

MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION
CROSS-SECTION IN THE DILEPTON CHANNEL USING LEPTON PLUS
TRACK SELECTION AND IDENTIFIED B-JETS

by

Teresa Spreitzer

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Copyright © 2008 by Teresa Spreitzer

Abstract

Measurement of the Top Quark Pair Production Cross-Section in the Dilepton Channel
using Lepton plus Track Selection and Identified b-jets

Teresa Spreitzer

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2008

Using 1.0 fb^{-1} of data collected by the Collider Detector at Fermilab (CDF) from Run II of the Fermilab Tevatron, we measure the $t\bar{t}$ production cross-section in events with two leptons, significant missing transverse energy, and ≥ 2 jets, at least one of which is identified as a b -jet. As the Run II dataset grows, more stringent tests of Standard Model predictions for the top quark sector are becoming possible. The dilepton channel, where both top quarks decay $t \rightarrow Wb \rightarrow \ell\nu b$, is of particular interest due to its high purity. Use of an isolated track as the second lepton significantly increases the dilepton acceptance, at the price of some increase in background, particularly from $W + \text{jets}$ events where one of the jets is identified as a lepton. To control the increase in background we add to the event selection the requirement that at least one of the jets be identified as a b -jet, reducing the background contribution from all sources. Assuming a branching ratio of $BR(W \rightarrow \ell\nu) = 10.8 \%$ and a top mass of $m_t = 175 \text{ GeV}/c^2$, the measured cross-section is $\sigma(p\bar{p} \rightarrow t\bar{t}) = 10.5 \pm 1.8(\text{stat.}) \pm 0.8(\text{syst.}) \pm 0.6(\text{lumi.}) \text{ pb}$.

Dedication

To my children,
thank you for the sacrifices you have made so early in your lives.

Acknowledgements

It has been said before that any work in high energy physics is a collaborative one. I first wish to sincerely thank the entire CDF family for all the countless hours of labour that went into producing a successful experiment.

I must thank my supervisor Pierre Savard. I don't think he understands how much he helped me to build confidence, in addition to teaching me so much about thinking critically. He was always motivated me by his example. I must especially thank him for his unwavering support while I was starting my family. I hope he knows how much I appreciate him.

I also wish to thank the other professors at the University of Toronto working on CDF, Pekka Sinervo and William Trischuk for their years of support and thoughtful suggestions.

Professors Bob Holdom and David Bailey have spent a good amount of time in helping me get to this point, as members of my Ph. D. Committee. I have always found our meetings helpful and informative. A big thank you goes to both of them.

I wish to thank present and past members of the lepton plus track analysis group. Joe Incandela, Peter Wittich, Petra Merkel and Julia Thom. I've learned a lot more from listening to them than I think they realize. I appreciate their guidance.

I would be remiss if I didn't thank some other members of the CDF group for their help, Kevin Lannon, Anyes Taffard, Beate Heinemann and Robin Erbacher. As conveners these people have helped me mature as a scientist, and supported me along the way.

When I first arrived at the University of Toronto I was very warmly welcomed. I can't thank everyone enough for all the time you took in teaching me so much that first year, and for making me feel like I belonged. Thank you Jean-Francois Arguin, Bernd Stelzer, Oliver Stelzer-Chilton, Simon Sabik, Kostas Kordas, Shabnaz Pashapour and Dan MacQueen. Stan Lai was about the best office mate one could have. To the Atlas newcomers, thanks for helping make finishing less painful, Mini-Pier, Saminder, Jaspreet,

Bin, Behi, Peter, Sing, and Dominique, I'll always remember you letting me talk for so long. Thanks Leslie Groer for all your computing help.

To Corrinne Mills....what can I say? I would never be where I am without having sheltered under your wing for so long. I have never enjoyed working with someone so much, or had physics bring me such a good friend.

To my loving husband Darrel, I can't thank you enough for understanding how important timely completion of my thesis was to me, and for putting your career on hold for a time to let me do this. Thanks for traveling with me across this big country, as I'm still not tired of physics. Thank for you for being such a good husband, and loving father. I didn't think that I could love you more each day, but I do. I can't wait for our family to continue to grow together, as I realize now the best really is yet to come.

Contents

1	Introduction	1
1.1	The Standard Model	2
1.2	The Top Quark	4
2	Experimental Apparatus	13
2.1	The Tevatron	13
2.2	The CDF Detector	16
2.2.1	The Tracking System	17
2.2.2	The Calorimeter System	20
2.2.3	The Muon System	24
2.2.4	Cherenkov Luminosity Counter	25
2.3	Event Triggers and Data Acquisition	26
2.3.1	The CDF Trigger System	27
2.3.2	Triggers Used in this Analysis	28
3	Pretag Event Selection	31
3.1	Lepton Definition	32
3.1.1	Electron Selection	34
3.1.2	Muon Selection	37
3.1.3	Track Selection	40
3.2	Jet Definition	41

3.2.1	Jet Clustering	41
3.2.2	Jet Corrections	42
3.2.3	Jets for Analysis	42
3.3	Missing Transverse Energy Reconstruction	43
3.3.1	Track Corrections	43
3.3.2	Jet Corrections	43
3.4	Event Selection	44
3.4.1	Basic Event Selection	44
3.4.2	$\Delta\phi$ Cuts	46
3.4.3	Drell-Yan Veto	48
3.4.4	Event Classification	48
3.5	Data Sample	48
4	Pretag Top Dilepton Acceptance	50
4.1	Simulated $t\bar{t}$ Sample	50
4.2	Corrections to Simulation	51
4.2.1	Primary Lepton Identification Scale Factor	51
4.2.2	Track Lepton Scale Factors	52
4.2.3	χ^2 probability Cut efficiency	53
4.2.4	Event Trigger Efficiencies	54
4.3	Acceptance for Pretag Top Dilepton Events	55
4.4	Systematic Uncertainties on the Acceptance	56
4.4.1	Primary Lepton Identification Efficiency	56
4.4.2	Track Lepton Identification Efficiency	57
4.4.3	Jet energy scale	58
4.4.4	Initial and Final State Radiation (ISR/FSR) modeling in simulation	58
4.4.5	Parton Distribution Functions	58
4.4.6	Monte Carlo Generator	59

5	<i>b</i>-Jet Identification	61
5.1	SECVTX Algorithm	62
5.2	Event Tagging Efficiency	64
5.2.1	Simulation Method	65
5.2.2	Factorization Method	65
5.2.3	Final Event Tagging efficiency	67
6	Pretag Backgrounds	69
6.1	Diboson and $Z/\gamma^* \rightarrow \tau\tau$ Processes	69
6.1.1	Systematic Uncertainties	71
6.2	$Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$	73
6.2.1	Normalizations from Data, $n_{\{25,40\}}$	74
6.2.2	Non- $Z/\gamma^* \rightarrow ee/\mu\mu$ contributions, $n_{\{25,40\}}^{feeddown}$	75
6.2.3	Ratios from Simulation, R_i	76
6.2.4	Corrections to Simulation, C_{Zmass}	76
6.2.5	Final Drell-Yan Background Estimate	77
6.2.6	Systematic Uncertainties	77
6.3	W + jets with a Fake Lepton	78
6.3.1	The Fake Rate	79
6.3.2	Fake Estimate Normalization	82
6.3.3	Contribution from $t\bar{t}$ Events with a Fake Lepton	85
6.3.4	Primary Lepton Fake Rates	85
6.3.5	Opposite-Charge Fraction	86
6.4	Summary	89
7	Tagged Background Estimate	91
7.1	Estimating Backgrounds with the Tag Rate Matrices	92
7.2	Tagged Background	93

7.2.1	Validation of events with low \cancel{E}_T	94
7.2.2	Validation of events with significant \cancel{E}_T	97
7.3	Physics and Mistag Backgrounds	100
7.3.1	Top dilepton signal component of the tag rate matrix estimate . .	101
7.4	Fake Lepton Background	101
7.5	Final Background Estimate	103
7.6	Systematic Uncertainties on the Background	103
8	Results	106
8.1	Tagged Cross-section Result	106
8.2	Pretag Cross-section Measurement	107
8.3	Cross-Checks	108
8.3.1	Double Tag Cross-section	108
8.3.2	Higher purity b tagging algorithm	110
8.4	Kinematic Features	110
9	Conclusions	118
	Bibliography	121

List of Tables

1.1	Fundamental fermions (quarks and leptons) of the Standard Model [6].	3
1.2	Fundamental bosons of the Standard Model [6].	4
1.3	Branching ratios of the $t\bar{t}$ experimental signatures.	6
3.1	Fraction (in %) of the (oppositely charged, ≥ 2 jet) acceptance that is attributable to each possible generated lepton pair. The two middle rows show the events divided according to whether they were reconstructed with an electron or muon as the primary lepton. The majority of events are accepted as electron plus track because there is more geometric acceptance for electrons, and because of the way in which events are categorized when both leptons can be fully reconstructed.	34
3.2	Integrated luminosity for each primary lepton type for both the tagged and pretag analysis. The uncertainty on the luminosity is 6% due to the limits of our understanding of the luminosity detectors and the $p\bar{p}$ inelastic scattering.	49
4.1	Efficiencies for identifying the different primary lepton types in data and simulation, and the ratio is used to correct the efficiency in $t\bar{t}$ simulation.	52
4.2	Data/simulation scale factors applied to the $t\bar{t}$ simulation to correct the pretag efficiency for the isolation efficiency cut on track leptons.	53

4.3	Data/simulation scale factors applied to the $t\bar{t}$ simulation to correct for data/simulation discrepancies.	54
4.4	Single lepton and full per-event trigger efficiencies for the four primary lepton types. Numbers in bold are inputs to the cross-section measurement. Uncertainties are negligible compared to other uncertainties on the acceptance.	55
4.5	Top dilepton signal acceptance, in percent, for $m_t = 175\text{GeV}/c^2$, measured in PYTHIA , corrected for all known discrepancies between data simulation. Uncertainty is statistical only.	56
4.6	Summary table of systematic uncertainties on the signal acceptance. The total systematic uncertainty is 3%.	57
5.1	Inputs and results of calculation of event tagging efficiency using the factorization method.	67
6.1	Fraction of events with a given jet multiplicity and scale factor correcting simulation to the level of data.	70
6.2	Counts from the data used to normalize the Drell-Yan estimate, and the expected contribution from other processes (“feeddown”).	76
6.3	Inputs to the Drell-Yan estimate from simulated $Z/\gamma^* \rightarrow ee/\mu\mu$ events. .	77
6.4	Predicted and observed number of fake leptons in the data and simulation cross-checks. The number of jets quoted is in addition to the jet which is misidentified as a track lepton. The only statistically significant discrepancy observed is in the 1 jet category in the simulation cross-check (top table). This discrepancy forms part of the basis of the systematic uncertainty on this background estimate.	82
6.5	Predicted background events in the pretag 1.1 fb^{-1} dataset, Systematic uncertainties are included.	90

7.1	Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of jet multiplicity. Statistical and systematic uncertainties predictions have been combined.	95
7.2	Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of Jet E_T . Uncertainties on the total predictions (bottom line) are statistical and systematic combined.	97
7.3	Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of track multiplicity in jets. Uncertainties on the predictions are statistical and systematic combined.	98
7.4	Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of Jet η . Uncertainties in the predictions are statistical and systematic combined.	99
7.5	Comparison of the observed and predicted tags with 1 jet data and simulated events. Errors on the predictions are statistical and systematic combined.	100
7.6	Tagged background calculation inputs.	104
7.7	Summary table of systematic uncertainties on the tagged background estimate.	105
8.1	The pretag and tagged cross-section as calculated at several input top masses. Theoretical prediction from Ref. [15]. The first uncertainty is statistical, the second is systematic. A common uncertainty of 6%, due to the uncertainty on the integrated luminosity, is omitted.	107
8.2	Predicted and observed events in the pretag 1.1 fb^{-1} dataset, with details of the background contributions. Systematic uncertainties are included in the prediction numbers.	108

List of Figures

1.1	Main $t\bar{t}$ production and decay process at the Tevatron	6
1.2	Missing transverse energy of $t\bar{t}$ dilepton candidates at CDF in Run I, compared to the expectation for combined signal and background. Taken from Ref. [20].	10
1.3	Recent results of the top quark cross-section measurements from CDF [24].	12
2.1	Schematic drawing of the Tevatron accelerator chain.	14
2.2	A pictorial representation of the Collider Detector at Fermilab.	17
2.3	Definition of CDF coordinate system. The arrows for p and \bar{p} indicate direction of travel of the protons and anti-protons around the Tevatron ring.	17
2.4	A schematic view of one quadrant of the CDF tracking system.	18
2.5	End view of the CDF silicon tracking system.	19
2.6	Diagram of a CEM wedge.	23
2.7	Four layers of drift chambers in the CMU.	25
2.8	Schematic representation of the data flow in the CDF trigger and data acquisition systems.	30
3.1	Kinematically allowed decay modes of the tau lepton. The τ is more massive than the estimated mass of the c quark, but no charmed meson is lighter than the τ , so $\bar{c}s$ and $\bar{c}d$ decays are not allowed.	32

3.2	Schematic diagram of the experimental signatures of different objects as seen by the CDF detector. Note how muons and electrons both leave isolated tracks in the tracker, while jets are identified by clusters of several tracks.	33
3.3	Drawing of the mechanism by which the decay in flight of a medium- to low- p_T kaon can be reconstructed as a high- p_T muon.	39
3.4	Missing Transverse Energy (\cancel{E}_T) in simulated $Z/\gamma^* \rightarrow \ell\ell$ events and dilepton top decays.	45
3.5	The angle between the \cancel{E}_T and the jet in simulated $t\bar{t}$ events and $Z/\gamma^* \rightarrow \ell\ell$ events.	47
5.1	Schematic diagram of a displaced vertex.	63
5.2	Drawing showing fake ($L_{xy} < 0$) and true ($L_{xy} > 0$) reconstructed secondary vertices.	64
5.3	The tagging efficiency for the SECVTX tagger for the two SECVTX tunings, as a function of the jet E_T and η . These are based on simulated datasets, but corrected for discrepancies between the simulation and data. The blue (lower efficiency) line is for the higher purity, lower efficiency version of the tagging algorithm. The red line is for the lower purity, higher efficiency version of the tagging algorithm, which is used in this analysis.	68
6.1	Diboson decays.	72
6.2	Two processes contributing to “fake” events. The left is the dominant W + jets production process in this sample, and the right is $t\bar{t}$ decay in the lepton + jets channel.	80

6.3	Predicted and observed number of isolated tracks in simulated $W + \text{jets}$ events, as a function of the number of jets in the event. The fake rate used is derived from photon+jets and (for events reconstructed with zero jets) $Z+\text{jets}$ simulation. The over-prediction for zero jet events is incorporated in the systematic uncertainty on the fake lepton background.	83
6.4	Predicted and observed number of isolated tracks in $Z + \text{jets}$ data, as a function of the number of jets in the event. The fake rate used is derived from photon+jets data.	84
6.5	Testing the applicability of the 80 GeV photon + jets fake rate to fake leptons from top in the lepton + jets mode, in simulation. No fake leptons are observed from events with one jet.	86
7.1	Kinematic distributions for pretagged $Z+\text{jets}$ candidate events for data and simulation. Note that the uncertainties on the data are statistical only.	95
7.2	Kinematic distributions for observed tags in data, and predicted tags in data and simulation. Note that the uncertainties on the observations in data are statistical only, and the predictions in data and simulation are subject to a 5% systematic uncertainty from the positive tag rate matrix.	96
8.1	Number of predicted events compared to the number observed in the pre-tag candidate data. The cross-hatched areas show the (1σ) uncertainties on the predicted numbers (statistical and systematic).	109
8.2	Summed scalar energy of the leptons, \cancel{E}_T and jets in the tagged candidate sample. The signal region is scaled to the measured cross-section.	111
8.3	Missing transverse energy of the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	112

8.4	Invariant mass of the lepton pair in tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	113
8.5	Primary lepton E_T in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	114
8.6	Track lepton p_T in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	115
8.7	The E_T of the tagged jets in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	116
8.8	The track multiplicity of the tagged jets in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.	117

Chapter 1

Introduction

In the last 50 years or so, particle physics has undergone a dramatic transformation. Before the bubble chamber experiments of the mid-century, scientists had experimentally observed only a few particles. With the bubble chamber experiments in the 1950's and 1960's came the discovery of many new particles within a short decade or so. This became known as a golden era for particle physics, as the results of experiments were shaping our evolving theories of the universe. The new observations were eventually understood in terms of a coherent theory. After the theory of quarks and leptons came the electroweak theory. Strong forces describing the quarks were soon included and together this description was called the Standard Model (SM) [1]. All presently observed phenomena (with the exception of the observed neutrino mixing [2]), are described by the SM. With the Large Hadron Collider at CERN [3] set to collide protons with a centre of mass energy of 14 TeV, we eagerly await either the discovery of, or perhaps fruitless search for, the Higgs boson, the last particle predicted by the SM. Unlike the last golden era of discovery, currently several theories predict what will be seen beyond our current knowledge frontier. The next generation of particle physics experiments are expected to tell us much about the universe.

In the meantime much important work has been, and continues to be, done. The

Tevatron at Fermilab is currently the world's highest-energy collider. In 1995 both Tevatron detectors, the Collider Detector at Fermilab (CDF) and DØ announced the discovery of the last quark predicted by the SM, the top quark [4], [5]. The top quark's large mass gives us clues to what we might find at the LHC, such as the mass of the Higgs boson, and serves as a good testing ground for learning how to do physics at the LHC.

1.1 The Standard Model

The Standard Model (SM) of particle physics describes the universe in terms of fundamental particles and their interactions. The SM encompasses spin-1/2 fermions, of which there are six quarks and six leptons. The fermions are organized into three generations, and make up all observable matter. For each fermion there is a corresponding antiparticle, similar to its matter partner, with the same mass and opposite charge. For reasons which are unknown, the universe we live in is composed almost entirely of matter, but at colliders matter and antimatter are produced, and generally treated, equally. In the SM there are also four spin-1 gauge bosons, which mediate the interactions between particles. These particles are all listed in Tables 1.1, 1.2. Photons mediate the electromagnetic force, W^\pm and Z bosons mediate the weak nuclear force, and gluons mediate the strong nuclear force. Gravity is not described by the SM, but since gravity is so weak its effect is negligible in high energy particle interactions.

Quantum Electrodynamics (QED) is based on the $U(1)$ symmetry group and describes the electromagnetic interaction among quarks and leptons through the exchange of massless spin-1 photons [7].

The W , Z , and photon bosons mediate the electroweak interaction, a $SU(2)_L \times U(1)$ symmetry group. The W and Z bosons are very massive, and thus can only act over short distances. One of the greatest achievements of the Standard Model is the unification of both the electromagnetic and weak interactions into electroweak theory [1]. Unification

Fundamental Fermions (Spin-1/2)					
Quarks			Leptons		
Flavour	Charge	Mass [MeV/c ²]	Flavour	Charge	Mass [MeV/c ²]
Up (<i>u</i>)	+2/3	1.5 to 3.0	Electron neutrino (ν_e)	0	$< 1.5 \times 10^{-6}$
Down (<i>d</i>)	-1/3	3 to 7	Electron (e^-)	-1	0.511
Charm (<i>c</i>)	+2/3	$(1.0 \text{ to } 1.4) \times 10^3$	Muon neutrino (ν_μ)	0	$< 1.9 \times 10^{-4}$
Strange (<i>s</i>)	-1/3	70 to 120	Muon (μ^-)	-1	105.7
Top (<i>t</i>)	+2/3	$(174.2 \pm 3.3) \times 10^3$	Tau neutrino (ν_τ)	0	< 18.2
Bottom (<i>b</i>)	-1/3	$(4.2 \text{ to } 4.7) \times 10^3$	Tau (τ^-)	-1	1777.1

Table 1.1: Fundamental fermions (quarks and leptons) of the Standard Model [6].

is achieved by introducing a scalar field, known as the Higgs field, whose presence breaks the symmetry of the $SU(2)_L \times U(1)$ electroweak symmetry. This mechanism also predicts the existence of a neutral scalar particle, known as the Higgs boson [8]. The Higgs boson is the only particle predicted by the SM which has yet to be observed. Predictions based on electroweak theory have been verified by experiment to very high precision [9].

The strong interaction among quarks is mediated by the exchange of eight massless spin-1 gluons [10]. This interaction is based on the $SU(3)$ symmetry group, called Quantum Chromodynamics (QCD). Quarks and gluons carry colour charges (red, green and blue) which are analogous to the electric charge in QED. The eight gluons exist in the following colour combinations:

$$R\bar{B}, R\bar{G}, B\bar{R}, B\bar{G}, G\bar{R}, G\bar{B}, (R\bar{R} - G\bar{G})/\sqrt{2}, (R\bar{R} + G\bar{G} - 2B\bar{B})/\sqrt{6}$$

The strong interaction binds quarks together into colourless states. Mesons are $q\bar{q}$ bound states, and qqq bound states are baryons.

Differences between QED and QCD are due to the fact that gluons carry colour charge and photons are electrically neutral. This gives rise to gluon self-coupling. At

Fundamental Bosons (Spin-1)			
Interaction	Name	Charge	Mass [GeV/c ²]
Electromagnetic	Photon (γ)	0	0
Weak	W boson (W)	1	80.403 ± 0.029
Weak	Z boson (Z)	0	91.1876 ± 0.0021
Strong	Gluon (g)	0	0

Table 1.2: Fundamental bosons of the Standard Model [6].

short distances or large momentum transfers, the strong force gets weaker. This property is known as asymptotic freedom [11]. On the other hand, at larger distances the strong force increases in strength. So, the potential energy of two coloured particles increases with the distance between them. This property results in confinement, and explains the creation of “jets” in high energy collisions. When a large amount of energy is transferred to a quark, such as a quark inside a proton at the Tevatron, the quark moves away from the other quarks in the proton to which it is bound, and the colour field grows until enough energy is present to create a particle-antiparticle pair out of the vacuum. This process repeats itself until the original energy is dissipated. The process is called hadronization because several hadrons are created, all of which are traveling in the same general direction as the original particle [12]. The top quark is an exception to this hadronization process, as it almost always decays before hadronization [13].

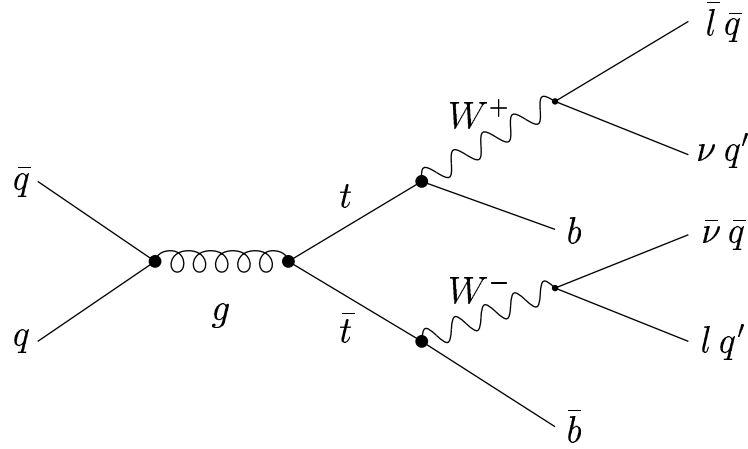
1.2 The Top Quark

The top quark is unique among the quarks. Its most striking feature is its large mass; the current best measurement of the top quark mass is $170.9 \pm 1.8 \text{ GeV}/c^2$ [14]. This is ≈ 35 times heavier than the next heaviest quark, the b quark. Due to its large mass, it has a lifetime of $\approx 10^{-25}$ s. Also due to its large mass the top quark decays to a real

W boson. Both of these facts make reconstruction of top event kinematics easier. Also because of its large mass, the top quark has the strongest coupling to the Higgs field. The most recent theoretical calculations of the top quark cross-section have uncertainties of less than 15% [15, 16]. Since the original discovery of the top quark during Run I of the Tevatron [4, 5], more luminosity has been delivered to the detectors CDF and DØ, and it is now possible to study the top quark sector in detail.

At the Fermilab Tevatron, top quark pairs are created through quark anti-quark annihilation about 85% of the time, and through gluon fusion about 15% of the time. In the Standard Model the top quark decays as $t \rightarrow Wb$ almost 100% of the time [6]. The W boson can decay into a charged lepton-neutrino pair, or as a quark anti-quark pair. Top quark pair decays are classified by the decay of the W bosons. If both W bosons decay hadronically it is referred to as the “all hadronic” channel. In the case where one W boson decays to leptons and the other to quarks we call it the “lepton + jets” channel. If both W bosons decay leptonically, as in the events selected in this thesis, we call the decay a “dilepton channel” decay. The signature of the dilepton channel is two charged leptons (except in the case where one of the leptons is a τ lepton, where we observe the τ decay products), two jets from the hadronization of the b and \bar{b} , and missing transverse energy (energy imbalance from the two undetected neutrinos, see Section 3.3). In addition to the expected two jets from b and \bar{b} decay, gluons can be radiated in the initial or final state, producing extra jets. The relative fraction of these decay channels is outlined in Table 1.3.

The goal of particle physics is to determine the fundamental constituents of the universe and the nature of their interactions. This thesis forms part of the effort to understand the role of the top quark in fundamental physics. It is important to study the top quark because in addition to the role it plays in the SM, it is also expected to give us clues to physics beyond the SM. Because the top quark is so heavy, the possibility exists that the top quark decays to yet undetected other particles. Significant deviation of the

Figure 1.1: Main $t\bar{t}$ production and decay process at the Tevatron

Channel	Branching ratio
All-hadronic	46%
Lepton+jets	29%
Dilepton	5%
Tau lepton+anything	20%

Table 1.3: Branching ratios of the $t\bar{t}$ experimental signatures.

measured cross-section from the predicted value could indicate a non-standard process such as the decay into supersymmetric particles [18, 19]. Even in the absence of clues for new physics, a solid understanding of the $t\bar{t}$ process is crucial for searches for physics beyond the Standard Model at future colliders, such as the LHC. Some supersymmetry signatures are predicted to have multilepton, multijet final states, and if so, the $t\bar{t}$ signal will be a background source. More generally, it is important to understand the mechanisms that generate missing energy in searches for new physics. In these respects the top quark cross-section becomes a testing ground for future searches for new physics.

Top Quark Cross-section Measurement

The top quark decay in the dilepton channel experimental signature contains two b quarks. The b hadrons are relatively long lived, and the jets produced by b quark decay

can be identified by the displacement of some tracks from the primary interaction. These displaced tracks form a secondary vertex, which is identified as a b jet. Because the process of identifying b jets is about 50% efficient, and the dilepton channel branching ratio is small, sufficient integrated luminosity has not been available to use b jet identification in a useful way to measure the top quark cross-section in the dilepton channel until now. The sample of top quarks in the dilepton channel with an identified b -jet is expected to be the most pure and may shed light on certain new physics scenarios.

The production cross-section for a process is a fundamental quantity, and is a measure of the expected rate of that process. In theory, it is easy to measure, basically amounting to counting the number of observed events passing the selection and estimating the contribution from background events. In practice, estimating the contribution from background processes complicates matters. The cross-section as measured in this thesis is given by

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{\epsilon_{tag} \times \mathcal{A} \times \int \mathcal{L} dt} \quad (1.1)$$

where N_{obs} is the number of observed events passing the selection. N_{bkg} is the number of background events, defined as the number of events not from the process of interest which pass the selection. \mathcal{A} is the acceptance, the fraction of events of interest which pass the selection. $\int \mathcal{L} dt$ is the integrated luminosity, the total amount of data examined. ϵ_{tag} is the event tagging efficiency, or the fraction of events passing the selection which have at least one jet which is identified as having come from heavy flavour decay.

There is no best way to measure the top quark cross-section. The all-hadronic channel poses difficulties with large QCD backgrounds, although this channel has the largest branching ratio, and thus the largest dataset with which to work. The lepton+jets channel has the second largest branching ratio, but it also has smaller background contributions, and those backgrounds are easier to quantify. The dilepton channel has the smallest branching ratio and as such has limited statistical power. But the dilepton experimental signature has the fewest backgrounds, it is the most pure sample. To combat

low event counts, one can use less stringent selection, increasing the size of the selected dataset. Instead of fully reconstructing two charged leptons, if an analysis accepts events with one fully reconstructed charged lepton and another partially reconstructed charged lepton, the event selection efficiency is increased, at the expense of sample purity, but with an overall benefit to the significance of the result [22]. This selection is called the lepton plus isolated track selection, and is used in the analysis presented in this thesis.

As stated in the previous section, study of the top quark sector is an important check on the predictions of the SM, as a tool to search for new physics, and as a testing ground for future work at new colliders. In addition to these considerations it is important to measure the top quark cross-section in all of the decay channels, because the SM predicts that the cross-section should be the same in each. Different cross-sections in different decay channels would indicate that the decay of the top quark is not as predicted by the SM. Measurement of the cross-section in all decay channels therefore tests the theoretical assumptions of how the top quark decays. In this regard it is also important to measure the top quark cross-section with and without the presence of identified b jets. For example in the dilepton channel if a significantly different cross-section were observed when requiring b identified jets, and making no such requirement on the jets, this would indicate that, possibly the decay of the top quark is not $t \rightarrow Wb$ almost 100% as is predicted by the SM. This thesis describes the first measurement using the CDF detector of the top quark cross section in the dilepton channel with identified b jets. As such, this work is an important step in the understanding of the top quark sector, and as a tool to look for new physics.

Previous Measurements

In Run I of the Tevatron the final CDF measurement of the top quark cross-section in the dilepton channel used just over 100 pb^{-1} of data. The dilepton cross section was $8.2^{+4.4}_{-3.4} \text{ pb}$ for a top mass of $175 \text{ GeV}/c^2$. That was not significantly higher than the theoretical

prediction of the cross-section for that top mass and $\sqrt{s} = 1.8$ TeV, $5.2^{+0.5}_{-0.7}$ pb [20]. The interesting part of the result was not in the measured cross-section, but in the kinematic distributions. Seven of the nine observed events were in the $e\mu$ final state, and those events also had larger than expected \cancel{E}_T (see Figure 1.2). DØ also measured the dilepton cross-section with about the same integrated luminosity and got a cross-section closer to the theoretical expectation ($5.69 \pm 1.21 \pm 1.04$ pb). But the DØ result in the $e\mu$ channel, as in the CDF result, was higher than in the ee and $\mu\mu$ channels [21].

