© 2008 Ulysses Allen Grundler

A MEASUREMENT OF THE $t\bar{t}$ PRODUCTION CROSS SECTION IN $p\bar{p}$ COLLISIONS AT $\sqrt{s} = 1.96$ TEV USING SOFT MUON TAGGING

BY

ULYSSES ALLEN GRUNDLER

B.S., University of Illinois at Urbana-Champaign, 2002

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2008

Urbana, Illinois

Doctoral Committee:

Professor Kevin Pitts, Chair
Professor Tony Liss, Director of Research
Professor James Eckstein
Professor Scott Willenbrock

Abstract

I present a measurement of the $t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV using 2034 pb^{-1} of CDF Run II data using events with a high transverse momentum electron or muon, three or more jets, and missing transverse energy. The measurement assumes a $t \to Wb$ branching fraction of 100 percent. Events consistent with $t\bar{t}$ decay are found by identifying jets containing heavy-flavor semileptonic decays to muons. The dominant backgrounds are evaluated directly from the data. Based on 248 candidate events and an expected background of 86.8±5.6 events, I measure a production cross section of $8.7 \pm 1.1^{+0.9}_{-0.8} \pm 0.6$ pb, in agreement with the Standard Model.

To my wife and sons.

Acknowledgments

I would like to thank Tony Liss who has been my advisor since I was an undergraduate at the University of Illinois. During that time, I have learned a great deal from him about CDF, and physics in general. He has always made time to give me advice when I have run into a difficult problem.

Thanks to Anyes Taffard who gave me my first piece of code, and who was always willing to answer any questions I had. Lucio Cerrito, thank you for sharing your expertise, especially on the art of QCD-background estimation. I also appreciate Xiaojian Zhang providing me with efficiency measurements and making cross checks of my work. I would also like to thank my fellow graduate students, in particular Chris Marino and Alice Bridgeman, for their help in classwork and research, and for lending an ear when needed.

I would like to thank my family for their all of their support. They have shown interest in my work even when they did not understand what I was talking about. My parents, in particular, have given me much guidance and encouragement over the years. My sons, Xavier and Ezekiel, have been a source of pride, and have kept a smile on my face when I was tired and frustrated. My lovely wife, Deirdre, deserves special thanks for her love and support. She has given of herself so that I could complete this thesis.

Finally, I thank my Lord Jesus Christ for creating such an awesome world and giving us the ability to explore it. Beyond this, he has given me many wonderful gifts and I know that "I can do all things through Christ who strengthens me."

This work was supported in part by the United States Department of Energy under grant DE-FG02-91ER40677.

Table of Contents

List of Tables . . . . . **viii**

List of Figures . . . . . **x**

Chapter 1 The Standard Model and the Top Quark **1**
1.1 Introduction 1
1.2 The Standard Model 1
1.3 The Top Quark 3
1.3.1 Top Quark Production at the Tevatron 3
1.3.2 Top Quark Decay 5
1.4 The Interest in the $t\bar{t}$ Production Cross Section 6

Chapter 2 The Fermilab Tevatron and the CDF II Detector **8**
2.1 The Tevatron 8
2.1.1 Proton Production 8
2.1.2 Antiproton Production and Recycling 9
2.1.3 Collisions 9
2.2 The CDF II Detector 11
2.2.1 Tracking Systems 12
2.2.2 Calorimetry 15
2.2.3 Muon Detectors 16
2.2.4 Trigger System 19

Chapter 3 Monte Carlo Modeling **21**
3.1 $t\bar{t}$ Event Modeling 21
3.2 W+Jets Modeling 21
3.3 Other Processes 22
3.4 Detector Simulation 22

Chapter 4 Soft Muon Tagging (SLT) **24**
4.1 The SLT Algorithm 24
4.1.1 Taggable Tracks 24
4.1.2 A χ^2 Function for Muon Identification 25
4.2 Parametrization of the Muon Matching Distribution Widths 26
4.3 Efficiency of the Soft Muon Tagger 29

Chapter 5 $t\bar{t}$ Event Selection and Acceptance **33**
5.1 High-p_T Electrons 33
5.1.1 Electron Triggers 33
5.1.2 Offline Electron Selection 34
5.2 High-p_T Muons 35
5.2.1 Muon Triggers 35

5.2.2 Offline Muon Selection . . . 35
5.3 Jet Selection . . . 36
5.4 Missing Transverse Energy . . . 37
5.5 Event Selection . . . 37
5.5.1 Kinematic Selection . . . 37
5.5.2 SLT-Tag Requirements . . . 38
5.6 Data Sample . . . 39
5.7 $t\bar{t}$ Acceptance and Tagging Efficiency . . . 40
5.7.1 Geometric and Kinematic Acceptance . . . 40
5.7.2 SLT Efficiency for $t\bar{t}$ events . . . 42

Chapter 6 Mistag Matrix . . . 44
6.1 Mistag Probability . . . 44
6.1.1 Track-based definition . . . 44
6.1.2 Data samples . . . 44
6.2 Event Reconstruction . . . 45
6.2.1 Reconstruction of $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+, \mathrm{D}^0 \to K^-\pi^+$ . . . 45
6.2.2 Reconstruction of $\Lambda^0 \to p\pi$ . . . 47
6.3 Mistag Matrix . . . 47
6.3.1 Sideband Subtraction . . . 49
6.3.2 Decay-in-Flight Correction . . . 51
6.3.3 Finalizing the Matrix . . . 54
6.4 Prediction of the Mistag Matrix and the Systematic Uncertainty . . . 55
6.4.1 Tracks in a Jet . . . 55
6.4.2 Test Samples . . . 57
6.4.3 The Mistag Rate and Dijet Balancing . . . 58
6.4.4 Heavy Flavor . . . 59
6.4.5 Determination of the systematic uncertainty . . . 60
6.4.6 Double tagged jets: a cross check . . . 63

Chapter 7 Background Estimation . . . 66
7.1 Mistags . . . 66
7.2 W+Heavy Flavor . . . 67
7.2.1 Heavy-Flavor Fractions . . . 68
7.2.2 Heavy-Flavor Tagging Efficiency . . . 69
7.3 QCD . . . 70
7.3.1 The Fraction of QCD Events Before SLT Tagging . . . 71
7.3.2 The Tag Rate of QCD Events . . . 73
7.3.3 The QCD Background Estimate . . . 74
7.4 Drell-Yan$\to \mu\mu$ . . . 74
7.5 Monte Carlo Driven Backgrounds . . . 77
7.5.1 Diboson Backgrounds . . . 77
7.5.2 Drell-Yan$\to \tau\tau$ Background . . . 78
7.5.3 Single Top Background . . . 79

Chapter 8 Systematic Uncertainties . . . 80
8.1 Uncertainties on Acceptance and Efficiency . . . 80
8.2 Uncertainties on Background Predictions . . . 83

Chapter 9 Results . . . 85
9.1 Cross Section Calculation . . . 85
9.2 Observed Tags and Measured Cross Section . . . 86
9.3 Kinematic Features of the SLT-Tagged Sample . . . 89

Chapter 10 Conclusions **95**

Appendix A SLT Tagger Parametrization **98**
A.1 Stub-Track Matching Variables in CMU 98
A.1.1 CMU dx 98
A.1.2 CMU dz 99
A.1.3 CMU dϕ_L 99
A.2 Stub-Track Matching Variables in CMP 99
A.2.1 CMP dx 100
A.2.2 CMP dϕ_L 100
A.3 Stub-Track Matching Variables in CMX 100

Appendix B Measurement of the Mistag Rate Using γ+Jets Events **102**
B.1 Mistag Probability 102
B.2 Mistag Matrix 102
B.3 Prediction of the Mistag Matrix and the Systematic Uncertainty 104
B.3.1 Test Samples 104
B.3.2 Summary of the comparisons and the systematic uncertainty 104
B.4 Fraction of Tags from Heavy-Flavor decays in the γ+Jets Sample 107

Appendix C CMP Ageing Studies **109**
C.1 ^{90}Sr Measurements 109
C.1.1 Setup and Procedure 109
C.1.2 Results 110
C.2 *In Situ* Gain Related Measurement 112

References **116**

Author's Biography **119**

List of Tables

1.1 Charge and mass of Standard Model quarks and leptons. 2
1.2 Charge and mass of Standard Model gauge bosons. 3

3.1 Summary of the Monte Carlo samples used for this analysis. 23

4.1 Summary of basic track-quality cuts set in the SLT algorithm. 25
4.2 $W \to \mu\nu$ selection cuts used in the determination of the SLT tagger parametrization. 27

5.1 Summary of event counts for 2 fb^{-1} of CDF Run II data. 40
5.2 Summary of lepton trigger and identification efficiencies. 41
5.3 Acceptance for $t\bar{t}$ events as a function of jet multiplicity. 41
5.4 $t\bar{t}$ event tagging efficiency for SLT muons as a function of jet multiplicity. 42

6.1 Mass windows used in determining the status of a D* or Λ^0 candidate. 47
6.2 Relevant numbers in the determination of bias against DIF. 53
6.3 Weights given to each separate particle matrix when combining them into one final matrix. 55
6.4 Checks of the mistag matrix in different jet $E_T^{corr.}$ bins. 62
6.5 Checks of the mistag matrix with different jet samples. 62
6.6 Fraction of taggable jets in W+3-or-more jets events that fall within a given E_T range. 63
6.7 Checks of the mistag matrix using second SLT tags in different jet samples. 64
6.8 The ratio of opposite-sign SLT tags to same-sign SLT tags in double-tagged jets. . . 65

7.1 Summary of background estimate from mistags in W+jets events. 67
7.2 The heavy flavor fractions for W+heavy flavor background. 68
7.3 Relative systematic errors for the determination of the heavy-flavor fraction of W+jets events before tagging. 69
7.4 The *real* tagging efficiencies for the W+heavy flavor background. 70
7.5 The fractions, F_{QCD}, of lepton-plus-jets events due to QCD multijet processes. . . . 72
7.6 Summary of the QCD background estimate. 74
7.7 Drell-Yan background summary. 77
7.8 Summary of diboson backgrounds. 78
7.9 Summary of Drell-Yan$\to \tau^+\tau^-$ background. 78
7.10 Summary of single top background. 79

8.1 Summary of lepton identification systematics. 81
8.2 Summary of systematic uncertainties. 84

9.1 Number of tagged events and the background summary. 87
9.2 Summary of components of the denominator for the cross section calculation. 88
9.3 Summary of $t\bar{t}$ cross sections split by primary-lepton channel. 88

B.1 Checks of the photon matrix with different samples. 105
B.2 Relevant numbers for the study of the heavy-flavor fraction, F_{HF}, of the photon matrix. 108

List of Figures

1.1 Parton distribution functions of quarks and gluons in the proton. 4
1.2 Leading order (tree-level) diagram of $t\bar{t}$ production via $q\bar{q}$ annihilation. 5
1.3 Leading order diagrams of $t\bar{t}$ production via gg fusion. 5

2.1 The Fermilab accelerator chain. 9
2.2 Tevatron peak luminosity. 10
2.3 Total luminosity gathered by the CDF detector. 11
2.4 One half of the CDF II detector from an elevation view. 12
2.5 One quadrant of the CDF tracking volume and forward calorimetry. 13
2.6 Coverage of the CDF muon systems in azimuth, ϕ, and pseudorapidity, η. 17
2.7 Number of interaction lengths as a function of η, averaged over ϕ acceptance, of the CMU, CMP, and CMX systems. 18
2.8 A block diagram of the CDF II data flow. 20

4.1 Width of the matching distributions ($\mathrm{d}x$, $\mathrm{d}z \times \sin\theta$, $\mathrm{d}\phi_L$) vs. p_T for each muon subsystem. 28
4.2 Efficiency of the SLT tagger as measured from J/ψ decays, as a function of the χ^2 quantity, L. 30
4.3 The SLT efficiency for CMU/CMP. 31
4.4 The SLT efficiency for CMX. 31
4.5 The SLT efficiency as a function of the number of tracks. 32

5.1 H_T distributions, normalized to unity, for $t\bar{t}$ and W+jets Monte Carlo events. 38

6.1 The $m(K\pi\pi) - m(k\pi)$ distribution for $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+, \mathrm{D}^0 \to K^-\pi^+$ candidates. . . . 46
6.2 The $m(p\pi)$ distribution for $\Lambda^0 \to p\pi$ candidates. 48
6.3 The tag rate as a function of $m(K\pi\pi) - m(k\pi)$. 50
6.4 The $m(p\pi)$ distribution for $\Lambda^0 \to p\pi$ candidates. 51
6.5 The measured and sideband-subtracted tag rates as a function of track p_T. 52
6.6 The relative abundances of taggable tracks of various particle types as found in W+jets Monte Carlo. 54
6.7 The mistag rate as a function of track p_T and detector η. 56
6.8 The SLT tag rate versus the difference between a jet's corrected E_T and the trigger E_T. 59
6.9 Distribution of the maximum d_0 significance of tracks. 60
6.10 Fraction of taggable b or c jets out of all taggable jets passing versus the maximum allowed $|d_0$ significance$|$. 61
6.11 Fraction of taggable tracks, in Monte Carlo, with a DIF. 61
6.12 Summary of the SLT mistag matrix prediction. 62
6.13 Corrected-E_T distribution for SLT taggable jets in ALPGEN W+jets events. 63

7.1 A diagram illustrating the regions defined in the missing-E_T vs. isolation plane. . . . 71
7.2 The ratio of observed to predicted tags as a function of $\not{E}_t$. 73

7.3 Comparison between data and ALPGEN Monte Carlo dimuon invariant mass distributions. 76
7.4 The product of the electromagnetic fraction of the SLT jet and the charge product of the SLT and primary muon. 76

8.1 Feynman diagram of top production and decay with ISR and FSR. 82

9.1 The number of SLT-taggable tracks in events with 3 or more jets. 86
9.2 The expected background and observed tags in 1, 2, 3, and 4-or-more jets events. . . 87
9.3 Comparison of the jet E_T distributions for tagged jets and for expectations from mistags, W+heavy-flavor, QCD and $t\bar{t}$ events. 90
9.4 Comparison of the H_T distributions for tagged events and for expectations from mistags, W+heavy-flavor, QCD and $t\bar{t}$ events. 91
9.5 p_T of the SLT tags. 92
9.6 The impact parameter (d_0) significance for tagged tracks. 93
9.7 The p_T^{rel} distribution for tags in data. 94

10.1 Standard Model predictions of the $t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV as a function of the top quark mass. 97

B.1 The mistag rate as a function of track p_T, seen using the γ+jets sample. 103
B.2 The SLT mistag prediction in Z+jets events. 105
B.3 Summary of the photon matrix prediction. 106

C.1 Data from an ^{55}Fe run. 111
C.2 Relative gain vs. integrated charge data. 113
C.3 Relative gain vs. integrated charge data for the second run. 114
C.4 Average width of CMP hits on J/ψ stubs vs. run number. 114
C.5 Average width of CMP hits on J/ψ stubs vs. run number. 115

Chapter 1

The Standard Model and the Top Quark

1.1 Introduction

The goal of particle physics is the understanding of the nature of the fundamental constituents of the universe and their interactions in a quantitative manner. Much progress has been made over the last half-century in uncovering these constituents and the rules governing their behavior. Today, the so-called "Standard Model" is the most successful theory of particle physics. Since its development, this theory has successfully predicted the outcome of a large variety of experiments and the existence of a number of particles (including the top quark). Despite the success of the Standard Model, it cannot be the final description of particles and their interactions—besides theoretical motivations to believe that a new framework must come into play at approximately the TeV scale, the measured neutrino masses cannot be accounted for in the Standard Model. Increasingly precise measurements of the properties of particles and their interactions and direct searches for new phenomena is the pursuit of modern experimental particle physics. Measuring the properties of the top quark is a significant part of this project.

1.2 The Standard Model

The Standard Model of particle physics is a theory that describes the fundamental constituents of the universe and their interactions [1]. In this theory, all matter is composed of spin-$\frac{1}{2}$ fermions, known as "quarks" and "leptons." These fermions interact via a few fundamental forces (electromagnetic, weak, and strong[1]) mediated by spin-1 bosons.

Quarks and leptons are arranged into three generations; the particles in each generation have similar relationships, but are of differing masses. There are six "flavors" of quarks, which can be classified into two general types, "up"-type and "down"-type after the prototypes in the first gen-

[1] Gravity is extremely weak in comparison to the other forces and is not included in the Standard Model.

eration. The quarks have electromagnetic, weak, and strong (or color) charges and so are affected by all the forces. Due to the nature of the strong force, quarks exist only in bound states called "hadrons" and are not detected as free particles.[2] There are also six leptons, three (electromagnetically) charged leptons, of which the electron is the prototype, and three neutrinos. The neutrinos participate only in weak interactions while the charged leptons interact electromagnetically as well. The quarks and leptons, along with some of their basic properties, are listed in Table 1.1. For each particle, there is a corresponding antiparticle with identical mass but opposite charge.

Generation	Particle (Symbol)	Charge (e)	Mass (MeV/c^2)
First	Electron (e)	-1	0.511
	Electron neutrino (ν_e)	0	$\neq 0$
	Up quark (u)	$+2/3$	1.5–3
	Down quark (d)	$-1/3$	3–7
Second	Muon (μ)	-1	106
	Muon neutrino (ν_μ)	0	$\neq 0$
	Charm quark (c)	$+2/3$	1250±90
	Strange quark (s)	$-1/3$	95±25
Third	Tau (τ)	-1	1777
	Tau neutrino (ν_τ)	0	$\neq 0$
	Top quark (t)	$+2/3$	$(172.6 \pm 1.4) \times 10^3$
	Bottom quark (b)	$-1/3$	$(4.2 \pm 0.07) \times 10^3$

Table 1.1: Charge and mass of Standard Model quarks and leptons. The mass values shown here are taken from the PDG review of particle physics [2] except for the top quark mass which is taken from a combination of CDF and D0 results [3]. The neutrinos have been shown to have mass, though the flavor eigenstates are a mixture of the mass eigenstates.

Two quantum field theories, electroweak theory and quantum chromodynamics (QCD), describe the interactions between quarks and leptons through the exchange of gauge bosons. Electroweak theory unifies the electromagnetic force, which is mediated by the massless photon, and the weak force, mediated by the massive charged $W^\pm$ bosons and the neutral Z boson. The masses of the W and Z bosons arise through the Higgs mechanism, which spontaneously breaks the symmetry of the electroweak Lagrangian. Also arising out of this mechanism is the Higgs boson, the sole remaining Standard Model particle to be detected. Strong interactions are described by QCD. Eight massless gluons mediate the strong force and themselves carry the corresponding charge (color). The basic properties of the gauge bosons are summarized in Table 1.2.

[2]The top quark is an exception in that it decays so quickly that it does not have a chance to hadronize. So, the top is thought to decay as a free quark.

Force	Particle (symbol)	Charge	Mass (GeV/c^2)
Strong	Gluon (g)	0	0
Electromagnetic	Photon (γ)	0	0
Weak (charged)	W Boson ($W^{\pm}$)	± 1	80.403±0.029
Weak (neutral)	Z Boson (Z)	0	91.1876±0.0021

Table 1.2: Charge and mass of Standard Model gauge bosons. The mass values shown here are taken from the PDG review of particle physics [2].

1.3 The Top Quark

The top quark was discovered in 1995 by the CDF [4] and D0 [5] collaborations and is the heaviest known elementary particle. Weighing in at over 170 GeV/c^2, the top quark is approximately 40 times the mass of the next heaviest quark. The Higgs boson is thought to be responsible for the top quark mass, and the Yukawa coupling between the two is ~ 1, giving rise to the question of whether the top quark plays a role in the electroweak symmetry breaking mechanism. The mass of the top quark has also given rise to speculations that it offers a unique window to search for physics beyond the Standard Model. Along with other features, such as the top quark's short lifetime, this makes the top quark a very interesting particle to study.

1.3.1 Top Quark Production at the Tevatron

At the Tevatron, top quarks are produced predominately in pairs via the strong interaction. To calculate the production cross section for $p\bar{p} \to t\bar{t}$, it is necessary to take into consideration the structure of the incoming protons. Because protons are not elementary particles, but are made of quarks and gluons, the initial state of the actual interaction is complicated. However, if the momenta of the incoming particles are high enough ($\gg \Lambda_{\mathrm{QCD}} \approx 200$ MeV), it is possible to calculate the cross section using perturbative QCD, treating the interaction as one between only two elementary particles. Unfortunately, the initial momenta of the particles (called partons) cannot be determined on an event-by-event basis. Each parton carries a fraction (x) of the proton (or antiproton) momentum according to a statistical distribution that depends on its type (gluon or quark flavor) and on the energy scale; this distribution is known as a parton distribution function (PDF). Figure 1.1 shows the proton PDFs for an energy scale of 100 GeV. The valence quarks (u and d) are most likely to carry a large fraction of the proton momentum, though gluons can carry a significant fraction as well. To theoretically evaluate the cross section, one must sum over all possible interactions,

weighted by their probability according to the PDFs. Thus, the $t\bar{t}$ cross section is given by:

$$\sigma = \sum_{i,j} \int dx_i dx_j f_i(x_i, \mu^2) f_j(x_j, \mu^2) \hat{\sigma}(p_i p_j \to t\bar{t}), \tag{1.1}$$

where the sum is over all possible initial parton states, $p_i p_j$. $f_i(x_i, \mu^2)$ is the PDF of a parton of type i, given a momentum fraction, x_i, and an interaction energy scale, μ. The cross section for the individual parton-parton interactions is represented by $\hat{\sigma}(p_i p_j \to t\bar{t})$.

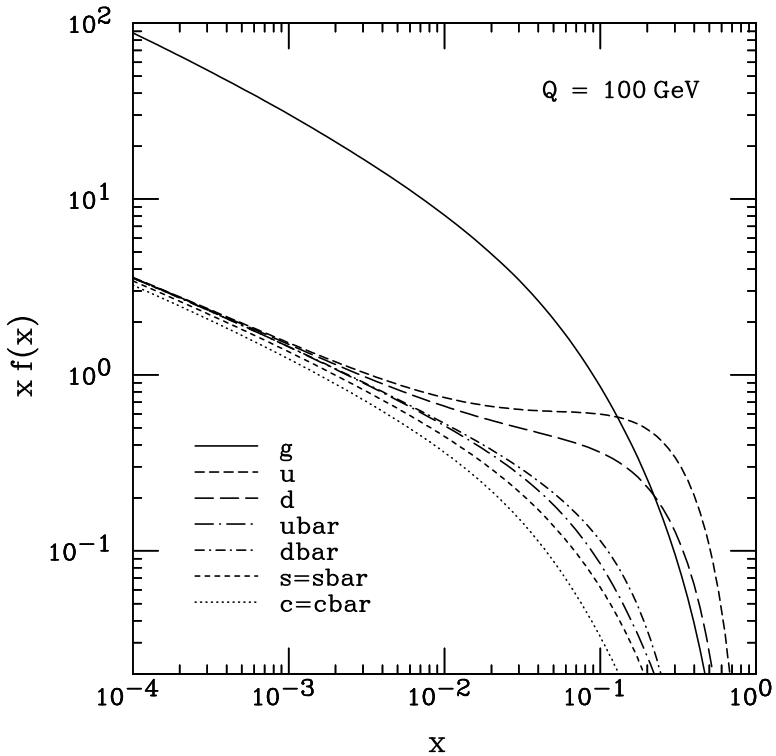


Figure 1.1: Parton distribution functions of quarks and gluons in the proton at an energy scale, $Q = 100$ GeV [6].

In the $p\bar{p}$ collisions of the Tevatron, $t\bar{t}$ pairs are produced through quark pair ($q\bar{q}$) annihilation or gluon-gluon (gg) fusion, the tree-level diagrams of which are shown in Figures 1.2 and 1.3, respectively. At the energy-scale of the Tevatron, production through $q\bar{q}$ annihilation accounts for about 85 percent of $t\bar{t}$ pairs produced while the remaining 15 percent are a result of gg fusion. This is because the energy needed to produce the massive top quarks requires the interacting partons to carry a significant fraction of the proton's (antiproton's) momentum and, as noted above, the u and d quarks carry most of the proton's momentum.

The total cross section for $t\bar{t}$ pair production has been calculated at next-to-leading order, in-

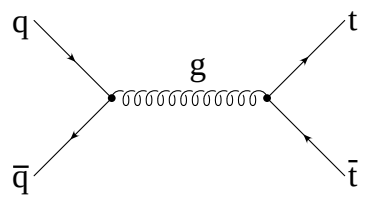


Figure 1.2: Leading order (tree-level) diagram of $t\bar{t}$ production via $q\bar{q}$ annihilation.

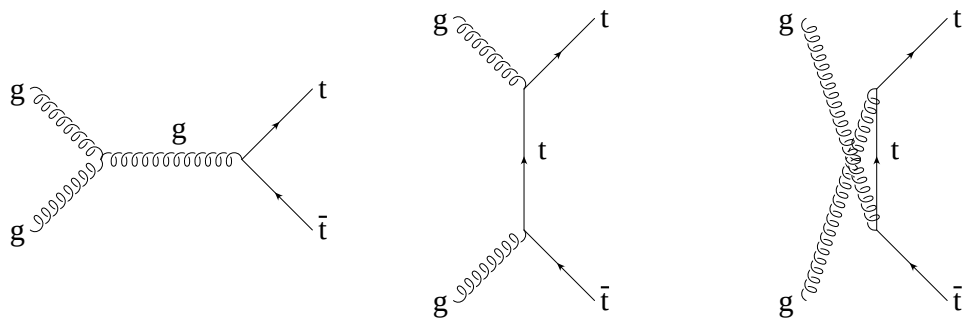


Figure 1.3: Leading order diagrams of $t\bar{t}$ production via gg fusion.

cluding next-to-leading-logarithm contributions, in the Standard Model to be $6.7^{+0.7}_{-0.9}$ pb[3] for a top quark mass of 175 GeV/c^2 [7, 8]. The uncertainty on this calculation is dominated by the PDF and α_s uncertainties. The large PDF uncertainty is a result of the sensitivity of the cross section calculation to the large-x gluon content of the proton, which is still poorly known. The next-to-leading-logarithm contributions have been shown to be small, but they improve the stability of the prediction with respect to changes of the scale, μ.

1.3.2 Top Quark Decay

In the Standard Model, top quarks decay into a b quark and a W boson nearly 100 percent of the time. This branching fraction is not well constrained experimentally, but can be predicted using other measurements and certain Standard Model assumptions. The Cabibbo-Kobayashi-Maskawa (CKM) matrix is a 3×3 matrix, V_{CKM}, that describes the couplings among the different quark flavors in charged-current $W^{\pm}$ interactions. A number of the couplings have been measured very well, so using a global fit to the data and constraining the matrix to be unitary, the magnitude of the nine CKM

[3]Cross sections are measured in units of barns (b), where 1 barn is 10^{-24} cm^2.

elements are [2]

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.9738 \pm 0.0002 & 0.227 \pm 0.001 & (3.96 \pm 0.09) \times 10^{-3} \\ 0.227 \pm 0.001 & 0.9730 \pm 0.0002 & (42.2^{+0.1}_{-0.8}) \times 10^{-3} \\ (8.1^{+0.3}_{-0.6}) \times 10^{-3} & (41.6^{+0.1}_{-0.8}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix}. \tag{1.2}$$

The decays $t \to W^+ d$ and $t \to W^+ s$ are suppressed relative to $t \to W^+ b$ by factors of $|V_{td}|^2$ and $|V_{ts}|^2$, respectively. The magnitude of V_{tb} is directly measurable by studying single-top production and the CDF and D0 collaborations have recently made those measurements. The CDF collaboration measured a value of $|V_{tb}| = 0.88 \pm 0.14(\text{experimental}) \pm 0.07(\text{theory})$ and placed a 95% confidence level lower limit of $|V_{tb}| > 0.66$ [9]. The D0 collaboration measured a value of $|V_{tb}| = 1.3 \pm 0.2$, and placed a 95% confidence level lower limit of $0.68 < |V_{tb}| \leq 1$ constraining the value to the Standard Model region [10].

The final state of $t\bar{t}$ pairs consists of the decays of the two W bosons and b quarks from the top decays. The b quarks hadronize and decay, and are detected as hadronic jets. Each W boson can decay either leptonically to a charged lepton and a neutrino or hadronically to a pair of quarks. Top-pair events are categorized by the decay of the W bosons; they may be "all-hadronic," "lepton-plus-jets," or "dilepton" events depending on whether neither, one, or both of the Ws decays leptonically.[4]

At leading order, the W has a leptonic branching fraction 1/9 for each channel and a hadronic branching fraction of 6/9; thus, the all-hadronic channel has the largest branching fraction of all $t\bar{t}$ event channels. However, the all-hadronic channel also has the largest background, which, along with the large uncertainties on jet energies measured in the calorimeter, makes studying the top quark in this channel very challenging. The dilepton channel, on the other hand, is very clean (i.e. contributions from backgrounds are small); however, this channel suffers from a low branching fraction. The lepton-plus-jets channel has a relatively high branching fraction of about 30 percent and though the backgrounds are considerably higher than in the dilepton channel, they are still manageable. The analysis presented in this thesis makes use of the lepton-plus-jets channel.

1.4 The Interest in the $t\bar{t}$ Production Cross Section

As noted above, the top quark is unique in a number of ways, and studying its properties is of great interest. In the Standard Model, the $t\bar{t}$ cross section is calculated with a precision of about

[4] The leptons in lepton-plus-jets and dilepton refer to e and μ. Top quark events containing a $W \to \tau\nu_\tau$ decay are difficult to identify.

15 percent. The Standard Model further predicts that the top quark decays to a W boson and a b quark with a branching fraction close to 100 percent. Measuring the cross section tests both the production and decay mechanisms of the top quark. The measured cross section could be enhanced by non-Standard Model production mechanisms, such as the production and decay of a heavy resonance into $t\bar{t}$ pairs [11]. Non-Standard Model top quarks decay, such as a decay into supersymmetric particles [12], could suppress the measured cross section by reducing the branching fraction of $t \to W^+b$, which is assumed to be 100 percent.

Besides the absolute measurement, deviations from the Standard Model could appear in the comparison of the cross section in different decay channels. Exotic top decays could enhance the measured cross section in one channel and suppress it in another. For example, a top quark decay to a charged Higgs boson would enhance the dilepton cross section measurement relative to the lepton-plus-jets one since a charged Higgs is expected to decay to predominantly to a τ lepton, which in turn decays leptonically with a large branching fraction.

The cross section is also sensitive to the top mass, decreasing, theoretically, by approximately 0.2 pb for each 1 GeV/c^2 increase in the value of the top mass over the range 170 GeV/$c^2 < M_{top} <$ 190 GeV/c^2. This dependence can be exploited to turn a cross section measurement into an indirect determination of the top quark mass. Conversely, the theoretically predicted dependence of the cross section on the top mass can be tested against measurements of the cross section and top quark mass.

At the Large Hadron Collider (LHC), top quark production is expected to be an important background to many possible new physics processes that will be searched for there. Precise measurements of the top quark at the Tevatron will help in understanding the top contribution at the LHC, a prerequisite to declaring any sort of discovery.

Chapter 2

The Fermilab Tevatron and the CDF II Detector

The data analyzed for this thesis were collected with the Run II Collider Detector at Fermilab (CDF), between March 2002 and May 2007, from proton-antiproton collisions produced by the Fermilab Tevatron. This chapter describes the Tevatron accelerator and the CDF II experimental apparatus.

2.1 The Tevatron

The Tevatron is a superconducting proton-antiproton ($p\bar{p}$) accelerator located at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois. Producing $p\bar{p}$ collisions at center-of-mass energies, $\sqrt{s}$, of 1.96 TeV, the Tevatron is currently the world's highest-energy particle accelerator and the only facility producing top quarks.

2.1.1 Proton Production

The Fermilab accelerator chain [13, 14] begins with a Cockcroft-Walton pre-accelerator. In the pre-accelerator, hydrogen gas is ionized to create H^- ions, which are then accelerated to an energy of 750 keV. Next, the ions are injected into a 150 m long linear accelerator, known as the Linac. A chain of radiofrequency cavities in the Linac accelerate the negatively charged ions from 750 keV to 400 MeV. Electrons are stripped from the ions when the 400 MeV beam is passed through a carbon foil, leaving only the hydrogen nuclei (protons). The protons are then sent to the Booster ring for the next stage of acceleration. The Booster is a circular accelerator, or synchrotron, that accelerates the protons to 8 GeV and gathers them into bunches. The next machine in the accelerator chain is the Main Injector, another synchrotron with a circumference several times that of the Booster. The Main Injector accelerates the protons to 150 GeV and coalesces them into a single bunch before injecting them into the Tevatron ring for the final stage of acceleration. In a separate process, the Main Injector is also used to accelerate protons to 120 GeV and to be used for the production of antiprotons. A schematic of the Fermilab accelerator complex is shown in Figure 2.1.

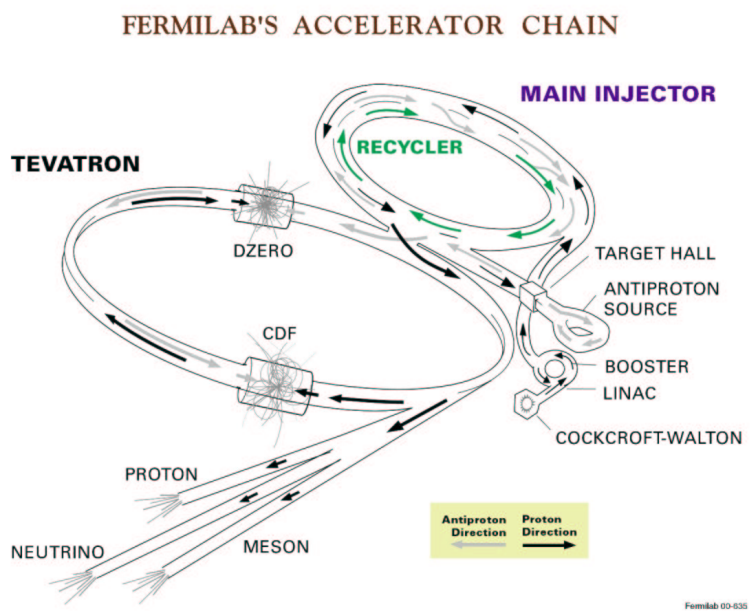


Figure 2.1: The Fermilab accelerator chain [15].

2.1.2 Antiproton Production and Recycling

To produce antiprotons, 120 GeV protons from the Main Injector collide with a nickel-copper target. The collisions produce a multitude of secondary particles, some of which are antiprotons. A cylindrical lithium conductor, called the Collection Lens, produces a solenoidal magnetic field that focuses the secondary particles into a parallel beam. A pulsed dipole magnet is then used to select negatively charged particles with with momenta of the order of 8 GeV from the spray. In this manner, antiprotons are harvested from the collision with the target [1]. The momentum spread of the beam is reduced inside the Debuncher through the application of betatron (transverse) stochastic cooling and momentum (longitudinal) cooling. After cooling, the antiprotons are injected into the Accumulator for temporary storage. From the Accumulator, antiprotons may also be stored in the Recycler Ring, which provides additional storage. Limiting the stack size in the Accumulator allows an optimization of antiproton accumulation rate—the primary limiting factor in Tevatron running.

2.1.3 Collisions

When enough antiprotons have been accumulated a transfer to the Main Injector is initiated. Here, both protons and antiprotons are accelerated to 150 GeV. From the Main Injector, 36 bunches

[1] This process is inherently inefficient. Typically, only 1 or 2 antiprotons are captured and stored for every 10^5 protons striking the target.

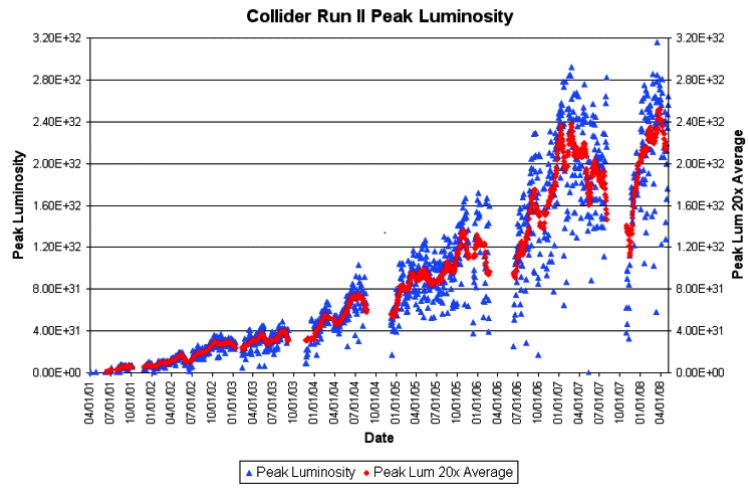


Figure 2.2: Tevatron peak luminosity between April 2001 and April 2008 [16].

of counter-rotating protons and antiprotons are distributed around the 6 km circumference of the Tevatron ring for the final stage of acceleration. Each beam is accelerated to 980 GeV inside the Tevatron, which gives center-of-mass energies, $\sqrt{s}$, for collisions of 1.96 TeV.

The 36 proton bunches, each containing 10^{11} protons, cross 36 bunches of antiprotons, each containing 10^{10} antiprotons, every 396 ns. The counter-rotating proton and antiproton bunches are kept apart in electrostatically-separated helical orbits so they pass each other undisturbed as they travel around the ring at $0.9999996c$. Only at the desired interaction points B0 and D0, where the CDF and D0 experiments reside, are the beams focused by quadrapoles to collide together.

