects.

In this chapter, the reconstruction of physics objects from clusters will be briefly described.

4.1   Charged Tracks

Charged particles produced in $p\bar{p}$ scattering travel through the 2 T solenoidal field surrounding the tracking detectors. Lorentz forces cause the particles to follow a helical path from the interaction point. The solenoidal magnetic field lines are carefully mapped, including fringe effects near the edges of the tracking volume, and the paths of charged particles can be predicted. Using this predictive power, algorithms are applied to group clusters in the

SMT and CFT into *seed tracks*. The algorithms are iterated on these seeds to group more tracking clusters until the addition ceases to improve the track measurement. Once tracks are found, a fit is performed to determine the kinematic track parameters and their errors. A more complete description of the central track algorithms can be found in [19].

4.2   Primary Vertexing

The spatial points where $p\bar{p}$ collisions occur are characterized by the emergence of many charged particles. For each event, such a point is the origin of all the particles and is referred to as the primary vertex (PV). The reconstructed tracks in the event are used to find the location of the PV. For each track, the distance of closest approach (dca) to a common initial point (nominally the geometrical center of the detector) is calculated. All tracks with dca significance $dca/\sigma_{dca} < 5$, at least two SMT hits, and $p_T > 0.5$ GeV/c are used to create a vertex with the position determined via a fit. If the χ^2/ndof (where ndof is the number of degrees of freedom in the fit) is larger than 10, the track with the highest χ^2/ndof contribution to the vertex fit is removed. This process is iterated until the vertex fit χ^2/ndof falls below 10 or the number of tracks in the vertex is less than two. After this initial pass, the origin of the vertex is used as a seed in a second pass of the algorithm in which the track dca significance is required to be less than 3. After a vertex is found, the process is repeated using the remaining unvertexed tracks until no more vertices can be fit.

For the instantaneous luminosities at which the data for this analysis was recorded, an average of 0.5 additional inelastic $p\bar{p}$ collisions are expected per event in addition to the hard-scatter interaction[1]. The tracks from inelastic collisions, or *minimum bias* interactions, have smaller transverse momenta than tracks from hard-scatter interactions. The $\log_{10}(p_T)$ distribution of tracks in primary vertices is used to define a probability for a track to come from a minimum bias vertex. For each vertex, a likelihood is calculated with the minimum bias probabilities. The vertex with the lowest minimum bias likelihood is chosen as the event's PV. This vertex serves as the center of the physics coordinates for the event.

4.3   Secondary Vertexing

Particles with a finite lifetime tend to travel away from their production vertex before decaying. The distance they travel depends on their momentum, and can easily reach lengths which are measurable by the central tracking detectors. A good example of this phenomena is the production and decay of B mesons. When a bare b quark is produced, the strong force will cause it to immediately hadronize, often forming a B meson. The lifetime of these mesons is well-measured at about 1.5 ps. Thus, a B meson with an energy of 30 GeV will travel a distance of roughly 3 mm in the laboratory frame[2]. The decay of such a meson generally results in several charged particles, each creating a track originating at the point

---

[1]These additional collisions arise from multiple interactions in the bunch crossing. The average refers to the mean of a Poisson distribution.

[2]B mesons have a mass of roughly 5 GeV. A B meson with an energy of 30 GeV has a Lorentz boost $\gamma \simeq 5.7$, giving it a dilated lifetime of $\simeq 8.5$ ps ($c\tau \simeq 2.8$ mm).

of decay. Thus, the reconstruction of vertices which are spatially separated from the PV (*secondary vertices*) is a means of identifying long-lived particles. Due to the particular interest in b quark physics and the comparatively long lifetimes of B mesons, this process one of several algorithms commonly referred to as b-quark tagging. Secondary vertices are reconstructed after the event's PV is identified. Several combinations of algorithm parameters are defined to provide different efficiency/purity operating points, but we will only describe the TIGHT secondary vertex selection. The LOOSE and MEDIUM versions have identical algorithms, with less stringent reconstruction parameters.

The secondary vertex (SV) algorithm begins by creating a list of tracks not included in the PV. Using this list, several algorithm steps are undertaken:

1. The list of tracks is searched for two-track vertices compatible with originating from light-quark hadrons (for example K_s^0 or Λ), which can be identified via the characteristic angles between the two tracks and their known masses. Any tracks identified in this initial search are removed from the list.

2. Next, *proto-vertices* are formed by clustering tracks in this list into 3-dimensional cones[3] defined by an angular ring of $\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.5$. The tracks clustered are required to have $dca < 0.15$ cm, $zdca < 0.40$ cm, at least two SMT hits, and $p_T > 0.5$ GeV/c, where the $zdca$ is the distance of closest approach along the beam direction.

3. The proto-vertices are then searched for two-track vertex *seeds* constructed with tracks with a dca significance greater than 3.5, $p_T > 1.0$ GeV/c, and the χ^2 of the track's fit (from hits in the tracking detectors) less than 3.0. These seed tracks are fit to a vertex with the requirement that the fit $\chi^2 \leq 100$. Additional tracks from the proto-vertex are attached to the seed vertex if the χ^2 contribution to the fit is less than 15, thus forming preliminary SV's.

4. The final list of SV's for each event is selected using the following criteria

   - The transverse decay length, $L_{xy} = \left|\vec{L}_{xy}\right|$, of the SV must be less than 2.6 cm. $\vec{L}_{xy}$ is the two-dimensional vector from the PV to the origin of the SV, projected into the plane transverse to the beam direction.

   - The collinearity of the SV must be greater than 0.9. The collinearity is defined as the projection of the SV transverse momentum, $\vec{p}_{xy}$, along $\vec{L}_{xy}$, normalized by their product. The value $\vec{p}_{xy}$ is calculated as the transverse momentum sum over all SV tracks.

   - The decay length significance, $L_{xy}/\sigma_{L_{xy}}$, of the SV is required to be greater than 7. The error value $\sigma_{L_{xy}}$ is evaluated from the vertex fit by propagating the errors associated with each SV track.

---

[3]A requirement of constant angular size over a range of radii can be used to define the three-dimensional geometry of a cone.

4.4 Electron and Photon Identification

Electrons and photons can be initially grouped together in the identification (ID) process. As described in Section 3.2.4, electrons and photons will create electromagnetic (EM) showers in the calorimeters via bremsstrahlung and photon conversion. Thus, EM identification begins by searching the calorimeter layers for clusters of adjacent cells with positive energy deposition. After initial clustering, the EM object is defined with the cells in a circle of radius $\mathcal{R} = 0.4$ around the energy-weighted centroid of the cluster. To be accepted as an electron or photon candidate, the shape and energy of the EM object must pass additional quality criteria defined by EM-ID variables. For reference, the full EM-ID process is documented in [41].

First, a crude requirement is imposed requiring that the fraction of the EM object's energy that is deposited in the EM section of the calorimeter (layers 1-4) should be at least 90% as the transverse extent of the EM layers of the calorimeter was designed to fully contain high energy EM showers. Next, the isolation of the EM object's energy deposition is calculated as the fraction of energy in an annular ring around the centroid of the cluster $0.2 < \mathcal{R} < 0.4$ to the energy within $\mathcal{R} = 0.2$ of the centroid: $Iso = \frac{E_{\mathcal{R}<0.4} - E_{\mathcal{R}<0.2}}{E_{\mathcal{R}<0.2}}$. Small values of isolation are characteristic of properly reconstructed electron showers (rather than jets or $\pi^o \to \gamma + \gamma$ decays), as the calorimeter was designed such that $\mathcal{R} = 0.1$ towers should contain a circle of the EM particle's Moliere radius[4]. The Iso variable is required to be less than 0.15. Finally, a multi-variable tool (referred to as the H-matrix) is used to further discriminate between characteristic EM shower shapes and hadronic showers. The H-matrix is calculated as the inverse of the covariance matrix calculated with the following variables:

- The energies deposited in the four EM layers of the calorimeter.

- The energies deposited in the preshower and fine-hadronic calorimeter layer.

- The energy-weighted widths of the shower at the EM3 layer in r-ϕ and z.

The H-matrix is trained against simulations of electron showers and then used as an error matrix for a χ^2-like calculation[5]. This method provides a quantitative measurement of the probability that a calorimeter cluster arises from a single electron or photon. A cut is placed on the H-matrix χ^2 at 15-25 depending on the purity desired.

Finally, electrons and photons can be distinguished by requiring a track in the central tracking detectors that projects to the position of the shower in the calorimeter. This ensures that the shower is initiated by a charged particle and not a photon or other neutral particle.

4.5 Muon Identification

Muons are reconstructed from hits in the muon system wire chambers and scintillators. Two broad classes of muons can be defined by the number of subdetectors which are used to

---

[4]The Moliere radius is a parameter defined by multiple scattering theory as the radius containing about 90% of the shower's energy

[5]The H-matrix variables are not, in general, Gaussian and thus such a calculation will not follow a standard χ^2 distribution exactly.

identify the muon. Muons which are identified using only the information from the muon system are called *local* muons. Muons which are also matched to a track from the central tracking detectors, greatly increasing momentum precision and accuracy, are referred to as *central* muons. A more complete discussion of muon reconstruction can be found in [20].

For the purposes of muon reconstruction, the three physical muon system layers are divided into two categories: the A-layer (before the toroid) and the B- and C-layers (outside the toroid). Signals from the wire chambers and scintillators in each category are combined into *segments*. Segments from different categories are joined in a fit, yielding a measurement of the muon's path (position and direction) and a measurement of the muon's momentum as its path bends in the toroidal field. A minimum of two nearby wire chamber hits is required to make a segment, which is the result of a linear fit to these wire chamber hits. Matching scintillator hits can be added to the segment, provided they overlap in η with the segment. Detector inefficiency motivates the categorization of these segments, and this categorization is summarized in the *nseg* variable. The absolute value $|nseg| = 1$, 2, or 3 respectively indicates that the local muon was detected with A-layer hits only (inside the toroid), BC-layer hits only (outside the toroid), or both A- and BC-layer hits. Local muons are given a negative *nseg* value, while central muons are given positive *nseg* values. Reconstructed muons are given a fit quality grade to describe the reliability to the muon fit

- **Tight Muons** have at least two wire-chamber hits in the A-layer (which is four chambers deep) with a matching A-layer scintillator hit. In the BC-layers, a total of three wire chamber-hits are required (the B- and C-layers each have three chambers), again with a matching scintillator hit. Finally, tight muons require a converged fit for a central track match.

- **Medium Muons** are similar to tight muons except for requiring only two wire chamber hits in the BC-layers. Also, medium muons do not explicitly require a central track match, although a match can be required separately from the definition.

- **Loose Muons** have at least one scintillator hit in either the A- or BC-layers. They must also have at least two wire chamber hits in one of the layers.

The signature of a muon can be mimicked by a charged pion which does not deposit all of its energy in the calorimeter, and thus reaches the muon system. However, such pions are generally produced in the hadronization of quarks from the hard-scatter interaction, and characteristically will be surrounded by additional energy deposition. Two isolation variables are constructed to discriminate against such *fake* muons

- **Halo Isolation:** The transverse energy deposited in the calorimeter in an annular ring (or halo), defined by $0.1 < \mathcal{R} < 0.4$ around the axis of the muon, is required to be smaller than 2.5 GeV.

- **Track Isolation:** The sum of the p_T of all tracks in a cone of $\mathcal{R} = 0.5$ around the muon direction is required to be less than 2.5 GeV/c.

Finally, cosmic-ray muons can penetrate the detector and leave a signature in one or more parts of the muon system. The muon scintillation detectors allow for a precise time

measurement of interacting muons. The time measurement for the scintillators is calibrated such that muons from $p\bar{p}$ collisions arrive at $t = 0$. Muons scintillator times far from this value are rejected. The angular incidence of cosmic-ray muons is roughly uniform and most will not appear to originate from the event's primary vertex (PV). Constraining the distance of closest approach (dca) for the path of the muon to the PV provides additional discrimination.

4.6 Jet Identification

As mentioned in sections 3.2.4 and 3.3.3, quarks originating in $p\bar{p}$ interactions will hadronize and subsequently deposit their energies in the calorimeter. Jets are reconstructed from this energy deposition, allowing for both a position (η, ϕ) and energy measurement of the initial quark. In this analysis, we use jets reconstructed with a simple cone algorithm. These cones are defined as being confined to an angular circle of the specified constant size, and are constructed using the calorimeter readout-towers. Each tower points radially toward the geometrical center of the detector, but its transverse energy (E_T) is calculated with respect to the PV, using the physics coordinate η'. The jet reconstruction procedure can be found in [10].

Jet reconstruction begins by seeding the algorithm with calorimeter towers with $E_T > 0.5$ GeV and a total $E_T > 1.0$ GeV within $\mathcal{R} = 0.3$. The cone is then defined with the desired radius (both $R = 0.5$ and $R = 0.7$ for this analysis) and the E_T-weighted centroid is calculated. The cone is then recentered at this centroid and the energies are recalculated. This process is iterated until the cone axis becomes stable, and the resulting cone is saved as a *proto-jet*. Next, the midpoints between proto-jet pairs are used as seeds for the cone algorithm. The list of proto-jets is then pruned by removing entries with adjacent axes and entries with $E_T < 8$ GeV. Next, the *split-merge* process begins by matching proto-jets with overlapping boundaries[6]. If the E_T contained in the overlapping region is greater than half the E_T of either proto-jet, they are merged and the cone is recalculated. Otherwise, they are split, with the towers being reassigned to the spatially closest cone axis, and the energies and axes of each modified cone are recalculated.

After the preliminary reconstruction, quality criteria are applied to remove jets due to calorimeter noise and EM showers. These criteria are applied to all jets, regardless of cone size.

- The fraction of the jet's energy deposited in the EM calorimeter layers is required to lie in the range $0.05 <$ EM Fraction < 0.95. Jets interact primarily via the strong interaction and tend to deposit energy uniformly in the calorimeter, with roughly one third of the jet's energy deposited in the EM layers. Very small values are indicative of calorimeter noise, while values above 0.95 are consistent with the signature of an electron or photon.

- The fraction of the jet's energy deposited in the CH calorimeter layers is required to be less than 0.4. The CH layers tend to exhibit more noise than the rest of the calorimeter

---

[6]The algorithm allows for two proto-jets to share calorimeter towers, which must be corrected to allow for proper event reconstruction.

and, as jets tend to deposit little of their energies in the CH layers, this criterion helps to reject calorimeter noise.

- The ratio of the energies of the first and second-most energetic calorimeter cells clustered within the jet cone is required to be less than 10. A ratio of 10 or above is a good indication that this jet is clustered around a *hot cell* caused by electronic noise.

- The number of calorimeter towers that together contain 90% of the jet's energy is required to be greater than one. A value of one indicates a jet reconstructed from an EM object or a hot cell.

- The transverse energy recorded by the L1 calorimeter trigger for the jet is required to be greater than 40% of the reconstructed jet's transverse energy in the EM and FH calorimeter layers (excluding the CH layers). When the jet is located in the ICD region, this cut is lowered to 20%. This criterion further discriminates against CH noise.

The jets which have been clustered with the specified cone must have a series of scale corrections applied to reflect the original parton's energy, referred to as jet energy scale (JES) corrections. These corrections include

- A correction for average energy deposition in the calorimeter due to inelastic collisions other than the hard-scatter collision is made. This offset energy, which varies as a function of η, must be subtracted from the jet.

- A correction for the non-linear calorimeter response to energy deposition is made. Non-uniformities in the geometry of the calorimeter lead to variations in calorimeter response as a function of both η and jet energy.

- Corrections for jet energy which is not clustered in the jet cone are made. These corrections vary with jet cone size.

- The presence of a muon in the jet cone can indicate a semi-leptonic decay of a heavy-flavor quark via a virtual W boson. As muons deposit very little of their energies in the calorimeter, the jet energy must be recalculated to account for the muon and its associated neutrino.

Only jets which have a transverse energy $E_T > 15$ GeV and $|\eta| < 2.5$ have JES corrections applied. Jets below these thresholds do not deposit enough transverse energy to allow for a reliable correction and are generally not considered in analyses.

4.7 Missing Transverse Energy

The protons and antiprotons colliding at the Tevatron are prepared to have equal and opposite momenta, indicating that the total vector momentum sum in any event should be zero. However, as discussed previously, the hard-scatter interactions occur between the partons of the parent proton and antiproton. These partons share the total momentum of the parent

and are not constrained to any particular momentum value. However, the partons tend to carry very little momentum in the plane transverse to the beam path. Thus, to a good approximation, transverse momentum can be considered a conserved quantity.

As neutrinos interact very weakly with matter, their presence is inferred from any imbalance of transverse momentum in the event. This "missing" momentum is defined by the net imbalance in the transverse energy deposited in the calorimeter. The transverse missing momentum is assumed to point opposite to this vector, and is referred to as the missing transverse energy, or $\not{E}_T$[7]. The transverse vector sum of all calorimeter cells with positive energy, except those in the CH layers, are used to define the $\not{E}_T$. The CH cells are included only if they are clustered within a reconstructed jet.

Several corrections must be applied to this preliminary $\not{E}_T$. Foremost are those due to JES corrections. Because the measured jet energies are re-calibrated to reflect their true energies in the detector, this change in energy impacts the $\not{E}_T$. The JES corrections must therefore be taken into account in the calculation of $\not{E}_T$. We use cone jets with $R = 0.7$ for this correction in order to minimize the uncertainty on jet energy and thereby on the JES corrections. Only jets that pass the quality requirements described in Section 4.6 are used for this correction. In addition, muons, which usually behave as minimum-ionizing particles, deposit very little of their energy in the calorimeter and appear as missing energy. The momenta of identified muons can be subtracted vectorially from the $\not{E}_T$ to correct for this effect.

4.8 Luminosity

Although not a physics object, the instantaneous luminosity is a quantity that must be reconstructed by the detector. At each bunch crossing, an opportunity for an inelastic $p\bar{p}$ collision occurs. On the occasions that a collision occurs with a large enough momentum transfer to trigger the detector, the deflected hadron remnants are detected in coincidence by the scintillation counters of the luminosity system. The small timing resolution (~ 0.2 ns) allows for the z position of the collision to be determined to within about 6 cm, providing rejection for collisions between beam halo particles.

The instantaneous luminosity for each proton-antiproton bunch pair is measured using the fraction of bunch crossings with collisions. The probability of observing at least one collision (multiple interactions are not uncommon) is given by

$$P(n > 0) = 1 - e^{-\mu} \tag{4.1}$$

where μ is the average number of inelastic collisions per bunch crossing

$$\mu = \mathcal{L}\sigma_{eff}/f \tag{4.2}$$

where $\mathcal{L}$ is the instantaneous luminosity, σ_{eff} is the effective $p\bar{p}$ inelastic cross section, and f is the frequency of the bunch crossings. The probability $P(n > 0)$ can be rewritten as

---

[7]The calorimeter measures energies, not momenta. The momentum relation for calorimeter energy depositions is calculated based on the location of the event's primary vertex.

$\mathcal{R}_{lumi}/f$, where $\mathcal{R}_{lumi}$ is the rate of inelastic collisions measured by the luminosity system. Thus, the luminosity can be written as

$$\mathcal{L} = -\left(f/\sigma_{eff}\right)\ln\left(1 - \mathcal{R}_{lumi}/f\right) \qquad (4.3)$$

The bunch crossing frequency f is well measured at 7.58 MHz. The effective inelastic cross section (or luminosity constant) is the dominant source of uncertainty. Its value is approximately 43.1 ± 1.9 mb, as measured by experiments at CERN and in Run I of the Tevatron [21]. Approximately every four minutes, a *luminosity block* is written to a database, recording the average inelastic collisions rate from the luminosity system for that time period, the status of the data acquisition, and the data quality as determined from the state of the detector hardware during that time.

Chapter 5

Simulation

After collecting and reconstructing data events, the task remains for the physicist to understand the patterns observed in the data. These patterns can reflect many different phenomena such as production of SM particles like W bosons or quarks, events present in the data due to electronic noise in the detector mimicking a physics process, or possibly the signature of a particle or phenomena previously unobserved. This confusing overlap of patterns must be categorized in order to separate events of interest from the bulk of data. One tool used to aid in this categorization is the artificial simulation of events, commonly referred to as Monte Carlo simulation. This simulation allows the prediction of the ideal behavior of the detector, providing a model to which data can be compared to determine detector and reconstruction performance. It also provides a prediction of the detector's response to specific physical processes of interest, allowing detailed measurements of physical distributions and selection efficiencies.

The Monte Carlo method itself refers to the calculation of approximate solutions to a variety of mathematical problems by performing statistical sampling experiments. The method is named after the city in the Monaco principality, because of the city's role in the popularization of the roulette wheel, the invention of which is often credited to Blaise Pascal, a 17th century French mathematician. The roulette wheel represents a simple random number generator, and the heart of the Monte Carlo method lies in the random selection of outcomes based on predefined probabilities. However, the method can apply to problems with no probabilistic content as well as to those with inherent probabilistic structure.

In the context of particle physics, Monte Carlo (MC) represents the random generation of physical final states occuring after the collision of two (or more) initial-state particles. After generation of a final state, the particles in that state are propagated through space based on their initial momenta, ultimately interacting with the detector. This chapter will briefly discuss these two MC simulation steps. To simplify the breadth of the discussion, we shall constrain the topic to the relevant case of $p\bar{p}$ collisions.

5.1 Generation

The MC process begins with generation, or producing the description of one possible outcome of a $p\bar{p}$ collision. This process involves the random selection of an initial state and a final

state from the available kinematic phase space, and can be summarized in three steps

1. First, a physics process is chosen, constrained principally by the requirement that it must be a process that can be described physically by the SM. A further constraint is the choice of initial-state particles (generally two), which must be obtained from the colliding hadrons. With these constraints met, the matrix element for the chosen process must be calculated. For this example, we will choose the annihilation of a $q\bar{q}$ pair to a gluon propagator which subsequently decays to a $t\bar{t}$ pair, as shown in Fig. 5.1.
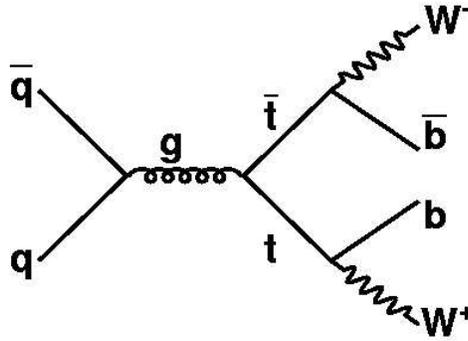
Figure 5.1: One possible Feynman diagram for $t\bar{t}$ production in $p\bar{p}$ collisions.

2. Second, the momenta of the initial-state quarks are determined by a simulation of the parton momentum distributions of the proton and anti-proton. These momenta are selected at random using a *parton distribution function* (PDF), which describes the motion of the quarks and gluons (partons) inside the hadrons. The PDF's are generally determined phenomenologically by a global analysis of a range of hard scattering processes, using a perturbative QCD framework.

3. Third, given the momenta of the initial-state quarks and the desired matrix element, the kinematic phase space for the decay of the gluon propagator can be determined. At random, one of the possible final states of the $t\bar{t}$ pair is selected.

The result is two sets of four-momenta and quantum numbers describing the $t\bar{t}$ pair, and is the final product of the generation process.

5.2 Propagation and Interaction

After obtaining a set of final-state particles and their momenta, the next step is to determine their behavior through the course of time. This step involves the determination of the trajectories of the particles as they pass into the detector, describing the decays of particles with finite lifetimes, predicting the process of energy loss as the particles interact with the detector, and gauging the realistic response of the detector to the energy depositions made by the particles.

This process necessarily starts with a detailed description of the detector geometry and materials. Such should include the specifics on how the detector materials[1] behave during the passage of charged particles and the details of any magnetic fields in the detector. The decay of short-lived particles is simulated, providing appropriate daughter particles based on the predictions of the SM. Most particles which will decay have already done so before exiting the beam pipe, resulting in primarily stable particles entering the detector. However, some particles enter the detector volume before decaying. At this point, the decays can be influenced by interactions with the detector materials, which must also be simulated.

The possibilities of initial- or final-state radiation can be included for all charged particles. This describes the emittance of photons for electomagnetically charged particles or gluons in the case of strongly charged particles, both impacting the momenta of the particles involved and introducing new particles to the final state. Futhermore, as quarks propagate, the strong force induces the phenomena of *hadronization* and *fragmentation*, which lead to the formation of hadronic showers or jets. These processes must all be simulated to obtain a proper description of particle behaviors.

Now all that remains is the propagation of the particles through the detector. The path of each particle is calculated based on magnetic fields encountered and the changes in trajectory caused by scattering from detector materials. The energy deposited in each part of the detector is converted to a signal by simulations of the detector's readout electronics. These signals are subsequently compiled into a full event as it would be written in the case of real data collection. At this point, the MC events are ready to be passed through the same reconstruction process as the data. Over time, this process is fine-tuned until the MC predictions can reproduce the gross features of the data, providing the desired predictive power.