These results pointed to the possibility that the top quark measured at the Tevatron wasn't exactly as the SM predicted. Since supersymmetry signatures are expected to look much like top decay, except with more \cancel{E}_T , these results were exciting. With the first 200 pb⁻¹ of Run II data, CDF measured the top dilepton cross-section using two selection methods. The first used two fully reconstructed leptons, the other used one fully reconstructed lepton, and one lepton only partially reconstructed. The combined result for $m_{top} = 175$ GeV/ c^2 was $7.0^{+2.4}_{-2.1}(\text{stat.})^{+1.6}_{-1.1}(\text{sys.})$ pb. DØ also used events with two fully reconstructed leptons and measured $8.6^{+3.2}_{-2.7}(\text{stat.}) \pm 1.1(\text{sys.})$ pb for a top mass of 175 GeV/ c^2 . Both of these results are in good agreement with the SM prediction of $6.7^{+0.7}_{-0.9}$ pb for a top quark mass of 175 GeV/ c^2 at $\sqrt{s} = 1.96$ TeV [15]. While these measurements represented a significant improvement in precision over the Run I results, no clear statement about content of the top dilepton decay signature could be made. Since these measurements, the DØ collaboration has measured the top quark cross-section in the dilepton channel again with 425 pb⁻¹ of Run II data [23]. Assuming a top quark mass of 175 GeV/ c^2 they find the top quark cross-section to be $7.4 \pm 1.4(\text{stat.}) \pm 1.0(\text{syst.})$ pb which is not significantly different from the theoretical prediction.

Both CDF and DØ have measured the top quark cross-section in all three decay channels, and have observed no significant deviation among the measurements, or from theoretical expectations. This consistency across decay channels indicates that within our current experimental precision, we observe that the top quark decays as predicted

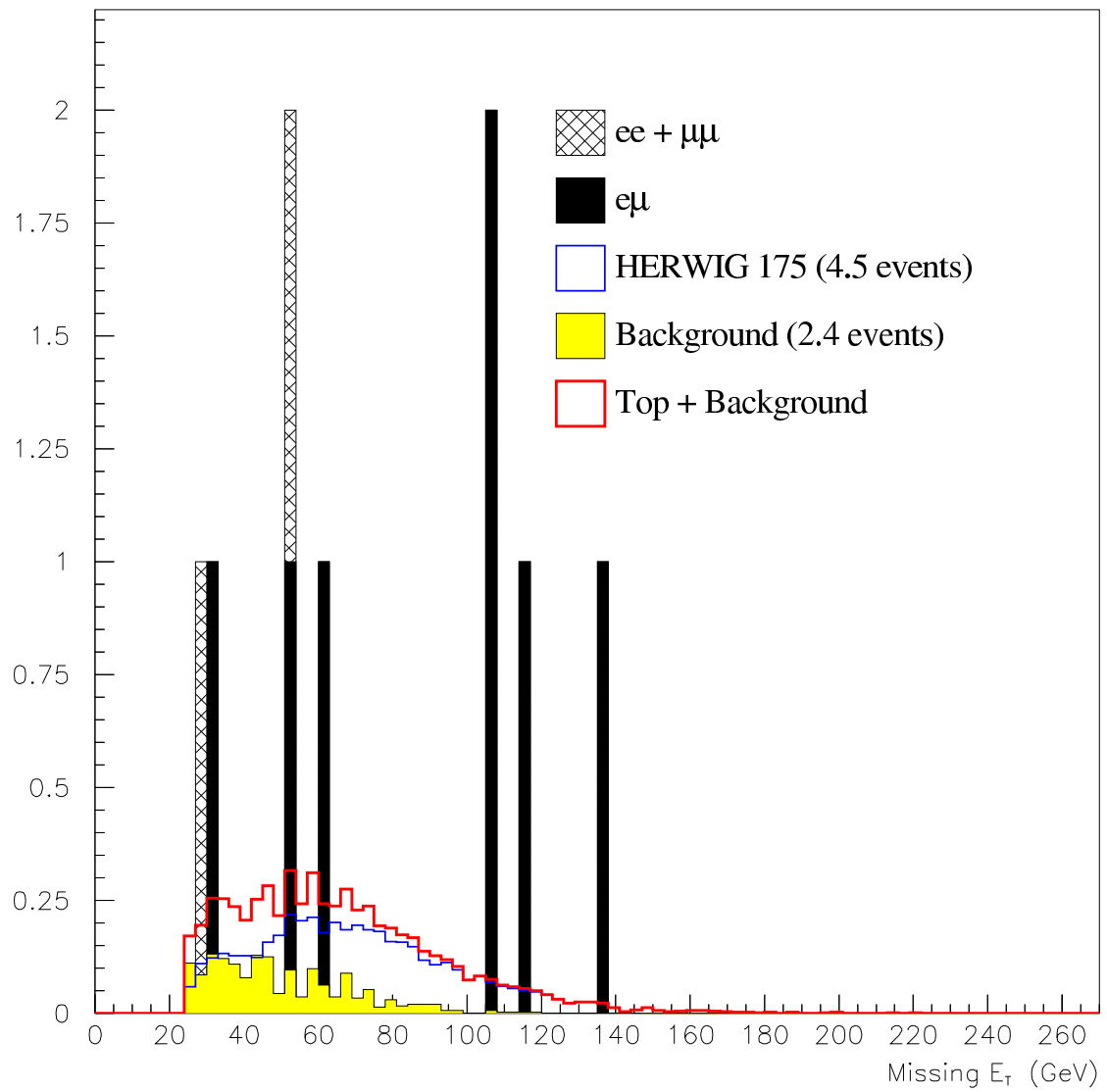


Figure 1.2: Missing transverse energy of $t\bar{t}$ dilepton candidates at CDF in Run I, compared to the expectation for combined signal and background. Taken from Ref. [20].

by the SM. Figure 1.3 shows the recent cross-section measurements from CDF.

This thesis describes the first measurement using the CDF detector of the top quark cross section in the dilepton channel with identified b jets. This work is closely related to the same measurement without the b -jet tagging requirement, which is referred to as the “pretag” measurement. The work contained in this thesis is referred to as the “tagged” measurement. Both analyses share the pretag selection and acceptance, and the tagged analysis uses pretag background estimates in the tagged background estimate. This thesis is organized as follows: The experimental apparatus is described in chapter two. Chapter three describes the pretag event selection, and chapter four the pretag acceptance, and validation. Chapter five outlines the event tagging algorithm and efficiency measurement. Chapter six describes the pretag background estimation, chapter seven, the tagged background estimate. Chapter eight describes the systematic uncertainties, and chapter nine the results and cross checks. The thesis concludes with chapter ten.

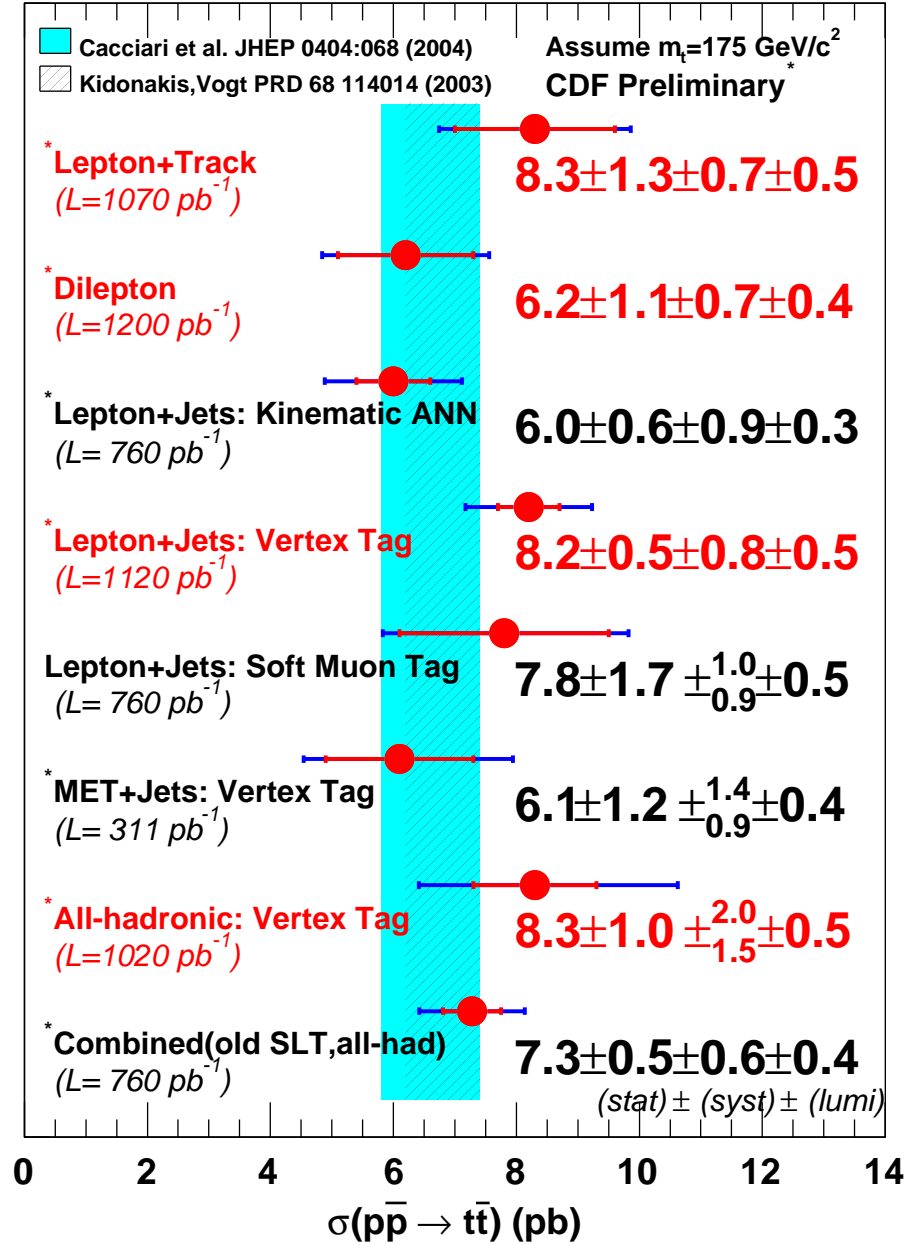


Figure 1.3: Recent results of the top quark cross-section measurements from CDF [24].

Chapter 2

Experimental Apparatus

Due to their large mass, top quarks can only be produced at today's most energetic particle accelerator, the Tevatron at Fermi National Accelerator Laboratory (Fermilab) in Illinois, USA¹.

The work contained in this thesis uses data produced by the collision of protons and anti-protons at centre of mass energy of 1.96 TeV at the Tevatron at Fermilab. The Collider Detector at Fermilab (CDF) is a multi-purpose detector surrounding one of two interaction points along the Tevatron ring, built to record information of the particles resulting from these collisions.

2.1 The Tevatron

The protons collided at the Tevatron begin as the nuclei of hydrogen atoms in a bottle of hydrogen gas in the corner of a room at Fermilab [25, 26]. These protons are accelerated to a final energy of 980 GeV in several steps. A schematic diagram of the Tevatron accelerator complex is shown in Figure 2.1.

The steps are:

¹and starting from 2008, at the Large Hadron Collider at CERN

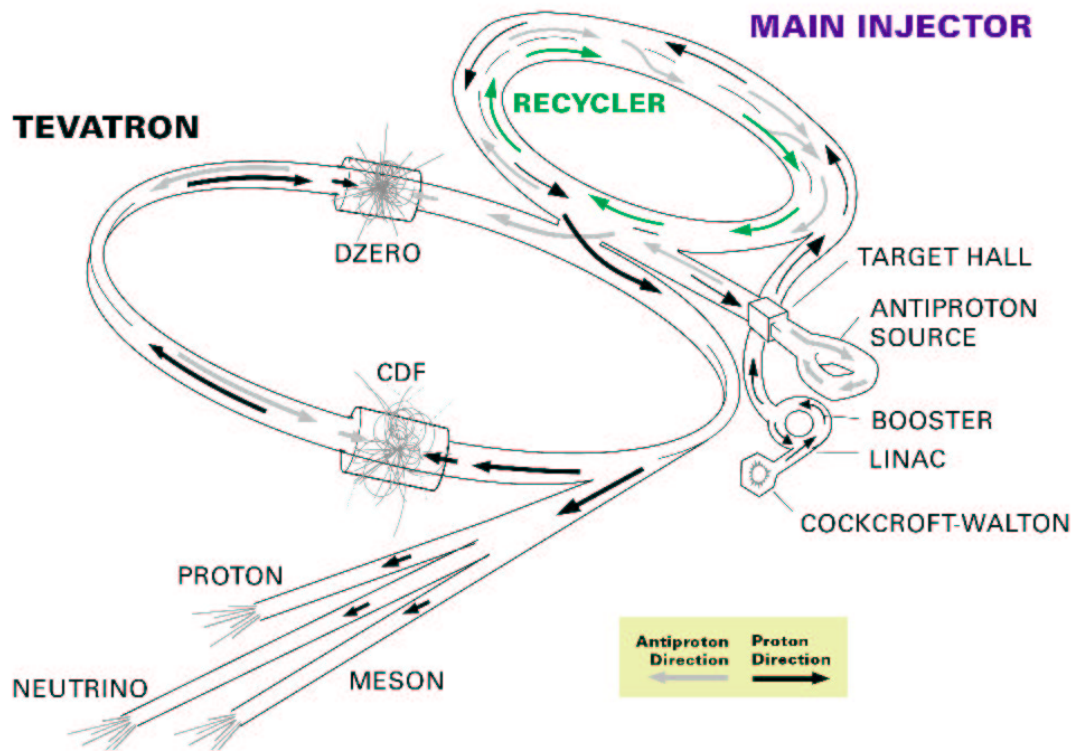


Figure 2.1: Schematic drawing of the Tevatron accelerator chain.

Cockcroft-Walton

The Cockcroft-Walton pre-accelerator provides the first stage of acceleration. H^- ions are produced inside the Cockcroft-Walton, and accelerated inside an electrostatic field to an energy of 750 keV [27].

Linac

H^- ions from the Cockcroft-Walton enter a 130 m linear accelerator (Linac). The Linac accelerates the ions using radio-frequency cavities to a final energy of 400 MeV. Once at this energy the H^- ions are passed through a carbon foil which strips off the electrons leaving bare protons.

Booster Ring

The protons are passed to the booster ring which is a standard circular accelerator (synchrotron) with a circumference of 475 m. The booster ring accelerates the protons to 8 GeV before they are transferred to the Main Injector.

Main Injector

The Main Injector is a synchrotron accelerator with a 3 km circumference. It accepts the 8 GeV protons from the booster, and accelerates a fraction of them to 150 GeV and injects them into the Tevatron. The Main Injector accelerates protons not sent into the Tevatron to 120 GeV and sends them to the anti-proton source.

The Anti-proton Source

Protons from the Main Injector strike a nickel target, producing anti-protons and other particles with a wide variety of momenta. The anti-protons are separated from the other resultant particles by a pulsed magnet, and are focused into a beam. The anti-proton beam is sent to a synchrotron called the Debuncher which reduces the spread of anti-proton energy. The anti-proton beam continues to the Accumulator, which is a storage ring used to accumulate anti-protons at a rate of about 15×10^{10} per hour. While in the Accumulator, the anti-protons also experience stochastic cooling [28] to reduce the beam emittance. Once the emittance and intensity of the anti-proton beam meet certain quality requirements, the anti-protons are then sent to the Tevatron for further acceleration. Anti-Proton production is an inefficient process and the ability to accumulate anti-protons and collect them into a tractable beam is the primary limit to the luminosity of the Tevatron.

Tevatron

The 150 GeV protons and anti-protons from the Main Injector are delivered to the Tevatron ring. The Tevatron is a superconducting synchrotron and accelerates the protons and anti-protons to their final energy of 980 GeV. The Tevatron primarily serves as a storage ring, storing the beams for several hours while they collide and the two experiments record data. The beams circulate in opposite directions while spiraling around each other in a helix in the same vacuum beampipe. Electrostatic separators keep the beam orbits separate.

2.2 The CDF Detector

The CDF detector surrounds one of the interaction points at the Tevatron with multiple layers of different types of detectors. Figure 2.2 shows a schematic diagram of CDF. CDF is an azimuthally and forward-backward symmetric apparatus designed to study $p\bar{p}$ collisions at the Tevatron. It is a general purpose solenoidal detector that combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection [29]. CDF uses a cylindrical coordinate system shown in Figure 2.3. The positive z -axis is defined along the direction of the protons. A trajectory of a particle is then described by the polar angle θ measured relative to the incoming proton beam, and the azimuthal angle ϕ around the beam axis. It is usually convenient to replace θ by the pseudo-rapidity, defined as $\eta = -\ln \tan(\theta/2)$ since it transforms linearly under boosts in the z -direction². Perpendicular to the beam, $|\eta|$ becomes 0 and extends to approximately 3.6 at the most forward part of the detector calorimetry system. A cone is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

²This is exactly true for the rapidity, defined as $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$, but also for the pseudo-rapidity in the limit of massless particles.

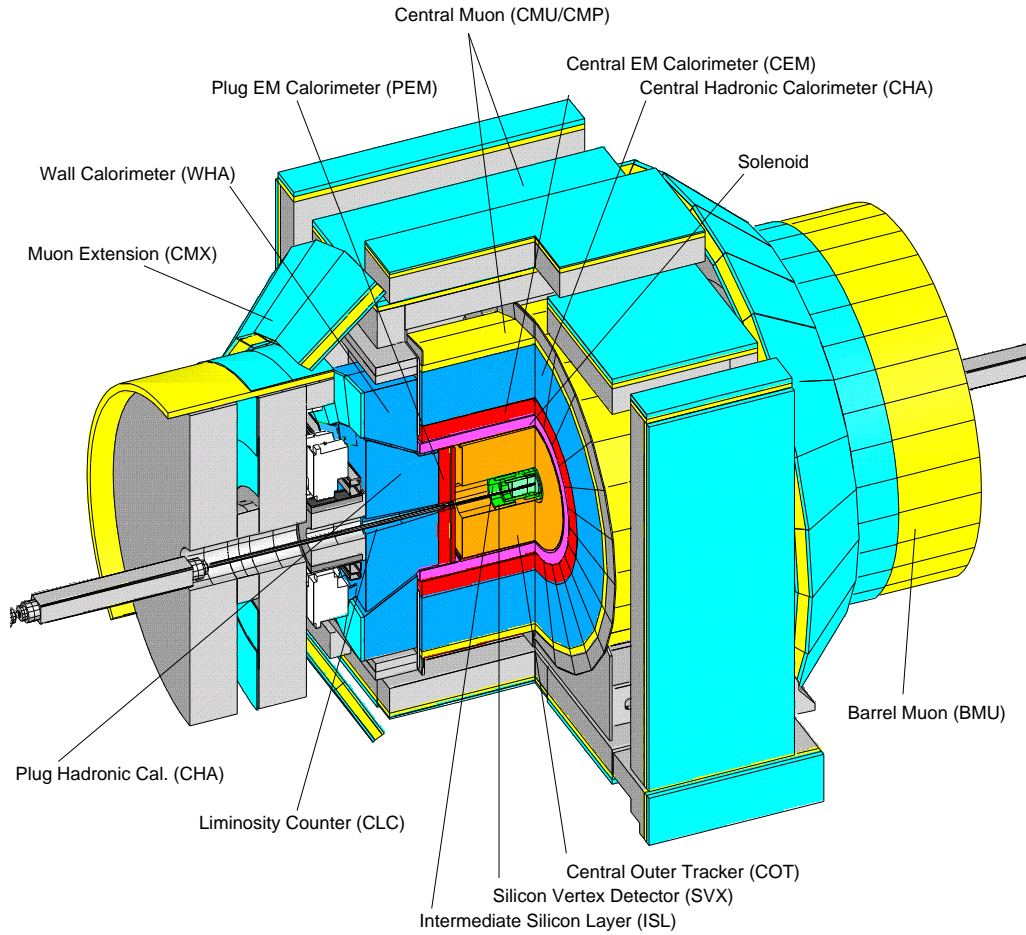


Figure 2.2: A pictorial representation of the Collider Detector at Fermilab.

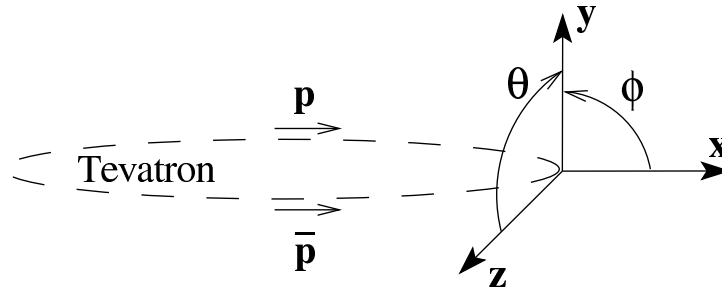


Figure 2.3: Definition of CDF coordinate system. The arrows for p and \bar{p} indicate direction of travel of the protons and anti-protons around the Tevatron ring.

2.2.1 The Tracking System

Particles emerging from the interaction point with significant transverse momentum first cross two tracking systems, the silicon semiconductor tracker and and open-cell wire

drift chamber, the Central Outer Tracker (COT). A schematic view of a quadrant of the CDF tracking system is shown in Figure 2.4. The entire tracking system is immersed in a 1.4 T magnetic field, produced by a superconducting solenoid. The magnetic field, parallel to the proton beam, causes charged particle trajectories to curve, allowing the tracking system to measure the trajectory and momentum of charged particles.

CDF Tracking Volume

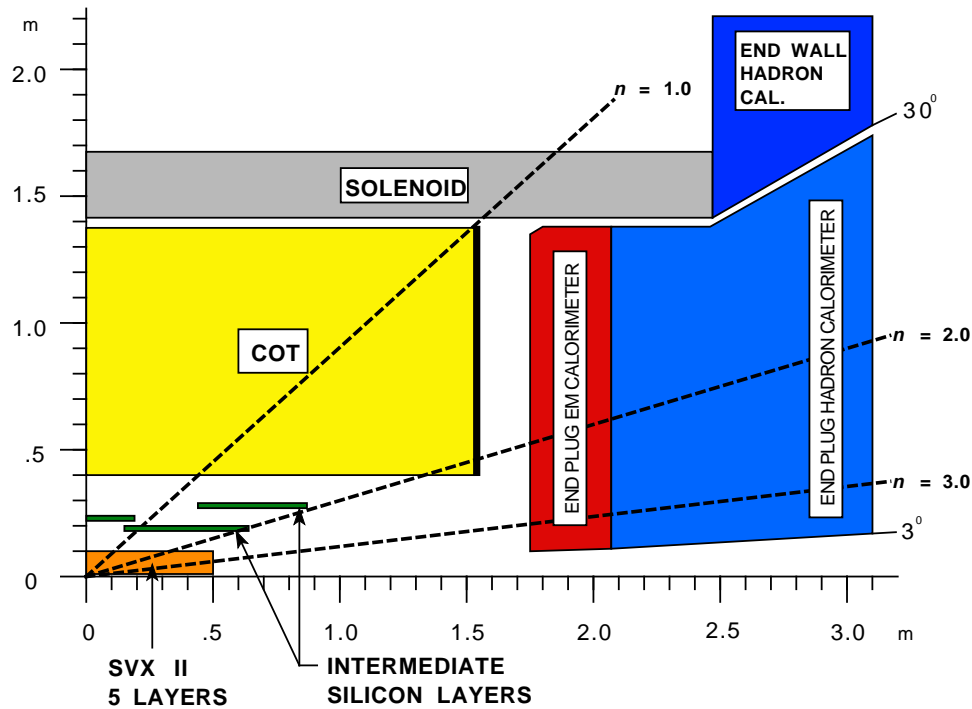


Figure 2.4: A schematic view of one quadrant of the CDF tracking system.

Silicon Tracking

The silicon tracking system consists of three parts. Layer00 (L00) is closest to the interaction region. Outside of Layer00 is the Silicon Vertex Tracker (SVX). The Intermediate Silicon Layers (ISL) lie in the region between the SVX and the COT. This arrangement is illustrated in Figure 2.5.

L00 is single-sided layer of radiation hard silicon microstrip arrays with a readout pitch of $50\text{ }\mu\text{m}$, which provides position measurements in the $r - \phi$ plane. This layer is

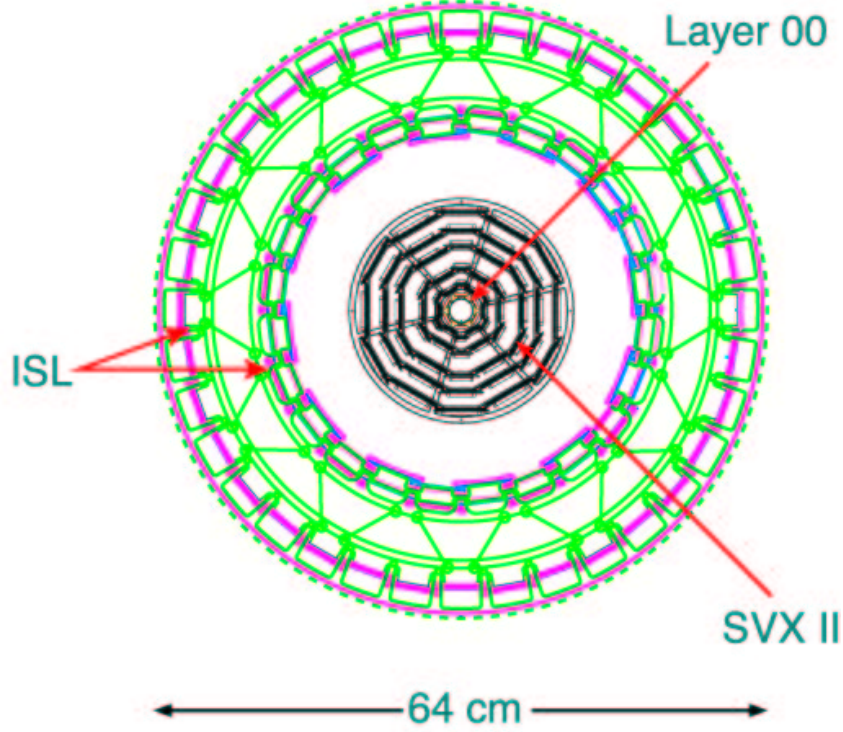


Figure 2.5: End view of the CDF silicon tracking system.

mounted directly onto the Tevatron beampipe.

The SVX is made up of five layers of double sided silicon wafers at radii of 2.4 to 10.7 cm from the beam, and is 96 cm long. Each layer is arranged in 12 wedges of strips in ϕ . One side of each wedge has its silicon strips aligned in the $r - \phi$ plane (axial), and the other side's strips are either parallel to the beam (stereo), or at a 1.2° angle to the beam (small angle stereo). The axial strip pitch is $55\text{--}65\ \mu\text{m}$, stereo strip pitch is $125\text{--}145\ \mu\text{m}$, and small angle stereo strip pitch is $60\text{--}75\ \mu\text{m}$. The SVX detector resolution for an axial measurement point is $\approx 12\ \mu\text{m}$.

The Intermediate Silicon Layers (ISL) are three layers of double-sided silicon detectors at radii 20 cm to 29 cm [30]. The ISL also have axial and small angle stereo strips with pitch $55\ \mu\text{m}$ (axial) and $73\ \mu\text{m}$ (small angle stereo). The ISL axial position resolution is $\approx 16\ \mu\text{m}$. The ISL provide an additional measurement point in the central region ($|\eta| < 1$) and two additional measurement points in the forward region ($1 < |\eta| < 2$).

Central Outer Tracker

The outermost part of the CDF particle tracking system is the Central Outer Tracker, a wire open-cell drift chamber. The gas in the COT is 50% argon and 50% ethane. It is 310 cm long and encloses the volume between an inner radius of 44 cm to an outer radius of 132 cm, providing charged particle tracking in the region $|\eta| \lesssim 1.0$ [31]. The COT consists of 96 concentric layers divided into 8 distinct groups called superlayers. Each superlayer has 12 layers of sense wires alternating with potential wires. The drift cells are arranged so that the wires are positioned in the middle of two gold-on-mylar cathode planes which are separated by 2 cm. Half of the superlayers are strung parallel to the beamline for axial measurements, the other half at 2° for stereo measurements. Axial and stereo superlayers alternate, starting with a stereo superlayer at the inner radius. Charged particles passing through the chamber ionize the gas, and a strong electric field causes the electrons to drift toward the sense wires, resulting in electric pulses on the wires. Both the timing and position of the pulse are read out electronically and recorded as a “hit” on the wire.

The COT is housed inside a solenoid providing a magnetic field of 1.4 T. The COT single-hit resolution is measured *in situ* using $Z \rightarrow \mu\mu$ decays. The single-hit resolution of the COT is 180 μm . Hits are combined to form track candidates. The transverse momentum of track candidates is then calculated by measuring their helical trajectories inside the magnetic field. The momentum resolution has been measured to be $\sigma_{p_T}/p_T^2 \approx 1.7 \times 10^{-3} [\text{GeV}/c]^{-1}$.

2.2.2 The Calorimeter System

The primary purpose of the CDF calorimetry system is to measure energy and position of charged and neutral particles. The central calorimeters surround the tracking detectors and lie outside of the solenoid. The calorimeters are divided into two regions. The

central barrel calorimeters cover the area from $|\eta| \leq 1$, and the plug calorimeters cover from $1.1 < |\eta| < 3.6$ [29]. CDF calorimeters are sampling scintillator calorimeters which measure the energy of particles that enter their volume. The calorimeters consist of alternating layers of active, scintillating material with absorbing material such as lead and steel. The absorbing material is chosen to favour the development of particle showers [32]. The shower is detected by the scintillator and converted to a light pulse with an amplitude proportional to the energy of the particles. Light guides collect the photons and pass them to photo-multiplier tubes (PMTs). The calorimeters are segmented into towers that are read out independently. The towers are arranged in such a way that they point toward the center of the CDF detector. This arrangement is known as “projective geometry”, and means that the boundaries between towers are at a fixed angle relative to the center of the detector.

The calorimeter system is divided into electromagnetic and hadronic calorimeter subsystems. The electromagnetic calorimeter is closer to the interaction region, and consists of the Central Electromagnetic Calorimeter (CEM) with pseudorapidity coverage of $|\eta| < 1.1$ and the Plug Electromagnetic Calorimeter (PEM) with coverage between $1.1 < |\eta| < 3.6$. The hadronic component of the calorimeter surrounds the electromagnetic component, and consists of the Central Hadronic Calorimeter (CHA), the Plug Hadronic Calorimeter (PHA), and the Wall Hadronic Calorimeter (WHA). The physical placement of these subsystems can be seen in Figure 2.2.

Electromagnetic Calorimeter

The electromagnetic calorimeter is the innermost layer of the CDF calorimetry system. It uses lead as the absorbing medium, and is intended to stop all electrons and photons. High-energy electrons lose most of their energy through *bremsstrahlung*, or the emission of a photon carrying a fraction of the energy of the electron. High-energy photons lose most of their energy through the production of positron-electron pairs. Both of these effects

take place under the influence of nuclear electric field of nearby atoms. All materials have an intrinsic *radiation length* (X_0) which is the mean distance over which an electron will lose all but $1/e$ of its energy to photon radiation [32]. This is also $7/9$ of the mean free path of a photon before it pair produces. The radiation length scales with the inverse square of the atomic number [6]. CDF uses lead in the electromagnetic calorimeters because of lead's large atomic number, and therefore short radiation length.

