A SEARCH FOR TOP QUARKS IN $\overline{P}P$ COLLISIONS USING DILEPTONS

BY

ANDREW JAMES MARTIN

B.S., Carnegie Mellon University, 1988
M.S., University of Illinois at Urbana-Champaign, 1990

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1996

Urbana, Illinois

A SEARCH FOR TOP QUARKS IN $\overline{P}P$ COLLISIONS USING DILEPTONS

Andrew James Martin, Ph.D.
Department of Physics
University of Illinois at Urbana-Champaign, 1996
Professor Tony Liss, Advisor

A search for $t\bar{t}$ pairs in $\overline{p}p$ collisions with a center of mass energy $\sqrt{s} = 1.8$TeV is presented. The analysis looks for $t\bar{t}$ pairs which decayed through the chain $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\overline{\nu}b\ell^-\nu\bar{b}$ where one of the b quarks is explicitly identified. In a sample of 109.4 ± 7.2 pb^{-1} of data, 6 tags in 4 events are found with an expected background of $1.53 \pm .33$ events. This results in an estimated production cross section of $4.6^{+4.4}_{-3.1}$ pb at $m_{top} = 175$ GeV/c^2.

I dedicate this to my parents, Francis C. and Mary E. Martin

Acknowledgements

There are many people to whom I owe a great debt in making this work possible. Without the countless hours the members of the top group on CDF have put in I could not have done this. To Wei Ming Yao, Doug Glenzinski, Richard Hughes, Brian Winer, and to many others to recall here, thanks for developing the vertex finding algorithms. To Claudio Campagnari, Avi Yagil, and Dave Kestenbaum, thanks for developing the SLT algorithm. To Rob Roser and Tom LeCompte, thanks for answering all the questions, no matter how trivial. Most importantly, thanks to my advisor Tony Liss for showing me how to take the idea and turn it into a finished analysis.

To Alain, who helped introduce me to the world of CDF code and helped get the geometry and simulation running. To Phil, Vic, Hovhannes, Randy, and many others who helped introduce me to life on CDF. To Lee Holloway, thanks for giving me the opportunity to get involved in the muon trigger, it was a lot of work but even more fun. Thanks to all those who spent many hours in the trigger room, especially Bill for making it possible to find the interesting events.

Thanks to Dave and Jerry for assembling a terrific VAX cluster and keeping in running. To Tom, Shirley, and Sue for making all the arrangements for things like travel and housing. To the engineering and technical support staffs both at the UofI and at FermiLab for designing and making a multitude of parts. Thanks to Fred for making all the gadgets which made it possible to build the chambers. To all the undergrads who helped to assemble the chambers, especially those lunch-time card players.

Special thanks to Professors Jack Ranney and Michael Cameron for the assistance in studying the physics of vibrating strings. Thanks also to all the Illinois bass players over the years whom I have had the privilege of playing with especially Jon, David, Kris,

and Catalin. Thanks to Randy, Jeremy, and Dave for the Sunday afternoons rockin' in HEPL. Also thanks to all the members of the CDF band for all the loud music and good times. Thanks to Mike and Merry Star for the friendship over the years. Thanks to Jodi violin and double bass duets just won't be the same with anyone else.

A special thank you to Val for reminding me of the importance of just enjoying an evening in the company of a good friend.

This thesis was supported in part by the United States Department of Energy under contracts DE-AC02-76ER-01195 and DE-FG02-91ER-40677 and by the Alfred P. Sloan foundation.

Table of Contents

List of Tables x

List of Figures xii

1 Introduction and Theory of $t\bar{t}$ Production 1
- 1.1 Production 2
- 1.2 Decay 4

2 The CDF Detector 8
- 2.1 Tracking Systems 8
- 2.2 Calorimetery 11
- 2.3 Muon Detection 13
- 2.4 Trigger 17
- 2.5 Luminosity Monitoring 18

3 Identifying b Quarks 19
- 3.1 Jet Clustering 19
- 3.2 Identifying b Jets 21

4 Signal Sample 23
- 4.1 Lepton Identification 23
  - 4.1.1 Electron Identification 24
  - 4.1.2 Muon Identification 26
  - 4.1.3 Isolation 30
  - 4.1.4 Charge 31
- 4.2 Optimization of Kinematic Cuts 32
  - 4.2.1 Monte Carlo Programs Used 33
  - 4.2.2 Variables to Optimize 33
  - 4.2.3 Optimization Method 39
  - 4.2.4 Results of Optimization 42

5 Backgrounds 44
- 5.1 Method 1 45
  - 5.1.1 SECVTX 45

5.1.2 SLT 48
5.1.3 Combined 52
5.2 Method 2 53
5.2.1 Mistags 54
5.2.2 Fake Primary Leptons 56
5.2.3 Drell-Yan Background 59
5.2.4 W^+W^- Background 64
5.2.5 Method 2 Total 67
5.3 Summary 67

6 Acceptance 69
6.1 Geometric Acceptance 70
6.2 Lepton Identification Efficiency 71
6.3 Kinematic Cuts 75
6.4 Trigger Efficiencies 76
6.5 b-tagging 77
6.6 Total Acceptance 78

7 Results 80

8 Conclusion 89

Appendix 93

A Design and Performance of the CMP 93
A.1 Chamber Design 94
A.2 Chamber Performance 96
A.2.1 Gas Tightness 96
A.2.2 Dark Currents 98
A.2.3 Wire Tension 98
A.2.4 Cosmic Ray Testing 100

B The CDF Muon Trigger Scheme 104
B.1 Level 1 104
B.2 Level 2 108

C Neutrino momentum 109

D Tagging Algorithms 111
D.1 SECVTX 111
D.2 SLT 115

Bibliography 117

Vita 120

List of Tables

1.1 Decay modes for a $t\bar{t}$ pair and their branching ratios assuming charged-current decays. The symbol $q\bar{q}'$ stands for $u\bar{d}$, $\bar{u}d$, $c\bar{s}$, or $\bar{c}s$. 6

2.1 Summary of CDF calorimeter resolutions. The symbol $\oplus$ signifies that the constant term is added in quadrature in the resolution. The resolutions in the electromagnetic calorimeters are for incident electrons and photons, while the hadronic calorimeter resolutions are for isolated pions. Energy is given in GeV and E_T is the projection of the energy onto the plane transverse to the beamline ($E_T = Esin\theta$). 13

4.1 Identification cuts for both tight and loose CEM electrons as well as the cuts applied in the Level 3 trigger. Note that the χ^2_{strip} cut is not applied for loose CEM electrons and the E/P and Z_{vertex} cuts are not required in Level 3. 26

4.2 CMUO muon identification cuts applied in the analysis and in the Level 3 trigger. Note that the Δx cuts are only applied if that particular detector has hits used. These cuts are used for both the loose and tight CMUO muon categories. 27

4.3 CMIO muon identification cuts applied in the analysis. Note that there is no trigger for a stiff track in Level 3. 30

4.4 Results of optimization method using the ISAJET sample with m_{top} = 160 GeV/c^2. The signal is the number of events expected after tagging in 100 pb^{-1}, the background is the expected number of background events expected in 100 pb^{-1}, and $F(\geq 3\sigma)$ is the fraction of experiments with a greater than or equal to 3σ discovery. 43

5.1 Number of tags predicted and found in a sample of Z's as a function of the number of jets with E_T>10 GeV and $|\eta| < 2.0$. Here the errors only take into account the error on the fake tag probability. Also shown is the number of jets in each bin before tagging. 47

5.2 Fake rates for isolated and non-isolated leptons derived for this analysis. Note that the CMIO category always requires isolation. 58

5.3 Expected background due to fake primary leptons for each b-tagging algorithm separately. 59

5.4 Method 2 backgrounds for the dilepton+b-tag search shown for each tagging algorithm separately and for the combination. 67

6.1 Geometrical acceptance for two leptons with $p_T >$ 20 GeV/c determined using the PYTHIA sample as a function of m_{top}. Note that the CEMU and MUCE categories are equivalent at this point. 70
6.2 Lepton identification efficiency for $m_{top} = 175$ GeV/c^2. Note that there is no overlap between the CEMU and MUCE categories After the lepton identification cuts are made. Most event which contain a tight CEM electron and a tight CMUO muon are classified in the CEMU category for historical reasons. 71
6.3 Single lepton identification efficiency measured using data and Monte Carlo Z events. The ratio of efficiencies for Data/Monte Carlo, $R(\ell)$, is also given. 75
6.4 Lepton identification and isolation efficiency as a function of m_{top}. 76
6.5 Kinematic acceptance for PYTHIA $t\bar{t}$ Monte Carlo as a function of the top quark mass. 76
6.6 Trigger efficiencies for $t\bar{t}$. 77
6.7 Total acceptance for $t\bar{t}$ events from PYTHIA $t\bar{t}$ Monte Carlo as a function of top quark mass. 78

7.1 Number of event remaining after each of the cuts in the analysis. 80
7.2 Measured cross section for $t\bar{t}$ production as a function of m_{top}. 87

List of Figures

1.1 Tree level QCD diagrams for $q\bar{q}$ annihilation (top diagram) and gluon-gluon fusion (lower diagrams) to produce a $t\bar{t}$ pair. 3
1.2 Two examples of the Feynman diagrams illustrating the $gg \rightarrow gt\bar{t}$ process which contributes to the $\mathcal{O}(\alpha_S^3)$ corrections to the $t\bar{t}$ cross section. 4
1.3 Theoretical total production cross section for $t\bar{t}$ pairs up to order α_S^3 from recent calculations. 5
1.4 Production of a $t\bar{t}$ pair through $q\bar{q}$ annihilation and the subsequent decay of the top quarks into the dilepton final state. 5

2.1 Cross section of one quadrant of the CDF detector. The nominal interaction point is in the lower left hand corner. 9
2.2 Isometric view of the CDF detector. 9
2.3 Isometric view of one half of the SVX detector. 10
2.4 Schematic view of a CEM calorimeter wedge. 12
2.5 The number of hadronic absorbtion lengths as a function of θ between the interaction region and the muon detectors at ϕ slightly greater than 0° to avoid the crack between wedges. The muon coverage extends to about 40° in θ. 14
2.6 Cross section of a CMU module. Each module subtends an angle of 5° in ϕ. 15
2.7 Cross section of a pair of CMP stack showing half-cell stagger. 15
2.8 CMX wedge view. 16

3.1 Examples of QCD production of heavy quarks Q. The upper left diagram is flavor excitation, the upper right is direct production, and the bottom diagram is gluon splitting. 22

4.1 Distributions of several of the reconstructed variables used to select CEM electrons. The vertical lines are located where the cuts are made. The plots are made from the second electron in $Z \rightarrow e^+e^-$ decays. 25
4.2 Distributions for some CMUO muon identification variables made from the second CMUO in $Z \rightarrow \mu^+\mu^-$ events. The vertical lines are located where the cuts are made. 28