---

[1]In the case of the DØ detector, this includes everything from the beryllium beam pipe and the silicon of the SMT to the last wires in the outer layer of the muon system, including any non-instrumented support structures.

Part II

Problem and Analysis

Foreword to Part II

After the introductory discussions in Part I, the second part of this document deals with the data analysis and the motivations behind it. We have examined the tools, both experimental and theoretical, used to predict and observe the interactions of elementary particles. However, such tools as they are, the level of their accuracy will ultimately decide their utility. A detailed test of the predictions made by the Standard Model (SM), in the form of simulated SM processes, is a necessary step in the study of elementary particles. In particular, there are a few specific physical signatures that are of prominent contemporary interest at the Fermilab Tevatron Collider.

The foremost of these signatures is the production of top quarks. In 1995, the two Fermilab Tevatron experiments (DØ and CDF), discovered the top quark, thus completing the third generation of matter particles as predicted by the SM. Following this triumph, great efforts have been made to measure attributes of the top quark, such as its mass and decay properties. In addition to this, the production of Higgs bosons is an area receiving much attention. Although predicted by the SM, such a particle has yet to be observed. Complicating the search, the SM does not explicitly predict the mass of a hypothetical Higgs boson, but merely provides a range of masses compatible with other SM observables.

The observable signature of these processes is that of b quarks in the final state. The large mass of the top quark ($m_{top} \simeq 178$ GeV/c^2) causes it to decay very quickly to a W boson and a b quark. The phenomenologically likely mass of a Higgs boson is of the order of $100 - 150$ GeV/c^2 and the dominant decay mode of a Higgs boson in this mass range is into two b quarks. Furthermore, the most observable production mechanism for a Higgs boson at the Tevatron is in association with W or Z bosons, providing a signature somewhat similar to the production of $t\bar{t}$ pairs. A large background process to both of these signals is the production of $b\bar{b}$ pairs via gluon splitting in association with a W boson. The focus of this thesis is therefore the careful measurement of inclusive heavy-flavor (t, b, c) quark production in association with W bosons. This analysis takes the form of a search for such production beyond the predictions of the SM. The second part of this thesis is devoted to the description of the search, the motivations driving it, and the analysis itself.

Chapter 6

Statement of Problem

As mentioned in the foreword, this analysis represents a search for anomalous production of heavy-flavor quarks in association with W bosons at the Fermilab Tevatron. This chapter discusses the motivations for a search of this type and the design of the analysis.

6.1 Motivations

The heavy-flavor (HF) quark (t, b, c) content of jets produced in association with a W boson provides a test of standard model (SM) predictions. Many physical measurements and searches at the Tevatron depend greatly upon the ability to predict HF quark production, such as measurements of top properties and searches for Higgs bosons. These issues alone motivate a careful measurement of HF production, but an additional issue was introduced recently when the CDF collaboraton announced an observation of an anomalously large rate of HF quark production during Run I of the Tevatron. Lacking the ability to efficiently perform HF-jet taging, the DØ collaboration was unable to test this observation.

A measurement of the HF quark production rate can be interpreted as a search for anomalous production, providing both a test of the SM and sensitivity to new physics. All of these processes are observed experimentally as a HF quark jets in association with a W boson, and we will thus focus on these signatures. I will attempt to more completely describe these motivations in the following.

6.1.1 Top Quark Production

Since the discovery of the top quark during Run I, Fermilab physicists have been attempting to make precision measurements of the top's properties. These measurements include the top quark mass m_{top}, the rate of top decays to HF quarks $\mathcal{R}_{t \to b}$, and the helicity of W bosons from the decays of top quarks $\mathcal{H}_{top}$. Each of these measurements is a probe of the SM predictions of the top quark properties and also of the SM itself. The study of electroweak (EW) interactions in the SM has reached a precision both theoretically and experimentally that quantum corrections become important. As the most massive SM particle, the top quark is the largest contributor to these corrections, and a precision measurement of the

top mass can lead to an indirect inference of the Higgs mass via radiative corrections[1]. The rate of top decays to HF quarks, $\mathcal{R}_{t \to b}$, is directly proportional to the square of the CKM matrix element V_{tb} [25]. This value is expected to be very near unity, and improved measurements can help to advance the constraints on other parameters of the CKM matrix, which is currently the focus of much effort in particle physics. The helicity of W bosons in top quark decays provides a third test of the SM. The top quark decays via a vector/axial-vector (V-A) charged current interaction. This parity-violating interaction limits decays of top quarks into W bosons with longitudinal and left-handed helicity states. A measurement of the left- and right-handed helicity of W bosons from top decays probes both the top mass and the SM description of EW decays [26].

The two dominant production modes for top quarks at the Tevatron are in $t\bar{t}$ pairs as illustrated in Fig. 6.1. The decays of the W bosons define the observable final states, with fully hadronic (six-jet) ($\sim 44\%$) and semi-leptonic (four-jet) ($\sim 44\%$) the most probable final states. However, the QCD multijet backgrounds to the six-jet signature are large and irreducible, leaving the leptonic W decays as the more promising signature. Thus, the leptonic decay of one of the W bosons along with the hadronic decay of the second W boson provides the dominantly observed final state.
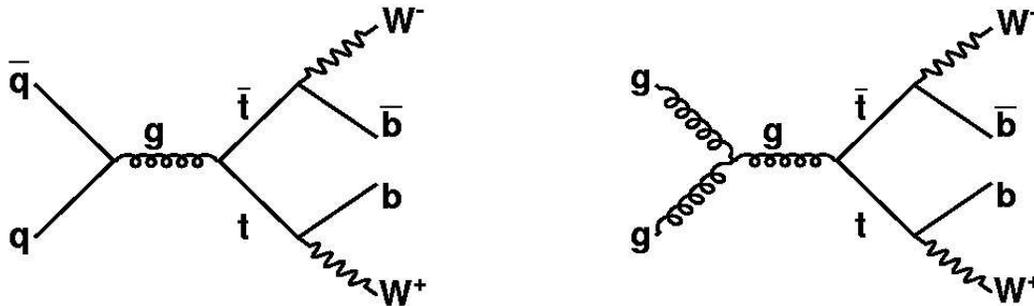
Figure 6.1: Feynman diagrams for $t\bar{t}$ production in $p\bar{p}$ collisions. The left diagram (quark-antiquark production) is dominant, but the right diagram (gluon fusion) contributes $\sim 5 - 10\%$ to the total cross section.

An additional contribution can come from single top quark production, as shown in Fig. 6.2. These two processes, referred to as the s-channel and t-channel processes, may be characterized by the momentum exchange of the participating W-boson: Q_W^2.

- t-channel W-exchange ($Q_W^2 < 0$): This process has the largest cross section of the single top processes. The t-channel process is commonly referred to as "qtb", and includes the $tq\bar{b}$, $\bar{t}\bar{q}b$, tq, and $\bar{t}\bar{q}$ final states.

- s-channel W-exchange ($Q_W^2 > 0$): The s-channel process is referred to as "tb", and includes both $t\bar{b}$ and $\bar{t}b$ final states. The s-channel process represents $\simeq 31\%$ of the single top cross section.

These signatures define two- and three-jet final states in association with a W boson.

---

[1]The current measurement of the top-quark mass from Run I of the Tevatron is 178.0 ± 4.3 GeV/c^2 [24]
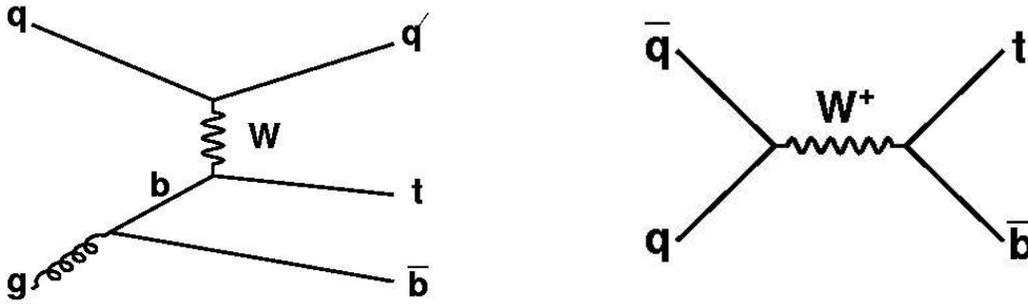
Figure 6.2: The dominant Feynman diagrams for t-channel (left) and s-channel (right) single top quark production.

6.1.2 Higgs Boson Production

Searches for a Higgs boson have continued for many years, the most comprehensive of which was performed at LEP[2]. Figure 6.3 displays the results of the combined LEP II search for a SM Higgs boson as a function of Higgs mass. The lower limit on the Higgs mass of 114.4 GeV/c^2 obtained from LEP can be combined with several other precision EW measurements (including the measurement of the top mass) to yield a constraint on the range of Higgs masses compatible with the formulation of the SM. A plot of this constraint in the form of a least χ^2 fit to global EW data is shown in Fig. 6.4. Thus, the indications leading into a Higgs search at the Tevatron are that a Higgs boson is most likely to have a mass near $115 - 130$ GeV/c^2, allowing for a more concerted search effort.

At the Tevatron, the dominant production mechanism is gluon fusion, as seen in Fig. 6.5. After gluon fusion, the largest production mode is associated production with a W or Z boson. Figure 6.6 outlines the decay branching fractions for a SM Higgs boson as a function of mass, showing $H \rightarrow b\bar{b}$ to be the dominant decay mode for masses below $\simeq 135$ GeV/c^2. In a similar fashion to the top quark signature, fully hadronic final states resulting from gluon fusion production are difficult to observe and searches generally focus on the leptonic decays of the associated W and Z bosons. This Higgs production and decay process defines a two-jet final state in association with a W or Z boson, as shown in Fig. 6.7.

The observation of a Higgs boson would occur in two ways. First, as the topology defines a two-jet final state, a Higgs signal would appear as an excess of events in the exclusive jet spectrum. However, this observation method is not very sensitive and is eclipsed by searches capitalizing on the dijet invariant mass of the Higgs boson. Although this search method is complicated without a predicted Higgs mass, it is ultimately the method which provides the most realistic prospects for Higgs discovery. However, without a very detailed understanding of the predicted SM backgrounds for this process, any observed signal can be mistakenly interpreted as an upward fluctuation.

---

[2]The Large Electron Positron (LEP) collider was a synchrotron at the CERN facility outside Geneva, Switzerland.
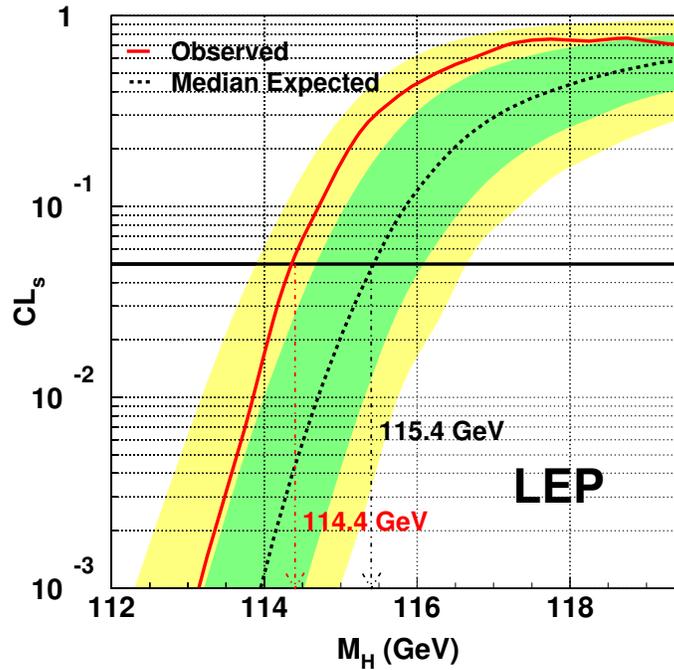
Figure 6.3: The 95% confidence level limits for the LEP II combined SM Higgs search.

6.1.3 CDF Anomalous Result

In 2002, the CDF collaboration reported evidence for anomalous production of HF jets in association with a W boson, based on events recorded during Run I of the Tevatron in which $p\bar{p}$ collisions occured at $\sqrt{s} = 1.8$ TeV [22], and subsequently reconfirmed in 2004 [23]. In this study of roughly 11,000 $W \to \ell\nu$ ($\ell = e$ or μ) inclusive decays[3] , two different HF-quark tagging algorithms were applied separately to identify jets arising from HF quarks. These two algorithms, secondary-vertex and soft-lepton tagging, resulted in selections which were roughly compatible with SM predictions. The CDF data observations and SM expectations for the secondary-vertex tagging and soft lepton tagging selections are shown in Figs. 6.8. In this figure, the exclusive number of jets for events with at least one jet tagged with the specified HF-tagging algorithm is plotted. The study then defined a *supertag* as a jet with both a secondary vertex and SLT tag. Jets tagged in this way are then referred to as *superjets*[4]. Upon selection of superjets, a marked deviation between data and SM expectations was observed, as seen in Fig. 6.9. The W boson plus superjet selection observed 13 with a SM expectation of 4.4 ± 0.6 events in the W plus 2- and 3-jet bins. The CDF collaboration reported a $<\sim 1\%$ probability that this excess is compatible with the predicted result.

This evidence for anomalous HF production can be interpreted in several ways. The

---

[3]The term *inclusive* refers to a selection without further discrimination on the remaining particles in the event.

[4]The prefix "super" is used to indicate high quality and is not meant as a reference to supersymmetry.
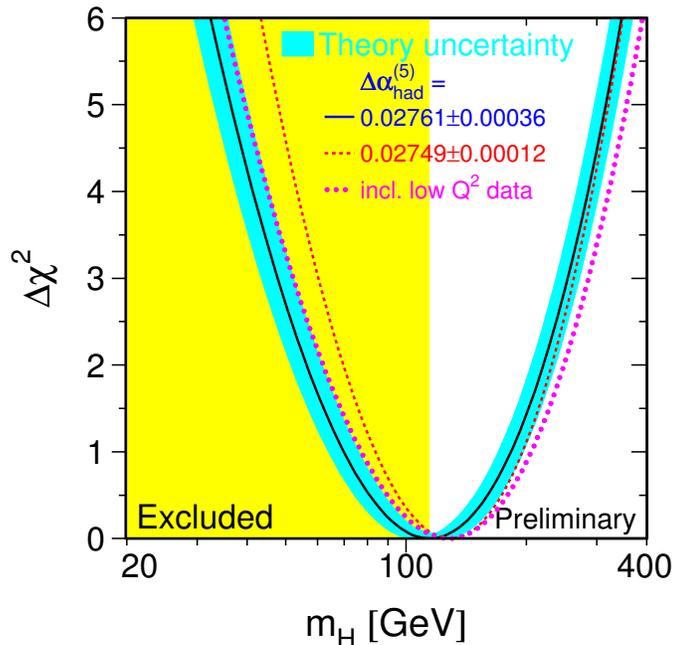
Figure 6.4: Least χ^2 fit for precision EW data constraining the mass of a SM Higgs boson.

possibilities of an upward fluctuation in data or a systematic error in the SM expectation cannot be dismissed in this study. However, one must also consider the possibility that a physical process not described by the SM is contributing to the observed number of events. In Run I of the Tevatron, the DØ experiment lacked a tracking detector with the resolution necessary to perform secondary-vertex tagging and an independent measurement of this phenomena was not made. However, the upgraded DØ detector makes this possible and a drives the motivation to attempt to reproduce the CDF result. Indeed, in addition to probing the accuracy of contemporary SM predictions, a search for anomalous HF jet production in association with W bosons provides the opportunity to investigate the CDF result.

6.1.4 Additional SM Backgrounds

In addition to the top and Higgs contributions to a W plus HF jet final state, there are several SM processes that are significant. These processes, although interesting alone, are considered a background to top measurements or Higgs searches. The dominant backgrounds are $Wb\bar{b}/c\bar{c}$, with the quark pair arising from gluon splitting, shown in Fig. 6.10, and WZ (with $Z \to b\bar{b}$) production, shown in Fig. 6.11, respectively. As seen in Fig. 6.10, the $Wb\bar{b}$ process can include additional jets arising from gluon emission (which can subsequently split to two quarks). In this way, the $Wb\bar{b}$ background can contribute to two or more jets final states. In the case of WZ production, the Z boson decays to $b\bar{b}$ in $\sim 15\%$ of decays, producing a two-jet final state. For both of these backgrounds, the analog processes of $Zb\bar{b}$ and ZZ can contribute. In this scenario, the Z boson decays to a dilepton final state ($Z \to \ell\ell$) and one of
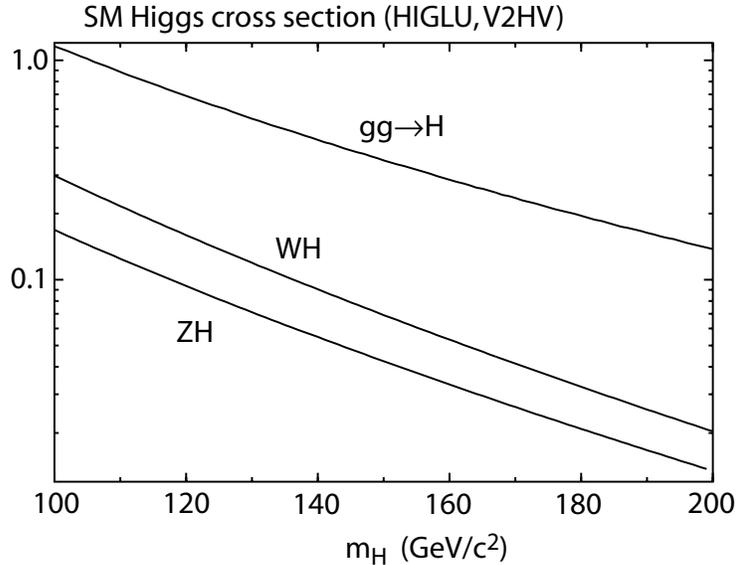
Figure 6.5: Production cross sections for a SM Higgs boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV as a function of hypothesized Higgs mass. These values were based on a calculation by T. Han and S. Willenbrock [28].

the leptons either escapes the detector or is lost due to detector/reconstruction inefficiencies. The missing lepton can mimic a neutrino signature of missing transverse energy.

HF quarks are also produced by QCD-only multijet interactions. These events can become background events if the signature of a W boson is imitated by a fluctuation in measurements made by the detector. For example, a quark or gluon can shower early in the calorimeter and create an energy deposition satisfying the electron reconstruction requirements. In addition, a hadron could contain pass through the calorimeter and be observed in the muon detectors. The behavior of a neutrino can also be "faked" by an upward or downward fluctuation in the measurement of the jets in the calorimeter, thus creating a net imbalance in transverse energy. If these two circumstances occur together with QCD $b\bar{b}$ production, the W-boson plus HF-jet topology is satisfied. Also as mentioned above, another class of backgrounds is defined in the event that a jet arising from a light-flavor quark (u, d, s) is misidentified as a HF quark. Although the probability for these two background types to pass into the selected sample is small, the rate at which they occur is very large and thus represents a significant background.

6.2 Analysis Design

The strategy for measuring the rate of HF jets in association with W bosons naturally begins with the selection of W bosons. In this analysis, only two of the three possible W decays are chosen: $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$. The $W \rightarrow \tau\nu$ decays are not a desirable topology as the τ itself decays dominantly to hadrons. The leptonic τ decays also do not represent an appreciable efficiency for selection as the energy of the resulting electron or muon is generally too small to be selected. The W-boson candidate selection proceeds by identifying
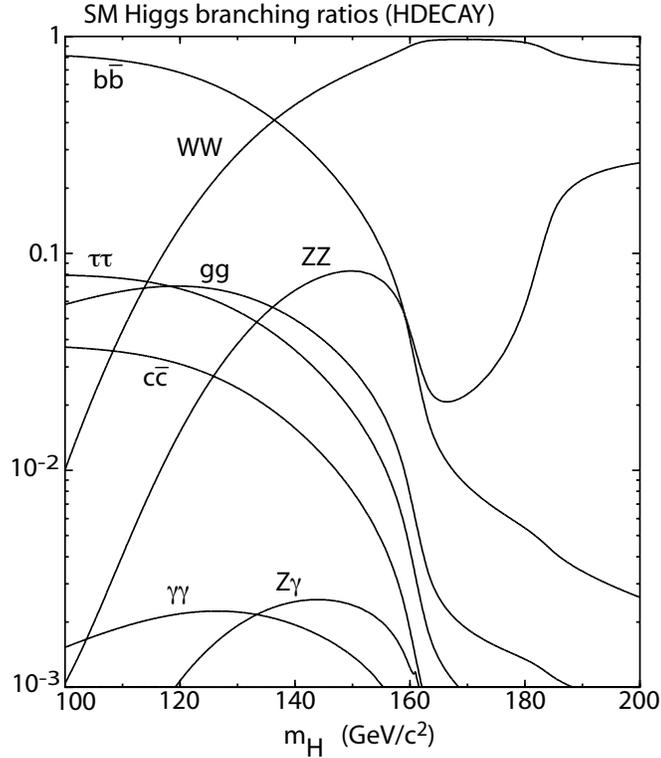
Figure 6.6: Branching fractions for a SM Higgs boson. These values were calculated using the HDECAY program [29].

an energetic electron or muon within the triggering acceptances of the detector, and requiring the event to exhibit sufficient missing transverse energy to be consistent with the neutrino from a W decay. The second step in the analysis is to evaluate the distribution of jets in the selected events. Jets are selected in regions of the calorimeter which are well-understood and which allow for proper energy scale corrections. The distribution can be categorized by the number of jets in the event, which is then referred to as the exclusive jet spectrum. After measuring this spectrum, the heavy-flavor quark content is evaluated via HF-quark tagging, or simply HF-tagging. This final sample of HF quarks in association with W bosons allows for a comparison of the SM predictions to the data observation. The following chapters explain the steps in the process of selecting and analyzing this desired sample of events.