The intensity of the beams, called "luminosity," can be calculated using the formula

$$\mathcal{L} = nf\frac{N_p N_{\bar{p}}}{4\pi\sigma_x\sigma_y}, \tag{2.1}$$

where n denotes the number of bunches, f is the frequency of revolution, N_p ($N_{\bar{p}}$) is the number of protons (antiprotons) per bunch, and σ_x and σ_y are the beam dimensions in the plane transverse to the beam at the interaction point. The instantaneous luminosity degrades over time as particles are lost due to collisions and orbital variations. An important aspect of the Tevatron Run II operations has been increased luminosity through more efficient storing, cooling, and transferring of antiprotons. Figure 2.2 shows the improvement of the Tevatron's peak luminosity over time.

It is through the high-energy proton-antiproton collisions that $t\bar{t}$ pairs are produced. The data used in this analysis represent an integrated luminosity, $\int \mathcal{L}dt \approx 2$ fb^{-1} collected by the CDF

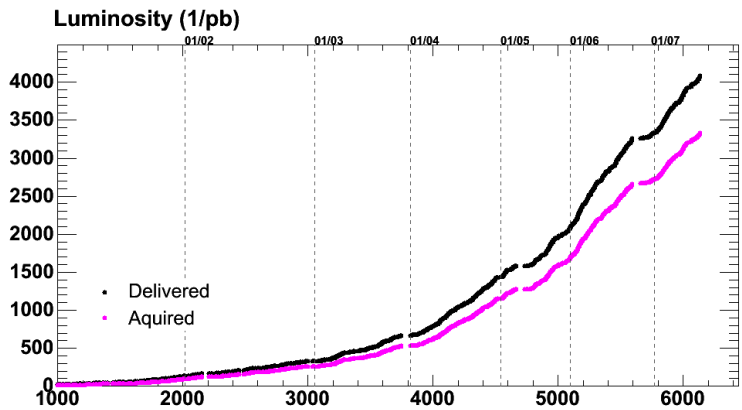


Figure 2.3: Total luminosity gathered by the CDF detector. The black curve shows the luminosity delivered by the Tevatron and the pink curve is the luminosity written to tape by CDF.

detector. Figure 2.3 shows the integrated luminosities delivered by the Tevatron and recorded by CDF as a function of time.

2.2 The CDF II Detector

CDF is a general purpose solenoidal detector [17] built around one of the $p\bar{p}$ collision points (B0) of the Tevatron accelerator. Actually a collection of several sub-detectors, it is designed to measure particles produced in $p\bar{p}$ collisions at the Tevatron. The CDF II detector (Figure 2.4) is designed around a central superconducting solenoid, which provides a constant 1.4 T magnetic field, and hence inherits a cylindrical geometry. The detector is azimuthally and forward-backward symmetric with the entire apparatus centered about the $p\bar{p}$ interaction point. The CDF coordinate system is (z, ϕ, η), where the z-axis runs parallel to the proton beam direction. The azimuthal angle is ϕ, and η is the "pseudorapidity"[2] variable given as $\eta = -\ln(\tan(\frac{\theta}{2}))$, where θ is the polar angle (pseudorapidity transforms linearly under Lorentz boosts in the z-direction and $\Delta\eta$ is an invariant). The range of $|\eta|$ is from 0 at 90° to the beam direction to ~ 3.5 at the most forward region of the detector (see Figure 2.5). The central region of CDF falls in the range of $0 \leq |\eta| \leq 1.0$ while the forward detectors are located in the range $1.0 < |\eta| \leq 3.0$. As mentioned above, CDF is a collection of detectors that work in concert to identify and measure the properties of particles produced in $p\bar{p}$ collisions. Detector subsystems are placed at various radial distances with tracking information recorded closest to the interaction point, muon detection done furthest out, and measurement of

[2]For relativistic particles ($p \gg m$), η is a good approximation of the true rapidity. The rapidity, $y \equiv \tanh^{-1} p_z/E$. The pseudorapidity of a particle can easily be measured even though the mass and momentum are unknown.

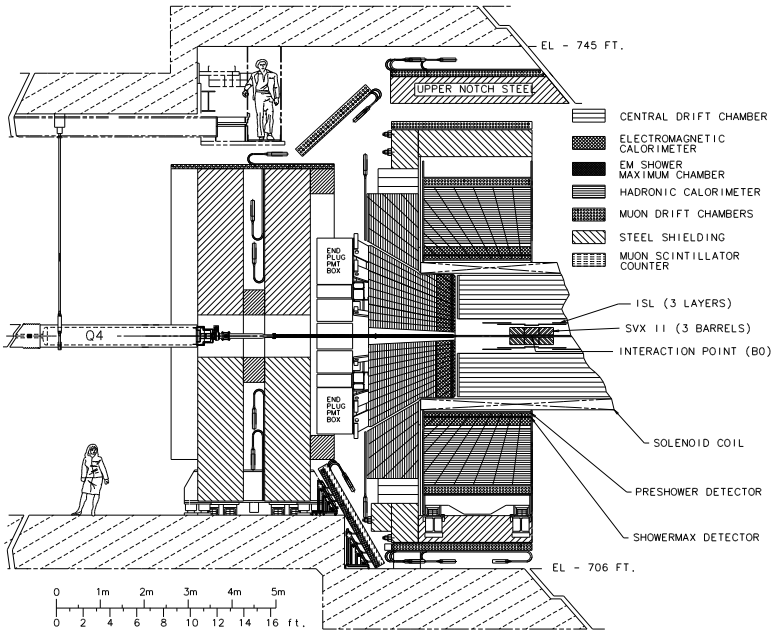


Figure 2.4: One half of the CDF II detector from an elevation view. The various sub-detector systems are symmetric both azimuthally and forward-backward.

particle energy performed between these two regions. In general, the CDF subsystems can be divided into several categories: particle tracking systems, calorimetry, and muon detectors. Data is collected using a triggering system that uses the information from the CDF subdetectors to determine what events are of interest for further study. CDF is composed of many subsystems, so the remainder of this chapter will only focus on those components of the detector that are relevant to the analysis described in this thesis. A detailed description of the the entire CDF II apparatus can be found in [17].

2.2.1 Tracking Systems

Tracking information is crucial for the measurement of particle momenta and identification of particle charge, and CDF employs several tracking systems to record the paths of charged particles as they pass through the detector. The tracking systems are contained inside a solenoid with a uniform 1.4 T magnetic field that runs parallel to the beam axis to make the measurement of particle charge and momentum possible. Charged particles passing through the solenoid's magnetic field follow a helical path. The radius of this helix is related to the transverse momentum of the particle:

$$p_T = \frac{0.3B}{2C}, \tag{2.2}$$

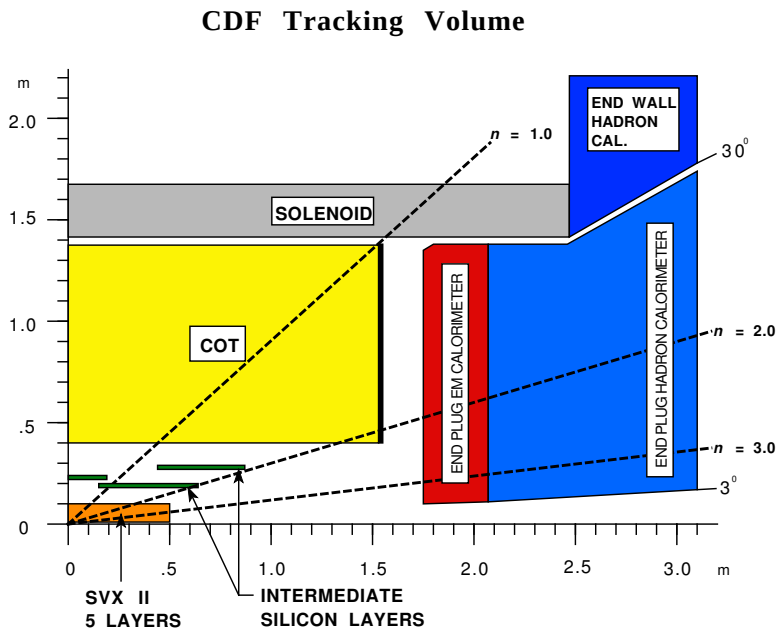


Figure 2.5: One quadrant of the CDF tracking volume and forward calorimetry with η superimposed.

expressed in units of GeV/c, where B is the magnitude of the magnetic field in Tesla and the half-curvature, C, is in m^{-1} and is related to the radius of the helix by $C = 1/(2R)$. The helix of a charged track is described by the following parameters:

- $\cot\theta$, the cotangent of the polar angle of the particle at the point of closest approach to the beamline.
- C, the half-curvature. This parameter has the same sign as the charge of the particle.
- z_0, the z position of the particle at the point of closest approach to the beamline.
- d_0, the impact parameter. The distance between the helix and the beamline at the point of closest approach to the beamline.
- ϕ_0, the azimuthal angle of the particle trajectory at the point of closest approach to the beamline.

CDF's tracking systems are arranged concentrically inside the solenoid magnet at various radii. The inner-most system is a silicon microstrip detector [18], known as the silicon vertex detector (SVX), which extends from a radius of 1.5 cm to 10.7 cm from the center of CDF. The SVX silicon microstrip detectors are arranged in a barrel geometry consisting of multiple layers and provide track information to $|\eta| < 2$. The microstrips are semiconductor detectors created from a

bulk layer of lightly-doped n-type silicon sandwiched between a strongly-doped n electrode and a lightly-doped p region. A positive voltage is applied to the n electrode, depleting the bulk layer of electrons and creating an electric field within its volume. When a charged particle passes through the detector causes ionization and, in turn, electron-hole pairs. The electrons and holes drift in the electric field of the microstrip detector inducing a signal.

The five layers of SVX microstrip detectors are double-sided with one side providing measurements in the r-ϕ plane and the other in the r-z plane. With an impact parameter resolution $\sigma_\phi < 30$ μm for central-high-momentum tracks, the SVX detector allows for the reconstruction of secondary vertices displaced from the primary vertex, which come from the decay of heavy-flavor hadrons. While not used extensively in this analysis, the ability to reconstruct secondary vertices is critical for an algorithm used in a complementary analysis for the identification of b quarks.

Outside the SVX lie the Intermediate Silicon Layers (ISL) [19]. The ISL detectors are similar in design to those in the SVX. The ISL tracking detector is comprised of three concentric layers of silicon microstrip detectors. In the central region a single layer of silicon placed at a radius of 22 cm improves the efficiency of 3D tracking by providing an extra track measurement that connects the track segments measured in the SVX and the Central Outer Tracker (COT). Two additional layers of microstrip detectors are positioned in the high-η region at radii of 20 cm and 28 cm and significantly improves the momentum resolution in the region where COT coverage is incomplete or missing.

The Central Outer Tracker (COT) is an open-cell-drift chamber that provides track information at large radii in the region $|\eta| \leq 1.0$ [20] (see Figure 2.5). The COT contains 30,240 gold-plated tungsten sense wires arranged in eight "superlayers" providing up to 96 track measurements between radii of 40 cm and 137 cm. Four "axial" superlayers run parallel to the beamline and provide r-ϕ information. The remaining four superlayers are positioned with a tilt of 2° with respect to the beamline to provide "stereo" information. The eight superlayers are enclosed in an air-tight chamber filled with a 50/50 gas mixture of Ar and ethane (C_2H_6) bubbled through isopropyl alcohol. Charged particles passing through the COT ionize the gas mixture along their path and the freed electrons drift toward the sense wires in an electric field created by voltages on the sense (anode) wires and cathode "field" panels. In the high electric field near a sense wire electron-atom collisions cause a multiplication (avalanche) of charges that is registered on the wire as a pulse. The drift time, measured as the time between the collision and the first arrival time of electrons on a wire, is converted into a distance of closest approach of the particle to the sense wire. Three-dimensional

reconstruction of the path of the charged particle is then possible using the distance of closest approach measurements on the axial and stereo wires. The COT chamber has a single hit position resolution of approximately 140 μm and a p_T resolution of $\delta_{p_T}/p_T^2 \simeq 0.15\%\ (GeV/c)^{-1}$.

Track hit information from the silicon systems and the COT are correlated using an "outside-in" algorithm. Tracks from the COT are extrapolated back into the volume of the silicon detectors and matched to silicon hits in the r-ϕ plane. At each step of this extrapolation, multiple scattering considerations determine the region in which the algorithm looks for a track. Once a reconstructed track is found in the r-ϕ plane, the z-position information from the silicon detector is added to the track.

2.2.2 Calorimetry

Outside of the solenoid volume and surrounding the tracking system, CDF employs scintillator-based sampling calorimetry to measure the kinetic energy deposited by particles observed in the detector in the region $|\eta| \leq 3.6$. The calorimetry is divided into two systems, providing separate measurements for electromagnetic (EM) and hadronic (HAD) energies. The EM calorimeter lies just outside the solenoid and is used primarily for the identification of electrons and photons. Further out, the HAD calorimeter measures the kinetic energy of hadronic particles in high-energy jets. Both components use a series of scintillators with metal absorber layers. Light from the scintillator is collected by a light-pipe and a wavelength shifting fiber that directs the energy into photomultiplier tubes for detection. The calorimeters are segmented in η and ϕ in a tower geometry with each segment pointing back to the nominal interaction region. Each tower covers a solid angle of 0.1 by 15° in $\eta \times \phi$ space.

Electrons lose energy by radiating photons in an electromagnetic shower while being decelerated in the Coulomb field of an atomic nucleus. The mean distance over which a high-energy electron loses all but $\frac{1}{e}$ of its energy is called a radiation length (denoted by X_0). The CDF EM calorimeter is designed in terms of these radiation lengths. The EM calorimeter [21] is composed of alternating layers of scintillating polystyrene and and Pb absorber and has a total thickness of 21 radiation lengths worth of material. The energy resolution of the central EM calorimetry is $13.5\%/\sqrt{E_T} \oplus 2\%$[3], while the resolution of the plug EM calorimeter is $16\%/\sqrt{E_T} \oplus 1\%$ [22]. The innermost layer of the calorimeter also has wire chambers that act as a calorimeter pre-shower detector (CPR), and sample the early development of an electromagnetic shower within the solenoid. Further out, located at

[3]The symbol $\oplus$ indicates that the errors are to be added in quadrature.

a depth of approximately 6 radiation lengths, is a shower maximum detector (also known as the CES), which is a system of proportional wire chambers. The CES provides an accurate position measurement of the electromagnetic shower, which is used to match the calorimeter deposits with electron tracks, as well as providing information on the shower shape.

Hadrons lose energy through ionization and collisions with the nuclei of the material through which they pass. The thickness of the hadronic calorimeter is, therefore, described in terms of the pion nuclear interaction length, λ_0, rather than radiation lengths. The central HAD calorimeters [23] are comprised of layers of acrylic scintillators and iron, totaling a thickness of approximately $4.5\lambda_0$ of material (the plug calorimeter is $7\lambda_0$ thick). The energy resolution of the hadronic calorimeters is $74\%/\sqrt{E_T} \oplus 4\%$.

2.2.3 Muon Detectors

The muon systems are the furthest detectors out from the interaction point. The CDF muon detectors are composed of layers of single-wire proportional drift chambers backed up by scintillator counters for fast timing. Together, the muon subsystems provide coverage for $|\eta| \leq 2.0$ and nearly complete coverage in ϕ. Figure 2.6 shows the coverage of CDF's muon detectors. Being positioned so far from the interaction point, muon detectors are shielded by a significant amount of steel from the hadronic calorimeter, the solenoid return yoke, and additional steel to reduce hadronic punch-through (non-muons that reach the muon chambers). High-p_T muons are approximately minimum ionizing particles and, therefore, travel many interaction lengths before stopping. The muon chambers record hits along the path of the muon through the detector from which muon track segments (stubs) are reconstructed.

Four subdetector systems comprise the muon detectors. The inner-most subdetector, the central muon detector (CMU) [24], is located behind the central calorimeter at a radius of 347 cm from the beam. Muons in the region $|\eta| \leq 0.6$ with $p_T > 1.4$ GeV/c, the minimum momentum required to get through the calorimeter, can be detected by the CMU. The CMU is segmented into 24 wedges of 12.6° in ϕ with a gap of 2.4° between each wedge. The CMU is also divided into East ($\eta > 0$) and West ($\eta < 0$) halves with a gap at $\eta = 0$. Each wedge consists of three modules with four layers of four rectangular drift cells for a total of 2304 drift chambers in the CMU. Four sense wires, one from each layer, make up a CMU tower. Two of the four sense wires, from alternating layers, lie on a radial line that passes through the beamline. The remaining two wires of the tower lie on a radial line that is offset by 2 mm from the first.The radial alignment of the wires allows for a crude

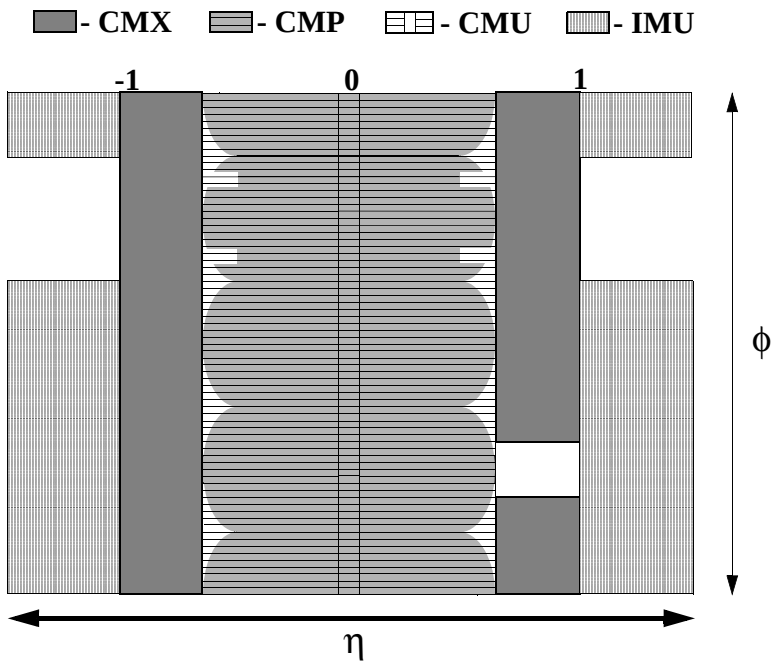


Figure 2.6: Coverage of the CDF muon systems in azimuth, ϕ, and pseudorapidity, η.

momentum measurement that is used by the CDF trigger system. The CMU has a hit resolution of about 250 μm in the r-ϕ plane. Additionally, the CMU chambers provide information on the z-position of a muon, with a resolution of approximately 1 mm, from charge deposition on each end of the sense wires.

The central muon upgrade (CMP) lies behind an additional 60 cm of steel, which reduces hadronic punch-through by a factor of 10. Like the CMU, the CMP has four layers of drift chambers, but with a half-cell stagger instead of the small offset of the CMU. The CMP has a box geometry, as opposed to the cylindrical geometry of the rest of the CDF detector, and thus has a coverage in η that varies as a function of ϕ, as seen in Figure 2.6. A layer of scintillation counters on the outer surface of the CMP chambers helps identify what beam crossing (collision) the muon track should be matched to because the maximum drift time (1.4 μs) is long compared to the bunch crossing rate. The CMP is able to detect muons with $p_T > 2.8$ GeV/c (at 90° to the beam, this increases as the direction becomes more forward) in the $|\eta| \leq 0.6$ region. The single-hit position resolution of the CMP is about 300 μm in the r-ϕ plane. The CMP does not provide any information on the z position of a muon. Because of the overlap in coverage, muons may leave stubs in both the CMU and CMP. Such muons are known as CMUP muons and have a lower rate of particles that are misidentified as muons.

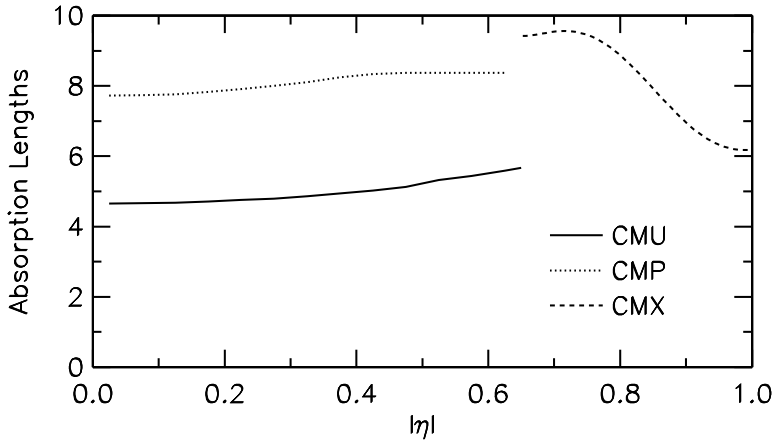


Figure 2.7: Number of interaction lengths as a function of η, averaged over ϕ acceptance, of the CMU, CMP, and CMX systems.

A third muon subsystem is the central muon extension (CMX), which, as its name implies, extends central muon coverage with chambers in the region $0.6 \leq |\eta| \leq 1.0$ and provides nearly complete coverage in ϕ, as shown in Figure 2.6. The large angle through the hadron calorimeter and CDF support structures yields a considerable amount of absorber material in front of the CMX, though the amount varies, as shown in Figure 2.7. The CMX chambers are similar in design to those in the CMP, only shorter. However, the chambers are arranged in a conical geometry and muons pass through between four and eight layers of wires depending on the position in η. The conical geometry of the CMX also allows the z-position of a muon passing through the detector to be determined. The CMX is actually comprised of three subsystems. The CMX "Arches" are the oldest and largest sections; they cover all but the top 30° and bottom 90° of the detector in azimuth. The two other pieces, known as the "Keystone" and "Miniskirt," were commissioned during Run II. The Keystone covers the top 30° on the West side ($\eta < 0$) of the detector (the East side contains the cryogenics system for the CDF solenoid). The Miniskirt instruments the bottom 90° of the detector and has a different geometry than the rest of the CMX detector due to space constraints.

The fourth muon subsystem, the barrel muon upgrade (BMU), covers the range $1.0 \leq |\eta| \leq 2.0$. Information from the BMU is not used for this analysis.

In general, muon identification (for isolated, primary muons) at CDF proceeds by matching these stubs to tracks in the central tracker and determining if energy deposition in the calorimeters is consistent with minimum ionization. This analysis is based around another method of muon identification that identifies muons in jets (see Chapter 4).

2.2.4 Trigger System

Given the high instantaneous luminosity and collision frequency at B0, it is impossible for CDF to record and analyze every event. At the Tevatron, proton-antiproton collisions occur at the rate of about 2.5 MHz (once every 396 ns) and the average size of each of these events is 250 kB. To reduce the amount of data recorded and stored for future analysis, CDF uses a three-tier trigger system that filters out most of the less-interesting events based on measured kinematic information and particle identification. The CDF trigger system was designed to be "deadtimeless," meaning that the detector is able to process an event even if it is still busy with the preceding one.

Figure 2.8 shows a block diagram of the data flow at CDF. The first stage in the trigger system, Level 1, collects information from the calorimeters, COT and muon systems and is capable of making a decision whether to accept or reject an event based on this information. Level 1 has a storage buffer that is 14 clock cycles deep (with 396 ns between bunch crossings) and arrives at a decision about any given event within 5.5 μs, before the event falls out of the pipeline. The Level-1 trigger reduces the event rate down to about 20 kHz. Events that pass Level 1 are transferred to the Level-2 trigger which has four local buffers and is asynchronous with a decision time of about 30 μs per event. The Level-2 trigger is able to reconstruct calorimeter clusters, and to use the CES information for jet reconstruction and identification of electrons and photons. Level 2 accepts events and passes them on to Level 3 at a rate of approximately 300 Hz (this has recently been upgraded to $\sim$ 1 kHz). Level 3 is the final stage in the trigger system and is a farm of nearly 300 CPU nodes that assemble and analyze each data event in greater detail. The Level-3 trigger algorithms perform a simplified version of the event reconstruction that is used in offline analysis. Once an event is accepted by Level 3 it is passed on to a data-logger subsystem that sends the event on to be written for permanent storage. The triggers applied at Level 3 reduce the event rate to about 100 Hz, which is written to tape.

The $t\bar{t}$ candidate events examined in this analysis were triggered on a single high-p_T lepton (electron or muon). The details of these triggers are given in Sections 5.1.1 and 5.2.1 of this thesis. Several other datasets were also used for the determination of background contributions. The triggers for these datasets are also briefly detailed in the relevant sections.

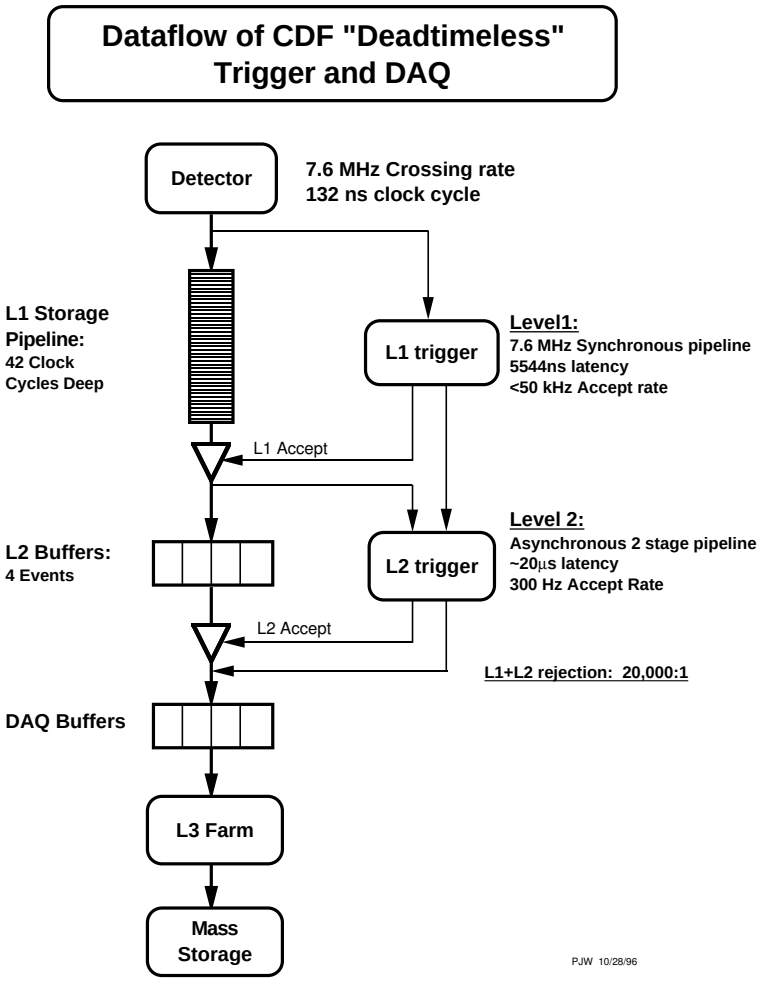


Figure 2.8: A block diagram of the CDF II data flow [17]. The diagram depicts a Level 1 storage pipeline 42 clock cycles deep (for a 132 ns clock cycle). The current Tevatron configuration produces collisions on a 396 ns clock cycle; with this configuration the Level 1 pipeline is 14 clock cycles deep.

Chapter 3

Monte Carlo Modeling

3.1 $t\bar{t}$ Event Modeling

The detector acceptance of $t\bar{t}$ events is modeled using PYTHIA v6.216 [25] and HERWIG v.6.510 [26]. This analysis uses the former for the final cross section estimate and the latter to estimate the systematics resulting in the modeling of $t\bar{t}$ production and decay. These generators employ leading order matrix elements for the hard parton scattering process and parton showering to simulate radiation and fragmentation effects. The PYTHIA event generator has been tuned using jet data to better model the effects of multiple interactions and remnants from the break-up of the proton and antiproton. The generators are used with the `CTEQ5L` pardon distribution functions [27]. Decays of b and c hadrons are modeled using `EvtGen` v9.1 [28].

3.2 W+Jets Modeling

Events with a W boson produced in association with multiple jets are modeled using ALPGEN v2.1 [29], with parton showering provided by PYTHIA v6.326 and heavy-flavor hadron decays handled by `EvtGen`. ALPGEN calculates exact matrix elements at leading order for a large set of parton level processes in QCD and electroweak interactions. Multiple samples are generated, forcing a certain heavy-flavor configuration as needed and allowing different numbers of light-flavor partons (denoted with a 'p') in the matrix element. In other words, before showering, there are separate samples $W+N\text{p}$, $Wb\bar{b}+N\text{p}$, $Wc\bar{c}+N\text{p}$, and $Wc+N\text{p}$, where the number, N, of light-flavor partons ranges from 0 to 4 (with a maximum of 4 total heavy- and light-flavor partons allowed).

The showering in PYTHIA may result in multiple ALPGEN samples covering the same phase space (e.g. an event with a W and 2 jets may come from the W+2p sample or the W+1p sample with an extra jet from the shower) and this is corrected for during generation using what is known as MLM matching. The MLM matching checks that every parton is matched to a jet in the event

and that the number of jets is equal to the number of generated partons (except for the highest Np samples, which are allowed to have extra jets). Similarly, the same heavy-flavor final state may arise in multiple samples. This overlap must be removed "by hand." For this analysis, the heavy-flavor overlap is removed by enforcing the appropriate heavy-flavor content for the dedicated heavy samples ($Wb\bar{b}$, $Wc\bar{c}$, and Wc), allowing additional heavy flavor from the shower only when it is lighter than the primary generated partons (i.e. charm pairs in $Wb\bar{b}$ are allowed, not vice versa) or when it fails the kinematic filter used at generation.

Once all the samples have been generated, they must be properly combined to simulate W+jets data. The samples are weighted so they all have equivalent integrated luminosities. Each sample generated has a number of events, n, and a cross section, σ_A, calculated by ALPGEN. Therefore, the integrated luminosity of a sample is given by $\int \mathcal{L} dt_{\text{sample}} = n/\sigma_A$. To give every sample the same integrated luminosity, each is weighted by a factor $1/\int \mathcal{L} dt_{\text{sample}} = \sigma_A/n$. The samples may then be combined by addition.

3.3 Other Processes

Estimates of backgrounds from diboson production (WW, WZ, and ZZ) and Drell-Yan/$Z \to \tau\tau$ are derived using PYTHIA. Drell-Yan to $\mu\mu$ events are model using ALPGEN with PYTHIA showering while single-top production is modeled with `MadEvent` [30], also with PYTHIA showering. The various Monte Carlo samples used in this analysis are also listed in Table 3.1.

3.4 Detector Simulation

The CDF detector simulation reproduces the response of the detector and uses the same detector geometry database as the event reconstruction. The simulation takes into account that certain elements of the detector have not always been on and operational; for example, certain layers of the SVX were not consistently operating early in Run II. Tracking of particles through matter is performed with `GEANT3` [31]. The drift model for the COT uses a parametrization of a `GARFIELD` simulation [20] with parameters tuned to match COT collider data. The calorimeter simulation uses the `GFLASH` [32] parametrization package interfaced with `GEANT3`. The `GFLASH` parameters are tuned to test beam data for electrons and high-p_T pions and checked by comparing the calorimeter energy of isolated tracks in the collision data to their momenta as measured in the COT. The output of the CDF simulation are events in the same format as real data, meaning the same code can be used to

Process	Generator
Signal samples	
$t\bar{t}$ ($M_{top} = 175$ GeV/c^2)	PYTHIA
$t\bar{t}$ ($M_{top} = 175$ GeV/c^2)	HERWIG
W+Jets	
$W \to e\nu_e + N\mathrm{p}$ (N=0,..,4)	ALPGEN+PYTHIA
$W \to \mu\nu_\mu + N\mathrm{p}$ (N=0,..,4)	ALPGEN+PYTHIA
$W \to e\nu_e + b\bar{b} + N\mathrm{p}$ (N=0,1,2)	ALPGEN+PYTHIA
$W \to \mu\nu_\mu + b\bar{b} + N\mathrm{p}$ (N=0,1,2)	ALPGEN+PYTHIA
$W \to e\nu_e + c\bar{c} + N\mathrm{p}$ (N=0,1,2)	ALPGEN+PYTHIA
$W \to \mu\nu_\mu + c\bar{c} + N\mathrm{p}$ (N=0,1,2)	ALPGEN+PYTHIA
$W \to e\nu_e + c + N\mathrm{p}$ (N=0,..,3)	ALPGEN+PYTHIA
$W \to \mu\nu_\mu + c + N\mathrm{p}$ (N=0,..,3)	ALPGEN+PYTHIA
Other backgrounds	
single top (s-channel)	`MadEvent`+PYTHIA
single top (t-channel)	`MadEvent`+PYTHIA
$DY \to \tau\tau$ ($M_Z > 30$ GeV)	PYTHIA
$Z \to \mu\mu + N\mathrm{p}$ (N=0,..,4) ($M_Z = [75, 105]$ GeV)	ALPGEN+PYTHIA
$DY \to \mu\mu + N\mathrm{p}$ (N=0,..,4) ($M_Z = [20, 75]$ GeV)	ALPGEN+PYTHIA
$DY \to \mu\mu + N\mathrm{p}$ (N=0,..,4) ($M_Z = [105, 600]$ GeV)	ALPGEN+PYTHIA
WW	PYTHIA
WZ	PYTHIA
ZZ	PYTHIA
For mistag studies	
Dijet $p_T > 18$ GeV/c	PYTHIA
Dijet $p_T > 40$ GeV/c	PYTHIA
Dijet $p_T > 60$ GeV/c	PYTHIA
Dijet $p_T > 90$ GeV/c	PYTHIA
$B \to \mathrm{D}^*\pi \to \mathrm{D}^0\pi \to K\pi\pi$	`Bgenerator`

Table 3.1: Summary of the Monte Carlo samples used for this analysis.

analyze data and Monte Carlo events. Further details on the CDF simulation can be found in [33].

Chapter 4

Soft Muon Tagging (SLT)

The purpose of the soft muon tagger (SLT) is to identify (tag) b-jets in $t\bar{t}$ events by looking for muons in jets resulting from the semileptonic decay of a b hadron [34, 35]. In general, muon identification at CDF relies on the presence of a track segment (stub) in the muon chambers, associated with a track in the central tracking systems and energy deposition in the calorimeters consistent with minimum ionization. However, with the goal of identifying muons in jets, calorimetry information is less useful (the energy from a muon can not be sorted out from the energy of other particles in the jet), and the SLT uses only information provided by the muon chambers and the central tracker for muon identification. Moreover, leptons from b decays in $t\bar{t}$ events have a broad momentum spectrum, typically spanning a range from a few GeV/c to over 50 GeV/c. With this momentum spread, multiple scattering (MS) in the material of the CDF detector can cause a deflection in the muon path that ranges from about half a meter to a few millimeters. Any algorithm to identify these muons must, therefore, be designed to maintain the muon ID efficiency across a wide range of conditions.

The soft muon tagger is based on a χ^2 function. The idea is to combine all the available information about the SLT candidate and compare it with distributions obtained from a sample of muons derived from the data.

4.1 The SLT Algorithm

4.1.1 Taggable Tracks

The first step in the SLT algorithm [36] is determining whether a track is an SLT candidate, that is, if the track is "taggable." The algorithm starts with high-quality reconstructed tracks in the COT, selected by requiring at least 3 axial and 2 stereo superlayers that have at least 5 track hits. Some rejection for pion and kaon decays in flight is achieved by requiring that the impact parameter of the reconstructed track be less than 2 mm with respect to the beamline. The track is

also required to originate within 60 cm in z of the center of the detector. These requirements are summarized in Table 4.1. The algorithm determines the extrapolation of the track (not including

Variable	Cut
$\|d_0\|$	< 0.2 cm
$\|z_0\|$	< 60 cm
# Axial superlayers	≥ 3 with ≥ 5 hits
# Stereo superlayers	≥ 3 with ≥ 5 hits

Table 4.1: Summary of basic track-quality cuts set in the SLT algorithm.

multiple scattering) to the muon chambers. To take into account the fact that the muon may not have had enough momentum to reach the CMP or CMX detectors, the track p_T is compared against a range-out threshold of 3.0 GeV/c. Tracks that are above range-out and extrapolate to a fiducial volume (which extends to a distance $3\sigma_{\mathrm{MS}}$ outside the physical edges of the muon chambers) are considered taggable.

4.1.2 A χ^2 Function for Muon Identification

Part of muon ID at CDF is the matching of tracks extrapolating to the the muon chambers with track segments reconstructed in the muon chambers. Matching is done in the following observables ("matching variables"): extrapolated position along the muon chamber drift direction (x), the longitudinal coordinate along the chamber wires (z) when such information is available, and the extrapolated slope compared to the slope of the reconstructed muon stub (ϕ_L). Tracks are paired with stubs based on the best match in x for those stubs that are within 50 cm of an extrapolated track. In what follows, I refer to the difference between the extrapolated and measured positions in x and z as dx and dz, respectively, and the extrapolated and measured slope as dϕ_L. The distributions of these variables over an ensemble of events are referred to as the "matching distributions".

Candidate muons are selected with the SLT algorithm by constructing a global χ^2 quantity, L, based on a comparison of the measured matching variables with their expectations. The first step in constructing L is taking a sum, Q, of individual χ^2 variables:

$$Q = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2}, \tag{4.1}$$

where μ_i and σ_i are, respectively, the expected mean and width of the distribution of the matching variable X_i. The sum is taken over n selected variables as described below. The quantity, L, is then

constructed by normalizing Q according to

$$L = \frac{(Q - n)}{\sqrt{\mathrm{var}(Q)}}, \tag{4.2}$$

where the variance, $\mathrm{var}(Q)$ is calculated using the full covariance matrix for the selected variables. The normalization is chosen to make L independent of the number of variables n.