4.3 Distributions for CMIO muon identification variables taken from the second muon in $Z \rightarrow \mu^+\mu^-$ events with one tight CMUO type muon and one CMIO type muon. The vertical lines are located where the cuts are made. 29
4.4 I_{trk} vs. I_{cal} for leptons from $b\bar{b}$ production (left) and from $t\bar{t}$ production (right). The $b\bar{b}$ sample was generated using a modified version of ISAJET which increases the efficiency for producing semi-leptonic decays of the b quark. The $t\bar{t}$ sample was generated using standard ISAJET with $m_{top} =$ 160 GeV/c^2. 32
4.5 Invariant mass spectrum from a Drell-Yan and Z boson Monte Carlo sample generated using PYTHIA after only the lepton identification cuts. . . 34
4.6 Invariant mass spectrum from a $t\bar{t}$ Monte Carlo sample generated using ISAJET after only the lepton identification cuts. The solid histogram was generated m_{top} = 160 GeV/c^2 while the dashed histogram used m_{top} = 180 GeV/c^2. This plot only shows events categorized as either ee or $\mu\mu$. . 35
4.7 $\not\!\!E_T$ spectrum from ISAJET $t\bar{t}$ dilepton events with m_{top} = 160 GeV/c^2 (solid) and m_{top} = 180 GeV/c^2 (dashed). 36
4.8 $\not\!\!E_T$ spectrum from W^+W^- production (dashed) and Drell-Yan plus Z production (solid), both generated using PYTHIA. The relative normalization of the plots is arbitrary, it is the shape which matters. These plots are after the lepton identification cuts have been made. 36
4.9 $\not\!\!E_T$ vs. $\Delta\phi(\not\!\!E_T,\ell)$ for $Z \rightarrow \tau^+\tau^-$ events. 37
4.10 $\not\!\!E_T$ vs. $\Delta\phi(\not\!\!E_T,\ell \text{ or } jet)$ for $t\bar{t}$ events generated using ISAJET with m_{top} = 160 GeV/c^2. 38
4.11 E_T spectrum of the two leading jets in $t\bar{t}$ events with $|\eta| < 2.0$. These events were generated using ISAJET with m_{top} = 160 GeV/c^2 (solid) m_{top} = 180 GeV/c^2 (dashed). 38
4.12 E_T spectrum of the two leading jets in Drell-Yan (solid) and W^+W^- (dashed) events generated using PYTHIA. The normalization is arbitrary, the important feature is the shape. 39

5.1 SECVTX tag rates in inclusive jets as a function of jet E_T and N_{SVX} using the parameterization. The particular E_T and N_{SVX} bins shown were arbitrarily chosen. 46
5.2 Ratio of the number of $+L_{xy}$ tags observed to the number predicted in the $\sum E_T$ sample as a function of various quantities. $|\Delta\phi_{min}|$ (ΔR_{min}) is the minimum $\Delta\phi$ (ΔR) to the closest jet, and Δz is the distance along the beamline between two good primary vertices. 49
5.3 Tagging rate parameterizations for the SLT algorithm. The CMU muon, CMX muon and CEM electron rates are shown separately. 51
5.4 ΔR distribution between a lepton and the nearest jet in $\eta - \phi$ for a $b\bar{b}$ Monte Carlo sample generated using ISAJET. The leptons are required to have $p_T > 15$ GeV/c. 53

5.5 Ratio of the number of $-L_{xy}$ tags observed to the number predicted in the $\sum E_T$ sample as a function of various quantities. $|\Delta\phi_{min}|$ (ΔR_{min}) is the minimum $\Delta\phi$ (ΔR) to the closest jet, and Δz is the distance along the beamline between two good primary vertices. 55
5.6 Comparison of the E_T^1, E_T^2, $\not\!\!E_T$, and N_{jets} spectra between PYTHIA Drell-Yan Monte Carlo and the data. Only the $\not\!\!E_T$ spectrum show much disagreement. 61
5.7 Comparison of invariant mass of two leptons between data and PYTHIA Drell-Yan Monte Carlo as a function of the number of jets with $E_T >$ 10 GeV and $|\eta| < 2.0$. The arrows mark the region cut by the invariant mass cut. 62
5.8 Comparison of invariant mass of two leptons between data and PYTHIA Drell-Yan Monte Carlo as a function of the number of jets with $E_T >$ 10 GeV and $|\eta| < 2.0$. For this figure, the Monte Carlo is normalized in the mass range 40 GeV/c^2 $< M_{\ell\ell} <$ 80 GeV/c^2 which is between the arrows. 63
5.9 Comparison of the E_T^1, E_T^2, $\not\!\!E_T$, and N_{jets} spectra for the ISAJET and PYTHIA W^+W^- Monte Carlos. 66

6.1 Plots showing $(N^i_{MC} - N^i_{data})/N_{data}$ which emphasize the discrepancy between data and Monte Carlo for some of the lepton identification variables for CEM electrons and CMUO muons. 72
6.2 Invariant mass distributions for ee pairs (left) and $\mu\mu$ pairs (right) used to calculate the single lepton identification efficiency. One lepton passes the tight lepton identification cuts while the second leg has only fiducial, an opposite charge, and p_T>20 GeV/c or E_T<20 GeV cuts applied. . . . 73

7.1 CDF event displays showing end view and lego plot of the calorimeter for Run 41540 Event 127085. 81
7.2 CDF event displays showing end view and lego plot of the calorimeter for Run 57621 Event 45230. 82
7.3 CDF event displays showing end view and lego plot of the calorimeter for Run 63700 Event 272140. 83
7.4 CDF event displays showing end view and lego plot of the calorimeter for Run 69808 Event 639398. 84
7.5 Comparison of measured $t\bar{t}$ production cross section vs theory. Note that the errors are highly correlated. 88

8.1 Comparison of the cross section from this analysis with those from other measurements. The CDF combined cross section does not include this analysis which is highly correlated with other CDF analyses. Also the D$\emptyset$ cross section is calculated at $m_{top} = 180$ GeV/c^2. 90

8.2 The point is located with at the current values of m_W and m_{top} with the uncertainties being those expected in Run II. The curves are from a calculation of the dependence of the mass of the W boson on the mass of the top quark and Higgs boson[36] . The uncertainty bands on the curves are obtained from the uncertainty on $\Delta\alpha(M_Z^2)$ ($\pm.0004$). 92

A.1 CMP stack in an exploded view showing the endplate and wire support. The scale is not exactly correct in this figure. 94
A.2 Electric field lines (top) and equipotential lines (bottom) inside a CMP chamber. 95
A.3 CMP preamp schematic. 97
A.4 CMP dark current distribution at anode voltage of 5600 V measured using good stacks. 98
A.5 Measured CMP wire tensions. The nominal tension is 250 g. 99
A.6 CMP hit residual distribution for a single chamber. 100
A.7 CMP hit resolutions measured using good stacks. 101
A.8 CMP single hit efficiency measured for a typical chamber. The chamber position is the distance between the wire and extrapolated/interpolated track. 101
A.9 CMP hit efficiency measured near the wire support. The position is the distance along the wire from the wire support to the extrapolated/interpolated track. 102
A.10 CMP hit efficiency measured near the end of the chamber. The chamber position is the distance from the end of the chamber to the extrapolated/interpolated track position. 102
A.11 CMP drift velocity distribution measured in a typical chamber. The mean is expected to fall between 45 and 50 mm/μs. 103

B.1 Organization of the CDF central muon trigger system. 105
B.2 Cross section of a CMU module showing the different drift times for chambers in alternate layers for a non-radial track. 106

D.1 Transeverse momentum spectrum of b quarks from top decays in $t\bar{t}$ Monte Carlo generated with PYTHIA at $m_{top} = 175$ GeV/c^2. 112
D.2 Diagram showing both negative (top) and positive (bottom) L_{xy} vertices. 113
D.3 L_{xy} distribution for tagged jets selected with a 50 GeV trigger threshold. 114

Chapter 1

Introduction and Theory of $t\bar{t}$ Production

The Standard Model[1] of high energy physics consists of a unified theory of the electroweak interactions plus Quantum ChromoDynamics (QCD). The electroweak interaction is itself a combination of the electromagnetic interactions as described by Quantum ElectroDynamics (QED) and the theory of the weak interaction which governs processes such as nuclear β-decay. The Standard Model contains twelve fermions, six quarks and six leptons. The quarks: up, down, charm, strange, top, and bottom (u, d, c, s, t, b) interact through both the electroweak interaction and QCD while the charged leptons: electron, muon, tau (e, μ, τ) and their associated neutrinos, only interact via the electroweak interaction. The quarks and leptons are further divided into the three generations shown below with the lightest on the left and the heaviest on the right.

$$\textit{Leptons} \qquad \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$$

$$\textit{Quarks} \qquad \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$$

In this arrangement, the two leptons or quarks which are grouped together form a weak isospin doublet.

The Standard Model also provides a description of the interactions between the fermions listed above. In the Standard Model particles interact by exchanging a gauge boson. The electromagnetic interactions are governed by the exchange of a photon (γ) between electrically charged particles, which includes all the quarks and leptons except the neutrinos. The weak interaction is described by the exchange of a massive vector gauge boson, $W^{\pm}$ or Z^0, which couples primarily to members of one of the quark or lepton doublets above. The quarks also interact via the exchange of a gluon (g) which is the mediator of the QCD color interaction. A feature of the QCD interaction is that all observed particles must have a no net color charge. This implies that the bare quarks and gluons are not observable, only combinations with no net color, called hadrons, are observable.

The Standard Model has been very successful over the last 20 years at describing the observations made at many detectors around the world operating at many different energies. All of the fermions and bosons had been found except the elusive top quark, whose mass is not directly predicted by the Standard Model. Recently even this has changed as the first observation of $t\bar{t}$ pairs was announced[2, 3]. These papers represent a huge effort and are the beginning of the difficult work needed to fully explore the top quark. These papers contain descriptions of some of the methods used to search for the top quark, the algorithms used to measure it's mass, and the first measurements of the production cross section. The analysis presented here is an extension of one of the analyses presented in Reference [2].