The photons emitted by an energetic electron may in turn pair produce, and the electrons and positrons from pair production may then emit photons. The resulting cascade of particles originating from a single electron or photon is called a shower. The radiation length for a material also gives the typical length scale for a shower. The CDF central (plug) electromagnetic calorimeter is 19 (20) interaction lengths deep, to ensure that most of the energy of electromagnetic showers will be fully contained in this part of the calorimeter [29].

The resolution of the electromagnetic calorimeter is expressed by $\sigma_E/E = 0.135/\sqrt{E_T}$ for the central component and $\sigma_E/E = 0.144/\sqrt{E} \oplus 0.7\%$ for the plug component (E measured in GeV).

Electromagnetic showers are maximally developed at a depth of $\approx 6X_0$. A multi-wire proportional chamber (CES) is embedded in the CEM and a scintillator strip detector (PES) in the PEM at this depth to improve the identification of electrons and photons by measuring their shower profile, see Figure 2.6. The CES and PES measure the $\phi - \eta$ position of the shower with a precision of ≈ 2 mm. The electron position resolution is also improved by the CES and PES by comparing the shower location with the extrapolation of the electron track candidate.

Presampling shower detectors located in front of the electromagnetic calorimeters are used in the central (CPR) and plug region (PPR) to improve photon identification. They are not used in this analysis.

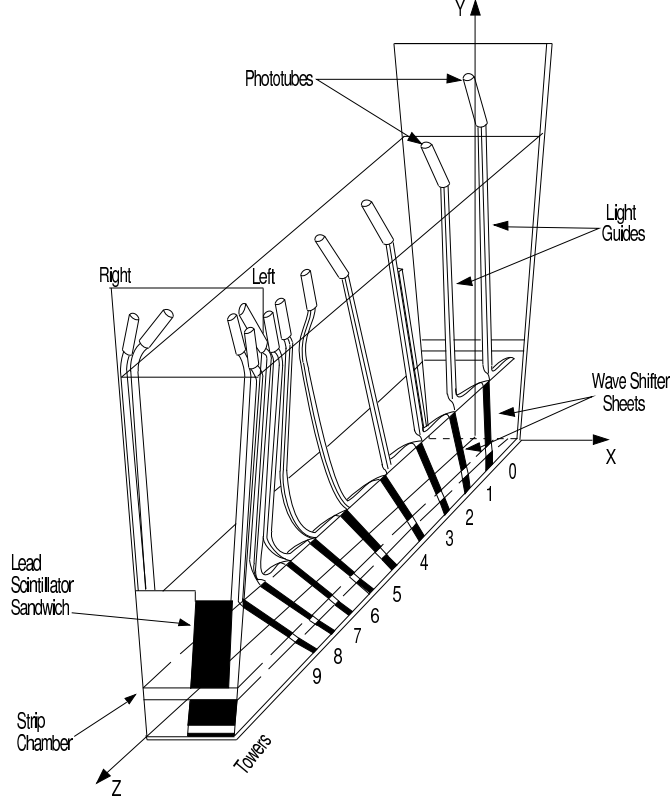


Figure 2.6: Diagram of a CEM wedge.

Hadronic Calorimeter

The hadronic calorimeter lies outside of the electromagnetic calorimeter, and consists of 23 alternating layers of absorbing and active layers. It uses iron as the absorbing material. The quantity interaction length, denoted λ is the mean free path of a particle before undergoing an interaction that is neither elastic or quasi-elastic in a given medium [32]. The electromagnetic calorimeter is only $\approx 1\lambda$ deep, so $\approx 2/3$ of the time the hadronic shower is only starting to develop. Hadrons, both charged and neutral, lose energy through ionization and collisions with the nuclei of the material through which they pass. The hadronic calorimeter is 4.5 interaction lengths deep in the central region and 7 interaction lengths deep in the plug.

The energy resolution of the hadronic calorimeter is given by $\sigma_E/E = 0.75/\sqrt{E}$ for the central component and $\sigma_E/E = (0.74/\sqrt{E}) + 3.8\%$ for the plug component (E measured

in GeV).

2.2.3 The Muon System

Muons are minimum ionizing particles. So unlike electrons they do not generally initiate particle showers, but lose only small amounts of their energy as they pass through large amounts of matter. CDF identifies muons by taking advantage of this fact. Drift chambers are positioned outside of the calorimeters, and also behind additional shielding. Most of the particles reaching these detectors are muons with small amounts of contamination from hadrons.

The CDF muon detection system is divided into four independent subsystems. The Central Muon detector (CMU) and the Central Muon Upgrade (CMP) give muon detection coverage of $|\eta| < 0.6$. The Central Muon Extension (CMX) detects muons with $0.6 < |\eta| < 1.0$. The Barrel Muon Detector (BMU) covers the pseudorapidity range $1.0 < |\eta| < 1.5$ but this system is not used in this search.

The CMU system consists of four layers of rectangular cells of drift chambers as shown in Figure 2.7. Each cell consists of a drift tube with a single wire in the centre. The chambers are filled with a mixture of argon and ethane gases. Muons passing through the muon chambers ionize the gas in the drift chambers. The hit position in the drift cells are determined from the drift time of the ions to the wire. One or more layers of scintillator are laid on one or both sides of the layers of drift tubes. These provide timing information, which allows for the association of muons with a particular interaction, since the drift time in the chambers is longer than the time between beam crossings. This is used to reject “out-of-time” muons both offline and for triggering.

The CMP system also consists of four layers of wire drift chambers. The design is identical to the CMU, except that the layers are staggered by a half cell per layer. The chambers are located outside of the CMP with 60 cm of steel between the CMU and CMP. This additional shielding reduces hadronic background. The CMU and CMP together

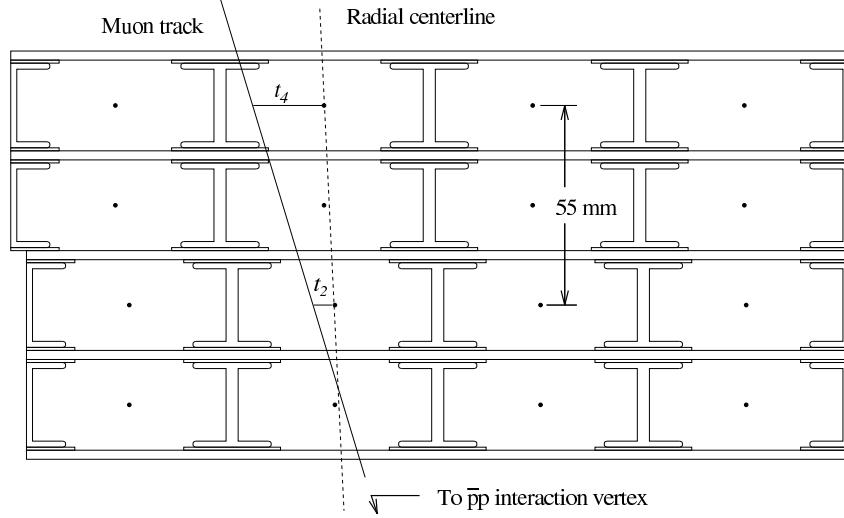


Figure 2.7: Four layers of drift chambers in the CMU.

form a rectangular box around the detector. The CMP is used together with the CMU detector in the central region, and the combined subsystem is referred to as the CMUP muon detector. The CMX system consists of eight layers of wire drift chambers placed in a conical pattern. A layer of scintillator plates is placed on the outermost edge of the CMP and CMX to provide additional timing information.

Hits from the layers of wires in the muon systems are connected into short tracks, or “stubs”. The matching of a stub to a track from the main tracking detectors are the basis of muon identification at CDF.

2.2.4 Cherenkov Luminosity Counter

At CDF the Cherenkov Luminosity Counters (CLC) measure the instantaneous luminosity of the $p\bar{p}$ interactions [33, 34]. The CLC is made up of modules of Cherenkov counters located at each end of the CDF detector, covering $3.6 < |\eta| < 4.6$. Each module is a conical unit, containing 48 smaller mylar cones which are filled with isobutane at atmospheric pressure. These collect Cherenkov light from particles emerging from the inelastic $p\bar{p}$ scattering. The amount of light detected by the CLC is directly proportional

to the number of particles generated from the $p\bar{p}$ collision.

The cross-section of inelastic $p\bar{p}$ scattering is known to about 3%. By measuring the scattering rate the instantaneous luminosity can be calculated as:

$$\mu f_{\text{BC}} = \sigma_{\text{in}} \mathcal{L}$$

In the above, σ_{in} is the inelastic $p\bar{p}$ cross-section, f_{BC} is the average bunch-crossing frequency at the Tevatron (2.5 MHz), \mathcal{L} is the instantaneous luminosity, and μ is the average number of $p\bar{p}$ interactions per bunch crossing. The luminosity detectors measure μ by counting the number of beam crossings where *no* scattering is observed. The number of $p\bar{p}$ interactions follows a Poisson distribution. By measuring the probability of no interaction occurring, the mean μ of the distribution can be determined. The probability of N interactions occurring is $\mathcal{P}(N, \mu) = e^{-\mu} \mu^N / N!$, so the probability of no interactions is $\mathcal{P}_0 = e^{-\mu}$. The uncertainties on the luminosity are from the understanding of the acceptance for the detectors for $p\bar{p}$ inelastic scattering as well as the 3% uncertainty on the value of σ_{in} , and combine to 6%.

2.3 Event Triggers and Data Acquisition

The Tevatron collides proton and anti-proton bunches at a rate of 2.5 MHz. The various CDF detector systems together produce 250 kilobytes of information per event, so recording this amount of data for each event is not feasible. Only a small fraction of all the collisions produce events of interest, so a “trigger” system selects these interesting events to be recorded for further study. The trigger and data acquisition system are outlined schematically in Figure 2.8.

2.3.1 The CDF Trigger System

Level 1

The Level 1 trigger is hardware based, and selects events using the least sophisticated methods. Calorimeter, track and muon detector information is used to determine if a given event is kept or discarded. This decision is made in $5.5 \mu\text{s}$ [29]. The custom built hardware quickly reconstructs simple quantities. Calorimeter information at this stage consists of the number of towers with recorded energy above an energy threshold, the summed transverse energy over all towers and the single trigger tower energy. Tracking is done using only axial information from the COT and compares hit information to a set of predetermined patterns. The tracking algorithms then output the p_T and ϕ of the track. Tracks are extrapolated to the calorimeter and muon systems so they can be used in electron and muon triggers. Track segments from the muon drift tubes and scintillator information are both used for muon identification. Level 1 accepts events at a rate of up to 50 kHz.

Level 2

Level 2 is also a hardware-based trigger, which takes events passing the Level 1 trigger as input. Level 2 has a maximum acceptance rate of 300 Hz, so there is up to $20 \mu\text{s}$ processing time for each event. This means that more sophisticated reconstruction is possible. Calorimeter information is clustered using a simple algorithm, to be used for photon, electron and jet identification. Information from the shower maximum detectors and silicon detectors is also used.

Level 3

The Level 3 trigger does a full event reconstruction at the software level. Events passing the Level 2 trigger are processed on a dedicated computing farm using a simplified version

of the object identification software that is used offline for analysis. Event processing at Level 3 can take up to one second to complete. Event classification is done using objects such as photons, electrons, muon, and jets. Selected events are written to tape at a maximum rate of 75 Hz.

2.3.2 Triggers Used in this Analysis

Over time operating conditions of the experiment change, and trigger definitions change as well to optimize data collection. An example of these changing run conditions is the increasing instantaneous luminosity, especially at the beginning of stores. But the high- p_T lepton triggers used in this analysis have, for the range of data used here, remained relatively stable.

Central high- E_T Electrons

The central high- E_T electron trigger selects electron-like objects in the central calorimeter. At Level 1 it requires a tower in the central calorimeter with at least 8 GeV of E_T , the ratio of hadronic to electromagnetic energy to be less than 0.125, and a matched track with $p_T > 8.3$ GeV/ c . At Level 2 the calorimeter energy is defined in terms of an energy cluster, which must be seeded by a tower with at least 8 GeV of energy. The cluster must have $E_T > 16$ GeV. At Level 3 the standard electromagnetic object clustering is used, and the E_T is required to be at least 18 GeV. The cluster must pass minimal shower shape requirements and be well matched to a track with $p_T > 9$ GeV/ c .

Forward high- E_T Electrons

Electrons in the forward region do not cross as much of the tracking system as do electrons in the central region. So the fast tracking in the trigger system can't match tracks to calorimeter clusters as well. This leads to a higher rate of misidentification for electrons in the forward region. To combat these "fakes", the plug electron trigger also requires \cancel{E}_T in

the event to preferentially choose electrons from W decays (which includes W decays from top events.) At Level 1, a cluster with $E_T > 8$ GeV, the ratio of electromagnetic to hadronic energy be less than 0.125. All towers with at least 1 GeV of energy are used to calculate the \cancel{E}_T which must be at least 15 GeV. At Level 2 the electron object is refined by requiring the energy cluster to have $E_T > 20$ GeV. At Level 3 the electron is required to have $\cancel{E}_T > 15$ GeV.

Central high- p_T Muons

The CMUP trigger path starts at Level 1 with the requirement of stubs in both the CMU and CMP, which are matched to a track with $p_T > 4$ GeV/ c . The only additional requirement at Level 2 is that the track have $p_T > 8$ GeV/ c . At Level 3, the muon stub is required to be matched to a fully reconstructed track with $p_T > 18$ GeV/ c . The CMX trigger requirements are identical except that the stub is in the CMX and the p_T thresholds at Level 1 and Level 2 are higher: 8 and 10 GeV/ c , respectively.

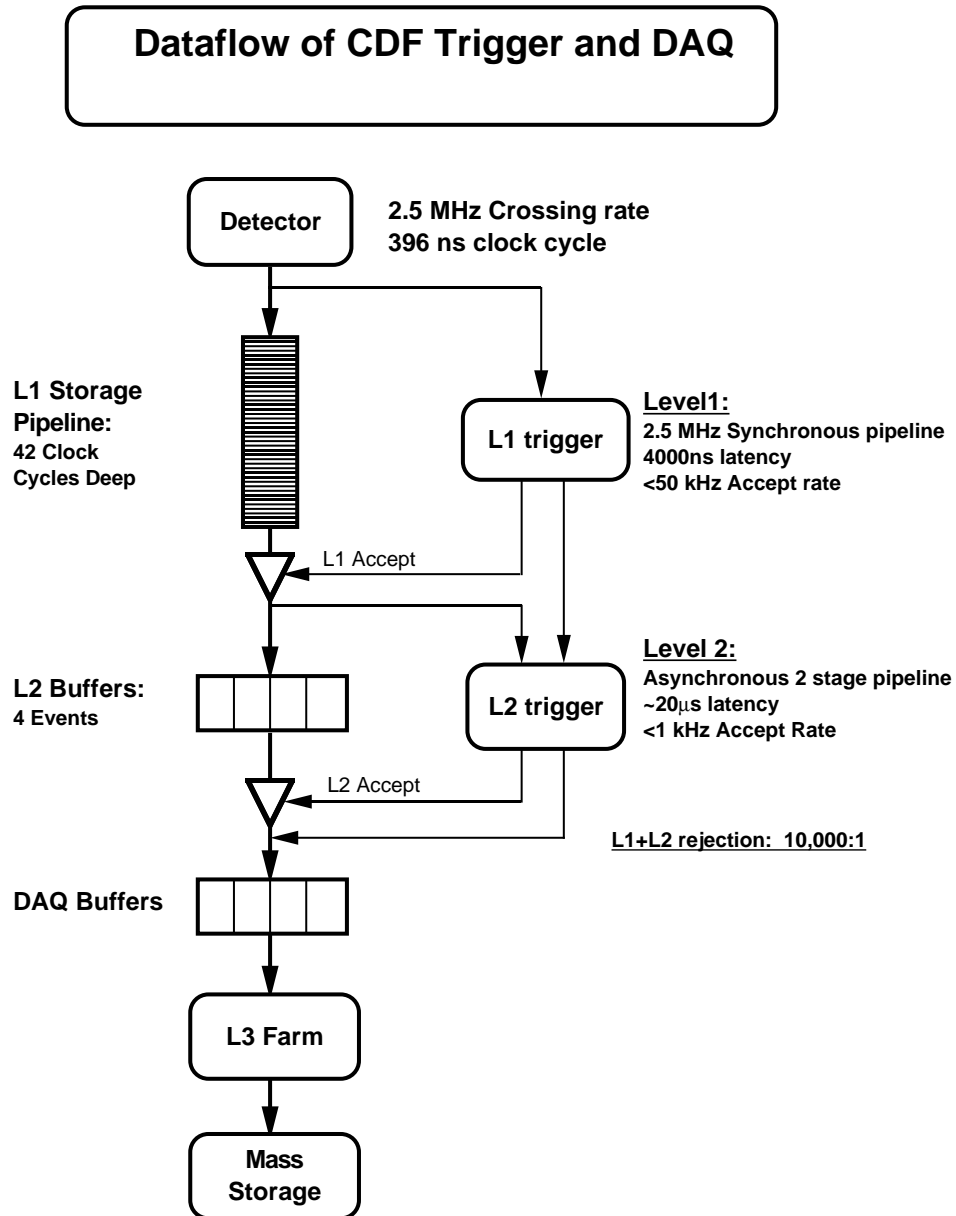


Figure 2.8: Schematic representation of the data flow in the CDF trigger and data acquisition systems.

Chapter 3

Pretag Event Selection

The goal of this analysis is to identify $t\bar{t}$ events in the dilepton decay channel, while rejecting other events with the same experimental signature as often as possible. As such, the first step in event selection is the identification of two leptons with opposite charge from W boson decay and significant \cancel{E}_T . A top quark decay produces a W boson and a b quark, which hadronizes. Thus we also select events with at least two jets, one of which must be identified as coming from b decay by a “b-tagging” algorithm. We restrict the angular distribution of these objects to reduce the contribution to the selected events by backgrounds such as Drell-Yan ($Z/\gamma^* \rightarrow \ell\ell$).

This thesis describes an analysis of dilepton events, with at least one identified b -jet, and is called the “tagged” cross section. This is closely related to the “pretag” cross section, where no requirement relating to the flavour of the jets is made. These two analyses share the event selection up to the tagging requirement, therefore the pretag acceptance is an input to the tagged cross-section calculation. The background estimate of the pretag selection is used in the calculation of the tagged background estimate. This chapter describes the pretag event selection, chapter 5 describes the b -jet tagging algorithm and selection. There are several advantages to using a common pretag acceptance, with a separate calculation for the event tagging efficiency

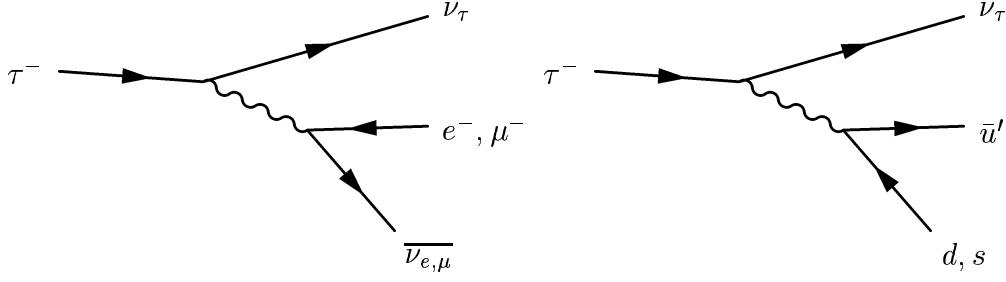


Figure 3.1: Kinematically allowed decay modes of the tau lepton. The τ is more massive than the estimated mass of the c quark, but no charmed meson is lighter than the τ , so $\bar{c}s$ and $\bar{c}d$ decays are not allowed.

- The work of validating the acceptance calculation only has to be done once.
- The pretag acceptance can easily be compared to previous iterations of the pretag analysis.
- The event tagging efficiency can be validated against other measurements of event tagging efficiencies.

3.1 Lepton Definition

The lepton selection criteria were developed to isolate the high p_T lepton decays from W bosons, while rejecting jets that can fake a lepton in the detector. One lepton is identified clearly as either an electron or a muon. Because τ leptons decay quickly (see Figure 3.1) this analysis does not identify τ leptons explicitly, but they do pass the selection cuts when they decay to either electrons or muons, as in Figure 3.1 or to a single hadronic track.

The second lepton is identified only as a track which is isolated from other activity in the tracking system (“isolated track”). By selecting these leptons without the use

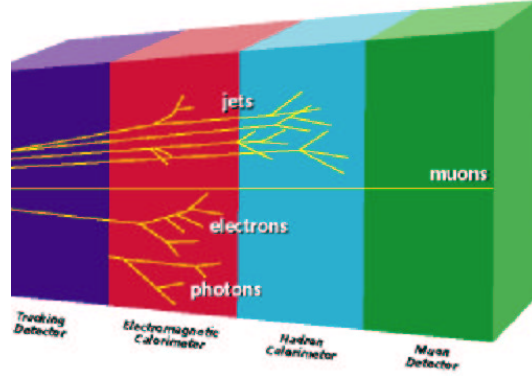


Figure 3.2: Schematic diagram of the experimental signatures of different objects as seen by the CDF detector. Note how muons and electrons both leave isolated tracks in the tracker, while jets are identified by clusters of several tracks.

of calorimetry or the muon systems, the geometric acceptance to leptons outside of the calorimeter and muons systems is increased. As illustrated in Figure 3.2 various particle types penetrate the detector volume to varying depths, and leave differing signatures. Electrons, muons and jets all leave tracks in the tracking volume. But electrons shower in the electromagnetic calorimeter and have single tracks. Muons deposit little energy in either calorimeter and as well leave single tracks in the tracking volume. These properties are exploited to reject jets, which have several tracks in close proximity, and have wide showers in both the electromagnetic and hadronic calorimeters. So by selecting for isolated tracks we reject jets, and yet increase our acceptance to leptons which fall in the cracks of the calorimetry.

Most significantly, this definition includes a large subset of events where one or both W s have decayed to a τ and ν_τ . Because the W is equally likely to decay to any one of the three generations of $\ell + \nu$ pairs, 5/9 of top events where both W s decay to leptons contain one or two tau leptons. By their nature any dilepton selection will accept leptonic τ decay products, which are 35% of all decays, although there will be some inefficiency

Primary Lepton Type	ee	$e\mu$	$\mu\mu$	$e\tau$	$\mu\tau$	$\tau\tau$	total
electron	20.0	34.0	0.0	8.9	2.6	0.5	66.2
muon	0.0	7.3	19.3	0.6	6.3	0.5	33.8
total	20.0	41.3	19.3	9.5	8.9	1.0	100.0

Table 3.1: Fraction (in %) of the (oppositely charged, ≥ 2 jet) acceptance that is attributable to each possible generated lepton pair. The two middle rows show the events divided according to whether they were reconstructed with an electron or muon as the primary lepton. The majority of events are accepted as electron plus track because there is more geometric acceptance for electrons, and because of the way in which events are categorized when both leptons can be fully reconstructed.

because some of the momentum of the original tau decay will be lost to the two neutrinos produced. Including both leptonic and hadronic decay modes, 85% of τ decays have just a single charged track in the final state [6]. Using an isolated track for the second lepton therefore significantly increases the acceptance. About 20% of the total acceptance with our selection is from events with one or two τ s. Table 3.1 shows the acceptance for both electron and muon candidates as the primary lepton by the generated lepton type.

3.1.1 Electron Selection

Electrons selected in this thesis are of two types. Central electrons are found in the region $|\eta| < 1.0$, have a track in the central tracker (COT), and deposit their energy in the central electromagnetic calorimeter (CEM). Plug electrons in the region $1.2 < |\eta| < 2.0$ have tracking information only from the silicon tracking system (the COT does not reach this far forward), and deposit their energy in the PEM calorimeter.

Both plug and central electron identification is done using the same principles. Electrons should deposit the majority of their energy in the electromagnetic portion of the calorimeter. Electrons leave a trail of ionization that is detected in the tracking system,

and points to the energy deposit in the calorimeter. Electron shower shapes are narrow compared to hadronic showers which are broad.

Central Electrons

Geometric and Kinematic Selection Central electrons are required to deposit their energy in the fiducial region of the CEM calorimeter. Electromagnetic clusters are required to have at least 20 GeV of transverse energy, and have an associated track with at least 10 GeV/ c of momentum. The ratio of the transverse energy to the momentum must be at least 2:1. These looser requirement on the momentum is for electrons which emit photons via the *bremsstrahlung* process. These bremsstrahlung photons typically deposit their energy in the same calorimeter tower as the electron, so the transverse energy measured in the calorimeter is typically the same as an electron alone, although the electron's momentum is reduced.

Track Quality Selection Tracks are required to have a reconstructed vertex within 60 cm of the interaction point. This keeps as many tracks as possible, but still requires tracks to be in the fiducial volume of the tracking system. Tracks must have at least 5 COT hits in each of 3 axial segments and 2 stereo segments. Track segment requirements ensure enough information is available to make a reliable p_T measurement.

Track Cluster Matching The track and electromagnetic cluster are matched by extrapolating the track to the shower maximum detector embedded in the electromagnetic calorimeter. A well measured electron will have a short distance between the shower centroid measured along the beamline, and the place where the extrapolated track crosses the plane of the shower maximum detector. We require this distance to be less than 3 cm. We further require that the product of the track charge (Q), and the distance between the shower centroid measured perpendicular to the beamline and the place where the extrapolated track crossed the plane of the shower maximum detector (ΔX), be -3.0 cm

$< Q\Delta X < 1.5$ cm. The asymmetric limit is due to the bending of the electron track in the magnetic field of the solenoid. Backgrounds from photons and $\pi \rightarrow \gamma\gamma$ have charged tracks which tend to point to the photon cluster.

Shower Shape and Isolation The composition of the electromagnetic calorimeter is selected to favour the development of electromagnetic showers, and is designed to be deep enough to contain these showers. As such, very little energy of a high energy electron reached the hadronic portion of the central calorimeters. We require all but 5.5% of the electron transverse energy to be deposited in the electromagnetic portion of the central calorimeter, with an extra allowance of $0.00045 \times E$ (in GeV) for very high transverse momentum electrons. The lateral shape of the shower is required to be narrow, and match expectation for electrons. The total amount of transverse energy in the surrounding calorimeter towers is required to be less than 10% of the total transverse energy.

Conversion Veto Photons can produce electron-positron pairs. These leptons can be identified in this analysis, but are not from the decay of a W or Z boson. These electrons are selected against by looking for tracks close to the primary leptons in θ , with the opposite charge, and which appear to originate from the same point.

Plug Electrons

The selection requirements for plug electrons are similar to central electrons, and take advantage of the same behaviours.

Geometric and Kinematic Selection Plug electrons are required to have $E_T > 20$ GeV, and be in the region $1.2 < |\eta| < 2.0$. The cut-off at $|\eta| = 2.0$ is motivated by the fact that top decays are distributed approximately equally in a sphere, while the electron identification becomes less efficient, and has higher fake rates at high values of $|\eta|$.

Silicon Track Reconstruction The tracks associated with plug electrons are reconstructed differently than the tracks reconstructed from the COT-dependent tracking algorithm used for central tracks. The COT extends out to $|\eta| < 1.0$ but tracks in the high $|\eta|$ region do cross parts of the COT. Tracking in the forward region is done using only information from the silicon system and the electromagnetic calorimeter.

Track Quality Selection The track z_0 must be within 60 cm of the center of the detector, as with central electrons. Tracks for plug electrons are required to have at least three silicon hits out of 6 or 7 possible hits, depending on which part of the silicon system the trajectory crosses.

Shower Shapes and Isolation Plug electrons are required to meet the same calorimeter-based isolation requirements as central electrons. As such, the plug electron must make up at least 90% of the energy in a cone of $\Delta R = 0.4$ around the plug electron center. The energy cluster must have at least 95% of it's total transverse energy in the electromagnetic part of the calorimeter. Plug electrons are required to conform to expected shower shapes, although the exact shape is different than for central electrons because of the different geometry of the plug calorimeter and a different shower maximum detector. Plug electron calorimeter clusters must have $\chi^2 < 10$ compared to the expected cluster profile for electrons. Also, the centroid of the cluster of calorimeter towers, and the centroid of the shower maximum detectors must be close to each other.

3.1.2 Muon Selection

Muon selection requirements are guided by the fact that muons are minimum ionizing particles. They usually do not initiate particle showers in the calorimeter, and exit the detector leaving little ionization. Muon detectors are therefore placed outside the calorimeter and extra shielding, where no other charged particles from the $p\bar{p}$ interaction

should reach. In the CDF detector a fully reconstructed muon is characterized by a track which points to a segment (“stub”) in the muon detectors. As explained in Section 2.2.3 the CMU and CMP muon detectors lie in concentric volumes with shielding between them. A muon which is reconstructed as passing through both CMU and CMP detectors is called a CMUP muon. A CMX muon is reconstructed in the CMX muon detector.

Kinematic and geometric requirements The muon track must have $p_T > 20 \text{ GeV}/c$. CMX muon tracks are also required to have passed through all the measurement layers of the COT, which ensures that the trigger is fully efficient. We reject muons which are not fully contained in the fiducial volume of the relevant muon detectors to reduce edge effects.

Muon detector signature To veto hadronic showers that leak into the muon chambers, CMUP muons are required to have stubs in both the CMU and CMP detectors. CMX muons only require a stub in the CMX detector. The quality of the match between the position of the muon track and the stub is evaluated by measuring the distance between the stub and track, minimum thresholds are imposed.

Track quality requirements Muon track selection is designed to reject backgrounds specific to muons. Because the only measure of the muon momentum is the track momentum, stringent cuts are made to the track quality. Analogous to electron track quality, muons are required to have at least 3 axial and 2 stereo segments in the COT, and the track $|z_0|$ must be less than 60 cm. Additionally, we require the track impact parameter $|d_0| < 250 \mu\text{m}$. Timing information is used to reject cosmic rays, which are real muons passing through the detector. The cosmic veto algorithm flags tracks that are not consistent with the event interaction vertex and which cross the detector out of time with the beam crossing.