To address the anomalous result from the CDF collaboration, a further selection is made by requiring the simultaneous overlap of two different tagging algorithms for the same jet. Jets selected by more than one algorithm have a larger probability for originating from a HF quark, and thus represent a more pure sample. With this selection, limits on the anomalous production of HF jets in association with W bosons is extracted.
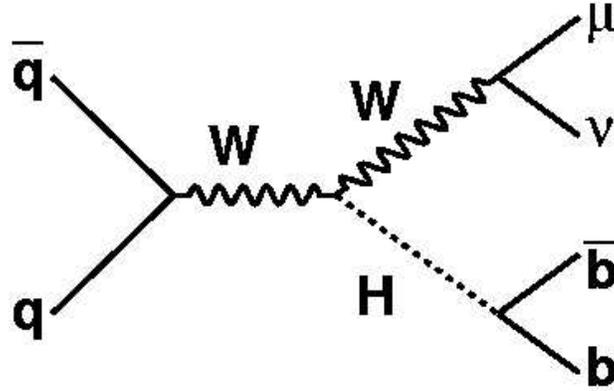
Figure 6.7: Feynman diagram of Higgs production in association with a W boson, with a $WH \to \mu\nu b\bar{b}$ final state.
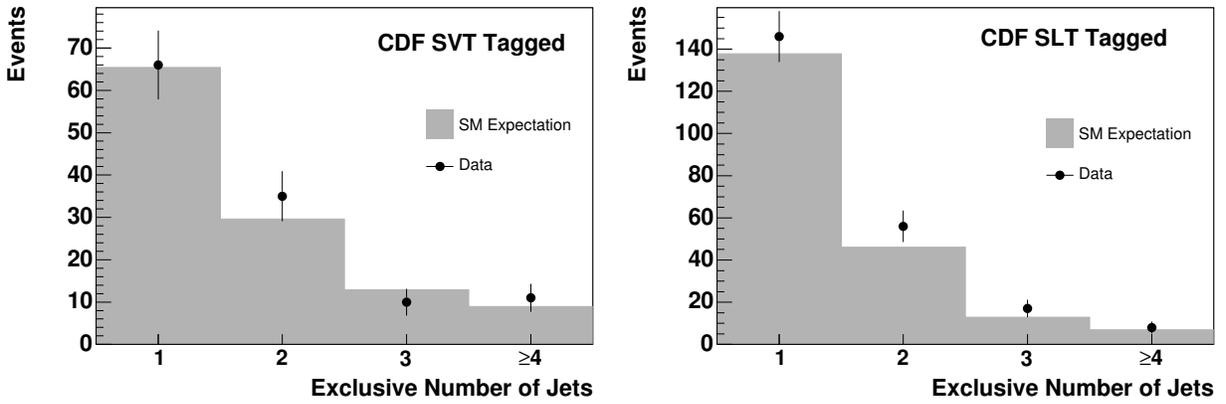
Figure 6.8: The exclusive number of jets in events with a selected W boson. One of the jets is required to have a SVX tag (left) or a SLT tag (right).
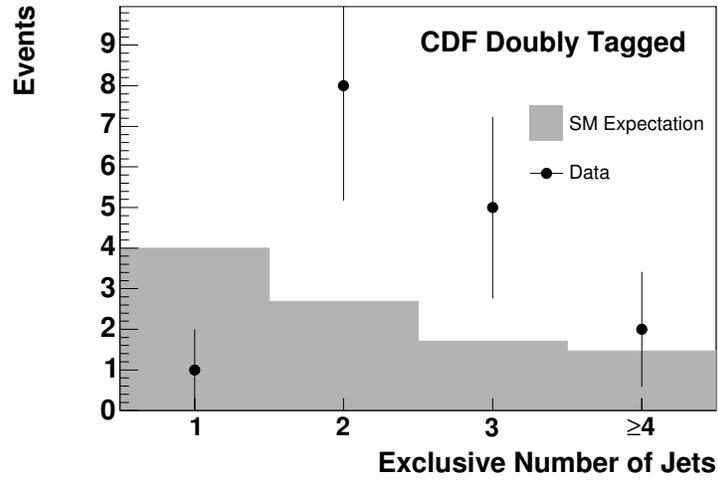
Figure 6.9: The exclusive number of jets in events with a selected W boson. One of the jets is required to have both a SVX tag and a SLT tag.
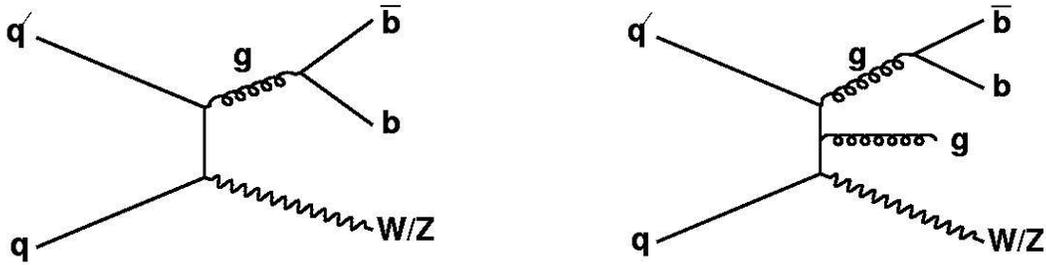
Figure 6.10: Primary Feynman diagrams contributing to the $W b\bar{b}$ (left) and $W b\bar{b}$+jets (right) final states.
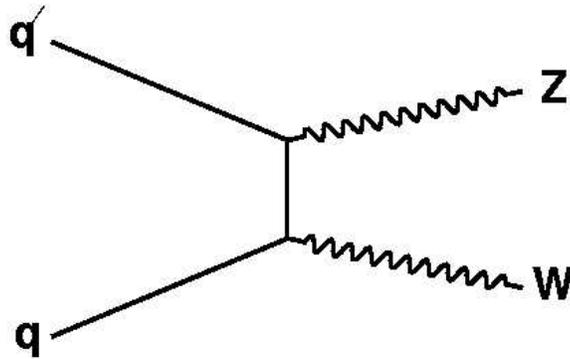
Figure 6.11: Dominant Feynman diagram for $q\bar{q} \rightarrow WZ$ production at the Tevatron.

Chapter 7

Data and Simulation

In addition to the data collected for this analysis, simulated samples of SM processes were generated in the form of Monte Carlo events. This chapter outlines the data sample used and the details of the simulation.

7.1 Data Sample

The DØ detector records events at an average rate of 50 Hz. As discussed in Sec. 3.3.7, these events are a small subset of the $p\bar{p}$ collisions initiated by the Tevatron, filtered down to match the maximum physical recording rate available using a number of three-tiered triggers. This final *rate to tape* consists of events triggered by a large menu of triggers representing the current physics priorities of the experiment. The first two trigger levels can accommodate 128 trigger bits (each trigger represents one bit). The third level concatenates these bits and refines the output to make a final trigger decision. This decision is made by comparing to a predefined list of *global triggers*, which consist of a requirement for each of the three physical trigger levels. Futhermore, the data collected for this analysis were recorded over a period of time which spans several global trigger lists (or versions), and we must therefore include triggers from both lists. This analysis relies on global triggers which select single, isolated electrons and muons consistent with the decays of W bosons. In this chapter, we will briefly discuss the triggers used and their corresponding efficiencies and luminosities.

In the selection of $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ decays, we require a single-electron or a single-muon trigger to have fired for each event. The full data sample is reduced to two smaller sub-samples, referred to in the following as the 1EMLOOSE and 1MULOOSE samples, which are selected as follows:

- **1EMloose:** One offline reconstructed electromagnetic shower with $p_T^e > 15$ GeV/c.

- **1MUloose:** One offline reconstructed muon with $p_T^\mu > 8$ GeV/c.

7.1.1 Single Electron Triggers

To increase trigger efficiency, we require $W \rightarrow e\nu$ candidate events in the 1EMLOOSE sample to have passed one of several similar triggers: the logical OR of triggers EM_HI,

EM_HI_SH, EM_MX, and EM_MX_SH for trigger versions below v12. For trigger version v12 and above, we include the logical OR of triggers E1_SHT20, E2_SHT20, E3_SHT20, E1_SH30, E2_SH30, and E3_SH30. The trigger requirements for the dominant triggers in these groups are as follows:

- **EM_MX_SH**

  - Level 1: CEM(1,15): At least one L1 calorimeter trigger tower with at least 15 GeV of transverse energy in the EM layers.
  - Level 2: No L2 trigger requirement is made for this trigger.
  - Level 3: L3Ele(1,20,sh): A single electron candidate with transverse energy above 20 GeV and within $|\eta_{em}| < 3.0$. This trigger term also places shower shape requirements on the reconstructed electron candidate.

- **E1_SHT20**

  - Level 1: CEM(1,11): At least one L1 calorimeter trigger tower with at least 11 GeV of transverse energy in the EM layers.
  - Level 2: No L2 trigger requirement is made for this trigger.
  - Level 3: L3Ele(1,20,sh): A single electron candidate with transverse energy above 20 GeV and within $|\eta_{em}| < 3.0$. This trigger term also places shower shape requirements on the reconstructed electron candidate.

The trigger efficiencies are obtained from Ref. [31]. The p_T-weighted trigger efficiency for $W \rightarrow e\nu$ decays is found to be $\epsilon_{trig} = 96.7 \pm 0.3\%$ for trigger versions below v12 and $\epsilon_{trig} = 97.7 \pm 0.4\%$ for trigger version v12. This yields a luminosity-weighted average trigger efficiency of $\epsilon_{trig} = 97.0 \pm 1.6\%$.

7.1.2 Single Muon Triggers

We use the following triggers for the 1MULOOSE data sample: MUW_W_L2M3_TRK10, MUW_A_L2M3_TRK10, MT10W_L2M5_TRK10, and MU_W_L2M5_TRK10. Because certain triggers were prescaled for different run periods, we take care to rely only on unprescaled triggers. Because of this, we used MU_W_L2M5_TRK10 for runs 162512-179555, and the other three triggers for all other runs. The trigger requirements for the dominant triggers are as follows:

- **MUW_W_L2M3_TRK10**

  - Level 1: mu1ptxwtlx: A wide region tight scintillator trigger with loose requirements on hits in the muon wire chambers.
  - Level 2: One medium muon, identified at L2, with $p_T > 3$ GeV/c.
  - Level 3: One track, reconstructed at L3, with $p_T > 10$ GeV/c.

- **MU_W_L2M5_TRK10**

Trigger	L1	L2	L3	Total
MUW_W_L2M3_TRK10	$86.7 \pm 0.7\%$	$97.7 \pm 0.4\%$	$74.3 \pm 4.1\%$	$63.0 \pm 4.2\%$
MU_W_L2M5_TRK10	$84.0 \pm 0.7\%$	$86.3 \pm 0.9\%$	$80.4 \pm 3.1\%$	$58.3 \pm 3.3\%$

Table 7.1: Trigger efficiencies for the dominant single-muon triggers used in this analysis [30].

- Level 1: mu1ptxwtxx_fz: A wide region tight scintillator trigger with a coincidence in the forward luminosity scintillators.
- Level 2: One medium muon, identified at L2, with $p_T > 5$ GeV/c.
- Level 3: One track, reconstructed at L3, with $p_T > 10$ GeV/c.

Table 7.1 details the efficiencies for the two main triggers at the three trigger levels, and the overall efficiencies. For the MUW_W_L2M3_TRK10 and MUW_A_L2M3_TRK10 triggers, we use the efficiency found for MUW_W_L2M3_TRK10. In the case of the MT10W_L2M5_TRK10 and MU_W_L2M5_TRK10 triggers, we apply the efficiency found for MU_W_L2M5_TRK10. The trigger efficiencies for these dominant single-muon triggers can be found in Table 7.1.

7.1.3 Luminosity

The luminosity can be calculated for each global trigger. This is done by summing the average instantaneous luminosity values recorded in the luminosity blocks (LB) for the data, as each LB contains a record of the triggers exposed during its corresponding time period. However, before this luminosity integration can proceed, one must remove all LB's from the list which correspond to data-taking periods in which the data was corrupted in any way. This data corruption can happen for several reasons, the most common of which are

- The DAQ system can experience problems which cause data flow to stop. These instances can arise from any of the three levels of the trigger and result in a LB for which no data is written (or written at low efficiency).

- After each run, the data quality from each subdetector is recorded. For exampole, some runs may consist of events in which the muon system is not operating properly, but the tracking detectors and calorimeter are. Such a run would be useful for analyses not sensitive to muon information, but useless for others.

- After the data is recorded, offline checks of the data quality can reveal instances of detector noise or otherwise degraded events. A detailed analysis is generally performed and a list of "bad" LB's is compiled.

To ensure a normalizable data set, we remove all corrupted LB's. Futhermore, to ensure the quality of the data, we remove all runs which have exhibit degraded detector performance for the following subdetectors: SMT, CFT, calorimeter, and muon system. Finally, we reject

Trigger	p14.03	p14.05	p14.06	Total
EM_MX_SH	32.6	87.5	7.6	127.7
E1_SHT20	42.7	0.0	0.0	42.7
Total	75.3	87.5	7.6	170.4

Table 7.2: Integrated luminosity for electron triggers, before good run/LBN selection, for the 1EMLOOSE data sample. The units presented are pb^{-1}.

the list of LB's which have been marked as containing calorimeter noise by offline analysis. The integrated luminosity for the dominant triggers before the removal bad runs and bad LB's are listed in Tables 7.2 and 7.4. The results after the removal of bad runs/LB's are shown in Tables 7.3 and 7.5. Because triggers can overlap, the total recorded luminosity does not have to equal the sum of the separate contributions in each column.

Trigger	p14.03	p14.05	p14.06	Total
EM_MX_SH	29.9	84.5	7.5	121.9
E1_SHT20	42.4	0	0	42.4
Total	72.3	84.5	7.5	164.3

Table 7.3: Integrated luminosity for electron triggers, after good run/LBN selection, for the 1EMLOOSE data sample. The units presented are pb^{-1}.

7.2   Simulated Samples

To study event rates in data, we produced full event-level and detector-level simulations of nearly all SM processes which may result in a W boson plus HF jets signature.

Trigger	p14.03	p14.05	p14.06	Total
MUW_W_L2M3_TRK10	75.5	42.9	0.1	118.5
MT10W_L2M5_TRK10	57.9	52.0	0.2	110.1
MU_W_L2M5_TRK10	0.1	30.0	0.9	31.0
MUW_A_L2M3_TRK10	23.2	29.6	0	52.8
Total	75.6	72.9	1.0	149.5

Table 7.4: Integrated luminosity for muon triggers, before good run/LBN selection, for the 1MULOOSE data sample. The units presented are pb^{-1}.

Trigger	p14.03	p14.05	p14.06	Total
MUW_W_L2M3_TRK10	72.4	42.4	0.1	114.9
MT10W_L2M5_TRK10	54.9	43.7	0.2	98.8
MU_W_L2M5_TRK10	0.1	28.8	0.5	29.4
MUW_A_L2M3_TRK10	20.9	28.8	0.0	49.7
Total	72.5	72.2	0.6	145.3

Table 7.5: Integrated luminosity for muon triggers, after good run/LBN selection, for the 1MULOOSE data sample. The units presented are pb^{-1}.

The Monte Carlo (MC) for all the processes of interest were generated at $\sqrt{s} = 1.96$ TeV, using the CTEQ5L [32] parton distribution functions (PDFs). A Poisson-distributed minimum-bias overlay, with an average of 0.5 events, was included for all events. The $t\bar{t}$ MC events were generated with $m_{top} = 175.0$ GeV. A complete list of the MC event samples can be found in Tables 7.6-7.8. These tables list both the number of events and the effective cross section value for the process. This effective cross section is given in terms of the total cross section for the generated process (with the appropriate generator level cuts) multiplied by the branching fractions appropriate for the specified final state.

MC Process	Generator	$\sigma \times B(pb)$	Events
$t\bar{t} \rightarrow \ell\nu b\, q\bar{q}\bar{b}$	ALPGEN	2.36	46k
$t\bar{t} \rightarrow \ell\nu b\, \ell\nu\bar{b}$	ALPGEN	0.56	53k
$t\bar{t} \rightarrow q\bar{q}b\, q\bar{q}\bar{b}$	ALPGEN	2.51	32k
tb (s-channel), $(W \rightarrow e, \mu\nu)$	COMPHEP	0.23	30k
qtb (Wg-fusion), $(W \rightarrow e, \mu\nu)$	COMPHEP	0.52	32k
$WZ \rightarrow \ell\nu q\bar{q}, q\bar{q}\ell\ell$	ALPGEN	0.72	48k
$ZZ \rightarrow \ell\ell q\bar{q}$	ALPGEN	0.21	24k
$WW \rightarrow \ell\nu q\bar{q}$	ALPGEN	1.20	23k
$W + n$-jet $(W \rightarrow \ell\nu)$	PYTHIA	9162	300k
$Z + n$-jet $(Z \rightarrow \ell\ell)$	PYTHIA	882	300k

Table 7.6: Monte Carlo samples simulated for $t\bar{t}$, single-top, and diboson SM processes.

The samples correspond to leading-order (LO) matrix elements for parton configurations. The $t\bar{t}$, ZZ, WZ, and WW were generated with ALPGEN 1.2 and PYTHIA 6.2 was used to simulate fragmentation and hadronization. For the W/Z+jets samples, the events were interfaced to PYTHIA to simulate the remainder of the event, initial state radiation (ISR), fragmentation, and hadronization. There was no matching done between the generated partons and the parton-level jets. These samples do not include $W \rightarrow \tau\nu$ or $Z \rightarrow \tau\tau$ decays. To ensure proper representation of these decays, the efficiency for the existing samples was

MC Type	Generator	$\sigma \times B(pb)$	Events
$Wb\bar{b}$, $(W \to e\nu)$	ALPGEN	3.35	100k
$Wb\bar{b}$, $(W \to \mu\nu)$	ALPGEN	3.35	100k
$Wb\bar{b} + 1$-jet	ALPGEN	3.0	45k
$Wb\bar{b} + 2$-jet	ALPGEN	1.53	44k
$Wc\bar{c}$	ALPGEN	5.89	41k
$Wc\bar{c} + 1$-jet	ALPGEN	3.76	49k
$Wc\bar{c} + 2$-jet	ALPGEN	2.01	20k
Wc	ALPGEN	64	20k
$Wc + 1$-jet	ALPGEN	40.3	20k
$Wc + 2$-jet	ALPGEN	15.5	20k
$Wc + 3$-jet	ALPGEN	5.3	20k
$W + 1$-jet	ALPGEN	1652	215k
$W + 2$-jet, $(W \to e\nu)$	ALPGEN	287.3	100k
$W + 2$-jet, $(W \to \mu\nu)$	ALPGEN	287.3	100k
$W + 3$-jet	ALPGEN	228	35k
$W + 4$-jet	ALPGEN	76	46k

Table 7.7: Monte Carlo samples simulated for select W+jets SM physics processes.

increased by a factor corresponding to the efficiency of the selection for $W \to \tau\nu$ and $Z \to \tau\tau$ decays in each channel. For example, in the case of $W \to e\nu$ decays, the efficiency was scaled by a factor of $1 + \epsilon_{W \to \tau\nu}/\epsilon_{W \to e\nu} = 1.042$, where the ratios for $W \to e\nu$ and $W \to \tau\nu$ events were evaluated using PYTHIA. The $W \to \mu\nu$, $Z \to ee$, and $Z \to \mu\mu$ events were also scaled in the same manner.

Because the ALPGEN W/Z+jets samples do not have matching between the generator partons and the final state jets enforced, there is no guarantee that the final state n-jet spectrum matches the generated event spectrum. To avoid incorrect calculations of cross sections, we eliminate events in which there are extra jets in the final state. We achieve this by applying two requirements to the events:

- Original partons are matched to parton-level jets, and only one parton-level jet is accepted for any generated parton.

- Events with a greater number of final-state jets than were generated are removed. In practice, this means that an event generated as $W+3$ jets and reconstructed with 4 or more jets will be rejected, and similarly for the full n-jet spectrum.

The cross sections for different samples are determined as follows, and are specific to the specified final state. Unless otherwise specified, weak-boson decays were forced into the leptonic modes.

- For the ALPGEN samples, we use the generator cross sections. The exception is the $t\bar{t}$ samples, which are normalized to 6.77 pb [36].

MC Type	Generator	$\sigma \times B(pb)$	Events
Zb, $(Z \to ee)$	PYTHIA	0.77	30k
Zb, $(Z \to \mu\mu)$	PYTHIA	0.77	30k
$Zb\bar{b}$, $(Z \to ee)$	ALPGEN	0.539	100k
$Zb\bar{b}$, $(Z \to \mu\mu)$	ALPGEN	0.539	100k
$Zb\bar{b}$ + 1-jet	ALPGEN	0.44	14k
$Zb\bar{b}$ + 2-jet	ALPGEN	0.22	10k
$Zc\bar{c}$	ALPGEN	3.07	30k
$Zc\bar{c}$ + 1-jet	ALPGEN	1.14	15k
$Zc\bar{c}$ + 2-jet	ALPGEN	0.41	10k
Z + 1-jet, $(Z \to ee)$	ALPGEN	90.4	100k
Z + 1-jet, $(Z \to \mu\mu)$	ALPGEN	90.4	100k
Z + 2-jet, $(Z \to ee)$	ALPGEN	29.1	100k
Z + 2-jet, $(Z \to \mu\mu)$	ALPGEN	29.1	100k
Z + 3-jet	ALPGEN	21.4	15k
Z + 4-jet	ALPGEN	2.3	-

Table 7.8: Monte Carlo samples simulated for select Z+jets SM physics processes.

- For the PYTHIA samples, we also use the generator cross sections.

- The single-top cross sections (tb and qtb) are from the results of NLO calculations [37].

- For the W/Z+jets samples, the generator cross sections are scaled by a k-factor of 1.3. This value is obtained by comparing LO and NLO cross sections using the MCFM generator [38].

The contribution of the inclusive process $W + b$ is estimated using MCFM to calculate a cross section ratio relative to $W + b\bar{b}$ production. This ratio is $R_{W+b/W+b\bar{b}} = 0.21$ for events in which all jets are required to have $p_T > 15$ GeV/c. The cross section for $W + b$ is taken as this ratio multiplied by the NLO-normalized $W + b\bar{b}$ cross section. In addition, the p_T spectrum for $W + b$ was measured using inclusive W production generated with PYTHIA. The expected efficiency for $W + b$ is then calculated using the efficiency of $W + b\bar{b}$ events in which only one jet is reconstructed convoluted with the efficiency for the jet p_T requirements.

7.3 Estimation of Multijet Background

The data sample in this analysis contains events originating from multijet production that mimics a W-boson signature. As this background arises predominantly from pure QCD interactions, it is commonly referred to as simply QCD background. Currently, MC simulations of these processes do not reproduce features of data well, mostly due to the calculational difficulties in the non-perturbative QCD model. However, much of the data collected arises

from QCD processes and provides a reliable means of predicting this background. To esti-
mate multijets backgrounds, we use the so-called matrix method [39]. This method proceeds
by defining two samples, "loose" and "tight". The loose sample should contain a large frac-
tion ($\geq 80\%$) of multijets background, and the tight sample should represent the selected
analysis sample. These values can be expressed as (for the example of W-boson selection)

$$N_{loose} = N_{QCD} + N_W$$
$$N_{tight} = \epsilon_{QCD} N_{QCD} + \epsilon_W N_W \tag{7.1}$$

where N_{QCD} is the number of multijets events in the loose sample and N_W is the number
of W-boson events in the loose sample. The efficiencies ϵ_{QCD} and ϵ_W represent the rate
at which multijets and W bosons pass the tight selection, respectively. This system of two
equations can be solved to yield the number of multijets events for a particular selection

$$N_{QCD} = \epsilon_{QCD} \frac{\epsilon_W N_{loose} - N_{tight}}{\epsilon_W - \epsilon_{QCD}} \tag{7.2}$$

To measure the value ϵ_{QCD}, events are selected in the $\not{E}_T < 15$ GeV region. Here, the
fraction of real W boson events becomes negligible, and, as seen in Eq. 7.1, the ratio of the
tight to loose selections returns a measurement of ϵ_{QCD}. The measurement of ϵ_W depends
on the difference between the tight and loose selections and will be described below.

In the case of the $W \to e\nu$ selection, the tight and loose samples differ by the requirement
of a matching charged track for the electron. Often, tracks are accidentally matched to small
shower depositions due to soft gluon radiation or other hadronic activity. Here, the value
for ϵ_W can be measured as the electron tracking efficiency, and is found in $Z \to ee$ events.
Events are selected with at least one electron with a matching track and a total of two
electrons which reconstruct to the Z mass within 10 GeV/c^2, thus increasing the likelihood
of originating from a Z decay. The electron track matching efficiency is taken as the fraction
of events in which the second electron also has a matching track. For the $W \to \mu\nu$ selection,
the tight and loose selections differ by the requirement on muon halo and track isolation. In
this scenario, ϵ_W is measured in $Z \to \mu\mu$ events in the same way as in the $W \to e\nu$ selection.

e + jets		
n-jets	ϵ_{sig}	ϵ_{fake}
≥ 1	0.77±0.04	0.028±0.001
≥ 2	0.79±0.05	0.029±0.001
≥ 3	0.79±0.06	0.030±0.002
≥ 4	-	0.028±0.007

Table 7.9: The "tight" selection efficiencies for signal and multijet events in the e + jets
sample.

The extracted values of ϵ_W and ϵ_{QCD} used are given in Tables 7.9 and 7.10. They are
calculated for inclusive multiplicities of N $\geq$ 1, 2, 3, and 4 jets. The statistics for N$\geq$ 4 are
too low to provide a reliable estimate, and we therefore use the values from N$\geq$ 3 for the
larger jet multiplicities.

μ + jets		
n-jets	ϵ_{sig}	ϵ_{fake}
≥ 1	0.89±0.03	0.14±0.02
≥ 2	0.88±0.03	0.16±0.02
≥ 3	0.86±0.04	0.15±0.03
≥ 4	-	0.15±0.05

Table 7.10: The "tight" selection efficiencies for signal and multijet events in the μ + jets sample.

Chapter 8

Analysis Methods

This chapter focuses on the details of the selection and evaluation of the W boson plus heavy-flavor (HF) jets sample. We describe the procedures for identifying W-boson candidates, the treatment of jets in the candidate events, and the HF-quark tagging methods. We also include a discussion of the systematic uncertainties associated with the analysis methods and the means used to make a comparison between the data and SM predictions.

8.1 W-Boson Selection

Selection of events which contain a W boson begins by searching for the decay products. Hadronic W-boson decays are dominant, but the final state signature is rendered unobservable by the large QCD multijet background at the Tevatron. As mentioned earlier, we focus on two of the semileptonic W boson final states: $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$. Here we outline the process of reliably selecting one of these final states.

8.1.1 Primary Vertex Selection

W-boson production begins at the center of the detector in $p\bar{p}$ collisions. The protons and antiprotons are brought into collision in bunches with some finite spatial extent, and thus tend to collide over an extended region in space. As described previously, the actual point where the collision occurs is referred to as the event's primary vertex (PV). As the PV serves as the center of the event's physics coordinate system, a poor measurement of the origin can cause errors in the calculations of p_T, η, and the $\not{E}_T$.