The selected variables are the full set of matching variables, x, z, ϕ_L in the CMU, CMP and CMX with the following two exceptions: The CMP chambers do not provide a measurement of the longitudinal coordinate z, and matching in ϕ_L is not included for stubs in the muon chambers that have only three hits. Because of their significantly poorer resolution, track segments reconstructed only in the CMU or only in the CMP chambers with only three hits are rejected (if the SLT candidate has stubs in both CMU *and* CMP, then a stub with only three hits is allowed). Note that a muon that traverses both the CMU and the CMP chambers yields two sets of matching measurements in x and ϕ_L and one z matching measurement, and is referred to as a CMUP muon. All available matching variables are used in the calculation of L for a given muon candidate.

4.2 Parametrization of the Muon Matching Distribution Widths

The quantity, L, is based on the comparison of a measured variable and its expected value. In order to measure the expected widths of the distributions of the matching variables in the data, samples of muons covering a broad p_T spectrum were used. The p_T range of ~ 1.5 GeV/c to ~ 10 GeV/c is covered using muons from J/ψ decays. The p_T range of ~ 20 GeV/c to ~ 50 GeV/c is covered using muons from Z^0 and $W \to \mu\nu$.

J/ψ candidates are selected by requiring two oppositely charged muons with stubs in the CMU, CMP, or CMX. The muon tracks must satisfy the track cuts listed in Table 4.1. To improve the resolution of the invariant mass, the tracks are also required to have silicon hits attached. Muon pairs whose invariant mass lie in the signal region between 3.02 GeV/c^2 and 3.15 GeV/c^2 are used to measure the matching distributions. Sidebands are defined in the ranges 2.80–2.95 GeV/c^2 and 3.20–3.35 GeV/c^2. For each matching distribution, the background contribution under the peak is subtracted by using the shape obtained from muons in the sideband regions.

Z candidates are also selected by requiring oppositely charged muon pairs. The tracks are not

required to have silicon hits attached, but are constrained to originate at the interaction point. At least one leg is required to have a stub in the CMU, CMP, or CMX. To increase the sample size, the second leg can either have a stub in those muon chambers, or pass minimum ionizing cuts: $E_{EM} < 2$ GeV and $E_{had} < 6$ GeV, where E_{EM} and E_{had} are the energies deposited in the electromagnetic and hadronic calorimeters, respectively. Only the legs with valid stubs are used to fill the muon matching variables whenever the invariant mass of the muon pair falls within the signal region between 75 GeV/c^2 and 105 GeV/c^2. No sideband subtraction is applied.

$W \to \mu\nu$ candidates require a high-p_T muon with a good quality track and large $\not\!\!E_t$ [37]. Due to the large QCD background minimum ionizing and isolation cuts have also been applied. The selection criteria for $W \to \mu\nu$ candidates are summarized in Table 4.2.

Variable	Cut
p_T	> 20 GeV
$\not\!\!E_t$	> 20 GeV
$\lvert z_0 \rvert$	< 60.0 cm
# Axial SL	≥ 3 with ≥ 7 hits
# Stereo SL	≥ 3 with ≥ 7 hits
$\lvert d_0 \rvert$	< 0.02 cm (silicon hits), < 0.2 cm (no silicon hits)
χ^2/ndf	< 2.0
E_{EM}	< 2+max(0, (P-100)·0.0115) GeV
E_{Had}	< 6+max(0, (P-100)·0.0280) GeV
$E_T^{\Delta R=0.4}/p_T$	< 0.1

Table 4.2: $W \to \mu\nu$ selection cuts used in the determination of the SLT tagger parametrization.

The mean values (μ_i in Equation 4.1) of the matching distributions are typically 0, except for a small offset in the CMU dz distribution. The widths are parametrized as a function of p_T as well as ϕ for the CMP and η for the CMX. These variables describe to first order the effects of multiple scattering in the detector. For the CMU, where the amount of material traversed by a muon is approximately independent of η and ϕ, the widths are parametrized by second-order polynomials in $1/p_T$ with an exponential term describing the low-p_T range. The box shape of the CMP creates a ϕ dependence on the widths of the matching distributions in addition to the p_T dependence on the effects of multiple scattering. There is an η dependence on multiple scattering in the CMX detectors due to variations in the amount of absorber between $\eta = 0.6$ and $\eta = 1.0$ (See Figure 2.7), which is included in the parametrization for CMX. The measured widths of the matching distributions as a function of p_T, overlaid with their fits, are shown in Figure 4.1. Details of the fit functions are summarized in Appendix A.

Some of the variables that compose the quantity, Q, are correlated (e.g. the CMU stub angle,

$\mathrm{d}\phi_L^{\mathrm{CMU}}$, is a good indicator of the position match in the CMP, $\mathrm{d}x^{\mathrm{CMP}}$), so the correlation coefficients have to be measured to normalize Q correctly. The covariance of two variables, x and y, in the quantity, Q, is:

$$\sigma^2_{x^2+y^2} = \sigma^2_{x^2} + \sigma^2_{y^2} + 2\rho\sigma_{x^2}\sigma_{y^2}, \tag{4.3}$$

where ρ is the correlation coefficient (with values between 0 and 1 for uncorrelated variables and $x = y$ respectively). After measuring $\sigma^2_{x^2+y^2}$, $\sigma^2_{x^2}$, and $\sigma^2_{y^2}$, Equation 4.3 is solved for ρ. The SLT tagger uses separate correlation coefficients, ρ^+ and ρ^-, for the cases $x \cdot y > 0$ and $x \cdot y < 0$, which increases the rejection of background because the correlation is really between signed values of x and y.

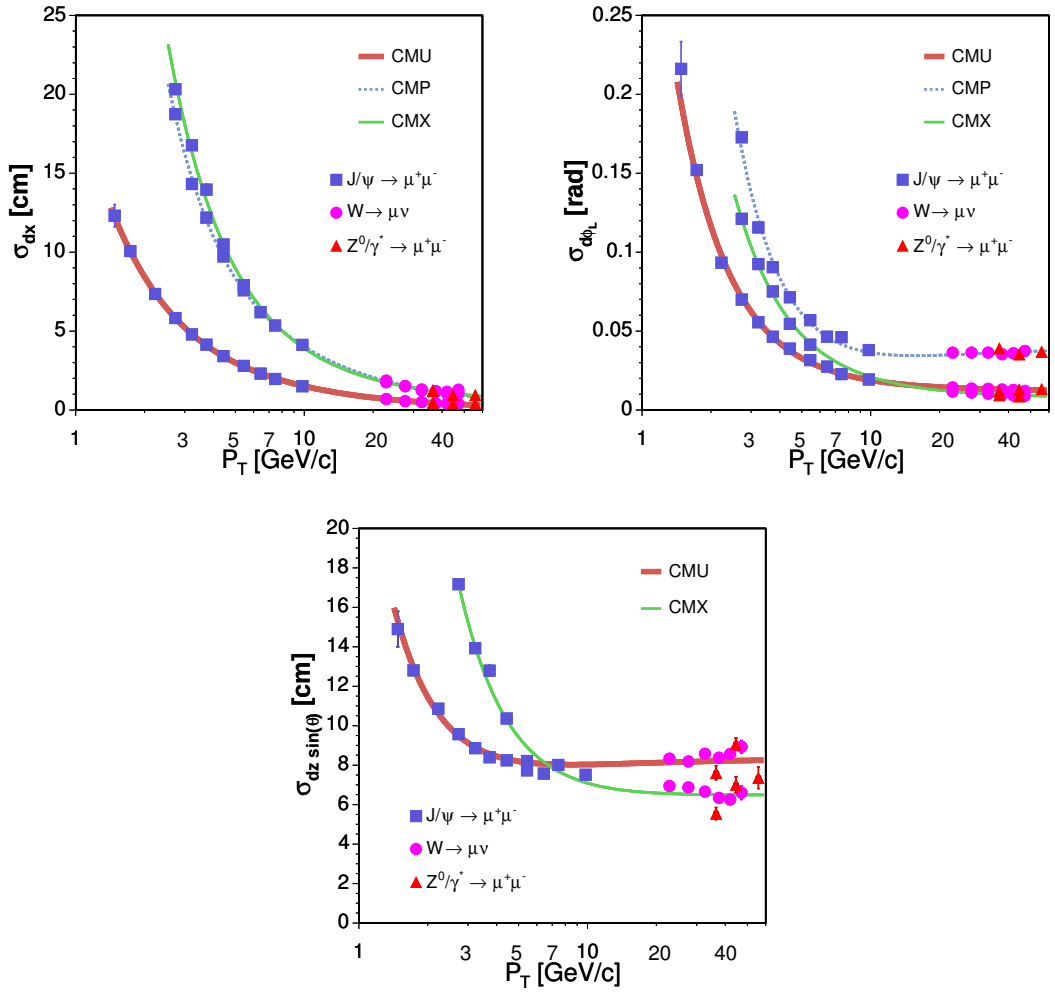


Figure 4.1: Width of the matching distributions ($\mathrm{d}x$, $\mathrm{d}z \times \sin\theta$, $\mathrm{d}\phi_L$) vs. p_T for each muon subsystem.

4.3 Efficiency of the Soft Muon Tagger

The overall efficiency of the tagger is measured using samples of J/ψ and Z^0 decays. For the J/ψs, samples triggered on a single muon are used and the tagger is applied to the non-trigger muon (probe leg). If both legs pass the trigger, only one of them is used. In the Z^0 sample, the tagger is similarly applied to the non-trigger leg or one of the two if both legs pass the muon trigger. To reduce the background in the Z^0 sample, the leg that is not used to measure the efficiency is also required to pass $E_{EM} < 2$ GeV, $E_{Had} < 6$ GeV and $E_T^{\Delta R=0.4}/p_T < 0.1$ where $E_T^{\Delta R=0.4}$ is the total transverse energy in a cone of $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ around the track. In other words, the leg must be minimum ionizing and isolated.

The efficiency of the SLT is defined as:

$$\epsilon = \frac{\text{Number of tagged muons}}{\text{Number of taggable muons with a stub}}, \tag{4.4}$$

where taggable is defined in Section 4.1.1. To decouple the tagging efficiency from the muon reconstruction efficiency, the muon is also required to have a stub in the subsystem where the efficiency is being measured. The efficiency is measured by computing the ratio of J/ψ and Z^0 signal events whose probe leg is SLT-tagged and those whose probe leg is SLT-taggable. The number of taggable/tagged signal events is obtained by subtracting background events from the J/ψ and Z^0 peaks, which is determined by a linear fit to the sideband regions. A tagged muon is defined as a taggable muon with $|L| < 3.5$. This cut on L is determined by measuring the muon tagging efficiency as a function of the cut; The particular value of 3.5 for the cut is chosen because the muon tagging efficiency plateaus at that point as seen in Figure 4.2.

Figures 4.3 and 4.4 show the SLT tagging efficiency in the central and CMX detectors as a function of p_T. The efficiency loss versus p_T in the central region is mainly due to the non-Gaussian tails of the $\mathrm{d}\phi_L$, and to a lesser extent the $\mathrm{d}x$, distributions of both CMU and CMP. Tagging in the CMX is more efficient because most of the stubs have 6 hits attached (as opposed to the maximum 4 hits in the CMU or CMP) allowing for better stub reconstruction. The efficiency in the central region is parametrized as a function of p_T with the following formula:

$$\epsilon_{\text{Central}} = a + b \cdot \ln p_T, \tag{4.5}$$

with $a = 87.95 \pm 0.23$ and $b = -5.97 \pm 0.11$. For the CMX, the efficiency is parametrized differently

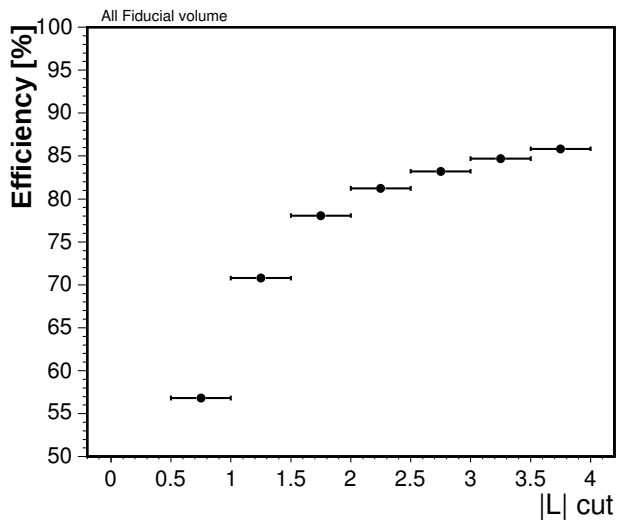


Figure 4.2: Efficiency of the SLT tagger as measured from J/ψ decays, as a function of the χ^2 quantity, L.

in two regions of the detector. For SLT candidates in CMX Arches, the efficiency is parametrized as:

$$\epsilon_{\text{Arch}} = \max(a + b \cdot p_T, a + b \cdot 70), \tag{4.6}$$

where $a = 92.67 \pm 0.18$ and $b = -0.09 \pm 0.01$. The efficiency in the CMX Miniskirt/Keystone regions is given as:

$$\epsilon_{\text{MK}} = a + b \cdot \ln p_T, \tag{4.7}$$

with $a = 103.73 \pm 0.94$ and $b = -12.70 \pm 0.46$.

The majority of muons from J/ψ and Z^0 decays are well isolated, but muons from the decays of B hadrons in $t\bar{t}$ events are not. Therefore, the possibility of efficiency dependence on the muon isolation has been checked. The number of tracks (N_{tr}) with $p_T > 1$ GeV/c surrounding the probe leg within a cone, $\Delta R = 0.4$ is taken as an indicator of the muon track isolation. The efficiency loss as a function if N_{tr} is measured after correcting for the efficiency dependence on muon p_T to remove the correlation between N_{tr} and p_T. Figure 4.5 shows the efficiency change (from the nominal value at $N_{tr} = 0$) as a function of N_{tr}. The efficiency change is fitted to a linear function with an intercept of (0.09 ± 0.10) percent and a slope of (-0.20 ± 0.15) percent per N_{tr}. Given the uncertainty on the fit, there is no sign of a dependence on track isolation.

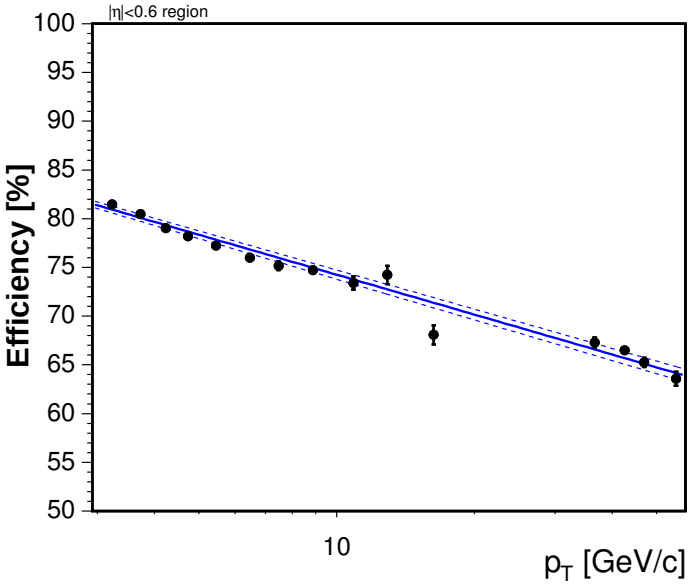


Figure 4.3: The SLT efficiency for CMU/CMP as a function of p_T as measured from J/ψ and Z^0 data for $|L|$ <3.5. The solid line is the fit to the data and the dashed lines indicate the uncertainty on the fit.

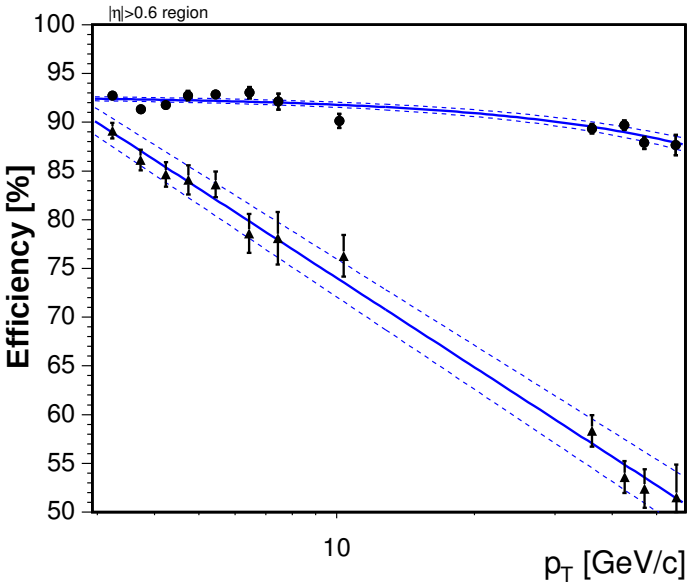


Figure 4.4: The SLT efficiency for CMX Arches (circles) and Miniskirt/Keystone (triangles) as a function of p_T as measured from J/ψ and Z^0 data for $|L|$ <3.5. The solid curves are the fits to the data and the dashed lines indicate the uncertainties on the fits.

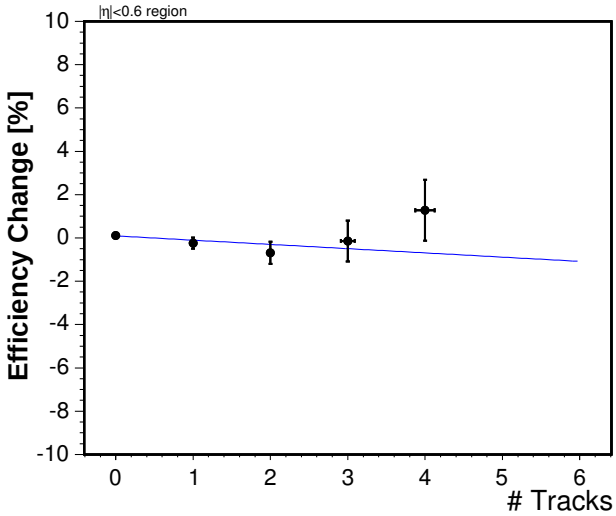


Figure 4.5: The SLT efficiency as a function of the number of tracks in a cone $\Delta R < 0.4$ surrounding the muon track for CMU/CMP. There is no loss in efficiency with increasing numbers of tracks.

Chapter 5

$t\bar{t}$ Event Selection and Acceptance

The goal of event selection is to select $t\bar{t}$ events while keeping contributions from other processes as small as possible. This analysis considers events consistent with $t\bar{t}$ events in which the W boson from one top quark decay has decayed leptonically (to an electron or muon) and the W from the second top quark has decayed hadronically. This event topology is known as the "lepton-plus-jets" channel because the signature of these events is a high-p_T lepton (also known as the primary lepton), large missing transverse energy due to the neutrino from the W decay, and several high-energy hadronic jets (ideally four jets: two from the hadronically decaying W and two from the b quarks in the event). This chapter details the selection criteria for the lepton-plus-jets events used in this analysis and the efficiency for $t\bar{t}$ events to pass this selection.

5.1 High-p_T Electrons

An electron is expected to leave a track in the COT and deposit most of its energy in a single tower of electromagnetic calorimeter. With all of its energy deposited in the calorimeter, the ratio of an electron's energy to its momentum, $E/p \approx 1$.[1] An electron should also have a well-defined transverse shower profile in the CES. Electrons from the decay of a W boson in $t\bar{t}$ events tend to have a large transverse momentum as well. Primary electrons are selected based on these considerations. Only electrons detected in the central calorimeter region ($|\eta| < 1$) are considered in this analysis.

5.1.1 Electron Triggers

The selection of events with a high-energy electron begins during data-taking with the online CDF trigger requirements (see Section 2.2.4). The trigger path for high-energy electrons is comprised

[1] The value E/p may be larger than 1 if the electron emits photons while traveling through the tracker. The emission reduces the electron's momentum, but since the photons tend to be collinear with the electron, the total energy deposited is the electron's original energy.

of three triggers, one at each level of the CDF trigger system.[2] At Level 1 events are selected by requiring a track with $p_T > 8$ GeV/c matched to an calorimeter tower with $E_T > 8$ GeV. The calorimeter tower is also required to have a hadronic to electromagnetic energy ratio, $E_{HAD}/E_{EM} < 0.125$. The calorimeter requirements at Level 2 are defined in terms of a cluster, which must have $E_T > 16$ GeV and be matched to the track found at Level 1. At Level 3 the electron candidate is built using the same clustering algorithm used offline and required to have $E_T > 18$ GeV. The cluster must also pass some minimal shower shape requirements and be matched to a track with $p_T > 9$ GeV/c.

5.1.2 Offline Electron Selection

Offline, selection requirements are made tighter than those used in the triggers during data collection. The requirements reduce the number of background events contained in the set of $t\bar{t}$ candidate events. Primary electrons are required to leave a track with $p_T > 10$ GeV/c that is well matched to an electromagnetic energy cluster with $E_T > 20$ GeV and hadronic to electromagnetic energy ratio $E_{HAD}/E_{EM} < 0.055 + 0.00045 \cdot E$. Additionally, the ratio of energy to momentum, E/p, must be less than 2.0 unless the track has $p_T > 50$ GeV/c. Electron tracks are required to have at least 5 hits on each of 3 axial superlayers and 2 stereo superlayers of the COT and $|z_0| < 60$ cm. The z_0 requirement incorporates as much of the interaction region as possible, while still keeping the track fiducial to the tracking detectors.

Matching between the track and the electromagnetic cluster is done by extrapolating the track to the shower maximum detector. To be considered a good match, the distance between the extrapolated track position and the shower maximum segment must be small. The distance, Δz, along the beam direction must satisfy $\Delta z < 3$ cm and the track charge, Q, times the distance, Δx, in the transverse direction must satisfy -3.0 cm$< Q \cdot \Delta x <$1.5 cm. The distribution of energy between the calorimeter towers in the cluster and the shape of the shower in the shower maximum detector are also required to be consistent with what is expected from an electron.

Electrons are required to be isolated as well. That is, the total energy in the towers surrounding the tower containing the electron shower is required to be less than 10 percent of the electron energy. This requirement may be written as $I \equiv \frac{E_T^{\Delta R<0.4} - E_T^e}{E_T^e} < 0.1$, where $E_T^{\Delta R<0.4}$ is the total transverse energy in a cone of $\Delta R < 0.4$ around the electron candidate and E_T^e is the transverse energy of the candidate.

[2]The trigger path for this dataset is `ELECTRON_CENTRAL_18` and the corresponding triggers are `L1_CEM8_PT8`, `L2_CEM16_PT8`, and `L3_ELECTRON_CENTRAL_18`.

Photons that interact with matter are capable of producing electron-positron pairs. These electrons, known as "conversion" electrons, can contaminate the sample when they are mistaken for the charged lepton from W boson decay. Events with conversion electrons are cut from the sample by removing events where the primary lepton has a partner track with an angular separation in the x-y plane less than 0.2 and a separation in z of less than 0.004 cm.

5.2 High-p_T Muons

A muon candidate is expected to leave a track in the COT that points to hits in the muon chambers. Muons are minimum ionizing particles, hence the energy deposited by a muon in the calorimeter is expected to be only a few GeV.

5.2.1 Muon Triggers

As for high-energy electrons, selection of events with a high-p_T muon begins during data-taking with the online CDF trigger requirements. There are two trigger paths used for high-p_T muons, depending on whether the muon candidate has stubs in the CMU and CMP, or in the CMX.[3] For CMUP muons, events are selected at Level 1 by requiring a track with $p_T > 4$ GeV/c matched to a stub in the CMU with $p_T > 6$ GeV/c and a stub in the CMP. At Level 2 the track p_T requirement is raised to 8 GeV/c. To pass the Level 3 requirement, there must be a track with $p_T > 18$ GeV/c matched to stubs in both the CMU and CMP detectors. The distance between the extrapolated track and the stub position, $|\mathrm{d}x|$ must be less than 10 cm for CMU stubs and less than 20 cm for CMP stubs. The triggers for CMX muons are similar, with Level 1 requiring a track with $p_T > 8$ GeV/c matched to a stub in the CMX and hits in the CSX detectors. At Level 2 the track p_T requirement is raised to 10 GeV/c. Level 3 requires a track with $p_T > 18$ GeV/c matched to a stub in the CMX detector with $|\mathrm{d}x| < 10$ cm.

5.2.2 Offline Muon Selection

Offline, primary muons are required to leave a track with $p_T > 20$ GeV/c. Energy deposition in the calorimeters must be consistent with minimum ionization; the electromagnetic energy must satisfy $E_{EM} < \max(2.0, 2.0 + 0.0115 \cdot (p - 100.0))$ GeV and the hadronic energy must pass $E_{HAD} < \max(6.0, 6.0 + 0.0280 \cdot (p - 100.0))$ GeV. The extrapolated track must be within $|\mathrm{d}x| < 7$ cm of a

[3] The trigger paths for this dataset are `MUON_CMUP18`—with corresponding triggers `L1_CMUP6_PT4`, `L2_CMUP6_PT8`, and `L3_MUON_CMUP_18`—and `MUON_CMX18`—with corresponding triggers `L1_CMX6_PT8`, `L2_CMX6_PT10`, and `L3_MUON_CMX_18`.

CMU stub and $|\mathrm{d}x| < 5$ cm of a CMP stub for CMUP muons or within $|\mathrm{d}x| < 6$ cm of a CMX stub for CMX muons.

Similar to electron selection, muon tracks are required to have at least 5 hits on each of 3 axial superlayers and 2 stereo superlayers of the COT and $|z_0| < 60$ cm. For CMX muons the track is required to have a COT exit radius greater than 140 cm to ensure that it passed through a sufficient number of COT superlayers. The impact parameter, d_0, of tracks without silicon hits attached must satisfy $|d_0| < 0.2$ cm while those with silicon hits must satisfy $|d_0| < 0.02$ cm. The cut on impact parameter helps with the rejection of cosmic rays that pass through the CDF detector, as well as muons from pions or kaons that decay in the tracking systems. These pion and kaon decays-in-flight (DIF) are also reduced by placing a cut on the χ^2 of the reconstructed track.

Primary muons are required to be isolated, just as primary electrons are. However, because muons leave little energy in the calorimeters, the isolation is determined differently. Muons must have an isolation $I \equiv \frac{E_T^{\Delta R<0.4}}{p_T^{\mu}} < 0.1$, where $E_T^{\Delta R<0.4}$ is the total transverse energy in a cone of $\Delta R < 0.4$ around the muon candidate and p_T^{μ} is the transverse momentum of the candidate.

Cosmic rays are detected by the CDF detector and are a source of background to the lepton-plus-jets dataset. Most cosmic rays pass through the CDF detector leaving two track segments in the muon detectors separated in ϕ by 180° and separated in time as the muon passes through the detector. Events with cosmic rays are vetoed by cutting on the $\Delta\phi$ between the reconstructed muon and any other muon stub found in the detector and considering the timing information obtained from the calorimeters and tracking detectors.

5.3 Jet Selection

The quarks and gluons produced in an event are not detected as single particles, but rather hadronize and manifest themselves in the detector as showers of neutral and charged particles that deposit their energy in the calorimeter. This shower of hadrons is known as a jet.

Jets are identified using an algorithm that groups the individual calorimeter energy deposits in neighboring towers into a cluster. The algorithm takes all towers with energy greater than 1 GeV as "seed" towers, around which a jet can be reconstructed. Calorimeter towers belonging to any electron candidate are not used. Clustering begins with the highest-E_T tower in a given event. All towers with $E_T > 1$ GeV surrounding that seed tower within a cone $\Delta R < 0.4$ are included in the cluster. Once all the neighboring towers are included, a new energy-weighted centroid is calculated

and the algorithm repeats using the centroid as the new geometric center of the cone until the centroid is stable. The process is repeated for all possible seed towers, in order of decreasing E_T, until all are used.

The energy in the towers of the jet must be corrected to get a good measure of the energy of the parton generating the jet. The response of the CDF detector to jets has been studied and correction factors measured that bring the average measured jet energy to the energy of the parton[38]. These factors make the calorimeter response uniform in η and also correct the absolute energy scale. This analysis counts jets with a corrected $E_T > 20$ GeV and $|\eta| < 2.0$ (called "tight" jets).

5.4 Missing Transverse Energy

The protons and antiprotons colliding at the Tevatron have zero momentum transverse to the beam, so the total vector sum of the transverse energy of each event should be close to zero[4] due to momentum conservation. In each event, the vector sum of the transverse energy in all calorimeter towers is taken and the difference between zero and this vector sum is defined as the missing transverse energy, $\not{E}_t$. Some $\not{E}_t$ is a result of energy mis-measurement from particles that escape the detector through cracks between calorimeter wedges, but for $t\bar{t}$ events in the lepton-plus-jets sample, a large amount of $\not{E}_t$ is expected from the neutrino, which passes through the detector without having its energy directly measured.

The $\not{E}_t$ measured in the event is corrected to adjust for the difference between the raw and corrected jet energies. This is done for all jets, which after correction, have $E_T > 8.0$ GeV and $|\eta| < 2.4$. If the primary lepton is a muon, the $\not{E}_t$ is also corrected by adding the muon momentum to the sum of calorimeter tower energies (corrected for the small amount of energy deposited by the muon).

5.5 Event Selection

5.5.1 Kinematic Selection

The candidate event sample is defined in terms of the objects described above. Exactly one high-p_T lepton (electron or muon), as defined in Sections 5.1 and 5.2 above, must exist in the event. There must also be at least 3 tight jets in the event (1- and 2-jet events are used as a control sample) and

[4]The individual partons may have some small transverse momentum, so the transverse momentum of an event may not equal zero exactly.

$\not\!\!E_t > 30$ GeV. Events in which the primary lepton and a second lepton candidate—of the same type and opposite charge—in the event form an invariant mass between 76 GeV/c^2 and 106 GeV/c^2 are rejected to reduce background from events where the primary lepton comes from the decay of a Z boson.

The total energy, H_T, of an event is defined as the scalar sum of the electron E_T or muon p_T, the event $\not\!\!E_t$ and the jet E_T for jets with $E_T > 8$ GeV and $|\eta| < 2.4$. Due to the large mass of the top quark, $t\bar{t}$ events are expected to have large H_T compared to the dominant background of the sample, $W+ \geq 3$-jets events, as illustrated in Figure 5.1. We therefore select events with $H_T > 200$ GeV as a first stage of background reduction, rejecting approximately 30 percent of the background while retaining more than 99 percent of the $t\bar{t}$ signal.

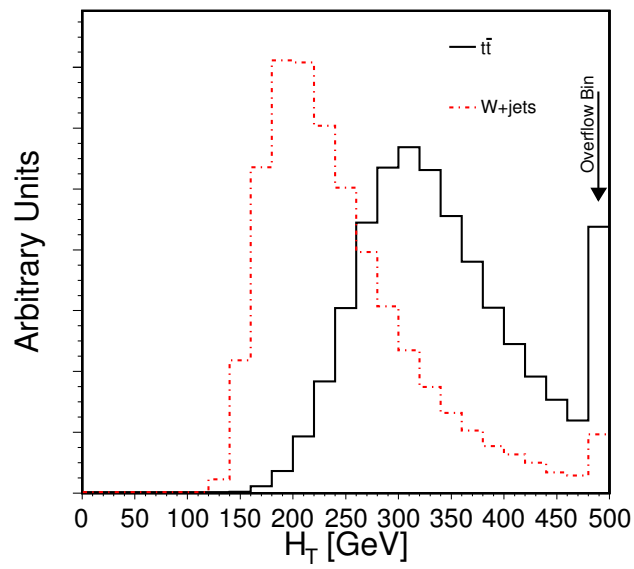


Figure 5.1: H_T distributions, normalized to unity, for $t\bar{t}$ and W+jets Monte Carlo events with three-or-more jets after the event selection described in the text.

5.5.2 SLT-Tag Requirements

Even after the H_T cut, the expected ratio of signal to background in the lepton-plus-three-or-more-jets sample is only of the order of 1:7. To further improve the signal to background ratio, events with one-or-more b-jets are identified by searching inside jets for semileptonic decays of b hadrons into muons. In other words, at least one jet in the event is required to be SLT-tagged (i.e. the jet contains an SLT muon within a cone, $\Delta R < 0.6$).

If the primary lepton is a muon, I remove events in which an SLT tag has opposite charge to the primary muon and together with that muon has an invariant mass between 8 and 11 GeV/c^2 or between 70 and 110 GeV/c^2. This rejects events in which an Υ or Z boson decays to a pair of muons, one of which becomes the primary lepton while the other ends up in a jet and is tagged by the SLT. Whether the primary is an electron or a muon, events where the invariant mass is less than 5 GeV/c^2 are also removed to prevent sequential double-semileptonic $b \rightarrow c \rightarrow s$ decays (where the the primary lepton and the SLT tag are from these semileptonic decays, rather than the primary lepton being from the decay of a W boson) from entering the sample, as well as events with a J/ψ decay.

Events with a Z boson in them may enter the lepton-plus-jets sample in another way, as well. If the Z decays to a pair of muons and one of the the muons subsequently radiates a high-energy photon ($E_T > 20$ GeV), the event will have a primary muon and what appears to be a jet containing a muon that can be SLT tagged. The invariant mass between the primary muon and the SLT muon is likely to fall outside of the Z-mass window and so not be removed by the above cuts. To reduce this background contribution events are removed in which the tagged jet has an electromagnetic energy fraction above 0.8 and just one track with $p_T > 1.0$ GeV/c within a cone of $\Delta R = 0.4$ about the jet axis. Modeling of these events is discussed in more detail in Section 7.4.

Three levels of selection are defined in this analysis. Events passing the kinematic cuts and the dilepton and radiative-Z vetoes, but do not necessarily have an SLT-taggable track in them, comprise the "pretag" sample. Pretag events that have an SLT-taggable track ($p_T > 3$ GeV/c, passing quality cuts, extrapolated to muon chambers) within $\Delta R < 0.6$ of tight jet are called taggable events. Finally, the subset of SLT-taggable events that have at least one SLT-tagged jet are called tagged events.

5.6 Data Sample

This analysis is based on an integrated luminosity of 2034±120 pb^{-1} [39] collected with the CDF detector (1993±117 pb^{-1} with the CMX detector operational). Table 5.1 shows the number of pretagged, taggable and tagged events in the electron and muon channels in this dataset as a function of jet multiplicity.

	H_T ≥0 GeV		H_T ≥200 GeV		
	1 jet	2 jets	3 jets	≥4 jets	≥3 jets
Electrons					
Pretag	79348	13068	1615	660	2275
SLT Taggable	43005	10479	1518	648	2166
SLT Tagged	519	224	85	64	149
CMUP Muons					
Pretag	38165	6320	719	325	1044
SLT Taggable	20162	4921	673	312	985
SLT Tagged	224	105	41	34	75
CMX Muons					
Pretag	23503	3672	422	162	584
SLT Taggable	12428	2864	396	160	556
SLT Tagged	149	55	16	8	24
CMUP+CMX Muons					
Pretag	61668	9992	1141	487	1628
SLT Taggable	32590	7785	1069	472	1541
SLT Tagged	373	160	57	42	99
Electrons+Muons					
Pretag	141016	23060	2756	1147	3903
SLT Taggable	75595	18264	2587	1120	3707
SLT Tagged	892	384	142	106	248

Table 5.1: Summary of event counts for 2 fb^{-1} of CDF Run II data for the event selection described in this chapter.

5.7 $t\bar{t}$ Acceptance and Tagging Efficiency

With the event selection defined, the efficiency for identifying $t\bar{t}$ events in the lepton-plus-jets channel can be determined. The efficiency is factorized into the geometric times kinematic acceptance and the SLT tagging efficiency. The acceptance includes all the cuts described in the preceding sections and is evaluated assuming a top mass of 175 GeV/c^2. The tagging efficiency is the efficiency for SLT-tagging at least one jet in events that pass the geometric and kinematic selection. Each piece is described below.

5.7.1 Geometric and Kinematic Acceptance

The acceptance of $t\bar{t}$ events in this analysis is measured in a combination of data and PYTHIA Monte Carlo simulations. The acceptance is measured in PYTHIA and then corrected for effects that are not sufficiently well modeled in the simulation: the lepton trigger efficiencies, the fraction of the $p\bar{p}$ luminous region well-contained in the CDF detector (i.e. the z-vertex cut efficiency), the difference between data and Monte Carlo efficiencies for track reconstruction and lepton identification. The efficiency of the z-vertex cut, $z_0 < 60$ cm, is measured to be 96.3 ± 0.2 percent. The correction factor for the different track reconstruction efficiencies is 1.014 ± 0.002.

The efficiencies for triggering on and identifying primary leptons are both measured (separately) in Z-boson decays acquired with the triggers described in Sections 5.1.1 and 5.2.1. The efficiencies are measured by first finding one fully reconstructed lepton in the data that passed the relevant trigger and then looking for a second, opposite-sign, lepton candidate of the same flavor that meets minimal kinematic and identification criteria. If the invariant mass of the leptons is in the range 81 GeV/c^2 $< M_{\ell\ell} <$ 101 GeV/c^2 (i.e. close to the Z resonance peak), the second lepton is checked to see if it passes all of the trigger/identification requirements. The efficiency is calculated as the ratio of the number passing the requirements to the total number of second leptons. The ratio of the lepton-identification efficiency in data to that in simulation is applied as a correction factor to the acceptance calculated from simulation. Events in the Monte Carlo are not required to pass any trigger, so the acceptance is multiplied by lepton trigger efficiency to take this into account. The measured lepton trigger and identification efficiencies and correction factors for each of the primary lepton types can be found in Table 5.2.