A rare but difficult background to estimate are kaons which decay *inside* the tracking

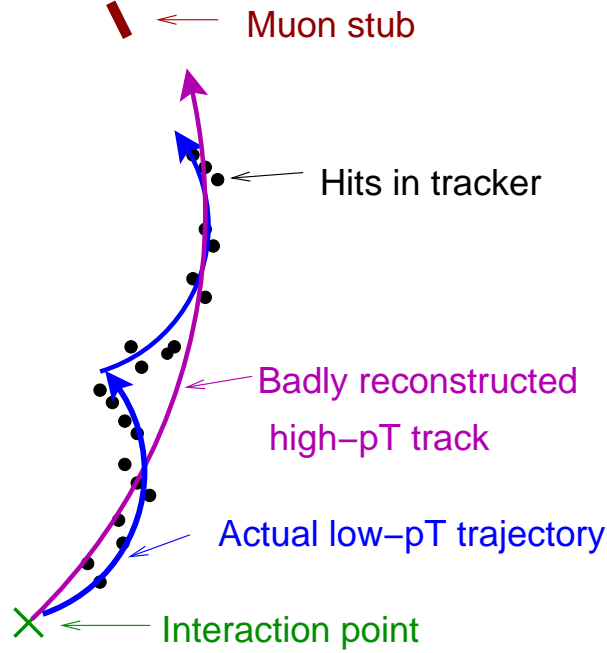


Figure 3.3: Drawing of the mechanism by which the decay in flight of a medium- to low- p_T kaon can be reconstructed as a high- p_T muon.

volume. As depicted in Figure 3.3 the resulting track is curved with a kink. The tracking algorithm attempts to reconstruct every track as a single helix and fits the kinked track as such. The resulting track has higher p_T than the true p_T , and poor fit quality. To combat these background events, muon tracks are required to have a χ^2 fit result such that the probability for the track to have a worse χ^2 for the same number of degrees of freedom in the fit is smaller than (10^{-8}) . The χ^2 probability cut is made in data only, and is not applied in the simulation due to an error in the simulation.

Calorimeter signature As already stated, minimum ionizing particles such as muons will deposit very little energy in the calorimeter. Thus we impose maximum energy thresholds of $E_{had} < (6 + \max(0, 0.028 \times (p - 100)))\text{GeV}$, and $E_{em} < (2 + \max(0, 0.0115) \times (p - 100))\text{GeV}$. The little energy deposited in the calorimeter is also required to be isolated from other energy deposits to reject muons from the semi-leptonic decay of a b hadron. At least 90% of the E_T in the towers surrounding the muon must be attributed to the

muon.

3.1.3 Track Selection

The selection of the second lepton using only tracking information is a unique feature of this analysis. Since no calorimeter information is used, the geometric acceptance in the central region is increased, and by not explicitly identifying the lepton as either an electron or muon, the acceptance to τ leptons is increased.

Geometric and Kinematic Selection Tracks are required to have transverse momentum of at least 20 GeV/ c . There is no explicit fiducial requirement, but since the track must pass through the tracking system, the tracks must effectively be in the region $|\eta| \lesssim 1.1$.

Track Quality Selection Tracks must be “well measured” to reduce the number of jets which are misidentified as isolated tracks. This means that there must be enough information from the tracking system to adequately measure the track momentum. We require at least 5 hits on each of at least 24 axial COT segments, and at least 20 stereo COT segments. If the track passes through at least three layers of the silicon tracking system, it is required to have at least three silicon hits. This requirement is released for events in which the silicon system is not operating optimally. Just as with the identified muon candidates, tracks are required to have $|d_0| < 250 \mu\text{m}$ and pass the same χ^2 probability requirement as primary muons.

Track Isolation Just as with the primary lepton, isolation is a powerful tool to reject background leptons. To calculate isolation, all the tracks with minimal quality requirements and transverse momentum of at least 0.5 GeV/ c within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ are summed. If the transverse momentum of the track

accounts for at least 90% of the total transverse momentum in the cone, the track is considered isolated.

Interaction Vertex Selection The primary lepton and track lepton are required to be distinct, but their origin along the beamline, z_0 , must be consistent with having come from the same interaction. We require $\Delta z_0 < 5$ cm.

3.2 Jet Definition

As described in Section 1.1, jets are produced from QCD interactions among coloured partons. As the separation of two partons becomes greater, the energy of the system increases until new quarks and gluons are created from the vacuum to create colourless objects, forming jets of particles traveling in the same general direction as the original parton.

Jets are inherently messy objects. It is somewhat arbitrary to define which particles belong to a jet, and to accurately measure the jet’s energy. Even so, there are several standard ways to define a jet, and to correct the measured energy to obtain an estimate of the original parton energy.

Jets are reconstructed by grouping adjacent towers in the calorimeter which register energy deposition above a certain threshold. The energy deposit in these towers are then treated as a group, or cluster. The total energy in the cluster is not an adequate representation of the true energy of the jet, and corrections are made to this energy based on several factors that will be described later.

3.2.1 Jet Clustering

Jets in this analysis are defined using the “JetClu” algorithm [35] which means that jets are based on the clustering of energy deposition in the calorimeter. Jet clustering begins

with a single tower called a “seed” tower. All towers in an event with $E_T > 1$ GeV are considered seed towers, and these are ordered in decreasing E_T . All adjacent towers with $E_T > 1$ GeV and within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ ¹ of the seed tower are included in the jet cluster. This process is repeated for all seed towers. Jets sharing towers are merged if more than 50% of the towers overlap, otherwise jets are split apart so each tower belongs to only one jet. Jet energies are the summed energy of the jet towers, with the jet direction being the E_T -weighted centroid of the calorimeter cluster.

3.2.2 Jet Corrections

The energy measured by the calorimeter is not necessarily the total energy of the jet. For example jets may contain photons, neutrons or muons. These particles do not create particle showers and deposit their energy in the active region of the calorimeter in the same way as other charged particles. The largest deviation from this behaviour is for muons, which, deposit small amounts of energy in the calorimeter, and usually pass through almost unaffected. The jet energy also does not account for the binding energy of nuclei, which is not measured by the calorimeter when nuclei break apart [32].

Jet energies are corrected for non-uniformity in detector response as a function of jet energy, angle and the fraction of total jet energy deposited in the electromagnetic calorimeter.

3.2.3 Jets for Analysis

A jet used in the pretag analysis must have $E_T > 20$ GeV and $|\eta| < 2.0$. Jets overlapping with either the primary or track lepton are excluded. Section 5.1 will describe the selection of tagged jets in the tagged analysis.

¹For this analysis we consider only jets with $\Delta R < 0.4$, although any cone size may be chosen. Typical cones used in CDF analyses are 0.4, 0.7, 1.0.

3.3 Missing Transverse Energy Reconstruction

Missing transverse energy is a measurement of the transverse energy imbalance in the detector. Thus, when all objects in the detector have been measured as well as possible, momentum conservation should be observed in the x and y directions. The remaining imbalance of transverse momentum in the event is interpreted as a neutrino escaping the detector. Missing transverse energy, \cancel{E}_T , is the vector which balances measured transverse energy sources in the detector.

$$\cancel{E}_{\{x,y\}} = - \sum_i (E_{\{x,y\}})_i \quad (3.1)$$

Where the sum is over all calorimeter towers. Some objects such as muons carry considerable energy, but do not deposit most of their energy in the calorimeters. So \cancel{E}_T needs to be re-calculated to account for known sources of mis-measurement.

3.3.1 Track Corrections

As mentioned earlier, muons leave little energy in the calorimeter, and thus do not contribute to the calculation of \cancel{E}_T . Other minimum ionizing particles also leave small energy deposits in the calorimeter when in fact they have considerable momentum. Other particles which pass through a gap in the calorimeter system also do not contribute to \cancel{E}_T in the event. \cancel{E}_T is corrected for muons by subtracting the muon p_T from the \cancel{E}_T . For all reasonably well-measured tracks in the event, with $p_T > 10 \text{ GeV}/c$, we check the amount of energy in the 3x3 block of calorimeter towers to which the track points. If the ratio of track p_T to the energy in the block is smaller than 0.7, we correct the \cancel{E}_T by subtracting the track p_T from the energy in the block of towers.

3.3.2 Jet Corrections

For all jets in the event with $E_T > 10 \text{ GeV}$ in equation 3.1 above, we replace the raw energy with the corrected energy.

3.4 Event Selection

To optimize event selection for dilepton channel decays of $t\bar{t}$, we require an ensemble of objects unique to the dilepton channel decays. We require one primary lepton and a track lepton, which must be oppositely charged, one from each W boson decay. We require two reconstructed jets, one from each b quark. In Chapter 5 we will discuss how jets are identified as coming from the hadronization of b quarks. We additionally require a significant amount of \cancel{E}_T which we attribute to neutrinos escaping the detector. We additionally select for relationships between these objects to reject background sources with similar decay signatures.

3.4.1 Basic Event Selection

The primary lepton type seeds the event selection and classifies the event type. Event selection is done using a hierarchy of primary lepton types. We first search for a primary lepton which is an electron in the central region (CEM electrons), then for muons in the central region (CMUP muons), then muons in the CMX detector (CMX muons), and finally electrons in the plug region (PHX electrons). For each primary lepton type, we order leptons in descending E_T or p_T and continue with all selection criteria until the event has failed a selection, or the event passes all selection criteria.

Once a primary lepton has been selected, we search for tracks passing the track kinematic, quality, and isolation cuts, in order of decreasing p_T . If no suitable track leptons exist we go to the next primary lepton.

Armed with a primary and track lepton, \cancel{E}_T is corrected, and we reject events with $\cancel{E}_T < 25$ GeV. This requirement is a good tool for rejecting backgrounds with two leptons and no neutrinos, such as Drell-Yan ($Z/\gamma^* \rightarrow \ell\ell$) events, see Figure 3.4.

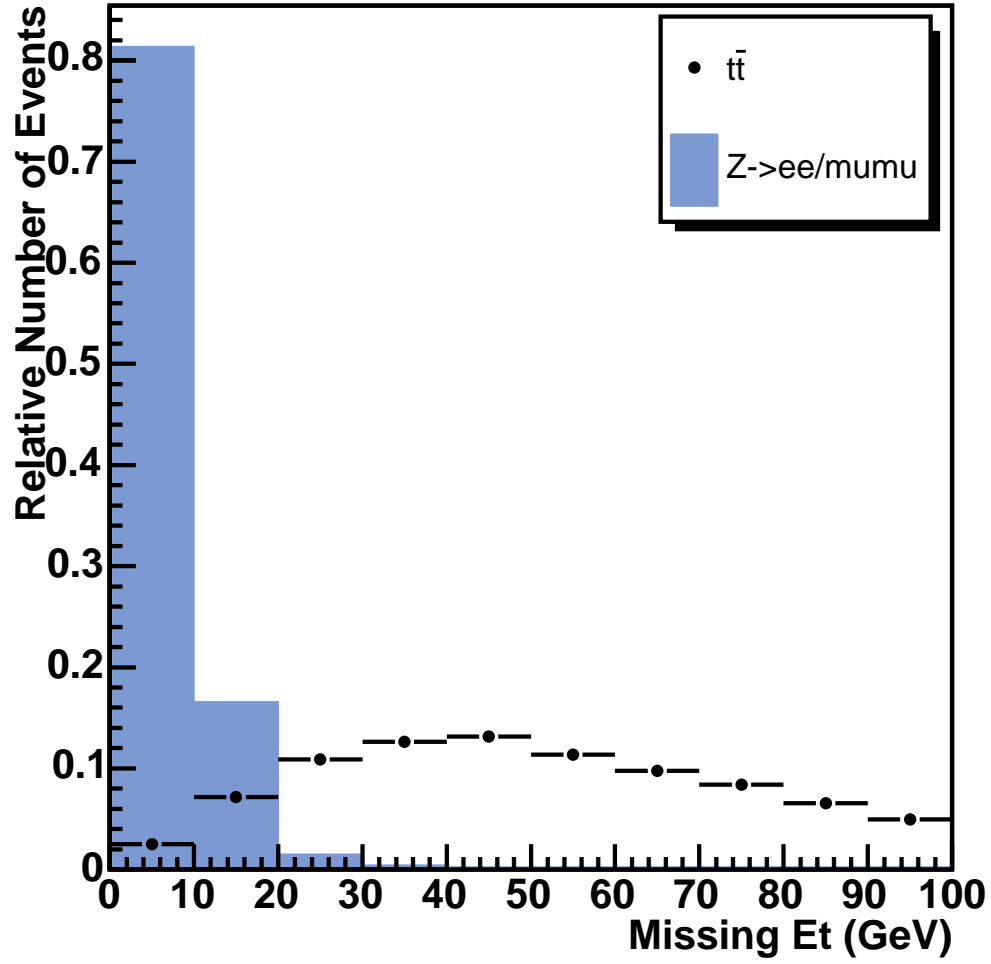
Missing Et in $t\bar{t}$ and DY MC events

Figure 3.4: Missing Transverse Energy (E_T) in simulated $Z/\gamma^* \rightarrow \ell\ell$ events and dilepton top decays.

3.4.2 $\Delta\phi$ Cuts

Escaping undetected particles can result in energy imbalances in the detector, \cancel{E}_T , but so can energy mismeasurement. For example, consider that an event contains two jets, which travel back-to-back from the center of the detector. If one of the jets falls in a gap in the calorimetry, that jet will not be reconstructed. This event will reconstruct with only one jet and a \cancel{E}_T vector along the lost jet direction, with the same E_T as the jet. This is an example of false \cancel{E}_T due to mismeasurement.

In order to combat these effects, we apply several $\Delta\phi$ cuts. We reject events with \cancel{E}_T pointing along the direction in ϕ of an object such as a lepton or jet. Specifically we reject events where

- the angle in ϕ between the primary lepton and the \cancel{E}_T is less than 5° : $\Delta\phi(lep, \cancel{E}_T) > 5^\circ$
- the track lepton points within 5° at or away from the \cancel{E}_T : $5^\circ < \Delta\phi(trklep, \cancel{E}_T) < 175^\circ$
- Any jet with $E_T > 10$ GeV which points within 25° of the \cancel{E}_T , unless $\cancel{E}_T > 50$ GeV.

The cut on the primary lepton is meant to combat a particular event type. A small fraction of $Z/\gamma^* \rightarrow \mu\mu$ events will have a muon which emits a photon via the bremsstrahlung process. The photon then produces an electron and positron. The electron/positron and muon travel in approximately the same direction. In this case, well measured tracks and an energy cluster in the electromagnetic calorimeter exist, and the object is reconstructed as an electron, with lower E_T than the original muon. This loss of energy from the escaping muon will result in \cancel{E}_T reconstructed along the direction of the reconstructed “electron”. The cut on the track lepton is designed to be more general. The track lepton p_T can be over or under measured, so a veto on \cancel{E}_T parallel or anti-parallel is necessary. Typically, the scale of jet energy mismeasurement is about 10 GeV, but since jets cover a

larger solid angle than leptons, the angular correlation between mismeasured jet energy and \cancel{E}_T is not as strong. Thus, the angular region vetoed is larger than for leptons. This requirement is not necessary for events with $\cancel{E}_T > 50$ GeV because the scale of the mis-measurement is on the order of 10 GeV.

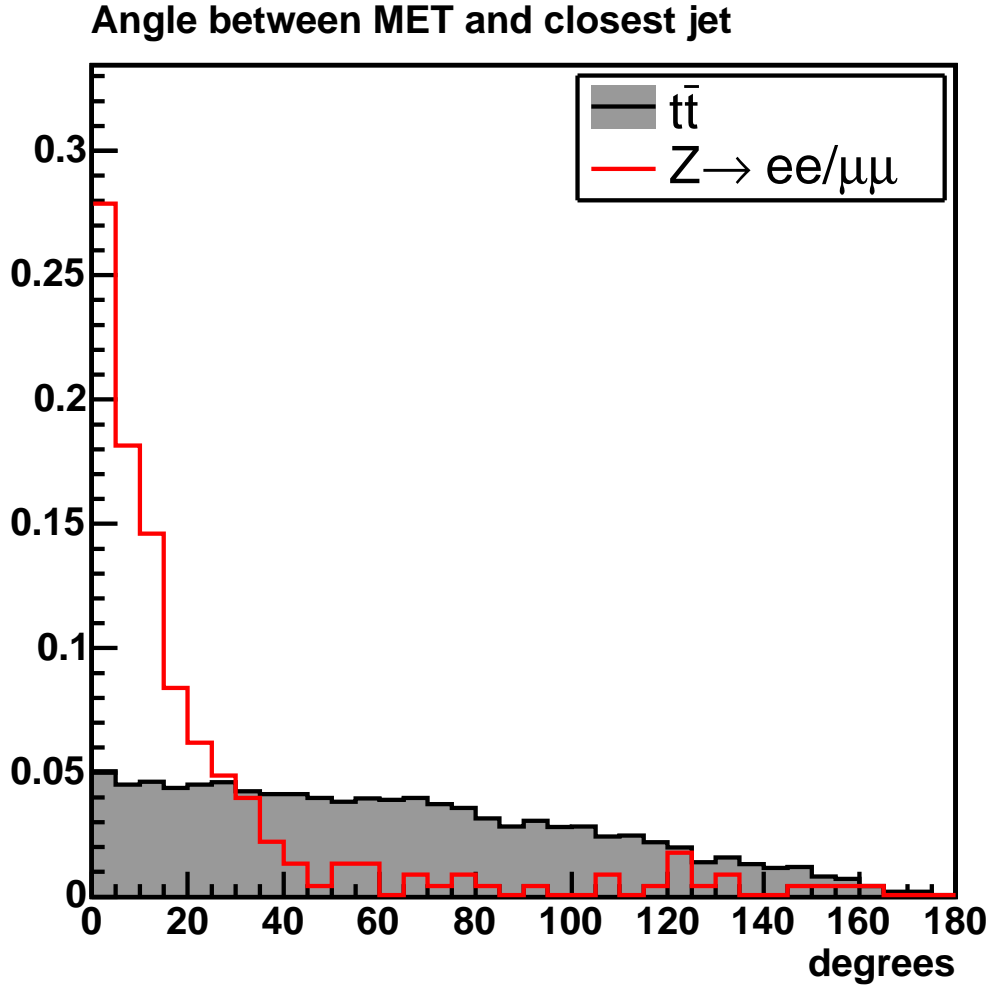


Figure 3.5: The angle between the \cancel{E}_T and the jet in simulated $t\bar{t}$ events and $Z/\gamma^* \rightarrow \ell\ell$ events.

If any event fails these cuts, the next primary lepton in the list, if any, is used to seed the event selection.

3.4.3 Drell-Yan Veto

Drell-Yan events, $Z/\gamma^* \rightarrow \ell\ell$, have a distinct invariant mass peak at ≈ 91 GeV. In order to produce at least 25 GeV of \cancel{E}_T , some p_T smearing must occur, but a large portion of events will still have a dilepton invariant mass of $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. To reduce contributions to the data sample from Drell-Yan events we increase the required \cancel{E}_T for events with lepton invariant mass $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$ to $\cancel{E}_T > 40$ GeV. If any event fails this veto, the next primary lepton is used to seed the event selection.

3.4.4 Event Classification

Events are classified by the primary lepton type, the number of jets, and the charges of the primary and track leptons. Events with opposite charge leptons and zero and one jet are expected to have very few signal events, and form a control sample. All events with the same charge leptons also form a control sample. We compare the predicted background to the observed number of events in the control regions as a way of checking the background estimation. Events with two or more jets and oppositely charged leptons form the signal region, where excesses over the predicted background events are interpreted as top quark decay in the dilepton channel.

3.5 Data Sample

Data collected by the CDF detector, selected by the triggers described in Section 2.3.2 are used as input to this analysis. We further require that all data runs pass offline quality checks, known as the “Good Run List”. This checks that all relevant detectors were in good operational order. For different primary lepton types the minimum detector requirements change. For the pretag analysis acceptance of central electrons (CEM) requires the calorimetry to be working correctly, but for plug electrons (PHX) we further require that the silicon system be functioning. For muons (CMUP and CMX) we require

lepton type	Pretag luminosity (pb^{-1})	Tagged luminosity (pb^{-1})
CEM	1097	1015
CMUP	1047	971
CMX	1027	958
PHX	1015	1015

Table 3.2: Integrated luminosity for each primary lepton type for both the tagged and pretag analysis. The uncertainty on the luminosity is 6% due to the limits of our understanding of the luminosity detectors and the $p\bar{p}$ inelastic scattering.

the calorimetry systems and the muon systems to be operating. The tagged analysis relies on the silicon system for the identification of b -jets, so the silicon system must be operating for all primary lepton types. The luminosity for both the tagged and pretagged analysis are found in Table 3.2.

Chapter 4

Pretag Top Dilepton Acceptance

The pretag acceptance represents the fraction of signal events expected to pass all selection cuts, except for the tagging requirement, and can be thought of as an efficiency which combines geometric and kinematic effects as well as other traditional efficiencies, such as identification and reconstruction efficiencies.

Chapter 3 described the event selection for the pretag analysis, which is also used in the tagged analysis. For the tagged analysis we then must account for the efficiency to find at least one tagged jet in the event. Factorizing the acceptance and efficiency into two sections this way is advantageous for several reasons. The pretag acceptance has been studied and validated previously. Keeping the event tagging efficiency as a separate measure makes it easier to compare against other analyses which also measure event tagging efficiency. Combining pretag acceptance with the event tagging efficiency would require large amounts of extra work. This chapter describes the calculation of the pretag acceptance.

4.1 Simulated $t\bar{t}$ Sample

To measure the acceptance we use the PYTHIA Monte Carlo generator with CTEQ5L parton distribution functions [36]. The top quark mass is set to $175 \text{ GeV}/c^2$. PYTHIA

uses leading-order matrix elements for top quark pair-production to produce kinematic information for final state quarks and leptons. PYTHIA outputs a list of particles and their 4-momenta, which are then passed through a full simulation of the CDF detector. The detector simulation is done with the GEANT 3 [38, 51] program. The program is given detailed information on the shape, spatial extent, and material composition of all detector sections. The program then executes a full simulation of the detector response to particles passing through the detector material. Most interactions such as ionization, showering and energy deposition is modeled using the standard GEANT algorithms. However, charged particle ionization and drift properties in the COT are parameterized and tuned to data. Additionally, the behaviour of the calorimeter showers is modeled by the use of the parameterized shower development package called GFLASH [40]. The parameters of the electromagnetic and hadronic showers are tuned to inclusive jet data. Identical event reconstruction to that used on data is used on simulated data, and data and simulation are analyzed using the same techniques and software.

4.2 Corrections to Simulation

Simulation models particle decays well, but not perfectly, and corrections need to be made to the acceptance calculated from simulation.

4.2.1 Primary Lepton Identification Scale Factor

The efficiency for identifying a primary lepton is measured in the $Z/\gamma^* \rightarrow \ell\ell$ sample. This sample is used because the $Z/\gamma^* \rightarrow \ell\ell$ decay is free from many backgrounds and is thus more pure than the top decay sample. The $Z/\gamma^* \rightarrow \ell\ell$ sample is also considerably larger.

The efficiency for identifying a primary lepton is calculated by first finding a high quality lepton, and then searching for a second, opposite sign lepton of the same flavour

Lepton type	$\epsilon(\text{data})$	$\epsilon(\text{simulation})$	data/simulation
CEM electron	0.793 ± 0.002	0.808 ± 0.001	0.981 ± 0.004
PHX electron	0.715 ± 0.003	0.764 ± 0.001	0.935 ± 0.010
CMUP muon	0.833 ± 0.004	0.900 ± 0.001	0.926 ± 0.004
CMX muon	0.904 ± 0.004	0.919 ± 0.001	0.984 ± 0.005

Table 4.1: Efficiencies for identifying the different primary lepton types in data and simulation, and the ratio is used to correct the efficiency in $t\bar{t}$ simulation.

which passes minimal selection cuts. The invariant mass of the pair must be close to the Z boson mass to reduce background contamination. The number of second leptons passing minimal selection cuts forms the denominator of the efficiency calculation. The second lepton is then tested against the primary lepton selection, and if passed, is counted toward the numerator of the efficiency. The efficiency is then the fraction of lepton like objects which is expected to be reconstructed as primary leptons.

This efficiency is calculated in both data and simulation, and the ratio of these efficiencies is called the identification scale factor. The scale factor is multiplied by the acceptance calculated in simulation, and corrects the acceptance for the imperfect modeling of the primary lepton identification in the simulation. The primary lepton scale factors are calculated separately for each primary lepton type, and are applied to the acceptance for each lepton type separately as well. The primary lepton efficiencies and scale factors are given in Table 4.1.

4.2.2 Track Lepton Scale Factors

Track leptons are required to be isolated from other activity in the tracking volume. Similarly to the primary lepton identification scale factor, we measure the ratio of the efficiency of the isolation requirement in the data to that in the simulation. The efficiency is measured using a sample of $Z/\gamma^* \rightarrow \ell\ell$ decays with opposite sign leptons, in the mass

Jet Multiplicity	data/simulation
0	1.004 ± 0.001
1	1.002 ± 0.003
≥ 2	0.965 ± 0.011

Table 4.2: Data/simulation scale factors applied to the $t\bar{t}$ simulation to correct the pretag efficiency for the isolation efficiency cut on track leptons.

range $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. We select a primary lepton and then apply all the track lepton selection cuts save the isolation on the second lepton. The isolation cut efficiency is then the ratio of the number of tracks which pass the isolation requirement, to the total number of tracks. We measure the isolation efficiency as a function of the number of jets in the event, and find the efficiency drops with jet multiplicity due to increased activity in the detector. The efficiency is $\approx 95\%$ for events with zero jets, and $\approx 90\%$ for events with two or more jets. The scale factors applied to the simulation are found in Table 4.2.

The efficiency of the impact parameter cut is also examined in data and simulation. We measure the efficiency of the track impact parameter cut efficiency, in $Z/\gamma^* \rightarrow \ell\ell$ data and simulation, and correct for the observed differences. The total observed efficiency in data is 0.909 ± 0.003 and 0.9185 ± 0.0007 in simulation. This leads to a scale factor of 0.989 ± 0.003 .

4.2.3 χ^2 probability Cut efficiency

As mentioned in Section 3.1.2 we apply a χ^2 probability requirement on muon primary leptons and all track leptons to reject kaon decays in flight. This requirement is not applied in the simulation thus the efficiency must be measured in the data and applied to the acceptance.

The χ^2 probability requirement efficiency is measured in the well-reconstructed plus

Event type	efficiency
electron+track	0.962 ± 0.001
CMUP+track	0.994 ± 0.001
CMX+track	0.951 ± 0.002

Table 4.3: Data/simulation scale factors applied to the $t\bar{t}$ simulation to correct for data/simulation discrepancies.

isolated track $Z/\gamma^* \rightarrow \ell\ell$ sample. Because we impose the cut on both primary leptons and track leptons for muon decays the cut efficiency is correlated between the two leptons. So the efficiency to apply the cut to both primary and track leptons is not equal to the product of the efficiencies on each lepton type separately. We measure the efficiencies for electron+track events separately from muon+track events. The measured efficiencies are displayed in Table 4.3.

4.2.4 Event Trigger Efficiencies

As outlined in Section 2.3.2, data used in this analysis is collected with triggers that select events with at least one high- p_T lepton candidate. The trigger thresholds are designed to be very efficient for fully-reconstructed leptons, but they are not 100% efficient. This small inefficiency is due to the fact that some offline quantities such as energy and tracking information are different than in the fast reconstruction software used by the trigger system.

The probability for a dilepton event to fire one of the high- p_T lepton triggers is higher than the single lepton trigger efficiency since each event has two chances to fire one of the triggers. Since the second lepton is not fully reconstructed the event trigger efficiency is not simply a combination of the single-lepton trigger efficiencies. We calculate a per-event trigger efficiency by creating a list of triggerable primary leptons ordered in a way that is consistent with our acceptance code, CEM, CMUP, CMX and PHX. For each primary

	CEM	CMUP	CMX	PHX
single lepton	0.971	0.918	0.910	0.918
$t\bar{t}$	0.975	0.916	0.937	0.918
WW	0.976	0.917	0.937	0.935
WZ	0.981	0.921	0.939	0.925
ZZ	0.983	0.926	0.941	0.924

Table 4.4: Single lepton and full per-event trigger efficiencies for the four primary lepton types. Numbers in bold are inputs to the cross-section measurement. Uncertainties are negligible compared to other uncertainties on the acceptance.

lepton, we generate a random number between zero and one and check it against the relevant single-lepton efficiency to decide if that primary lepton was “passed” by the trigger. If none of the primary leptons pass the trigger, the event is rejected; otherwise, it is accepted. We repeat this procedure for every event and obtain an average trigger efficiency for a sample and primary lepton type. Table 4.4 shows the trigger efficiency per sample derived by this method. For a given lepton type, the efficiencies are very similar for the different physics processes, so the $t\bar{t}$ value is used.

4.3 Acceptance for Pretag Top Dilepton Events

The acceptance for pretag top dilepton events using the lepton+isolated track selection is measured in a PYTHIA Monte Carlo sample of about 4 million inclusive $t\bar{t}$ decays, with the top quark mass set at $175\text{GeV}/c^2$. Table 4.5 shows the measured acceptance for events with oppositely charged leptons with all jet multiplicities and all correction factors described above included. We measure an acceptance for the signal region of 0.845%. The 0.004% statistical uncertainty is negligible.

	0 jets	1 jet	≥ 2 jets
CEM	0.009 ± 0.000	0.134 ± 0.002	0.472 ± 0.003
CMUP	0.004 ± 0.000	0.056 ± 0.001	0.196 ± 0.002
CMX	0.002 ± 0.000	0.023 ± 0.001	0.073 ± 0.001
PHX	0.002 ± 0.000	0.029 ± 0.001	0.104 ± 0.001
Total	0.017 ± 0.000	0.242 ± 0.003	0.845 ± 0.004

Table 4.5: Top dilepton signal acceptance, in percent, for $m_t = 175\text{GeV}/c^2$, measured in PYTHIA, corrected for all known discrepancies between data simulation. Uncertainty is statistical only.

4.4 Systematic Uncertainties on the Acceptance

Systematic uncertainties reflect limits of our understanding or ability to estimate the $t\bar{t}$ acceptance. They can arise from limitations of our theoretical knowledge of the production process, or limits on our ability to identify the final state objects experimentally. All systematic uncertainties on the signal acceptance are discussed individually below and summarized in Table 4.6.

4.4.1 Primary Lepton Identification Efficiency

Primary lepton identification efficiencies and scale factors are derived using the large, high-purity $Z/\gamma^* \rightarrow \ell\ell$ sample with mostly zero-jet events. The primary lepton identification systematic uncertainty quantifies the discrepancy associated with using factors derived from a quiet event environment in a “busy” one, with at least two jets.

To quantify the systematic uncertainty the primary lepton identification scale factor is measured as a function of the distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the lepton candidate and the nearest jet. This scale factor is re-weighted with the distribution of distances between the primary lepton and the closest jet. The systematic uncertainty is taken

Source	uncertainty assigned
primary lepton identification efficiency	1.1%
track lepton identification efficiency	1.1%
jet energy scale	1.3%
initial state QCD radiation	1.6%
final state QCD radiation	0.5%
parton density function	0.5%
Monte Carlo generator	1.5%

Table 4.6: Summary table of systematic uncertainties on the signal acceptance. The total systematic uncertainty is 3%.

as the larger of the discrepancy between the inclusive and re-weighted scale factors and the statistical uncertainty on the re-weighted scale factor. In most cases the statistical uncertainty is larger than the discrepancy.