When a PV is reconstructed with many tracks, it's position measurement becomes much more reliable because the PV fit has more constraints. As such, we reject all events with a PV with fewer than three associated tracks. With a well-reconstructed PV, the distribution of the longitudinal PV position assumes a Gaussian distribution with a mean value near the geometrical center of the detector. In some instances, tracking inefficiencies or mismeasurements can cause the reconstructed PV position to fall far from it's nominal value. In addition, multiple interactions in each bunch crossing can create additional vertices and tracks that can confuse the PV-finding algorithm, again resulting in anomalously large position measurements. To mitigate these effects, the PV acceptance is limited by requiring the

PV z-positions (along the beam) to be within $|z_{PV}| > 60$ cm from the geometrical center of the detector. The distributions of these variables are shown in Fig. 8.1, illustrating the ~ 25 cm Gaussian width of the z_{PV} distribution and the number of tracks used to reconstruct the PV.
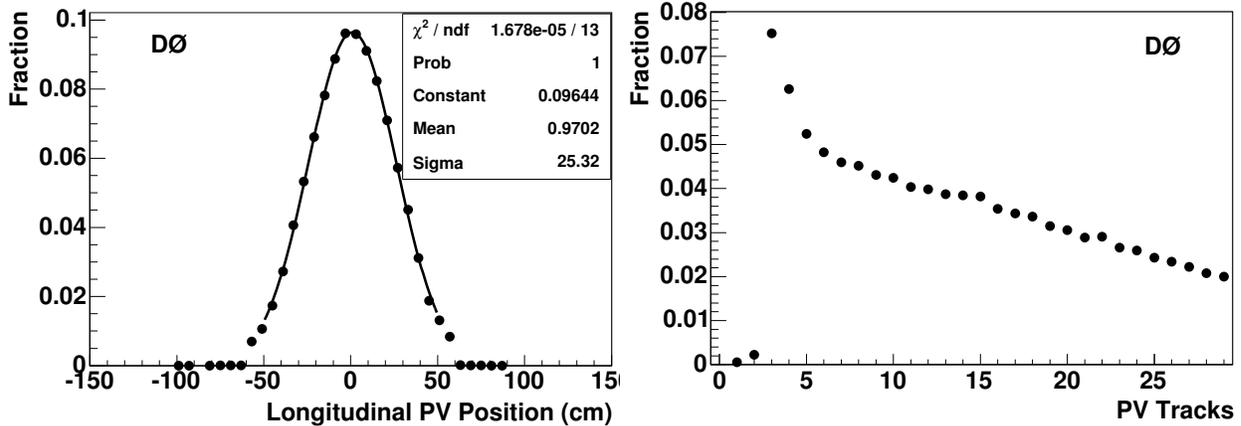
Figure 8.1: The distribution of the primary vertex z position (left) and the number of tracks used to reconstruct the vertex (right).

8.1.2 Lepton Selection

To select leptonic decays of W bosons, events are required to contain one well-identified lepton. To reject events arising from $Z \to \ell\ell$ decays in which one lepton is not detected, as well as contamination from WW, WZ, and ZZ backgrounds, a veto on any additional isolated leptons with $p_T > 8$ GeV/c in the event is imposed. For electrons, the specific reconstruction and quality requirements are:

- One isolated electron passing EM reconstruction criteria;

- Electron shower shape fit $\chi^2 < 20$;

- The trajectory of a charged track from the central tracking detectors matched to the position of energy clustered for the electron in the calorimeter;

- The electron must fall within an acceptable fiducial volume;

- $p_T > 20$ GeV/c, $|\eta| < 1.1$.

In the context of this selection, a lepton is considered isolated if it is separated from all jets in the event by an angular radius $\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.5$. The electron shower shape is compared in the longitudinal and transverse directions with that expected for real electrons. Placing a requirement on the χ^2 of this fit increases the reliability of the calorimeter measurements. Requiring a central track which matches the electron's calorimeter position provides discrimination against particles which do not originate from the PV. In addition,

this requirement helps to eliminate electrons that are faked by neutral hadrons and gluons (which leave no tracks), which are characteristic of QCD multijet backgrounds. As the calorimeter is constructed of three liquid-argon cryostats that are azimuthally segmented into wedges, the areas between the wedges exhibit slightly poorer energy resolution and efficiency. To eliminate systematics from this feature, the fiducial volume surrounding these regions is removed from the acceptance. In addition, electrons are only accepted within the region $|\eta_e| < 1.1$ (measured from the geometrical center of the detector) to ensure they are within regions instrumented for calorimeter triggers, and the requirement on the electrons p_T removes electrons not efficiently triggered upon. This last requirement increases the reliability of the luminosity measurement for electron triggers.

For muons, the reconstruction and quality requirements are:

- One isolated muon passing the medium muon reconstruction criteria;

- The muon must produce hits in all three muon system layers;

- The trajectory of a charged track matched to the position of hits in the muon detectors with minimum track-hit requirements in the central tracker: at least 2 SMT hits and 7 CFT hits;

- Cosmic ray muon rejection via timing requirements in the muon system scintillator detectors;

- The transverse energy deposited in the calorimeter in an annular ring, defined by $0.1 < \mathcal{R} < 0.4$ around the axis of the muon, must be smaller than 2.5 GeV;

- The sum of the p_T of all tracks in a cone of $\mathcal{R} = 0.5$ around the muon direction be less than 2.5 GeV/c. This excludes the track matched to the muon trajectory;

- $p_T > 20$ GeV/c, $|\eta_\mu| < 1.6$;

- For $|\eta_\mu| < 1.1$, further reject the region of $4.25 < \phi_\mu < 5.15$.

Muons reconstructed with fewer than three hits in the muon system tend to have a poorer momentum measurement and a higher incidence of detector noise. To ensure a reliable measurement, we require the muon to have produced hits in all three muon detector segments. In addition, a matching central track rejects faked muons from neutral hadrons and gluons which penetrate the calorimeter, as well as providing a high-resolution momentum measurement. To ensure a good central track match, minimum track-hit requirements are imposed. This requirement is more necessary than in the electron selection because the distance to the muon detectors is much greater than the distance to the calorimeter (from the tracking detectors). In addition, muons have a high incidence of contamination due to cosmic ray muons, which is reduced via the timing cuts. These timing cuts reject muons that are consistent with originating outside the detector. The muon fiducial volume is further reduced in the region $|\eta| < 1.1$ to remove the uninstrumented region where the detector is supported from below. Finally, the p_T and $|\eta|$ requirements ensure the muon is in a triggerable region of high efficiency, thus increasing the reliability of the luminosity measurement.

8.1.3 Missing Transverse Energy

The neutrino from the W-boson decay does not appreciably interact with the matter in the detector as it passes through. Thus, we rely on the transverse momentum imbalance, or $\not{E}_T$, to provide information on the neutrino's momentum. Both selections, $W \to e\nu$ and $W \to \mu\nu$, require a minimum $\not{E}_T$ of 20 GeV. The $\not{E}_T$ value is taken from the event reconstruction and verified to contain proper corrections for all muons in the event, including those that are not isolated. Even after corrections, the $\not{E}_T$ value can still deviate from it's actual value because of mismeasurements in the event. Such mismeasurements could arise from poor measurements of the jets or leptons in the event. In the case that the lepton in the event is grossly mismeasured, the $\not{E}_T$ will tend to point in the same azimuthal direction as the lepton, while in real W-boson events it will point in the opposite direction. Thus, to help reject against poorly reconstructed events and QCD multijet backgrounds, the separation in ϕ between the $\not{E}_T$ and the lepton is required to be at least $\pi/8$. Figure 8.2 shows the resolution of the $\not{E}_T$ in $W \to e\nu$ and $W \to \mu\nu$ events. The results are divided into events with one or two jets, and events with three or more jets. Here, the effects of the jet energy scale corrections are evident as the calorimeter response broadens with an increasing number of jets. These distributions represent the difference between the transverse neutrino momentum in W decays (taken from the MC parton information) and the reconstructed $\not{E}_T$. The distributions have been fit to Gaussian functions, and the fit parameters are shown in the plots. In events with one or more jets, the $\not{E}_T$ resolution varies from approximately 6-13 GeV.
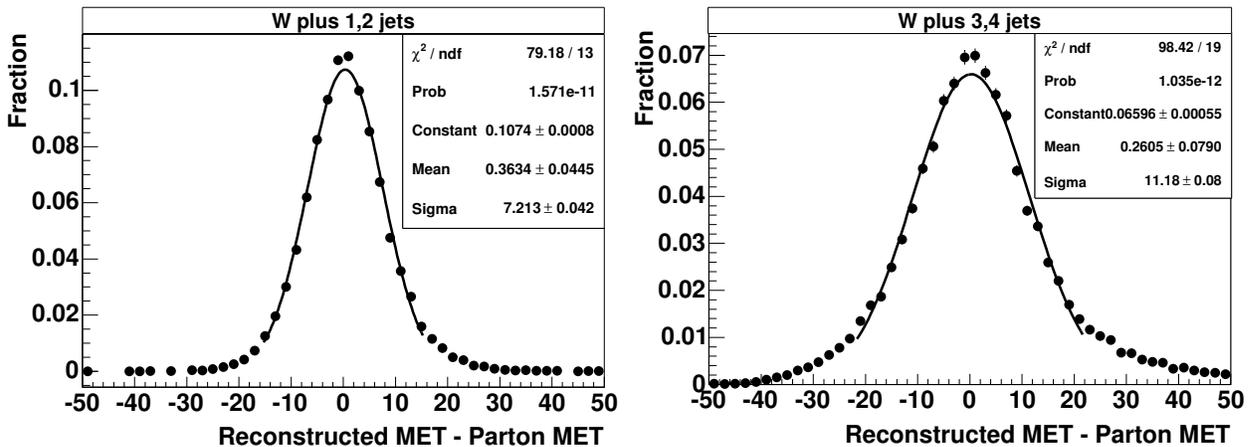
Figure 8.2: The $\not{E}_T$ resolution in $W \to \ell\nu$ plus jets events. The distributions have been divided into events containing 1 or 2 jets (left) and 3 or 4 jets (right).

8.1.4 W Boson Transverse Mass and Efficiencies

After lepton and $\not{E}_T$ selection, the next step is to reconstruct the mass of the W boson to ensure a robust W sample. Because there is no information about the longitudinal momentum of the neutrino from the $W \to \ell\nu$ decay, the W boson mass cannot explicitly be recovered. However, if we assume that the $\not{E}_T$ observed in the event corresponds to the transverse momentum of the neutrino, the projection of the W mass into the transverse plane can be

calculated as follows:

$$M_{W_T} = 2\, p_T^\ell\, \not{E}_T \left(1 - \cos\left(\phi_\ell - \phi_{\not{E}_T}\right)\right) \tag{8.1}$$

where p_T^ℓ is the transverse momentum of the lepton (e or μ), and $\phi_{\not{E}_T}$ is the azimuthal angle of the transverse $\not{E}_T$ vector. The full W mass is well measured by many experiments and the current world average is 80.425 ± 0.038 GeV/c [11]. Thus one expects a Jacobian mass peak, distorted by a varying distribution of transverse W momentum [40]. When the neutrino is produced with low p_z and the $\not{E}_T$ accurately represents its full momentum, M_{W_T} will reconstruct very near the full W mass, widened by the natural width of the W and the detector $\not{E}_T$ and lepton resolutions. With the W width measured to be 2.124 ± 0.041 GeV [11] and the nominal $\not{E}_T$ resolution approximately 10 GeV (see Fig. 8.2), one expects the upper edge of the M_{W_T} distribution to extend to 90-95 GeV/c^2 (higher in cases of poor $\not{E}_T$ or lepton measurements). The lower edge of the distribution will tail off to lower transverse mass with larger neutrino p_z values.

A constraint is placed on this reconstructed mass of $40 < M_{W_T} < 120$ GeV/c^2. On the lower cutoff, this requirement primarily rejects QCD multijet backgrounds which have produced a $W \to \ell\nu$ signature by chance and have a mass distribution exponentially falling from zero. On the upper end, the mass cut rejects poorly reconstructed W bosons and events arising from detector fluctuations. Figure 8.3 shows the distributions of transverse momentum for electron candidates and $\not{E}_T$ in both data and MC for the $W \to e\nu$ selection. Also included in the figures is the estimated QCD multijet background. Figure 8.4 shows the distribution of the transverse mass and p_T of the W, derived from the momentum of the electron and $\not{E}_T$. In Figs. 8.5 and 8.6, we show the analogous distributions for the $W \to \mu\nu$ selection. All four distributions do not involve jet selection or HF-tagging and represent a normalization of the inclusive PYTHIA $W/Z + n$-jet MC to data. These distributions are dominated by events with no jets present and the observed shapes can be seen to be in adequate agreement with the MC expectation.

The cumulative efficiencies for $W \to \mu\nu$ and $W \to e\nu$ MC events are shown for each successive W boson selection step in Tables 8.1 and 8.2, respectively. The final efficiencies represent the values used in the inclusive W selections.

ID Cut:	$W + n$-jet ($W \to e\nu$)	$Wb\bar{b}$ ($W \to e\nu$)	$t\bar{t} \to e\nu qq$
1 EM shower, p_T, and η criteria	29.4±0.2%	32.8±0.3%	33.2±0.2%
Corrected $\not{E}_T > 20$ GeV	25.1±0.1%	27.4±0.2%	30.9±0.2%
$\Delta\phi(p_{Te}, \not{E}_T)$	24.9±0.1%	27.1±0.2%	28.0±0.2%
Isolated lepton veto	24.8±0.1%	27.0±0.2%	27.8±0.2%
Transverse W mass window	24.1±0.1%	25.8±0.2%	25.6±0.2%

Table 8.1: Efficiencies for $W \to e\nu$ selections in Monte Carlo for channels of interest. Events in the three categories were generated using the ALPGEN generator. The efficiencies do not include the impact of the branching ratio for $W \to e\nu$.
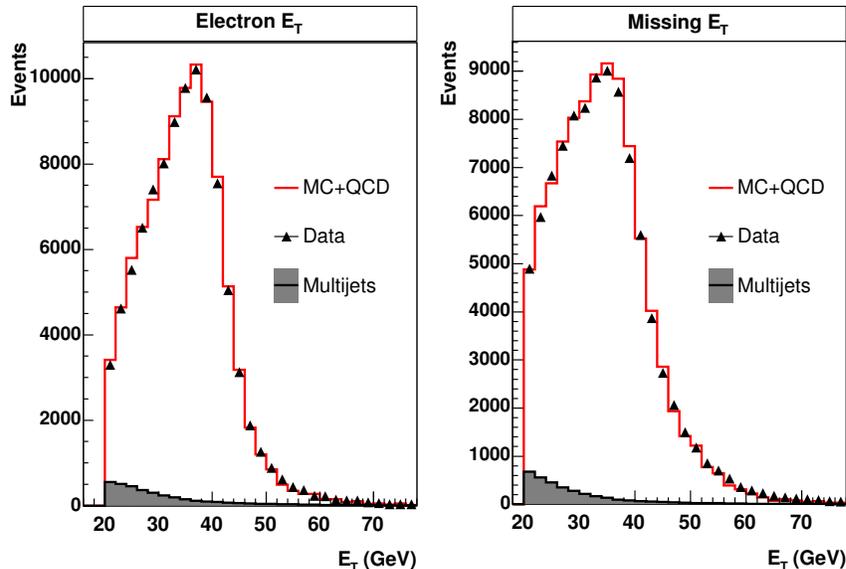
Figure 8.3: Electron p_T and missing transverse energy in the $W \to e\nu$ channel, prior to jet selection.

ID Cut:	$W + n$-jet ($W \to \mu\nu$)	$Wb\bar{b}$ ($W \to \mu\nu$)	$t\bar{t} \to \mu\nu qq$
1 Good muon, p_T, and η cuts	36.9±0.2%	36.9±0.3%	33.3±0.2%
Corrected $\not{E}_T > 20$ GeV	32.0±0.1%	31.7±0.2%	31.3±0.2%
$\Delta\phi(p_{T\mu}, \not{E}_T)$	31.7±0.1%	31.2±0.3%	28.7±0.2%
Isolated lepton veto	31.4±0.1%	31.0±0.3%	28.1±0.2%
Transverse W mass window	30.0±0.1%	28.9±0.2%	25.3±0.2%

Table 8.2: Efficiencies for $W \to \mu\nu$ selections in Monte Carlo for channels of interest. Events in the three categories were generated using the ALPGEN generator. The efficiencies do not include the impact of the branching ratio for $W \to \mu\nu$.

8.2 Jet Selection and HF-Quark Tagging

After completing a robust W-boson selection, the next steps in event selection is jet selection and HF-quark tagging. In this step, it is very important to select well-understood jets and perform the HF-tagging in a manner that provides high purity.

8.2.1 Jet Selection

As mentioned earlier, jets are defined using an iterative seed-based cone algorithm (including mid-points), clustering calorimeter energy within $\mathcal{R} = 0.5$, which is subsequently corrected for the jet energy scale (determined from momentum balance in photon+jet events) [43]. To ensure proper energy-scale (JES) corrections, all jets are required to lie within $|\eta_j| < 2.5$, where η_j is measured with respect to the center of the detector. Also, jets are required to
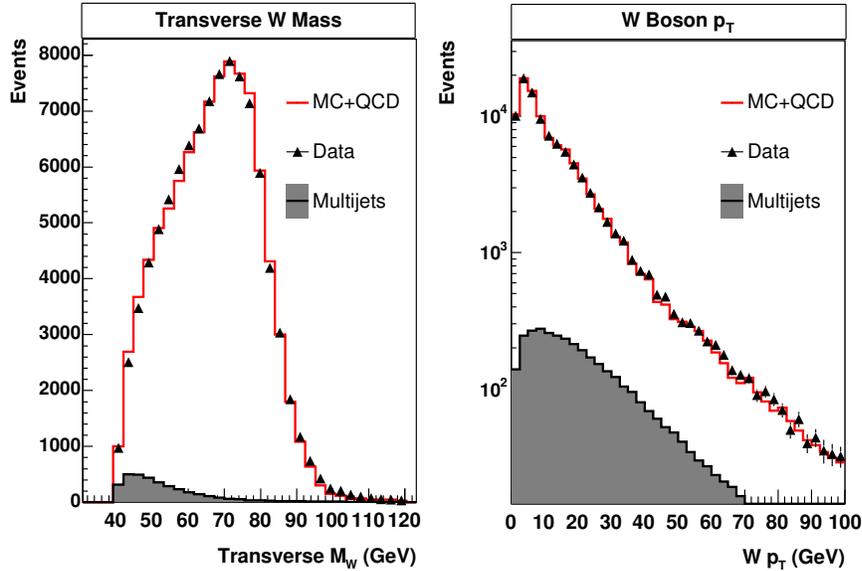
Figure 8.4: Transverse mass of W bosons, and their p_T in the $W \to e\nu$ channel, prior to jet selection.

satisfy $p_T^j > 25$ GeV/c. This constraint selects jets that have well-measured JES corrections and that are have a high reconstruction efficiency. At the cost of ambiguity for efficiency, the jet reconstruction algorithm can reconstruct jets which may also have been reconstructed as electrons, as they both appear as clustered calorimeter energy. To minimize this ambiguity, we therefore do not consider reconstructed jets that are within $\mathcal{R} = 0.5$ of the axis of any electron.

8.2.2 HF-Quark Tagging

After the sample of jets in association with a $W \to e\nu$ or $W \to \mu\nu$ decay is selected, the HF content of the jets is evaluated. This evaluation is performed using two different HF-tagging algorithms: secondary-vertex tagging and soft-lepton tagging.

Secondary-Vertex Tagging

The secondary-vertex tagging (SVT) algorithm relies upon the displaced decay vertices of long-lived hadrons. For this analysis, the TIGHT SVT operating definition is used. Only the jets that pass more stringent quality requirements are considered for SVT tagging. These criteria are not used to eliminate jets from an event, but only to choose tagging candidates for the SVT algorithm. The jets that do not pass these requirements are retained to characterize the total number of jets in the event. Jets used in the SVT algorithm must be matched in (η, ϕ) to *track jets* within a separation of $\mathcal{R} = 0.5$. Track jets are defined by:

- Cone size $\mathcal{R} = 0.5$ in (η, ϕ)

- Extent in the z-direction of $\Delta z = 2.0$ cm
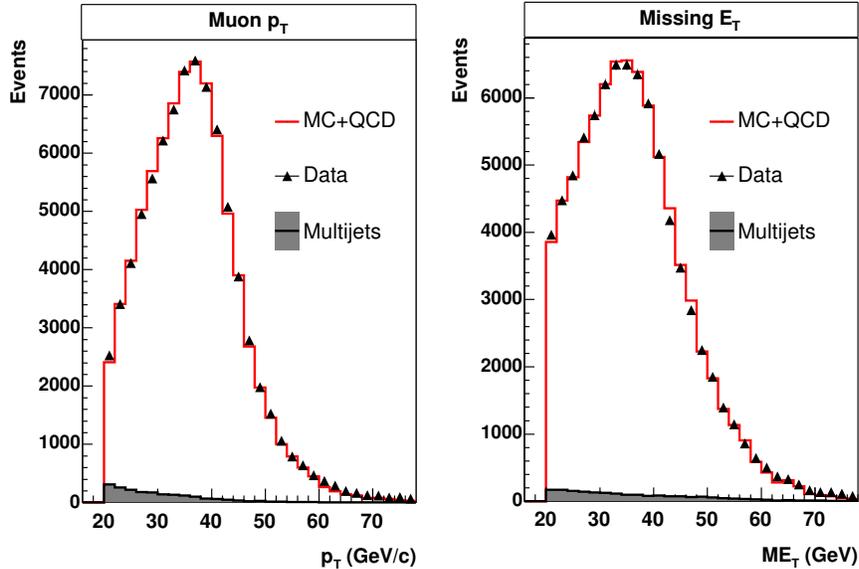
Figure 8.5: Muon p_T and missing transverse energy in the $W \to \mu\nu$ channel, prior to jet selection.

- A $p_T > 0.5$ GeV/c for any included charged track

- Tracks must have at least 2 hits in the SMT

- At least 1 track with $p_T > 1.0$ GeV/c (to be used as the seed)

- Individual track dca of less than 0.2 cm in the transverse plane and less than 0.4 cm in the z direction.

A list of calorimeter jets whose trajectories match to track jets is made. The efficiency for this matching is approximately 85% and is flat as a function of both η and p_T, as seen in Figs. 8.7 and 8.8. These jets are then matched to the reconstructed secondary vertices (described in Sec. 4.3). Jets are deemed SVT tagged if their axis matches to within $\mathcal{R} = 0.5$ of the axis of a secondary vertex. Figure 8.9 shows the SVT HF-quark and light-quark efficiencies for the TIGHT SVT definition. The operating point chosen for this analysis has an average b-quark tagging efficiency of $(33.4 \pm 2.7)\%$, and $(0.22 \pm 0.03)\%$ for light quarks.

Soft-Lepton Tagging

The second method employed for HF-tagging involves low-p_T muons from decays of HF-quarks through virtual W-boson intermediaries. About 18% of HF-quark decays result in a final state with this kind of muon [11], the detection of which provides a means to tag a jet of heavy-flavor. For this HF-tagging algorithm, no requirements are placed on the jets, as they are for the SVT algorithm. Jets with muons close them in (η, ϕ)-space are considered candidates for this type of HF-tag. The muons used for this evaluation must pass the medium muon reconstruction requirements. No muon track match requirement is made, but the momentum of the track is used to describe the muon's momentum in
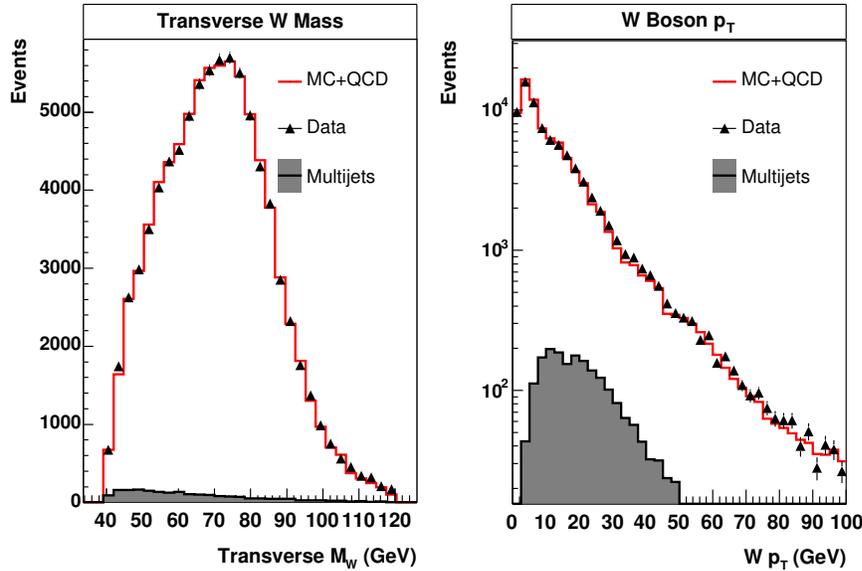
Figure 8.6: Transverse mass of W bosons, and their p_T in the $W \to \mu\nu$ channel, prior to jet selection.

the event of a match. In addition, only muons with $p_T^\mu > 4.0$ GeV/c and $|\eta_\mu| < 2.0$ are considered for soft-muon tagging. Muons that pass certification and the kinematic criteria are matched with jets based on their separation in (η, ϕ)-space. Jets with muons within $\mathcal{R} = 0.5$ are considered SLT tagged. In the case of ambiguity, the muon closest to the jet axis is chosen as the tagging muon. The basic efficiency operating point for this tagging is determined by the lower muon momentum cutoff. Lowering the momentum cutoff lets light-quark background into the algorithm. This occurs because a low-p_T muon is much more easily faked because of it's large curvature in the muon toroids. This feature is compounded by a higher combinatorial probability for low-p_T tracks in jets to match the faked muon hits position. The operating point chosen for this analysis results in a SLT b-tagging efficiency of $4.4 \pm 0.3\%$ and $0.85 \pm 0.08\%$ for light-quark jets.