Quantity	Electron	CMUP Muon	CMX Muon
Trigger efficiency	0.966±0.005	0.917±0.005	0.925±0.007
Lepton ID efficiency (data)	0.789±0.004	0.829±0.006	0.893±0.006
Lepton ID efficiency (MC)	0.806±0.001	0.896±0.001	0.916±0.002
Lepton identification correction	0.978±0.005	0.926±0.007	0.975±0.007

Table 5.2: Summary of lepton trigger and identification efficiencies.

The raw acceptance is defined as the number of pretag events divided by the total number of $t\bar{t}$ events in the PYTHIA sample. The acceptance after correcting for the differences between data and simulation is shown in Table 5.3 as a function of the number of identified tight jets.

	1 jet	2 jets	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
Electron (%)	0.163±0.002	0.858±0.004	1.63±0.01	2.08±0.01	3.71±0.01
CMUP Muon (%)	0.088±0.001	0.472±0.003	0.909±0.004	1.142±0.005	2.05±0.01
CMX Muon (%)	0.042±0.001	0.219±0.002	0.414±0.003	0.532±0.003	0.946±0.004
Combined (%)	0.292±0.002	1.544±0.005	2.946±0.008	3.743±0.009	6.69±0.01

Table 5.3: Acceptance for $t\bar{t}$ events as a function of jet multiplicity from PYTHIA Monte Carlo sample, corrected for the data/MC ratio for tight lepton ID efficiencies and the primary lepton trigger efficiency as well as the z-vertex cut efficiency and track reconstruction efficiencies. In the combined acceptance, the fact that the CMX detector was not operating early in Run II is accounted for. The uncertainties listed are statistical only.

5.7.2 SLT Efficiency for $t\bar{t}$ events

The detector simulation does not properly reproduce the non-Gaussian tails of the muon matching distributions. Therefore, when evaluating the efficiency for tagging a $t\bar{t}$ event, the measured SLT efficiencies (see Section 4.3) are applied directly to a generated muon in the Monte Carlo sample. This is done by throwing a random number and comparing the number to the efficiency for tagging a muon with a given p_T. This accounts for the tagging of semileptonic decays in $t\bar{t}$ events. Events from $t\bar{t}$ can also be mistagged if a tag results from hadronic punch-through or a decay-in-flight. This effect is accounted for by allowing events to be tagged by charged tracks other than muons using the tagging probabilities from the mistag matrix described in Chapter 6. The overall efficiency for finding at least one SLT tag in a $t\bar{t}$ event is shown in Table 5.4. Mistags account for 21 percent of the $t\bar{t}$ tagging efficiency.

	1 jet	2 jets	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
Electron (%)	7.0±0.3	11.4±0.2	12.9±0.1	14.9±0.1	14.0±0.1
CMUP Muon (%)	5.6±0.3	10.7±0.2	11.8±0.1	14.1±0.1	13.1±0.1
CMX Muon (%)	6.7±0.5	11.2±0.3	12.3±0.2	14.2±0.2	13.4±0.2
Average (%)	6.5±0.2	11.2±0.1	12.5±0.1	14.6±0.1	13.6±0.1

Table 5.4: $t\bar{t}$ event tagging efficiency for SLT muons as a function of jet multiplicity from PYTHIA Monte Carlo sample. The lepton category refers to the primary lepton. The average tagging efficiency is determined by weighting each channel by the acceptances found in Table 5.3. The uncertainties listed are statistical only.

In the determination of the SLT-tagging efficiency, the track-reconstruction efficiency is taken directly from Monte Carlo simulation. The reconstruction efficiencies of muon chamber track segments are also taken from the simulation and scaled to the values measured in the data using the muon in Z-boson decays unbiased by the trigger.

As noted in Section 4.3, the SLT efficiency has been parametrized using muons that tend to be isolated from other activity. To further check that this efficiency measurement is representative of muons in or near jets, a high-purity $b\bar{b}$ sample, derived from events triggered on 8 GeV electrons or muons. These events are enriched in semileptonic b-hadron decays. Events are selected that have two tight jets, one of which is required to be within a cone $\Delta R < 0.4$ of the primary electron or muon (the "lepton jet"). For jets associated with muons, the energy is corrected to account for the muon p_T. The second jet (the "away jet") is chosen as the jet with maximum separation in azimuth ($\geq$ 2 radians) from the lepton jet. Both jets are required to have a secondary vertex reconstructed and tagged by the SecVtx algorithm. This selection results in a $b\bar{b}$ sample with a

purity of approximately 95 percent [40, 41]. The SLT acceptance times efficiency for semileptonic decays to muons in the away jet is measured in a PYTHIA dijet Monte Carlo sample. The Monte Carlo events are subject to the same event selection used for the $b\bar{b}$ data sample. The efficiency parametrization measured from the data is applied in the same way as in the $t\bar{t}$ Monte Carlo sample. The derived efficiency times acceptance per b-jet is applied to the data to predict the number of SLT tags in the away jet. The tags in the data are well-predicted using this technique. Thus it is concluded that the efficiency for SLT-tagging muons from semileptonic decays of heavy flavor in jets is well modeled by the procedure used to calculate the efficiency in $t\bar{t}$ events..

Chapter 6

Mistag Matrix

Tracks in light-quark jets can be mis-identified as muons from the semileptonic decay of heavy-flavor hadrons. This is typically due to hadrons that have not deposited all their energy in the calorimeter and reach the muon chambers (punch-throughs), or due to decays-in-flight (DIF) of hadrons to muons far from the interaction point.

6.1 Mistag Probability

6.1.1 Track-based definition

I define the mistag rate as the number of tracks tagged as muons, divided by the number of taggable tracks (see Section 4.1.1). A track is tagged if the χ^2 value returned by the SLT is $|L| \leq 3.5$.

6.1.2 Data samples

I measure the mistag probability using samples of kaons, pions, and protons. To identify kaons and pions I reconstruct $\mathrm{D}^* \to \mathrm{D}^0\pi \to K\pi\pi$ decays. This dataset is collected using a two-track trigger that requires two oppositely charged tracks with $p_T \geq 2$ GeV/c. The tracks are also required to have a scalar sum $p_{T1} + p_{T2} \geq 5.5$ GeV/c, an opening angle between them of $2^\circ \leq |\Delta\phi| \leq 90^\circ$, and originate from a displaced vertex.

A sample of protons is obtained by reconstructing $\Lambda \to p\pi$ decays. These events are collected using another two-track trigger similar to the one described above. There is a different opening angle requirement of $20^\circ \leq |\Delta\phi| \leq 135^\circ$ and the invariant mass of the track pair (assumed to be pions) is required to be 4 GeV/$c^2 \leq M(\pi, pi) \leq 7$ GeV/c^2.

Although I am using the kaons, pions, and protons to determine the background to $t\bar{t}$, for the purposes of this chapter they will be generally referred to as the signal. I do this because, in the determination of the mistag rate, these samples are what I am looking at and other particles, especially muons, are "background" that affect the determination of the mistag rate.

6.2 Event Reconstruction

In the reconstruction of D^* and Λ^0 decays[1] I apply the following track quality criteria:

- the number of axial superlayers with ≥ 5 hits is ≥ 3;
- the number of stereo superlayers with ≥ 5 hits is ≥ 2;
- the track has $|z_0| \leq 60$ cm.

6.2.1 Reconstruction of $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+, \mathrm{D}^0 \to K^-\pi^+$

To reconstruct D^* decays I examine the mass difference $\Delta m = m(K\pi\pi) - m(K\pi)$. In addition to the above track quality requirements I also apply the following selection criteria:

- the kaon must have opposite charge to each of the two pions;
- $|z_0| \leq 5$ cm between any two tracks;
- the soft pion from the $\mathrm{D}^* \to \mathrm{D}^0\pi$ decay must have $p_T \geq 0.5\mathrm{GeV}/c$;
- the kaon and pion from the D^0 decay must each have $p_T \geq 2\mathrm{GeV}/c$;
- they must have impact parameter, $|d_0| \leq 0.2$ cm;
- $|m(K\pi) - m(\mathrm{D}^0)| \leq 0.03\mathrm{GeV}/c^2$;
- at least one of these two must be SLT taggable (including having $p_T \geq 3\mathrm{GeV}/c$).

As shown in Figure 6.1, a clean D^* signal is obtained for the right sign Δm distribution.

Once a D^* candidate meeting the above criteria is found I look at the mass difference $\Delta m = m(K\pi\pi) - m(K\pi)$ of the candidate. If Δm falls within the D^* signal region (defined in Table 6.1) the SLT taggable tracks are used for determining the K/π tag rate, otherwise if it falls within the sideband region it is used for determining the tag rate in the sidebands. If multiple D^* candidates are reconstructed with a given taggable track, the one with Δm closest in value to $\Delta M = M_{\mathrm{D}^*} - M_{\mathrm{D}^0}$ is used in determining the status of that track.

[1] The selection criteria I use for D^* reconstruction are taken, with some modification, from [42]. The criteria for Λ^0 reconstruction are similarly taken from [43].

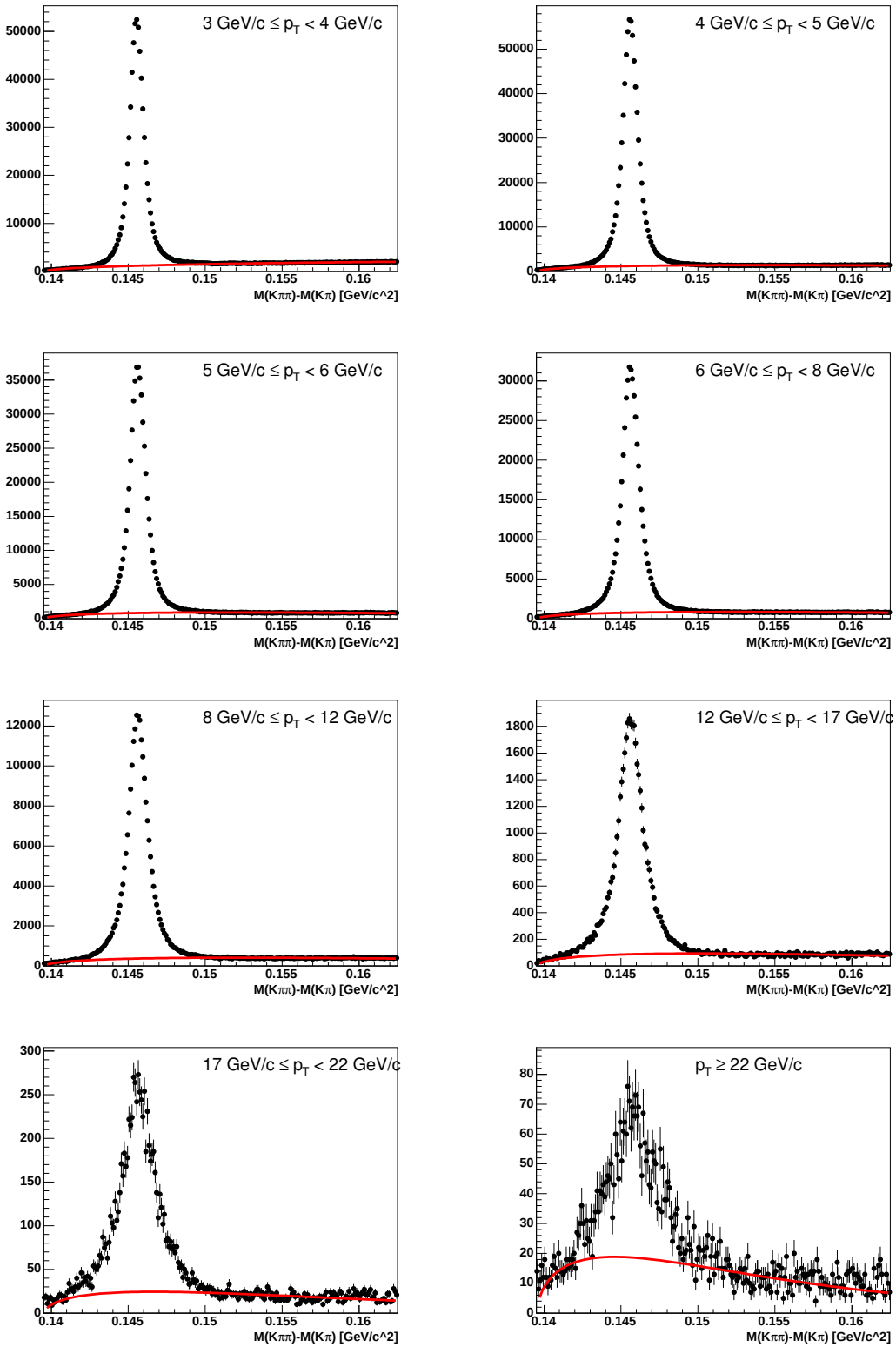


Figure 6.1: The $m(K\pi\pi) - m(k\pi)$ distribution for $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+, \mathrm{D}^0 \to K^-\pi^+$ candidates in different SLT-track-p_T bins. The red line in each plot represents the fit to the sideband regions.

Region	Mass (MeV/c^2)
D^* Signal	$142.421 < \Delta m < 148.421$
D^* Sidebands	$139.6 < \Delta m < 141$ or $152 < \Delta m < 162.5$
Λ^0 Signal	$1109.683 < m < 1121.683$
Λ^0 Sidebands	$1090 < m < 1105.683$ or $1125.683 < m < 1170$

Table 6.1: Mass windows used in determining the status of a D^* or Λ^0 candidate.

6.2.2 Reconstruction of $\Lambda^0 \to p\pi$

In the reconstruction of Λ decays I apply the following selection criteria:

- the pion and proton must have opposite charge;
- $|z_0| \leq 2$ cm between the two tracks;
- the χ^2 of the vertex fit must be ≤ 10;
- the vertex must have decay length $L_{xy} \geq 0.5$ cm;
- the proton p_T is greater than the pion p_T
- the pion must have $p_T \geq 0.4$GeV/c;
- the proton must have $|d_0| \leq 0.2$ cm;
- the proton must be SLT taggable (including having $p_T \geq 3$GeV/c).

Figure 6.2 shows the invariant mass distribution in the $p\pi$ mass hypothesis.

Whether the SLT taggable track is used in determining the tag rate from protons or in the sidebands is decided much like it was in the D^* case. If $m(p\pi)$ is in the Λ^0 mass window (as defined in Table 6.1) the track is considered a proton and if it falls within the sideband windows it is used for sideband subtraction.

6.3 Mistag Matrix

The mistag matrix is a parametrization of the tag rate that is used to predict the number of tags given a set of taggable tracks. I parametrize the rate as a function of track p_T and detector η. The mistag matrix has 8 bins in p_T and 9 in η. This makes the mistag matrix a 72-bin, two-dimensional

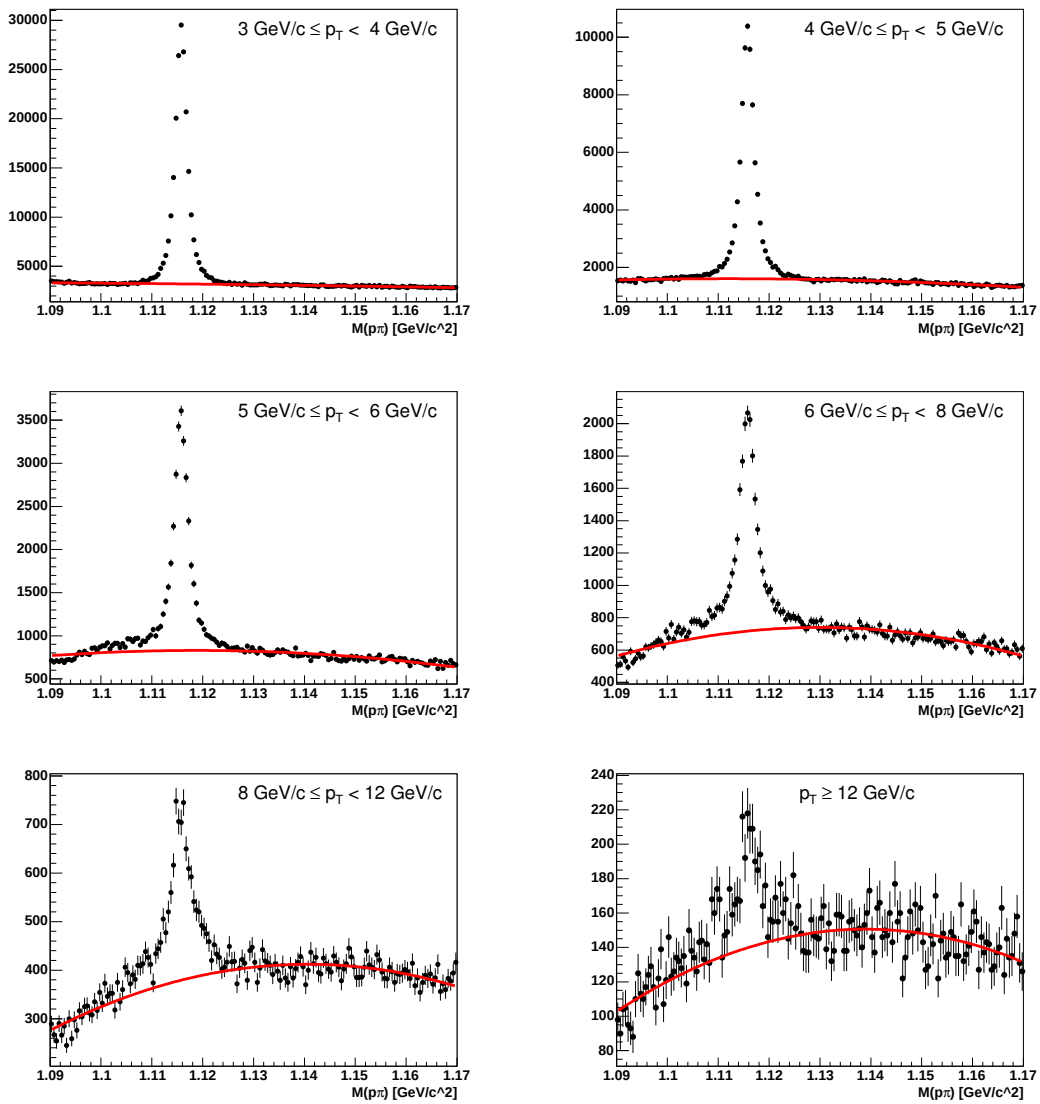


Figure 6.2: The $m(p\pi)$ distribution for $\Lambda^0 \to p\pi$ candidates in different SLT-track-p_T bins. The red line in each plot represents the fit to the sideband regions.

probability matrix, nominally defined as:

$$M_{ij} = \frac{N_{ij}^{\text{Tags}}}{N_{ij}^{\text{Taggables}}}, \tag{6.1}$$

where N_{ij}^{x} are the number of SLT tagged or taggable tracks in the ith p_T bin and jth η bin.

To calculate the matrix, I use samples of kaons, pions, and protons obtained in the manner described in Sec. 6.2. I start by counting, separately for each particle type, the number of tagged tracks and taggable tracks in each bin of the matrix and dividing. Once this is done, several corrections need to be made, and I need to weight the separate pion, kaon and proton tag probabilities to form a matrix for a generic track.

6.3.1 Sideband Subtraction

While these samples are fairly clean after the reconstruction requirements, they are not entirely pure and contain some background, in particular muons from the semileptonic decay of heavy-flavor hadrons may be present. The tag rate of this background is not the same as that of the signal as can be seen in Figure 6.3. The sideband regions contain various particles, in particular muons from the semileptonic decay of D^* mesons. In the signal region, however, the contribution to the tag rate from muons is suppressed so the tag rate is lower. I correct for this contamination by subtracting the sidebands. I determine the tag rate, R_S, of the K, π, or p as:

$$R_S = \frac{R_M - R_B \cdot f_B}{1 - f_B}. \tag{6.2}$$

where R_M and R_B are the tag rates measured in the signal mass window and the sideband regions, respectively, and f_B is the fraction of background under the signal peak. I determine f_B by fitting the sidebands to a background shape: a quadratic for the Λ^0 background and

$$a\sqrt{\Delta m - m_\pi} \cdot e^{-c(\Delta m - m_\pi)} \tag{6.3}$$

for the D^* background [42]. Once I have the background shape, f_B is the integral of that shape under the signal mass window divided by the number of D^* or Λ^0 candidates found in the window.

f_B varies with the SLT taggable track p_T, particularly for Λ^0, as can be seen in Figure 6.4. For this reason I do the sideband subtraction in various p_T regions that correspond to the binning of the matrix to get a better estimate of the tag rate as a function of p_T. Due to a small sample

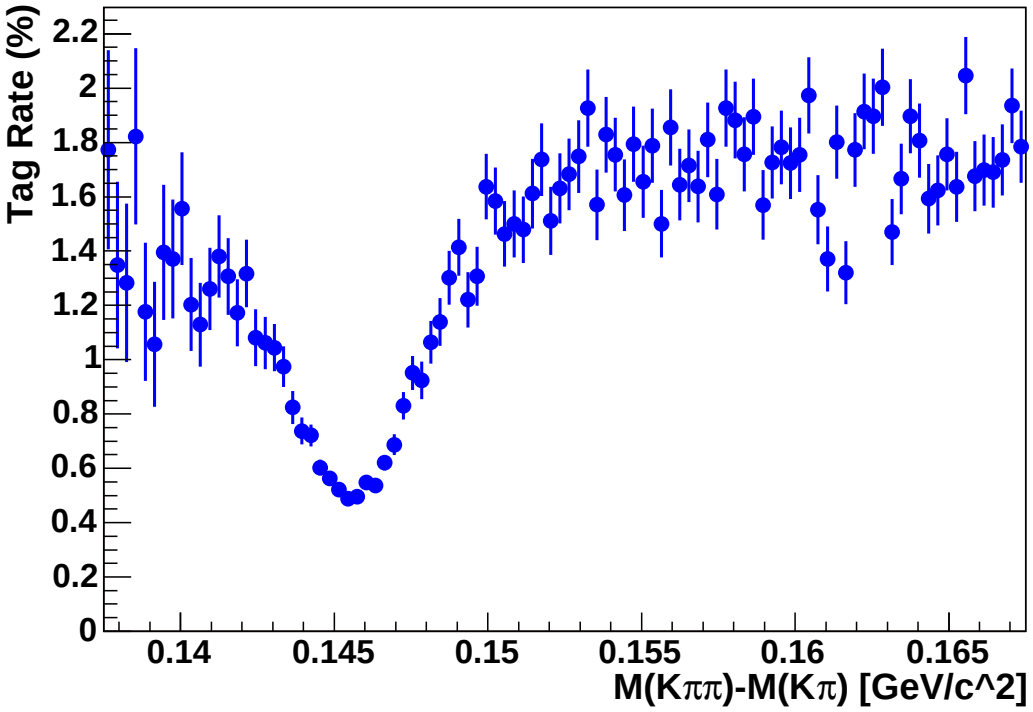


Figure 6.3: The tag rate as a function of $m(K\pi\pi) - m(k\pi)$. Outside of the D* signal region, which is dominated by kaons and pions, I see a higher tag rate probably due to muons from the semileptonic decay of heavy-flavor hadrons in the data sample.

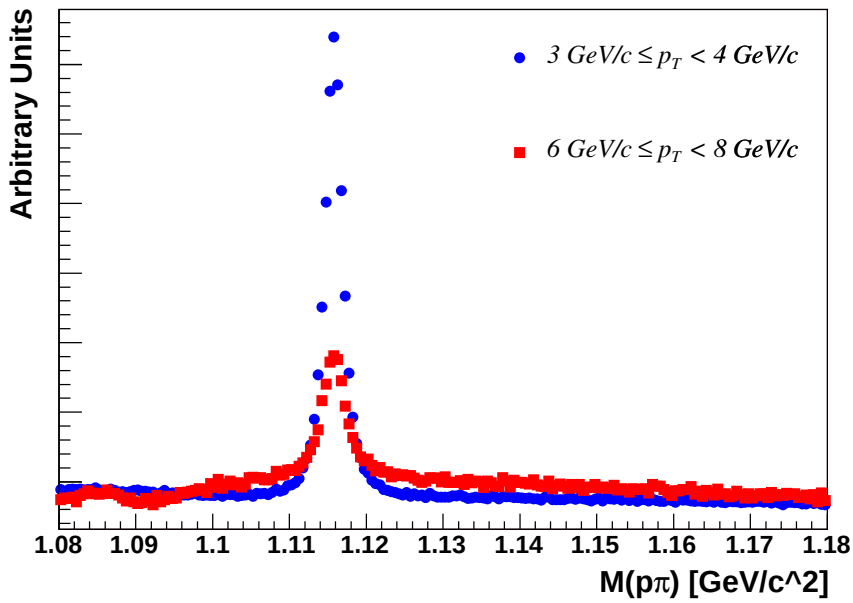


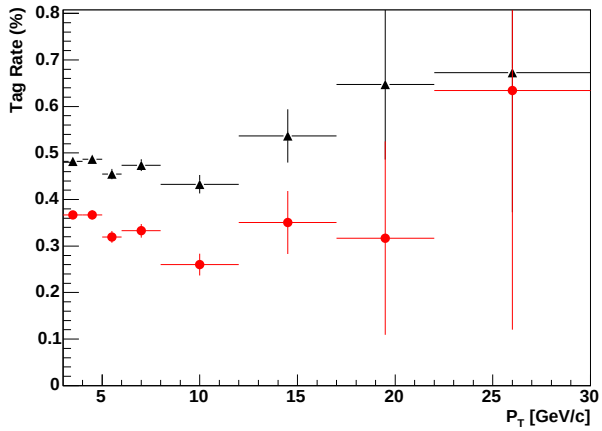
Figure 6.4: The $m(p\pi)$ distribution for $\Lambda^0 \rightarrow p\pi$ candidates in several SLT taggable track p_T bins. The background fraction, f_B, is basically a measure of the relative peak-to-sideband height ratio—the taller the peak compared to the sideband, the lower f_B. As the figure shows, this peak-to-sideband height ratio, and thus f_B, depends on the p_T of the taggable track.

size, I have a single bin above 12 GeV/c for Λ^0. Figure 6.5 shows the originally measured and sideband-subtracted tag rates vs. track p_T.

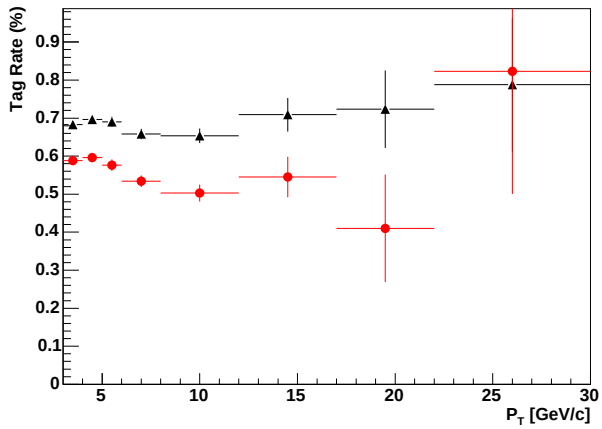
6.3.2 Decay-in-Flight Correction

Kaons and pions can sometimes decay in flight and the matrix is intended to predict the tag rate from these DIF as well as punch-throughs. Reconstructing D^* decays, however, creates a bias against DIF because DIF have a greater tendency to fail the reconstruction criteria than do punch-throughs. This bias is not present in W + jets where I apply the matrix. Since I do not expect the tag rate from DIF to be the same as from punch-through, this bias will affect the tag rate and a correction must be made.

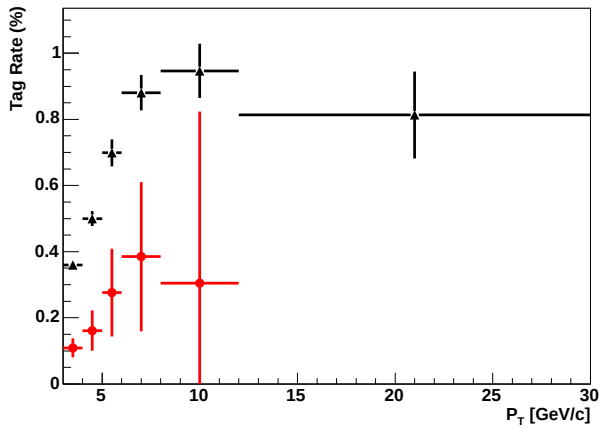
To determine what effect the D^* reconstruction has on the measured tag probability I go to the Monte Carlo. I use a sample with $B \rightarrow \mathrm{D}^*\pi, \mathrm{D}^* \rightarrow \mathrm{D}^0\pi, \mathrm{D}^0 \rightarrow K\pi$ for this purpose. First, I measure the effect of the reconstruction requirements by measuring the fraction of taggable tracks coming from DIF with all the D^* reconstruction requirements in place and again with just the standard SLT taggable track requirements. These quantities can be found in Table 6.2. The fraction of tracks



(a)



(b)



(c)

Figure 6.5: The measured (black triangles) and sideband-subtracted (red circles) tag rates as a function of track p_T for (a) pions, (b) kaons, and (c) protons.

	p_T [GeV/c]	Frac. DIF [%]	After Reco. [%]	DIF Tag rate [%]	R_T/R_m
π	3–4	0.40 ± 0.05	0.20 ± 0.04	44.6 ± 2.0	1.25 ± 0.08
	4–6	0.25 ± 0.04	0.16 ± 0.03	60.1 ± 2.3	1.16 ± 0.08
	> 6	0.20 ± 0.04	0.16 ± 0.04	75.6 ± 2.6	1.09 ± 0.14
K	3–4	0.99 ± 0.09	0.77 ± 0.08	10.2 ± 1.3	1.04 ± 0.02
	4–6	0.65 ± 0.06	0.51 ± 0.06	10.8 ± 1.3	1.02 ± 0.02
	> 6	0.39 ± 0.05	0.23 ± 0.04	18.2 ± 1.8	1.05 ± 0.02

Table 6.2: Relevant numbers in the determination of bias against DIF caused by reconstructing D* decays. These include the fraction of taggable tracks with DIF, the same fraction after reconstruction, and the tag rate of DIF. The final column gives the correction factor, R_T/R_m, applied the the measured rate.

coming from DIF is p_T dependent, so ideally I would like to measure the fraction in each p_T-bin of the matrix, but due to small size of this Monte Carlo sample, I only break the fraction into three bins: 3 GeV/$c \leq p_T <$ 4 GeV/c, 4 GeV/$c \leq p_T <$ 6 GeV/c, and $p_T \geq$ 6 GeV/c.

Once I have determined the fractions of DIF I must calculate how the bias affects the measured tag rate. Given a fraction, f_m, of DIF in the matrix (*i.e.* including the D* reconstruction requirements) I can determine the measured tag rate in the matrix, R_m, as:

$$R_m = f_m \cdot R_{DIF} + (1 - f_m) \cdot R_{PT}, \tag{6.4}$$

where R_{DIF} and R_{PT} are the tag rates of DIF and punch-throughs, respectively. The true tag rate, R_T, that would be measured without the reconstruction requirements can similarly be found by substituting the true fraction of DIF, f_T, in the sample for f_m in Equation 6.4 above. With some algebraic manipulation to the two equations I find

$$\frac{R_T}{R_m} = \frac{1 - f_T}{1 - f_m} + \frac{f_T - f_m}{1 - f_m} \cdot \frac{R_{DIF}}{R_m}. \tag{6.5}$$

Equation 6.5 gives the correction factor that must be applied to the measured tag rate to obtain the true rate.

The only piece remaining to be determined is the tag rate, R_{DIF}, of taggable tracks that are DIF, which is also measured in the Monte Carlo. These numbers, along with the resulting correction factors can be found in Table 6.2.

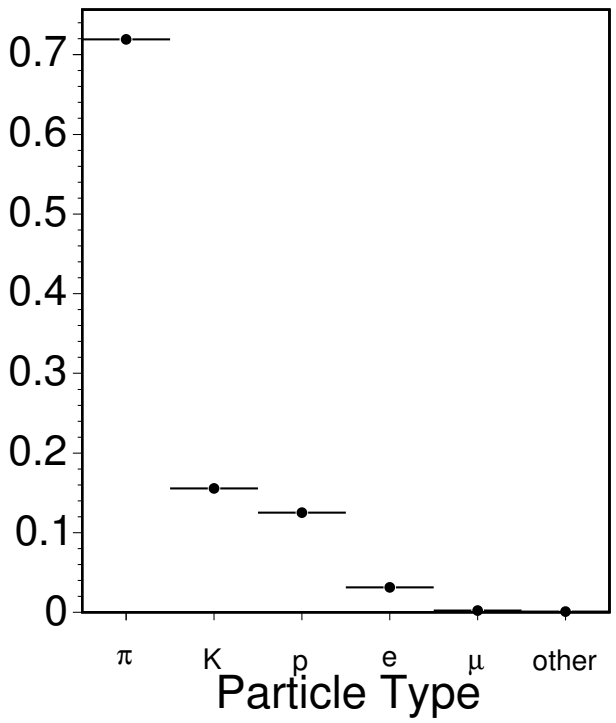


Figure 6.6: The relative abundances of taggable tracks of various particle types as found in W+jets Monte Carlo. The taggable tracks are mostly pions with some kaons and protons as well. Other types particles are not commonly found among the taggable tracks.

6.3.3 Finalizing the Matrix

At this point I have measured the tag rates and built matrices for kaons, pions, and protons separately and corrected for sideband contributions and a bias against DIF. The final step is to combine these matrices in a reasonable way. I combine the matrices by weighting them according to the relative abundance of each particle in Alpgen W + jets events. Thus the final matrix is

$$M_{ij} = W_\pi \cdot M^\pi_{ij} + W_K \cdot M^K_{ij} + W_p \cdot M^p_{ij}. \tag{6.6}$$

The weights, W, are estimated using W+jets MC. I use the Monte Carlo truth information to determine what type of particle is associated with the SLT-taggable track. The relative abundances of each particle type are then used as weights in the combined matrix. The results can be found in Figure 6.6 and Table 6.3.

The combined matrix has an average mistag rate of (0.41 ± 0.01) percent. Figure 6.7 shows the final tag rate vs. track p_T and detector η binned as the matrix.

Mistag Type	Matrix Weight [%]
π	71.9 ± 0.1
K	15.6 ± 0.1
p	12.5 ± 0.1

Table 6.3: Weights given to each separate particle matrix when combining them into one final matrix. These weights are determined from W+jets Monte Carlo.

In a previous version of this measurement [34, 35], the tags in γ+jets events were used to model the mistag rate (see Appendix B). The estimates of that method and the one presented here are not directly comparable due to the inclusion of heavy-flavor tags in the other method. However, at that time, the heavy-flavor content of the matrix was estimated and from that, a light-flavor tag probability. The light-flavor tag rate estimated with the previous matrix is $(0.42 \pm 0.02 \pm 0.01)$ percent, remarkably close to the mistag rate described above.

6.4 Prediction of the Mistag Matrix and the Systematic Uncertainty

To test the accuracy of the mistag matrix prediction, the matrix is applied to tracks in generic jets, in data, similar to those found in W+jets events. The prediction from the matrix is then compared to the number of tags observed in these jets. If the comparison is good, this test validates the measurement of the mistag probability. Otherwise, improvements can be made to the measured mistag probability until it predicts the tag rate in generic jets to a satisfactory level.

I check the mistag matrix using several jet samples by comparing the number of predicted tags to the number of measured tags. In using these jet samples several things must be considered:

- the mistag rate is connected with the way jets are selected and
- all jet samples contain some heavy flavor, which the matrix is not designed to predict.

These considerations will be discussed in the following sections as well as a description of the jet samples used for the test and the results.

6.4.1 Tracks in a Jet

In addition to the taggable track requirements outlined in Section 4.1.1, I require tracks to be within $\Delta R < 0.6$ of a tight-jet axis (see Section 5.3). This condition of tracks in a jet undermines the idea

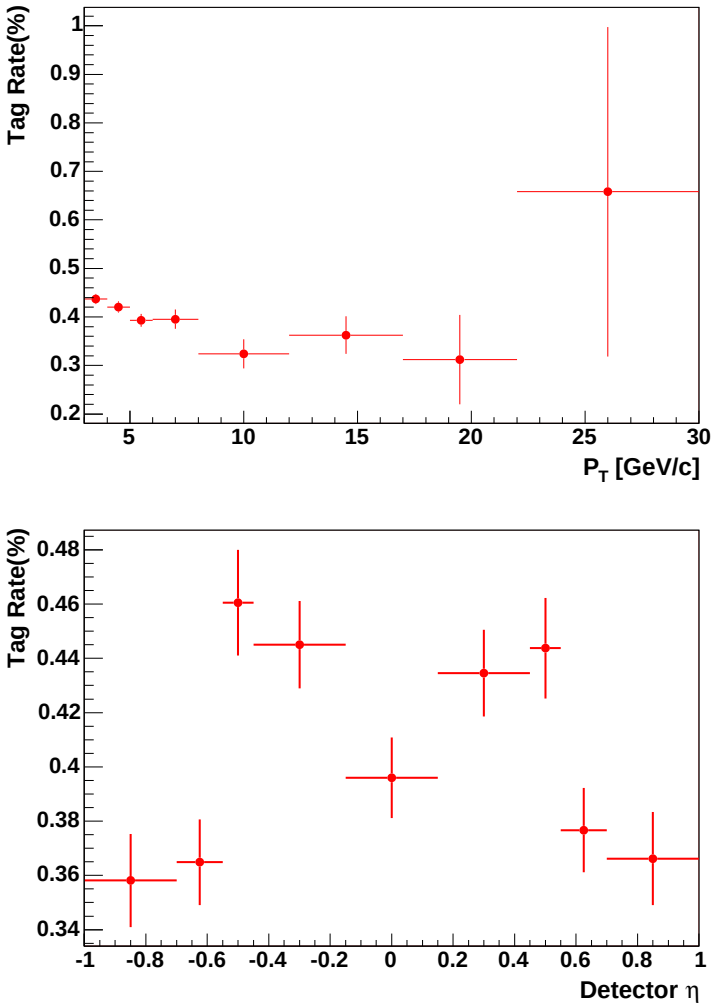


Figure 6.7: The mistag rate as a function of track p_T and detector η. The histogram binning matches that of the matrix.

of a purely track-based mistag probability. Since the measured energy of the jets depends on the presence of calorimeter gaps, as well as real muons, low interacting hadrons etc., which all are correlated with the probability of finding a SLT tag in that jet, the mistag rate is connected with the way the jets are selected. That is, the mistag rate is not an invariant number under selection of jets. In Sections 6.4.2–6.4.2 below I describe the selection of jets similar to those present in W+jets events.