The total systematic uncertainty is the weighted average of the uncertainties on the individual lepton types, where the weights are the top dilepton acceptances for each lepton category. The resulting uncertainty is 1.1%.

4.4.2 Track Lepton Identification Efficiency

This uncertainty is based on the extent of our understanding of how well the simulation models the track isolation requirement in an environment with added activity with many jets. Section 4.2.2 explains the derivation of a scale factor correcting the track isolation efficiency in simulation to that observed in data and we take its associated statistical uncertainty in the two or more jet bin, 1.1%, as the uncertainty on track lepton identification.

4.4.3 Jet energy scale

To estimate the uncertainty associated with the jet energy scale [35], we remeasure the signal acceptance in PYTHIA $t\bar{t}$ simulation with the jet energy corrections varied within their one standard deviation uncertainties. For events with at least two jets, the measured pretag acceptance increases by 1.2% for the plus one standard deviation variation and decreases by 1.4% for the minus one variation. We take half the full difference between these, 1.3%, as our systematic uncertainty.

4.4.4 Initial and Final State Radiation (ISR/FSR) modeling in simulation

Additional jets can be produced in association with the $t\bar{t}$ pair through QCD radiation of one or more gluons from the initial or final state particles. We can measure the dependence of the acceptance on QCD radiation by comparing the central value of the acceptance to acceptances calculated in PYTHIA $t\bar{t}$ samples identical to those used to calculate the central value, except that the rate of initial and final state radiation has been varied by one standard deviation. The acceptance increases in the sample with more initial-state radiation, and decreases in the sample with less. We take half the full difference, 1.6%, as the systematic uncertainty. The results with the more- and less- final-state radiation samples are less conclusive, as the measured acceptances do not differ significantly from the nominal value. We take the larger of the two observed discrepancies, 0.5%, as the systematic uncertainty. It does not hurt to be conservative here since this contribution has a negligible impact on the total uncertainty.

4.4.5 Parton Distribution Functions

Parton Distribution functions (PDFs) describe the fraction of the proton momentum carried by each of the constituent partons. Because the PDFs affect the kinematics of

the $t\bar{t}$ decay products, the PDFs also affect the acceptance.

To evaluate the systematic uncertainty from the PDFs we examine both the known uncertainty on the PDFs, as well as consider the differences between PDF sets and uncertainty on the value of α_s . To evaluate the uncertainty arising from the PDFs used to generate the simulation used to calculate the top dilepton acceptance, we recalculate the acceptance with events re-weighted using PDFs which differ from the nominal in a specific way. We identify the incoming partons from the generator level truth information in the simulation, and the probability of their momenta is varied. The event weight is the ratio of the product of the varied to nominal probabilities:

$$\text{weight} = \frac{p_1(\text{variation}) \cdot p_2(\text{variation})}{p_1(\text{nominal}) \cdot p_2(\text{nominal})}$$

We repeat this process for each variation in the PDFs and record the change in the acceptance.

For the CTEQ PDFs, the quadratic sums of the observed differences in the acceptance are +0.3% and -0.7% from the nominal. We also compare the acceptance as calculated using the nominal CTEQ PDFs to the acceptance calculated using PDFs determined independently by a second collaboration (MRST) [41]. The difference, 0.2%, is smaller than the uncertainties derived using only CTEQ PDFs and therefore ignored. The uncertainty obtained by varying α_s is 0.1%, which is added in quadrature with the CTEQ uncertainties for a total uncertainty of +0.3% and -0.7%. Symmetrizing this gives a systematic uncertainty of 0.5%.

4.4.6 Monte Carlo Generator

To check for dependence of the measured acceptance on the simulation event generator, we remeasured the $t\bar{t}$ acceptance using the HERWIG Monte Carlo with $m_t = 175$ GeV and compared it to the nominal acceptance using PYTHIA simulation. We compare the two acceptance calculations relative to the total number of generated dilepton events to

exclude the effect of the different $W \rightarrow \ell\nu$ branching ratios used in the two generators ($1/9 = 0.111$ in Herwig and 0.108, the LEP measured value, in PYTHIA). The resulting difference of 1.5% is taken as a systematic uncertainty.

Chapter 5

b-Jet Identification

The reconstruction of tracks displaced from the primary interaction vertex, originating from B hadron decay, which infer the existence of a b quark, is called “ b -tagging”. The SECVTX algorithm identifies b -jet candidates based on the determination of the primary event vertex and the reconstruction of additional secondary vertices using displaced tracks associated with jets [42]. If such a secondary vertex is displaced far enough away from the primary vertex, then the jet is identified (“tagged”) as a b -jet candidate.

As discussed in Chapter 3, this thesis describes the measurement of the top dilepton cross-section with the identification of at least one b tagged jet. This analysis is closely related to the “pretag” analysis as these two analyses share the pretag acceptance calculation, equation 1.1, presented here again.

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{\epsilon_{tag} \times \mathcal{A} \times \int \mathcal{L} dt} \quad (5.1)$$

This chapter describes the event tagging efficiency (ϵ_{tag}), which is unique to the tagged analysis. The requirement of a tagged b -jet increases the purity of the sample from $\sim 1 : 1$ for the pretag measurement to about $\sim 5 : 1$ for the tagged measurement.

There are two versions of the SECVTX algorithm, one which imposes stricter quality cuts resulting in a more pure sample with a lower rate of mis-identified jets, at the cost of lower tagging efficiency. The other version of the algorithm uses less stringent quality

cuts resulting in higher tagging efficiency at a cost of lower purity in the tagged sample. This analysis uses the higher efficiency, lower purity version of the SECVTX algorithm. This choice is motivated by the dilepton channel's relatively small contributions from background events, and the fact that it suffers from a low branching ratio to two leptons, so the significance of the measurement is limited mainly by the size of the sample.

5.1 SecVtx Algorithm

The distance a particle travels in the detector frame of reference is a product of the particle's velocity ($\beta \cdot c$) and the time dilated lifetime of the particle ($\tau \cdot \gamma$).

$$distance = \gamma \cdot \tau \cdot \beta \cdot c$$

At CDF the average transverse momentum of *b*-jets from $t\bar{t}$ decays is ≈ 65 GeV, so we have $\beta \sim 1$ and $\gamma \sim 15$. The average lifetime of a *B* hadron is 1.5×10^{-12} s. So we see the average *B* hadron can travel ≈ 7.5 mm before decaying. The perpendicular point of closest approach to the primary vertex is called the impact parameter, (d_0). The average value of d_0 is $c \cdot \tau$, and for *B* hadrons this value is about $450 \mu\text{m}$. CDF has developed an algorithm, SECVTX, which takes advantage of this displaced secondary vertex of the *B* hadron decay to identify *b*-jets [42]. The SECVTX algorithm first obtains a primary interaction vertex for each event, for which the resolution is $\approx 10 - 20 \mu\text{m}$. The SECVTX algorithm then considers tracks in the event by examining the tracks associated with a jet, and applying basic quality criteria on the number of hits in the SVX, impact parameter limits, and track helix χ^2 per degrees of freedom. The algorithm then attempts to resolve a secondary vertex which is displaced from the primary vertex, by vertexing tracks with large impact parameter significance, see Figure 5.1.

The SECVTX algorithm resolves a vertex using a two-pass algorithm. The first pass of the algorithm calculates a “seed” vertex from the two most displaced tracks, then adds tracks to the seed vertex. Pairs of tracks with invariant masses consistent with the K^0

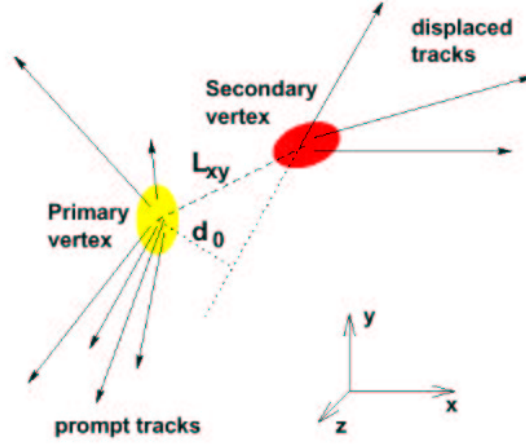


Figure 5.1: Schematic diagram of a displaced vertex.

and Λ mass are removed. It attempts to find a vertex with at least three tracks, each with impact parameter significance, $d_0/\sigma_{d_0} > 2.5$ and transverse momentum $p_T > 0.5$ GeV/ c . At least one of the three tracks must have $p_T > 1.0$ GeV/ c . The final vertex must have a χ^2 of less than 120 to be accepted. If no such vertex is found, the next seed vertex is attempted, until a vertex is resolved, or the seed list is exhausted.

If no three-track (or “first-pass”) vertex is found, the algorithm attempts to find a vertex using only two tracks, where the quality requirements of those two tracks is increased. Again, pairs of tracks whose invariant mass is consistent with the K^0 and Λ masses are removed. Tracks included in the two-track (or “second-pass”) vertex must have $d_0/\sigma_{d_0} > 3.0$ and $p_T > 1.0$ GeV/ c , with at least one track with $p_T > 1.5$ GeV/ c . The χ^2 of the vertex must be less than 2000.

With a secondary vertex in hand, SECVTX calculates the distance between the primary and secondary vertices in the plane perpendicular to the beamline. This vector is then projected on the jet axis

$$L_{xy} = (r_{PV} - r_{SV}) \cdot \hat{p}_{jet}$$

where r_{PV} is the position of the primary vertex, r_{SV} is the position of the secondary vertex, \hat{p}_{jet} is the jet direction, and L_{xy} is the two-dimensional decay length along the

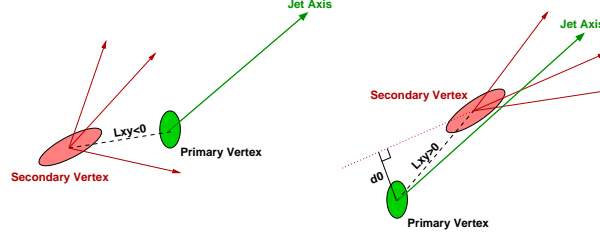


Figure 5.2: Drawing showing fake ($L_{xy} < 0$) and true ($L_{xy} > 0$) reconstructed secondary vertices.

jet axis. SECVTX defines a “displaced” (or “tagged”) vertex as one with $|L_{xy}/\sigma_{L_{xy}}| > 3.0$. The sign of L_{xy} determines the type of tag; $L_{xy} > 0$ is called a *positive* tag, and $L_{xy} < 0$ is called a *negative* tag. This is illustrated in Figure 5.2. Long lived quarks enhance the positive tag distribution. The negative tag distribution is interpreted as the distribution of “mistags”, or tagging mistakes due to resolution effects, and the jets originating in the detector material.

5.2 Event Tagging Efficiency

The efficiency of the SECVTX tagger to tag an individual jet in a top quark decay event is shown in Figure 5.3 as a function of the jet E_T and η .

The event tagging efficiency is the efficiency for tagging at least one of the two *b*-jets in a $t\bar{t}$ dilepton decay event using the SECVTX tagger. To find the event tagging efficiency we use $t\bar{t}$ PYTHIA Monte Carlo with $m_{top} = 175 \text{ GeV}/c^2$. The event tagging efficiency is measured using two independent methods, the simulation method and the Factorization method. As is traditional, we use the result from the factorization method as the event tagging efficiency used in this analysis. Corrections to the event tagging efficiency are made to account for the reconstruction of a jet corresponding to a *B* hadron, and the efficiency to tag light quark jets as heavy flavour jets.

5.2.1 Simulation Method

The event tagging efficiency can be derived by using simulated $t\bar{t}$ events. Events selected as pretag candidate events with opposite-sign leptons, and at least two jets, form the denominator of the efficiency. To account for the difference in tagging efficiency between the simulation and data, we do a random rejection technique. The data/simulation scale factor is found to be $\epsilon_{SF} = 0.94 \pm 0.06$. Events are examined for the presence of a B hadron in the Monte Carlo generator information. The B hadron must be matched in a cone of $\Delta R < 0.4$ to a jet accepted in the analysis, eg. has $E_T > 20$ GeV and $|\eta| < 2.0$. Jets with these criteria and which are positively tagged by the SECVTX algorithm as b -jets are then subjected to the random rejection. A random number between 0 and 1 is generated, and if the random number is lower than $1 - \epsilon_{SF}$ the jet is rejected. The numerator of the efficiency is then all events with at least one jet with a positive SECVTX tag, which also survives the random-rejection technique.

The systematic uncertainty on the event tagging efficiency using the simulation method is due to the systematic error on the SECVTX scale factor used. We obtain the systematic on the event tagging efficiency by re-calculating the event tagging efficiency with the SECVTX scale factor varied by \pm one standard deviation and take half of the difference between these two results as the systematic uncertainty. We obtain a systematic uncertainty of 4.1% using this method.

The event tagging efficiency using the simulation method is then found to be $\epsilon_{tag} = 0.673 \pm 0.028$.

5.2.2 Factorization Method

The event tagging efficiency can be factorized as the sum of the single jet tagging efficiency, ϵ_b' , times the fraction of events with one and two taggable b -jets, denoted F_{1b} and F_{2b} . The taggable jet fractions describe the fraction of events with one (two) jets which

come from the hadronization of a *b* quark, and are able to be tagged. The denominator is defined as the number of pretag candidate events in runs with the silicon systems functioning properly. The numerator is the subset of events with one (two) jets matched to a *B* hadron at generator level and that have at least two tracks which pass the SECVTX track quality cuts. The jets are also required to fall within the fiducial region of the SECVTX algorithm, but all jets considered in this analysis pass that requirement. The single jet tagging efficiency, $\epsilon_{b'}$ is the fraction of taggable jets with a positive SECVTX tag, ϵ_b , multiplied by the SECVTX data/simulation scale factor, ϵ_{SF}

$$\epsilon_{tag} = F_{1b} \cdot \epsilon_{b'} + 2F_{2b} \cdot \epsilon_{b'}(1 - \epsilon_{b'}) + F_{2b} \cdot \epsilon_{b'}^2 \quad (5.2)$$

$$\epsilon_{b'} = \epsilon_b \times \epsilon_{SF}$$

The event tagging efficiency using the factorization method is found to be $\epsilon_{tag} = 0.668 \pm 0.037$ including the systematic uncertainty from the systematic uncertainty on the data/simulation scale factor. Table 5.1 shows the inputs used in the calculation of the event tagging efficiency.

Efficiency of matching *B* hadrons to jets

The event tagging efficiency needs to account for the fraction of *B* hadrons that are not matched to reconstructed jets. We measure the efficiency for matching a *B* hadron to a reconstructed jet in PYTHIA simulated data, with $m_{top} = 175 \text{ GeV}/c^2$. We find the efficiency to match a *B* hadron to a reconstructed jet to be $\epsilon_{bmatch} = 0.9889 \pm 0.0004$, where the uncertainty is statistical only. We multiply the per jet tagging efficiency by the efficiency to match a *B* hadron to a jet to account for the inefficiency to match a reconstructed jet to a *B* hadron.

Efficiency to tag a light quark jet

The per jet tagging efficiency needs to account for the efficiency to tag a non-*b* (light quark) in top events. We interpret the negative tags on jets in PYTHIA top simulation

F_{1b}	0.321 ± 0.003
F_{2b}	0.611 ± 0.003
ϵ_b	0.591 ± 0.002
ϵ_{SF}	$0.94 \pm 0.01(\text{stat.}) \pm 0.06(\text{syst.})$
$\epsilon_{lightquark}$	$0.013 \pm 0.001 \text{ (stat)}$
ϵ_{bmatch}	$0.9889 \pm 0.0004(\text{stat})$
ϵ_{tag}	0.669 ± 0.037

Table 5.1: Inputs and results of calculation of event tagging efficiency using the factorization method.

events as tags on light quark jets. We find the probability of tagging a light quark jet in top events of $\epsilon_{lightquark} = 0.013 \pm 0.001$. We account for the tagging of light quark jets by dividing the per jet tagging efficiency by $(1 - \epsilon_{lightquark})$.

5.2.3 Final Event Tagging efficiency

The per jet tagging efficiency including the efficiency to match *B* hadrons to jets, and to tag light quark jets in top events can be expressed in analytic form:

$$\epsilon_{tag} = F_{1b} \cdot \epsilon_b I + 2F_{2b} \cdot \epsilon_b I (1 - \epsilon_b I) + F_{2b} \cdot \epsilon_b I^2$$

$$\epsilon_b I = \frac{\epsilon_b \times \epsilon_{SF} \times \epsilon_{bmatch}}{(1 - \epsilon_{lightquark})}$$

Table 5.1 gives a summary of the inputs used to calculate the final event tagging efficiency. Using the above per jet tagging efficiency in equation 5.2, we obtain an event tagging efficiency of 0.669 ± 0.037 , giving a 5.5% systematic uncertainty due to event tagging.

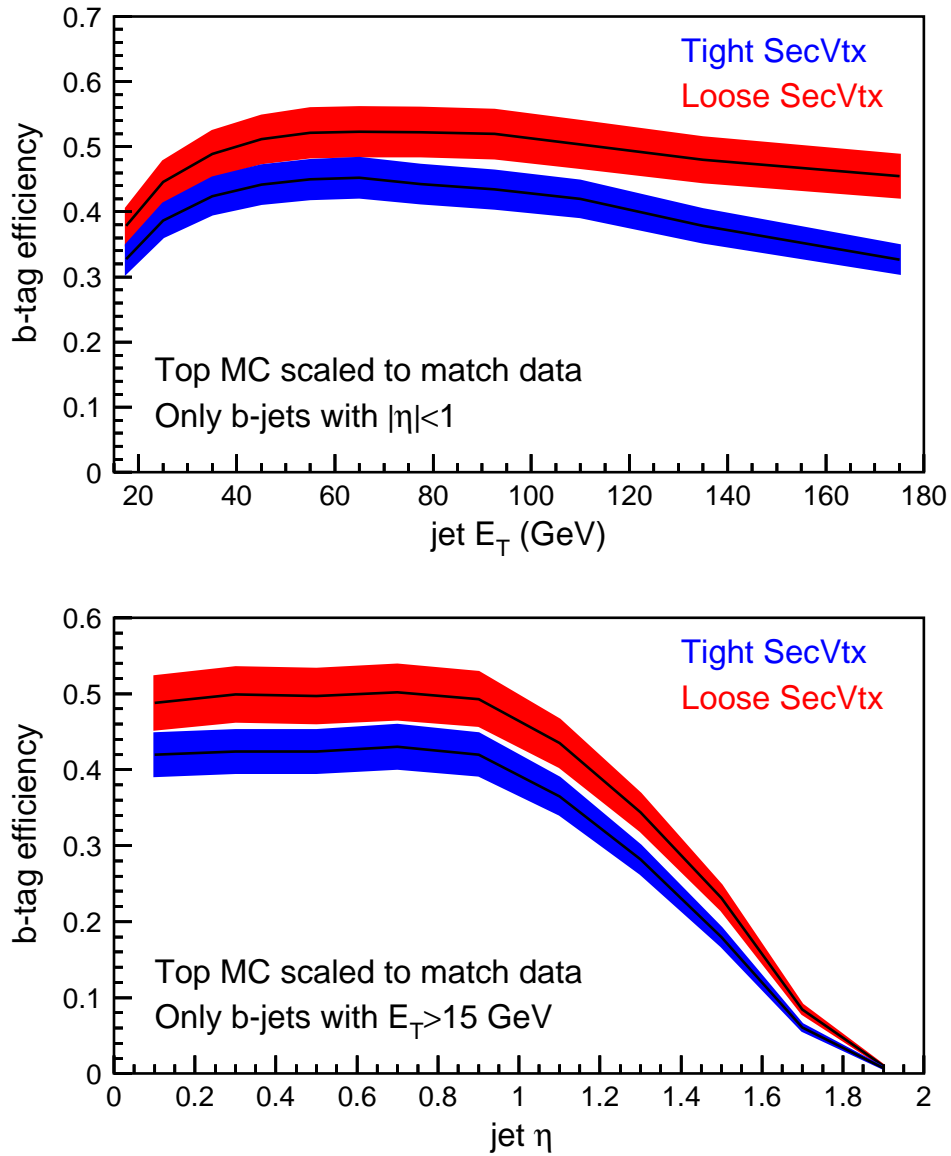


Figure 5.3: The tagging efficiency for the SECVTX tagger for the two SECVTX tunings, as a function of the jet E_T and η . These are based on simulated datasets, but corrected for discrepancies between the simulation and data. The blue (lower efficiency) line is for the higher purity, lower efficiency version of the tagging algorithm. The red line is for the lower purity, higher efficiency version of the tagging algorithm, which is used in this analysis.

Chapter 6

Pretag Backgrounds

The pretag background processes are of two classes. The first class is physics processes that have the same event signature as $t\bar{t}$ decay in the dilepton channel. These are processes such as diboson decays and $Z/\gamma^* \rightarrow \tau\tau$ processes. These processes need to have enough extra jets produced to substitute for b -jets, but otherwise the event signatures are the same as for the $t\bar{t}$ dilepton process. The second class of backgrounds are those that require some mis-measurement to mimic the $t\bar{t}$ signal. Non-negligible processes included in this analysis are $W + \text{jets}$ processes, where one of the jets is misidentified as a second lepton, and Drell-Yan ($Z/\gamma^* \rightarrow \ell\ell$) events where \cancel{E}_T is produced by the mis-measurement of objects in the event.

The background estimate for the tagged analysis, which is the main result presented in this thesis uses the pretag background estimates to normalize the tagged background estimates. As such, it is necessary to discuss briefly the method of estimating the pretag backgrounds but it should be noted that these are not the primary focus of this thesis.

6.1 Diboson and $Z/\gamma^* \rightarrow \tau\tau$ Processes

The final states of both the diboson (WW , WZ , and ZZ , see Figure 6.1), and $Z/\gamma^* \rightarrow \tau\tau$ processes can have two, oppositely charged leptons, missing transverse energy that is

	0 jets	1 jet	≥ 2 jets
Data	0.882 ± 0.002	0.102 ± 0.002	0.0153 ± 0.0008
simulation	0.8771 ± 0.0006	0.1086 ± 0.0005	0.00142 ± 0.0002
scale factor	1.006 ± 0.002	0.940 ± 0.018	1.08 ± 0.06

Table 6.1: Fraction of events with a given jet multiplicity and scale factor correcting simulation to the level of data.

attributed to the escaping neutrinos and several jets. Even though the production cross-sections for the diboson processes are small, because the event signature contains two true leptons and \cancel{E}_T from escaping neutrinos, diboson and $Z/\gamma^* \rightarrow \tau\tau$ processes still contribute to the $t\bar{t}$ signal region. Luckily because the event signatures contain objects which are well-modeled by our simulation, we can use it to estimate their contribution to the signal region.

We measure the acceptance for diboson and $Z/\gamma^* \rightarrow \tau\tau$ processes using the PYTHIA event generator and a detector simulation based on GEANT and GFLASH [40] of ≈ 2.4 million events. We correct the acceptance measured in simulation for all known discrepancies between data and simulation. These corrections are the same as the correction factors in the $t\bar{t}$ acceptance, except for corrections to the jet distribution. The event generator PYTHIA does not correctly model the number of non- b -jets in an event, and we apply an additional correction factor. We derive the correction factor by comparing $Z +$ jets data to simulation. We select events with two oppositely charged, fully reconstructed electrons or muons, with an invariant mass in the interval $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. We find the fraction of events with zero, one and two-or-more jets, and derive a scale factor (the ratio of the fraction in data to the fraction in simulation) for each jet multiplicity. The resulting scale factors are shown in Table 6.1. These scale factors do not apply to $t\bar{t}$ decays because few events have additional non- b -jets.

We use the theoretical cross-sections given in [43], which are 13.5 pb for WW , 4.0

pb for WZ , and 1.6 pb for ZZ . With these cross-sections, along with the acceptances described above we estimate the total diboson contribution in our sample by:

$$N_{diboson} = \mathcal{A} \times \sigma_{diboson} \times \int \mathcal{L} dt$$

Because there is little contribution from $t\bar{t}$ decays with zero and one jets we estimate the background contributions in these jet bins, and compare to the observed number of events as a validation of the background estimate.

6.1.1 Systematic Uncertainties

Primary and Track Lepton Identification

As in the $t\bar{t}$ dilepton acceptance, the acceptance for diboson and $Z/\gamma^* \rightarrow \tau\tau$ events depends on correctly identifying leptons in the simulation in the same way as in data. A systematic uncertainty of 1.1% is applied to the acceptance from uncertainty in the correction factor.

Jet energy scale

The diboson and $Z/\gamma^* \rightarrow \tau\tau$ acceptance also depends on the jet energy scale. To estimate this uncertainty, we use the same method as employed for the $t\bar{t}$ acceptance, but this time in a sample of WW events generated by PYTHIA . We vary the jet energy scale by ± 1 standard deviation, and take as the uncertainty one half the difference in the acceptance. We measure a systematic uncertainty of 5.8%. The effect of the jet energy scale is larger for diboson acceptance than for the signal acceptance. This is due to the E_T distribution of jets selected in this sample which are produced through QCD radiation, which falls exponentially. The spectra of selected b -jets from the $t\bar{t}$ signal decay rises slowly, because these jets get a boost from the large mass of the top quark. As a result, diboson events are more sensitive to changes in the jet energy than $t\bar{t}$ events.

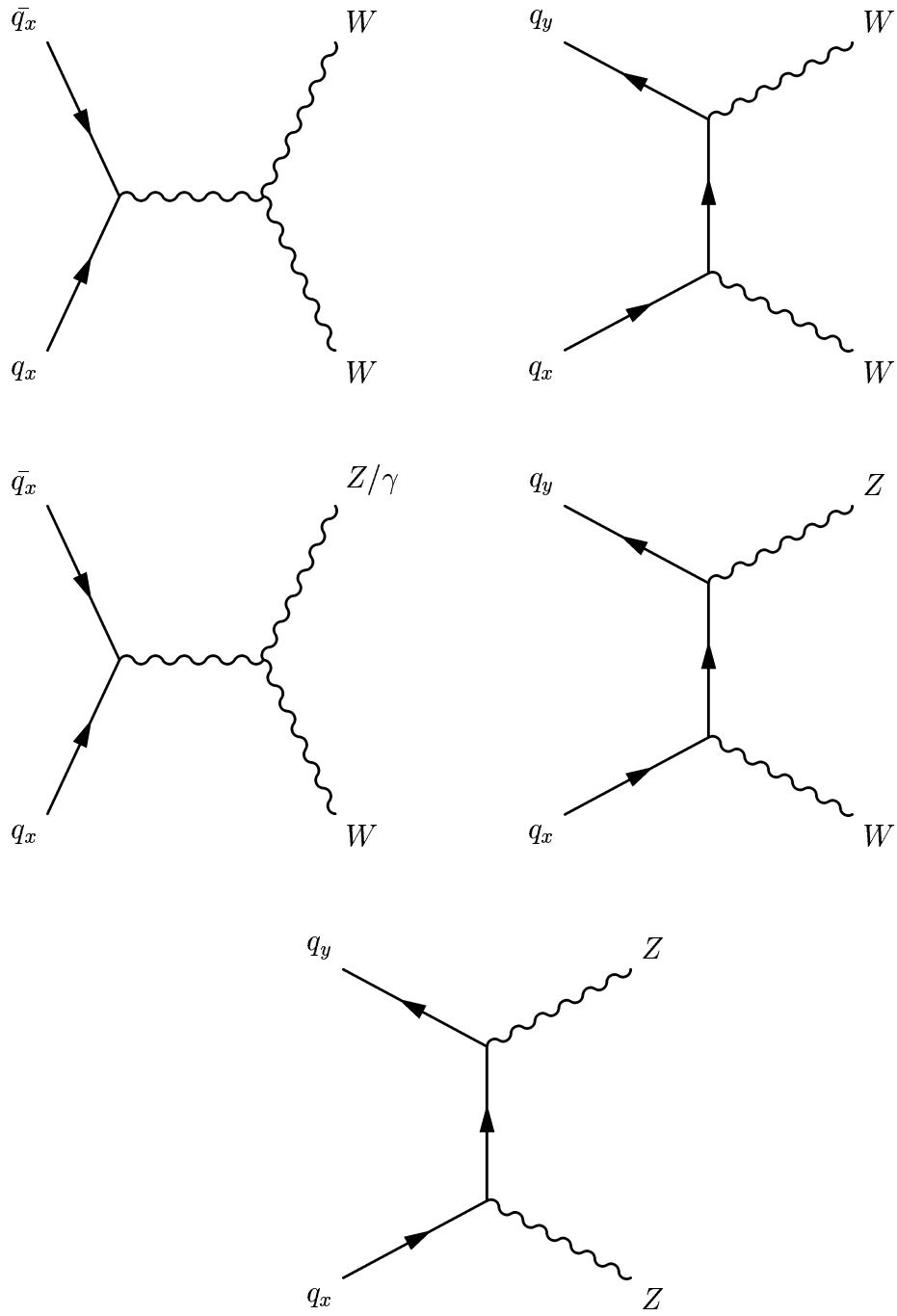


Figure 6.1: Diboson decays.

Jet Multiplicity Correction

The jet multiplicity correction factor (1.08 ± 0.06) has a non-negligible statistical uncertainty of 5.5%. This is taken as a systematic on the diboson and $Z/\gamma^* \rightarrow \tau\tau$ estimate acceptance.

6.2 $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$

The Drell-Yan processes, $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$, represent significant sources of background events because of their large cross-section and the fact that they contain two real leptons in the final state. Because there are no neutrinos in the final state of $Z/\gamma^* \rightarrow ee$ or $Z/\gamma^* \rightarrow \mu\mu$, such events pass our selection only when they have false \cancel{E}_T due to the mis-measurement of one or more objects in the event. For example, if an electron goes into a crack in the calorimeter, not all of its transverse energy will be counted in the \cancel{E}_T calculation. Another example is a muon which emits a bremsstrahlung photon may appear to be a perfect electron of much lower energy than the original muon. In this situation, the muon is not reconstructed, and the p_T of the muon is omitted from the \cancel{E}_T calculation.

These sources of spurious \cancel{E}_T are mitigated in two ways. First, the \cancel{E}_T corrections account for known sources of false \cancel{E}_T (See Section 3.3 for a description of the corrections). We require that no physics object in the event, primary or track lepton, or jet – point at the \cancel{E}_T . Additionally, for the $t\bar{t}$ cross-section measurement itself we use only events with two or more jets, further reducing the proportion of the candidates attributable to Drell-Yan events.

The background from Drell-Yan events is difficult to estimate because false \cancel{E}_T is generated by mis-measurement of some object in the event. This is not accurately modeled by the simulation. It is difficult to disentangle actual Drell-Yan events from other sources of lepton pairs (such as diboson production or $t\bar{t}$ processes) in data, so it would be

preferable to use simulation to predict the number of events expected from these sources, as is done for the diboson and $Z/\gamma^* \rightarrow \tau\tau$ backgrounds.