In instances in which one of the muons from a $Z \to \mu\mu$ decay by chance lies within a jet, the event can easily mimic the signature of a $W \to \mu\nu$ plus SLT tagged jet. To reject any $Z \to \mu\mu$ backgrounds, we require $p_T^\mu < 15.0$ GeV/c. This requirement is performed for both the $W \to e\nu$ and $W \to \mu\nu$ samples to eliminate any bias between the two. In practice, no events are rejected from the $W \to e\nu$ data sample due to this cut.

Events with a SLT contain a muon which can impact trigger efficiency. The single-electron triggers used for $W \to e\nu$ decays depend on calorimeter trigger inputs, and do not receive input from the muon detectors. However, the single-muon triggers used for $W \to \mu\nu$ decays can be impacted by the presence of an additional muon. This effect has been studied and shown to increase single-muon trigger efficiency from $(62.1\pm3.4)\%$ to $(68.4\pm3.5)\%$. The details of this study can be found in Appendix A. This increased trigger efficiency will be used in the normalization of MC events in which a SLT tag is required.
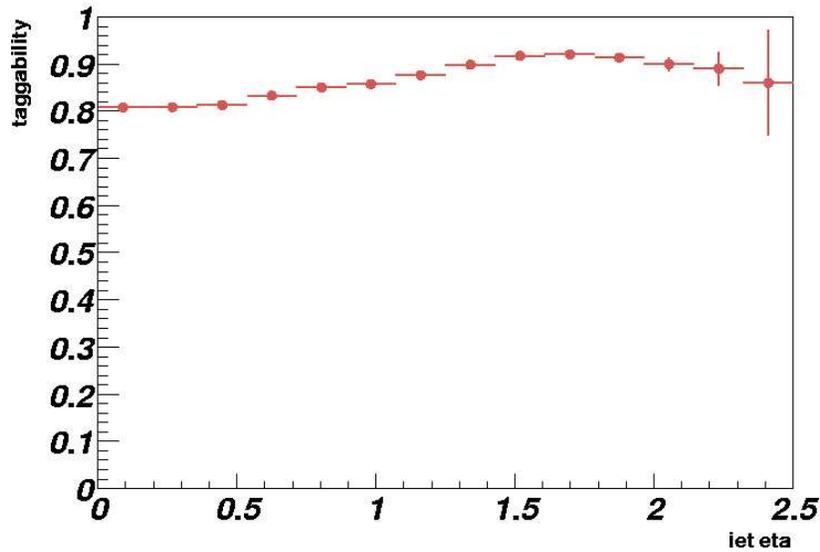
Figure 8.7: The SVT jet taggability efficiency as a function of η, with no requirement on p_T.

8.3 Monte Carlo / Data Normalization

Reconstruction algorithms are never fully efficient, neither on data nor on MC. However, it is important for these efficiencies to match in order to extract reliable measurements from data. With this in mind, we normalize all selection efficiencies in MC to data to account for any discrepancies. Similarly, distributions in variables used for selection criteria may not agree between data and MC, and these must consequently be adjusted to match data. We identify below the instances in which we modify efficiencies or distributions of variables in MC.

8.3.1 Electron Selection

The reconstruction efficiency for electrons in data is defined by the DØ EM-ID group [41], and is found in data to be $\epsilon_{reco}^{data} = 85.0 \pm 1.96\%$ for electrons in the central cryostat ($|\eta| < 1.1$) and for the electron reconstruction criteria outlined in Section 4.4. In MC, this efficiency is found to be $\epsilon_{reco}^{MC} = 96.54 \pm 1.98\%$. From these values we derive a data/MC efficiency correction factor of $f_{reco}^{em} = 0.880 \pm 0.024$, which is used to normalize (downward) the electron reconstruction MC efficiency.

The track-matching efficiency in MC also differs from that found in data. The EM-ID group has measured the electron track-matching efficiency as $\epsilon_{trk}^{data} = 76.91 \pm 2.67\%$ in data [41]. For MC, the electron track-matching efficiency is found to be $\epsilon_{trk}^{MC} = 82.51 \pm 1.96\%$ [41]. We must therefore correct (downward) the electron track-matching efficiency in the MC by the ratio of these two numbers: $f_{trk}^{em} = 0.932 \pm 0.037$.
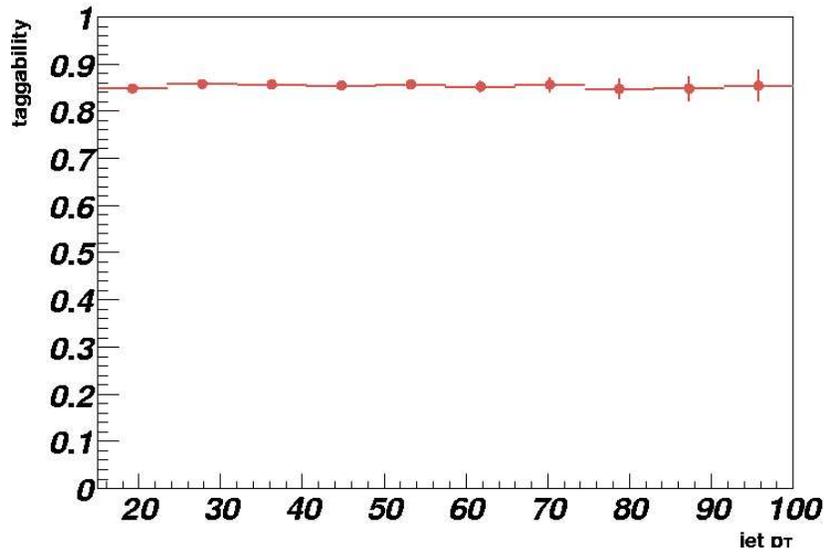
Figure 8.8: The SVT jet taggability efficiency as a function of p_T, with no requirement on η.

8.3.2 Muon Selection

The data/MC efficiencies for muons have been studied extensively in [42]. These studies indicate that the reconstruction efficiency for muons is the same for data and MC, within the uncertainty of the measurement. As such, there is no normalization for this efficiency. The isolation efficiency is found to be $\epsilon_{iso} = (87.0 \pm 0.4)\%$ in data and $\epsilon_{iso} = (89.0 \pm 0.3)\%$ in MC, resulting in an efficiency normalization factor of $f^\mu_{iso} = 0.978 \pm 0.005$. The track-matching efficiency for muons is found to be $\epsilon_{trk} = (92.1 \pm 0.5)\%$ in data and $\epsilon_{trk} = (98.4 \pm 0.2)\%$ in MC, resulting in a track matching efficiency normalization factor of $f^\mu_{trk} = 0.936 \pm 0.005$.

8.3.3 HF-Tagging

Differences in tracking efficiencies between MC and data must be kept at a minimum to be able to extract reliable results for HF-tagging. Any detector inefficiency not modeled in the MC will widen discrepancies, and to mitigate these kinds of deficiencies, we always normalize the MC HF-tagging results to match data.

Primarily because of differences in tracking efficiency between data and MC, the SVT tagging algorithm is more efficient in MC. To account for this discrepancy, the B-ID group provides functions parameterized in (p_T, η) for normalizing the MC SVT tagging results to data. These functions act as scale factors, effectively reducing the efficiency for MC. The details of this procedure can be found in [44].

For the SLT tagging algorithm, differences between MC and data are corrected by scaling the MC muon selection efficiency to match data. We correct for the muon identification efficiency in jets and for the track matching efficiency in jets for muons with a matched track. The corrections are applied per tag and the average efficiency for these corrections in MC is $f^{SLT} = 0.842 \pm 0.056$.
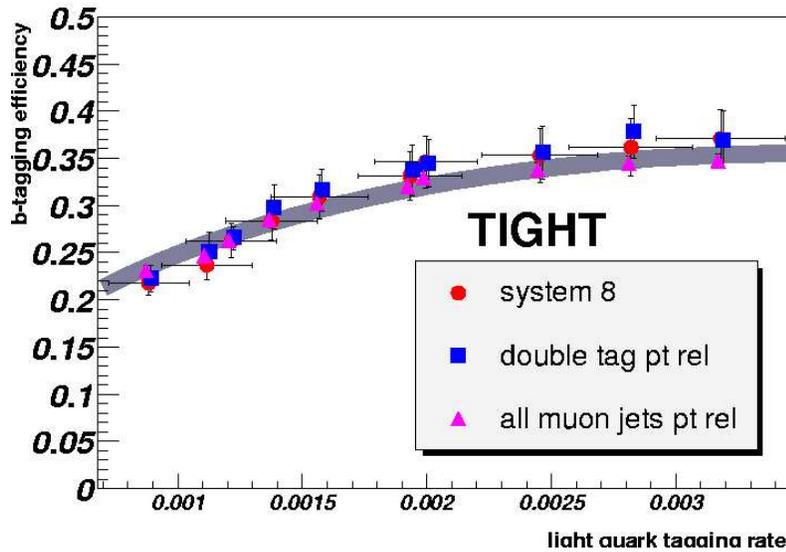
Figure 8.9: The heavy-flavor and light-quark tagging efficiencies for the SVT algorithm.

8.3.4   Smearing of Lepton p_T

To obtain agreement between electron and muon p_T spectra for data and MC, the MC must be smeared to match the resolution in data. To perform this smearing for electrons, we modify the electron energy as follows:

$$E' = \alpha_e \times E \times (1.0 + G) \tag{8.2}$$

where E' is the resulting energy, E is the energy before smearing, α_e is an η-dependent scaling factor, and G is a random number selected from a Gaussian distribution centered at zero with an η-dependent standard deviation σ_e. For electrons with $|\eta_e| < 1.1$, $\alpha_e = 1.003$ and $\sigma_e = 0.045$. All 3 components of electron momentum are scaled accordingly to maintain the electron mass. For muon p_T smearing, we modify the muon momentum components as follows:

$$\frac{1}{p_T'} = \frac{1}{\alpha_\mu \times p_T} + G \tag{8.3}$$

where p_T' is the resulting transverse momentum, p_T is the transverse momentum before smearing, α_μ is an η-dependent scaling factor, and G is a random number selected from a Gaussian distribution centered at zero and η-dependent σ_μ. For muons with $|\eta_\mu| < 1.6$, $\alpha_\mu = 0.991$ and $\sigma_\mu = 0.00226$. For muons with $|\eta_\mu| > 1.6$, we use $\alpha_\mu = 0.999$ and $\sigma_\mu = 0.00465$. The remaining component of the muon momentum (p_z) and the muon energy are both scaled accordingly to maintain the muon mass. The parameters in Eqs. 8.2 and 8.3 reflect the shifts needed in absolute scale as well as resolution to match MC to data.

8.4 Systematic Uncertainties

We classify the sources of systematic uncertainty in the analysis into two categories: those that impact the overall normalization of the selection (i.e., the number of W bosons we select), and those that affect the shape of the selected jet distributions. Of course, some sources of error will fall into both categories, and are considered accordingly.

8.4.1 Systematics from Normalization

The analysis begins by selecting an inclusive W-boson sample and comparing the normalization and shapes of distributions of interest of data with MC. Uncertainties in selection procedures propagate to the normalization, and can influence the expected number of events. We consider the impact of the following sources of systematic uncertainty:

- Efficiencies for electron ID (em_{ID}) and track matching (em_{trk})

- Efficiencies for muon ID (μ_{ID}) and p_T (δp_T^μ) resolution

- Measurement of the absolute luminosity

- Trigger Efficiency

The systematic uncertainties in variables used for lepton selection are available in [46]. In particular, the uncertainty in em_{ID} is $\pm 2.1\%$, the uncertainty in em_{trk} is $\pm 3.1\%$, in μ_{ID} it is $\pm 0.8\%$, and $\pm 3.0\%$ for δp_T^μ. These parameters do not depend significantly on jet multiplicity. For the integrated luminosity, we assign a systematic uncertainty of 6.5%, associated with the absolute uncertainty on the inelastic $p\bar{p}$ cross section [45]. The uncertainties on the trigger efficiencies are taken from Sec. 7.1.

We furthermore consider the systematic uncertainties associated with the MC generation:

- MC cross sections

- $W \to \tau\nu, Z \to \tau\tau$ Factors

We associate a conservative uncertainty of 15% on the cross sections used to normalize the $W + X$ ALPGEN MC samples (where X includes all quark flavors). We use a value of 6.4% [36] for the samples, and 18% for the single-top samples. Futhermore, we associate a 1% uncertainty with the inclusion of $W \to \tau\nu, Z \to \tau\tau$ efficiencies in MC samples that do not include these decays (see Sec. 7.2).

8.4.2 Systematics from Discrepancies in Shapes

After selecting the W-boson sample, we check whether the jets in each event can be HF-tagged. Tagging defines our final selection, and requires full understanding of the shape of distributions of variables most affected by criteria that cause events to shift from bin to bin, or to be rejected. The sources of systematics we consider are:

- Jet Energy Scale (JES)

- Jet Identification

- SVT tagging efficiency

- SLT tagging efficiency

- MC jet-parton matching

The uncertainty in the JES impacts the measurement of $\not{E}_T$, which also impacts the normalization. Changes in jet p_T can affect the acceptance because requirements on jet p_T in the analysis can shift events from bin to bin, or reject them outright. The effect of the uncertainty in JES is studied by changing the p_T of each jet by $\pm$ one standard deviation ($\pm 1\sigma$), where 1σ is defined by the systematic uncertainty in the JES. The systematic uncertainty associated with jet identification is measured to be 4% in [47]. To gauge the impact of a change in HF-tagging efficiency on the shape of distributions in HF-tagged events, we also vary the SVT HF-tagging efficiency by $\pm 1\sigma$ for each tagged jet, where 1σ is defined by the systematic uncertainty in the normalization of the SVT data and MC efficiencies. For the SLT HF-tagging efficiency, we include a systematic uncertainty of 8% obtained by varying the $\mathcal{R} = 0.5$ tagging cut by $\pm 15\%$. The uncertainty associated with the matching of MC partons to reconstructed jets is taken as 3%, which is obtained by varying the $\mathcal{R}$ matching value by $\pm 15\%$.

All major systematic uncertainties associated with this analysis are summarized in Table 8.3, as a function of the exclusive number of jets in each event (including any tagged jets).

Systematic	$N_{jet} = 1$	$N_{jet} = 2$	$N_{jet} = 3$	$N_{jet} \geq 4$
em_{ID}	2.1%	2.1%	2.1%	2.1%
em_{trk}	3.1%	3.1%	3.1%	3.1%
μ_{ID}	0.8%	0.8%	0.8%	0.8%
μ_{pT}	3.0%	3.0%	3.0%	3.0%
Lumi	6.5%	6.5%	6.5%	6.5%
Electron Trigger	2.0%	2.0%	2.0%	2.0%
Muon Trigger	5.2%	5.2%	5.2%	5.2%
MC Cross Section	10 /18%	10 /18%	10 /18%	10 /18%
$W \to \tau\nu, Z \to \tau\tau$ Factors	1.0%	1.0%	1.0%	1.0%
JES	2.0%	2.0%	2.0%	2.0%
Jet ID	4.0%	4.0%	4.0%	4.0%
SVT tagging	9.5%	9.5%	9.7%	9.7%
SLT tagging	8.0%	8.0%	8.0%	8.0%
Parton matching	3.0%	3.0%	3.0%	3.0%

Table 8.3: Summary of systematic uncertainties associated with results for HF-tagged jets, as a function of the total exclusive number of jets, including any tagged jets.

Chapter 9

Results and Discussion

This chapter focuses on the details of the evaluation of the W boson plus heavy-flavor (HF) jets sample. This measurement is performed on the selected W+jets sample using secondary-vertex and soft-lepton tagging algorithms, all of which are described in the previous Chapter 8. We compare these selections to SM predictions and the previous CDF anomalous W+HF-quark production result.

Results reported in this section correspond to 164 pb^{-1}of integrated luminosity in the e+jets selection, and 145 pb^{-1}in the μ+jets selection. As detailed in the previous section, the MC are normalized to these luminosities and corrected for differences between data and MC in HF-tagging and lepton ID efficiencies. Also, any discrepancy in trigger efficiency between MC and data is taken into account for each set of selections (e + jets and μ + jets). Because these selections have different integrated luminosities, we treat each one separately before combining. After a the normalizations, we sum the two samples to improve the statistical uncertainty of the analysis.

9.1   Event Selection Results

To ensure a robust analysis, we examine the event sample at different points in the selection process. First, we will discuss the W+jets sample before HF-tagging has been performed. After this step, we will impose the HF-tagging requirements to produce the final selected event samples.

9.1.1 W plus Jet Selection

Figure 9.1 shows the exclusive number of jets in events with a selected W boson. The fourth bin in the plot includes the sum of four or more jets and the plot legend indicates the expected SM processes contributing to the observed data sample. The numbers of expected and observed events for this distribution are summarized by source in Table 9.1, demonstrating reasonable agreement between data and SM predictions at this stage of the analysis. At this point, it is prudent to test whether the observed W-boson selection variables agree with the SM predictions for the exclusive jet distributions. These results are discussed in Appendix B, and show adequate agreement.
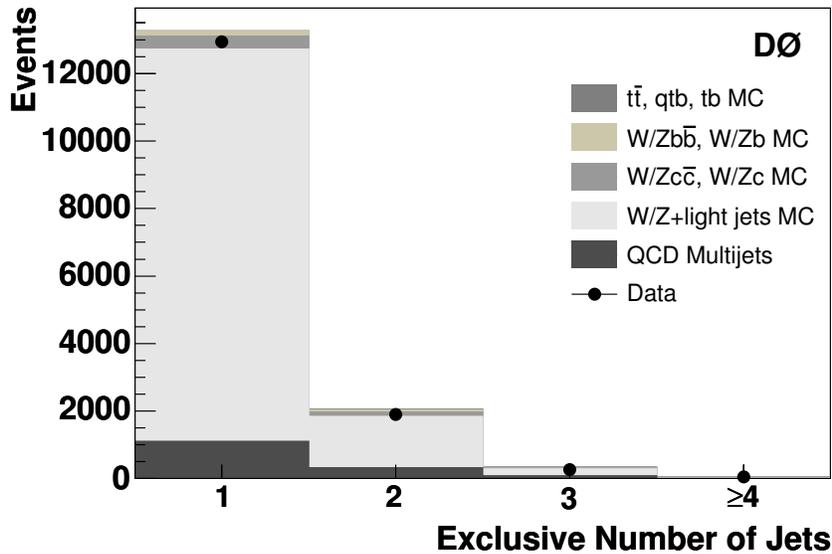
Figure 9.1: The exclusive number of jets with $p_T > 25$ GeV/c in events with a selected W boson, prior to requiring HF-tagging.

9.1.2 W plus HF-Quark Selection

Figure 9.2 shows the exclusive number of jets in events with at least one jet tagged with the SVT algorithm. The last bin in the plot contains the sum of four or more jets and there can be more than one SVT-tagged jet in any event. Figure 9.3 shows the p_T distribution of SVT-tagged jets. Any SLT tag is simply ignored in these plots and can be included if the event also contains a SVT tag. The distributions of expected and observed events with at least one SVT-tagged jet are summarized by source in Table 9.2.

Figure 9.4 shows the exclusive number of jets in events with at least one SLT-tagged jet. The format for these plots is the same as that for the SVT plots in the previous section. Figure 9.5 shows the p_T distribution of SLT-tagged jets. The distribution of expected and observed events with at least one SLT b-tagged jet is summarized by source in Table 9.3.

It is important to ensure that the process of HF-tagging does not introduce a significant bias in the transverse W-boson mass distribution for selected events. Such a bias could emanate from the primary lepton momentum or the $\not{E}_T$. In practice, the decays of b quarks will contain neutrinos which can impact the $\not{E}_T$ value. This effect can, in turn, impact the reconstructed W boson mass. Figure 9.6 shows the distribution of the measured transverse W boson mass for events in which at least one jet was HF-tagged with either the SVT or SLT tagging algorithms. Here we see a small enhancement of the high-M_{WT} tail, as expected for inclusive b-quark decays. Aside from this, no significant departure from the transverse mass distributions obtained before HF-tagging.

Jets tagged with both algorithms should provide a cleaner sample of heavy-flavor jets. The results for events with at least one "doubly-tagged" jet, supposedly enriched in heavy-quark content, is shown in Fig. 9.7. The distribution of expected and observed events with at least one jet tagged by both the SLT and SVT algorithms is summarized in Table 9.4.

Source	W+1jet	W+2jets	W+3jets	W+$\geq$4jets
W/Z+jet	11146±2190	1448±284	203±40	25±5
Multijet	1113±391	323±114	80±28	23±8
WW,WZ,ZZ	11.2±1.8	14.9±2.4	1.2±0.2	0.1±0.01
Wc	274.9±54.0	83.8±16.5	13.6±2.7	1.6±0.3
$W/Zc\bar{c}$	97.5±19.2	37.7±7.4	6.2±1.2	0.5±0.1
W/Zb	20.7±4.1	7.8±1.5	1.2±0.2	0.1±0.02
$W/Zb\bar{b}$	127.1±25.0	60.9±12.0	11.1±2.2	0.9±0.2
$t\bar{t}$, Single top	6.6±1.0	21.2±3.4	21.8±3.4	15.5±2.5
SM prediction	12796±2685	1997±441	338±78	66±16
Data	12928	1899	289	58

Table 9.1: Summary of the exclusive number of jets with $p_T > 25$ GeV/c in events with a selected W boson, prior to requiring HF-tagging. The fourth bin represents the integral of four or more jets.

The correlation between the SVT and SLT algorithms has been studied in MC. These studies find no significant correlation between algorithms, and the details about these correlations can be found in Appendix C.

9.2 Limits on Anomalous Heavy-Flavor Quark Production

Upon finding adequate agreement between data and SM MC predictions in the W+HF-tagged jet samples, we can derive limits on anomalous HF-quark production in association with W bosons. We have three selections from which to derive limits (SVT, SLT, & doubly-tagged). However, only the doubly-tagged jet sample has the sensitivity to yield limits on the scale of the dominant SM contributions.

9.2.1 Model Independent Limits

Because we have not suggested a possible model for the production of anomalous events, we cannot base limits on any model for new physics. In the absence such a model, we quote limits on the number of expected events per exclusive jet bin. To determine limits, we calculate the 95% confidence level (C.L.) for additional event production in each bin using a modified Frequentist method, also referred to as the CL_s method [48].

We begin by defining the values N_s and N_b as the number of anomalous (signal) and expected SM (background) events, respectively, for each distribution. The value N_{s+b} is then defined as the sum of these two numbers: $N_{s+b} = N_s + N_b$. To evaluate a limit, we assume a Poisson distribution for repeated trials of each measurement. We begin by populating a normalized distribution with a large number ($\sim 10^8$) of Poisson trials with a

Source	W+1jet	W+2jets	W+3jets	W+$\geq$4jets
W/Z+jet	56.1±11.0	7.2±1.4	3.8±0.7	0.6±0.1
Multijet	13.4±4.7	6.8±2.4	1.5±0.5	0.6± 0.2
WW,WZ,ZZ	0.26 ±0.04	0.71±0.11	0.08±0.01	0.01±0.01
Wc	10.7±2.1	3.8±0.8	0.8±0.2	0.1±0.02
$W/Zc\bar{c}$	2.2±0.4	1.8±0.4	0.41±0.08	0.05±0.01
W/Zb	2.8±0.6	1.8±0.4	0.32±0.06	0.03±0.01
$W/Zb\bar{b}$	13.9±2.7	12.4±2.4	2.7±0.5	0.23±0.04
$t\bar{t}$, Single top	1.0±0.2	5.0±0.8	6.2±1.0	5.1±0.8
SM prediction	100.3±21.7	39.5±8.6	15.7±3.1	6.6±1.2
Data	104	37	18	6

Table 9.2: Summary of observed and predicted W-boson candidate events with at least one SVT-tagged jet.

mean value given by N_{s+b}. For each entry in the Poisson distribution, the N_b value is drawn from a Gaussian distribution with a mean value given by the expected SM background found in this analysis and a width given by the error on that expectation. The value N_s is held fixed at the number of anomalous signal events being tested. Next, we define the value CL_{s+b} as the integral of the normalized Poisson distribution below the observed number of events in data for $N_s > 0$, and CL_b as the integral below the observed number of events for $N_s = 0$. The value of N_s is increased until the value

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{9.1}$$

falls below 0.05, or $CL_s < 0.05$. The N_s value at which this test is satisfied is defined as the 95% C.L. limit for the event rate for anomalous heavy-flavor production in association with a W boson. Table 9.5 shows the values extracted for each exclusive jet bin. These limits represent the number of additional events, per jet bin, that could be observed and still remain consistent with SM results at 95% C.L. Given a particular model, a researcher could extract cross-section limits by calculating the model acceptance in each exclusive jet bin, with the luminosities for each sub-channel (e + jets and μ + jets).