1.1 Production

At the Tevatron, protons and anti-protons are collided together at a center of mass energy ($\sqrt{s}$) of 1.8TeV. The proton (anti-proton) is the combination of three quarks (anti-quarks) and gluons which together are called partons. The partons inside the proton and

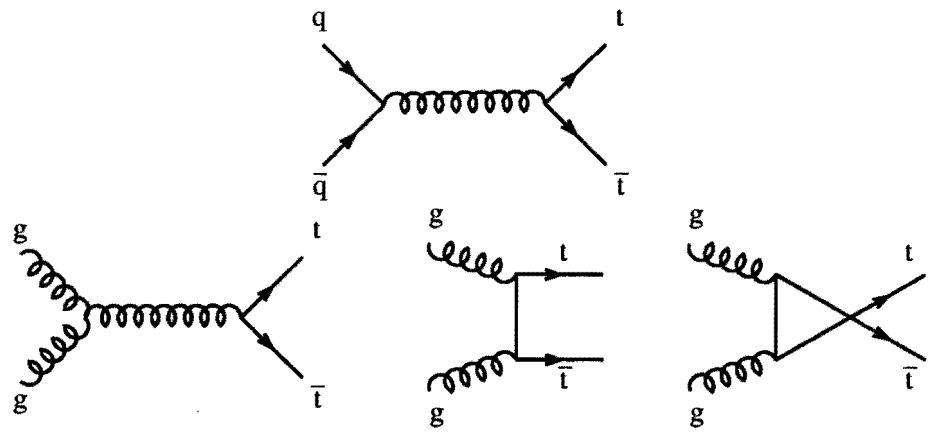

Figure 1.1: Tree level QCD diagrams for $q\overline{q}$ annihilation (top diagram) and gluon-gluon fusion (lower diagrams) to produce a $t\overline{t}$ pair.

anti-proton combine to carry the observed macroscopic properties (momentum, charge, ...) of the proton and anti-proton. It is these partons which actually scatter off of each other to produce the $t\overline{t}$ pairs.

The total $t\overline{t}$ production cross section can be written as[4]:

$$\sigma(s) = \sum_{i,j} \int dx_1 dx_2 \hat{\sigma_{ij}}(x_1 x_2 s, m_{top}^2, \mu^2) F_i^p(x_1, \mu) F_j^{\overline{p}}(x_2, \mu). \tag{1.1}$$

The functions F_i are the parton distribution functions which describe the probability of finding the ith parton inside the p ($\overline{p}$) with fraction x_i of the p's ($\overline{p}$'s) momentum. The parton-parton scattering cross section calculated in QCD is $\hat{\sigma}$. The renormalization and factorization scale is given by μ which is known to be of order the mass of the top quark.

The tree level diagrams which contribute to the first term in the perturbative expansion for $\hat{\sigma}$ are those due to $q\overline{q}$ annihilation and gluon-gluon fusion (see Figure 1.1). This is not the whole story though as the leading order corrections are known to be large[4]. This is based on the observation that the process $gg \rightarrow gt\overline{t}$, where one of the gluons has split into a $t\overline{t}$ pair (see Figure 1.2 for example Feynman diagrams), can be as important as the lowest order cross section, even though it is of order α_S^3. This occurs because at lowest order the $gg \rightarrow gg$ cross section is about 100 times the $gg \rightarrow t\overline{t}$ cross section.

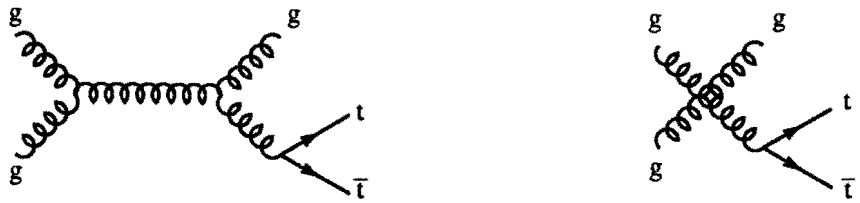

Figure 1.2: Two examples of the Feynman diagrams illustrating the $gg \rightarrow gt\bar{t}$ process which contributes to the $\mathcal{O}(\alpha_S^3)$ corrections to the $t\bar{t}$ cross section.

This large gluon production cross section combined with the rate of gluons fragmenting into a heavy quark pair, which goes like $\alpha_S(m_t^2)/2\pi$, means that this process is still competitive numerically with the tree level processes. The next-to-leading order corrections (order α_S^3) have been calculated by several different groups[4, 5, 6, 7, 8] using different techniques to calculate the soft gluon contribution. The total production cross section of $t\bar{t}$ pairs up to order α_S^3 as a function of top mass is plotted in Figure 1.3 as calculated in References [6], [7], and [8].

1.2 Decay

Within the framework of the Standard Model, the top quark decays into a real W boson and a real b quark nearly 100% of the time if the mass of the top quark is greater than the sum of the W and b masses (see Figure 1.4). Because the quarks themselves carry color and are thus not observable, each b quark pulls $q\bar{q}$ pairs out of the vacuum such that hadrons are formed. This process is called "hadronization". These hadrons follow the same basic direction as the initial quark and collectively are called a "jet" which is observed. The two W bosons each decay independently into either a light quark-antiquark pair or a lepton neutrino pair. The possible final states are given in Table 1.1 with the branching ratios found by assuming the Standard Model decay rates.

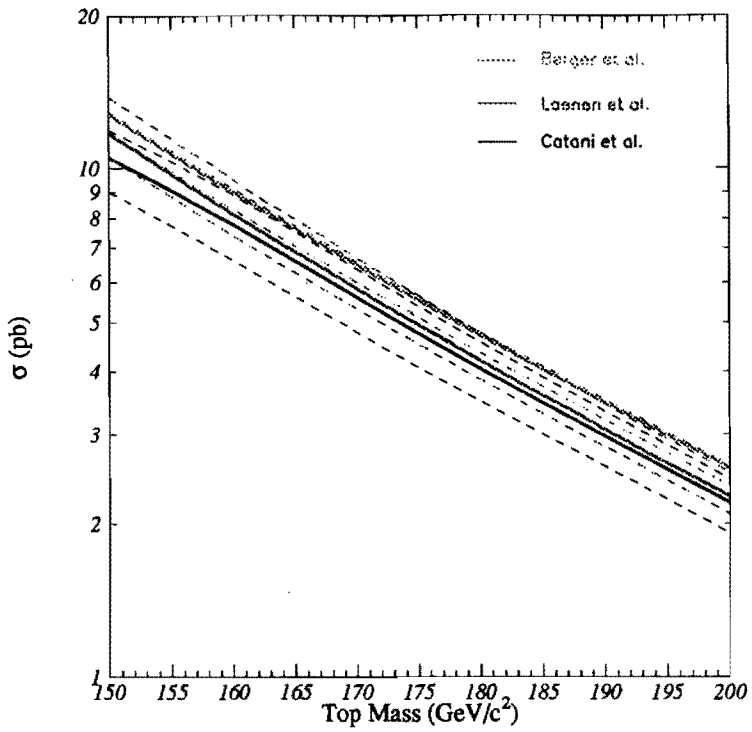

Figure 1.3: Theoretical total production cross section for $t\bar{t}$ pairs up to order α_S^3 from recent calculations.

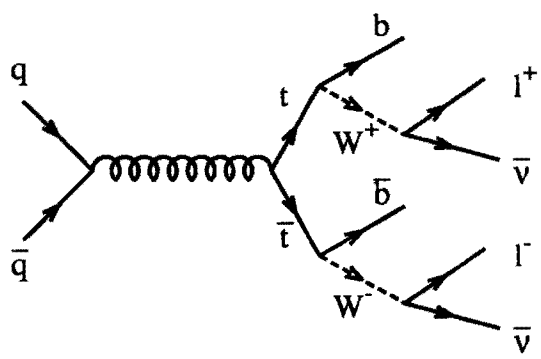

Figure 1.4: Production of a $t\bar{t}$ pair through $q\bar{q}$ annihilation and the subsequent decay of the top quarks into the dilepton final state.

Decay mode	Branching ratio
$t\bar{t} \longrightarrow q\bar{q}'bq\bar{q}'b$	36/81
$t\bar{t} \longrightarrow q\bar{q}'be\nu\bar{b}$	12/81
$t\bar{t} \longrightarrow q\bar{q}'b\mu\nu\bar{b}$	12/81
$t\bar{t} \longrightarrow q\bar{q}'b\tau\nu\bar{b}$	12/81
$t\bar{t} \longrightarrow e\nu b\mu\nu\bar{b}$	2/81
$t\bar{t} \longrightarrow e\nu b\tau\nu\bar{b}$	2/81
$t\bar{t} \longrightarrow \mu\nu b\tau\nu\bar{b}$	2/81
$t\bar{t} \longrightarrow e\nu be\nu\bar{b}$	1/81
$t\bar{t} \longrightarrow \mu\nu b\mu\nu\bar{b}$	1/81
$t\bar{t} \longrightarrow \tau\nu b\tau\nu\bar{b}$	1/81

Table 1.1: Decay modes for a $t\bar{t}$ pair and their branching ratios assuming charged-current decays. The symbol $q\bar{q}'$ stands for $u\bar{d}$, $\bar{u}d$, $c\bar{s}$, or $\bar{c}s$.

The all hadronic decay channel of the $t\bar{t}$ pair, where both W's decay into a $q\bar{q}'$ pair, has the largest branching fraction (36/81). However the QCD multijet cross section, which is the main background, is about 4-5 orders of magnitude larger than the $t\bar{t}$ cross section[9]. To see a respectable signal in this channel requires cutting hard on the kinematics of the event and requiring a jet consistent with having come from a b quark, which is called a b-tag. The lepton+jets decay modes (here 'lepton' includes only e or μ) have the next highest branching fraction (24/81). Again there is a large background due to QCD production of W+jets. To get a reasonable signal, either kinematic cuts or b-tagging information must be used. The τ decays into either $e\nu\bar{\nu}$ or $\mu\nu\bar{\nu}$ are implicitly included in searches involving electrons and muons while the hadronic τ decays are very difficult to differentiate from a QCD jet which fragments into just a few particles. As a result of this QCD background, hadronic τ decays are very difficult to include in search analyses.

This leaves the dilepton final states (again only considering the e and μ) which have a relatively small branching fraction (4/81) but have the advantage of having much lower backgrounds than the all hadronic, τ, and lepton+jets channels. The main backgrounds to be considered here are Drell-Yan production ($q\bar{q} \rightarrow \ell^+\ell^-$ including the Z peak), W^+W^- production, direct $b\bar{b}$ production, and events where a hadron mimics a lepton.

In this analysis at least one jet is required to be consistent with having come from a b quark. In this case the backgrounds are those listed above, but in addition there must be a jet in the event, typically a gluon jet, which is (mis)identified as a b-jet. This allows the backgrounds to be reduced without cutting too hard on the kinematics of the $t\bar{t}$ system resulting in an observable signal.

Chapter 2

The CDF Detector

The Collider Detector at Fermilab (CDF)[10] is a general purpose particle detector designed to record the results of $\overline{p}p$ interactions. The original detector was designed to have azimuthal symmetry as well as forward-backward symmetry (see Figures 2.1 and 2.2). Before data taking started in 1992, several major upgrades to the detector were performed including additional muon detectors, a new vertex time projection chamber and a high resolution silicon microstrip detector.