We estimate the background contribution from $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events by using a hybrid method, using both data and simulation. The data acts as a normalization of the overall estimate, and the simulation is used to divide the estimate between different jet multiplicities and projection outside of the Z mass window. We use simulation to divide the estimate across jet multiplicities because of the small number of data counts. The hybrid method's formula can be written analytically as:

$$N_{\text{DY}}^i = N_{\text{out}}^i + N_{\text{in}}^i \quad (6.1)$$

$$N_{\text{out}}^i = ((n_{25} - n_{25}^{\text{feeddown}}) \times (f'_i \cdot C_{\text{Zmass}})) \times R_i \quad (6.2)$$

$$N_{\text{in}}^i = ((n_{40} - n_{40}^{\text{feeddown}}) \times (f'_i \cdot C_{\text{Zmass}})) \quad (6.3)$$

where i represents the jet multiplicity, $n_{\{25,40\}}$ are the number of observed data events in the normalization sample with $E_T > 25$ and 40 GeV and $n_{\{25,40\}}^{\text{feeddown}}$ are the number of predicted events from other sources in the normalization sample. C_{Zmass} is a correction factor for the number of data events found in data versus the number found in simulation. f'_i is a correction to the simulation for the jet multiplicity observed in data. Each of these terms will be discussed in detail. Due to the low number of events observed in data (N_{out}^i and N_{in}^i), we add events passing the selection both inside and outside the Z resonance window to reduce the statistical uncertainty. We then estimate how many events from Drell-Yan backgrounds are inside (outside) the Z resonance window by subtracting the non-Drell-Yan contributions from data, and multiplying by the fraction of observed events expected inside (outside) of the window.

6.2.1 Normalizations from Data, $n_{\{25,40\}}$

For the data normalization we use the usual selection cuts with some modifications to enhance the Drell-Yan component. Two normalization variables are required, which corre-

spond to the separate \cancel{E}_T thresholds applied in the standard event selection for events inside and outside of the Z resonance mass window $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. To find $n_{\{25,40\}}$ we require that the lepton invariant mass falls in the window $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$, and no jet multiplicity requirement is applied. Because there are so few $Z/\gamma^* \rightarrow \mu\mu$ events, we collect them in one category, but $Z/\gamma^* \rightarrow ee$ events are separated into central and plug electron types.

6.2.2 Non- $Z/\gamma^* \rightarrow ee/\mu\mu$ contributions, $n_{\{25,40\}}^{\text{feeddown}}$

In the data, events from other sources such as $t\bar{t}$, dibosons, $Z/\gamma^* \rightarrow \tau\tau$, and events with a misidentified lepton can pass the selection for Drell-Yan events. These events artificially inflate this background estimate unless subtracted from n_{25} and n_{40} . We call the event counts contributing to n_{25} and n_{40} from non-Drell-Yan sources “feeddown”, n_{25}^{feeddown} and n_{40}^{feeddown} . We use the acceptance derived from simulation, theoretical cross-sections, and sample luminosity to predict the number of events from sources other than $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ that will pass that n_{25} and n_{40} selection, except for the contribution from events with a fake lepton. The fake lepton contribution is calculated in the same manner for the whole analysis, see Section 6.3, except that we use W boson events simulated with the PYTHIA event generator to find and apply the efficiency of the invariant mass and $\cancel{E}_T > 40$ criteria (8% with and 21% without the $\cancel{E}_T > 40$ requirement).

The feeddown calculation uses as a starting point the theoretical prediction of the $t\bar{t}$ cross-section corresponding to the top mass of $175 \text{ GeV}/c^2$, $\sigma_{t\bar{t}} = 6.7 \text{ pb}$. Since it is not appropriate for the $t\bar{t}$ cross-section measurement to depend on the $t\bar{t}$ theoretical cross-section, we recalculate the feeddown using the measured cross-section, and then recalculate the cross-section. This process is repeated until changes to the cross-section are undetectable ($\Delta\sigma_{\text{measured}} < 0.01 \text{ pb}$). The variations typically converge after the second iteration and total changes to the measured cross-section are only a few tenths of a picobarn. All of the data inputs to the Drell-Yan estimate, including the feeddown,

	Raw data counts		Initial feeddown		Final feeddown	
	n_{25}	n_{40}	$n_{25}^{feeddown}$	$n_{40}^{feeddown}$	$n_{25}^{feeddown}$	$n_{40}^{feeddown}$
CEM	148	54	71.	38.	75.	41.
CMUP+CMX	107	42	52.	27.	55.	29.
PHX	53	19	30.	14.	30.	15.

Table 6.2: Counts from the data used to normalize the Drell-Yan estimate, and the expected contribution from other processes (“feeddown”).

are collected in Table 6.2.

6.2.3 Ratios from Simulation, R_i

As is shown in equations 6.1, R_i is the ratio of the number of events inside the Z mass window, to the number of events outside the window, as a function of the number of jets in the event. We derive R_i from simulation of $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events using the PYTHIA event generator, requiring events to pass all lepton identification and $\Delta\varphi$ selections. All the inputs to the Drell-Yan estimate derived from simulation are collected in Table 6.3.

6.2.4 Corrections to Simulation, $C_{Z_{mass}}$

There are two more small correction factors in the Drell-Yan background calculation. The first is the difference in the fraction of events inside the mass window between data and simulation. This is significantly different from unity only for the CEM and PHX Z events, where the mass window cut in data only includes 98% and 94%¹ as many events as it does in simulation. This correction is applied as $C_{Z_{mass}}$ for the CEM (98%) and PHX (94%) in the calculation, correcting $R_{out/in}^i$. The second correction is the scale factor needed to bring the number of extra jets generated by PYTHIA up to match the

¹The statistical uncertainty of these numbers is insignificant.

	f_i , the fraction with each jet multiplicity		
	0 jets	1 jet	≥ 2 jets
CEM	0.62 ± 0.02	0.30 ± 0.02	0.08 ± 0.01
CMUP+CMX	0.56 ± 0.02	0.33 ± 0.02	0.11 ± 0.01
PHX	0.68 ± 0.03	0.27 ± 0.02	0.05 ± 0.01
	R_i , (number inside 76-106 GeV/ c^2)/(number outside)		
	0 jets	1 jet	≥ 2 jets
CEM	1.25 ± 0.09	0.96 ± 0.10	0.91 ± 0.20
CMUP+CMX	0.41 ± 0.04	0.31 ± 0.04	0.34 ± 0.08
PHX	0.50 ± 0.06	0.50 ± 0.09	0.89 ± 0.31

Table 6.3: Inputs to the Drell-Yan estimate from simulated $Z/\gamma^* \rightarrow ee/\mu\mu$ events.

data, as described in section 6.1. This is applied as a correction factor to R_i .

6.2.5 Final Drell-Yan Background Estimate

We assemble the inputs from Tables 6.2 and 6.3 and insert these into equations 6.1 to estimate the total background from $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events. The final $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ contributions are summarized in Table 6.5.

6.2.6 Systematic Uncertainties

The largest uncertainty on the $Z/\gamma^* \rightarrow ee/\mu\mu$ background estimate is the statistical uncertainty determined by the size of the input data and simulation samples. This uncertainty is 20%.

The scale factor correcting the number of extra jets produced by PYTHIA, which is applied as a correction to R_i , has a statistical uncertainty which also contributes to the overall systematic uncertainty. We assign a 5.5% systematic uncertainty.

We also consider that the reliability of the ratios measured in simulation depends on

the ability of the simulation to model the \cancel{E}_T from mis-measured objects. We make a quantitative comparison between data and simulation by comparing the fraction of events above the $\cancel{E}_T > 25$ GeV threshold in data and simulation. To do this we first find the fraction of events with $\cancel{E}_T > 25$ GeV in data. We then calculate what threshold x applied to the simulated data is necessary to find the same fraction of events over the threshold x . We find the necessary threshold, $x = 24$ GeV. The background is then recalculated with the new threshold value $x = 24$ GeV. The background with the new threshold is found to be 13.5% lower than the default estimate. We take the full difference between the background estimates as the systematic uncertainty.

Combining the three contributions in quadrature yields a total uncertainty of 25% on this background.

6.3 $W + \text{jets with a Fake Lepton}$

The $W \rightarrow l\nu$ with extra jets process can pass the selection if one of the jets mimics a lepton. The rate of jets fluctuating into objects identified as leptons is rare, but since the $W \rightarrow l\nu + \text{jets}$ cross-section is so large compared to $t\bar{t}$ decay, this process represents a considerable background to the dilepton selection. The rate at which jets “fake”, or are misidentified as leptons, is smaller for primary leptons than for track leptons. We consider both contributions to the background.

To estimate the fake background we require two inputs:

- The fake rate, the probability for a jet to be misidentified as a lepton. This is parameterized as a function of E_T and $|\eta|$.
- The number of jets associated with single W s. This gives us an overall normalization for the total fake background. This is parameterized as a function of E_T and $|\eta|$.

The fake rate is derived in data from jets in the high energy photon sample. The photon sample is used because it is similar to the $W \rightarrow l\nu + \text{jets}$ sample, and is a large enough sample to build fake rates. The best sample to use would be the $Z + \text{jets}$ sample, but there are not enough $Z + \text{jets}$ events with higher jet multiplicities. The total number of expected fakes is the product of the fake rate, and the number of jets associated with single W s, with corrections for the efficiency for selecting the jets associated with W s. These efficiencies are for the opposite-charge requirement, the restriction on the angle between the track lepton and \cancel{E}_T , and the Z boson veto.

6.3.1 The Fake Rate

There are some subtleties of the fake rate definition that need to be explained. As previously stated, the fake rate is the number of observed fake leptons divided by the total number of observed fakeable objects. Fakeable objects are any object which passes the event selections, that could be mistaken as a lepton. This includes all jets which pass the selection cuts, but also includes tracks which are not associated with a jet. This expansion of the denominator to include tracks and not just jets ensures that the fake rate numerator is a strict subset of the denominator.

The fake rate depends on the type of jet. Jets which originated from quarks have a higher fake rate than do jets from gluons splitting, and thus, fake rates must be derived from samples with the same quark/gluon content. Most jets in the kinematic region selected in our fakeable object sample associated with W boson production are quark jets, as seen in Figure 6.2. Multijet events are dominated by processes with mostly gluons, and thus most jets originate from gluons. Because the gluon must produce two quarks, which then further fragment, the average charged particle multiplicity in a gluon jet is higher than in a quark jet. So the chance of finding only one track is low, producing lower fake rates from gluon jets than in quark jets.

Because quarks and gluons produce fakes at such different rates, reliable simulation

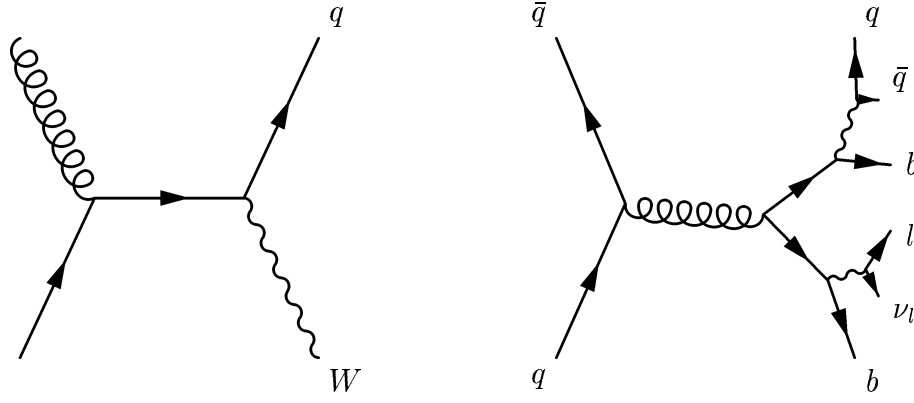


Figure 6.2: Two processes contributing to “fake” events. The left is the dominant W + jets production process in this sample, and the right is $t\bar{t}$ decay in the lepton + jets channel.

is required to model the fragmentation process correctly to accurately model the relative content of quark and gluon jets, and therefore the total fake rates. Since the simulation does not model the overall fake rate correctly we rely on data to build the fake lepton rates. We use a sample where the contamination from true leptons is negligible, as real leptons would inflate the fake rate. We select the photon + jets sample, in which the contributing processes are the same with the W and photons swapped. We use photon triggered data with the requirement of $E > 80$ GeV to produce similar kinematics and similar quark/gluon content as the W + jets sample.

Use of Z + jets Data

There is one issue in using the photon triggered sample for the construction of the fake rates. In the photon plus one jet sample, conservation of energy means that events with the E_T of the jet significantly lower than the photon E_T are rare. As a result there are few events with which to build the fake rate in the low E_T region. This only affects the zero jet event count prediction, and is not crucial to the overall fake estimate. But we can improve the situation by using the Z + one jet sample to build the fake rate in the low E_T region. The Z + jets sample is a near-perfect analog to the W + jets sample,

up to the difference in their masses. Unfortunately there are not enough $Z + \text{jets}$ events with which to build fake rates for higher jet multiplicities. But, most events have only one jet, with energy near the E_T threshold. Therefore, the total fake rate used to predict the number of fake leptons in the zero-jet dilepton sample is a combined rate from the photon and Z plus jets sample.

Validation of the Fake Rate

The validity of the track lepton fake rate can be tested in both simulation and data. We chose a lower kinematic threshold of 15 GeV for both the track lepton and jets to collect a larger sample with more statistical power. The lower kinematic threshold will add many jets to the sample due to the exponential falloff of the jet E_T distribution, and because the fake rate for 15 GeV track leptons is higher than for 20 GeV track leptons.

Test in Simulated Data Using simulated data we test the fake rate using the exact procedure used in data. We are able to measure the fake rate from photon events to predict the number of fake leptons in the jets in W events. We consider events with zero and one jets as control samples as well as events with two or more jets, in addition to the fake lepton. Figure 6.3 shows the predicted and observed fake leptons as a function of E_T of the misidentified jet, for each jet multiplicity. The total number of fake events is given in Table 6.4. The fake rate from jets associated with an 80 GeV photon could be a slight overestimate of the number of fake leptons observed in jets associated with a W , but this effect is not statistically significant except in the case of zero additional jets. The 18% discrepancy forms part of the basis of the systematic uncertainty assigned to the fake background. See Table 6.4.

Test in $Z + \text{jets}$ Data A direct test of the fake rate from photon data used in the analysis is possible. We predict the number of fake leptons using $Z + \text{jets}$ data, and compare this to the number of observed events with zero, one and two or more jets in

Predict simulated W+jets with simulated γ and Z+jets

	predicted	observed
1 jets	5473 ± 147	4480
2 jet	1332 ± 200	1047
≥ 3 jets	304 ± 67	226

Predict Z+jets data with γ +jets data

	predicted	observed
1 jets	100 ± 13	101
2 jet	28 ± 2	26
≥ 3 jets	12 ± 1	13

Table 6.4: Predicted and observed number of fake leptons in the data and simulation cross-checks. The number of jets quoted is in addition to the jet which is misidentified as a track lepton. The only statistically significant discrepancy observed is in the 1 jet category in the simulation cross-check (top table). This discrepancy forms part of the basis of the systematic uncertainty on this background estimate.

addition to the fake lepton. The event sample is small, but we observe no statistically significant discrepancy between predictions and observations. Figure 6.4 and Table 6.4 show the results of the test.

6.3.2 Fake Estimate Normalization

With the fake rate in hand, we estimate the number of fakes by multiplying the fake rate (as a function of jet multiplicity and $|\eta|$) by the normalization for the size of the data sample. The normalization is derived from data, based on the $W + \text{jets}$ sample. We select $W + \text{jets}$ events with any primary lepton type where the event passes all lepton+isolated track selection, except for cuts referring to the existence of the isolated

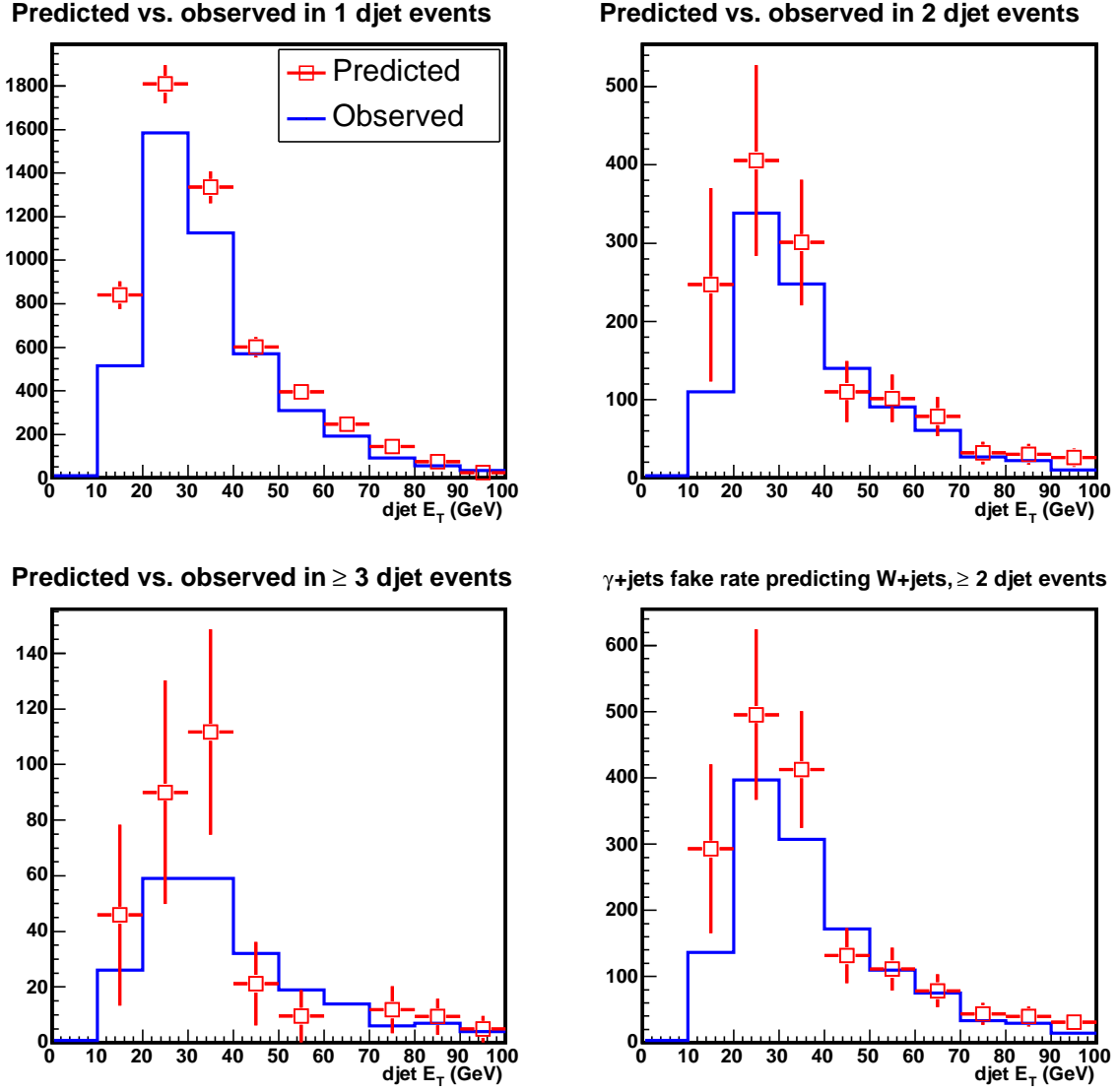
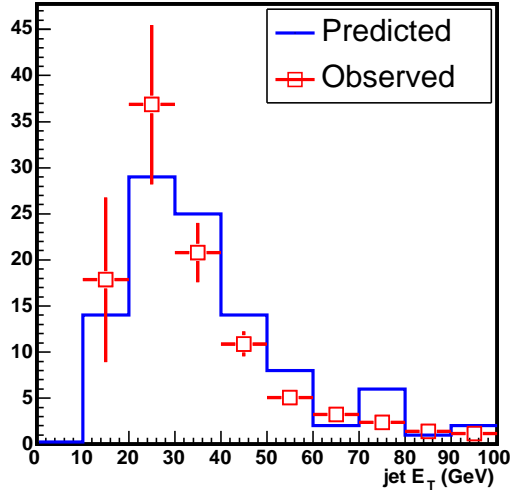


Figure 6.3: Predicted and observed number of isolated tracks in simulated $W + \text{jets}$ events, as a function of the number of jets in the event. The fake rate used is derived from photon+jets and (for events reconstructed with zero jets) Z +jets simulation. The over-prediction for zero jet events is incorporated in the systematic uncertainty on the fake lepton background.

Predicted vs. observed in 1 djet events



Predicted vs. observed in 2 djet events

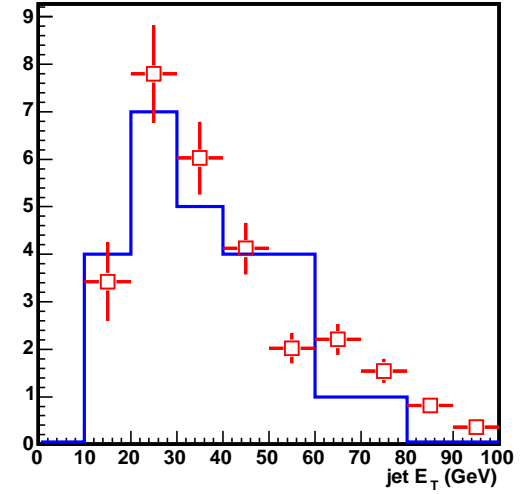
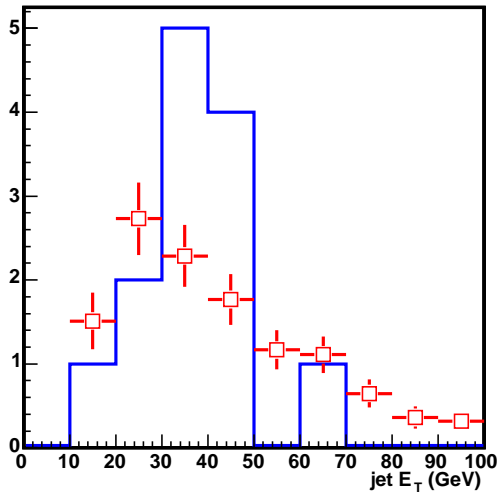
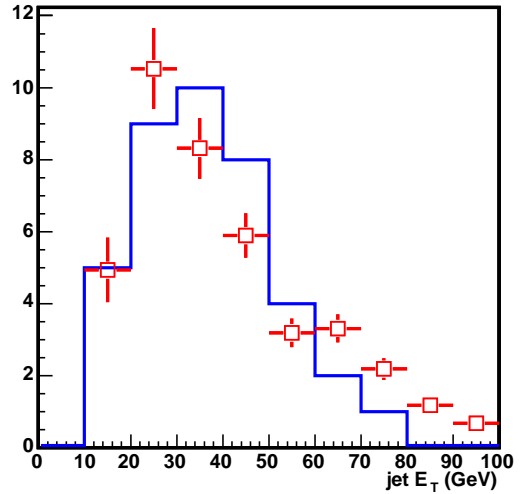
Predicted vs. observed in ≥ 3 djet eventsPredicted vs. observed in ≥ 2 jet events

Figure 6.4: Predicted and observed number of isolated tracks in $Z + \text{jets}$ data, as a function of the number of jets in the event. The fake rate used is derived from photon+jets data.

track. In short this means that there must be a primary lepton, $\cancel{E}_T > 25$ GeV which does not point at the primary lepton or jets. For each primary lepton type the E_T and $|\eta|$ of each jet in the event is noted, and the fake rate as a function of E_T and $|\eta|$ is multiplied by the normalization distributions.

6.3.3 Contribution from $t\bar{t}$ Events with a Fake Lepton

The dominant process contributing to fake events in our pretag sample is from W +jets. But there is also a non-negligible contribution from $t\bar{t}$ decay in the lepton + jets channel, where a primary lepton, \cancel{E}_T and jets are produced. The sample of events used to normalize the fakes estimate is based on W + jets selection, which naturally includes fakes from the $t\bar{t}$ decay in the lepton + jets channel. So no special treatment is necessary as long as the fake rate from $t\bar{t}$ events is the same as for generic QCD W + jets. We check the fake rate for $t\bar{t}$ lepton + jets in simulation. We predict 585 ± 128 fake events from $t\bar{t}$ lepton + jets events using the 80 GeV photon triggered fake rates. This is compared to 424 events observed, see Figure 6.5. This level of disagreement is not significant compared to the size of the statistical uncertainty combined with the 18% systematic uncertainty on the fake rate discussed previously. So we conclude we can use the photon + jets fake rate to predict the number of fake leptons, regardless of the source of events passing W + jets selection.

6.3.4 Primary Lepton Fake Rates

The fake rate for primary leptons is an order of magnitude smaller than for track leptons, but their contribution to the total pretag fake background is not negligible. We measure the pretag number of fake primary leptons using the same method as for track leptons, with the exception that a single fake rate is derived for all jet multiplicities. By combining all jet multiplicities into one fake rate we have enough data to avoid large statistical errors. Because the fake rate for primary leptons does not statistically vary among jet

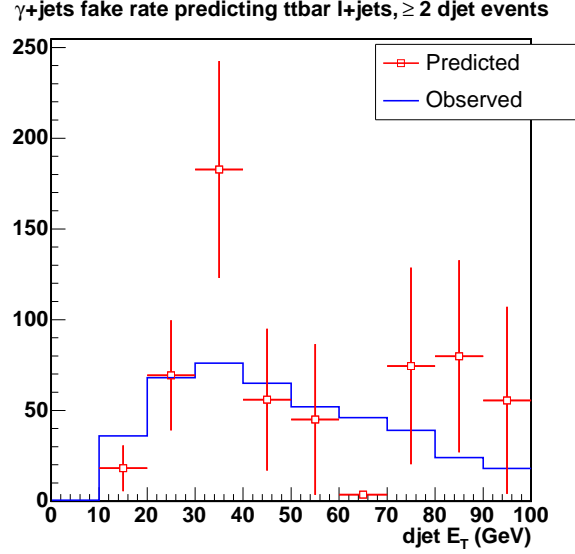


Figure 6.5: Testing the applicability of the 80 GeV photon + jets fake rate to fake leptons from top in the lepton + jets mode, in simulation. No fake leptons are observed from events with one jet.

multiplicities this is a valid strategy.

6.3.5 Opposite-Charge Fraction

Recall that the primary and track leptons are required to have opposite charge. Because the jet can fake a lepton of either charge, the final step in estimating the number of events with a fake lepton is to determine the fraction of fake events which have oppositely-charged leptons. It is necessary to consider the W and $t\bar{t}$ processes separately, as the opposite-charge fraction differs between the two.

First we will consider the opposite-charge fraction for W +jets events. Because leading diagrams for W + jets production have the W recoiling against a quark which becomes a jet in the event, the charge of the tracks in that jet can be expected to be correlated to the charge of the W .

In the ideal situation one would measure the opposite-charge fraction in simulated

$W + \text{jets}$ events. The problem is that, like the fake rate, the charge correlations depend on the details of the jet fragmentation process, which are not totally reliable in the simulation. In the simulation, the opposite-charge fraction is large, and decreases with increasing jet multiplicity. Using zero jet events, where contributions from $t\bar{t}$ processes are small, we compare the opposite-charge fractions in data and simulation and find that simulation overestimates the fraction of events with a fake lepton where the leptons have opposite charge.

Because we cannot use simulation, we measure the opposite-charge fraction in data in events with zero jets. We enhance the fraction of fake events with further quality cuts and subtract the estimated number of remaining zero-jet events from sources with two real leptons. We find $(67 \pm 3)\%$ of events have opposite-charge, which is taken as the opposite-charge fraction for all events with zero jets. We cannot use the same procedure to find the opposite-charge fraction in higher jet multiplicity events because the $t\bar{t}$ contribution is non-negligible. But, we know from simulation how the opposite-charge fraction varies as a function of jet multiplicity and use this knowledge to extrapolate the fraction in the zero-jet bin in data to that events with high jet multiplicities. We find opposite-charge fractions of 0.63 ± 0.03 for events with one jet, and 0.59 ± 0.03 for events with two or more jets.

Second we must consider fake events from the $t\bar{t}$ lepton + jets channel decays. We understand the charge correlation in this case by noting that the summed charge of the quark pair produced by one of the W bosons must be of opposite charge to the lepton from the other W boson. To produce a hadron of the same charge as the lepton requires the production of a large number of additional quark-anti-quark pairs, and is unlikely to result in an isolated track. In simulation, $78.9\%^2$ of events have oppositely-charged leptons. The opposite-charge studies in the previous $W + \text{jets}$ sample showed that we cannot rely on the simulation directly to measure the opposite-charge fraction. So we

²The statistical error on this number, and others in this section are negligible.

correct the opposite-charge measurement from simulation using the observed differences between data and simulation for $W + 1$ jet events with the jet identified as an isolated track. We find an opposite-charge fraction of 0.66 for events with at least three jets, in which one will be the jet faking the lepton.

So, to combine these two opposite-charge fractions we require the relative fractions of the $W + \text{jets}$ and $t\bar{t}$ fakes. We find the predicted fraction of $t\bar{t}$ events in the $W + \text{jets}$ data sample based on acceptance measured in simulation. This introduces a dependence on the input top cross-section. This dependence is iterated out in the same way as with the Drell-Yan background, and quickly converges. We find that 23% of the jets in the normalizing $W + \text{jets}$ sample are predicted to be from the $t\bar{t}$ leptons + jets process, and the re-weighted opposite-charge fraction is 0.61.

Final Fake Lepton Estimate

The final estimate of the number of events with N jets and a fake lepton is obtained by multiplying the jets from the $W + (N + 1)$ jet distribution by the isolated track fake rate. To include the contribution from events with a fake primary lepton, we multiply the same $W + \text{jets}$ distributions by the primary lepton fake rates. This yields a predicted number of events with two fully reconstructed leptons, where one is real, and the other fake. This is done for each possible permutation of lepton types. We then sum over the type of “true” lepton and rescale the number by the relative $W + \text{jets}$ acceptance as measured in simulation.

There are three small corrections that need to be made before we have our final total fake background estimate. There are three selection criteria that cannot be applied in selecting the $W + \text{jets}$ events, so we need to apply the efficiency for these selections. The first is the opposite-charge fraction, which has been discussed previously. The other two are the increased \cancel{E}_T threshold for events with dilepton invariant masses close to the Z resonance, and the track lepton- \cancel{E}_T opening angle veto. We measure the efficiency for

these requirements in simulation and apply them to the fake rate to yield the final fake background prediction.