6.4.2 Test Samples

Jet triggered samples

I test the matrix against several generic jets samples triggered on a single jet with (uncorrected) E_T greater than 20, 50, 70, or 100 GeV (called Jet20/50/70/100 for short). In these samples I first remove the trigger bias of the leading (trigger) jet by declaring tracks in that jet not taggable unless the second-highest-energy jet has uncorrected energy above the trigger threshold. Typically the trigger jet has a lower mistag rate than other jets in the event because the fact that it passed the high-E_T trigger means it is less likely to contain a any punch-throughs or DIFs that do not deposit all their energy in the calorimeter.

I also take care to avoid a second bias arising from dijet balancing; if the trigger jet passed a 100 GeV energy threshold, a 20 GeV jet opposite to it is highly biased (having probably been heavily mis-measured). This bias is discussed in greater detail in Section 6.4.3. In the Jet50/70/100 samples (Jet20 does not suffer significantly from this bias) I remove the bias in two distinct ways depending on the number of jets. For the trigger plus one jet case (dijets), I raise the energy threshold of taggable jets from 20 GeV to the sample's trigger threshold. For higher jet multiplicities, I require tracks to be away from the trigger jet and balancing jet, i.e. the separation, ΔR, between the taggable track and the trigger-jet axis must be between 0.7 and 2.6. Section 6.4.3 explains the selection of these cuts.

γ+jets sample

The photon sample is triggered on one photon with energy greater than 25, 50, or 70 GeV depending on the specific trigger, and contains typically one or more jets. Though it contains a well identified trigger object like the W plus jets sample, the trigger threshold creates a bias from dijet balancing. I avoid this bias just as I do in the jet samples by raising the jet-energy threshold for γ+1 jet events and require tracks to be away from the photon and balancing jet for higher jet multiplicities.

ΣE_T sample

The ΣE_T is a generic jet sample triggered on four jets with total measured E_T of 100 GeV. I check the measured prediction against measured tags in the cases of 2/3/4-or-more tight jets in the event. The 2-jets case is still largely biased, since the total 100 GeV energy is made of only two jets plus two soft (below 20 GeV) jets. A jet that is just above 20 GeV, out of the two tight jets, is balanced by a jet whose energy is the remaining up to 100 GeV. Therefore, the lower energy jet is most likely mis-measured, and with a higher mistag rate. This bias is reduced when requesting more jets in the event. I do not remove any bias in the ΣE_T sample.

6.4.3 The Mistag Rate and Dijet Balancing

As mentioned above, the mistag probability depends on the sample of jets in which the probability is measured. If I collect a sample of jets whose energy has been badly mis-measured and determine the mistag probability for tracks in those jets, it should not surprise us that (due to the correlation between calorimeter gaps, low interacting hadrons, real muons and SLT tags) the mistag rate differs somewhat from that of a more randomly chosen set of jets. To illustrate this effect, consider the case of dijets in the Jet100 sample. The trigger jet has energy above 100 GeV by definition. The away jet can have energy in the whole range between the tight jet definition (20 GeV) up to well above 100 GeV. Since they are balancing jets with energy above 100 GeV, the set of away jets with energy around 20 GeV is highly enriched in mis-measured energy jets. The mistag rate consequently starts rising, up to several times the normal rate, below the trigger threshold. Therefore, below the trigger threshold there is an excess of mis-measured jets that is not seen in W+jets events, where the trigger object is a 20 GeV lepton. A certain amount of mis-measured jets is obviously allowed and present in that sample, so I do not want to remove all mis-measured jets, but rather want to obtain a set of jets that resembles those in W+jets events.

The need for a change in the jet-energy threshold for dijet events has now been explained, leaving the question of to what it should be changed. Since the goal is to select a group of jets that are similar to those in W+jets events, particularly in regards to mistag probability, the cut should not be placed arbitrarily. Figure 6.8 shows the tag rate versus the difference between corrected jet E_T and trigger E_T in various jet samples. This figure illustrates the similar trends in each sample after adjusting for trigger thresholds (I consider 20 GeV to be the effective threshold of the W boson in W+jets events with a taggable-jet threshold of 20 GeV), including the similarity of all the test samples with the W+jets sample above the trigger threshold. From this I conclude that setting a

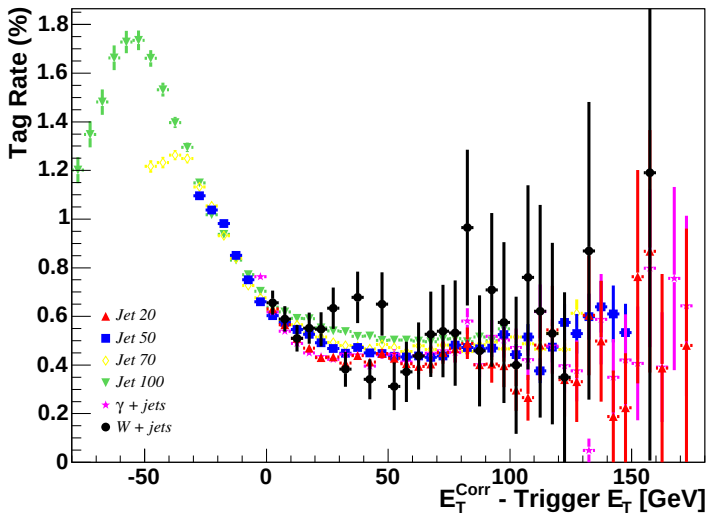


Figure 6.8: The SLT tag rate versus the difference between a jet's corrected E_T and trigger E_T in trigger plus 1-jet events in various samples. For W+jets events I use an effective trigger threshold of 20 GeV.

jet-energy threshold for dijet events at the trigger threshold of each sample leaves us with jets that are similar to those in W+jets events.

6.4.4 Heavy Flavor

All of the jet samples contain some heavy flavor for which the mistag matrix does not account. Since approximately 20 percent of measured tags (in γ+jets [35], see also Appendix B; this may be different in other samples) are from heavy flavor (HF) contributions, this means the mistag prediction will be significantly low if the samples are taken as is. Ideally, I would like to have a heavy-flavor-free sample on which to test, but no such sample exists and I must attempt to reduce the HF contribution in the samples I have. The logical place to start is to require that there be no secondary vertex (SecVtx) tags in the event. Alone, however, this requirement is not enough as there is still too much HF left in the samples since the SecVtx tagging efficiency is only about 50 percent. Therefore, I also make a requirement that all jets have a vertex mass below 0.3 GeV/c^2 for the few jets which have a vertex but were not SecVtx tagged, and finally, I set a cut on the d_0 significance (d_0/σ_{d_0}) of tracks in jets.

Figure 6.9 shows that there is a difference between the d_0 significance of tracks found in heavy-flavor jets and those in light-flavor jets. This difference is exploited to reduce the heavy-flavor content of the jet samples. I use jet MC to determine the optimal position to set a cut. Events in

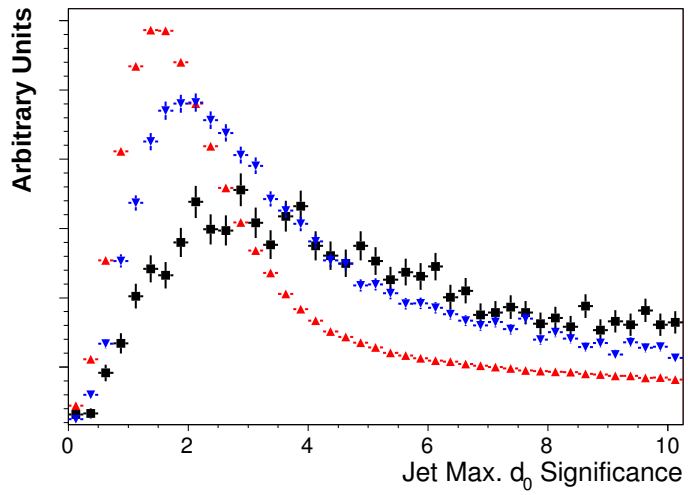


Figure 6.9: Distribution of the maximum d_0 significance of tracks in light-flavor (upward triangles), c (downward triangles), and b (squares) jets in simulation. Heavy-flavor jets contain tracks with larger d_0 significance than those in light-flavor jets.

which a jet contains a track with $|d_0 \text{ sig.}| \geq 2$ are removed because this point is about where the HF content of the sample is minimized, as can be seen in Figure 6.10. As this cut is track-based, it has the potential to create a bias against taggable tracks that decay in flight. I check this possibility in Monte Carlo and find that no bias is created based on Figure 6.11.

6.4.5 Determination of the systematic uncertainty

I assign a systematic uncertainty to the mistag prediction based on the difference between observed tags and the prediction of the matrix. After making the cuts to reduce the HF and to avoid trigger biases I suffer from a lack of events so I combine the samples arguing that, after removing biases and heavy-flavor jets, two jets of the same energy should have the same properties with respect to the mistag rate regardless of which sample they came from or the number of jets in the event. I then check the matrix prediction in various jet-E_T bins to determine the mistag systematic uncertainty. The results of these checks can be found in Table 6.4, which is visually represented in Figure 6.12. Table 6.5 has the same results broken up by jet sample.

The mistag matrix is used in the cross-section measurement to predict the W+jets mistag background in $W+ \geq 3$ jets. So it is appropriate to use that E_T spectrum to weight the results found in Table 6.4 to determine an overall uncertainty on the prediction. The jet E_T spectrum of W+jets events is shown in Figure 6.13 and the corresponding weights may be found in Table 6.6. I obtain a

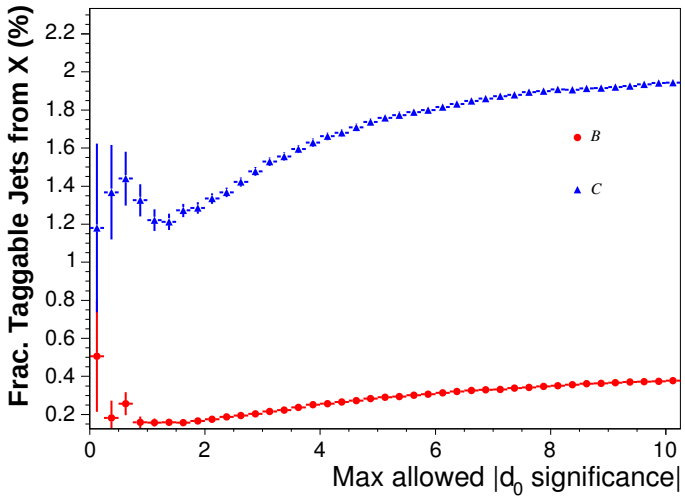


Figure 6.10: Fraction of taggable b or c jets out of all taggable jets passing versus the maximum allowed $|d_0$ significance$|$.

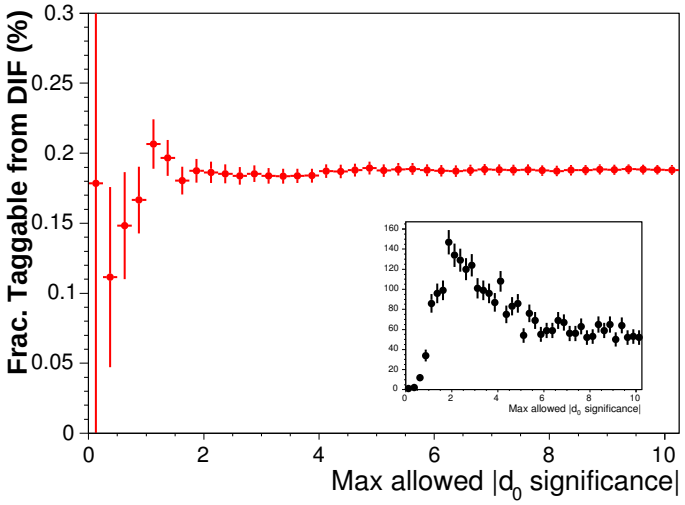


Figure 6.11: Fraction of taggable tracks, in Monte Carlo, with a DIF out of all taggable tracks passing versus the maximum allowed $|d_0$ significance$|$. The inset plots the number of taggable tracks with a DIF versus the maximum allowed $|d_0$ significance$|$.

Jet $E_T^{corr.}$ [GeV]	Observed	Predicted	Δ [%]
20-30	1892	1641 ± 29	-15.3 ± 3.3
30-45	1561	1693 ± 45	7.8 ± 3.4
45-65	701	768 ± 46	8.7 ± 6.4
65-90	464	462 ± 46	-0.5 ± 11.0
≥90	466	466 ± 76	-0.1 ± 16.9
≥20	5084	5029 ± 219	-1.1 ± 4.6

Table 6.4: Checks of the mistag matrix in different jet $E_T^{corr.}$ bins. The Δ =(Pred.-Obs.)/Pred. is determined in percent of the predicted numbers, the uncertainty is derived from the Poisson fluctuation of the observed numbers. These values of Δ are weighted using the W+3-or-more jets distribution (Table 6.6) to determine a systematic uncertainty on the mistag prediction.

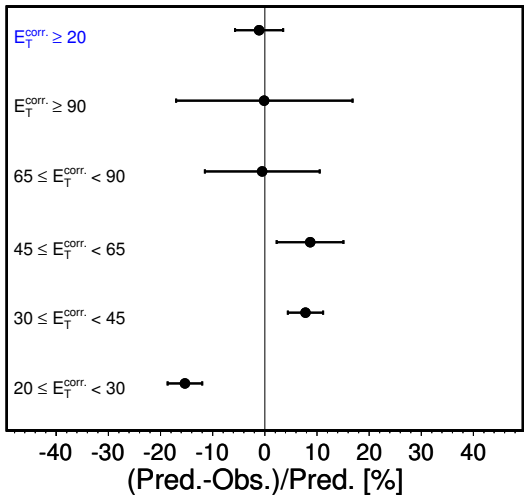


Figure 6.12: Summary of the SLT mistag matrix prediction compared to the observed tags in various jet-$E_T^{corr.}$ bins.

Sample	Observed	Predicted	Δ [%]
Jet 20	1525	1431 ± 29	-6.6 ± 3.5
Jet 50	375	369 ± 31	-1.5 ± 10.0
Jet 70	281	268 ± 33	-5.0 ± 14.4
Jet 100	240	225 ± 39	-6.6 ± 19.8
γ + Jets	2435	2496 ± 89	2.4 ± 4.0
ΣE_T	228	240 ± 9	4.9 ± 7.3
Combined	5084	5029 ± 219	-1.1 ± 4.6

Table 6.5: Checks of the mistag matrix with different jet samples. The Δ =(Pred.-Obs.)/Pred. is determined in percent of the predicted numbers, the uncertainty is derived from the Poisson fluctuation of the observed numbers.

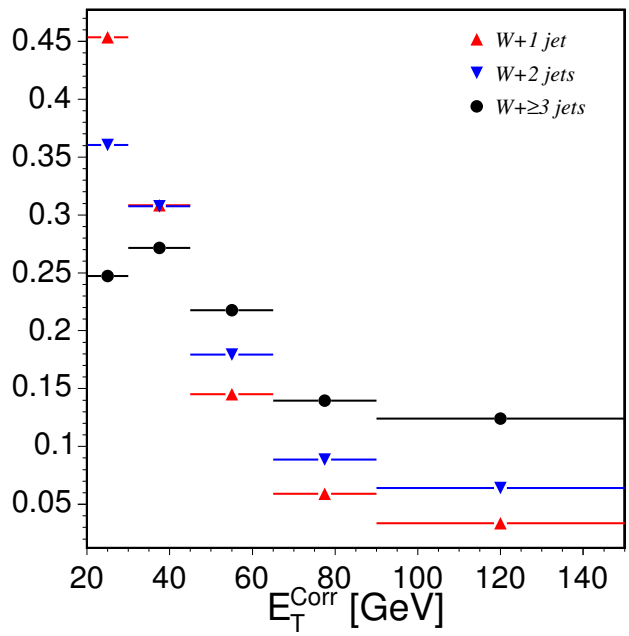


Figure 6.13: Corrected-E_T distribution for SLT taggable jets in ALPGEN W+jets events. I weight the jet E_T bins according to the distribution in W+3-or-more jets in determining the systematic uncertainty of the mistag matrix.

final weighted value for Δ =(Pred. - Obs.)/Pred. of (0.1 $\pm$ 4.4) percent with this method. Given the statistical uncertainty on Δ I assign a 5 percent systematic uncertainty on the matrix prediction.

6.4.6 Double tagged jets: a cross check

Unlike other b-taggers, the SLT does not identify whole jets as coming from the hadronization of b quarks, but rather identifies single tracks within a jet as being a muon track. This means that a single jet may contain multiple SLT tags. I use this fact and take the opportunity to look at the mistag prediction in a semi-independent sample from the one described above.

Jet $E_T^{corr.}$ [GeV]	Fraction of Jets [%]
20-30	25.1 $\pm$ 0.1
30-45	27.3 $\pm$ 0.1
45-65	22.1 $\pm$ 0.1
65-90	14.0 $\pm$ 0.1
$\geq$90	11.5 $\pm$ 0.1

Table 6.6: Fraction of taggable jets in W+3-or-more jets events that fall within a given E_T range. These fractions are used as weights when determining the systematic uncertainty.

Sample	Observed	Predicted	Δ [%]
Jet 20	528	427 $\pm$ 8	-23.6 $\pm$ 5.9
Jet 50	757	724 $\pm$ 33	-3.1 $\pm$ 6.0
Jet 70	1134	1093 $\pm$ 70	-3.7 $\pm$ 7.3
Jet 100	1927	1864$\pm$ 158	-3.4 $\pm$ 9.1
γ + Jets	1050	970 $\pm$ 42	-8.3 $\pm$ 5.7
ΣE_T	3574	2902$\pm$ 94	-23.2 $\pm$ 4.5

Table 6.7: Checks of the mistag matrix using second SLT tags in different jet samples. The Δ =(Pred.-Obs.)/Pred. is determined in percent of the predicted numbers, the uncertainty is derived from the Poisson fluctuation of the observed numbers.

For this method, I use the same jet samples described above and continue to avoid the biases from trigger thresholds and dijet balancing, but I do not make any cuts on SecVtx tags, vertex mass, or d_0 significance to reduce HF in the samples. Instead, I look at jets with at least one SLT tag and ask if I find others. Many of these jets were removed by the cuts in the primary systematic estimation, so I have a somewhat independent sample here. After identifying one tagged track I test the mistag prediction on all the other taggable tracks in the jet as compared to any second tags in the jet. Thus, I make the assumption that once an SLT tag is found in a jet, any other SLT tags in the jet are mistags. The results of this test are found in Table 6.7

There are some caveats to this double tag study, however. First, it is not guaranteed that the second tag in a jet is a mistag—there may have been a sequential decay, $B \rightarrow D \rightarrow \mu$, as well as semileptonic decay of a $B \rightarrow \mu$. Second, an extra mistag in a jet is more likely to be found once the first is identified; if a jet falls into a calorimeter gap and one track is tagged, the other tracks in that jet are more likely to be in the gap and be tagged as well. Attempting to avoid these two remaining biases would make the already small size of the double-tagged-jets sample abysmal, so I leave the sample alone and merely use it as a cross check.

There is generally good agreement between the measured and predicted tags though there is a systematic underprediction, which is expected given the caveats described above. The predictions in the Jet 20 and ΣE_T samples are particularly low (the γ+jets prediction is also somewhat low), so I made a quick check to see if the reason could be determined. Table 6.8 presents the ratio of jets in which the two tagged tracks have opposite-sign charge, to those with same-sign charge. Though the uncertainties are large, it appears that the ratio is largest in the same samples that are underpredicted. It is difficult to tell, but the results are consistent with there being somewhat more sequential decays in the samples where the underprediction is largest.

Sample	Ratio OS/SS tags
Jet 20	2.2±0.9
Jet 50	1.6±0.8
Jet 70	1.5±0.6
Jet 100	1.3±0.5
γ+Jets	1.9±0.7
ΣE_T	1.8±0.4

Table 6.8: The ratio of opposite-sign SLT tags to same-sign SLT tags in double-tagged jets.

Chapter 7

Background Estimation

The dominant background contribution to the top signal in this analysis comes from *fake* muon tags (mistags) in W+jets events. In this context, fake tags are particles identified as muons whose origin is not semileptonic heavy-flavor decay. This includes non-muons which are tagged, such as a pion faking a muon, as well as muons from pion or kaon decays-in-flight. Another smaller, yet still significant background comes from W bosons produced in association with with heavy flavor ($Wb\bar{b}$, $Wc\bar{c}$ and Wc). The estimate of the mistag background is described in Section 7.1, while the W+heavy-flavor background estimate is described in Section 7.2.

Other backgrounds that can produce a W boson and an SLT muon that are not accounted for by the mistag rate include dibosons (WW, ZZ, WZ), $Z \to \tau^+\tau^-$, single top, QCD multijet backgrounds including $b\bar{b}$, and residual Drell-Yan ($\mu\mu$) events not removed by the dimuon removal. QCD and Drell-Yan backgrounds are measured using the data, as described in detail in Sections 7.3 and 7.4 below. The remaining backgrounds are estimated from Monte Carlo as described in Section 7.5. I treat QCD independently of the calculation of mistags in W+jets events because events that enter our sample by mimicking the signature of a W boson can have a significantly larger tag rate than true W events. The enhanced tag rate arises due to the contribution of $b\bar{b}$ events to the QCD background and because of the correlation between the tag rate and measured $\not{E}_t$ in events in which the $\not{E}_t$ arises from jet mismeasurement or semileptonic heavy-flavor decay rather than from a neutrino in a W boson decay. In order to avoid double counting I correct the estimate of tags in W+jets events by $(1 - F_{QCD})$, where F_{QCD} is the QCD multijet fraction in the W+jets candidate sample.

7.1 Mistags

The background due to mistags is estimated using the track-based mistag rate described in Chapter 6. To predict the number of events from W+jets with at least 1 mistag, I apply the mistag

matrix to all pretag events in the signal region according to:

$$N_{\text{raw}}^{\text{Wjtag}} = \sum_{\text{events}} \left[1 - \prod_{i=1}^{N_{\text{trk}}} \left(1 - \mathcal{P}(p_{T_i}, \eta_i)\right) \right], \tag{7.1}$$

where the sum runs over each event in the pretag sample, and the product is over each taggable track in the event. $\mathcal{P}(p_{T_i}, \eta_i)$ is the probability from the tag matrix for tagging the ith track with parameters p_{T_i} and η_i. Note that the sum over the events in Equation 7.1 includes any $t\bar{t}$ events that are in the pretag sample. I correct for the resulting overestimate of the background at the final stage of the cross section calculation, since the correction depends on the measured tagging efficiency (see Section 9.1). It also includes W+heavy-flavor events, diboson events, etc. The inclusion of these other backgrounds is not corrected for. Instead, the mistag matrix is used to predict the mistags from those backgrounds, and the estimates of the contributions from those backgrounds are limited to the *real* tags in those events (i.e. tags of muons resulting from the decay of heavy-flavor hadrons or W or Z bosons).

A fraction, F_{QCD}, of the events in the signal region are QCD events for which the background is estimated separately. Therefore, I correct the prediction of Equation 7.1 according to

$$N_{\text{corr}}^{Wjtag} = (1 - F_{QCD}) \cdot N_{\text{raw}}^{\text{Wjtag}}. \tag{7.2}$$

The background estimate from the application of the mistag matrix is shown in Table 7.1. I list here both the raw prediction and that corrected by $(1 - F_{QCD})$. The calculation of F_{QCD} itself is described in Section 7.3.1 and the values given in Table 7.6. is given in Table 7.6 below, for the final mistag background.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jet	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
$N_{\text{raw}}^{\text{Wjtag}}$	641.1±32.1	237.7±11.9	54.97±2.75	32.56±1.63	87.53±4.38
$N_{\text{corr}}^{\text{Wjtag}}$	622.3±31.4	225.9±11.6	53.03±2.68	31.44±1.59	84.47±4.26

Table 7.1: Summary of background estimate from mistags in W+jets events.

7.2 W+Heavy Flavor

The estimation of background tags from the semileptonic decays of heavy-flavor partons in $Wb\bar{b}$, $Wc\bar{c}$ and Wc events is estimated by using the ALPGEN Monte Carlo program to determine the

fraction of W+jets events that contain heavy flavor at the pretag level and the tagging efficiency for these events. This is done, rather than using the cross sections for these processes, because, while the total cross sections for the $Wb\bar{b}$, $Wc\bar{c}$ and Wc processes can be calculated, the calculations have large theoretical uncertainties. The fraction of W events containing heavy-flavor jets has a smaller uncertainty because much of the uncertainty cancels in the ratio of W+heavy-flavor events divided by all W+jets events, and the total background yield can be constrained by the data sample.

7.2.1 Heavy-Flavor Fractions

The heavy-flavor fractions of events in the W+jets sample is determined by measuring the fractions in Monte Carlo and then multiplying those raw fractions by a factor, K, that corrects for the differences between data and Monte Carlo, as described below. Important to the interpretation of the heavy-flavor fractions is an understanding of what is meant by a b-jet and a c-jet. I define a b-jet, using Monte Carlo truth information, as a jet containing a bottom quark within a cone $\Delta R = 0.4$ of the jet axis that is also closer to the jet axis than any charm quark within the cone. A c-jet is defined similarly.

With this definition of heavy-flavor jets, determining the raw heavy-flavor fraction in Monte Carlo is straightforward. In each jet-multiplicity bin, I count the number of events containing one or two b-jets, or one or two c-jets, and divide by the total number of W+jets events. These calculated raw fractions are shown in Table 7.2.

Category	1 jet	2 jets	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
Raw Fractions (%)					
2 b		0.787±0.005	1.60±0.01	2.42±0.02	1.75±0.01
1 b	0.580±0.003	1.20±0.01	2.25±0.01	2.59±0.04	2.31±0.01
2 c		1.16±0.01	2.44±0.02	3.93±0.04	2.72±0.02
1 c	4.78±0.01	7.72±0.03	9.53±0.05	10.0±0.1	9.62±0.05
Corrected Fractions (%)					
2 b		0.9±0.2	1.8±0.5	2.8±0.7	2.0±0.5
1 b	0.7±0.2	1.4±0.4	2.6±0.7	3.0±0.8	2.7±0.7
2 c		1.3±0.4	2.8±0.7	4.5±1.2	3.1±0.8
1 c	5.5±1.5	8.9±2.4	11.0±2.9	11.5±3.1	11.1±2.9

Table 7.2: The heavy flavor fractions for W+heavy flavor background. The uncertainty on the corrected fractions include the systematics from the K factor and the ALPGEN parameters.

ALPGEN is a leading-order Monte Carlo and there is no guarantee that the heavy-flavor fractions are independent of NLO effects. Therefore, the raw fractions I measure are corrected by a multiplicative K factor [44]. Because of the limited size of the W+jets tagged sample it is not

possible to verify the ALPGEN heavy-flavor fraction in this sample. Instead, the heavy-flavor content of ALPGEN and data are compared in QCD multi-jet samples. For this comparison, ALPGEN QCD samples are prepared in the same way the W+jets samples are (see Section 3.2), and the fractions are calculated similarly. The heavy-flavor content of the single-jet-triggered Jet50 data sample is inferred from fits to secondary-vertex-tagged mass[1] templates. The K factor is defined as the ratio of heavy-flavor fractions in the data and simulation and determined to be 1.15 ± 0.15. The uncertainty on the K factor is determined from fit statistics and χ^2, repeated tests with an alternate template variable (pseudo-$c\tau$[2]), and the spread in results as a function of energy and jet multiplicity. The heavy-flavor fractions corrected by the K factor may be found in Table 7.2 along with the raw fractions.

In addition to the uncertainty on the K factor, I apply a systematic uncertainty to the measured heavy-flavor fractions that is due to the ALPGEN-parameter selection [44]. The uncertainties are determined by varying parameters in ALPGEN and determining the heavy-flavor fractions with the new settings. Several settings are varied including the MLM matching parameters (see Section 3.2), quark masses, and Q^2 scale, as well as the amount of ISR/FSR and the PDF. Table 7.3 lists the systematics on the determined heavy-flavor fractions.

Systematic Source	Relative Error (%)
MLM Matching	6
Q^2	20
Quark masses (c and b)	6
ISR	4
FSR	5
PDF	3
K	13
Total	26

Table 7.3: Relative systematic errors for the determination of the heavy-flavor fraction of W+jets events before tagging.

7.2.2 Heavy-Flavor Tagging Efficiency

The Monte Carlo is also employed to determine the efficiency for tagging a muon from a semileptonic heavy-flavor decay in W+heavy-flavor events. The calculation of these efficiencies is straightforward,

[1] The vertex mass is calculated as the invariant mass of all the particles whose tracks cross at the (secondary) vertex.

[2] The pseudo-$c\tau$ of a vertex is defined as the displacement, in the xy-plane, of the vertex from the primary event vertex times the vertex mass divided by the p_T of the vertex. Pseudo-$c\tau$ is a measure of the lifetime (τ) of the particle from which the tracks in the vertex originated.

being equal to the number of tagged events divided by the number of pretag events. As for the $t\bar{t}$ tagging efficiency described in Section 5.7.2, tags are assigned based on the SLT-muon-tagging efficiency measured in the data. The results are shown in Tab. 7.4. The pretag sample to which the

Tagging Efficiency (%)					
Category	1 jet	2 jets	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
2 b		7.8±0.2	8.4±0.2	8.5±0.3	8.4±0.2
1 b	3.54±0.05	4.30±0.06	5.5±0.1	5.8±0.2	5.53±0.09
2 c		3.1±0.1	3.6±0.1	3.5±0.2	3.6±0.1
1 c	1.52±0.02	1.70±0.03	2.04±0.07	2.05±0.06	2.04±0.06

Table 7.4: The *real* tagging efficiencies for the W+heavy flavor background.

HF fractions and efficiencies are applied must be corrected for other the other backgrounds (QCD, DY, MC driven backgrounds) in the pretag sample to avoid double counting. Therefore, once the fractions (F_{HF}) and efficiencies (ϵ_{HF}) have been determined, I estimate the number of tags from W+heavy-flavor events as:

$$N_{HF} = (1 - F_{QCD} - F_{other}) \cdot N_{pretag} \cdot F_{HF} \cdot \epsilon_{HF}. \tag{7.3}$$

where F_{QCD} is the fraction of events from QCD in the pretag sample and F_{other} is the fraction from the other backgrounds. As with the mistag prediction, correction for $t\bar{t}$ in the pretag sample is done during the final cross section calculation.

7.3 QCD

The background due to tags in QCD events that enter the signal sample is estimated by calculating the fraction of QCD events in the W+jets data and applying the standard mistag matrix times a multiplicative factor. The multiplicative factor is required because the tagging rate of QCD events that enter the pretag sample is higher than the corresponding tagging rate for W+jets events. The QCD background calculation then has two pieces: the calculation of the fraction of QCD events in the W plus jets sample, F_{QCD} and a measurement of the enhancement of the mistag rate for QCD events. The measurements of these two pieces are described in Sections 7.3.1 and 7.3.2 below.

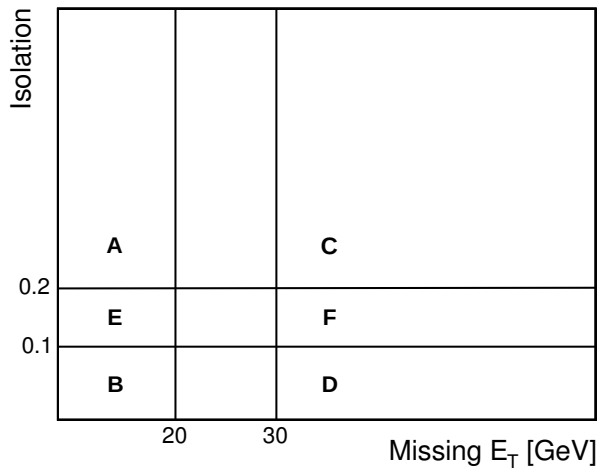


Figure 7.1: A diagram illustrating the regions defined in the missing-E_T vs. isolation plane.

7.3.1 The Fraction of QCD Events Before SLT Tagging

The fraction of QCD events before SLT tagging is determined using the so-called "MET versus Iso" technique. The missing-E_T (MET or $\not\!\!E_t$) of an event is assumed to be uncorrelated with the isolation (Iso) of the high-p_T lepton in QCD events. Therefore, events are counted in the $\not\!\!E_t$-Iso plane, divided in four kinematic regions (see Figure 7.1) according to the following boundaries:

$$\text{Region A} : \not\!\!E_t < 20 \text{ GeV}; \quad Iso > 0.2$$

$$\text{Region B} : \not\!\!E_t < 20 \text{ GeV}; \quad Iso < 0.1$$

$$\text{Region C} : \not\!\!E_t > 30 \text{ GeV}; \quad Iso > 0.2$$

$$\text{Region D} : \not\!\!E_t > 30 \text{ GeV}; \quad Iso < 0.1$$

For events with $\not\!\!E_t$ and Iso uncorrelated, the number of events in Region D (the signal region) is given by:

$$N_D^{QCD} = \frac{N_C}{N_A} N_B, \tag{7.4}$$

where N_X is the number of events counted in Region X. If the total number of events in Region D is N_D^{tot}, the fraction of QCD events in Region D, F_{QCD}, is given by $F_{QCD} = N_D^{QCD}/N_D^{tot}$. Table 7.5 shows this "raw" F_{QCD} in the electron and muon channels as a function of the number of jets in the event. The measured QCD fraction is also corrected for the contamination of W+jets and $t\bar{t}$

events that would introduce a correlation between $\not{E}_t$ and Iso. The $t\bar{t}$ contamination is removed by subtracting the number of events in each region that are expected for a $t\bar{t}$ cross section of 6.7 pb. The contamination from W+jets is determined by scaling W+jets Monte Carlo to the data in Region D—that is, $N_W^D = N_{\text{data}}^D \cdot (1 - F_{QCD}) - N_{t\bar{t}}^D$. As the scaling depends on F_{QCD}, an iterative process of scaling W+jets, subtracting its contribution to each region, and remeasuring F_{QCD} is used until F_{QCD} changes by less than 0.1 percent. The corrected QCD fractions are given in Table 7.5, along with the raw fractions.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jets	3 jets	≥ 4 jets	≥ 3 jets
Electron channel					
F_{QCD}^{raw}	0.0560±0.0009	0.086±0.002	0.071±0.004	0.086±0.008	0.074±0.004
$F_{QCD}^{\text{corrected}}$	0.0423±0.0009	0.070±0.002	0.049±0.003	0.056±0.006	0.051±0.003
Region F	0.95±0.04	0.97±0.06	0.84±0.10	1.06±0.24	0.89±0.09
Muon channel					
F_{QCD}^{raw}	0.0224±0.0005	0.036±0.001	0.051±0.005	0.043±0.006	0.049±0.004
$F_{QCD}^{\text{corrected}}$	0.0118±0.0004	0.020±0.001	0.013±0.004	0.007±0.004	0.011±0.003
Region F	0.58±0.05	0.65±0.07	0.31±0.09	2.27±4.25	0.45±0.13

Table 7.5: The fractions, F_{QCD}, of lepton-plus-jets events due to QCD multijet processes before SLT tagging, both raw and corrected for $t\bar{t}$ and W+jets contamination. Also shown is the measured QCD fraction in Region F, used to assign a systematic uncertainty on the F_{QCD} prediction.