CDF uses a right-handed coordinate system with the z direction in the direction of the proton beam. The y axis is chosen to be up and the x axis points radially out from the beamline. The azimuthal angle ϕ is defined as usual, as is the polar angle θ. However, pseudorapidity, $\eta = -\ln\tan(\theta/2)$, is usually used instead of θ because at relativistic energies differences in pseudorapidity are approximately Lorentz invariant.

2.1 Tracking Systems

The tracking detectors are the innermost detectors in CDF. Moving outward from the interaction point, a particle first meets the thin beryllium beampipe, followed by the silicon vertex detector (SVX), then the vertex time projection chamber (VTX), and finally the central tracking chamber (CTC). These three tracking detectors are designed to work together to provide high resolution on both the momentum of charged particles

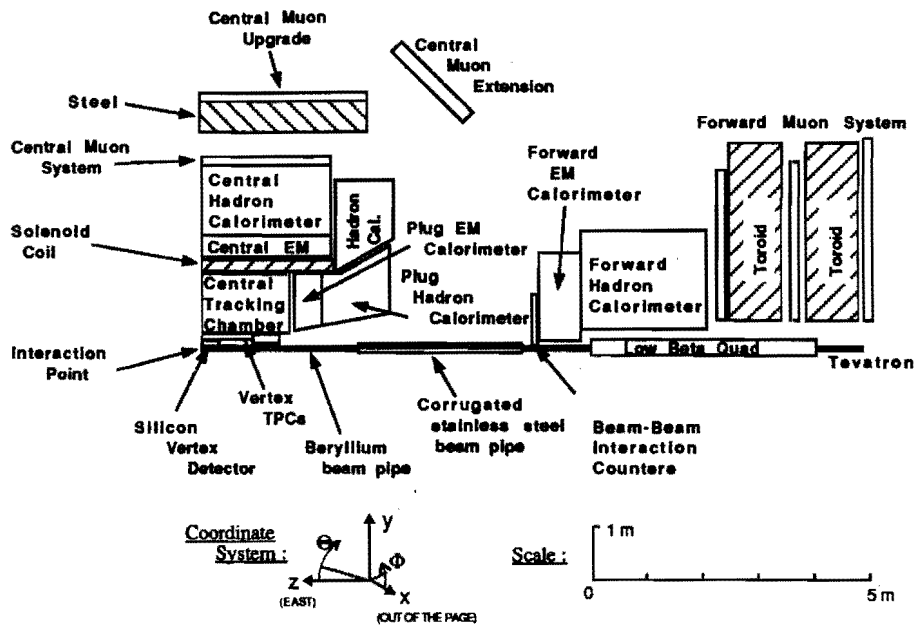

Figure 2.1: Cross section of one quadrant of the CDF detector. The nominal interaction point is in the lower left hand corner.

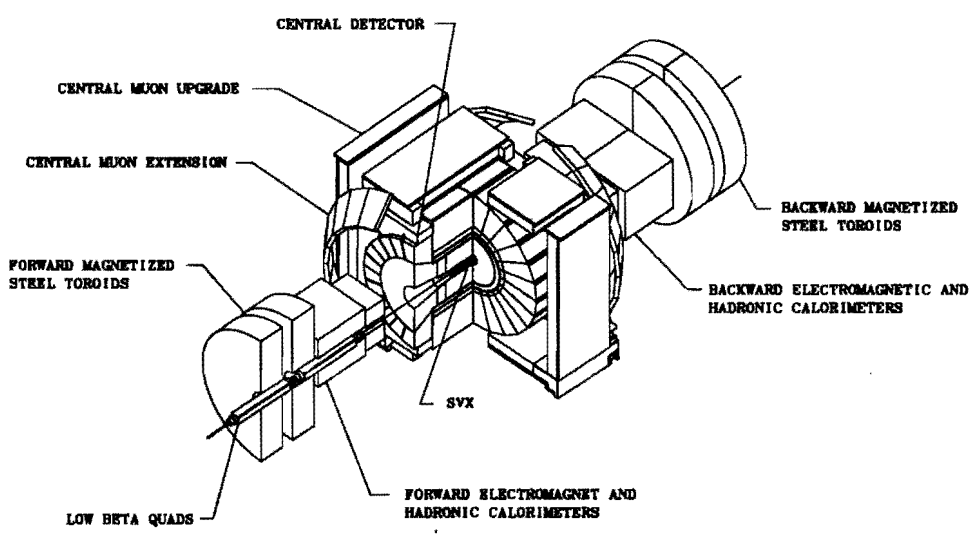

Figure 2.2: Isometric view of the CDF detector.

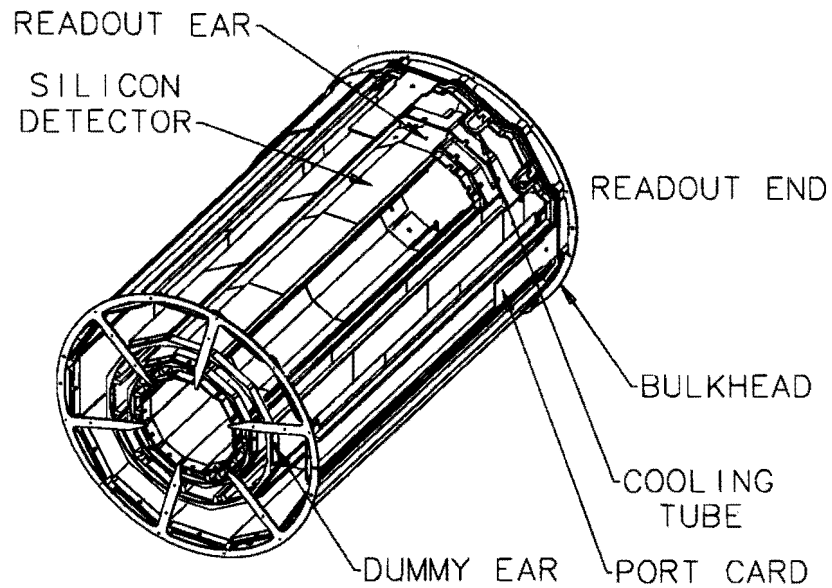

Figure 2.3: Isometric view of one half of the SVX detector.

and on the vertex position. To obtain a momentum measurement, the tracking systems are all contained in a superconducting solenoid of radius 1.5 m with an axial field of approximately 1.4 T. As a result of this field, charged particles will follow a helical trajectory which can be reconstructed. Measuring the curvature of the reconstructed track is equivalent to measuring the momentum of the particle transverse to the beamline, p_T. Measuring the θ of the track, allows the full momentum vector, $\vec{\mathbf{p}}$, of the particle to be reconstructed.

Immediately outside the the beampipe is a four layer silicon microstrip detector (SVX)[11]. The four layers are located at radii of 3.0, 4.2, 5.7, and 7.9 cm from the center of the beampipe (see Figure 2.3). The SVX provides r-ϕ information which gives a high resolution on the impact parameter of a reconstructed track. The precise information obtained using the SVX makes it possible to find tracks which do not come from

the primary vertex and to use those tracks to search for evidence of long lived particles such as b hadrons (see Section D.1).

The VTX, which is located between the CTC and the SVX, provides r-z information out to a radius of 22 cm with $|\eta| < 3.25$. This information combined with the SVX provides an excellent measure of the beam position within the CDF detector on a run by run basis. For each individual event the combined VTX and SVX detectors provide precision information about the location of the primary vertex.

Most of the tracking measurements start with the reconstruction of a track in the central tracking chamber (CTC)[12] which is a 3.2 m long drift chamber. This chamber has 84 layers of sense wires organized into 5 axial and 4 stereo "superlayers". The axial superlayers each have 12 wires parallel to the z axis which are used to measure the r-ϕ position of the track. Each stereo layer consists of 6 wires with a $\pm 3°$ stereo angle which allow a measurement of the r-z of the track. The combination of these measurements enables the reconstruction of tracks in three dimensions. The momentum of a charged particle as measured by the CTC has a resolution of $\delta p_T/p_T = [(0.0020 p_T)^2 + (0.0066)^2]^{1/2}$ where p_T is given in GeV/c. When used in conjunction with the SVX, the momentum resolution of the track improves to $\delta p_T/p_T = [(0.0009 p_T)^2 + (0.0066)^2]^{1/2}$.

2.2 Calorimetery

Surrounding the tracking systems at CDF are the calorimeters which cover 2π in azimuth and out to ± 4.2 in η. The CDF calorimeters are divided into three distinct regions in η: the central region covers $|\eta| < 1.1$, the plug covers $1.1 < |\eta| < 2.4$, and the forward calorimeters in the region $2.2 < |\eta| < 4.2$. Each region consists of an electromagnetic calorimeter with a lead absorber followed by an hadronic compartment which uses iron as the absorber. Also each detector is segmented in both η and ϕ into a projective tower geometry where each tower points back to the nominal interaction point.

The central region consists of the central electromagnetic calorimeter (CEM)[13] followed by the central hadronic (CHA) and wall hadronic calorimeters (WHA)[14] all of

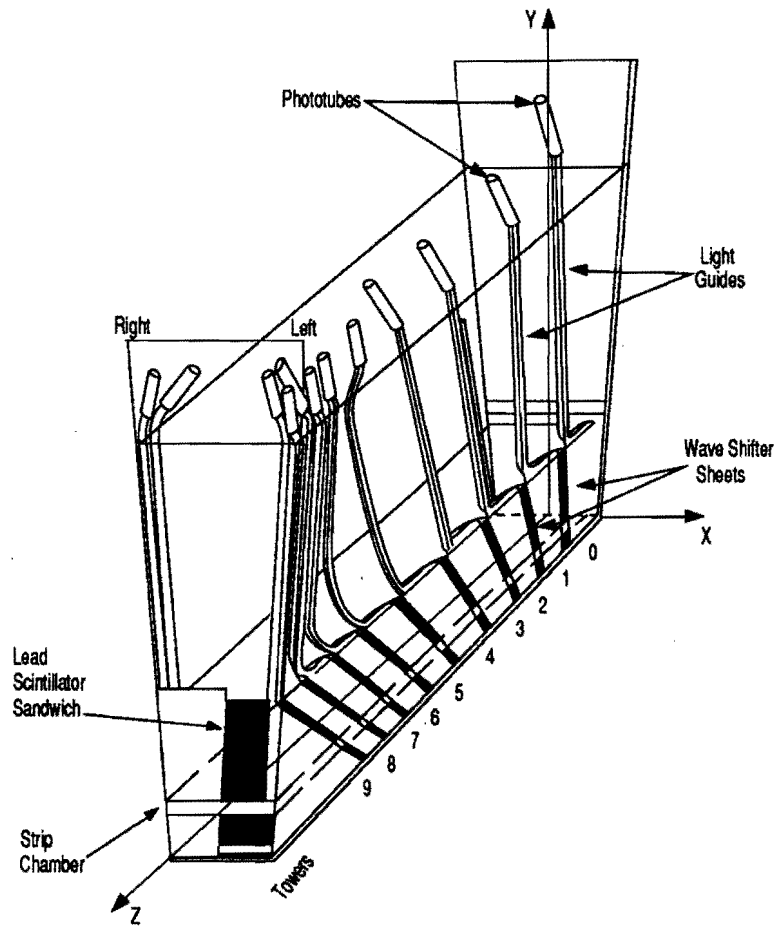

Figure 2.4: Schematic view of a CEM calorimeter wedge.