Systematic Uncertainties

There are three sources of uncertainty in the W + fake lepton estimate: the statistical uncertainty on the fake rate, which varies from 5-10% depending on the jet multiplicity, the systematic uncertainty on the overall normalization of the estimate, and the uncertainty on the fraction of events with a fake lepton where the leptons have opposite sign.

The uncertainty on the overall normalization comes from the largest observed discrepancy in the data and simulation cross-checks described above. This is the 18% difference between the observed and predicted number of events in simulated events with zero reconstructed jets, which also covers possible discrepancies between the fake rate for W + jets and the one for jets from $t\bar{t}$.

The uncertainty on the opposite-sign fraction is the statistical uncertainty on the calculated opposite-sign fraction in data propagated to the ≥ 2 jet bin and is 5%.

Including both systematics and the 6% statistical uncertainty in the ≥ 2 jet bin, the total uncertainty on this background for the cross-section measurement is 20%.

6.4 Summary

Table 6.5 gives a summary of all the estimated pretag background contributions.

	0 jets	1 jet	≥ 2 jets
WW	91.7 ± 7.5	16.0 ± 1.3	3.9 ± 0.4
WZ	10.0 ± 0.8	4.6 ± 0.4	1.4 ± 0.1
ZZ	2.4 ± 0.0	0.7 ± 0.0	0.3 ± 0.0
$Z/\gamma^* \rightarrow ee$	72.4 ± 15.8	25.9 ± 6.1	7.8 ± 2.2
$Z/\gamma^* \rightarrow \mu\mu$	18.9 ± 5.3	8.9 ± 2.7	3.4 ± 1.2
$Z/\gamma^* \rightarrow \tau\tau$	35.5 ± 3.2	26.5 ± 2.5	7.3 ± 0.9
Fakes	244.1 ± 6.4	76.8 ± 14.6	29.9 ± 5.9
All Backgrounds	475.0 ± 51.6	159.2 ± 16.9	54.0 ± 6.6

Table 6.5: Predicted background events in the pretag 1.1 fb^{-1} dataset, Systematic uncertainties are included.

Chapter 7

Tagged Background Estimate

The tagged background estimate is very different from the pretag background estimate. First, the nature of the backgrounds themselves are very different for the two analyses. In the pretag analysis the dominant background is W +jets events faking a lepton. In the tagged analysis, backgrounds from b -jets dominate, for example, $Z + b\bar{b}$ events, and events where the b -jet is misidentified. Second, in the tagged analysis we are able to estimate all backgrounds, except for the fakes, using one data driven technique, as opposed to using different techniques for each background type as in the pretag analysis.

The tagged background estimate is based on a data driven technique of estimating the tagging rates of generic QCD multijet production. We use the pretag background estimate as a normalization for the tag rates of tagged lepton+track selected events. Several corrections to the data driven estimate need to be made for tags from $t\bar{t}$, and fakes. Simulated events are used to estimate the size of the corrections.

7.1 Estimating Backgrounds with the Tag Rate Matrices

We use a data driven background technique to estimate the tagged backgrounds. Chapter 5 explained the concept of positive and negative tags, which are interpreted to be tags of long lived B hadrons and “mistags”, or mistakes due to resolution effects, respectively. Positive and negative tag rates of generic QCD jets are found to be dependent on five jet quantities, jet E_T , the number of tracks in the jet, jet η , the number of primary vertices in the event, and the total scalar sum of the E_T of all the jets in the event, which we call sum- p_T . We parameterize the tag rates of generic QCD jets as a function of these five quantities. These parameterizations are termed “tag matrices”. The generic QCD jet samples used to build the matrices contain real tags from B hadron decays, as well as mistags. So we can use the positive tag matrix to estimate the number of tagged events in our dataset from background processes with the same final event signature, and mistags based on the jets in pretag events.

The tag rate matrices are obtained by examining all jets with uncorrected jet $E_T > 10 \text{ GeV}^1$ within the fiducial volume of the silicon detectors; $|\eta| < 2.4$. Further jets are required to have at least two tracks that pass the tagging quality cuts described in Section 5.1, so that it is possible for the jet to be tagged. The tag rate matrices then are filled with the fraction of all taggable jets, which are in fact tagged, as a function of the five matrix parameters. We use the tag rate matrix by querying the matrix for each jet in all pretagged candidate events, and getting a tag rate for that jet based on the jet quantities; jet E_T , jet η , the number of tracks in the jet, and sum- p_T . These tag rates are combined together to find an ensemble predicted number of tags for the pretagged jet sample.

¹Uncorrected E_T means that the jet corrections described in Section 3.2.2 and the \cancel{E}_T corrections to jets described in Section 3.3 are not applied.

7.2 Tagged Background

The background of the tagged dilepton sample can be organized into two parts. The first is made up of processes with a decay signature similar to the signal $t\bar{t}$ events, such as $Z + b\bar{b}$ events, and mistags. In the $Z + b\bar{b}$ case, mis-measurement of some object leads to false \cancel{E}_T . In the mistag case, a jet is falsely identified as having come from a B hadron. The second category of background events are fakes, where a jet is falsely identified as a primary or track lepton.

The tag rate matrix technique of estimating backgrounds relies on the assumption that the tag rate matrices are applied to generic jets produced via QCD interactions, the same type of jets that are used to build the matrices. Therefore, using this method in dilepton top samples requires corrections to the expected number of background tags, N_{matrix} , for several effects: First, top signal events make up a considerable portion of the pretag events, and should not contribute to the background estimation. Second, a portion of the fakes background is not due to generic QCD processes, but are $t\bar{t}$ decays in the lepton + jets channel. Because these $t\bar{t}$ events do not have the same tagging rate as generic QCD events as predicted by the tag rate matrices, we need to subtract their (incorrect) contribution from the matrix estimate, and add back in the correct contribution they make to the background. It should be noted that the background from events with a fake lepton from QCD processes is well estimated by the tag rate matrices, and is included in N_{matrix} . Thus, the total dilepton tagged background is given by:

$$N_{bkg}^{tag} = N_{matrix} - N_{matrix}^{dil} - N_{matrix}^{LJ} + N_{fakes} \quad (7.1)$$

where N_{matrix} is the predicted number of tagged background events by the tag rate matrices, N_{matrix}^{dil} is the number of top dilepton events in the pretagged dilepton candidates, N_{matrix}^{LJ} is the number of events from N_{matrix} background due to $t\bar{t}$ decay in the lepton+jets channel that enter the dilepton acceptance as fakes, and N_{fakes} is the proper estimate of contamination from lepton+jets events in the fake background.

7.2.1 Validation of events with low \cancel{E}_T

We first demonstrate the ability of the tag rate matrices to predict the tag rates in Z +jet events. There is a possibility that differences between data and simulation in quantities such as sum- p_T could affect the predicted tag rates, and change the background estimation.

The predictive power of the tag rate matrices is tested with Z +jets events. Events are selected using a modified lepton+isolated track selection. We require $\cancel{E}_T \leq 20$ GeV to reduce the contribution from $t\bar{t}$ events, and do not apply any $\Delta\phi$ cuts between the \cancel{E}_T and any other object. The invariant mass of the lepton pair must fall in the range $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. We compare the number of:

- tagged events observed in data
- predicted tagged events in data
- predicted tagged events in MC

If the tag rate matrices are capable of predicting the tag rates in simulated data, these predictions should be compatible. First, we ensure that the distributions we wish to test are similar before applying the tag rate matrices. Figure 7.1 shows the pretagged distributions for the quantities for both data and simulation. We observe generally good agreement in these distributions. We then apply the tag rate matrices to the pretag data and simulation. Figure 7.2 shows the distributions of the number of observed tags in data, the number of predicted tags in data, and the number of predicted tags in simulation. Again, we see general agreement between the observed and prediction distributions. Tables 7.1- 7.4 show the results of these comparisons numerically. So we conclude that the data and simulation are similar enough in composition to be able to use the tag rate matrices to predict the background in events with low \cancel{E}_T .

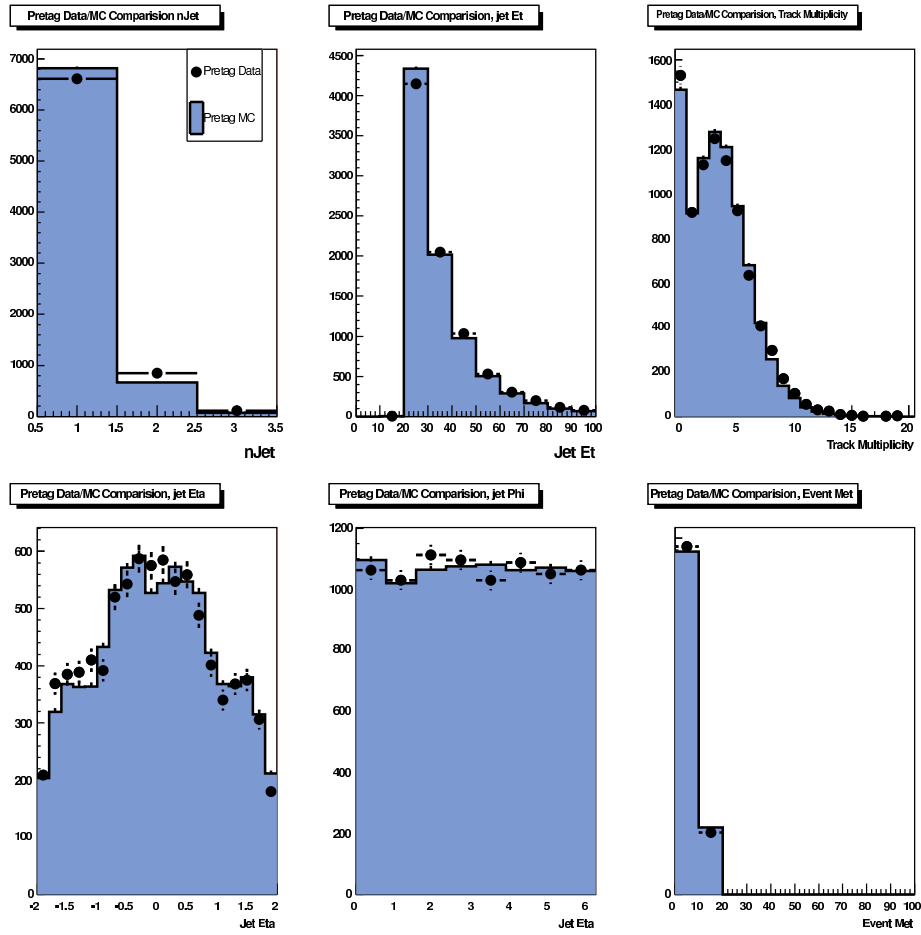


Figure 7.1: Kinematic distributions for pretagged Z +jets candidate events for data and simulation. Note that the uncertainties on the data are statistical only.

Sample	1 Jet	2 Jets	3 Jets	Total
Obs. Tags in Data	156 ± 12	43 ± 7	15 ± 4	214 ± 15
Pred. Tags in Data	165 ± 18	50 ± 8	12 ± 4	227 ± 24
Pred. Tags in simulation	159 ± 18	49 ± 8	11 ± 3	218 ± 23

Table 7.1: Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of jet multiplicity. Statistical and systematic uncertainties predictions have been combined.

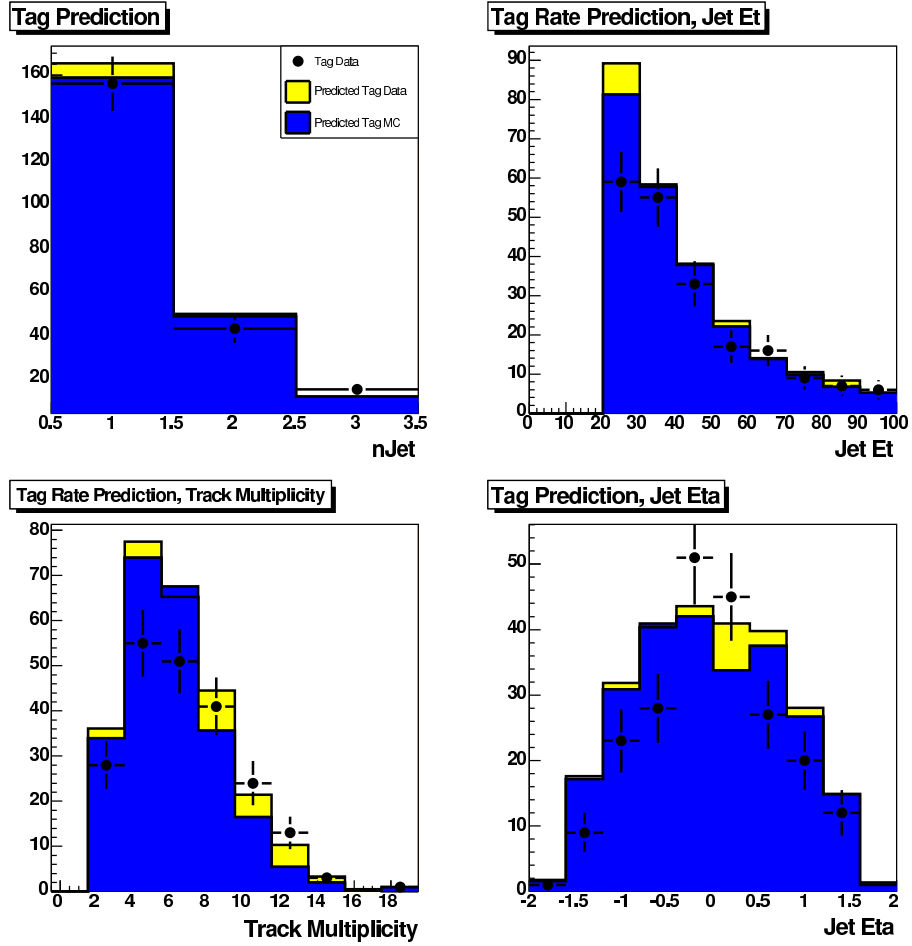


Figure 7.2: Kinematic distributions for observed tags in data, and predicted tags in data and simulation. Note that the uncertainties on the observations in data are statistical only, and the predictions in data and simulation are subject to a 5% systematic uncertainty from the positive tag rate matrix.

Jet E_T bin	Obs Tags in Data	Pred. Tags in Data	Pred. Tags in simulation
20 - 30 GeV	59 ± 8	89 ± 7	81 ± 6
30 - 40 GeV	55 ± 7	58 ± 5	58 ± 5
40 - 50 GeV	33 ± 6	38 ± 3	38 ± 3
50 - 60 GeV	17 ± 4	24 ± 2	22 ± 2
60 - 70 GeV	16 ± 4	14 ± 1	14 ± 1
70 - 80 GeV	9 ± 3	10 ± 1	10 ± 1
80 - 90 GeV	7 ± 3	8 ± 1	7 ± 1
90 - 100 GeV	6 ± 2	5 ± 1	5 ± 1
Total	202 ± 14	246 ± 10	236 ± 9

Table 7.2: Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of Jet E_T . Uncertainties on the total predictions (bottom line) are statistical and systematic combined.

7.2.2 Validation of events with significant \cancel{E}_T

A part of the background in the pretag analysis is from events where \cancel{E}_T is generated from some mis-measurement, mainly of jets. Since the lepton+isolated track selection requires that the \cancel{E}_T not point directly at either of the jets, both jets must be mis-measured to some extent. One example would be if the energy of both jets is underestimated because the jets partially overlap with an uninstrumented part of the calorimeter. In this case the jet quantities that are used by the tag rate matrices for predicting the tag rates would be incorrect. The measured jet energy would be lower than the true jet energy. So the tag rate matrix would return a different tagging rate than it would for a correctly measured jet. This kind of effect is not tested well by the Z +jets sample because there is no false \cancel{E}_T .

To test the predictive ability of the tag rate matrices in such events we select $Z + 1$

Track multiplicity bin	Obs Tags in Data	Pred. Tags in Data	Pred. Tags in simulation
2 - 3	28 ± 5	36 ± 4	34 ± 3
4 - 5	55 ± 7	78 ± 9	74 ± 8
6 - 7	51 ± 7	65 ± 7	68 ± 7
8 - 9	41 ± 6	45 ± 5	36 ± 4
10 - 11	24 ± 5	21 ± 3	16 ± 2
12 - 13	13 ± 4	10 ± 1	5 ± 1
14 - 15	3 ± 2	3 ± 1	2 ± 1
16 - 17	0 ± 1	0 ± 1	0 ± 1
18 - 19	1 ± 1	1 ± 1	0 ± 1
Total	216 ± 15	259 ± 13	235 ± 12

Table 7.3: Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of track multiplicity in jets. Uncertainties on the predictions are statistical and systematic combined.

Jet η bin	Obs Tags in Data	Pred. Tags in Data	Pred. Tags in simulation
-2.0 - -1.6	1 ± 1	2 ± 1	2 ± 2
-1.6 - -1.2	9 ± 3	18 ± 2	17 ± 1
-1.2 - -0.8	23 ± 5	32 ± 3	31 ± 2
-0.8 - -0.4	28 ± 5	40 ± 4	41 ± 3
-0.4 - 0	51 ± 7	44 ± 4	42 ± 3
-0 - 0.4	45 ± 7	41 ± 3	34 ± 3
0.4 - 0.8	27 ± 5	40 ± 3	37 ± 3
0.8 - 1.2	20 ± 4	28 ± 2	27 ± 2
1.2 - 1.6	12 ± 3	15 ± 1	15 ± 1
1.6 - 2.0	0 ± 0	1 ± 1	1 ± 1
Total	216 ± 15	259 ± 8	247 ± 7

Table 7.4: Comparison of the observed and predicted tags in data and simulated Z +jets events as a function of Jet η . Uncertainties in the predictions are statistical and systematic combined.

Sample	1 Jet
Obs. Tags in Data	9 ± 3
Pred. Tags in Data	8.3 ± 1.4
Pred. Tags in MC	8.1 ± 1.3

Table 7.5: Comparison of the observed and predicted tags with 1 jet data and simulated events. Errors on the predictions are statistical and systematic combined.

jet events to avoid samples with significant contributions from $t\bar{t}$ events. We require that the \cancel{E}_T and jet are aligned within 25° . The invariant mass of the lepton pair is again required to be $76 \text{ GeV}/c^2 < M < 106 \text{ GeV}/c^2$. As in the previous study we compare the number of observed tags in the data with the number of tags predicted by the tag rate matrices in data, and the number of tags predicted by the tag rate matrices in simulation. Results are in Table 7.5. Again we draw the conclusion that the tag rate matrices are able to predict the observed tag rate within errors.

7.3 Physics and Mistag Backgrounds

N_{matrix} is the number of background events returned by the tag rate matrices. For each jet in all pretagged candidate events the matrix predicts the tagging probability for that jet based on the jet quantities; jet E_T , jet η , the number of tracks in the jet, the number of primary vertices in the event, sum- p_T . These tag probabilities are combined for all jets. The statistical error on the final prediction takes into account the statistical error on the prediction from the number of events in the matrix bin from which the prediction is based. Correlations between tagging probabilities from the same matrix bins are taken into account.

7.3.1 Top dilepton signal component of the tag rate matrix estimate

The pretagged dilepton sample has a large fraction of $t\bar{t}$ events, which should not contribute to the background estimate. The contribution to N_{matrix} from dilepton channel $t\bar{t}$ decays is estimated using simulated $t\bar{t}$ events. We derive the tag rate (ϵ_{matrix}) by applying the tag rate matrix to the jets in the pretagged candidate events in $t\bar{t}$ simulated events, and divide by the total number of pretagged events. The total contribution to the background N_{matrix}^{dil} from $t\bar{t}$ dilepton events is then given by multiplying the $t\bar{t}$ tag rate by the estimated number of top events in the pretagged data.

$$N_{matrix}^{dil} = \epsilon_{matrix}^{dil} \times (N_{obs}^{pretag} - N_{bkg}^{pretag})$$

We find N_{matrix}^{dil} to be 8.6 ± 2.2 , values used in the calculation are found in Table 7.6.

7.4 Fake Lepton Background

As discussed in Chapter 6 the fake lepton background originates from the QCD radiation of extra jets from $W + \text{jets}$ process and from the lepton + jets decay channel of $t\bar{t}$ events. The first process is well modeled by the tag rate matrix method, because the jets are generic QCD jets. But the tag rate in $t\bar{t}$ events is higher than for generic jets, so fake events from $t\bar{t}$ decays need to be treated differently.

In Section 7.3.1 decays from $t\bar{t}$ were subtracted from N_{matrix} because decays from the signal process, $t\bar{t}$ decays in the *dilepton* channel should not contribute to the background counts. We need to treat the $t\bar{t}$ decays in the lepton+jets channel as background events because we reject these events when calculating the acceptance for our selection. We subtract the $t\bar{t}$ lepton + jets contribution in the same way as for dilepton events. Using simulated $t\bar{t}$ lepton + jets events we find the tag rate, and then multiply by the number

of pretag fake background events from the $t\bar{t}$ process

$$N_{matrix}^{LJ} = N_{fakes}^{pretag} \times f_{LJ}^{fakes} \times \epsilon_{matrix}^{LJ}$$

We find N_{matrix}^{LJ} to be 0.7 ± 0.1 .

The actual contribution that $t\bar{t}$ decays in the lepton + jets channel (N_{fakes}), makes now needs to be estimated and added back into the total background estimate, see equation 7.1. This boils down to finding the actual $t\bar{t}$ content which is tagged in events passing the $W + \text{jets}$ selection. N_{fakes} is factorized as:

$$N_{fakes} = N_{fakes}^{pretag} \times f_{W+\geq 3jets}^{tag} \times f_{tag}^{top}$$

N_{fakes}^{pretag} is the total estimated number of fakes in the pretagged sample, corrected for the opposite sign fraction in pretag and tagged fakes. $f_{W+\geq 3jets}^{tag}$ is the fraction of $W + \geq 3$ jet events in data that are tagged, and f_{tag}^{top} is the fraction of tagged $W + \geq 3$ jets events which are actual $t\bar{t}$ decays in the lepton + jet channel.

The fraction $f_{W+\geq 3jets}^{tag}$ is derived from data by counting the number of events satisfying $W + \geq 3$ jets selection which are also b tagged, and dividing by all $W + \geq 3$ jets events.

f_{tag}^{top} is estimated using both data and simulation via the equation

$$f_{tag}^{top} = \frac{\alpha_{LJ} \times \int \mathcal{L} dt \times \sigma_{t\bar{t}}}{N_{cand}^{LJ}}$$

where α_{LJ} is the acceptance in simulated $t\bar{t}$ events for tagged $W + \geq 3$ jet events, $\int \mathcal{L} dt$ is the total luminosity of the dataset, and $\sigma_{t\bar{t}}$ is the theoretical top quark cross-section, 6.7 pb. So the fraction of $W + \geq 3$ jets events which are from $t\bar{t}$ decays is estimated by multiplying the acceptance for $t\bar{t}$ decays in the lepton + jets channel using the pretag lepton plus isolated track selection, by the integrated luminosity of the dataset and the theoretical $t\bar{t}$ cross-section, giving the number of expected fake tagged $t\bar{t}$ lepton + jets channel decays in the sample. By subsequently dividing by the number of candidate lepton + jets events in the data we find the fraction of $t\bar{t}$ events that are tagged. Using

the theoretical $t\bar{t}$ cross-section in the calculation of f_{tag}^{top} is undesirable in the calculation of the $t\bar{t}$ cross section. We remedy the situation the same way we did when calculating the Drell-Yan pretag backgrounds, see Section 6.2. We eliminate this theoretical dependence by iterating over the value of f_{tag}^{top} . We first use the theoretical value of the cross-section, and calculate our first output cross-section. We then use the calculated $t\bar{t}$ cross section in the calculation of f_{tag}^{top} , and recalculate the cross section. This iterative procedure is continued until the final $t\bar{t}$ cross-section changes by less than 1%. It is only necessary to iterate twice. We find $N_{fakes} = 5.1 \pm 1.1$ events.

7.5 Final Background Estimate

All inputs to the calculation of the background are summarized in Table 7.6. The total background is estimated to be $N_{bkg}^{tag} = 9.5 \pm 2.7$ events.

7.6 Systematic Uncertainties on the Background

The systematic error on the background estimate consists of the errors on the pretag inputs, and the systematic error on tagging predictions made by the tag rate matrix. Statistical errors on quantities derived from simulation, such as ϵ_{M1}^{top} , and $f_{W+\geq 3}^{tag}$, are negligible. Tagging predictions made by the tag rate matrices have statistical uncertainties associated with them due to the sample size for each matrix bin, and these errors are returned by the matrix. For the matrix technique, we also apply an 8% systematic uncertainty associated with predictions from the mistag matrix. This systematic arises from charm and light flavour contamination in the datasets used to derive the tag rate matrices. Because we use the mistag matrix to find the contribution from top events to N_{matrix} , the systematic from the matrix is anti-correlated between N_{matrix} , ϵ_{matrix}^{dil} , and ϵ_{matrix}^{LJ} . Thus we only apply the systematic to the physics background portion of N_{matrix} , which is determined as $N_{matrix} - N_{matrix}^{dil} - N_{matrix}^{LJ}$.

Background Input	Value
N_{matrix}	13.7 ± 1.1
$N_{obs}^{pretag} - N_{bkg}^{pretag}$	70.1 ± 10.6
ϵ_{matrix}^{dil}	0.12 ± 0.03
N_{matrix}^{dil}	8.6 ± 2.2
N_{fakes}^{pretag}	30.7 ± 6.1
f_{LJ}^{fakes}	0.23 ± 0.02
ϵ_{matrix}^{LJ}	0.106 ± 0.003
N_{matrix}^{LJ}	0.7 ± 0.1
N_{fakes}^{pretag}	30.7 ± 6.1
$f_{W \geq 3j}^{tag}$	0.239 ± 0.008
f_{tag}^{top}	0.701 ± 0.027
N_{fakes}	5.1 ± 1.1
N_{bkg}^{tag}	9.5 ± 2.7

Table 7.6: Tagged background calculation inputs.

Source	Uncertainty on background	Uncertainty on cross-section
tag matrix technique	8%	1%
W + fake lepton	20%	2%

Table 7.7: Summary table of systematic uncertainties on the tagged background estimate.

The fake estimate N_{fakes} is subject to the same 20% systematic uncertainty as the pretag fake background estimate, and is 100% correlated between the pretag and tagged analyses.

The total tagged background systematic uncertainty is obtained by adding the total uncertainty on the matrix and fakes predictions and adding these contributions in quadrature. We find the overall systematic uncertainty on the background estimate to be 28%. The systematic uncertainties on all of the backgrounds, and the corresponding uncertainty contributed to the cross-section measurement, are collected in Table 7.7.

Chapter 8

Results

8.1 Tagged Cross-section Result

Using the selection criteria described in Chapters 3 and 5, we select a sample of events with a well reconstructed lepton, a track lepton, significant amount of missing transverse energy, and at least two jets, at least one of which is identified as having come from B hadron decay. We observe 69 such events. We calculate the pretag acceptance and event tagging efficiency in Chapters 4 and 5 as $(0.843 \pm 0.025)\%$ and 0.669 ± 0.037 , respectively. The total tagged background contribution is calculated in Chapters 6 and 7 and is estimated to be 9.5 ± 2.7 events. With an integrated luminosity of $(1000 \pm 60) \text{ pb}^{-1}$, the $t\bar{t}$ dilepton cross-section is calculated via the equation

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{\epsilon_{tag} \times \mathcal{A} \times \int \mathcal{L} dt} \quad (8.1)$$

to be

$$\sigma_{t\bar{t}} = 10.5 \pm 1.8(\text{stat.}) \pm 0.8(\text{sys.}) \pm 0.6(\text{lum.})\text{pb} , \quad (8.2)$$

where the first uncertainty is statistical, the second systematic, and the third due to the systematic uncertainty on the luminosity measurement. This result assumes that $m_t = 175 \text{ GeV}/c^2$. Table 8.1 gives results with other assumed values of the top quark mass.

Input m_t (GeV/ c^2)	Theoretical σ (pb)	$t\bar{t}$ Acceptance	Pretag Measured σ (pb)	Tagged Measured σ (pb)
170	$7.8^{+0.9}_{-1.0}$	$0.69 \pm .02$ %	$10.2 \pm 1.5 \pm 0.9$	$12.9 \pm 2.2 \pm 1.0$
175	$6.7^{+0.7}_{-0.9}$	$0.85 \pm .03$ %	$8.3 \pm 1.3 \pm 0.7$	$10.5 \pm 1.8 \pm 0.8$
180	$5.8^{+0.6}_{-0.8}$	$1.01 \pm .03$ %	$6.9 \pm 1.1 \pm 0.6$	$8.8 \pm 1.5 \pm 0.7$

Table 8.1: The pretag and tagged cross-section as calculated at several input top masses. Theoretical prediction from Ref. [15]. The first uncertainty is statistical, the second is systematic. A common uncertainty of 6%, due to the uncertainty on the integrated luminosity, is omitted.

To test the robustness of the analysis, we study the impact of modifying the selection criteria. We consider the variations to the tagging selections such as requiring two tagged b -jets, and using a b tagging algorithm with more stringent quality cuts. These variations drastically reduce the contributions from background events. We also show some kinematic features of the candidate event sample.

8.2 Pretag Cross-section Measurement

The cross-section measurement using the pretag selection is highly correlated to the tagged measurement. Table 8.2 shows the expected $t\bar{t}$ and background contributions in each of the zero and one jet control samples, as well as in the greater than or equal to two jet signal sample. This information is displayed graphically in Figure 8.1. The pretag selection produces a measurement of

$$\sigma_{t\bar{t}} = 8.3 \pm 1.3(\text{stat.}) \pm 0.7(\text{sys.}) \pm 0.5(\text{lum.})\text{pb} \quad (8.3)$$

which is consistent with the Standard Model prediction of $6.7^{+0.7}_{-0.9}$ pb for a top mass of 175GeV/ c^2 [15]. To test the consistency of the tagged and pretagged cross-section results we do one million pseudoexperiments, assuming that the true cross-section is the pretag

	0 jets	1 jet	≥ 2 jets
WW	91.7 ± 7.5	16.0 ± 1.3	3.9 ± 0.4
WZ	10.0 ± 0.8	4.6 ± 0.4	1.4 ± 0.1
ZZ	2.4 ± 0.0	0.7 ± 0.0	0.3 ± 0.0
$Z/\gamma^* \rightarrow ee$	72.4 ± 15.8	25.9 ± 6.1	7.8 ± 2.2
$Z/\gamma^* \rightarrow \mu\mu$	18.9 ± 5.3	8.9 ± 2.7	3.4 ± 1.2
$Z/\gamma^* \rightarrow \tau\tau$	35.5 ± 3.2	26.5 ± 2.5	7.3 ± 0.9
Fakes	244.1 ± 6.4	76.8 ± 14.6	29.9 ± 5.9
All Backgrounds	475.0 ± 51.6	159.2 ± 16.9	54.0 ± 6.6
$t\bar{t}, \sigma = 6.7$ pb	1.2 ± 0.1	17.3 ± 0.6	60.5 ± 1.9
Predicted	476.2 ± 51.6	176.5 ± 17.0	114.5 ± 7.0
Observed	443	187	129

Table 8.2: Predicted and observed events in the pretag 1.1 fb^{-1} dataset, with details of the background contributions. Systematic uncertainties are included in the prediction numbers.

measured value. In each pseudoexperiment we randomly vary the pretag and tagged signals and backgrounds within their uncertainties. We find a cross-section equal to, or higher than, the measured tagged cross-section measurement in 3% of pseudoexperiments.