9.2.2 SM Anomalous Production Limits

Assuming that anomalous heavy-flavor quark production has the same event topology as some SM process, the event limits derived above can be translated into upper limits on cross sections. To this end, we consider three scenarios which cover the full exclusive jet range:

- "$Wb\bar{b}$-like" production in which two b quarks are produced in association with a W boson. In this scenario, additional light quarks or gluons can be produced, and thereby shift the event topology to more than two jets. Jets not within the acceptance of the
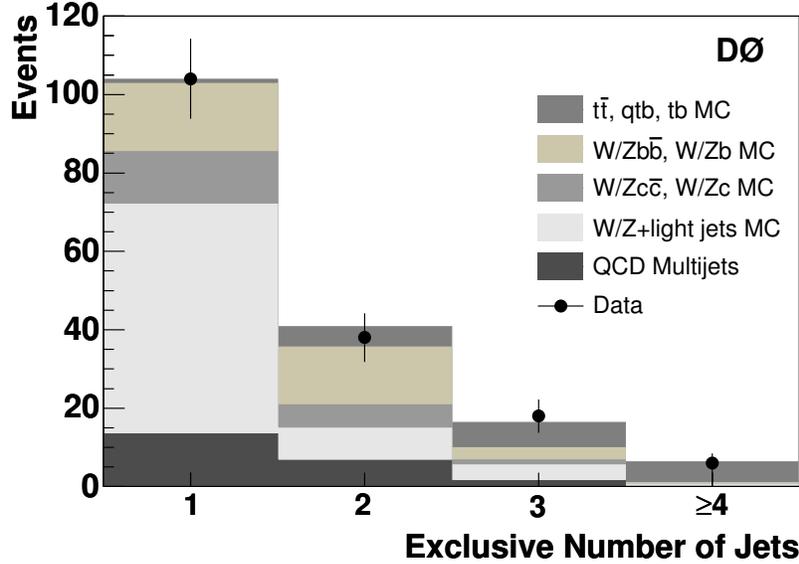
Figure 9.2: Exclusive jet multiplicity for events with at least one SVT-tagged jet. The fourth bin represents the integral of four or more jets.

    detector can also cause the event topology to drop to less than two jets. We model this production using the cross-sections and efficiencies for SM $W/Z + b\bar{b}$ production.

- "Single-top-like" production in which one heavy particle is produced in association with one or more quarks, that could possibly be possibly b quarks. The event can contain additional light or heavy quarks and gluons. We model this scenario using the cross-section and efficiency for SM single top production, dominated by two- and three-jet topologies.

- "$t\bar{t}$-like" production in which a two heavy particles are produced and decay to a W boson and a HF quark. The event can also contain additional light or heavy quarks and gluons. We model this scenario using the cross-section and efficiency for SM ttb production, dominated by three- and four-jet topologies.

    We evaluate an upper cross-section limit on exclusive jet production for each scenario, but first ignore the efficiency for reconstructing the predicted jets. The remaining efficiency represents W-boson selection and HF-tagging. To extract limits for a specific model, this efficiency must be multiplied by the probability to reconstruct the number of jets given in each exclusive jet bin. Given a model with W-boson production similar to that observed in one of the three scenarios, a researcher can calculate the specific jet acceptances in the exclusive jet bin(s) of interest. These results are presented in Table 9.6.

    To evaluate a limit on inclusive jet production for each scenario, we reintroduce the efficiency for reconstructing the predicted jets. For inclusive $Wb\bar{b}$-like anomalous production, we sum the first two exclusive $W + n$-jet bins, as the contribution from the remaining bins is negligible. For $t\bar{t}$-like and single-top-like anomalous production, we sum all $W + n$-jet bins,
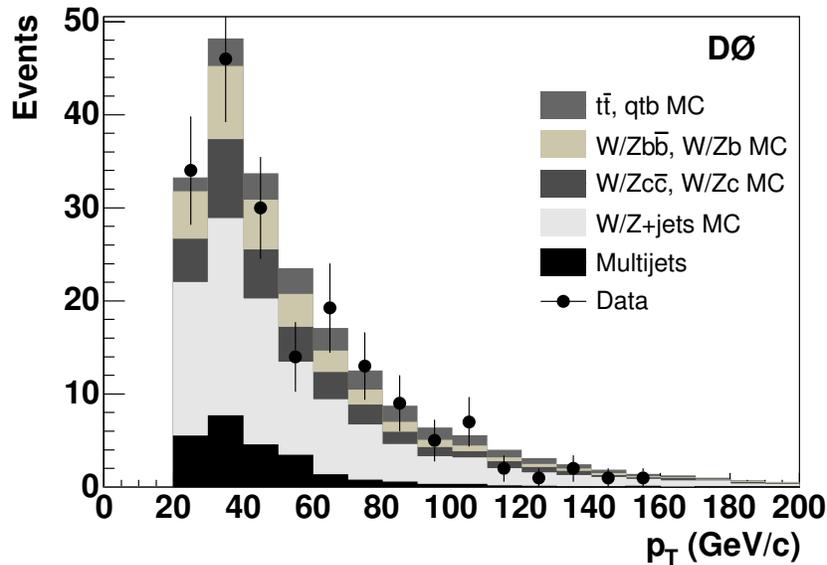
Figure 9.3: Transverse momentum for jets which have been tagged with the SVT tagging algorithm.

except for the $n = 1$ bin, where the contribution is also negligible. For these last two scenarios, we also report limits based on a sum of the events in the $W+2,3$ jet bins for comparison to the CDF anomalous result. Table 9.7 shows the 95% C.L. event limits for the combination of jet bins for these three hypotheses, and also the corresponding anomalous heavy-flavor quark production cross section limits. The jet reconstruction efficiency is included in the calculations and the limits contain the expected efficiency of the specified SM processes.

9.3 Discussion and Conclusions

9.3.1 Comparison to SM Predictions

We have measured the HF-tagging rates in data events that also contain a W boson and compared these results to SM MC predictions. The MC prediction for the SVT-tagged jet sample demonstrates a χ^2 deviation from the data observation of 0.94 for the four exclusive jet bins, while the SLT-tagged jet sample has a χ^2 sum of 3.38. In these selections, we find no significant departure from SM predictions. The largest effect in these two distributions is a small systematic data deficit in the $W + 2$-jet bin. This deviation could be explained by either the modeling of the processes in that bin, or by a statistical fluctuation in the data. In either case, the data observation is still within 1σ of the SM prediction. Furthermore, the W boson plus doubly-tagged jet sample demonstrates good agreement between data and SM MC processes. These tests increase our ability to limit the rate of rare processes in W plus heavy-flavor quark production.

Source	W+1jet	W+2jets	W+3jets	W+$\geq$4jets
W/Z+jet	42.2±10.3	10.4±2.0	3.2±0.6	0.26±0.05
Multijet	10.2±3.6	3.9±1.4	1.2±0.4	0.51± 0.18
WW,WZ,ZZ	0.15±0.02	0.36±0.06	0.04±0.01	0.01 ±0.01
Wc	8.6±1.7	2.1±0.4	0.68±0.13	0.05±0.01
$W/Zc\bar{c}$	2.1±0.4	1.47±0.29	0.31±0.06	0.02±0.003
W/Zb	1.2±0.2	0.84±0.16	0.13±0.03	0.02±0.003
$W/Zb\bar{b}$	5.9±1.2	5.3±1.0	1.1±0.2	0.13±0.02
$t\bar{t}$,Single top	0.36±0.06	1.8±0.3	2.2±0.4	1.9± 0.3
SM prediction	80.7±17.4	26.1±5.7	8.9±1.8	2.9±0.6
Data	81	21	8	2

Table 9.3: Summary of observed and predicted W-boson candidate events with at least one SLT -tagged jet.

Source	W+1jet	W+2jets	W+3jets	W+$\geq$4jets
W/Z+jet	2.2±0.4	0.05±0.01	0.15±0.03	0.0±0.0
Multijet	1.2±0.4	0.5±0.2	0.1±0.1	0.0± 0.0
W/Zb,W/Zc	0.58±0.11	0.18±0.03	0.04±0.01	0.02±0.003
$W/Zb\bar{b}$,$W/Zc\bar{c}$	0.91±0.18	0.87±0.17	0.20±0.04	0.02±0.004
$t\bar{t}$, Single Top	0.07±0.01	0.32±0.05	0.44±0.07	0.33±0.05
SM prediction	5.0±1.2	2.0±0.5	1.0±0.2	0.4±0.1
Data	5	1	1	0

Table 9.4: Summary of observed and predicted W-boson events with at least one jet tagged by both the SLT and SVT algorithms.

9.3.2 CDF Anomalous Result

Using the cross-section limits we've derived, the anomalous HF-quark production result from CDF can be addressed. The CDF result was introduced in Sec. 6.1.3 and is characterized by an excess of events in the W+2,3 jet bins, shown in Fig. 6.9. As is evident, there are several interpretations that can be made about the distribution of events in the CDF W plus doubly-tagged sample. We will address the three most relevant scenarios.

1. The excess seen in the first three bins of the CDF W+SLT distribution (see Fig. 6.8) indicates an underestimation of the SLT efficiencies between MC and data. The SLT-only distribution for CDF consists of approximately 75% light-quark backgrounds (mistagged jets), with the rest consisting of HF-quark processes. The CDF SLT efficiency measurement for HF quarks (used for MC normalization) could be underestimated by performing its measurement in a sample with an over-represented light-quark content. With the additional requirement of the SVT tag (for doubly-tagged jets), this
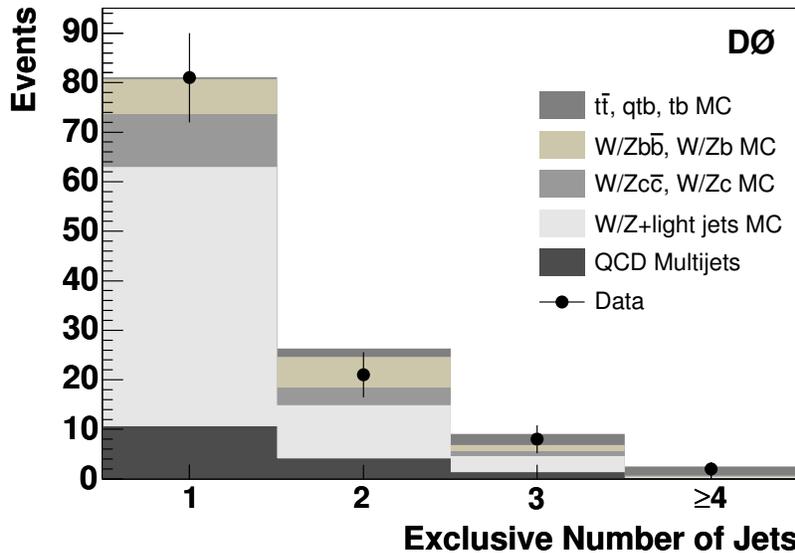
Figure 9.4: Exclusive jet multiplicity for events with at least one SLT-tagged jet. The fourth bin represents the integral of four or more jets.

Source	$W+$1jet	$W+$2jets	$W+$3jets	$W+\geq$4jets
Data observation	5	1	1	0
SM prediction	4.9±1.2	2.0±0.5	0.94±0.18	0.37±0.05
95% C.L. Limit (events)	6.68	3.86	4.14	3.00

Table 9.5: Observed and predicted W-boson events with at least one jet tagged by both the SLT and SVT algorithms. Also shown is the 95% C.L. limit in the form of additional expected events.

discrepancy would become more pronounced in the richer HF sample as the light-quark contribution is reduced. Assuming that the light-quark SLT rate is measured well, a simple calculation can be performed to determine the proper per-jet SLT HF-tagging efficiency:

$$f_{HF}(n) = (P_{false}/P_{true})^n \tag{9.2}$$

where $f_{HF}(n)$ is the fraction of the correct HF content for each $n-$jet bin, P_{false} is the incorrect per jet SLT HF-tagging efficiency, and P_{true} is the true efficiency[1]. This condition is most closely satisfied with a value of $P_{false}/P_{true} = 0.787$, determined by increasing the HF-contribution of each bin to eliminate the data excess observed and minimizing the deviation over the jet bins. Correcting for this factor and assuming no

---

[1]If one assumes that the data and MC predictions should agree, then the true efficiency is determined by the additional SLT rate needed to achieve agreement.
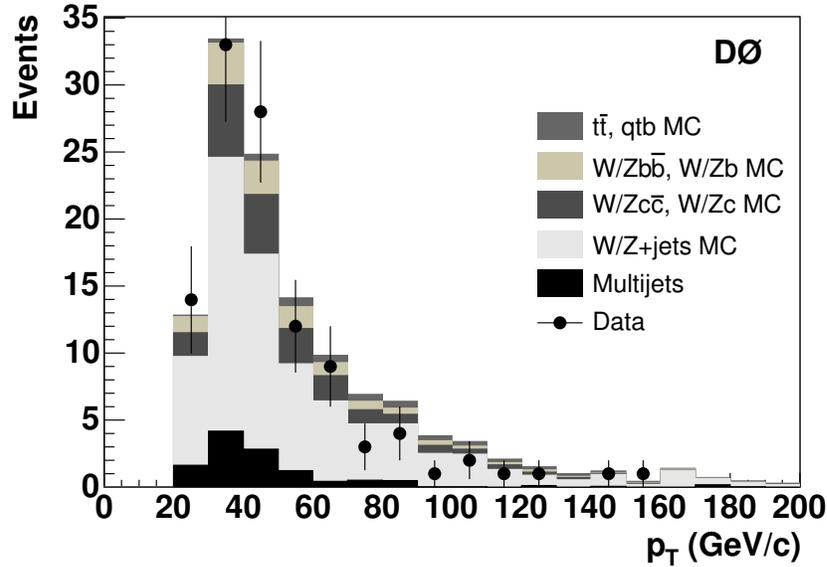
Figure 9.5: Transverse momentum for jets which have been tagged with the SLT tagging algorithm.

Model	$W + 1$ jet	$W + 2$ jets	$W + 3$ jets	$W+ \geq 4$ jets
Wbb-like	35.0	9.1	6.0	4.5
Single-top-like	33.3	16.2	12.5	8.5
$t\bar{t}$-like	26.4	10.2	11.7	12.6

Table 9.6: Cross-section limits in pb, based on the hypotheses of "top-like" anomalous production and "$Wb\bar{b}$-like" anomalous production of exclusive number of jets. Each value is corrected for the efficiency of reconstructing the predicted number of jets in each jet bin.

SVT/SLT correlations, the excess in the $W+2,3$ jet bins of the CDF doubly-tagged sample drops from 8.6 events to 4.7 events. As a fraction of the SM prediction, the excess drops from 194% to 54% of the SM prediction.

2. The one-jet bin in the CDF doubly-tagged distribution is deficit in data. Considering that the general shapes of the CDF and DØ MC distributions agree, such a deficit is unexpected. Such an offset can occur if there is a difference between the jet selection efficiencies between MC and data. In particular, differences in the jet energy scale correction could cause such an effect by causing more jets to be accepted by the p_T cut than would otherwise be accepted. If this bin is summed with the $W+2,3$ jet bins, the excess over the SM prediction falls from 8.6 to 5.6. As a fraction of the SM prediction, the excess falls from 194% to 66%.

3. The third interpretation, which has been used most by the CDF collaboration, is that the excess in observed events in the $W+2,3$ jet bins represents a contribution from a
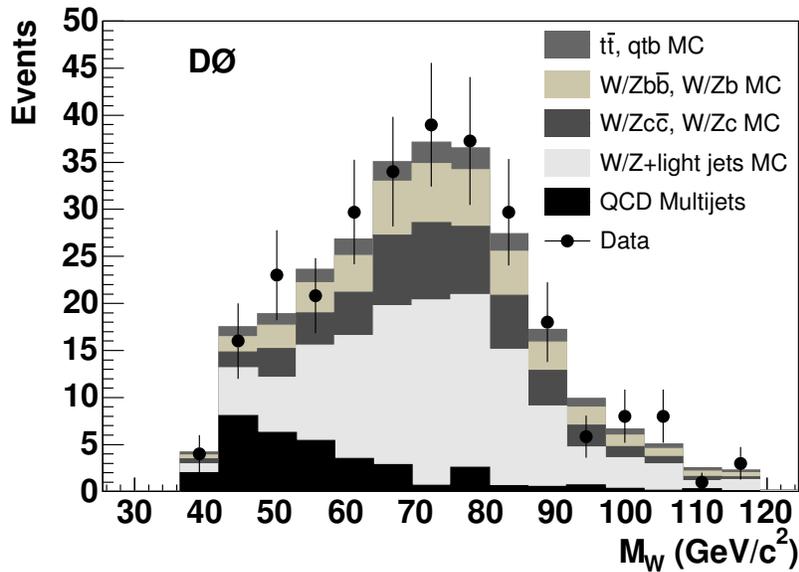
Figure 9.6: Transverse W-boson mass for events with at least one HF-tagged jet.

non-SM source of physics. This scenario cannot be dismissed without an independent measurement. Such a measurement has been discussed in this dissertation, and we will address the hypothesis of anomalous HF production in detail below.

Making a direct comparison of the results and limits described here to the CDF anomalous result is difficult. Differences in efficiencies, cross sections, and the center-of-momentum energy cause inconsistencies in comparisons between the two results. However, we can identify a scenario in which a direct comparison can be motivated: with the same method of calculating upper cross-section limits used in Sec. 9.2.2, we can calculate an anomalous cross section for the CDF sample using the SM production scenarios outlined above. These values can be compared directly to the limits found in this dissertation, and a comparison of these numbers can be found in Table 9.8. To calculate these numbers for the CDF result, the cross sections and selection efficiencies from the CDF study are used. This allows the same process to be compared with the luminosity, cross-section, and center-of-mass energy differences factored out. Table 9.8 indicates that using the $Wb\bar{b}$-like scenario, no exclusion of the comparable CDF observation can be made. However, for the single-top-like and $t\bar{t}$-like scenarios, an exclusion of the CDF result can be observed at a level of $CL_s = 2.2 \times 10^{-2}$ and $CL_s = 4.0 \times 10^{-2}$, respectively.

9.3.3 Conclusions

We have presented an analysis of events in which a W boson was selected in either the $W \to e\nu$ or $W \to \mu\nu$ decay channel. After this selection, we examined the jets in these events for possible HF-tags, using both secondary-vertex and soft-muon tagging algorithms. In the e + jets channel we analyzed 164 pb^{-1} of data, and 145 pb^{-1} of data in the μ + jets channel. At this time, we see no significant departure from the predictions of the standard
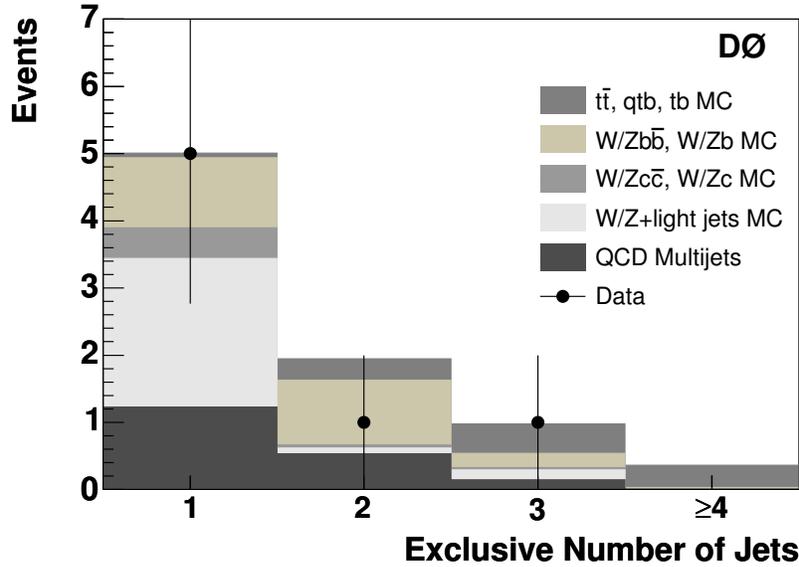
Figure 9.7: Exclusive jet multiplicity of events with at least one jet that has been tagged with both the SVT and SLT algorithms. The fourth bin represents the integral of four or more jets.

model (see Figs. 9.2-9.7). Using a W boson plus doubly-tagged jet sample, we set a 95% CL limit on the rate of anomalous production as a function of the number of jets in the events in which at least one jet is tagged with a simultaneous SLT and SVT tagging algorithm (see Table 9.5). Interpreting these results as anomalous SM production of $Wb\bar{b}$-like events, single-top-like events, and $t\bar{t}$-like events, we are able to set upper cross section limits of 26.2 pb, 15.6 pb, and 14.7 pb, respectively for such anomalous HF-quark production (see Table 9.7).

Finally, we made a comparison with the anomalous W boson plus HF-quark production result published by the CDF collaboration. Using the upper cross-section limits derived using the DØ data, we exclude the CDF result at greater than 99% C.L. in the W+2,3 jet bins, and in the $W+ \geq 2$ jet bins in single-top-like and $t\bar{t}$-like scenarios, respectively. We find we cannot exclude the CDF result in the W+1,2 jet bins for a $Wb\bar{b}$-like scenario, as these combined bins do not constitute a significant departure from the SM predictions. Based on these comparisons, and the further considerations discussed in Sec. 9.3.2, we conclude that the CDF result does not represent sufficient evidence to indicate a non-SM source of physics in W boson plus HF-quark production at the Fermilab Tevatron $p\bar{p}$ Collider.

Source	$W + 1, 2$ jets	$W + 2, 3$ jets	$W + \geq 2$ jets
Data observation	6	2	2
SM prediction	6.9±1.2	2.9±0.5	3.3±0.5
95% C.L. Limit (events)	6.58	4.51	4.41
Model			
Wbb-like	26.2 pb	–	–
Single-top-like	–	16.7 pb	15.6 pb
$t\bar{t}$-like	–	23.7 pb	14.7 pb

Table 9.7: 95% C.L. limits for the number of events summed over the indicated jet bins. Also shown are cross-section limits based on the hypotheses of $Wb\bar{b}$-like, single-top-like, and $t\bar{t}$-like anomalous production for the selected numbers of jets.

Source	$Wb\bar{b}$-like (1,2 jets)	Single-top-like (2,3 jets)	$t\bar{t}$-like (≥ 2 jets)
CDF Anomalous	4.3	20.7	15.6
DØ 95% C.L. Limit	26.2	16.7	14.7
Exclusion C.L.	–	2.2×10^{-2}	4.0×10^{-2}

Table 9.8: Comparison of CDF and DØ results for a W plus doubly-tagged jets selection. The three rows correspond to the CDF anomalous cross section, the DØ 95% C.L. upper cross section limit, and the C.L. at which the CDF result can be excluded.

Appendix A

Muon Trigger Enhancement in SLT Events

In this analysis, W boson selection begins with requiring a single-lepton trigger to fire, as discussed in Sec. 7.1. The efficiency for each trigger ultimately determines the maximum W-boson selection efficiency. The analysis further vetos events containing more than one isolated lepton, which helps to eliminate certain backgrounds, but also helps to reduce the complication of trigger overlaps. However, in the instances in which a jet is b-tagged using the SLT algorithm, an additional (non-isolated) muon is present in the event. This muon will not impact the $W \to e\nu$ triggering efficiency, as these events are selected by requiring a single-electron trigger and the muon bias in these triggers is negligible. However, there is a significant probability that such a muon can impact the $W \to \mu\nu$ triggering efficiency.