which use a plastic scintillator as the active sampling material (see Figure 2.4). These detectors are segmented into 24 "wedges", 15° in ϕ by 0.1 in η. The energy resolution ($\delta E/E$) of the calorimeters is summarized in Table 2.1. The central strip chambers (CES), located at 6.3 radiation lengths into the CEM which is approximately shower maximum for electrons, are a set of proportional chambers with both strip and wire readout to provide θ and ϕ information. The CES provides information about the position of the shower and it's transverse development. The CES has a position resolution of roughly ± 2 mm for electrons with a momentum of 50 GeV/c. A central pre-radiator (CPR), which is another set of proportional chambers located between the solenoid and the CEM, provide a measurement of the shower's early development in the material of the

System	Energy Resolution	Thickness
CEM	$13.7\%/\sqrt{E_T} \oplus 2\%$	$18X_0$
PEM	$22\%/\sqrt{E} \oplus 2\%$	$18 - 21X_0$
FEM	$26\%/\sqrt{E} \oplus 2\%$	$25X_0$
CHA	$50\%/\sqrt{E_T} \oplus 3\%$	$4.5\lambda_0$
WHA	$75\%/\sqrt{E} \oplus 4\%$	$4.5\lambda_0$
PHA	$106\%/\sqrt{E} \oplus 6\%$	$5.7\lambda_0$
FHA	$137\%/\sqrt{E} \oplus 3\%$	$7.7\lambda_0$

Table 2.1: Summary of CDF calorimeter resolutions. The symbol $\oplus$ signifies that the constant term is added in quadrature in the resolution. The resolutions in the electromagnetic calorimeters are for incident electrons and photons, while the hadronic calorimeter resolutions are for isolated pions. Energy is given in GeV and E_T is the projection of the energy onto the plane transverse to the beamline ($E_T = Esin\theta$).

solenoid. Also the CHA and WHA are equipped with TDC's which provide a measure of the time between the beam crossing and the particles passing through the hadron calorimeter.

The plug and forward calorimeters[15] use gas based proportional chambers as the sampling medium. The PEM and PHA are the electromagnetic and hadronic calorimeters in the plug region while the FEM and FHA are the same for the forward region. All of these calorimeters are segmented into towers 5° in ϕ by 0.1 in η. In the PEM, electromagnetic shower positions are measured using θ and ϕ-oriented strips giving a position resolution of 0.2 cm by 0.2 cm. The FEM uses cathode pads and anode wires to provide a position resolution which varies from about 1-4 mm depending on the location in the calorimeter.

2.3 Muon Detection

The muon detectors are the farthest out from the interaction region. There are three sets of central muon detectors at CDF in addition to the forward muon system (FMU). (The latter is not used in this analysis and will not be discussed further.) The central

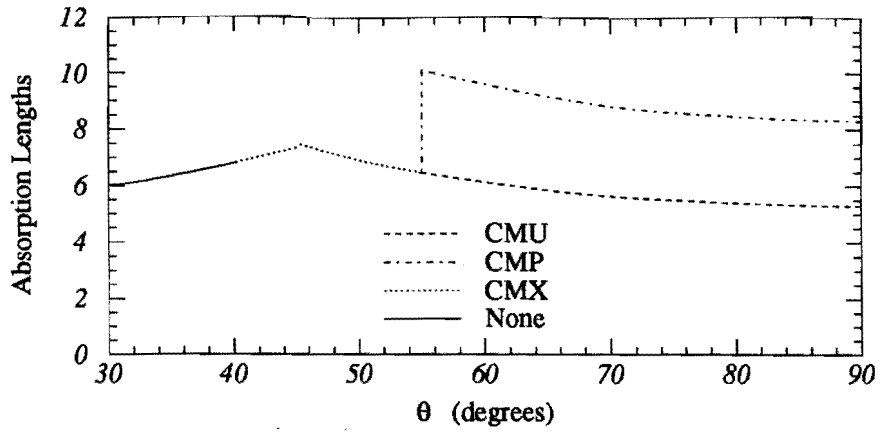

Figure 2.5: The number of hadronic absorbtion lengths as a function of θ between the interaction region and the muon detectors at ϕ slightly greater than 0° to avoid the crack between wedges. The muon coverage extends to about 40° in θ.

muon chambers (CMU)[16] were built and installed for previous runs while the central muon upgrade (CMP) and the central muon extension (CMX) were new detectors in this run[17].

The CMU and CMP chambers cover basically the same ranges in η ($|\eta| < 0.6$) and ϕ (2π). However, they are separated by $\sim$ 60 cm of steel which reduces the background in a muon sample by absorbing many of the hadrons which exit the back of the CHA and reach the CMU. The number of hadronic absorbtion lengths between the interaction region and the muon detectors is shown as a function of θ in Figure 2.5. The CMX extends the muon coverage from where the CMU ends, at $|\eta| = 0.6$ out to $|\eta| = 1.0$.

The CMU chambers (see Figure 2.6) consist of four layers of drift cells whose wires are aligned along the z axis. Alternating layers of wires in the CMU chamber are radially aligned to provide a rough momentum measurement which is used in the trigger (see Section 2.4). The CMU system is divided into the same 15° "wedges" in ϕ as the central calorimeter and is divided at $\eta = 0$. Gaps in ϕ between the wedges reduce the CMU coverage to about 84% of the solid angle for $|\eta| < 0.6$. Tracks can be reconstructed in the CMU using the measured drift times and the known time-to-distance relationship

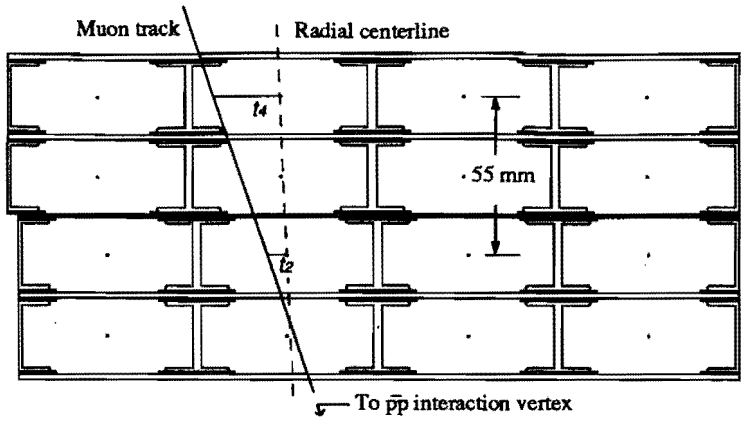

Figure 2.6: Cross section of a CMU module. Each module subtends an angle of 5° in ϕ.

in the drift direction. The use of a slightly resistive wire allows measurements of the z position of the track by comparing the amount of charge collected at each end of the wire. Resolutions of 250 μm in the r-ϕ direction and 1.2 mm in the z direction are found from cosmic ray studies.

The CMP also uses four layers of drift tubes but in this case a half-cell staggered geometry is used (see Figure 2.7). These stacks are mounted in a rectangular box around

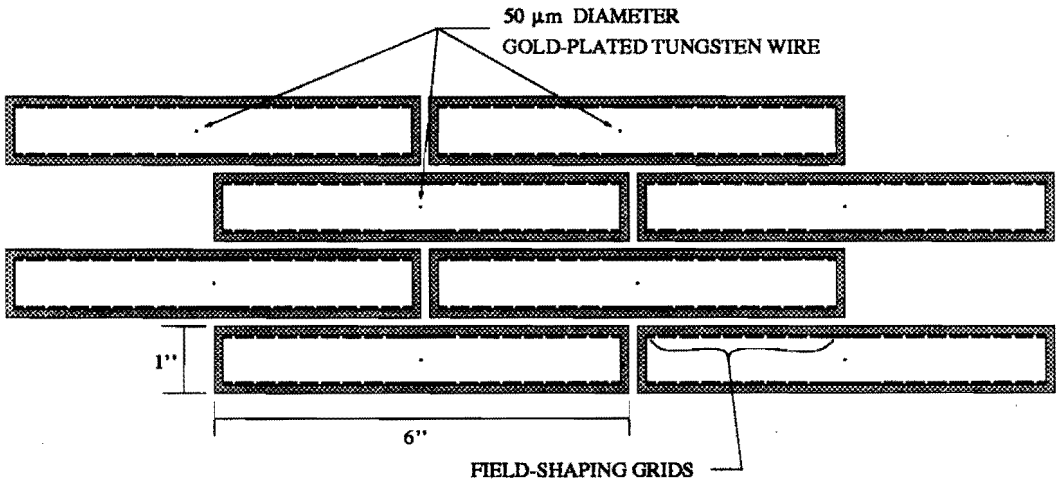

Figure 2.7: Cross section of a pair of CMP stack showing half-cell stagger.

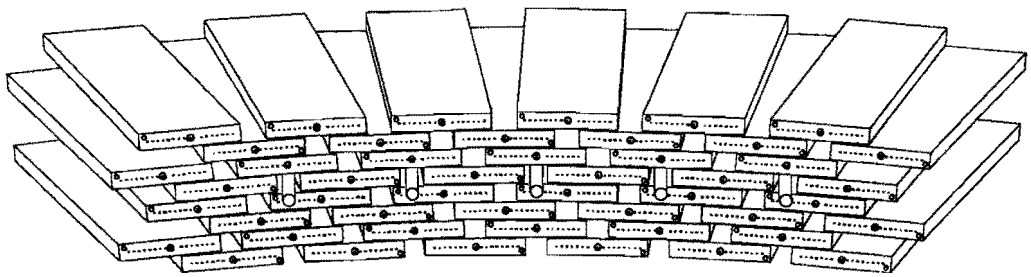

Figure 2.8: CMX wedge view.

the central detector. This results in a non-trivial $\eta - \phi$ coverage of the chambers. For the region $|\eta| < 0.6$, CMP covers 63% of the solid angle where most of the lost coverage is due to gaps in the magnet return yoke and CMU and CMP overlap in 53% of the solid angle. The acceptance lost by requiring CMP in conjunction with the CMU is outweighed by the gain in signal quality. (See Appendix A for more details on the CMP.)