To evaluate the accuracy of the $\not{E}_t$-Iso prediction, two complementary regions in the plane are defined as:

$$\text{Region E} : \not{E}_t < 20 \text{ GeV}; \quad 0.1 < Iso < 0.2$$

$$\text{Region F} : \not{E}_t > 30 \text{ GeV}; \quad 0.1 < Iso < 0.2$$

Region F is outside the signal region and, once contamination from W+jets and $t\bar{t}$ is removed, should have a QCD fraction, F_{QCD}^F, of approximately 1.0. Analogously to the determination of F_{QCD} given above, I calculate F_{QCD}^F as:

$$F_{QCD}^F = \frac{N_C \cdot N_E}{N_A \cdot N_F}, \tag{7.5}$$

and use the difference in F_{QCD}^F from 1.0 to estimate a systematic uncertainty on the MET vs. Iso technique. The results of my F_{QCD}^F are also given in Table 7.5. Given the deviation from 1.0 in the ≥ 3 jets data with $H_T > 200$ GeV, I assign an 11 percent (120 percent) systematic uncertainty to

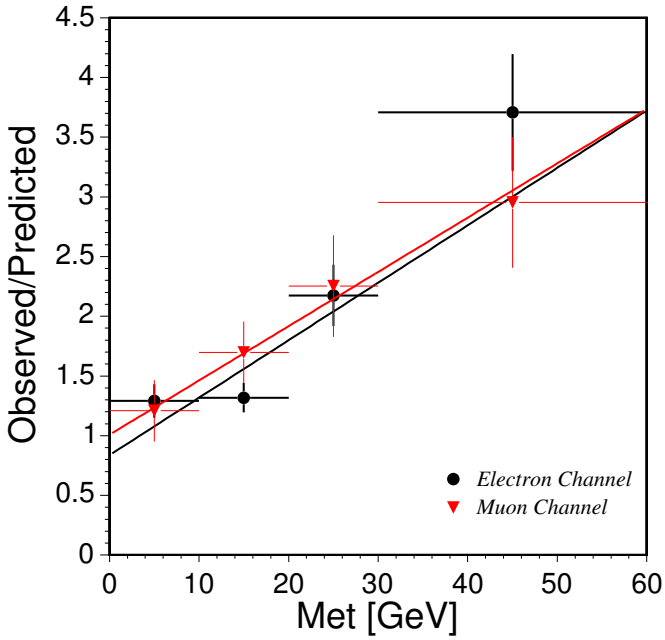


Figure 7.2: The ratio of observed to predicted tags as a function of $\not\!\!E_t$ in data with a non-isolated primary lepton ($Iso \geq 0.2$).

F_{QCD} for electrons (muons).

7.3.2 The Tag Rate of QCD Events

The tag rate in QCD events is not, *a priori* the same as that predicted by the mistag matrix, and in fact is found to be decidedly enhanced relative to the matrix. There are two sources for this enhancement. First, much of the $\not\!\!E_t$ in QCD events is due to mismeasurement of jet energies, which is correlated to the tag rate (see Section 6.4.3). As seen in Figure 7.2, the tag rate (shown as the ratio of observed to predicted tags) increases with increasing $\not\!\!E_t$. Second, QCD includes $b\bar{b}$ and $c\bar{c}$ events in which the high-p_T lepton comes from the semileptonic decay of one of the b or c quarks; if the other also decays semileptonically, it may be tagged by the SLT.

In the $\not\!\!E_t$-Iso plane, the region closest kinematically to the signal is the high-isolation and high-$\not\!\!E_t$ region, Region C. Region C has the same $\not\!\!E_t$ requirement that Region D does, and likewise it requires a high-p_T lepton in the event. Therefore, the tag rate measured in Region C is a good representation of that of QCD events in Region D.

The tag rate of QCD events is measured as an enhancement factor, k, times the W+jets mistag rate. k is determined simply by counting the observed number of events tagged in Region C and

dividing by the number of events predicted by the mistag matrix, also in Region C. The results are shown in Table 7.6.

7.3.3 The QCD Background Estimate

Having determined F_{QCD} and k, the QCD background may be given by:

$$N_{QCD} = N_{\text{raw}}^{\text{Wjtag}} \cdot k \cdot F_{QCD} \tag{7.6}$$

where $N_{\text{raw}}^{\text{Wjtag}}$ from Equation 7.1 represents the application of the mistag matrix to the events in Region D (the signal region: $\not{E}_t > 30$ GeV, $I < 0.1$).

The total systematic uncertainty on the QCD background estimate includes the systematic uncertainty on F_{QCD} added in quadrature with the statistical uncertainties on the k-factor determination and on F_{QCD}, all quoted in Table 7.6. The results are summarized in Table 7.6.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jets	3 jets	≥ 4 jets	≥ 3 jets
Electron channel					
F_{QCD}	0.042±0.005	0.070±0.008	0.051±0.006		
k	5.3±0.4	3.9±0.4	3.7±0.5		
N_{QCD}	82.63±11.90	38.05±6.13	6.28±1.20	3.62±0.69	9.90±1.89
Muon channel					
F_{QCD}	0.0118±0.0004	0.0205±0.0009	0.011±0.014		
k	2.9±0.4	3.4±0.4	3.0±0.5		
N_{QCD}	9.33±11.28	6.84±8.26	0.72±0.89	0.44±0.55	1.16±1.44
Combined channels					
Total N_{QCD}	91.95±16.51	44.89±10.35	7.00±1.50	4.06±0.89	11.06±2.39

Table 7.6: Summary of the QCD background estimate. The uncertainties on the QCD fractions, F_{QCD}, and number of QCD events, N_{QCD}, are systematic and statistical combined.

7.4 Drell-Yan$\rightarrow \mu\mu$

The Drell-Yan$\rightarrow \mu\mu$ background is suppressed in the analysis with two vetoes. First, the lepton + jets sample has a Z veto applied before heavy flavor tagging. This veto looks for a second lepton candidate in the event of the same flavor and opposite sign as the trigger lepton. If the invariant mass of the pair falls within the range of $76 < M_{\ell\ell} < 106$ GeV/c^2 the event is rejected. This veto considers only isolated objects for the second leg. Drell-Yan events with one non-isolated leg can

survive this veto. For the muon sample I apply an additional veto on opposite sign high-p_T muon and soft muon tagged pairs. The pair invariant mass is not allowed to be in the ranges $M_{\ell\ell} < 5$ GeV/c^2, $8 < M_{\ell\ell} < 11$ GeV/c^2, or $70 < M_{\ell\ell} < 110$ GeV/c^2. These windows are chosen to suppress J/Ψ's and sequential B decays, Υ's, and Z's respectively.

When a muon from a Z decay radiates a high-E_T photon, the event can fall outside the nominal Z-mass window. If the photon has $E_T > 20$ GeV, it will be counted as a jet, making an N-jet event appear to be an $(N+1)$-jet event. Furthermore, the radiating muon in these events typically remains close to the photon and appears as a muon inside a jet in the SLT analysis, making these events an important source of background. Figure 7.3 illustrates the effect of these radiative Z events on the dimuon invariant mass distribution. Among the features in the distribution, one notices the enhancement of events between about 60 and 80 GeV, which are mostly events where one muon has radiated a stiff photon. To separate the radiative Z events from other Drell-Yan outside the Z window, I plot events vs. electromagnetic fraction of the jet containing the SLT-tagged muon signed by the charge-product of the high-p_T muon and the SLT in Figure 7.4. The peak close to -1.0 are events where a muon has radiated a photon of $E_T > 20$ GeV. To reduce this contribution to the background, I remove events in which the tagged jet has an EM fraction above 0.8 and just one track with $p_T > 1.0$ GeV/c within a cone of $\Delta R = 0.4$ about the jet axis if the high-p_T muon and SLT have opposite charge.

The residual Drell-Yan component surviving these vetoes is calculated as follows: I use $Z \to \mu\mu$ events in the data and measure the number of events in each jet bin that pass all of the event selection cuts, except that I count, rather than remove, SLT tagged events where the SLT–primary-lepton invariant mass falls inside the Z-mass window. An ALPGEN $Z/\gamma^* \to \mu\mu$ Monte Carlo sample is then used to estimate the ratio of events inside the Z-mass window to events outside, $R^{out/in}$, and their respective tagging efficiencies, $\epsilon^{out,in}$. Events used to measure $R^{out/in}$ do not have the $\not\!\!E_t$ or H_T cuts applied, or the dilepton veto, and there is no requirement that the SLT taggable leg be inside of a jet and the ratio is expected to be consistent across jet multiplicity bins. The above requirements are accounted for in $\epsilon^{out,in}$. For the purposes of this calculation, 'inside' the Z-window is defined as invariant mass from 80-110 GeV/c^2, while 'outside' is defined as outside the region 70-110 GeV (and outside the other dimuon mass windows). The lower limit of 80 GeV/c^2 for the Z-window is made to remove biases from non-Z backgrounds in the window. The Drell-Yan background is then written:

$$N_{DY} = N^Z_{tag} \cdot R^{out/in} \cdot \frac{\epsilon^{out}(\not\!\!E_t, H_T, \text{dilep}, \text{SLT–jet})}{\epsilon^{in}(\not\!\!E_t, H_T, \text{dilep}, \text{SLT–jet})}, \tag{7.7}$$

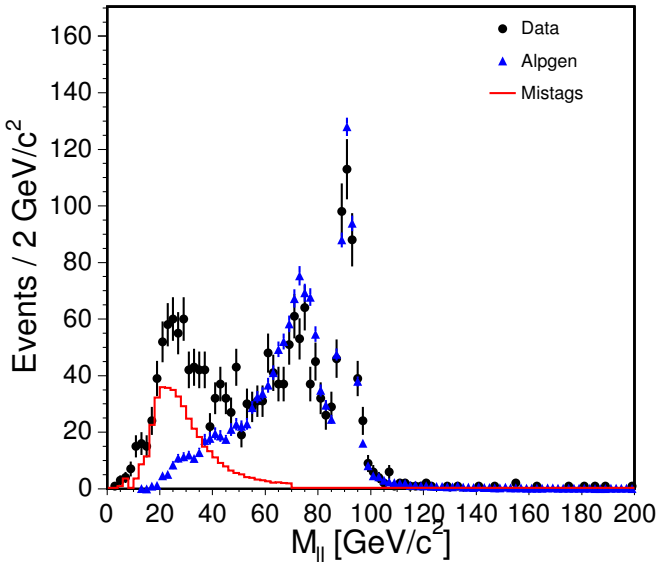


Figure 7.3: Comparison between data and ALPGEN Monte Carlo dimuon invariant mass distributions in events with 1 jet. The distributions are normalized to the same number of events under the Z peak.

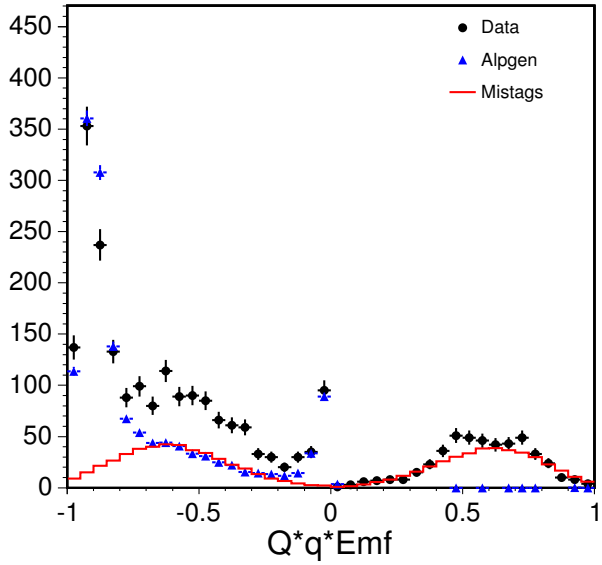


Figure 7.4: The product of the electromagnetic fraction of the SLT jet and the charge product of the SLT and primary muon.

with the first factor coming from the data and the last two from the Monte Carlo. The inputs and results of the Drell-Yan calculation are given in Table 7.7.

	H_T >0 GeV		H_T >200 GeV		
	1 jet	2 jets	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
N^Z_{tag} (data)	27	25	3	0	3
$R^{\frac{out}{in}}$	0.325±0.003	0.315±0.005	0.312±0.006		
ϵ^{out} (%)	0.71±0.04	1.62±0.10	2.85±0.24		
ϵ^{in} (%)	1.03±0.05	3.10±0.12	3.24±0.18		
$R^{\frac{out}{in}} \cdot \frac{\epsilon^{out}}{\epsilon^{in}}$	0.223±0.018	0.165±0.013	0.274±0.027		
Drell-Yan Total	6.02±1.25	4.12±0.88	0.82±0.44	0.00±0.19	0.82±0.48

Table 7.7: Drell-Yan background summary. Uncertainties are statistical only.

7.5 Monte Carlo Driven Backgrounds

Backgrounds from dibosons (WW,WZ,ZZ), $Z \rightarrow \tau^+\tau^-$, and single top are determined from Monte Carlo. Each of these MC background calculations requires measuring the acceptance and tagging efficiency for a particular kind of event. The acceptance and efficiency are measured in the same way as for $t\bar{t}$ events described in Section 5.7. The only exception is that we do not include in the efficiency the contribution from mistags because that contribution is included in the application of the mistag matrix to the pretag sample. For each of these backgrounds, the estimated number of tags is calculated as

$$N_i = \sigma_i \cdot A_i \cdot \epsilon_i \cdot \int \mathrm{d}t, \tag{7.8}$$

where σ_i is the theoretical cross section for the process, A_i and ϵ_i are the acceptance and efficiency measured in Monte Carlo, and $\int \mathrm{d}t$ is the integrated luminosity. A brief discussion of how each these backgrounds enter the SLT tagged sample and the results of the background predictions may be found below.

7.5.1 Diboson Backgrounds

Background from WW, WZ, and ZZ production may enter the signal sample in two ways. Either one boson decays into an e or μ and the other decays hadronically into charm or bottom quarks that subsequently produce an SLT tag through their decay, or the other boson also decays into an e or μ that ends up in a jet and is misidentified as an SLT. The theoretical cross sections used, and the acceptances and efficiencies computed, are given in Table 7.8 along with the estimated number

of tags for each process.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jet	3 jets	≥ 4 jets	≥ 3 jets
WW					
σ_{theory}	12.4±1.2 pb				
A (%)	2.44±0.01	2.62±0.01	0.403±0.004	0.121±0.002	0.524±0.005
ϵ (%)	0.49±0.03	0.76±0.04	0.88±0.09	1.56±0.23	1.06±0.09
N_{WW}	2.986±0.299	5.001±0.394	0.892±0.190	0.475±0.118	1.395±0.228
WZ					
σ_{theory}	3.96±0.40 pb				
A (%)	1.085±0.007	1.317±0.007	0.233±0.003	0.070±0.002	0.302±0.004
ϵ (%)	0.85±0.06	1.72±0.07	1.46±0.16	2.69±0.39	1.77±0.15
N_{WZ}	0.740±0.075	1.821±0.128	0.274±0.044	0.151±0.036	0.432±0.058
ZZ					
σ_{theory}	3.4±0.3 pb				
A (%)	0.104±0.002	0.097±0.002	0.060±0.001	0.012±0.001	0.042±0.001
ϵ (%)	1.0±0.2	2.4±0.3	2.3±0.5	1.6±0.5	2.1±0.4
N_{ZZ}	0.07±0.02	0.16±0.05	0.05±0.02	0.013±0.006	0.06±0.02
$N_{diboson}^{total}$	3.80±0.44	6.98±0.66	1.21±0.23	0.64±0.14	1.88±0.30

Table 7.8: Summary of diboson backgrounds. The theoretical cross sections [45] are inclusive. The acceptance includes the branching fraction to events with N jets.

7.5.2 Drell-Yan$\rightarrow \tau\tau$ Background

Drell-Yan$\rightarrow \tau^+\tau^-$ enters the sample when one τ produces a high-p_T electron or muon and either the other τ results in an SLT tag, or a jet produced in association with the Z boson produces an SLT tag. In the latter case (as with other backgrounds), the tag can be from heavy-flavor decay or from a mistag. The estimate of the Drell-Yan$\rightarrow \tau\tau$ background includes tags from the heavy-flavor decays, but not from mistags, which is accounted for by the mistag matrix. The theoretical cross section used, and the acceptance and efficiency computed, for the Drell-Yan$\rightarrow \tau\tau$ estimate are given in Table 7.9 along with the estimated number of tags.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jet	3 jets	≥ 4 jets	≥ 3 jets
σ_{theory}	333±4.2 pb				
A (%)	0.112±0.001	0.054±0.001	0.0058±0.0004	0.0014±0.0002	0.0073±0.0004
ϵ (%)	0.4±0.1	0.4±0.1	1.6±0.7	1.3±0.5	1.3±0.5
$N_{\mathrm{Drell-Yan}\rightarrow\tau\tau}$	2.65±0.57	1.54±0.43	0.65±0.28	0.13±0.05	0.65±0.27

Table 7.9: Summary of Drell-Yan$\rightarrow \tau^+\tau^-$ background. The theoretical cross section [46] is inclusive. The acceptance includes the branching fraction to events with N jets.

7.5.3 Single Top Background

The signature of single-top events, with additional jets from gluon radiation, is nearly identical to that of the $t\bar{t}$ signal. So, single top may be tagged the same way $t\bar{t}$ is, primarily through the semileptonic decay of a B hadron. The contribution from the s-channel and t-channel diagrams to the single-top background are computed separately and the theoretical cross sections used, and the acceptances and efficiencies computed, are given in Table 7.10 along with the estimated number of tags.

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
	1 jet	2 jet	3 jets	≥ 4 jets	≥ 3 jets
s-channel					
σ_{theory}	0.88±0.11 pb				
A (%)	3.37±0.02	7.97±0.03	2.15±0.01	0.61±0.01	2.76±0.02
ϵ (%)	5.0±0.1	9.7±0.1	10.2±0.2	11.1±0.4	10.4±0.2
$N_{s-\text{chan}}$	2.99±0.34	13.82±1.38	3.94±0.44	1.20±0.16	5.14±0.56
t-channel					
σ_{theory}	1.98±0.08 pb				
A (%)	5.72±0.02	6.30±0.03	1.04±0.01	0.170±0.004	1.21±0.01
ϵ (%)	4.38±0.09	5.19±0.09	5.96±0.24	7.37±0.64	6.16±0.22
$N_{t-\text{chan}}$	10.08±1.11	13.18±1.42	2.48±0.33	0.50±0.10	2.99±0.38

Table 7.10: Summary of single top background. The theoretical cross sections [47] are inclusive. The acceptance includes the branching fraction to events with N jets.

Chapter 8

Systematic Uncertainties

Systematic uncertainties in this analysis come from Monte Carlo modeling of the geometrical and kinematic acceptance, knowledge of the SLT tagging efficiency, the effect on the acceptance of the uncertainty on the jet energy scale, uncertainties on the background predictions, and the uncertainty on the luminosity.

8.1 Uncertainties on Acceptance and Efficiency

Acceptance systematics are caused by uncertainties in the efficiency of the trigger and lepton identification, or the efficiency for passing the $\not\!\!E_t$ and jet-E_T cuts. Sources of systematic uncertainty on the acceptance are listed below along with their relative effect on the $t\bar{t}$ acceptance:

- **Lepton Identification Efficiency.** As discussed in Section 5.7.1, corrections are applied to the acceptance measured in simulation to better represent the data. These include the corrections for the difference between data and PYTHIA lepton identification efficiencies and and track reconstruction efficiencies as well as the application of the trigger efficiencies and z-vertex efficiency Each of these carries an uncertainty. Additionally, there are uncertainties on the lepton isolation and the muon-track-χ^2 cut (see Section 5.2.2). These uncertainties are listed in Table 8.1, and the combination is estimated to contribute a 2.9 percent uncertainty to the $t\bar{t}$ acceptance.

- **Parton Distribution Functions.** The uncertainty in the distribution of the proton (and anti-proton) momentum among its constituent partons (PDFs) affects the relative contributions of the $q\bar{q}$ and gg processes to the $t\bar{t}$ production as well as the momentum of the $t\bar{t}$ system. The uncertainty due to PDFs is estimated by re-weighting the $t\bar{t}$ events generated with `CTEQ5L` for different sets of PDFs. Three different uncertainty estimates are made and then added in quadrature. First, in the `CTEQ` parametrization of the parton distribution functions are described by 20 independent eigenvectors. In a next-to-the-leading-order (NLO) version of PDFs,

	Electrons	CMUP Muons	CMX Muons
Source	Fractional Uncertainty (%)		
Trigger efficiency	0.5	0.6	0.8
Lepton identification	0.5	0.7	0.7
χ^2 cut	N/A	0.2	0.2
Lepton Isolation	0.8	1.0	1.4
Tracking efficiency	0.2		
z-vertex efficiency	0.2		
Total	2.9		

Table 8.1: Summary of lepton identification systematics.

`CTEQ6M`, a 90 percent confidence interval is provided for each eigenvector. A new acceptance is computed for the maximum and minimum value of each eigenvector by re-weighting the default PYTHIA $t\bar{t}$ sample. The differences between the weighted acceptances for the 20 eigenvectors with respect to the weighted acceptance from the central `CTEQ6M` value are added in quadrature. Second, the difference between the acceptance from an alternative PDF, `MRST` [48], and the default `CTEQ5L` sample is considered. Finally, the uncertainty from α_s is estimated by comparing the weighted acceptance for `MRST` with $\alpha_s = 0.1125$ and $\alpha_s = 0.1175$. Adding these three contributions in quadrature, an uncertainty of 0.9 percent on the $t\bar{t}$ acceptance is obtained.

- **ISR/FSR.** Jets due to initial-state and final-state gluon radiation (ISR and FSR, respectively) may be produced in addition to the jets from the top decay products (see Figure 8.1). The uncertainties associated with the modeling of ISR and FSR is estimated by measuring the acceptance with different levels of ISR and FSR. The range of variation for ISR levels was determined by taking the extremes of a range of ISR values determined by a study of Drell-Yan $Z \to \ell\ell$ events in data and Monte Carlo. The same variations are also applied to the FSR. A total uncertainty of 0.8 percent is estimated from these effects.

- **Jet Energy Scale.** The raw calorimeter energy of jets is corrected to better represent the energy of the originating parton (see Section 5.3), and this rescaling carries a systematic uncertainty. A shift in this jet-energy scale can affect not only the number of jets passing selection, but also whether an event passes the $\not\!\!E_t$ requirement. The effect of the uncertainty on the jet-energy scale is estimated by shifting the jet-energy scale by $\pm 1\sigma$ and measuring the change caused in the measured signal acceptance. The result is a 4.1 percent uncertainty on the acceptance.

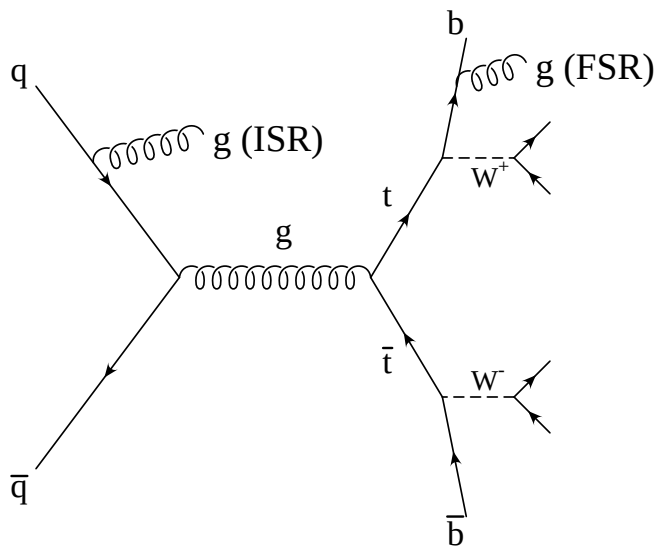


Figure 8.1: Feynman diagram of top production and decay with ISR and FSR.

- **Generator.** Effects of the modeling of $t\bar{t}$ production and decay are evaluated by using the difference in acceptance between signal samples generated by PYTHIA and HERWIG. This covers differences between the two in generated kinematics, shower model, and fragmentation. An overall shift in the acceptance of 2.4 percent is observed between the two generators and taken as a systematic uncertainty.

Adding the contributions in quadrature gives an overall systematic uncertainty on the acceptance of ± 5.7 percent.

Several factors contribute to the systematic uncertainty on the SLT tagging efficiency. The uncertainty due to the limited knowledge of the P_T dependence is determined by varying the efficiency curves used in the $t\bar{t}$ Monte Carlo for the tagging efficiency measurement according to the upper and lower bands in Figures 4.3 and 4.4. These bands are, in turn, determined by the 1 σ variations in the parameters of the fits. The tagging efficiency for $t\bar{t}$ events changes by ± 1.0 percent from its central value when varying the efficiency curves, and this full value is taken as a systematic uncertainty. An additional systematic uncertainty for the tagging efficiency comes from the fact that the Monte Carlo tracking efficiency is used implicitly for taggable tracks. As these tracks can be in dense environments in or near jets, the efficiency is expected to be somewhat less than for isolated tracks. Studies done by embedding Monte Carlo tracks in jets in both data and Monte Carlo events indicate that the Monte Carlo tracking efficiency in dense environments is a few percent higher than in data, but the uncertainties are large. A 5 percent systematic uncertainty is assigned

to the tagging efficiency to for this effect. Finally, the statistical uncertainty on the SLT tagging efficiency measurement also comes in as a systematic uncertainty. Adding these contributions in quadrature gives an overall systematic uncertainty of 5.1 percent. Note that the uncertainty on the tagging efficiency also affects the background determination. The reason is that $t\bar{t}$ events need to be subtracted from the pretag sample which is used in Equations 7.1 and 7.3 to determine the W+jets background. This is taken into account when calculating the uncertainty on the cross section.

8.2 Uncertainties on Background Predictions

Uncertainties on the mistag matrix are determined by the level of agreement between observed tags and predictions in a variety of samples across a range of jet energies, as described in Section 6.4. The uncertainty on the mistag prediction is ± 5 percent.

Three sources contribute to the uncertainty on the $Wb\bar{b} + Wc\bar{c}$+Wc background prediction: the selection of ALPGEN settings, the errors associated with the K factor that scales the heavy-flavor fraction in ALPGEN to the data, and the uncertainty on the tagging efficiency. The determination of the uncertainties on the ALPGEN settings and the K factor are described in Section 7.2.1. The ALPGEN settings contribute 23 percent to the W+heavy-flavor background uncertainty and the K factor another 13 percent. The uncertainty on the heavy-flavor tagging efficiency is the same as that for the $t\bar{t}$ tagging efficiency described in Section 8.1 above. Due to the correlation, I include the uncertainty on the heavy-flavor tagging efficiency with that of the $t\bar{t}$ tagging efficiency when determining the effect on the cross section measurement.

Uncertainties on the QCD background prediction are determined using the level of agreement predicted and measured events in 'Region F', as described in Section 7.3.1. From the results, I assign a systematic uncertainty on the F_{QCD} measurement of 11 percent for electrons and 120 percent for muons, given conservatively by the worst agreement of the Region F prediction in each case. I fold this in with the statistical uncertainty on the F_{QCD} determination, the uncertainty on the correction factor k, both given in Table 7.6, and the 5 percent systematic uncertainty due to the application of the mistag matrix. The total QCD background uncertainty is 124 percent and 19 percent for muons and electrons, respectively. I add in quadrature the separate effects on the cross section of the QCD uncertainties for electrons and muons. The estimate of the QCD background is correlated with the estimates of the mistags and W+heavy-flavor backgrounds (Equations 7.2, 7.3, and 7.6) and this is taken into account when determining the effect on the $t\bar{t}$ cross section.

The systematic uncertainty on the small Drell-Yan background is determined by the statistical uncertainty of the estimate. Uncertainties on the Monte Carlo background predictions come from uncertainties in the cross sections for the various processes and from the event sizes of the Monte Carlo samples. This uncertainty is reflected in the uncertainties quoted in Section 7.5. The combined uncertainty on the Drell-Yan and Monte Carlo backgrounds is 8 percent.

The systematic uncertainties are summarized in Table 8.2.

Source	Fractional Sys. Uncert. (%)	$\Delta\sigma_{t\bar{t}}$ (%)
Acceptance Modeling and SLT Tagging Efficiency	7.7	$^{+8.8}_{-8.0}$
Mistag Matrix Prediction	5	3.8
$Wb\bar{b} + Wc\bar{c}$+Wc Prediction	26	3.9
QCD Prediction	19 (e) 124 (μ)	1.2
Drell-Yan and other MC backgrounds	8	0.6
Total Systematic Uncertainty		$^{+10.4}_{-9.7}$

Table 8.2: Summary of systematic uncertainties.

The uncertainty on the luminosity is 5.9 percent, where 4.4 percent comes from the acceptance and operation of the luminosity monitor and 4.0 percent from the calculation of the total $p\bar{p}$ cross section [39].

Chapter 9

Results

9.1 Cross Section Calculation

The cross section is calculated as

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bckg}}{A_{t\bar{t}} \cdot \epsilon_{t\bar{t}} \cdot \int \mathcal{L} dt}, \tag{9.1}$$

where N_{obs} is the number of events with ≥ 3 tight jets that have at least 1 SLT tag, N_{bckg} is the corrected background, $A_{t\bar{t}}$ and $\epsilon_{t\bar{t}}$ are the $t\bar{t}$ event acceptance and tagging efficiency, and $\int \mathcal{L} dt$ is the integrated luminosity.

The estimated number of background events must be corrected for $t\bar{t}$ events in the pretag sample because I apply the mistag matrix to the events before tagging to estimate the mistag and QCD backgrounds and also use the pretag sample in the W+heavy-flavor background determination assuming no $t\bar{t}$ signal events in the pretag sample. A simple iterative procedure is used to correct the mistag, QCD, and W+heavy-flavor backgrounds (the other background estimates are made independent of the pretag data and so need no correction). The $t\bar{t}$ contribution to the pretag sample is first estimated as the number of tagged events minus the total background in 3-or-more-jets events, all divided by the $t\bar{t}$ tagging efficiency. This contribution is subtracted from the total number of events to which the mistag matrix is applied (after accounting for the difference in number of taggable tracks in W+jets events and $t\bar{t}$ events, explained below), and W+heavy-flavor background normalized. A new estimate for the tagged background is calculated and the number of $t\bar{t}$ candidates is re-evaluated. This procedure is repeated until the total background estimation changes by less than 0.01 events, typically after a few iterations. As noted above, the correction takes into account the different SLT-taggable-track multiplicities in generic W+jets events and $t\bar{t}$ events. This is done because the probability per event of an SLT mistag is approximately proportional to the number of taggable tracks per event. Comparison of W+jets and $t\bar{t}$ simulations indicate that the mean multiplicity of $t\bar{t}$ events is 53 percent larger than W+jet events when 3 or more jets are

required (see Figure 9.1).

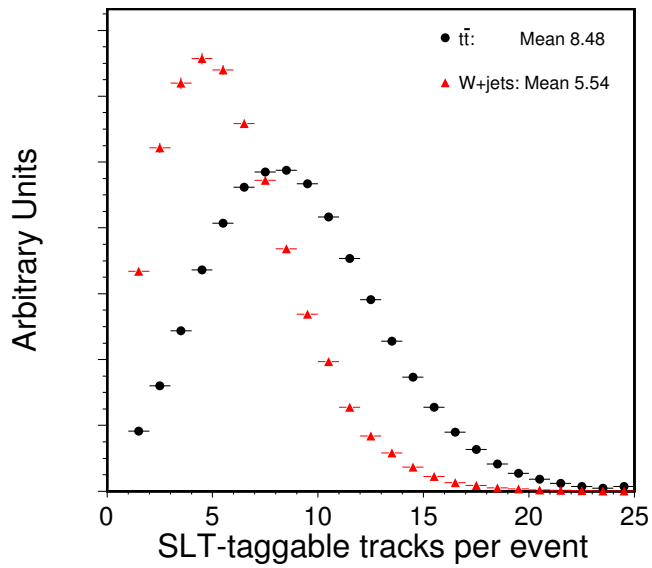


Figure 9.1: The number of SLT-taggable tracks in events with 3 or more jets from $t\bar{t}$ (circles) and W+jets (triangles) simulations. The mean number of taggable tracks in $t\bar{t}$ events is 8.48, 53 percent higher than the mean of 5.54 taggable tracks in W+jets events.

9.2 Observed Tags and Measured Cross Section

Prior to calculating the cross section, the background and signal estimates are compared to the total number of observed events. The contribution from $t\bar{t}$ in events with one or two jets is very small, so those events are used as a control sample to test the background predictions. Table 9.1 shows a summary of the background estimates and the number of SLT tagged events as a function of the number of jets. A tagged event is an event with at least one tagged jet. The total background and the $t\bar{t}$ expectation are also listed. The line labeled "Corrected Background" corresponds to the background after correcting for the $t\bar{t}$ content of the pretag sample, as described above. These results are also visually represented in Figure 9.2, which shows the number of tags in 1, 2, 3, and ≥ 4 jet events together with histograms representing the total corrected background and $t\bar{t}$ signal scaled to the measured cross section.

I calculate the cross section using Equation 9.1, given above. The acceptance and efficiency are discussed in Section 5.7. The acceptance and efficiency results for 3-or-more-jets events are summarized in Table 9.2. Together with the integrated luminosity they give a total denominator for

	$H_T \geq 0$ GeV		$H_T \geq 200$ GeV		
Background	1 jet	2 jet	3 jets	$\geq$ 4 jets	$\geq$ 3 jets
Taggable events	75595	18264	2587	1120	3707
Mistags	622.3±31.4	225.9±11.6	53.03±2.68	31.44±1.59	84.47±4.26
Wbb+Wcc+Wc	145.2±38.4	65.72±17.47	14.98±3.98	8.37±2.23	22.61±6.00
QCD multijet	91.95±16.51	44.89±10.35	7.00±1.50	4.06±0.89	11.06±2.39
WW+WZ+ZZ	3.80±0.44	6.98±0.66	1.21±0.23	0.64±0.14	1.88±0.30
Drell-Yan$\rightarrow \tau^+\tau^-$	2.65±0.57	1.54±0.43	0.65±0.28	0.13±0.05	0.65±0.27
Drell-Yan$\rightarrow \mu^+\mu^-$	6.02±1.25	4.12±0.88	0.82±0.44	0.00±0.19	0.82±0.48
Single top	13.08±1.16	26.99±1.98	6.42±0.55	1.71±0.19	8.13±0.68
Total Background	885.0±53.6	376.2±24.1	84.1±5.4	46.3±3.0	129.6±8.3
Corrected Background	–	–	86.8±5.6		86.8±5.6
$t\bar{t}$ Expectation (σ =6.70)	2.60±0.33	23.5±1.8	50.1±3.6	74.2±6.5	124.3±9.1
Total Background + $t\bar{t}$	887.6±53.6	399.6±24.2	211.1±10.7		211.1±10.7
Tagged events	892	384	142	106	248

Table 9.1: Number of tagged events and the background summary. The uncertainty on the total background is not a simple sum in quadrature of the individual backgrounds because of the correlation between the mistag, W+heavy-flavor, and QCD background predictions.

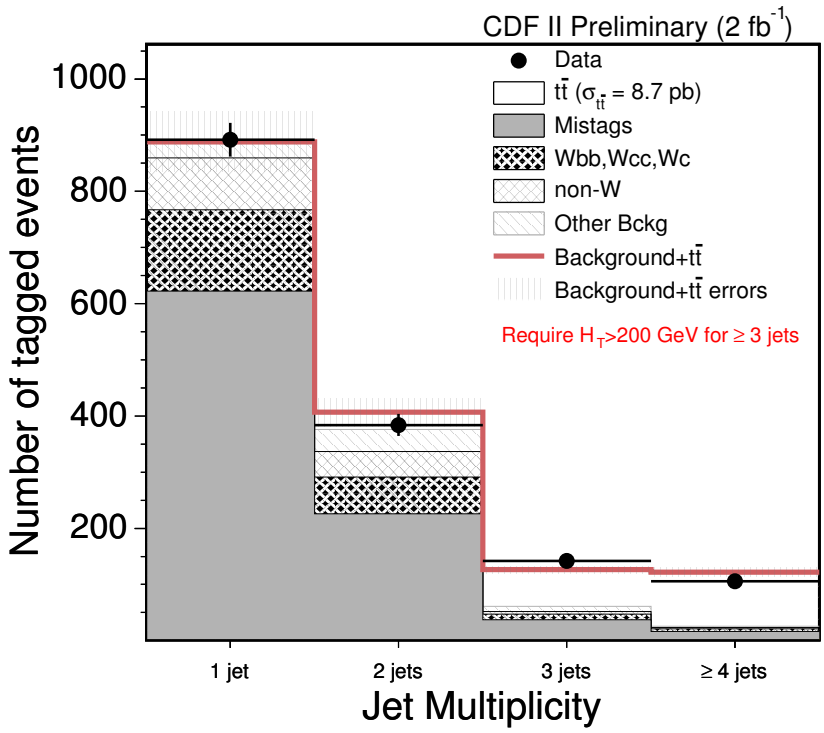


Figure 9.2: The expected background and observed tags in 1, 2, 3, and 4-or-more jets events. The background is corrected for the $t\bar{t}$ content of the pretagged sample. The expected $t\bar{t}$ contribution is normalized to the measured cross section.

the cross section calculation with a value of 18.6 pb^{-1}. From the number of candidate events with three or more jets and the expected background, I find a total $t\bar{t}$ cross section of

$$\sigma(p\bar{p} \to t\bar{t}X) = 8.7 \pm 1.1^{+0.9}_{-0.8} \pm 0.6 \text{ pb}, \tag{9.2}$$

where the first uncertainty is statistical, the second is systematic, and the third is from the luminosity. The statistical uncertainty on the measurement includes the usual counting uncertainty, but also includes an additional piece due to the background correction which also depends on the number of tags counted. The statistical uncertainty is evaluated by running pseudo-experiments. The Poisson uncertainty alone gives an uncertainty of 0.8 pb; the background correction inflates this to 1.1 pb. This cross section value uses acceptances and tagging efficiencies appropiate for a top mass of 175 GeV/c^2. The acceptances and efficiencies, and therefore the calculated cross section, change slightly for other assumed top masses. The calculated cross section is 3% higher assuming a top mass of 170 GeV/c^2, and 4% lower assuming a top mass of 180 GeV/c^2.