8.3 Cross-Checks

8.3.1 Double Tag Cross-section

Requiring two jets to be b -tagged almost totally eliminates any background events. We estimate the double tag event tagging efficiency from equation 5.2 using only the term $F_{2b}\epsilon'^2$, and find the double tagged event tagging efficiency to be (0.191 ± 0.011) . This result is verified using the simulation method. Using the data-driven tag rate matrix

Events Predicted vs. Number of Jets

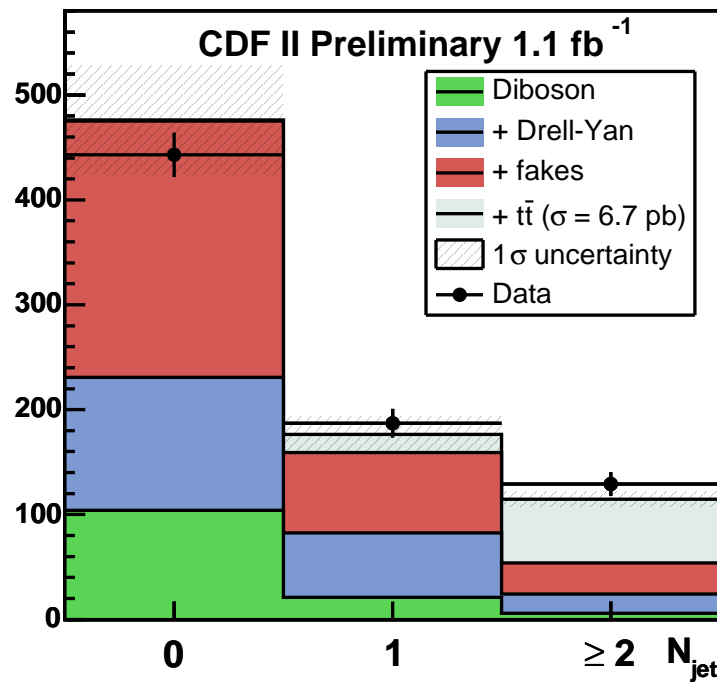


Figure 8.1: Number of predicted events compared to the number observed in the pretag candidate data. The cross-hatched areas show the (1σ) uncertainties on the predicted numbers (statistical and systematic).

technique used for the main tagged analysis we estimate the background contribution to be 2.1 ± 1.1 events. We observe 21 events passing the pretag event selection with 2 b -tagged jets. Using the same systematic uncertainties as the main tagged analysis we obtain a double tagged cross-section of

$$\sigma_{t\bar{t}} = 11.7 \pm 3.5(\text{stat.}) \pm 0.9(\text{sys.}) \pm 0.7(\text{lumi.})\text{pb.},$$

which is consistent with the main tagged cross-section result.

8.3.2 Higher purity b tagging algorithm

As discussed in Chapter 5 there are two versions of the SECVTX tagging algorithm. The main result of this thesis has used the lower purity, higher efficiency version. We perform a cross-check using the higher purity, lower efficiency version of the algorithm. We observe 59 pretag candidate events with at least one tagged b -jet. Background processes are expected to contribute 7.8 ± 1.5 events, and we find the event tagging efficiency to be 0.576 ± 0.035 . This gives a cross-section measurement of

$$\sigma_{t\bar{t}} = 10.5 \pm 1.8(\text{stat.}) \pm 0.8(\text{sys.}) \pm 0.6(\text{lumi.})\text{pb.},$$

which is consistent with the main tagged cross-section result.

8.4 Kinematic Features

It is important to examine the kinematic features of the events in our sample. Because the selected sample is expected to be rich in $t\bar{t}$ events, the events observed in data can be compared to that predicted by the simulation. The shapes of predicted $t\bar{t}$ events are from simulated $t\bar{t}$ decays with $m_{top} = 175 \text{ GeV}/c^2$, with the overall contribution normalized to the observed cross-section. Background shapes are from the same simulated events, normalized to measured background sizes. We define H_T , which is the scalar sum of the primary lepton E_T , track lepton p_T , \cancel{E}_T , and each jet E_T . The distribution of H_T in the

tagged candidate sample is shown in Figure 8.2. The large top quark mass leads to large values of H_T in top quark decay events, so H_T is an interesting kinematic variable. Good agreement between observed data and expected background distributions is observed.

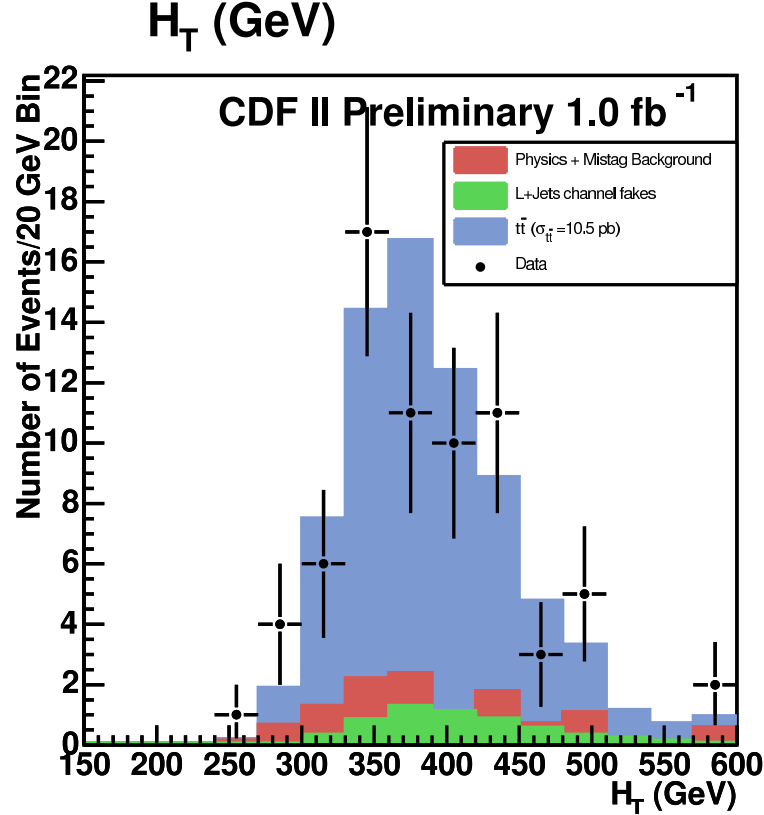


Figure 8.2: Summed scalar energy of the leptons, \cancel{E}_T and jets in the tagged candidate sample. The signal region is scaled to the measured cross-section.

In analogy to the pretag measurement the \cancel{E}_T distribution of the tagged sample is also interesting. Figure 8.3 shows the \cancel{E}_T distribution for the tagged candidate sample, and the expected background contribution. Again, good agreement is observed. We also show the lepton invariant mass in Figure 8.4, the primary lepton E_T in Figure 8.5 and the track lepton p_T in Figure 8.6. The E_T of the tagged jets in all candidate events is shown in Figure 8.7 and the track multiplicity in tagged jets is shown in Figure 8.8.

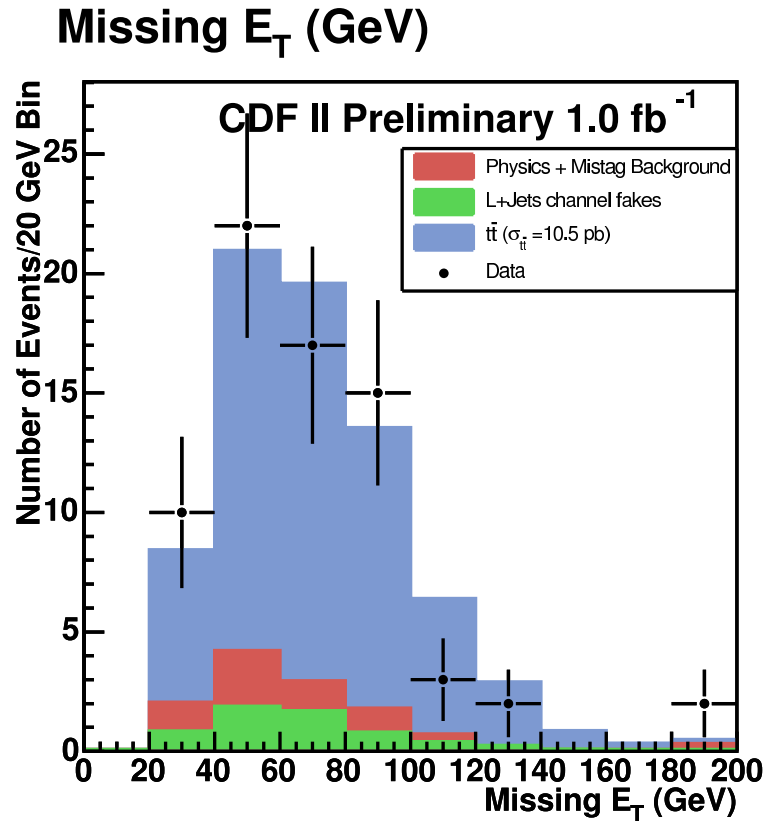


Figure 8.3: Missing transverse energy of the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

Lepton Inv Mass (GeV)

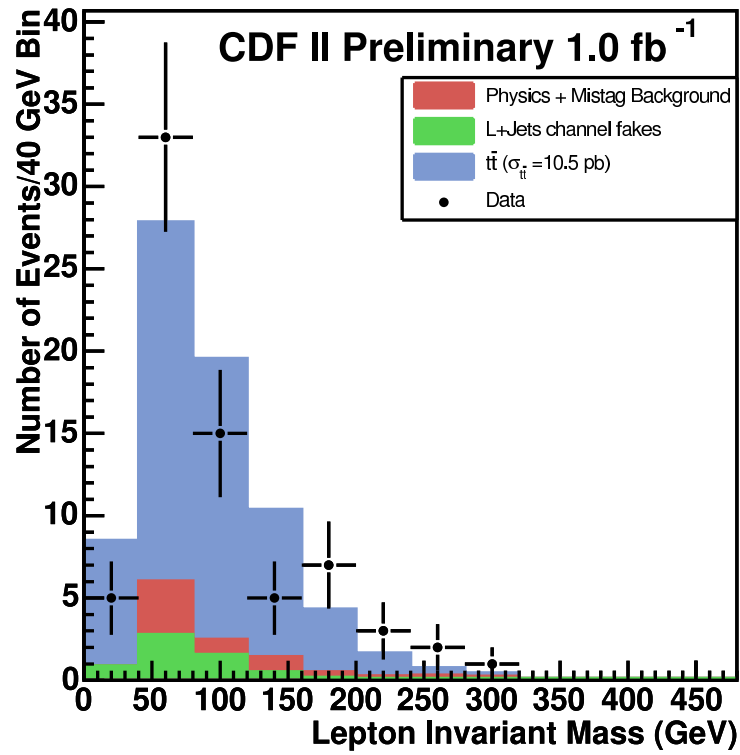


Figure 8.4: Invariant mass of the lepton pair in tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

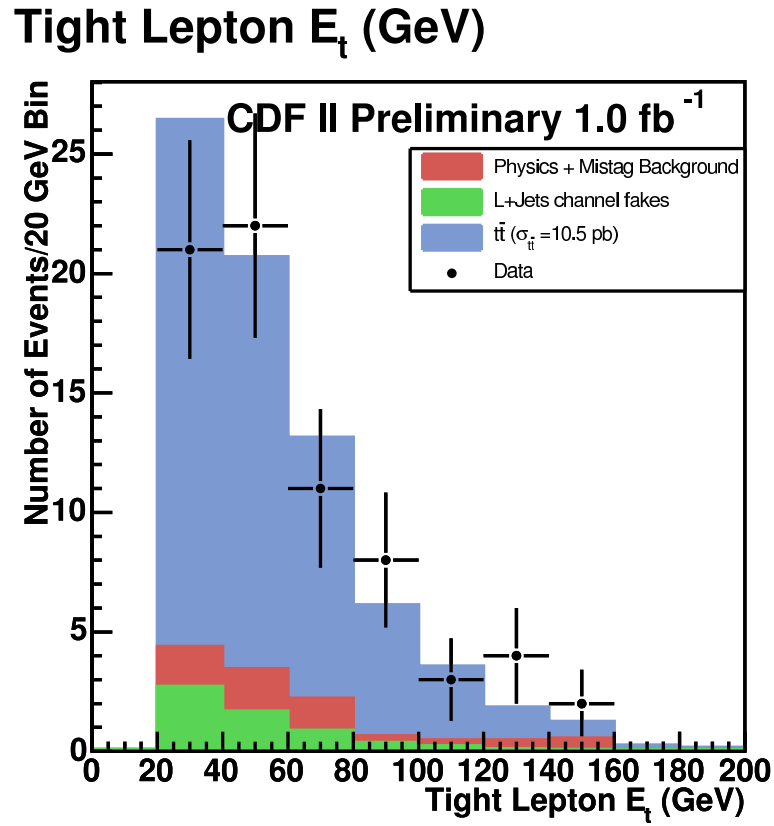


Figure 8.5: Primary lepton E_T in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

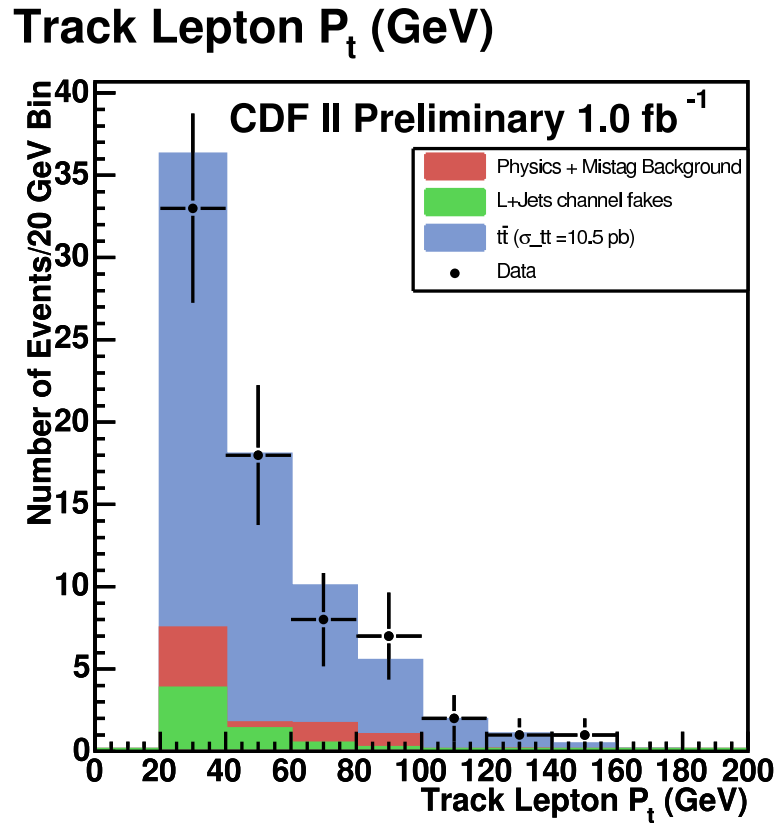


Figure 8.6: Track lepton p_T in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

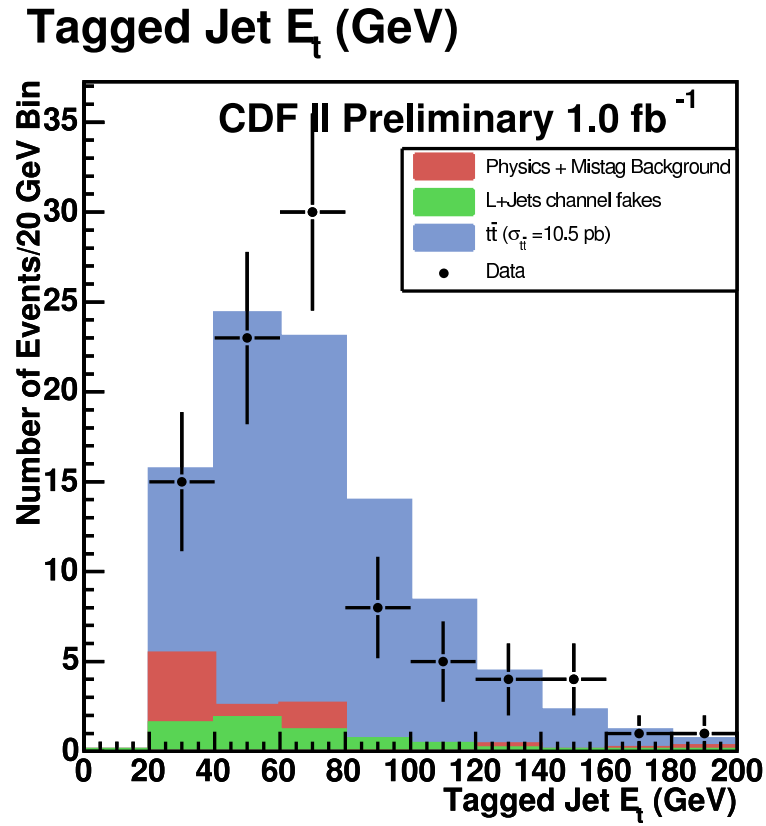


Figure 8.7: The E_T of the tagged jets in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

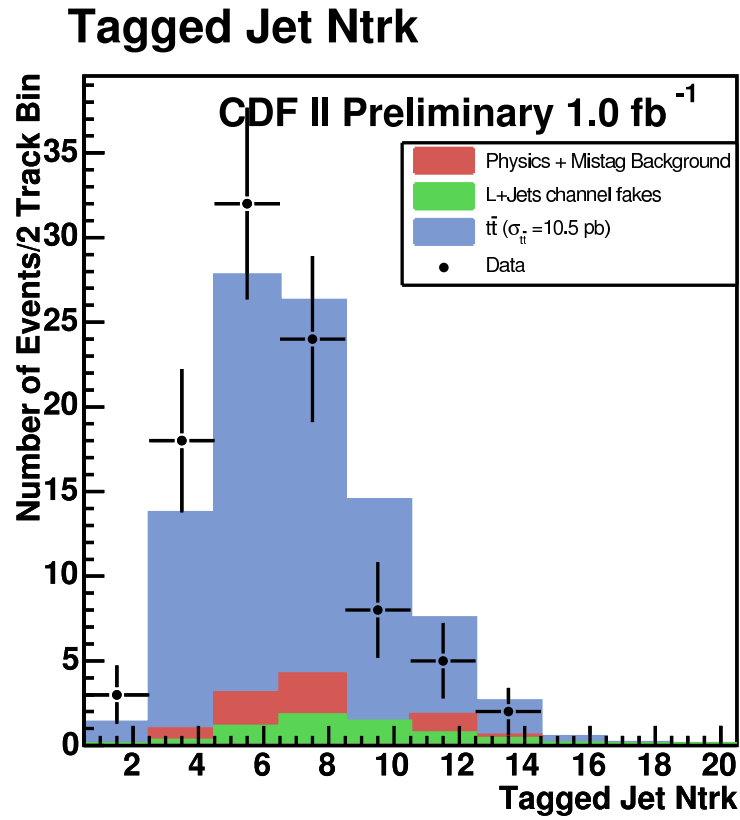


Figure 8.8: The track multiplicity of the tagged jets in the tagged candidate sample and predicted background contributions. The signal region is scaled to the measured cross-section.

Chapter 9

Conclusions

Since the upgraded Tevatron Collider resumed operation in 2001, both the CDF and DØ Collaborations have collected over 2 fb^{-1} of data. The Tevatron at Fermilab remains the only place in the world to study the top quark until the the LHC starts colliding beams in late 2008.

This thesis has presented the first measurement at CDF, of the $t\bar{t}$ production cross-section in the dilepton channel, using events identified with one fully reconstructed lepton, one isolated track and at least one identified b -jet. The result is

$$\sigma_{t\bar{t}} = 10.5 \pm 1.8(\text{stat.}) \pm 0.8(\text{sys.}) \pm 0.6(\text{lum.})\text{pb} ,$$

assuming a top quark mass of $175 \text{ GeV}/c^2$. For this mass the expected top quark cross-section is $6.7^{+0.9}_{-0.7} \text{ pb}$, and the result in this thesis is 1.7 standard deviations from the theoretical expectation. The current world average top mass is $170.9 \pm 1.8 \text{ GeV}/c^2$ [14]. For a top quark mass of $170 \text{ GeV}/c^2$ we find

$$\sigma_{t\bar{t}} = 12.9 \pm 2.2(\text{stat.}) \pm 1.0(\text{sys.}) \pm 0.8(\text{lum.})\text{pb}.$$

We have performed checks using a higher-purity tagging algorithm, and requiring two identified b -jets, and found consistent results, validating that the method is robust. Kinematic features of the candidate sample are qualitatively consistent with theoretical ex-

expectations for $t\bar{t}$ signal and background. This result is also consistent with the pretag analysis, in which no requirement is made on the flavour of the jets, 3% of the time.

As discussed in the introduction, study of the $t\bar{t}$ sample serves as a probe of possible new physics processes. Several theories exist, for example supersymmetric theories, in which the observed top quark cross-section could be different in different decay modes. Thus it is also important to measure the top quark cross-section in all three decay channels. Previous study of the top quark cross-section at both CDF and DØ have shown consistent results in all decay channels. At CDF the most recent measurement of the cross-section in lepton+jets channel is [24]

$$\sigma_{t\bar{t}} = 8.2 \pm 0.5(\text{stat.}) \pm 0.8(\text{sys.}) \pm 0.5(\text{lum.})\text{pb} ,$$

and in the all hadronic channel is [24]

$$\sigma_{t\bar{t}} = 8.3 \pm 1.0(\text{stat.})_{-1.5}^{+2.0}(\text{sys.}) \pm 0.5(\text{lum.})\text{pb} .$$

As the first measurement of the top quark cross-section using the tagged lepton plus isolated track selection, this measurement offers new insight into the top quark decays. Recall that in several measurements of a given quantity, one would expect about one third (32%) of the measurements to be more than one standard deviation from the mean. Taking into consideration this expected randomness, the result presented in this thesis is not unexpected. It also does not, in itself, indicate the existence of a non-Standard Model process. This result does however, warrant further study with more data. Additional data would verify if the current measurement represents a statistical fluctuation, or if indeed there is a non-Standard Model process present in the sample. Set against the current backdrop of previous measurements, this result is interesting.

Continuation of the study of the top quark sector at the Tevatron is useful for other reasons as well. Analysis of the top quark sector is important for searches for new physics at future colliders. $t\bar{t}$ signals need to be well understood as these will be one of the dominant backgrounds in searches for new physics at the LHC. For example, predicted

supersymmetry signatures may have multilepton, multijet final states, possibly with \cancel{E}_T from escaping dark matter candidates. In such a situation all top quark backgrounds, and the top quark decays themselves, will become background sources for at least some new physics. More generally than that, improved understanding of the mechanisms that generate \cancel{E}_T in the absence of neutrinos or dark matter candidates will be necessary to accurately estimate the contributions from known physics processes at future hadron colliders such as the LHC. Measurements of the top quark cross-section, especially in the dilepton channel, are currently the best testing ground available for our understanding of these backgrounds. Top quark cross-section measurements are crucial to our understanding of the fundamental particles and their interactions. This is a stepping stone to discovering what, if anything, will supersede the Standard Model.

This analysis is the first time at CDF that the dilepton cross-section analysis has imposed the identified b -jet requirement. This has improved the observed signal:background ratio from about 1:1 to about 5:1, although at a cost of a slightly larger statistical uncertainty due to the reduced size of the candidate sample, and slightly larger systematic uncertainty. As the Tevatron continues to deliver more proton-anti-proton collisions, and CDF builds on the over 2.0 fb^{-1} dataset it presently has, this selection will yield more precise results.

Bibliography

- [1] S. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, *Proceedings of the 8th Nobel Symposium*, ed. N. Svartholm (Almqvist and Wiksell, Stockholm), 367 (1968).
- [2] The Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Nucl. Rev. Lett. **81**, 1562-1567 (1998)
- [3] <http://public.web.cern.ch/Public/Welcome.html>
- [4] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995)
- [5] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995)
- [6] Particle Data Group, *The Review of Particle Physics*, <http://pdg.lbl.gov> (2004)
- [7] R. P. Feynman, *QED, The Strange Theory of Light and Matter*, Princeton (1985)
- [8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508-509 (1964)
- [9] A. Sirlin, *Thirty Years of Precision Electroweak Physics*, J. Phys. **G29**, 213-224 (2003)
- [10] R. K. Ellis, W. J. Stirling, B. R. Webber *QCD and Collider Physics*, Monographs on Particle Physics, Nuclear Physics and Cosmology (1996)

- [11] D. Gross, F. Wilczek, *Asymptotically Free Gauge Theories*, Phys. Rev. Lett. **30**, 1343 (1973)
- [12] G. Kane, *Modern Elementary Particle Physics*, Addison-Wesley (1993)
- [13] I. Bigi *et al.*, *Production and Decay of Ultraheavy Quarks*, Phys. Lett. **B181**, 157 (1986)
- [14] Tevatron Electroweak Working Group, *A Combination of CDF and D0 Results on the Mass of the Top Quark*, hep-ex/0703034, 2007.
- [15] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, G. Ridolfi, *The $t\bar{t}$ 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)
- [16] N. Kidonakis, R. Vogt, *Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction*, Phys. Rev. /bf D68, 114014 (2003).
- [17] C. T. Hill, S. J. Parke, *Top production: Sensitivity to new physics*, Phys. Rev. **D49**, 4454-4462 (1994).
- [18] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, Phys. Rept. **110**, 1, (1984).
- [19] H. E. Haber, K. L. Gordon, *The Search for Supersymmetry: Probing Physics Beyond the Standard Model*, Phys. Rept. **117**, 75-263, (1985).
- [20] CDF Collaboration, F. Abe *et al.* *Measurement of the top quark mass and $t\bar{t}$ production cross section from dilepton events at the Collider Detector at Fermilab*, Phys. Rev. Lett. **80**, 2779-2784 (1998).
- [21] DØ Collaboration, V. M. Abazov *et al.* *$t\bar{t}$ production cross section $p\bar{p}$ collisions at $\sqrt{1.8}$ TeV*, Phys. Rev. **D67**, 012004 (2003).

- [22] CDF Collaboration, D. Acosta *et al.* *Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{1.96}$ TeV using dilepton events*, Phys. Rev. Lett. **93**, 142001 (2004).
- [23] DØ Collaboration, V. M. Abazov *et al.* *Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions using dilepton events*, hep-ex/0706.0458v2, 2007.
- [24] CDF Collaboration, Top Group Public Results, http://www-cdf.fnal.gov/physics/new/top/public_xsection.html.
- [25] H. T. Edwards *The Tevatron Energy Doubler: A Superconducting Accelerator*, Ann. Rev. Nucl. Part. Sci. **35**, 605-660 (1985).
- [26] C. Gattuso *et al.* *Accelerator Concepts*, Fermilab Accelerator Division Documentation, http://www-bdnew.fnal.gov/operations/rookie_books/Concepts_v3.1.pdf, 2006.
- [27] C. Gattuso *et al.* *Accelerator Concepts*, Fermilab Accelerator Division Documentation, http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf, 2006.
- [28] D. Mohl, G. Petrucci, L. Thorndahl and S. Van Der Meer, *Physics and Technique of Stochastic Cooling*, Phys. Rept. **58**, 73 (1980)
- [29] *The CDF II Detector Technical Design Report*, FERMILAB-Pub-96/390E (1996); <http://www-cdf.fnal.gov/>
- [30] CDF Collaboration, A. Affolder *et al.* *Intermediate Silicon Layers Detector for the CDF Experiment*, Nucl. Instrum. Meth. **A453**, 84-88 (2000).
- [31] CDF Collaboration, A. Affolder *et al.* *CDF Central Outer Tracker*, Nucl. Instrum. Meth. **A526**, 249 - 299 (2004).
- [32] R. Wigmans *Calorimetry Energy Measurement in Particle Physics*, International Series of Monographs on Physics **107** (2000).

- [33] CDF Collaboration, D. Acosta *et al.*, *The CDF Cherenkov Luminosity Monitor*, Nucl. Instrum. Meth. A **461**, 540-544, (2001)
- [34] CDF Collaboration, D. Acosta *et al.*, *The Performance of the CDF Luminosity Monitor*, Nucl. Instrum. Meth. A **494**, 57-62, (2002)
- [35] A. Bhatti *et al.*, *Determination of the jet energy scale at the Collider Detector at Fermilab*, Nucl. Instrum. Meth. **A566**, 375 - 412, (2006).
- [36] T. Sjöstrand *et al.*, Comp. Phys. Commun. **135**, 238 (2001)
- [37] J. Pumplin *et al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07**, 012, (2002).
- [38] E.A. Gerchtein and M. Paulini, ECONF **C0303241**, TUMT005 (2003).
- [39] R. Brun *et al.*, CERN-DD/EE/84-1 (1984).
- [40] G. Grindhammer *et al.*, Nucl. Instrum. Meth. **A290**, 469 (1990).
- [41] A. D. Martin *et al.*, */em Uncertainties of predictions from parton distributions. II: Theoretical errors*, Eur. Phys. J. **C35**, 325-348, (2004).
- [42] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **64**, 032002 (2001).
- [43] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).
- [44] M. R. Whalley *HEPDATA*, *The Durham HEP Databases*,
<http://durpdg.dur.ac.uk/hepdata/pdf3.html>
- [45] H. L. *et al.*, Phys. Rev. D **55**, 1280, hep-ph/9606399 (1997);
H. L. *et al.*, Eur. Phys. J. **C 12**, 375, hep-ph/9903282 (2000)
- [46] T. Stelzer and W.F. Long, Phys. Commun. **81**, 337 (1994);
F. Maltoni, T. Stelzer, hep-ph/0208156,
<http://madgraph.hep.uiuc.edu/> (2002)

- [47] Fermi National Accelerator Laboratory, Tevatron Department,
<http://www-bdnew.fnal.gov/tevatron/>
- [48] V. Barger, R. Phillips, *Collider Physics*, Addison-Wesley, New York (1987)
- [49] <http://hepwww.rl.ac.uk/theory/seymour/herwig/>
- [50] S. Jadach, *et. al.*, CERN-TH-6793 (1992)
- [51] Application Software Group, Computing and Network Division, CERN,
<http://wwwasd.web.cern.ch/wwwasd/geant/>