The CMX uses a chamber design like that used in the CMP but the chambers are mounted in a projective geometry more like CMU (see Figure 2.8). The CMX wires are radially aligned with the interaction point allowing a simple relationship between the track angle and the particle's p_T. The chambers are arranged in a conical geometry which make possible position measurements in both the r-ϕ and r-z directions. The r-z measurement is possible because there are eight layers of chambers mounted in a conical shape resulting in wires which are not parallel and giving an effective stereo angle. Knowing the geometry makes it possible to reconstruct a track in all three spatial dimensions. CMX covers about 71% of the solid angle for $0.6 < |\eta| < 1.0$. Again the lost coverage is due to large gaps in ϕ due to obstructions. A layer of scintillation counters (CSX) were installed on both the inside and outside of the CMX chambers. These provide timing information to help remove backgrounds in the CMX.

2.4 Trigger

At the Tevatron, there is one $\overline{p}p$ beam crossing every 3.5 μs, which results in a crossing rate of over 285 kHz. The fraction of these which have an interaction depends upon the instantaneous luminosity. Meanwhile, the ability to write events to tape within the CDF environment is only a few Hz. Additionally, The process of reading the data out from the detector components takes about 3-5 ms which is "dead time" during events cannot be observed. To provide the reduction in rate needed, CDF uses a three level trigger system[18] which selects the most interesting events at each level for further processing and analysis. The trigger is designed to keep the detector ready to record an event for as many beam crossings as possible. The trigger at each of the three levels is the logical OR of a number of triggers designed to select various kinds of objects (muons, electrons, jets).

The lowest level trigger, Level 1, uses fast inputs from the calorimeters and the muon systems to make a decision in less than 3.5 μs resulting in no deadtime for the Level 1 trigger. The calorimeter based triggers, such as electron and jet triggers, all start off looking for a single tower with E_T greater than a threshold which can be different for each region of the calorimeter. The Level 1 muon trigger looks for the presence of hits in the muon chambers which appear to have come from a muon with a p_T above a programmable threshold (see Appendix B). If for a given beam crossing the Level 1 trigger for any subsystem is present the next level of trigger processing begins, otherwise the detector electronics is reset in preparation for the next beam crossing. The Level 1 accept rate is approximately 1 kHz.

The Level 2 trigger does more extensive processing of the data which takes less than about 10 μs. A hardware track finder, the central fast tracker (CFT)[19], processes hits in the CTC to find 2 dimensional r-ϕ tracks with a transverse momentum resolution of $\frac{\delta p_T}{p_T} = 0.035 \times p_T$ and an efficiency of $93.5 \pm 0.3\%$ for tracks with $p_T > 10$ GeV/c. If a CFT track extrapolates to a Level 1 muon trigger and the p_T measured by the CFT is greater than a programmed threshold, a Level 2 muon trigger is generated. For the jet triggers, a

hardware cluster finder looks for contiguous towers in the calorimeter with an E_T above some low E_T threshold. If the total E_T of the towers is above a programed threshold, a Level 2 jet trigger is generated. To look for electrons, the hardware cluster finder looks for energy clusters using only the electromagnetic calorimeters. In the central region, a CFT track is required to extrapolate to the same phi as the electromagnetic cluster for a Level 2 trigger. Outside the central region no track is required for the Level 2 electron triggers. If there is a Level 2 trigger from any subsystem, the data is readout before the detector electronics is reset to prepare for another event. The Level 2 accept rate is on the order of 12 Hz.

The last online trigger system, Level 3, is a software based trigger running on a farm of 8 commercially available Silicon Graphics multi-processor computers each with 8 processors installed. These computers run a portion of the standard event reconstruction software on the events coming out of the Level 2 trigger. The reconstruction time for a single event is about 1.5 seconds and is dominated by the reconstruction of tracks in the CTC. Cuts are then placed on quantities from this reconstruction to decide if the event passes a Level 3 trigger. If the event passes one of the Level 3 triggers, it is written out to magnetic tape and sent off for further analysis.

2.5 Luminosity Monitoring

The CDF luminosity measurement comes from looking at hits in the beam-beam counters (BBC). These are two planes of scintillators, one in the forward and one in the backward direction at $3.24 < |\eta| < 5.88$. Hits in both planes which arrive in coincidence with the proton and antiproton beams serve as both a minimum-bias trigger and as a luminosity monitor. The rate of coincidences in the BBC's divided by their effective cross section is the instantaneous luminosity. Using the direct measurements of the elastic and total cross sections made by the CDF collaboration[20], the BBC cross section is measured to be 51.2 ± 1.7 mb. The instantaneous luminosity varied during the run from about 1×10^{30} cm^{-2}s^{-1} up to a high of 25×10^{30} cm^{-2}s^{-1}.

Chapter 3

Identifying b Quarks

This analysis relies upon the ability to determine if there is at least one b quark in the event. The quarks have a color charge and are not directly observable, all observable particles have no net color charge. After the quark is created it "hadronizes" to make color singlet objects such as mesons and baryons which can be observed. This happens as the quark begins to move away from other quarks and gluons, the QCD potential energy grows until it is energetically favorable to produce a pair of light quarks from the vacuum. This process continues until all the quarks and gluon are bound up into color singlet states. The particles which are created from a quark during this hadronization process tend to move away from the interaction region along the direction of the initial quark and are reconstructed in the detector into an object called a "jet". Below is a description of how a jet is identified using the CDF calorimeters and a brief overview of how those jets originating from a b quark are identified.

3.1 Jet Clustering

The quantities of interest in looking at a jet is the 4-momentum of the initial parton from which the jet evolved. Since jets often contain electrically neutral particles such as γ, π^0, and η's, the tracking chamber cannot be used for either measurement and the information must be based on the calorimeter measurements. To identify jets CDF uses

a clustering algorithm with a fixed cone of $\Delta R = 0.4$ in $\eta - \phi$ space to define the towers included in a jet. This cone size was chosen because the jets in $t\bar{t}$ events tend to be high E_T and very collimated when m_{top} is large.

The jet identification algorithm (see Reference [21] for more details) starts by building a list of towers above an E_T threshold of 1.0 GeV. (In the plug and forward regions the towers are taken in groups of three in ϕ to provide the same 15° width as in the central region.) From this list contiguous towers are grouped to form a seed cluster. All towers within a cone of radius 0.4 around the seed cluster which have a measured E_T greater than 0.1 GeV are added to the jet and an E_T weighted centroid is calculated for the cluster. If the list of towers included in the new cluster is not the same as for the seed cluster, a new cone is drawn around the cluster centroid. All towers inside this new cone with E_T greater than 0.1 GeV are added to the new cluster and a new centroid is computed. It is possible for towers to be added or removed from the cluster at this point depending upon how much the centroid has moved. This process is repeated until the list of towers in the jet is unchanged.

While the jet clustering algorithm described above explains how the energy of a jet is reconstructed, the quantity of real interest is the energy and momentum of the parton which hadronized to make the jet. To reconstruct the parton energy from the measured jet energy, corrections must be applied to take into account mismeasurement of the jet E_T. These mismeasurement possibilities include calorimeter non-linearities, low p_T charged tracks curving in the magnetic field and not being inside the jet reconstruction cone, poor response in the calorimeters at the boundaries between modules and regions, contributions from the underlying event, out-of-cone losses due to radiation, and undetected energy of muons and neutrinos. These corrections are intended to reproduce the average jet E_T correctly, they cannot reduce the fluctuations about this mean.

The corrections for detector effects are derived by looking at the behavior of dijet events where one jet is required to be in the central calorimeter and the other is allowed to be anywhere in the detector. By requiring the transverse momentum of the two jets to balance, a correction for the E_T of second jet can be derived. Looking at many events

with the second jet in all regions of the detector gives the corrections as a function of η and ϕ. Additionally, the corrections have been checked by looking at a sample of direct photon events which contain only one other jet. These events are a subset of the two jet sample where most of the energy of one jet is due to a single γ, π^0, or η whose shower is fully contained in the CEM where the energy scale is well understood. A comparison between the true parton and the measured jet was done using Monte Carlo events. This makes it possible to get the average parton p_T which corresponds to the measured jet E_T. These corrections are typically about 30% and result in a total uncertainty of about 3-10% depending on the jet E_T.

3.2 Identifying b Jets

In CDF two basic methods are used to identify jets from the hadronization and decay of b quarks. The first method, SECVTX, looks for secondary vertices coming from the decay of a long lived b hadron. The second method, SLT, looks for low p_T electrons or muons coming from the semi-leptonic decays of the b quark. Both of these algorithms which are described more fully in Appendix D are briefly summarized here.

The SECVTX algorithm takes advantage of the lifetime of hadrons containing a b quark by looking for a reconstructed decay vertex which is displaced from the primary event vertex. This is possible because of the high resolution on the impact parameter, transverse to the beamline, of a track which is provided by the SVX. If a secondary vertex is found which has a significant displacement transverse to the beamline from the primary vertex, the jets is declared to be "tagged" as coming from a b quark.

The SLT algorithm takes advantage of the fact that roughly 35% of the time b hadrons will decay semileptonically either directly or through $b \rightarrow cX \rightarrow \ell\nu Y$[22]. These decays can be found by looking for leptons associated with a jet. Since the leptons coming from a b decay are often soft (low momentum), the minimum p_T required is only 2 GeV/c to keep the acceptance high for $t\bar{t}$. In addition, these soft leptons are required to be within a cone of 0.4 in $\eta - \phi$ of a jet to be called a tag.

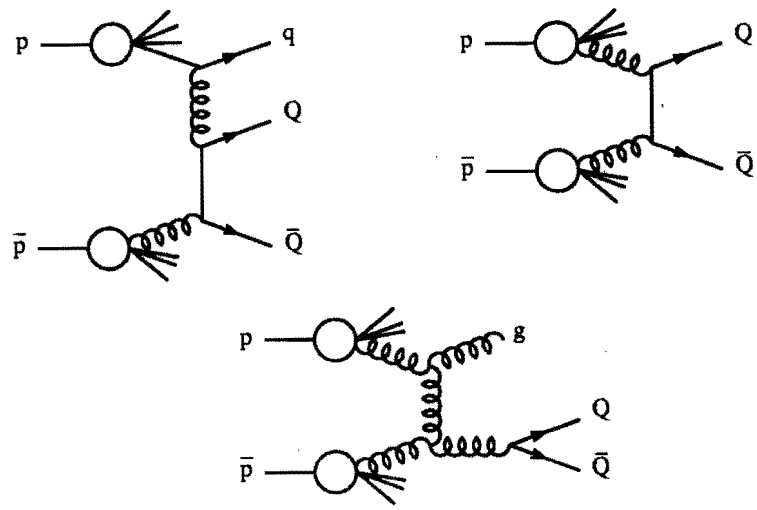

Figure 3.1: Examples of QCD production of heavy quarks Q. The upper left diagram is flavor excitation, the upper right is direct production, and the bottom diagram is gluon splitting.