	Electrons	CMUP Muons	CMX Muons
Raw Acc. no Tag (%)	4.02 ± 0.01	2.51 ± 0.01	1.090 ± 0.005
Corr. Acc. no Tag (%)	$3.71 \pm 0.01 \pm 0.21$	$2.05 \pm 0.01 \pm 0.14$	$0.946 \pm 0.004 \pm 0.050$
Event Tagging Eff. (%)	$14.02 \pm 0.08 \pm 0.72$	$13.07 \pm 0.10 \pm 0.67$	$13.38 \pm 0.16 \pm 0.68$
Acc. with Tag (%)	$0.520 \pm 0.003 \pm 0.039$	$0.268 \pm 0.002 \pm 0.022$	$0.127 \pm 0.002 \pm 0.009$
Luminosity (pb^{-1})	2033.6 ± 119.6	2033.6 ± 119.6	1992.5 ± 117.2
Denominator (pb^{-1})	$10.58 \pm 0.07 \pm 0.80 \pm 0.62$	$7.97 \pm 0.06 \pm 0.49 \pm 0.47$	
Total denominator (pb^{-1})	18.56 ± 0.09(stat.)± 0.94(sys.)± 1.09(lum.)		

Table 9.2: Summary of components of the denominator for the cross section calculation. The $t\bar{t}$ acceptance and tagging efficiency for 3-or-more-jets events is determined using PYTHIA Monte Carlo.

As a cross check I have also measured the cross section for the electron and muon channels separately. Table 9.3 summarizes the results. The cross section measurements in the electron and muon samples agree within their statistical uncertainties.

Channel	$\sigma_{t\bar{t}}$ (pb)
Electron	9.1±1.2
Muon	8.1±1.2
Combined	8.7±0.8

Table 9.3: Summary of $t\bar{t}$ cross sections split by primary-lepton channel. The uncertainties shown are the Poisson statistical uncertainties only.

9.3 Kinematic Features of the SLT-Tagged Sample

I have examined a number of kinematic distributions of the SLT tagged events and compared them with expectations based on the measured signal plus background. Figure 9.3 shows the E_T distribution of the tagged jets in 1- and 2-jet events (combined), and in the signal region of three-or-more-jets events. There is a small discrepancy at low E_T in the 3-or-more-jets sample. The discrepancy may be due to an underprediction in the mistags for jets with low E_T, which was seen in the tests of the mistag matrix (see Figure 6.12). As noted in Section 5.5, $t\bar{t}$ events have significantly larger total energy, H_T, than do W+jets events. Figure 9.4 shows the H_T distributions for tagged events.

The kinematic features of the SLT tags themselves are also of interest. The p_T distributions of the SLT tags is shown in Figure 9.5. Figure 9.6 shows the impact parameter significance, defined as the impact parameter divided by its uncertainty, for the SLT tracks and the expectation from signal plus background. The sign of the impact parameter is defined according to whether the track trajectory crosses the jet axis in front of or behind the primary event vertex. The long-lived component from semileptonic b-hadron decays is readily apparent in the shape of the positive impact parameter distribution. Finally, Figure 9.7 shows the momentum of the SLT tracks transverse to the jet axis, p_T^{rel}. Due to the large mass of heavy-flavor quarks, their decay products receive a bit of a kick which results in tracks with a larger p_T^{rel}, on average, than those in light-flavor jets.

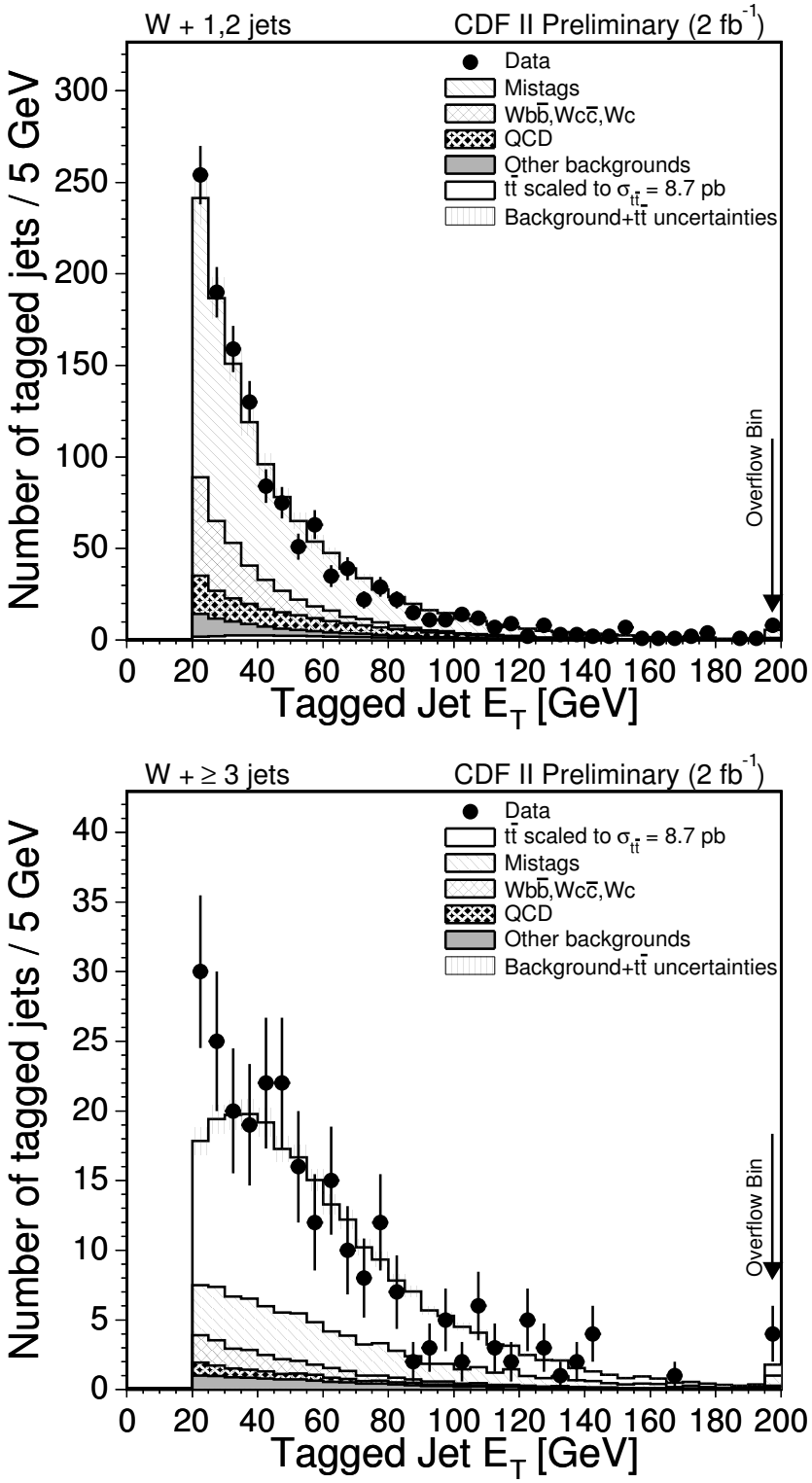


Figure 9.3: Comparison of the jet E_T distributions for tagged jets and for expectations from mistags, W+heavy-flavor, QCD and $t\bar{t}$ events. The upper plot is for 1- and 2-jet events and the lower plot for 3-or-more-jets events.

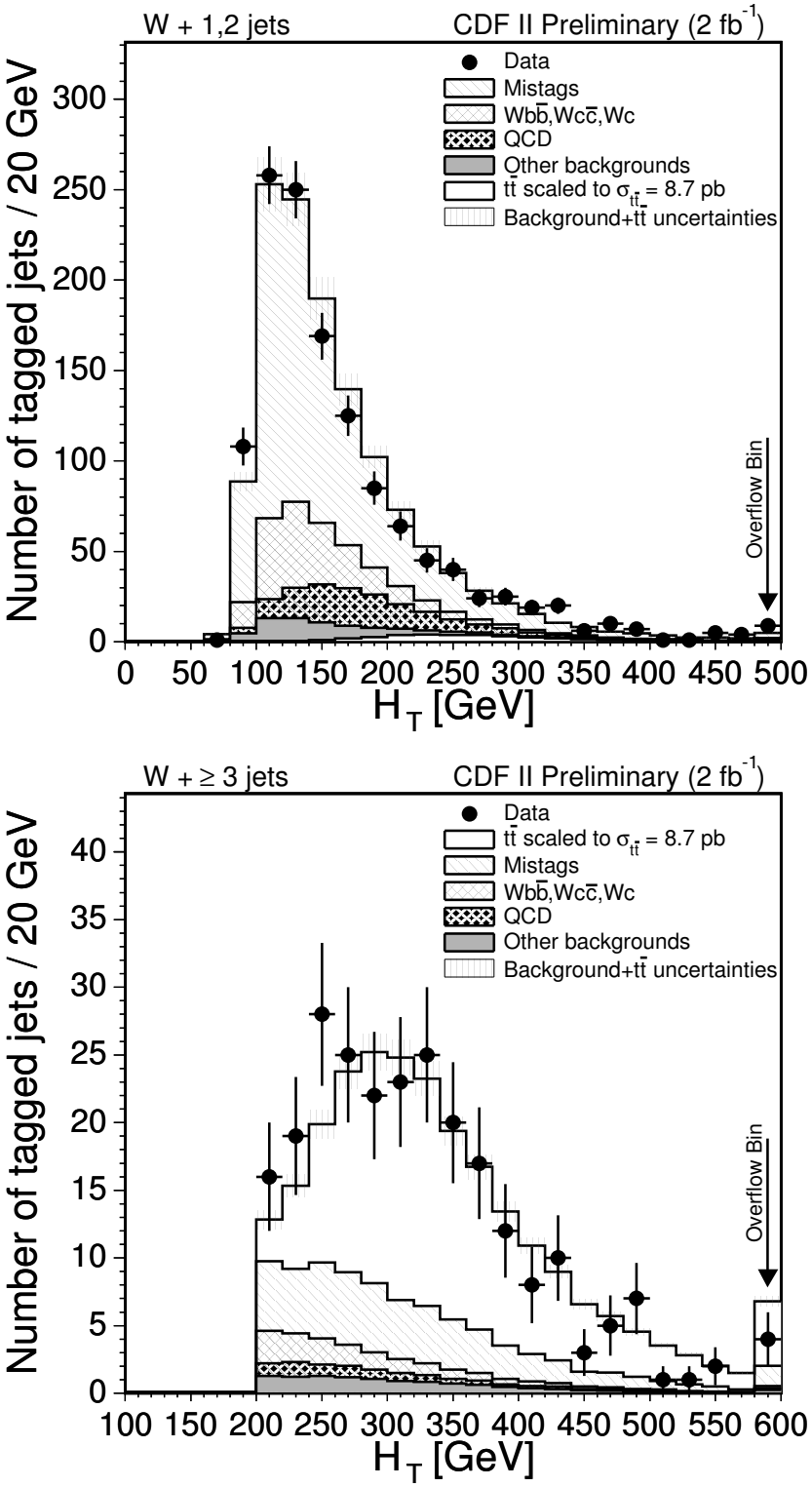


Figure 9.4: Comparison of the H_T distributions for tagged events and for expectations from mistags, W+heavy-flavor, QCD and $t\bar{t}$ events. The upper plot is for 1- and 2-jet events and the lower plot for $\geq$ 3-jet events.

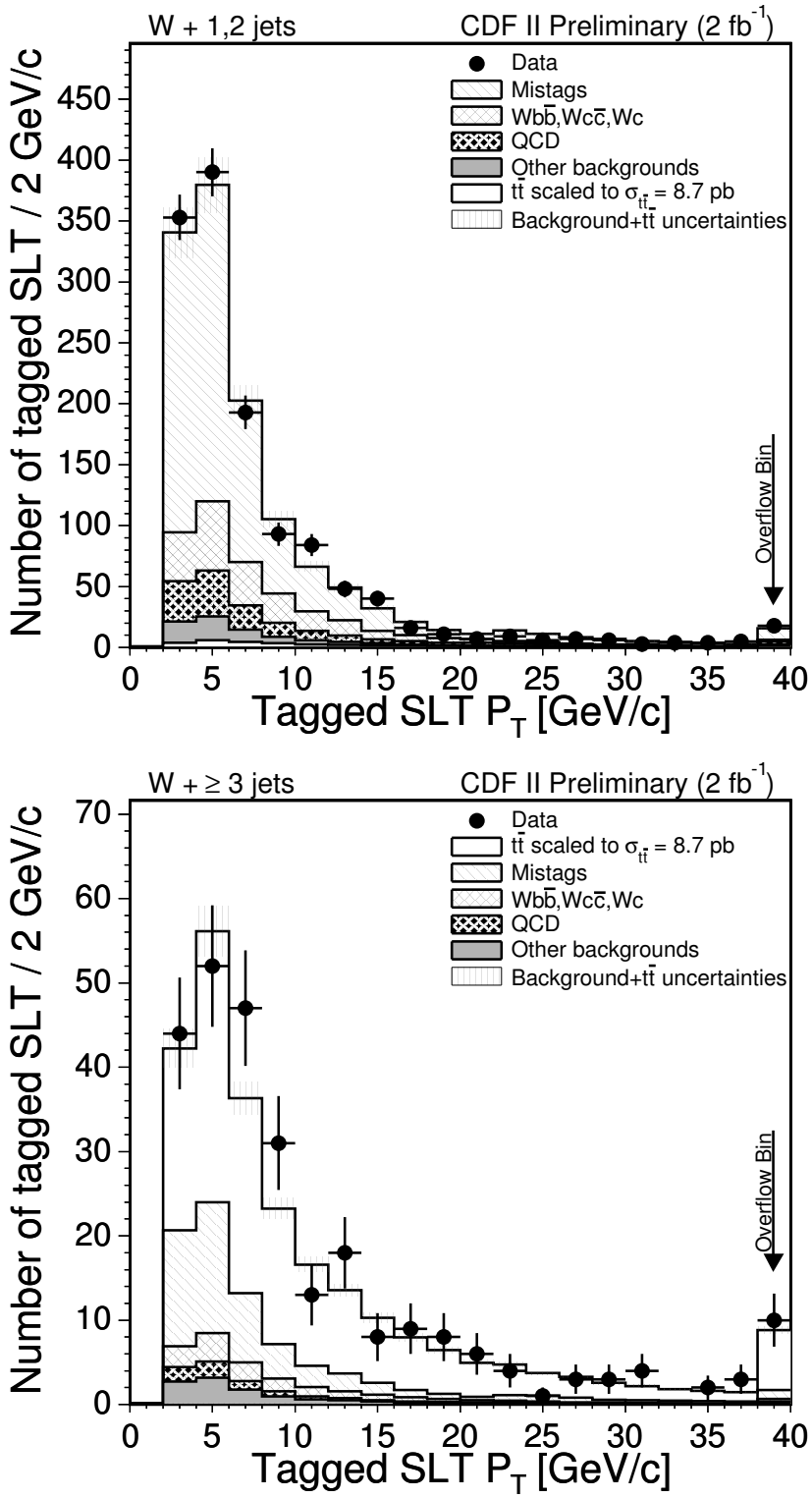


Figure 9.5: p_T of the SLT tags. The upper plot is for 1- and 2-jet events and the lower plot for $W+ \geq 3$-jet events.

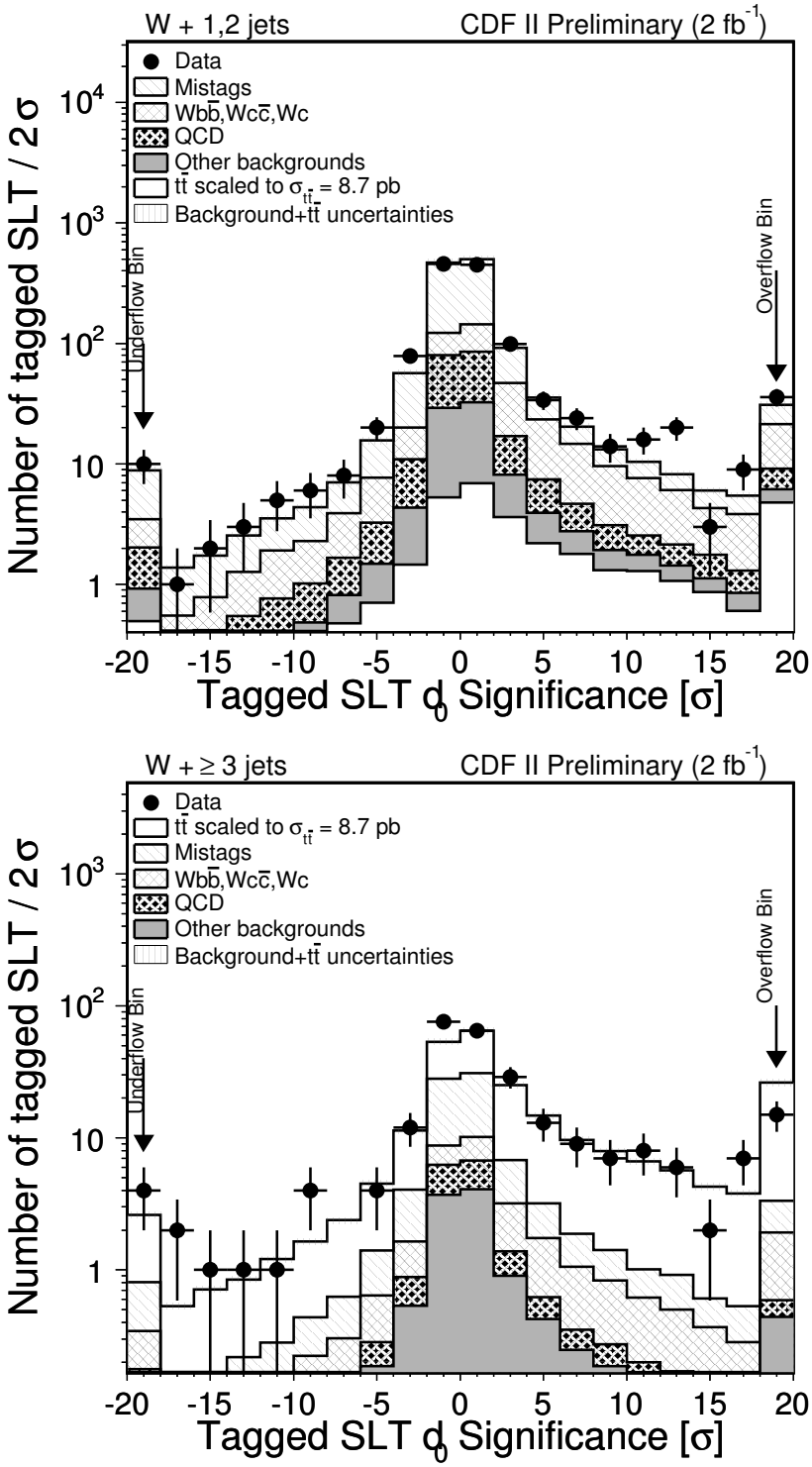


Figure 9.6: The impact parameter (d_0) significance for tagged tracks, compared with expectations from backgrounds plus $t\bar{t}$. The upper plot is for 1- and 2-jet events and the lower plot for $\geq$ 3-jet events.

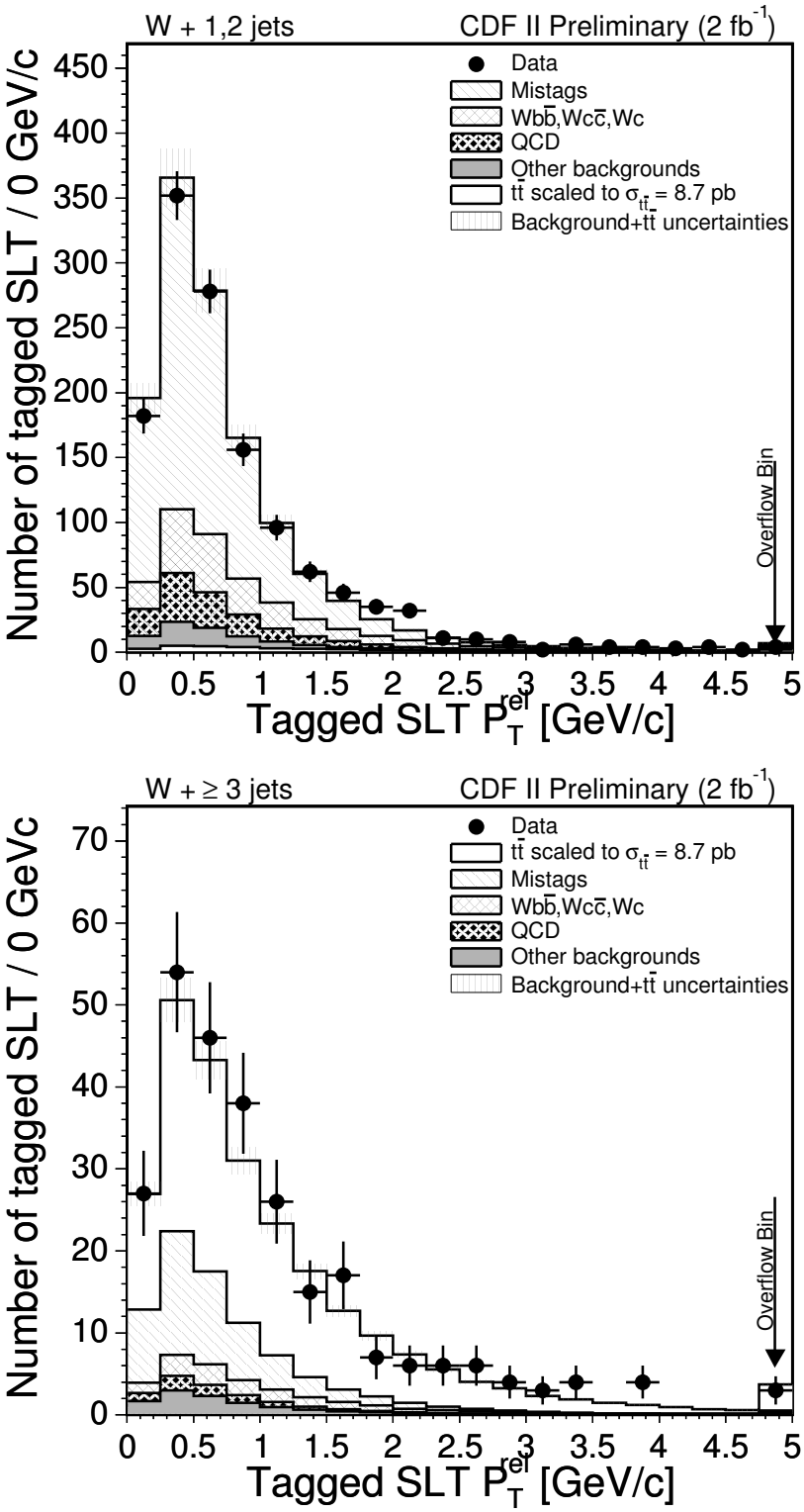


Figure 9.7: The p_T^{rel} distribution for tags in data, compared with expectations from backgrounds plus $t\bar{t}$. The upper plot is for 1- and 2-jet events and the lower plot for $\geq$ 3-jet events.

Chapter 10

Conclusions

Using ~ 2 fb^{-1} of data collected by the CDF detector, I have measured the total cross section for $t\bar{t}$ production in $p\bar{p}$ collisions with a center-of-mass energy, $\sqrt{s} = 1.96$ TeV. I examine events with a signature consistent with the decay of top pairs into an electron or muon plus multiple jets and separate signal from background by identifying semileptonic decays of b hadrons into muons. The measured $t\bar{t}$ cross section is

$$\sigma(p\bar{p} \rightarrow t\bar{t}X) = 8.7 \pm 1.1^{+0.9}_{-0.8} \pm 0.6 \text{ pb},^{1} \tag{10.1}$$

consistent with the expectation of $6.7^{+0.7}_{-0.9}$ pb for Standard Model production and decay of top quark pairs with a mass of 175 GeV/c^2. Assuming the cross section increases 0.2 pb for every GeV/c^2 decrease in the top mass, then at the world average top mass of 172.6 GeV/c^2 the theoretical cross section is approximately 7.2 pb. Using a linear fit to the mass dependence, the measured cross section can also be extrapolated to the world average top mass and is found to be 8.8 ± 1.6 pb. The kinematic distributions of the tagged sample are also consistent with Standard Model expectations. The observed number of tags in 1- and 2-jet events is in excellent agreement with expectations from background, indicating that the backgrounds are well understood.

For this analysis, I made a significant effort to precisely estimate the background from mistags, the largest background of the SLT-based cross section measurement. The level of understanding of the mistag background was one of the larger systematic uncertainties in the past. This is the first time that SLT mistags have been estimated separately from the $Wb\bar{b} + Wc\bar{c} + Wc$ backgrounds using data. Samples of pions, kaons, and protons were collected from D* and Λ^0 decays and used to measure the mistag rate from hadronic punch-through and decays-in-flight. The mistag matrix predicted well the number of tags in generic jets, and the precision of the estimate was limited by the statistical uncertainty of the mistag matrix. Even so, the uncertainty on the mistag background was

[1] At the time of this writing, there is an asymmetry in the charge of the primary leptons in tagged events that is under investigation. The asymmetry may be a result of a statistical fluctuation, but if not, it may slightly modify the cross section result.

cut in half, though at the expense of a larger uncertainty on the background from W+heavy-flavor events. As new data is collected, the mistag uncertainty will decrease, but the uncertainty may also be able to be reduced without waiting for new data by obtaining samples of pions, kaons, and protons, from other decays than those used in this analysis—for example, pions and kaons from D^+ decays.

Future measurements with the full Run II dataset of 4–8 fb^{-1} will provide a factor of up to approximately two in statistical precision. However, as data continues to accumulate, systematic uncertainties will soon become the limiting factor in the precision of the $t\bar{t}$ cross section measurement. The largest systematic uncertainties are on the modeling of acceptance and the SLT-tagging efficiency, along with the luminosity measurement. Efforts are ongoing to better understand systematic effects on the $t\bar{t}$ acceptance, in particular the jet-energy scale, the largest contributor to the uncertainty. The uncertainty on the SLT-tagging efficiency is dominated by the lack of precise knowledge of the tracking efficiencies in jets in data and in Monte Carlo. Better understanding and correction for this effect will be important to improving the precision of the cross section measurement using soft muon tagging. Improved precision is important for several reasons. Precise measurements of the $t\bar{t}$ cross section and the top mass can be checked for consistency with Standard Model top and QCD calculations of the dependence of the cross section on the top mass (see Figure 10.1). Measurements of the $t\bar{t}$ cross section made at the Tevatron will also be extrapolated to the energies of the LHC. The more precise the measurements now, the easier it will be to understand the top signal at the LHC where it may be an important background to new physics.

Whether or not the $t\bar{t}$ sample from the Tevatron holds evidence of physics beyond the Standard Model, careful study of the top quark and its properties is essential to the searches for new physics at future hadron colliders. The anticipated signatures for new physics are similar to those of the top quark, including high-energy leptons and multiple jets, and large $\not{E}_t$. This means that all of the backgrounds to the top signature, as well as top production itself, will be backgrounds to at least some new physics processes. Making measurements of the top-quark cross section is one of the best methods available for understanding these backgrounds. As we continue to push the energy frontier, top-quark measurements will remain vital to our understanding of the fundamental constituents of the universe and their interactions, and perhaps to uncovering what lies beyond the Standard Model.

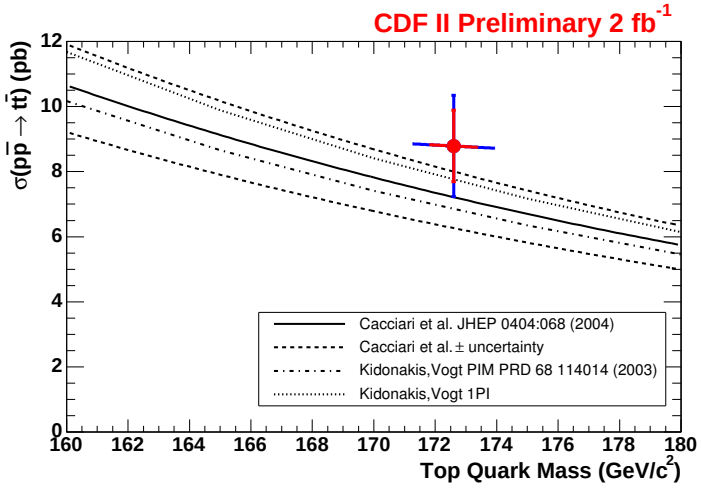


Figure 10.1: Standard Model predictions of the $t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV as a function of the top quark mass. Overlaid is the measured $t\bar{t}$ cross section extrapolated to the world average top mass of 172.6 GeV. The extrapolated cross section is $\sigma_{t\bar{t}} = 8.8 \pm 1.6$ pb.

Appendix A

SLT Tagger Parametrization

This appendix describes the parametrization of the widths of matching variables used in the SLT tagging algorithm. The width of the matching variables at low p_T is mostly determined by multiple scattering. Therefore, they are functions of the muon momentum (not p_T) and the amount of material traversed by the muon. At higher p_T, the intrinsic resolution of the muon detectors becomes the main determinant of the matching variable widths.

A.1 Stub-Track Matching Variables in CMU

A.1.1 CMU dx

For the CMU, the width of dx follows the reciprocal of p_T to first order. There is a slight exponential rise at very low p_T, so the width is parametrized as:

$$\sigma_{\mathrm{d}x}^{\mathrm{CMU}} = \frac{a + e^{b+c\cdot p_T}}{p_T} + d, \tag{A.1}$$

where the constant term accounts for the intrinsic x resolution of the CMU. The values of the fit parameters are:

$a = 14.65$
$b = 2.555$
$c = -0.889$
$d = 0.048$

No ϕ dependence is observed for dx in the CMU and only a small η dependence (less than 3 percent) for tracks with very low p_T. The small η dependence is due to the increased amount of material (and thus multiple scattering) traversed by a track with larger η. However, when requiring constant p_T, larger η means that $p = p_T/\sin(\theta)$ is larger, reducing multiple scattering. The result is that the width of the distribution remains approximately constant over η.

A.1.2 CMU dz

Since the dz is diagonal with respect to the flight direction, its distribution is not expected to be Gaussian. However, d$z \times \sin(\theta)$, which is the projection orthogonal to the direction of flight, is expected to be Gaussian. Therefore, d$z \times \sin(\theta)$ is parametrized as a function of p_T as:

$$\sigma^{\mathrm{CMU}}_{\mathrm{d}z\times\sin(\theta)} = a + \frac{b}{p_T} + \frac{c}{p_T^2}, \tag{A.2}$$

where the constant term accounts for the intrinsic CMU-z resolution and the p_T terms are for the multiple scattering. The values of the fit parameters are:

$a = 8.33$
$b = -5.25$
$c = 22.96$

A.1.3 CMU dϕ_L

The dϕ width is fit to the function:

$$\sigma^{\mathrm{CMU}}_{\mathrm{d}\phi_L} = a + \frac{b}{p_T} + \frac{c}{p_T^2}. \tag{A.3}$$

The SLT algorithm also accounts for a different dϕ resolution of 3-hit stubs versus 4-hit stubs. The parametrization of the width of dϕ_L for 3-hit stubs is multiplied by a factor of 1.3 since the width of dϕ_L in 3-hit stubs is seen to be 30 percent broader than 4-hit stubs at comparable p_T. The values of the fit parameters are:

$a = 0.0120$
$b = 0.0330$
$c = 0.352$

A.2 Stub-Track Matching Variables in CMP

The parametrization of widths in the CMP is a little more complicated due to its box geometry. The amount of material traversed depends on the ϕ angle of the muon as seen from the interaction point.

A.2.1 CMP dx

The p_T dependence of dx in the CMP is first measured at normal incidence, $|\phi| < 20°$. Then the ϕ dependence is measured. Including the track ϕ modulation, the width is parametrized as:

$$\sigma_{\mathrm{d}x}^{\mathrm{CMP}} = A(p_T) + B(p_T)[1 - \sin(4\phi + \pi/2)], \tag{A.4}$$

with $A(p_T) = \frac{a+e^{b+c\cdot p_T}}{p_T} + d$ and $B(p_T) = \max[(e + f \cdot p_T), 0]$. $B(p_T)$ is the p_T dependent amplitude of the ϕ modulation. The values of the fit parameters are:

$a = 39.38$
$b = 4.55$
$c = -0.782$
$d = 0.102$
$e = 2.8$
$f = -0.27$

A.2.2 CMP dϕ_L

The p_T dependent parametrization of the width of dϕ_L for the CMP is given as:

$$\sigma_{\mathrm{d}\phi_L}^{\mathrm{CMX}} = a + \frac{b}{p_T} + \frac{c}{p_T^2}. \tag{A.5}$$

As with the CMU-dϕ_L there is a difference in the resolution for 3-hit and 4-hit stubs. This is taken into account by multiplying the dϕ_L for 3-hit stubs by a factor of 1.2. The values of the fit parameters are:

$a = 0.04022$
$b = -0.1853$
$c = 1.43$

A.3 Stub-Track Matching Variables in CMX

The parametrization of the matching variables for the CMX has a dependence on η due to the decrease in material absorption length as η increases (see Figure 2.7). The width of each of the matching distributions—dx, dz, and dϕ_L—have been parametrized as a function of η and p_T with

the following formula:

$$\sigma^{\mathrm{CMX}}_{\mathrm{d}x,\mathrm{d}z,\mathrm{d}\phi_L} = \min(a + b \times p_T, 0) \times (|\eta| - 0.675) + c + \frac{d}{p_T} + \frac{e}{p_T^2}. \tag{A.6}$$

The values of the fit parameters for each matching distribution are listed below.

For the dx distribution of CMX:

$a = -43.58$
$b = 6.50$
$c = 0.366$
$d = 27.92$
$e = 77.71$

For the dz distribution of CMX:

$a = -33.01$
$b = 4.996$
$c = 6.522$
$d = -3.1$
$e = 87.91$

For the dϕ_L distribution of CMX:

$a = -0.2228$
$b = 0.04$
$c = 0.0074$
$d = 0.0695$
$e = 0.655$

Appendix B

Measurement of the Mistag Rate Using γ+Jets Events

In a previous measurement of the $t\bar{t}$ cross section using SLT tagged events backgrounds were estimated from mistags and from semileptonic heavy-flavor decays in $Wb\bar{b}$, $Wc\bar{c}$, and Wc events together. This was done by building a 'mistag' matrix using γ+jets events that was intended to predict both mistags and semileptonic decays. This appendix summarizes the results of that study as a comparison with those of the method presented in the main body of this thesis. To distinguish the two mistag matrices, I will refer to the matrix described in this appendix as the 'photon' matrix and the matrix described in Chapter 6 as the 'heavy-flavor-free' (or HF-free) matrix when necessary.

B.1 Mistag Probability

Just as with the HF-free matrix, the mistag rate for the photon matrix is defined as the number of tracks tagged as muons, divided by the number of SLT taggable tracks. Tracks are also required to be within a cone $\Delta R < 0.6$ of a tight jet axis (including when building the matrix, unlike the HF-free matrix). For the photon matrix (and the analysis in which it was applied), a tight jet has corrected measured energy greater than 15 GeV. I should note that for this previous measurement a different set of jet corrections (known as "Level 4" corrections) were applied than those used for the analysis presented in the main part of this thesis ("Level 5" corrections). The cut at 15 GeV with the Level 4 jet corrections is approximately equivalent to the 20 GeV cut with Level 5 corrections.

The photon matrix is built using jets in the γ+jets (photon) sample. The photon sample is triggered on one photon with energy greater than 15 or 25 GeV, and contains typically one or more jets. The sample has a well identified trigger object, of energy comparable to the boost of a W in a W plus jets event.

B.2 Mistag Matrix

The SLT mistag probability, as measured in the γ+jets sample, is about 0.53 percent per-track.

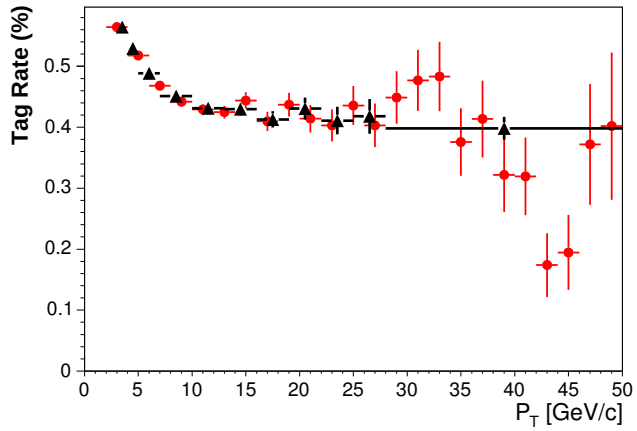


Figure B.1: The mistag rate (red circles) as a function of track p_T, seen using the γ+jets sample. The matrix prediction (black triangles) is overlaid with the histogram bins matching the binning of the matrix.