A.1   Single-Muon Trigger Enhancement

The dominant single-muon triggers used in this analysis are MUW_W_L2M3_TRK10 and MU_W_L2M5_TRK10, with the following trigger terms

- **MUW_W_L2M3_TRK10**

  - Level 1: mu1ptxwtlx (wide region tight scintillator trigger with loose requirements on hits in the wire chambers)
  - Level 2: 1 medium muon, identified at L2, with $p_T > 3$ GeV/c
  - Level 3: One L3 track, with $p_T > 10$ GeV/c

- **MU_W_L2M5_TRK10**

  - Level 1: mu1ptxwtxx_fz (wide region tight scintillator trigger with fast Z coincidence)
  - Level 2: 1 medium muon, identified at L2, with $p_T > 5$ GeV/c
  - Level 3: One L3 track, with $p_T > 10$ GeV/c

The muons resulting from the semi-leptonic decay of a B meson generally has a small p_T, as shown in Figure A.1. Therefore, the L3 requirement will generally be difficult to satisfy with these muons. However, the presents of one or more jets and another high-p_T muon will serve to augment this efficiency. Thus, the largest increase in single-muon triggers will occur at L1 and L2.
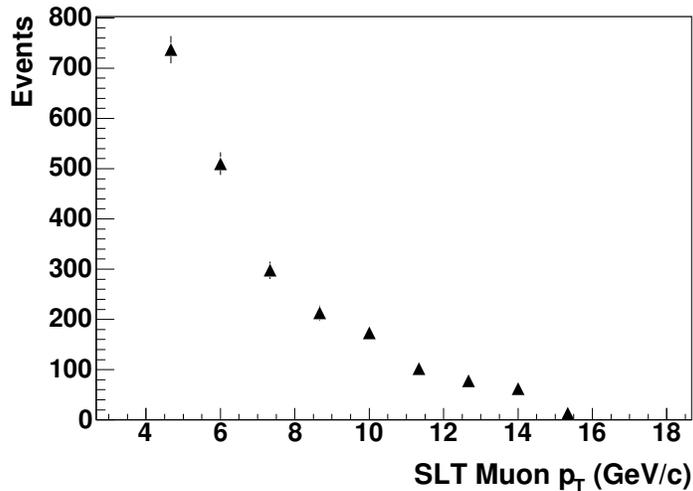
Figure A.1: Transverse momentum for SLT muons in data events.

To evaluate the impact of this effect, we use data events which have been selected using a single-electron trigger. The single-muon trigger efficiency for SLT muons will be measured by counting the frequency in which these muons satisfy the selected single-muon triggers in $W \rightarrow e\nu$ plus SLT events. To model most closely the $W \rightarrow \mu\nu$ plus SLT final state, we perform a loose $W \rightarrow e\nu$ selection with a soft-muon b-tag requirement. The $W \rightarrow e\nu$ selection performed is identical to that described in Sec. 8.1, with the following changes

- $p_T > 15$ GeV/c

- Corrected $\not{E}_T > 15$ GeV

- $30 < M_{W_T} < 150$ GeV

These changes are performed to increase the statistics of the test sample and should not introduce significant systematic errors. The SLT b-tagging requirement is described in Sec. 8.2.2. The shape of the jet distributions does not differ between the $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ selections, and the requirement of a matched-track for the electron will provide a similar L3 enhancement as in $W \rightarrow \mu\nu$ events. With these considerations, we do not associate a systematic uncertainty with this method of measuring the single-muon trigger enhancement due to SLT muons.

The first step in this evaluation is to measure the single-muon trigger efficiency in $W \rightarrow e\nu$ plus jets events in which no SLT is observed. This residual single-muon trigger efficiency is determined by counting the number of events selected as described above, with a veto on SLT b-tagged events. The fraction of these events satisfying a single-muon trigger is kept

as the residual efficiency. Each trigger was considered only during the runs in which it was active, as specified in Sec. 7.1. These results can be found in Table A.1.

Trigger	Total Evts	Passed Trigger	Efficiency
MUW_W_L2M3_TRK10	437327	515	0.12±0.005%
MU_W_L2M5_TRK10	98789	35	0.04±0.006%

Table A.1: Numbers of selected events for a loose $W \to e\nu$ + jets requirement and the number of events which also fired the specified trigger. All events which contain a SLT b-tagged jet are vetoed.

Next, the number of events which contain a SLT b-tagged jet is evaluated. The results of the measurement can be seen in Table A.2. The total number of events indicates the number of $W \to e\nu$ plus SLT events. The number of these events passing the specified trigger is evaluated and listed in the table. The errors listed are statistical errors.

Trigger	Total Evts	Passed Trigger	Efficiency
MUW_W_L2M3_TRK10	1631	317	19.4±1.2%
MU_W_L2M5_TRK10	378	27	7.1±1.4%

Table A.2: Numbers of selected events for a loose $W \to e\nu$ + SLT requirement and the number of events which also fired the specified trigger.

Given a single-muon trigger efficiency of P_1 and a SLT muon trigger efficiency of P_2, the total trigger efficiency for $W \to \mu\nu$ plus SLT events can be evaluated as

$$P_{tot} = 1 - ((1 - P_1)(1 - P_2)) \tag{A.1}$$

where the SLT muon trigger efficiency, P_2, is calculated by subtracting the residual single-muon trigger efficiency measured in $W \to e\nu$ plus jets events with a SLT veto from the efficiency found for events requiring a SLT b-tag.

$$P_2 = P_{SLT} - P_{noSLT} \tag{A.2}$$

Trigger	P_1	P_2	P_{tot}
MUW_W_L2M3_TRK10	63.0 ± 4.2%	19.3 ± 1.9%	70.1 ± 4.2%
MU_W_L2M5_TRK10	58.3 ± 3.3%	7.1 ± 2.1%	61.3 ± 4.1%

Table A.3: Combined selection efficiencies for single-muon triggers as given by Eq. A.1.

and the results are shown in Table A.3. Thus, the luminosity-weighted single-muon trigger efficiency for $W \to \mu\nu$ events rises from $62.1 \pm 3.4\%$ to $68.4 \pm 3.5\%$ in SLT b-tagged events.

Appendix B

W-Boson Selection in Exclusive Jet Distributions

Although good agreement was observed for the W-boson selection variables in Sec. 8.1, it is important for the transverse W-boson mass distributions to agree after requiring jets in the events. The distributions shown in Sec. 8.1 were dominated by events with no jets and were thus normalized to the large-statistics PYTHIA W/Z+jets MC samples. Here we will present these distributions for each exclusive jet selection, as described in Sec. 8.2.1. At this point, we will normalize the selections to the ALPGEN MC samples listed in Sec. 7.2.

B.1 $W \to e\nu$ Selections Variables in Exclusive Jet Distributions

The distributions for $W \to e\nu+$ 1, 2, 3, and ≥ 4 jets are shown in Figs. B.1- B.2. The distributions show a slow decrease in $\not{E}_T$ resolution with increasing number of jets, which leads to a broader M_{W_T} distribution. This feature fortifies the need for the upper cutoff on the transverse mass distribution. The agreement between data the MC predictions is good within the limits of statistics.

B.2 $W \to \mu\nu$ Selections Variables in Exclusive Jet Distributions

The distributions for $W \to \mu\nu+$ 1, 2, 3, and ≥ 4 jets are shown in Figs. B.3- B.4. The distributions follow the general trend seen in the $W \to e\nu$ decays, with a noticably lower QCD multijet contribution. Again, the agreement between data and MC predictions is good within statistics.
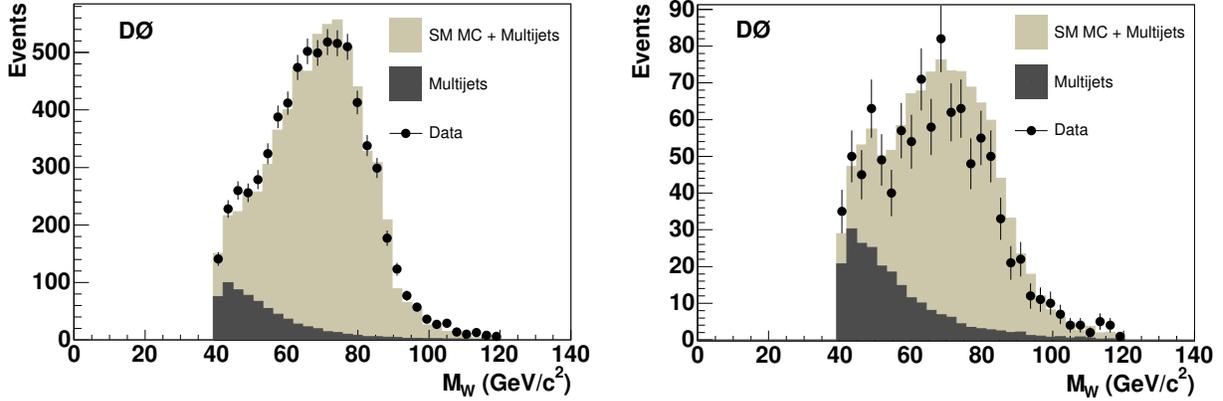
Figure B.1: Transverse W-boson mass distributions for events with exactly one (left) and two (right) jets with $p_T > 25\text{GeV/c}$ and $|\eta| < 2.5$. These events are selected $W \to e\nu$ decays.
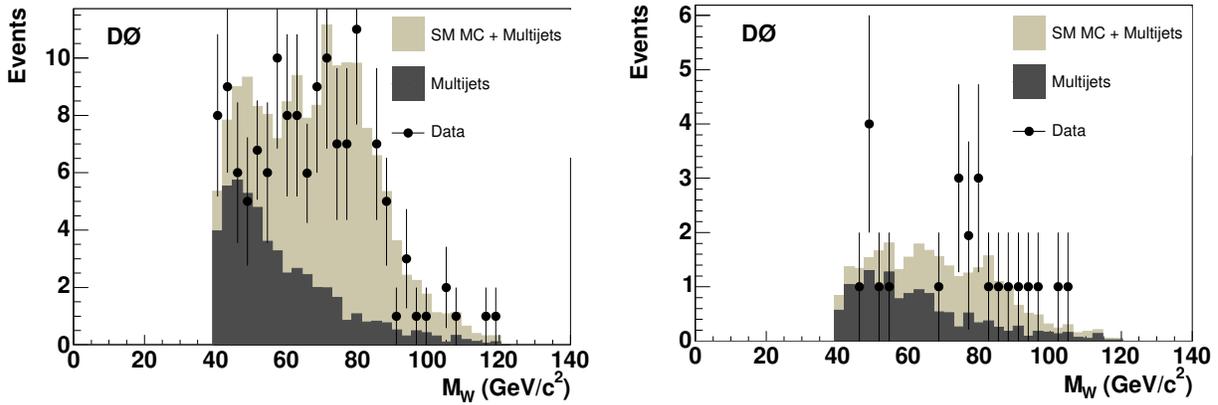
Figure B.2: Transverse W-boson mass distributions for events with exactly three (left) and four (right) jets with $p_T > 25\text{GeV/c}$ and $|\eta| < 2.5$. These events are selected $W \to e\nu$ decays.
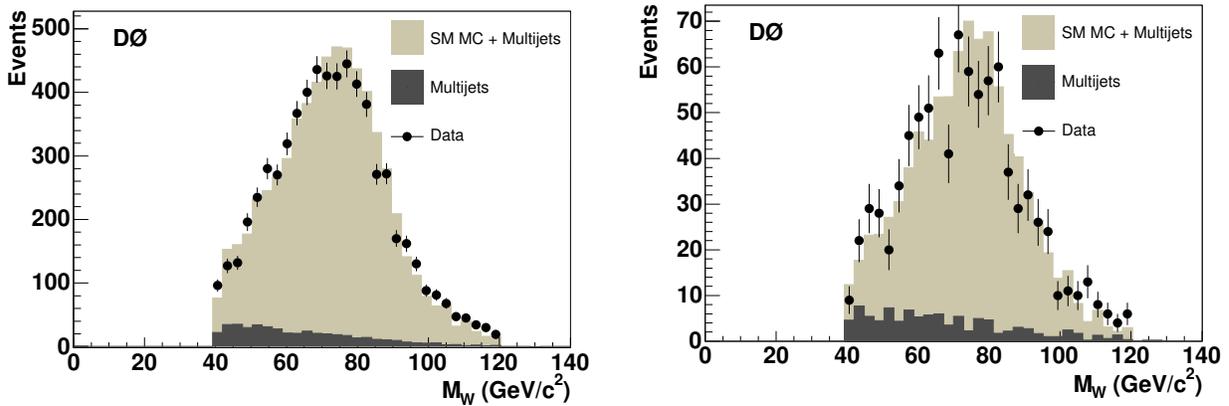
Figure B.3: Transverse W-boson mass distributions for events with exactly one (left) and two (right) jets with $p_T > 25\text{GeV/c}$ and $|\eta| < 2.5$. These events are selected $W \to \mu\nu$ decays.
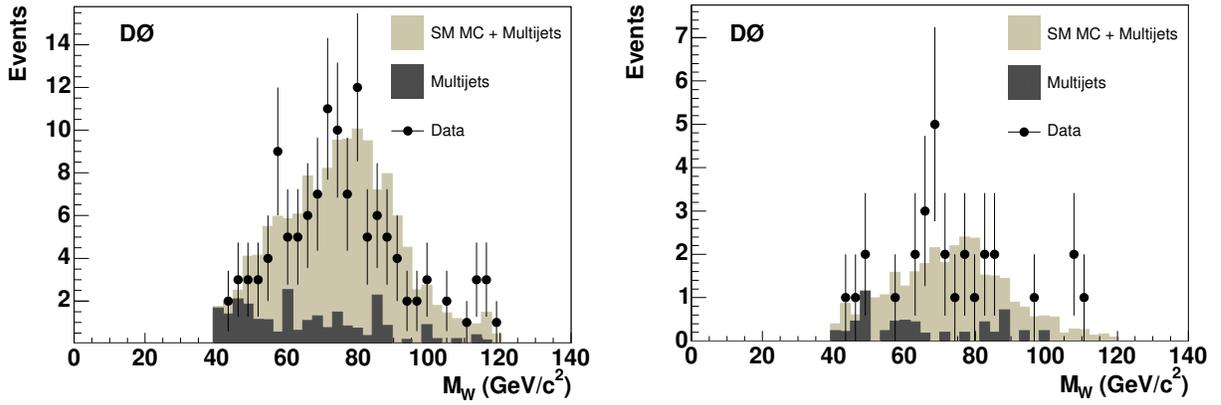
Figure B.4: Transverse W-boson mass distributions for events with exactly three (left) and four (right) jets with $p_T > 25\text{GeV/c}$ and $|\eta| < 2.5$. These events are selected $W \rightarrow \mu\nu$ decays.

Appendix C

Correlations between Tagging Algorithms

The heavy-flavor quark content of events can be enhanced by applying an HF-tagging algorithm to the jets in the event. Even after this requirement, a significant number events containing only light-flavor quarks can contaminate the sample. Requiring the simultaneous overlap of two or more tagging algorithms can significantly increase the purity of such a selection, although at the cost of efficiency. It is important to understand the correlations between the two tagging algorithms when requiring such an overlap. In this Appendix, we evaluate the correlations between secondary-vertex tagging (SVT) and soft-muon (SLT) tagging. We define jets which have been tagged by both of these two algorithms as doubly-tagged jets.

Because SLT and SVT algorithms tag heavy-flavor quarks decaying into different (but not orthogonal) ways, we expect the efficiency for each tagger to depend on different properties of the jets. The SVT algorithm relies on charged tracks from the decay of heavy-flavor quarks. The number of these tracks is larger for fully hadronic HF-decays than for semi-leptonic decays. However, the SLT algorithm intrinsically requires a muon present in the jet (ideally, from $b \rightarrow \mu\nu c$ or $b \rightarrow c \rightarrow \mu\nu q$ decays). This difference should cause the SVT efficiency to be lower for jets that are SLT tagged (i.e., semi-leptonic HF decays) and can be observed as a negative correlation between the tagging algorithms.

To study the performance of the SVT and SLT algorithms, we first measure the efficiencies per jet as a function of p_T. We perform this study on two groups of MC samples:

- **Top-like:** the first group are events that contain high p_T b-quarks from decays of heavy particles. We include the following event types: $t\bar{t}$, qtb (W-g fusion), and tb (s-channel).

- $Wb\bar{b}$-**like:** the second group are events in which the b-quark p_T is lower than in the first group. Here, we include the following event types: $Wb\bar{b}$, $Zb\bar{b}$, and Zb.

We ensure that any muons used for the SLT algorithms arise from HF-quark decays and not decays of primary W or Z bosons by matching with generator partons.

To measure the tagging effiencies, we match jets in each event to b-quark partons using generator information. Next, the tagging algorithms are applied to each jet. The p_T is recorded for each sample, before and after applying the tagging algorithms. The tagging

effiency is then given by the ratio of the two p_T distributions:

$$\epsilon_{tag} = \frac{\text{Tagged Jets Matched to a } b\text{-parton}}{\text{Total Jets Matched to a } b\text{-parton}} \tag{C.1}$$

In the following, this efficiency is measured using good jets (described in Sec. 4.6) with $p_T > 20$ GeV/c, good PV selection (described in Sec. 4.2), and no additional event selection. The measured SVT efficiencies include the inefficiency associated with jet taggability requirements (see Sec. 4.6).

C.1 HF-tagging Efficiencies for Inclusive b Decays

Figures C.1 and C.2 show the SVT, SLT, and doubly-tagged jet efficiencies, for top-like and $Wb\bar{b}$-like inclusive b-quark decays, respectively. To measure the correlation between the taggers, we divide the doubly-tagged efficiencies by the product of SVT and SLT efficiencies. These distributions are shown in Figs. C.3 and C.4, for top-like and $Wb\bar{b}$-like samples, respectively. In the event of no correlations, we expect this distribution to be equal to unity for all p_T values. We fit the normalized efficiencies to a straight line:

$$f(p_T) = a_0 + a_1 \times p_T \tag{C.2}$$

Table C.1 shows the fit parameters for the two event samples. A negative correlation is observed in these samples, indicating that the doubly-tagged efficiency is lower than expected from the SVT and SLT efficiencies alone.

Sample	a_0	a_1
Top-like	0.931 ± 0.034	-0.0001 ± 0.0004
$Wb\bar{b}$-like	0.801 ± 0.048	0.0016 ± 0.0007

Table C.1: Parameters for a linear fit to the normalized doubly-tagged efficiency for inclusive b-decays.

C.2 HF-tagging Efficiencies for Semi-Leptonic b Decays

Figures C.5 amd C.6 show the SVT, SLT, and doubly-tagged jet efficiencies for the top-like and $Wb\bar{b}$-like samples for semi-leptonic b-quark decays. We determine the correlation in the same way as in the previous section. The normalized doubly-tagged efficiencies are shown in Figs. C.7 and C.8 for top-like and $Wb\bar{b}$-like samples respectively. The results for a linear fit are given in Table C.2. As expected, limiting the measurement of tagging efficiencies to jets matched to semi-leptonic decays of b-quarks produces a correlation much closer to unity.
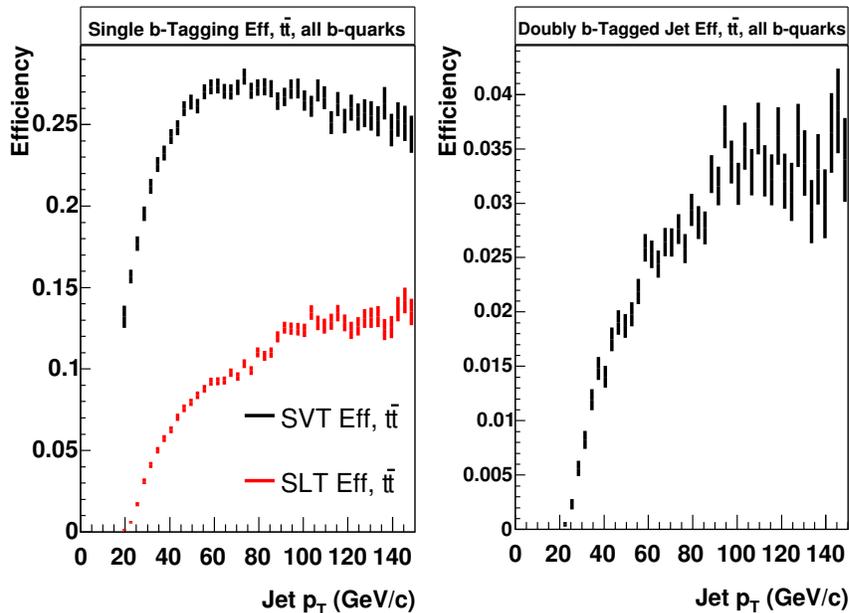
Figure C.1: Tagging efficiencies for SVT, SLT, and doubly-tagged algorithms in the top-like sample (denoted $t\bar{t}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for inclusive decays of b-quarks.

Sample	a_0	a_1
Top-like	1.008 ± 0.057	0.0003 ± 0.0007
$Wb\bar{b}$-like	0.988 ± 0.075	0.0004 ± 0.0013

Table C.2: Parameters for a linear fit to the normalized doubly-tagged efficiency for semi-leptonic b-decays.

C.3 Drop in SVT Efficiency for Semi-leptonic b Decays

We have observed a negative correlation for the overlap of SVT and SLT algorithms in jets not constrained to arise from semi-leptonic b-quark decays. Provided this negative correlation can be explained by a decrease in the number of tracks produced by b decays, we should observe a corresponding decrease in SVT efficiency for the two samples. At the same time, we should be able to measure a decrease in the number of tracks per secondary vertex in MC and data.

Figures C.9 and C.10 show the decrease in SVT efficiency when restricted to semi-leptonic b decays as a function of p_T for top-like and $Wb\bar{b}$-like samples, respectively. Also shown is the ratio of efficiencies as a function of p_T. Table C.3 shows the results of a linear fit to this ratio. The values we find are very close to those observed in the negative correlation of SVT and SLT efficiencies in inclusive b decays, indicating that this decrease in SVT efficiency is the source for the negative correlation.
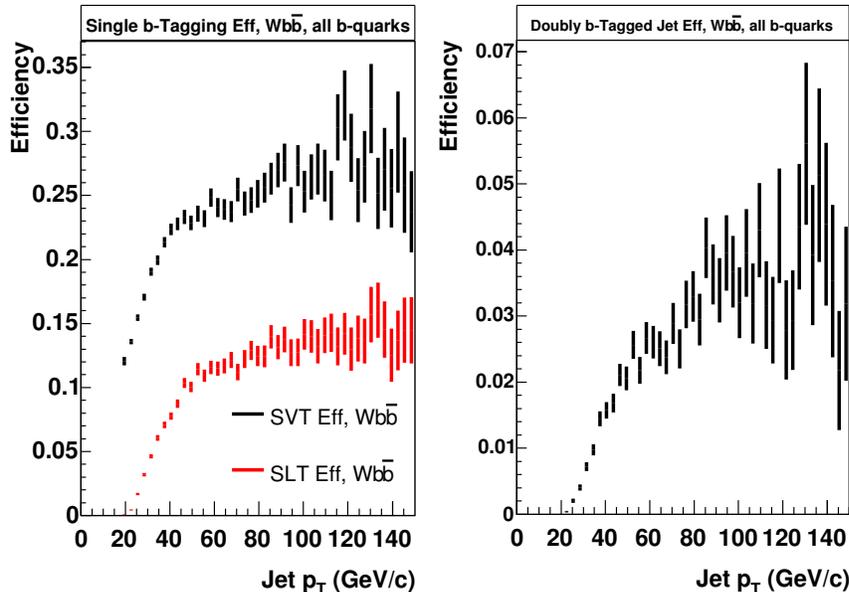
Figure C.2: Tagging efficiencies for SVT, SLT, and doubly-tagged algorithms in the $W b\bar{b}$-like sample (denoted $W b\bar{b}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for inclusive decays of b-quarks.

Figure C.11 shows the difference in the number of tracks per jet between jets that have been tagged with the SVT algorithm and the SLT algorithm for the top-like and $W b\bar{b}$-like MC samples. The same distribution is shown for data in Fig. C.12. In both plots, there is a marked decrease in the number of tracks per jet for semi-leptonic b-decays. The decrease in the mean value of the distributions as a percentage of the SVT mean value is 32.3%, 24.7%, and 27.2% for $W b\bar{b}$, $t\bar{t}$, and data, respectively. The number of tracks in jets tagged with the SVT algorithm starts at two, as that is part of the requirement for forming secondary vertices (see Sec. 4.6).

Sample	a_0	a_1
Top-like	0.926 ± 0.027	-0.0001 ± 0.0003
$W b\bar{b}$-like	0.846 ± 0.032	0.0007 ± 0.0006

Table C.3: Parameters for a linear fit to the decrease in SVT efficiency for semi-leptonic b decays.