The backgrounds and tagging rates for the b identification algorithms are often studied by looking at samples of inclusive jets. These jets were selected using the jet triggers as described in Section 2.4 with thresholds of 20, 50, 70, and 100 GeV for the total jet energy. These jet samples contain a mixture of gluon, light quark, and heavy quark jets where the heavy quarks come from processes like direct production (*e.g.* $gg \rightarrow b\bar{b}$), flavor excitation (an initial state gluon branching to a heavy flavor pair), and final state gluon splitting (see Figure 3.1). These processes are the source of the b jets in the background samples.

Chapter 4

Signal Sample

The data samples used in this analysis were derived from two periods of data taking with the CDF detector. The first period, Run 1A, resulted in an integrated luminosity of 19.3 ± 0.68 pb^{-1} during 1992-1993. After a brief shutdown, CDF resumed taking data for Run 1B which lasted from 1994 through 1995 and resulted in an additional 90.1 ± 7.21 pb^{-1} of data. These samples have been combined to give a total integrated luminosity of 109.4 ± 7.2 pb^{-1}.

In the analysis presented here, events were required to have two leptons (e, μ) (see Section 4.1) and to be consistent with the kinematics of a $t\bar{t}$ pair decaying into dileptons (see Section 4.2). Lastly one of the identified jets was required to be consistent with having come from a b quark by requiring a b-tag in a jet. The number of events remaining after each of the cuts which are described in detail below, is shown in Table 7.1.

4.1 Lepton Identification

This search begins by looking for events which contain two lepton candidates (ee, $e\mu$, $\mu\mu$). One lepton is required to pass a tight set of cuts while the second lepton is required to pass a looser set of cuts. After identifying these two lepton candidates in an event, the lepton passing the tight cuts is required to be isolated and the two leptons are required

to have opposite charge. The selection criteria used to select dilepton events is described below.

4.1.1 Electron Identification

The algorithm used to identify high E_T electrons begins by looking for clusters of energy in the CEM with an $E_T > 20$ GeV and comparatively little energy in the hadronic portion of the calorimeter, CHA or WHA, (E_{HAD}/E_{EM}). The cluster is required to be in a fiducial region of the CEM, which means that it is far enough from the edges of a tower that the shower is well contained in the detector. A reconstructed track pointing to the cluster is required to have a measured $p_T > 10$ GeV/c and be consistent with the CEM energy measurement (E/P). (E/P should be approximately equal to 1 for electrons of these energies but may differ due to bremsstrahlung and other effects.) The track is also required to come within 5 cm in z of a primary vertex which is required to be centrally located. The data from the CES chambers is used to make cuts on how well the extrapolated track matches to the location of the cluster in both the r-ϕ and r-z planes (Δx and Δz cuts). A cut is made on the comparison between the observed lateral shower profile and parameterized lateral shower profiles derived from electron test beam data,

$$L_{SHR} = 0.14 \cdot \sum_i \frac{E_i^{meas} - E_i^{pred}}{\sqrt{0.14E^2 + \sigma^2(E_i^{pred})}} \tag{4.1}$$

where the sum is over the towers adjacent to the seed tower for the electron, and E_i^{pred} is the prediction based on test beam electrons. A cut is made on the χ^2 of a fit of the CES strip data to electron test beam data, χ^2_{strip}. In addition cuts are made to remove electrons which appear to have come from a photon conversion. These cuts are summarized in Table 4.1 for both tight and loose CEM electrons and for the Level 3 trigger. Distributions of several of these variables are shown in Figure 4.1 for the second electron in $Z \rightarrow e^+e^-$ events. The events were chosen to have one CEM electron passing the tight cuts, the second was required to have the opposite charge and the pair had to have an invariant mass in the range 75 GeV/c^2 $< M_{\ell\ell} <$ 105 GeV/c^2.

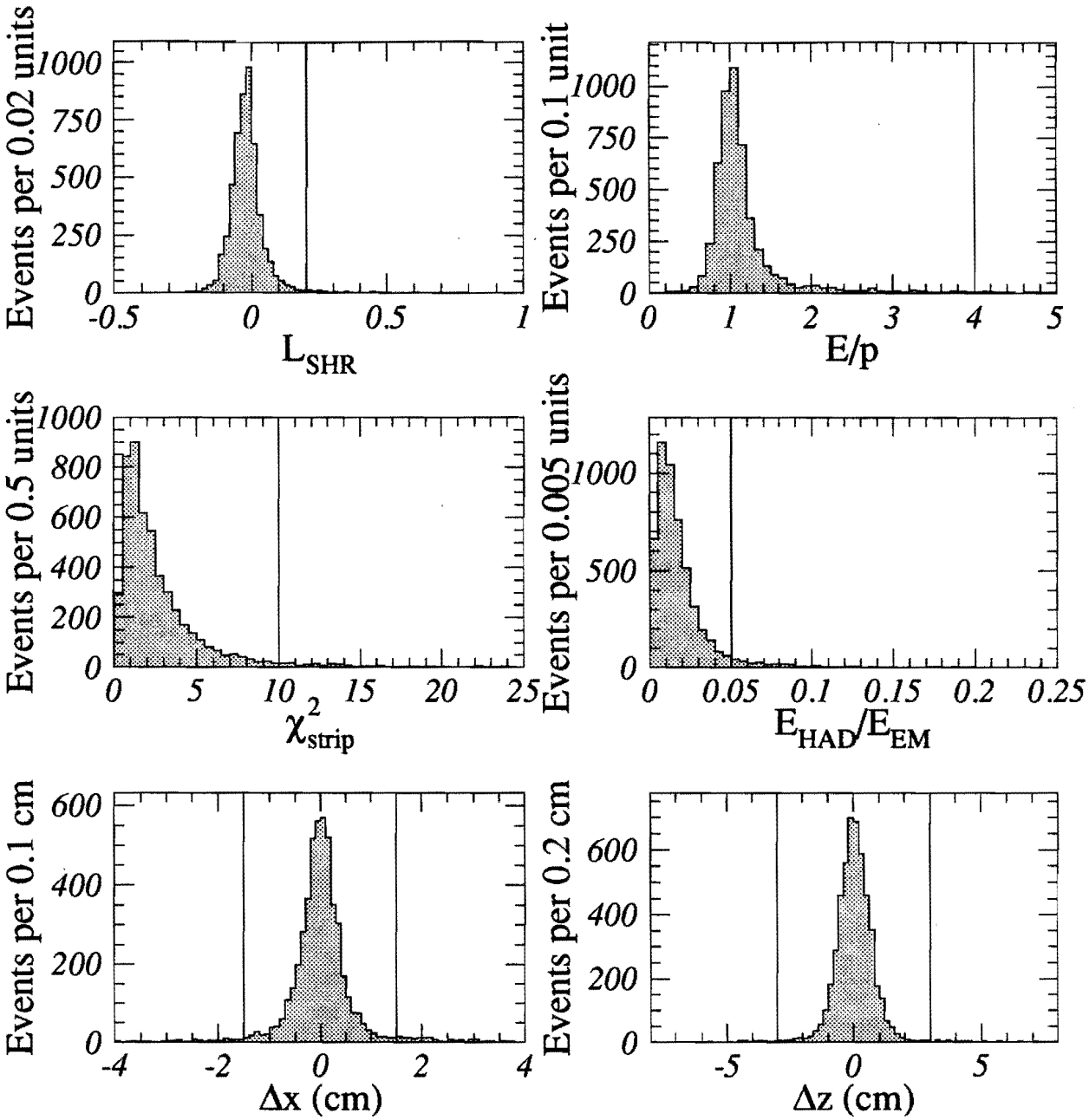

Figure 4.1: Distributions of several of the reconstructed variables used to select CEM electrons. The vertical lines are located where the cuts are made. The plots are made from the second electron in $Z \rightarrow e^+e^-$ decays.

	Tight CEM	Loose CEM	Level 3
E_T	> 20.0 GeV	> 20.0 GeV	> 18.0 GeV
p_T	> 10.0 GeV/c	> 10.0 GeV/c	> 13 GeV/c
E/P	< 1.8	< 4.0	
E_{HAD}/E_{EM}	< 0.05	< 0.055+0.045E/100	< 0.125
L_{SHR}	< 0.2	< 0.2	< 0.2
$\|\Delta x\|$	< 1.5 cm	< 1.5 cm	< 3.0 cm
$\|\Delta z\|$	< 3.0 cm	< 3.0 cm	< 5.0 cm
χ^2_{strip}	< 10.0		< 10.0
$\|z_{track} - z_{vertex}\|$	< 5.0 cm	< 5.0 cm	
$\|z_{vertex}\|$	< 60.0 cm	< 60.0 cm	

Table 4.1: Identification cuts for both tight and loose CEM electrons as well as the cuts applied in the Level 3 trigger. Note that the χ^2_{strip} cut is not applied for loose CEM electrons and the E/P and Z_{vertex} cuts are not required in Level 3.

4.1.2 Muon Identification

The first part of the muon identification algorithm is designed to look for muons which actually hit one the muon systems, called CMUO muons. This algorithm begins by looking for a track in at least one of the muon systems with a high p_T track reconstructed in the CTC pointing at it. Cuts are made on the quality of the match (Δx) between the CTC track after it is extrapolated to the muon chambers and the position of the track reconstructed in the muon chambers. Also the energy deposited in the electromagnetic (E_{EM}) and hadronic (E_{HAD}) calorimeters is required to be consistent with what is expected for a muon based upon test beam data. The total energy measured in the calorimeter towers traversed by the muon ($E_{EM} + E_{HAD}$) is required to be non-zero to remove candidates with a poorly reconstructed track which points to the wrong calorimeter tower. To remove cosmic rays and poorly reconstructed tracks, a cut has been made on the impact parameter transverse to the beam line (d_0). Also a cut is made on the distance in z between the intercept of the track and the beamline and the nearest primary vertex in the event. The cuts used for the analysis and in the Level 3 trigger are summarized in Table 4.2. Distributions of several of these variables are shown in Fig-

	CMUO Muon Cuts	Level 3 Cuts
p_T	> 20 GeV/c	> 18 GeV/c
E_{EM}	< 2.0 GeV	
E_{HAD}	< 6.0 GeV	< 6.0 GeV
$E_{EM} + E_{HAD}$	> 0.1 GeV	
Impact parameter	< 3.0 mm	
$\lvert z_{track} - z_{vertex} \rvert$	< 5.0 cm	
$\lvert z_{vertex} \rvert$	< 60.0 cm	
$\lvert \Delta x \rvert_{CMU}$	< 2.0 cm	< 5.0 cm
$\lvert \Delta x \rvert_{CMP}$	< 5.0 cm	< 10.0 cm
$\lvert \Delta x \rvert_{CMX}$	< 5.0 cm	< 10.0 cm

Table 4.2: CMUO muon identification cuts applied in the analysis and in the Level 3 trigger. Note that the Δx cuts are only applied if that particular detector has hits used. These cuts are used for both the loose and tight CMUO muon categories.

ure 4.2 for the second CMUO muon in $Z \rightarrow \mu^+\mu^-$ events. The Z events were required to have one tight CMUO muon, opposite charge, and the invariant mass in the range 75 GeV/$c^2 < M_{\ell\ell} <$ 105 GeV/c^2.