The mistag matrix is a parametrization of the tag rate used to predict the number of tags given a set of taggable tracks. The rate is parameterized as a function of track p_T (shown in Figure **??**), detector η and ϕ. The photon matrix has 11 bins in p_T, nine in η, and seven in ϕ. Therefore, the photon mistag matrix is a 693-bin, three-dimensional probability matrix, nominally defined as:

$$M_{ijk} = \frac{n_{ijk}^{\text{Tags}}}{N_{ijk}^{\text{Taggables}}}. \tag{B.1}$$

To calculate the matrix, I use a sample of jets from the γ+jets dataset. I require 1 or more tight jets to be present in the event. While there are triggers that require higher-E_T photons, I only make use of the 15 and 25 GeV photon triggers to better match the E_T spectrum of the triggered object to that of the W in W+jets events. Also, to reduce the bias in the rate due to the trigger requirement, I raise the jet-energy threshold from 15 GeV to 20 GeV for events triggered on 25 GeV photons. I have found this to be a good match to jets counted above 15 GeV in a W sample.

Using a sample of ∼10.3M taggable tracks I find that the average rate across jet bins ranges from (0.4872±0.0025) percent in the trigger + 1-jet bin to (0.5812±0.0112) percent in the trigger + ≥3-jets bin. As trigger + 1-jet events dominate the sample, the tags in the trigger + ≥3-jets bin would be underpredicted if, I simply counted up all the taggable tracks and all the tagged tracks to calculate the mistag rate. Instead, I do this counting in each jet bin and take a simple average of

these rates, meaning that the final matrix is defined as:

$$M_{ijk} = \frac{M_{ijk}^{1-jet} + M_{ijk}^{2-jet} + M_{ijk}^{\geq 3-jet}}{3}. \tag{B.2}$$

where the M^{bin} are the probability matrices for each jet bin determined as in Equation. B.1. The average mistag rate I measure in this matrix is (0.5317±0.0073) percent.

B.3 Prediction of the Mistag Matrix and the Systematic Uncertainty

B.3.1 Test Samples

The mistag matrix is checked with several different samples and the number of predicted tags is compared to the number of measured tags. There is no need to remove heavy flavor from the jet samples when testing the photon matrix. For the gamma plus jets sample, which is used to calculate the mistag matrix, I check the prediction separately in events with a photon plus 1, 2 and ≥3 tight jets.

The same biases noted in Section 6.4.2 for the Jet20/50/70/100 samples apply here. The steps taken to avoid these biases are the same as noted in that section, except that for dijet events I raise the jet-energy threshold only to 5-GeV-below the trigger threshold. The ΣE_T sample is also used in this study with no measures taken to reduce any bias. A summary of the numbers from this study can be found in Table B.1.

The prediction of the photon matrix was also tested in Z+jets events. Checking the tags versus the expected number in the Z+jets sample is a useful exercise to directly test the mistag matrix to a sample very similar to W+jets, for which I want to measure the contribution as background to $t\bar{t}$ events. Figure B.2 shows the predicted and observed tags as a function of the number of jets. Unfortunately the test is statistically limited. Of the 39 observed tags the predicted number is 38.7 ± 0.5.

B.3.2 Summary of the comparisons and the systematic uncertainty

Table B.1 is visually summarized in Fig. B.3. The central line is the mistag matrix prediction. The measured tags always fall within ±10 percent of the estimated number of tags. Therefore a ±10 percent systematic uncertainty is assigned on the mistag prediction in the W+jets sample.

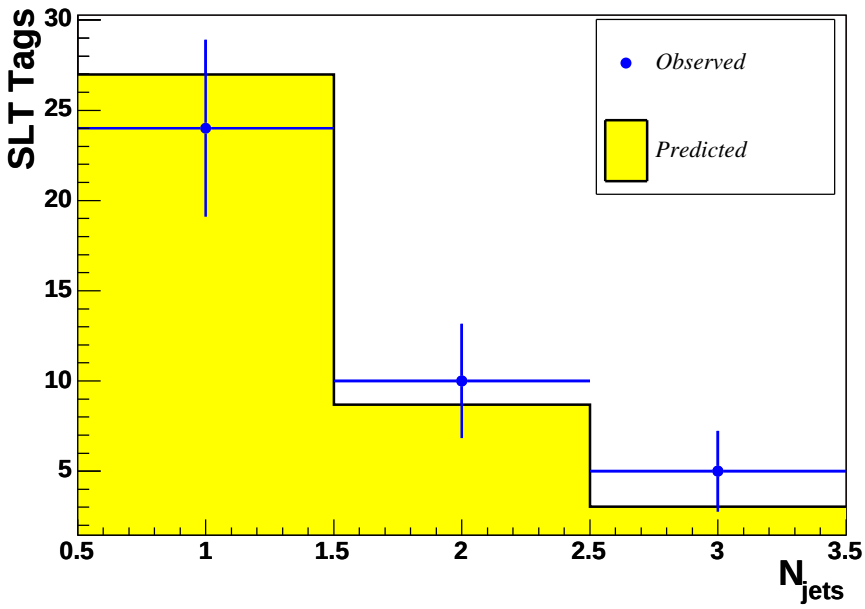


Figure B.2: The SLT mistag prediction in Z+jets events compared to the observed tags versus the number of jets in the event.

Sample		Trig.+1 jet	Trig.+2 jets	Trig.+≥3 jets
Jet 20	Observed	21863	3010	514
	Predicted	22498 ± 314	2977 ± 41	495 ± 7
	Δ (%)	2.8 ± 1.5	-1.1 ± 2.3	-3.8 ± 4.8
Jet 50	Observed	24191	2033	1557
	Predicted	25569 ± 362	2029 ± 29	1503 ± 21
	Δ (%)	5.4 ± 1.5	-0.2 ± 2.6	-3.6 ± 3.0
Jet 70	Observed	24588	2330	2722
	Predicted	25439 ± 383	2243 ± 31	2599 ± 36
	Δ (%)	3.3 ± 1.6	-3.9 ± 2.6	-4.7 ± 2.5
Jet 100	Observed	34383	3236	5597
	Predicted	33922 ± 563	3079 ± 43	5062 ± 70
	Δ (%)	-1.4 ± 1.8	-5.1 ± 2.4	-10.6 ± 2.1
γ+jets	Observed	39118	9628	2683
	Predicted	42692 ± 587	9700 ± 133	2450 ± 34
	Δ (%)	8.4 ± 1.3	0.8 ± 1.7	-9.5 ± 2.6
ΣE_T	Observed	9937	5677	20877
	Predicted	9999 ± 138	5491 ± 77	19493 ± 271
	Δ (%)	0.6 ± 1.7	-3.4 ± 2.0	-7.1 ± 1.7
Z+jets	Observed	24	10	5
	Predicted	26.99 ± 0.41	8.68 ± 0.16	3.03 ± 0.07
	Δ (%)	11 ± 18	-15 ± 36	-65 ± 74

Table B.1: Checks of the photon matrix with different samples. Within a sample, jets are selected in events with n-jets. The Δ is determined in percent of the predicted numbers, the uncertainty is derived from the Poisson fluctuation of the observed numbers.

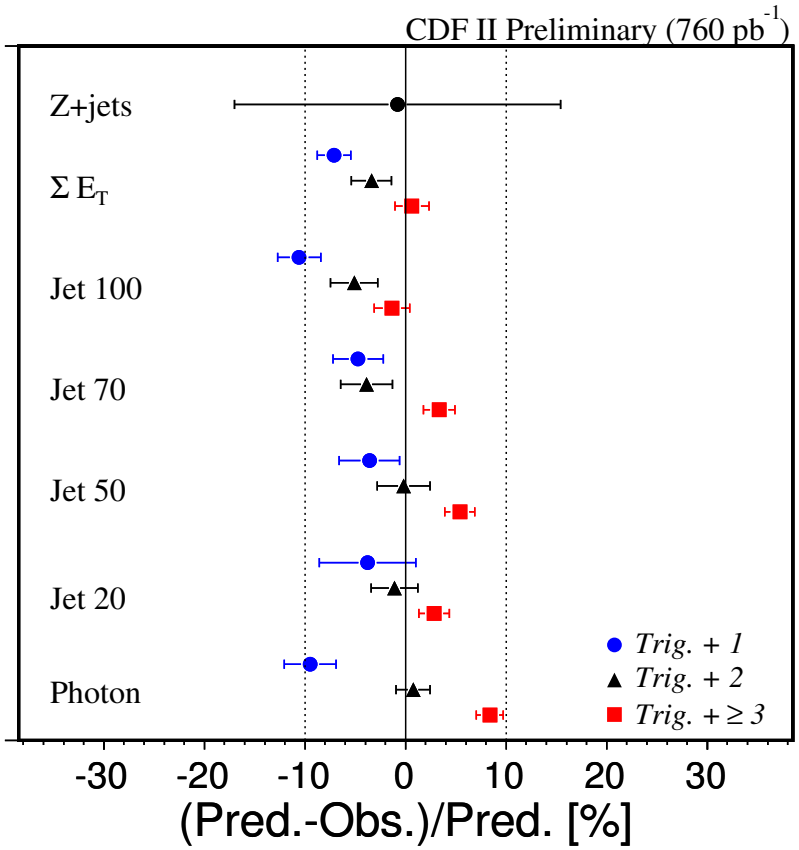


Figure B.3: Summary of the photon matrix prediction compared to the observed tags in several test samples. The square/triangle/circle markers indicate there are 1/2/3-or-more tight jets in the event in addition to the trigger object (photon or jet).

B.4 Fraction of Tags from Heavy-Flavor decays in the γ+Jets Sample

Method A

The first of two methods I use for determining the fraction of heavy-flavor tags, $F_{\rm HF}$, uses an iterative process for the calculation. I count events with both secondary vertex tags (SecVtx) and SLT tags, N^{SLT}_{SecVtx}, and want to subtract from that the predicted SLT mistags in SecVtx tagged jets, after accounting for heavy flavor, $N^{SLTmistag}_{SecVtx} \cdot (1 - F_{\rm HF})$. The iteration comes in because, of course, $F_{\rm HF}$ is unknown. The heavy-flavor fraction is calculated using this method as

$$F^{i+1}_{\rm HF} = \frac{N^{SLT}_{SecVtx} - N^{SLTmistag}_{SecVtx} \cdot (1 - F^{i}_{\rm HF})}{N_{SLT} \cdot \epsilon^{SecVtx}_{SLT}}, \tag{B.3}$$

where $F^{i}_{\rm HF}$ is 0 at the first iteration, N_{SLT} is the total number of SLT tags in the sample, and ϵ^{SecVtx}_{SLT} is the efficiency of SecVtx tagging an SLT tagged heavy-flavor jet (derived from a Jet50 Monte Carlo sample). $F_{\rm HF}$ converges after a few iterations and the results are found in Table B.2.

Method B

An alternative method starts again by first counting tags in the overlap of SecVtx and SLT tags. I then subtract from that overlap the predicted mistags of the SecVtx, $N^{SecVtxmistag}_{SLT}$.What remains after the subtraction represents a set of heavy-flavor jets that are tagged by the SLT. The SLT may mistag even a heavy-flavor jet (that is, it can identify a track as a muon, even though the track is not a muon from the semileptonic decay of a b or c hadron) so not all these tags are *real*. I can use the fraction real tags in a heavy flavor jet, ($k_{\rm real}$), known from Monte Carlo, to eliminate the light flavor tags, so that:

$$F_{\rm HF} = \frac{(N^{SLT}_{SecVtx} - N^{SecVtxmistag}_{SLT}) \cdot k_{\rm real}}{N_{SLT} \cdot \epsilon^{SecVtx}_{SLT}}. \tag{B.4}$$

The results are again found in Table B.2 and agree well with the first method.

The two methods are correlated and so cannot be trivially combined. Therefore, I use $F_{\rm HF}$ as determined by Method B, because it has a smaller statistical uncertainty, and add a systematic uncertainty of 2.1 percent to cover the difference between methods. With $F_{\rm HF}$ determined, the light-flavor-only probability, $P_{\rm LF}$, of the mistag matrix can now be estimated as the overall probability lowered by the factor $(1-F_{\rm HF})$, and the result is $0.53 \cdot (1-0.214) = 0.42 \pm 0.02(\text{stat.}) \pm 0.01(\text{syst.})$ percent. This probability, along with the probability in each jet bin, is shown in Table B.2. The

	Trig.+1 jet	Trig.+2 jets	Trig.+≥3 jets
SLT Tags	39118	9628	2683
SecVtx+SLT Tags	2856	906	283
SLT mistags (in SecVtx jets)	1043	334	113
SecVtx mistags (in SLT jets)	262	92	31
ϵ_{SLT}^{SecVtx} (%)	35.2±1.9	31.0±2.4	28.8±3.6
$k_{\rm real}$ (%)	62.2±1.9	80.3±2.0	80.3±3.1
$F_{\rm HF}$ - Method A (%)	16.2±2.0	24.6±3.3	29.6±5.6
$F_{\rm HF}$ - Method B (%)	12.7±1.3	23.5±2.9	28.1±4.9
$P_{\rm LF}$ (%)	0.43±0.01±0.02	0.40±0.02±0.01	0.42±0.03±0.01
Matrix			
$F_{\rm HF}$ - Method A (%)	23.5±3.9		
$F_{\rm HF}$ - Method B (%)	21.4±3.3		
$P_{\rm LF}$ (%)	0.42±0.02±0.01		

Table B.2: Relevant numbers for the study of the heavy-flavor fraction, $F_{\rm HF}$, of the photon matrix.

light-flavor-only probabilities, agree well between jet multiplicity bins.

The mistag rate in the photon matrix is not directly comparable to that in the HF-free matrix because of the inclusion of tags from heavy-flavor decays. However, the estimated light-flavor tag probability of $(0.42 \pm 0.02 \pm 0.01)$ percent is in excellent agreement with the mistag rate of (0.41 ± 0.01) percent of the HF-free matrix.

Appendix C

CMP Ageing Studies

During Run 1 and the early part of Run 2 the anode current draw of the CMP chambers was everywhere less than 1 μA, and typically only a few tenths of a μA . In late August or early September 2001 the CMP currents made a sudden jump, in response to a change in accelerator conditions (for which the cause remains unknown). Since that time, the currents on the walls and the top of CMP have measured between about 4 μA and about 12 μA depending on location on the detector and beam conditions. Only the well shielded part of CMP below the detector has retained currents below 1 μA, although there have been times when bad accelerator conditions have caused increased currents on the bottom as well. When there is no beam in the Tevatron, CMP currents return to their Run 1 levels.

A single CAEN high-voltage pod distributes anode high voltages to typically 40 or more wires, so the measured currents correspond to the summed currents on the wires of a single pod. In Run 1 the current draw was independent of the luminosity, an indication that the current was not from ionization in the chambers but from leakage current on the high-voltage distribution boards. The increase in currents in Run 2 therefore represents a very large increase in the ionization rate in the chambers (much larger than the ratio of Run 1 to Run 2 currents). This has raised concern about possible ageing of the CMP chambers (and probably CMX as well). In this appendix, I report on measurements of ageing effects in the CMP chambers. This study includes both direct measurements of ageing in a test stand via strontium-90 irradiation, and monitoring of any possible ageing in the chambers in B0 by tracking the width of hits on J/ψs.

C.1 ^{90}Sr Measurements

C.1.1 Setup and Procedure

Direct measurements of CMP chamber ageing characteristics were made by irradiating test chambers with 1 mCi ^{90}Sr sources and periodically monitoring the chamber gain with an ^{55}Fe source. Three

chambers were used in this test. Two were single CMX cells[1] (CMX and CMP drift cells are identical in construction) and the third was one cell of a four-cell CMP test stack that was built at the same time as the rest of CMP but is only six feet long, for ease of use in a test stand. The three chambers were used as follows. The CMP cell was used as a control chamber to monitor gain variations due to temperature, pressure and gas mixture. Regular ^{55}Fe gain runs were done on the control chamber, but it was not irradiated with ^{90}Sr. All ageing measurements were done by taking ratios of gains with respect to the control chamber gain. The two CMX cells were irradiated with ^{90}Sr and the anode wire currents were monitored vs. time in order to track the integrated charge on the wire. One of the CMX cells was run with pure 50-50 argon-ethane while the other was run with the argon-ethane bubbled through isopropyl alcohol at 0° C, which is how the CMP chambers are run in B0.

^{90}Sr irradiation was done through a slot milled in the aluminum wall of the chambers directly above the anode wire. Nearly all of the aluminum was removed in this slot in order to reduce the attenuation of the source. Because the drift cells have copper-clad Glassteel cathode pads on the top and bottom of the chamber, gain monitoring with ^{55}Fe could not be done through this slot because attenuation of the x-rays in the Glassteel was too severe. Instead a small hole was drilled in the side wall of each drift cell and the ^{55}Fe source mounted over this hole for gain monitoring.

During ^{90}Sr irradiation only the anode current is recorded. Read out of the anodes occurred only during ^{55}Fe runs. Due to the non-existent budget for this study, the data acquisition was something of a kludge. The chambers were read out through the standard CMP front-end preamp card. The preamp output was fed, through a special adapter board, to a VME-based CMU ASDII board. The analog output of the ASDII was fed to an oscilloscope and the pulses read out, at about 0.5 Hz, using LabWindows. LabWindows-based software, written for this study, was used to find and histogram the peak voltage of the oscilloscope pulse. Data from a typical ^{55}Fe run is shown in Figure C.1.

This very slow, one channel at a time, DAQ system did the job (barely), though a faster, multichannel DAQ with a proper ADC would have allowed us to conclude this study in half the time.

C.1.2 Results

Figure C.2 shows the results of the first set of gain vs. integrated charge measurements. The uncertainties plotted for each point are estimated from repeated measurements at the same integrated

[1] Thanks to Michael Schmitt and Abraham Gallas for donating the two CMX cells.

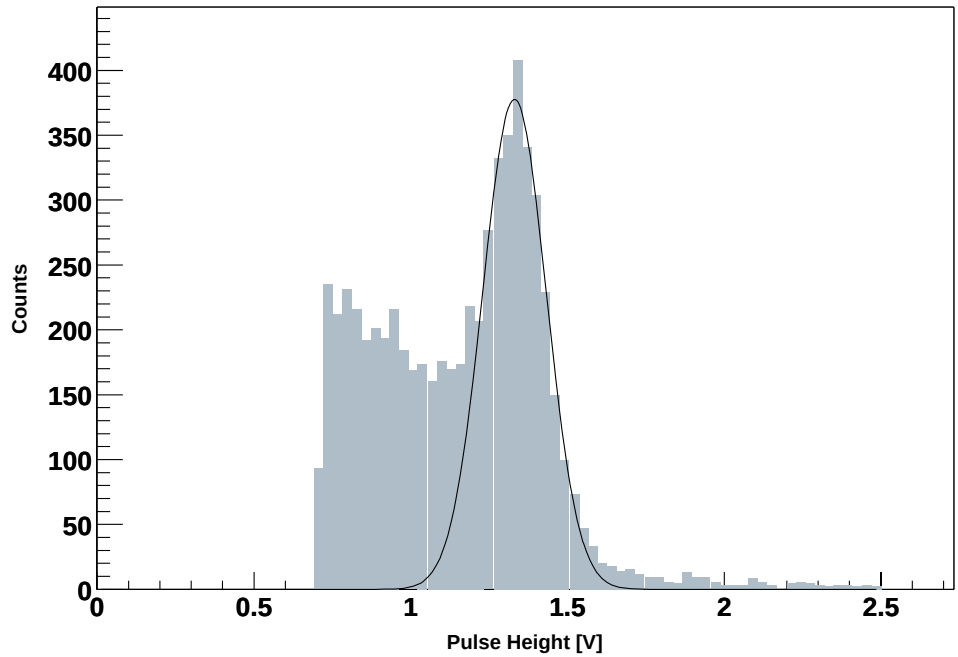


Figure C.1: Data from an ^{55}Fe run. The escape peak is fit to a Gaussian to determine the relative gain of the chamber.

charge values. In those cases where more than one measurement was made, the average value is plotted and the uncertainty is correspondingly smaller. The two sets of points correspond to the relative gains, with respect to the control chamber, of test chambers with and without isopropyl alcohol in the gas. One can see a clear decrease in the gain, approximately linear with integrated charge up to about 0.4 coulombs, for the chamber that was run with no alcohol. After integration of 0.4 coulombs, the gain reduction is about 60 percent, and it appears to plateau there, or even recover a little. No gain decrease was observed for the chamber with alcohol. The abscissa in these plots is coulombs of integrated charge, rather than coulombs/cm . The irradiating source sat approximately 0.4 inches above the wire and behind the Glassteel cathode. Converting the measurement to coulombs/cm requires a model of the illumination of the chamber by the source. Since there is no simple way to verify the validity of any such model, I plot what was measure directly. The goal is to determine if the chambers in B0 are in any danger due to ageing, and in this respect the

measurement represents something of a worse case scenario since the illumination by the source is almost certainly over a much more limited region of the anode wire than the illumination by beam spray at B0. The results are encouraging, given that no ageing effect is observed in the chamber run with alcohol. For comparison an 8 μA current, typical for an anode pod under current running conditions at B0, corresponds to 48 CMP wires and therefore 2.6×10^{-3} Coul/cm in a 10^7 sec year of running, assuming uniform illumination of the chambers. One has to invent a pretty unlikely pattern of illumination to get anywhere close to the integrated charge measured in these ageing tests.

During the first set of measurements the charge integrated between gain measurements was large, as much as 0.2 coulombs. The resulting gain loss of the non-alcohol chamber was rapid. The measurement was then repeated in a more controlled way. The non-alcohol chamber was re-strung with a new anode wire and the procedure repeated. The results are shown in Figure C.3. Qualitatively the results are the same: the chamber without alcohol suffers an almost linear decrease in gain with integrated charge, followed by something of a plateau, while the chamber with alcohol shows no effect. The slope of the gain decrease is, however, much smaller than in the original measurement with just a 5 percent. decrease in gain after 0.4 coulombs. In addition, the gain appears to be relatively constant after about 0.5 coulombs of integrated charge in both sets of measurements.

The two sets of measurements do not agree quantitatively and I do not have a convincing explanation for the difference. Qualitatively the results do agree and, based on the second set of measurements, there is no observed ageing in the chamber run with alcohol up to an integrated charge of 0.9 coulombs.

C.2 *In Situ* Gain Related Measurement

The muon TDCs record both the leading edge time and the width of the digital output of the muon ASD cards. The leading edge time is determined by the time at which the analog signal from the chambers crosses the discriminator threshold in the ASD, while the width measures the time-above-threshold of the analog signal. The time-above-threshold is determined by the amplitude of the signal and is therefore sensitive to the gain of the chambers. To check for any ageing effects in the chambers in B0 J/ψ events are used as a source of bona fide muons, and the average width of hits from J/ψ muons is measured as a function of run number. The results are shown in Figure C.4.

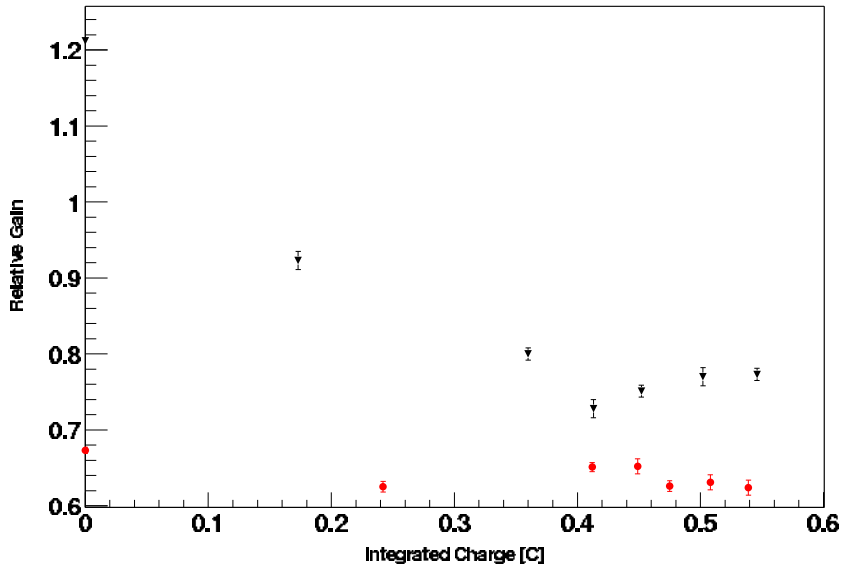


Figure C.2: Relative gain vs. integrated charge data for the first run. The triangles are for the chamber run with no alcohol and the circles are the chamber run with alcohol.

Since the chambers located on the top of the detector and at the top of the walls draw significantly more current than the chambers on the lower parts of the wall and underneath the detector, the upper and lower chambers are also analyzed separately. Figure C.5 shows the average width on J/ψ stubs as a function of run number for the upper half of the detector and for the lower half. No obvious ageing is observed so far.

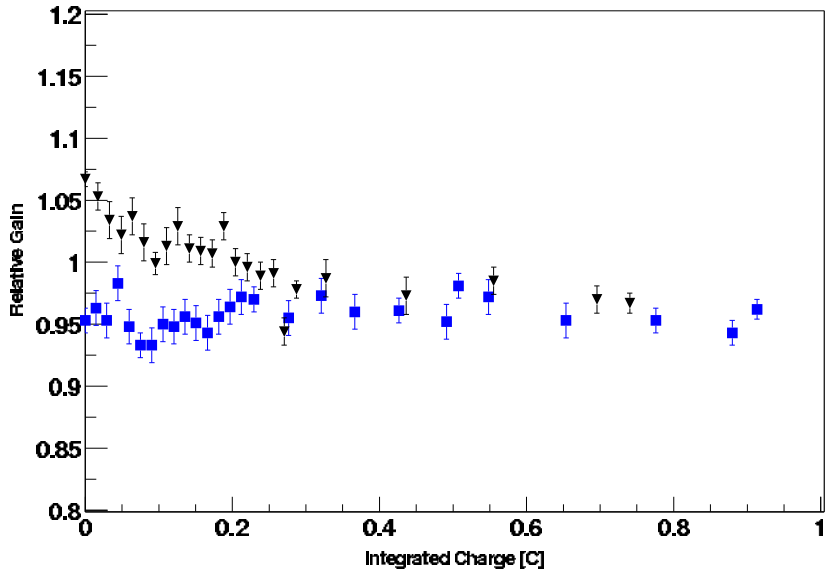


Figure C.3: Relative gain vs. integrated charge data for the second run. The triangles are for the chamber run with no alcohol and the circles are the chamber run with alcohol.

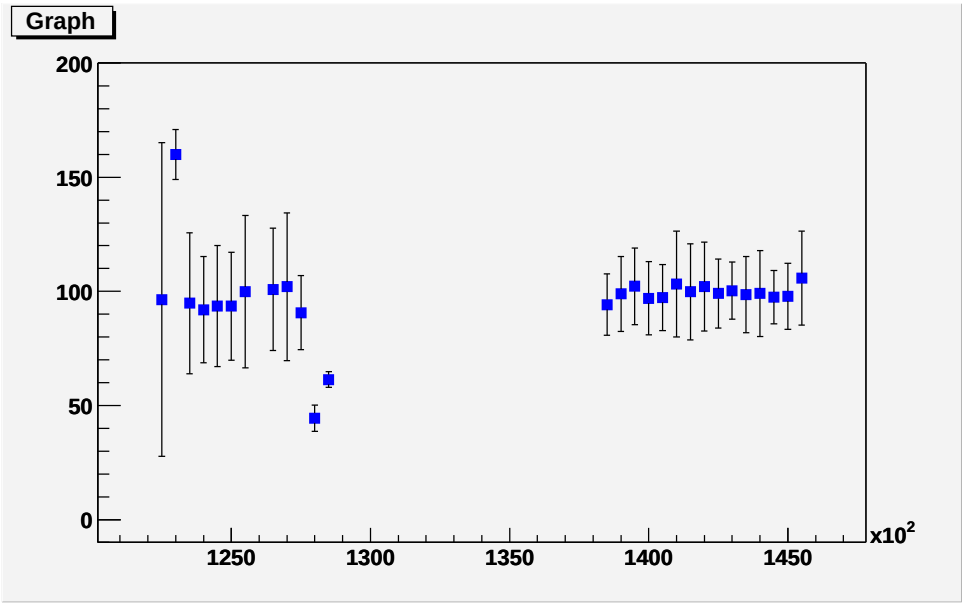


Figure C.4: Average width of CMP hits on J/ψ stubs vs. run number.

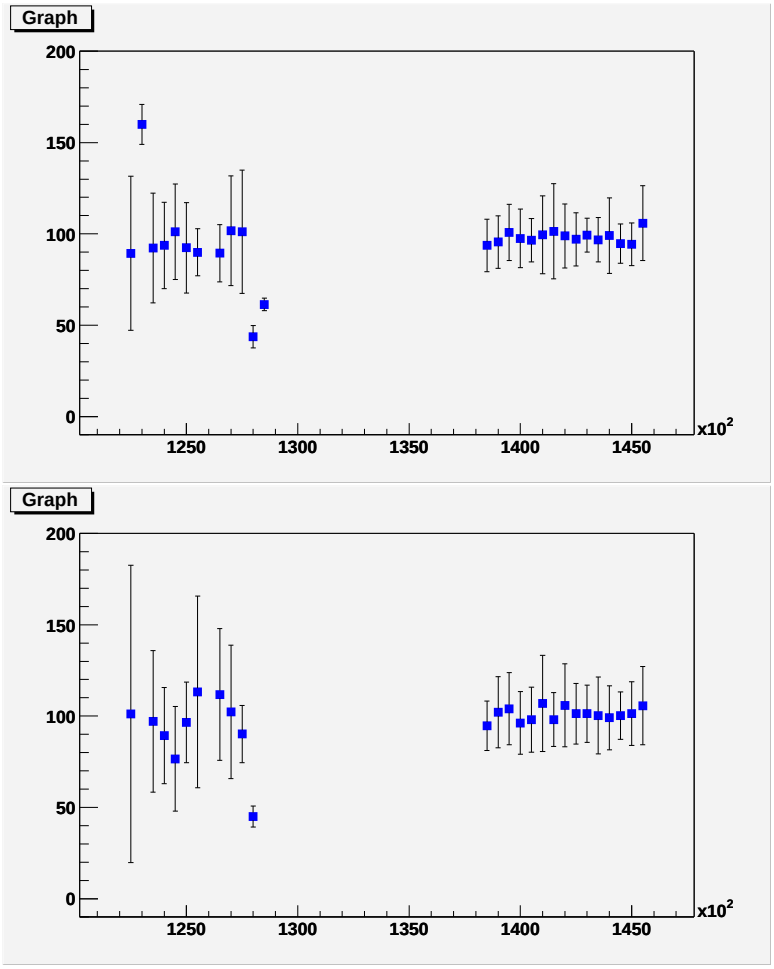


Figure C.5: Average width of CMP hits on J/ψ stubs vs. run number. The upper plot is for stubs between ϕ=0° and 180° and the lower plot is for stubs between ϕ=180° and 360°.

References

[1] There are a number of excellent texts describing the Standard Model of particle physics. Several that I have found helpful include:
F. Halzen and A. Martin, Quarks & Leptons: An Introductory Corse in Modern Particle Physics, John Wiley & Sons, Inc. (1984);
M. Peskin and D. Schroeder, An Introduction to Quantum Field Theory, Perseus Books (1995);
B. Povh *et al.*, Particles and Nuclei: An Introduction to the Physical Concepts, 2nd ed., Springer-Verlag (1999).

[2] W.-M. Yao et al. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006.

[3] The Tevatron Electroweak Working Group et al. A Combination of CDF and D0 Results on the Mass of the Top Quark. 2008.

[4] F. Abe et al. Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[5] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[6] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[7] M. Cacciari et al. The t anti-t cross-section at 1.8-TeV and 1.96-TeV: A study of the systematics due to parton densities and scale dependence. *JHEP*, 04:068, 2004.

[8] N. Kidonakis and R. Vogt. Next-to-next-to-leading order soft-gluon corrections in top quark hadroproduction. *Phys. Rev.*, D68:114014, 2003.

[9] The CDF Collaboration. Combination of cdf single top quark searches with 2.2 fb^{-1} of data. `http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf9251_stcomb2fbpub.pdf`, 2008.

[10] V. M. Abazov et al. Evidence for production of single top quarks and first direct measurement of —V(tb)—. *Phys. Rev. Lett.*, 98:181802, 2007.

[11] C. Hill and S. Parke. Top production: Sensitivity to new physics. *Phys. Rev.*, D49:4454–4462, 1994.

[12] H. P. Nilles. Supersymmetry, Supergravity and Particle Physics. *Phys. Rept.*, 110:1–162, 1984.

[13] `http://www.fnal.gov/pub/inquiring/physics/accelerators/chainaccel.html`.

[14] `http://www-bdnew.fnal.gov/operations/rookie_books/Concepts_v3.1.pdf`.

[15] `http://www.fnal.gov/pub/inquiring/physics/accelerators/00-635D.jpg`.

[16] `http://www-cdfonline.fnal.gov/ops/opshelp/stores/`.

[17] F. Abe et al. The CDF II technical design report. 1996. FERMILAB-PUB-96-390-E.

[18] A. Sill. CDF Run II silicon tracking projects. *Nucl. Instrum. Meth.*, A447:1–8, 2000.

[19] A. Affolder et al. Intermediate silicon layers detector for the CDF experiment. *Nucl. Instrum. Meth.*, A453:84–88, 2000.

[20] A. Affolder et al. CDF central outer tracker. *Nucl. Instrum. Meth.*, A526:249–299, 2004.

[21] L. Balka et al. The CDF Central Electromagnetic Calorimeter. *Nucl. Instrum. Meth.*, A267:272, 1988.

[22] M. G. Albrow et al. The CDF plug upgrade electromagnetic calorimeter: Test beam results. *Nucl. Instrum. Meth.*, A480:524–546, 2002.

[23] S. Bertolucci et al. The CDF Central and Endwall Hadron Calorimeter. *Nucl. Instrum. Meth.*, A267:301, 1988.

[24] G. Ascoli et al. CDF Central Muon Detector. *Nucl. Instrum. Meth.*, A268:33, 1988.

[25] T. Sjostrand et al. High-energy-physics event generation with PYTHIA 6.1. *Comput. Phys. Commun.*, 135:238–259, 2001.

[26] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[27] H. L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur. Phys. J.*, C12:375–392, 2000.

[28] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[29] M. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[30] F. Maltoni and T. Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 02:027, 2003.

[31] R. Brun and F. Carminati. Cern programming library long writeup w5013, 1993.

[32] G. Grindhammer, M. Rudowicz, and S. Peters. THE FAST SIMULATION OF ELECTROMAGNETIC AND HADRONIC SHOWERS. *Nucl. Instrum. Meth.*, A290:469, 1990.

[33] E. Gerchtein and M. Paulini. CDF detector simulation framework and performance. 2003.

[34] D. Acosta et al.

[35] The CDF Collaboration. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ tev using lepton plus jets events with soft muon b-tagging. `http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf8565_SLTxs_public.ps`, 2006.

[36] The soft muon tagging algorithm employed for this analysis was developed primarily by two former Illinois postdocs, Lucio Cerrito and Anyes Taffard.

[37] A. Abulencia et al. Measurements of Inclusive W and Z Cross Sections in p-pbar Collisions at sqrts =1.96 TeV. *J. Phys.*, G34:2457–2544, 2007.

[38] A. Bhatti et al. Determination of the jet energy scale at the Collider Detector at Fermilab. *Nucl. Instrum. Meth.*, A566:375–412, 2006.

[39] S. Klimenko, J. Konigsberg, and T. M. Liss. Averaging of the inelastic cross sections measured by the CDF and the E811 experiments. FERMILAB-FN-0741.

[40] D. E. Acosta et al. Measurements of $b\bar{b}$ azimuthal production correlations in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev.*, D71:092001, 2005.

[41] K. P. Lannon. A measurement of B hadron correlations in p anti-p collisions at s**(1/2) = 1.8-TeV. UMI-31-01893.

[42] C. Chen. A measurement of the direct charm meson production cross section at CDF II. UMI-30-87379.

[43] M. Campanelli et al. Time-of-flight studies with λ and k_s^0 decays. CDF Internal Note 6757, 2004. unpublished.

[44] `http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf8795_SecVtxXSPublic.ps`. The K factor and ALPGEN parameter uncertainties used in the heavy-flavor background estimate were determined as part of a separate analysis by M. Franklin, S. Grinstein, J. Guimarães da Costa, K. Lannon, T. Schwarz, D. Sherman, A. Taffard, and I. Zaw.

[45] J. M. Campbell and R. K. Ellis. An update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999. The Tevatron has a slightly different center-of-mass energy than assumed by Campbell and Ellis. Further, the ZZ cross section I quote includes off-shell Z bosons. Therefore, the cross sections I quote are not identical to those calculated by Campbell and Ellis.

[46] D. Acosta et al. First measurements of inclusive W and Z cross sections from Run II of the Tevatron collider. *Phys. Rev. Lett.*, 94:091803, 2005. The measured Z cross section is for a Z mass between 66 GeV/c^2 and 116 GeV/c^2. I quote a larger cross section that is extrapolated to all dilepton masses greater than 30 GeV/c^2.

[47] Z. Sullivan. Understanding single-top-quark production and jets at hadron colliders. *Phys. Rev.*, D70:114012, 2004.

[48] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Parton distributions: A new global analysis. *Eur. Phys. J.*, C4:463–496, 1998.

Author's Biography

Ulysses Allen Grundler was fascinated with science from an early age, his interest in physics in particular began with his high school physics classes. During his final year as an undergraduate physics major, participation in several research projects with the CDF collaboration convinced him that he wanted to study particle physics. He received a B.S. in Physics from the University of Illinois at Urbana-Champaign in 2002 and decided to remain in Urbana for graduate study. After studying top quark physics, he graduated with a Ph.D. in Physics in 2008. He accepted a position as a postdoctoral research position with Iowa State University to work on the ATLAS experiment at CERN in Geneva, Switzerland.