C.4 Conclusions on SVT/SLT Correlations

In this Appendix, we studied efficiencies for secondary-vertex (SVT) and soft-muon (SLT) tagging. We have seen that there is a negative correlation when a jet is required to be tagged simultaneously with both algorithms. Futhermore, we showed that nearly all of this
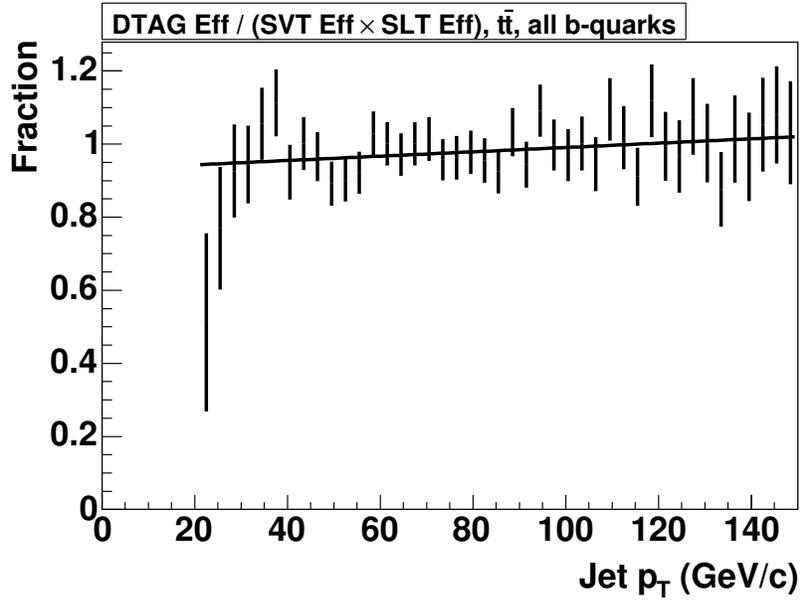
Figure C.3: Doubly-tagged jet efficiency normalized by SVT and SLT efficiencies in the top-like sample (denoted $t\bar{t}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for inclusive decays of b-quarks.

correlation can be attributed to a decrease in the SVT efficiency for jets that have a SLT, and that this decrease in SVT efficiency can be accounted for by a decrease in the number of tracks in jets that results from semi-leptonic decays of heavy-flavor quarks relative to fully hadronic heavy-flavor quark decays.
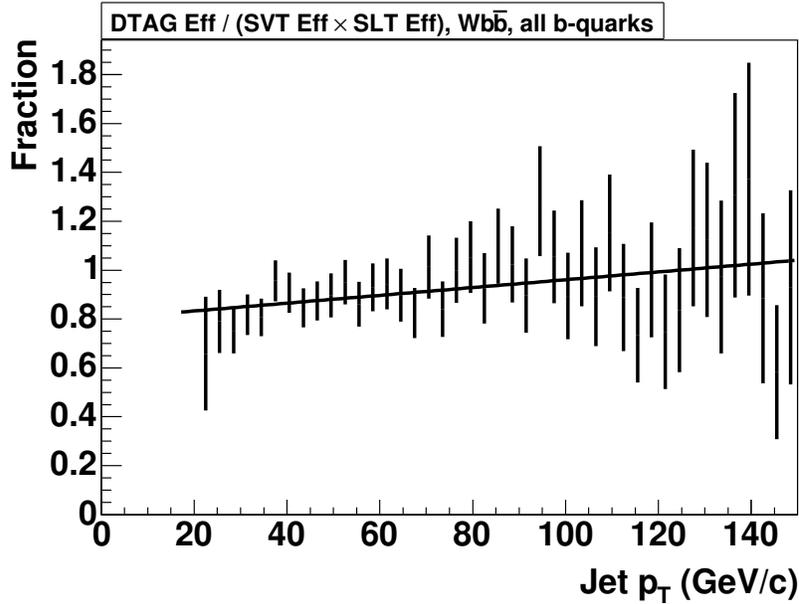
Figure C.4: Doubly-tagged jet efficiency normalized by SVT and SLT efficiencies in the $W b\bar{b}$-like sample (denoted $W b\bar{b}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for inclusive decays of b-quarks.
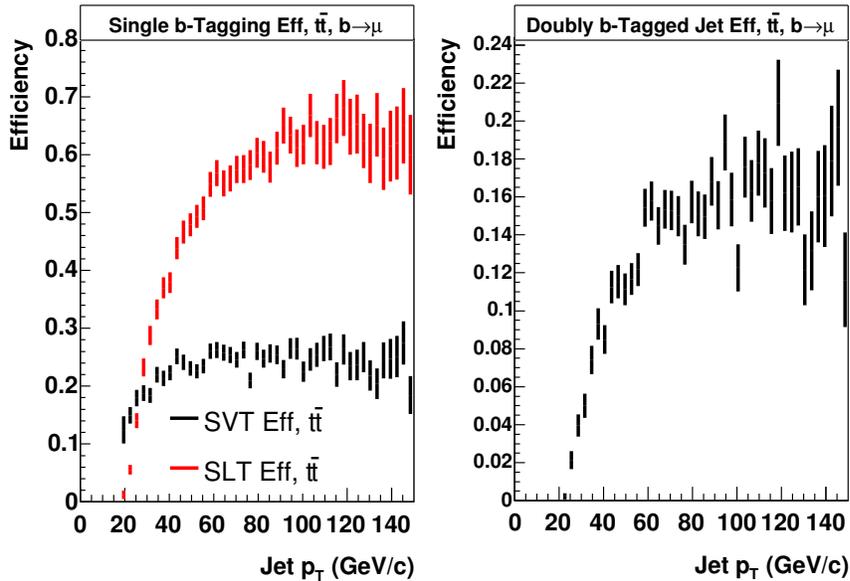
Figure C.5: Tagging efficiencies for SVT, SLT, and doubly-tagged algorithms in the top-like sample (denoted $t\bar{t}$ in the plot). The efficiency is derived for semi-leptonic decays of b-quarks.
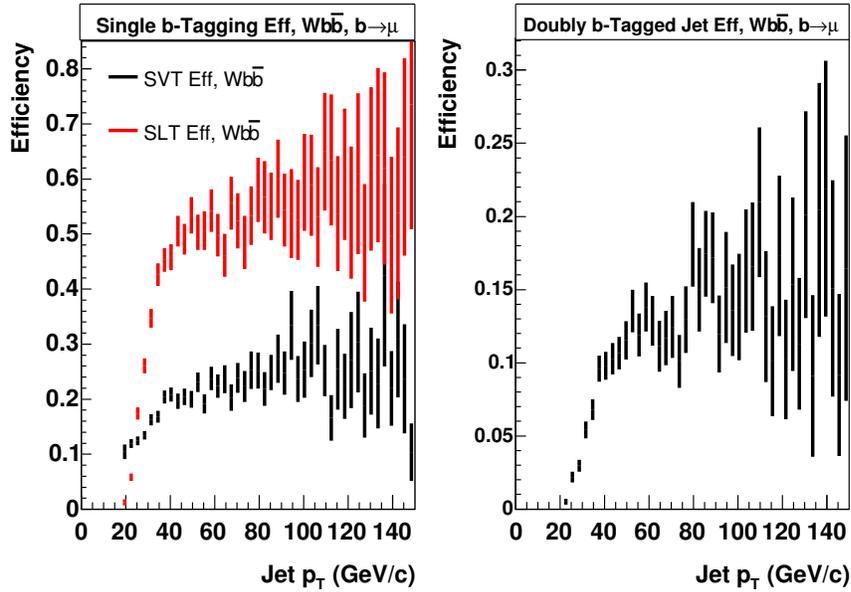
Figure C.6: Tagging efficiencies for SVT, SLT, and doubly-tagged algorithms in the $W b \bar{b}$-like sample (denoted $W b \bar{b}$ in the plot). The efficiency is derived for semi-leptonic decays of b-quarks.
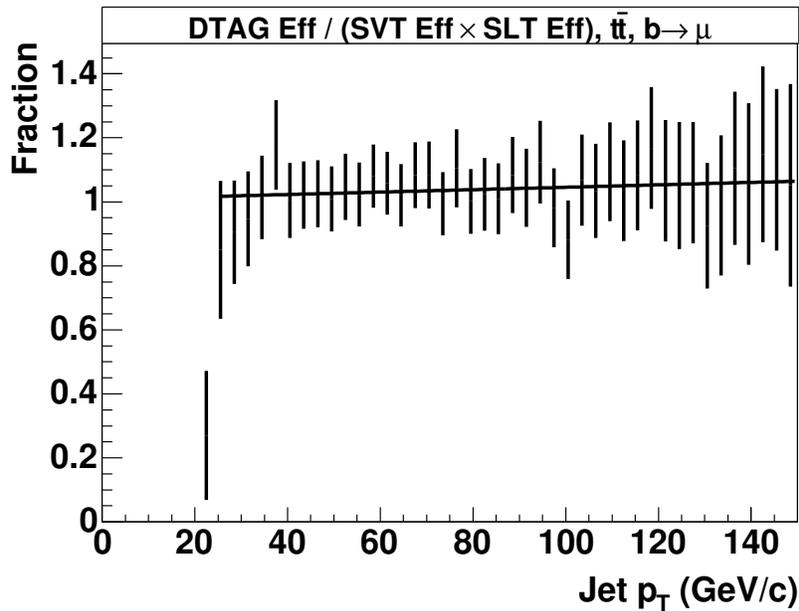
Figure C.7: Doubly-tagged jet efficiency normalized by SVT and SLT efficiencies in the top-like sample (denoted $t \bar{t}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for semi-leptonic decays of b-quarks.
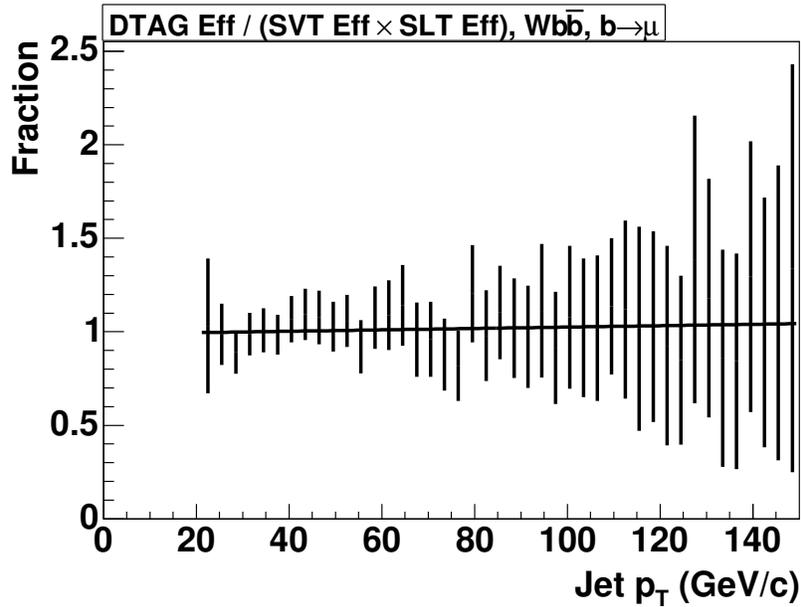
Figure C.8: Doubly-tagged jet efficiency normalized by SVT and SLT efficiencies in the $W b\bar{b}$-like sample (denoted $Wb\bar{b}$ in the plot) as a function of tagged jet p_T. The efficiency is derived for semi-leptonic decays of b-quarks.
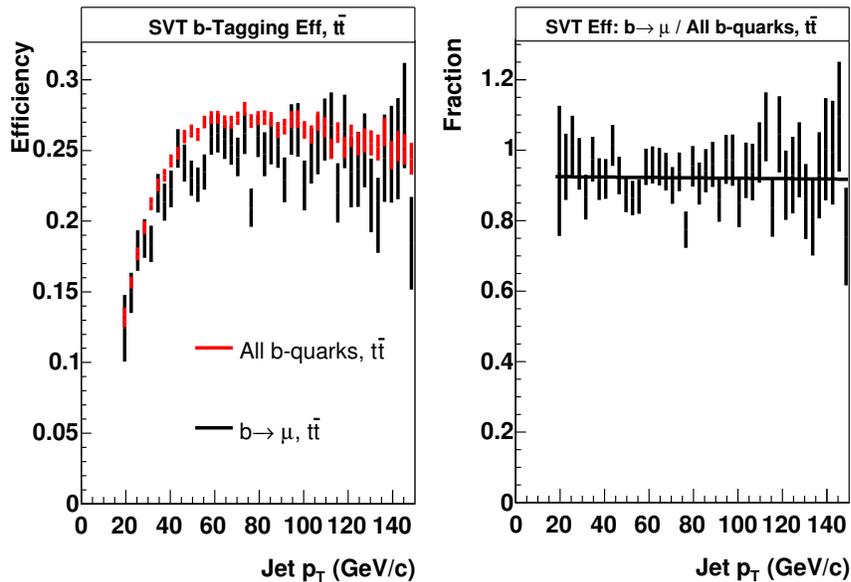
Figure C.9: Left: SVT efficiency for inclusive b decays and semi-leptonic b decays as a function of p_T. Right: The ratio of SVT efficiencies for these two scenarios. These figures correspond to the top-like event sample.
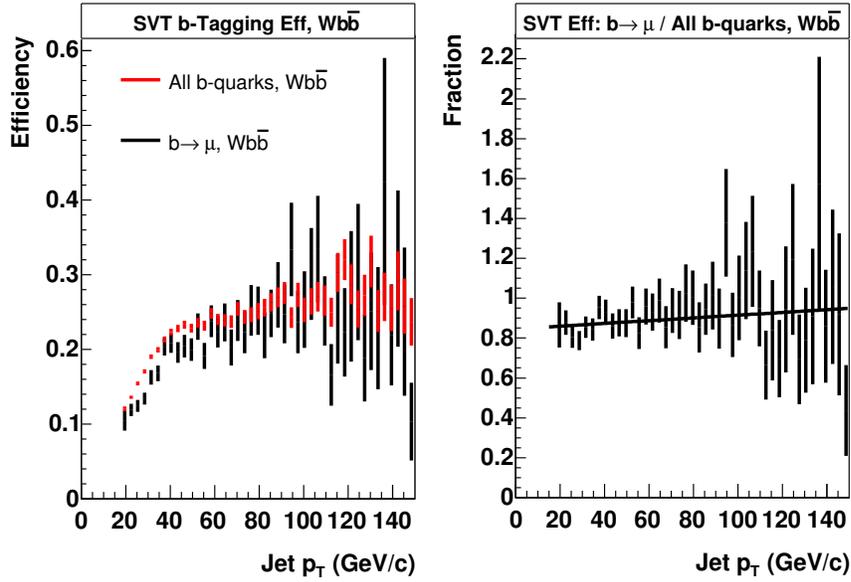
Figure C.10: Left: SVT efficiency for inclusive b decays and semi-leptonic b decays as a function of p_T. Right: The ratio of SVT efficiencies for these two scenarios. These figures correspond to the $W b \bar{b}$-like event sample.
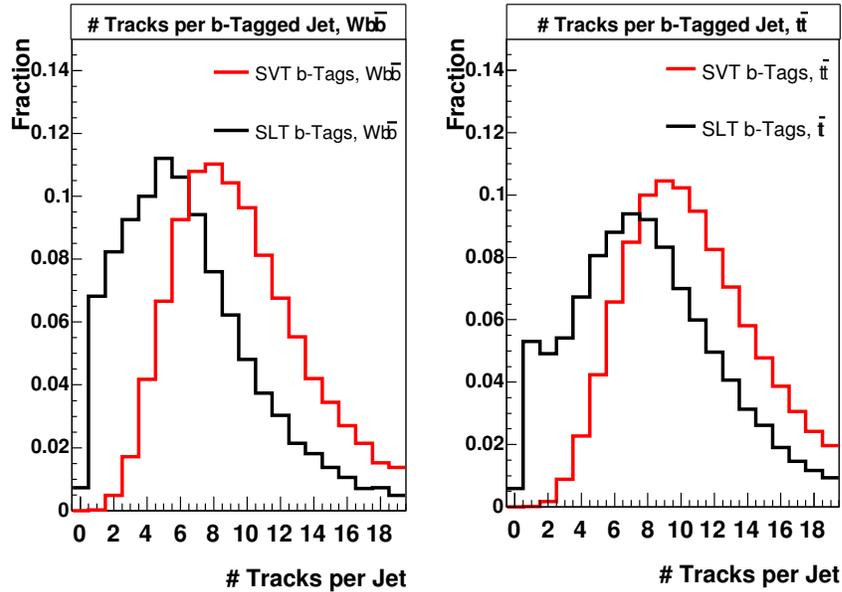
Figure C.11: Left: Number of tracks per tagged jet in the $W b \bar{b}$-like MC sample. Right: Number of tracks per tagged jet in the top-like MC sample. Plots are shown for SVT and SLT tagged jets separately.
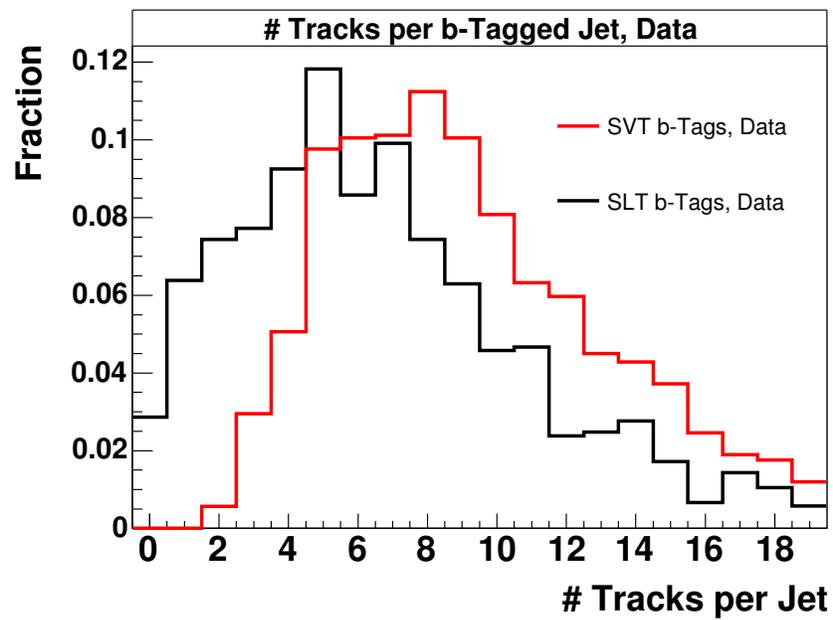
Figure C.12: Number of tracks per tagged jet in data for SVT and SLT tagged jets.

Bibliography

[1] Sheldon L. Glashow. Nuclear Physics, **22** 579 (1961).

[2] A. Salam. *Elementary Particle Theory*, ed. N. Svarthold (Almqvist and Wiksells, Stockholm, 1969), p.367.

[3] Steven Weinberg. Physics Review Letters, **19** 1264 (1967).

[4] M. Kaku, *Quantum Field Theory*, Oxford University Press, 1993.

[5] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, New York, 1995.

[6] Direct communication with Joel K. Erickson, August 17 2004.

[7] S. Dawson, arXiv:hep-ph/9612229.

[8] H. Davoidiasl, *et al*, arXiv:hep-ph/0405097, 2004.

[9] http://www-bd.fnal.gov/runII/index.html

[10] A. Haas, Ph. D. Thesis, University of Washington, 2004 (unpublished).

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

[12] S. Abachi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).

[13] DØ Collaboration, F. Last, Nucl. Instrum. Methods A **123**, 456 (2004).

[14] DØ Collaboration. "The DØ Upgrade: The Detector and Its Physics." July 30, 1996.

[15] DØ Collaboration, "DØ Silicon Tracker Technical Design Report", DØ Note 2169, 1997 (unpublished).
http://d0server1.fnal.gov/projects/silicon/www/tdr_final.ps

[16] DØ Collaboration, "The DØ Upgrade Central FiberTracker: Technical Design Report", DØ Note 4164, 1997 (unpublished).
http://d0server1.fnal.gov/users/stefan/www/CFT_TDR/CFT_TDR.ps

[17] M. Adams *et al.*, "Design Report of the Central Preshower Detector for the DØ Upgrade", DØ Note 3014, 1996.

[18] A. Goprdeev *et al.*, "Technical Design Report of the Forward Preshower Detector for the DØ Upgrade", DØ Note 3445, 1998 (unpublished).

[19] G. Hesketh, "Central Track Extrapolation Through the D0 Detector", DØ Note 4079 (2003).

[20] G. Hesketh, Ph. D. Thesis, University of Manchester, 2003 (unpublished).

[21] B. Abbot *et al.* (DØ Collaboration), Phys. Rev. D **58**, 4769 (1999).

[22] D. Acosta *et al*, "Study of the heavy flavor content of jets produced in association with W bosons in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV ", Phys. Rev. D **65**, 052007 (2002).

[23] D. Acosta *et al.* " Heavy flavor properties of jets produced in $p\bar{p}$ interactions at $sqrts = 1.8$ TeV ." Phys. Rev. D **69**, 072004 (2004).

[24] The CDF Collaboration, the DØ Collaboration, and the TEVATRON Electro-Weak Working Group. Combination of CDF and DØ Results on the Top-Quark Mass. arXiv:hep-ex/0404010.

[25] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963);
M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 552 (1973).

[26] M. Fischer, S. Groote, J.G. Korner, and M.C. Hauser, Phys. Rev. D **63** 031501 (2001).

[27] DØ Collaboration, arXiv:hep-ex/0406031.

[28] Phys. Lett. B **273** 167 (1991).

[29] A. Djouadi, J. Kalinowski, and M. Spira, Comp. Phys. Commun. C **108** 56 1998.

[30] V.M. Abazov *et al.*, "Measurement of the Cross Section for W Boson + Photon Production at $\sqrt{s} = 1.96$ TeV", submitted to PRL.

[31] Doug Chapin, *et al.* "Measurement of $Z \rightarrow e^+e^-$ and $W \rightarrow e^\pm\nu$ Production Cross Sections with $|\eta| < 1.05$", submitted to PRL.

[32] H. L. Lai *et al.* [CTEQ Collaboration], "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions,"Eur. Phys. J. C **12** (2000) 375 arXiv:hep-ph/9903282.

[33] ALPGEN, a generator for hard multiparton processes in hadronic collisions, M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, JHEP 0307:001,2003, arXiv:hep-ph/0206293.

[34] CompHEP, A.Pukhov, E.Boos, M.Dubinin, V.Edneral, V.Ilyin, D.Kovalenko, A.Kryukov, V.Savrin, S.Shichanin, and A.Semenov. Preprint INP MSU 98-41/542, arXiv:hep-ph/9908288.

[35] Pythia 6.154: T. Sjstrand, P. Edn, C. Friberg, L. Lnnblad, G. Miu, S. Mrenna and E. Norrbin, Computer Phys. Commun. 135 (2001) 238 (LU TP 00-30, arXiv:hep-ph/0010017).

[36] Nikolaos Kidonakis and Ramona Vogt, "Theoretical status of the top quark cross section", arXiv:hep-ph/0410367.

[37] Tim Tait and C.–P. Yuan, arXiv:hep-ph/9710372;
Tim M.P. Tait, arXiv:hep-ph/9909352.

[38] The Monte Carlo for Femtobarn Processes, http://mcfm.fnal.gov/

[39] Emanuela Barberis *et. al.*, "The Matrix Method and its Error Calculation", DØ Note 4564.

[40] J. Smith, W.L. van Neerven, and J.A.M. Vermaseren, Phys. Rev. Lett. **50**, 22 (1983).

[41] V. Buescher and J. Zhu, "em_cert: EM Certification Tools", DØ Note 4171.

[42] M. Hohlfeld, "Search for Associated Chargino and Neutralino Production in emu+l Final States in D0 data from Run II ", DØ Note 4543.

[43] DØ Collaboration, V. Abazov *et al.*, arXiv:hep-ex/0409040.

[44] DØ *b*-Tagging Group, http://www-d0.fnal.gov/phys_id/bid/d0_private/certification/p14-certification.html

[45] DØ Luminosity ID Group, http://www-d0.fnal.gov/phys_id/luminosity/

[46] M. Agelou *et. al.*, "Search for Single Top Quark Production at DZero in Run II", DØ Note 4510.

[47] V.M. Abazov *et al.*, "A Search for W$b\bar{b}$ and WH Production in pbarp Collisions at $\sqrt{s} = 1.96$ TeV", submitted to Phys. Rev. Lett., arXiv:hep-ex/0410062, Fermilab-Pub-04/288-E. PS;
S. Beauceron and G. Bernardi, "Evidence for W$b\bar{b}$ Production at $\sqrt{s} = 1.96$ TeV", DØ Note 4394.

[48] T. Junk, Nucl. Instr. Meth. Phys. Res. **A434** (1999) 435;
A. Read, in *1st Workshop on Confidence Limits*, CERN-2000-005.