Loose muon candidates are identified in regions where the muon chambers do not exist, called CMIO muons, by looking for a high p_T track pointing to an isolated energy cluster consistent with being from a muon. The cuts for a CMIO muon (see Table 4.3) are almost the same as for a CMUO muon, except the Δx cuts are removed and additional track quality and fiducial cuts are applied. The fiducial cuts remove tracks which point to cracks in the calorimeter which would result in a mismeasurement of the particle's energy deposition. Distributions of several of these variables are shown in Figure 4.3 for the second lepton in $Z \rightarrow \mu^+\mu^-$ events. The first muon is required to be a tight CMUO muon, the two muons have opposite charge and an invariant mass in the range 75 GeV/$c^2 < M_{\ell\ell} <$ 105 GeV/c^2.

In the $\mu\mu$ channel there remains a background due to cosmic rays with a small impact parameter transverse to the beam line overlapping a real $\overline{p}p$ collision. A cosmic ray event will have two back-to-back tracks in the tracking chamber which can be removed by the following procedure. If a $\mu\mu$ event has two CMUO's or a CMUO-CMIO pair which

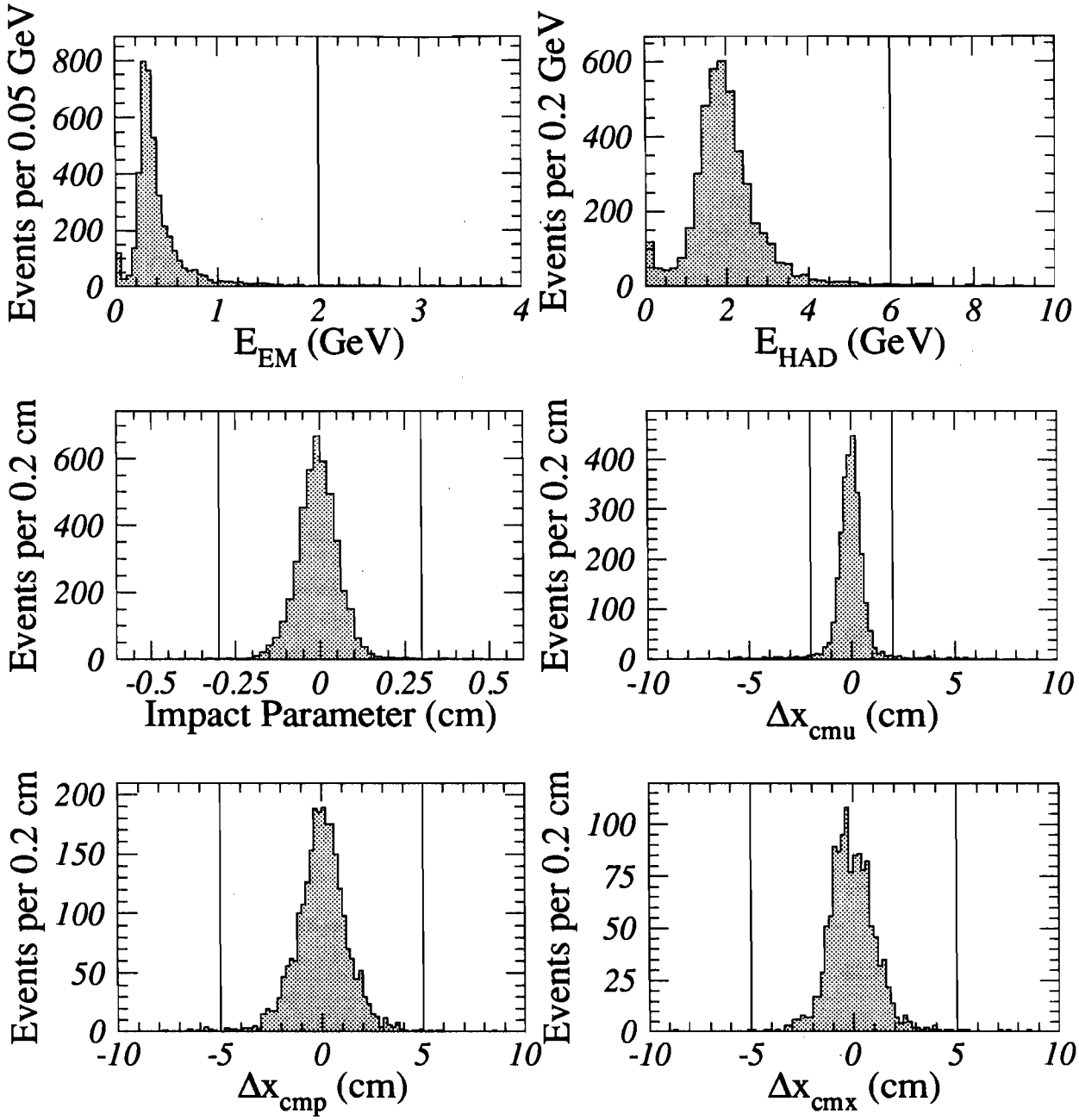

Figure 4.2: Distributions for some CMUO muon identification variables made from the second CMUO in $Z \rightarrow \mu^+\mu^-$ events. The vertical lines are located where the cuts are made.

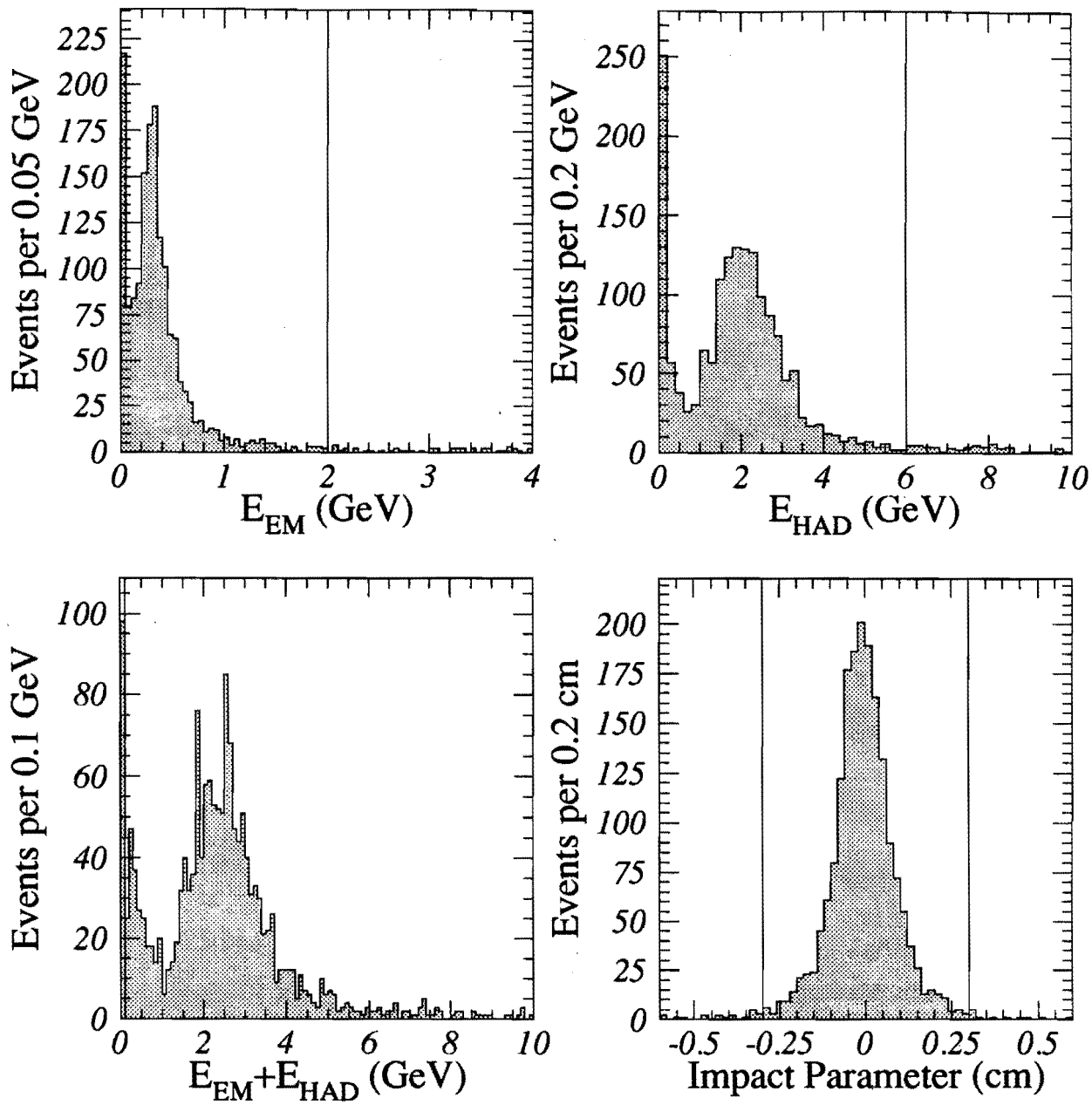

Figure 4.3: Distributions for CMIO muon identification variables taken from the second muon in $Z \rightarrow \mu^+\mu^-$ events with one tight CMUO type muon and one CMIO type muon. The vertical lines are located where the cuts are made.

p_T	> 20 GeV/c
E_{EM}	< 2.0 GeV
E_{HAD}	< 6.0 GeV
$E_{EM} + E_{HAD}$	> 0.1 GeV
Impact parameter	< 3.0 mm
$\lvert z_{track} - z_{vertex} \rvert$	< 5.0 cm
$\lvert z_{vertex} \rvert$	< 60.0 cm
number of axial CTC SL hit	≥ 3
number of stereo CTC SL hit	≥ 2
total number of CTC SL hit	≥ 6

Table 4.3: CMIO muon identification cuts applied in the analysis. Note that there is no trigger for a stiff track in Level 3.

are have $|\Delta\phi|$ >178.5° and $|\eta_1 + \eta_2|$ <0.1, the event is removed. This is not completely efficient because sometimes one of the reconstructed tracks will have its z intercept pulled toward the primary vertex by the tracking algorithm thus changing the true η of the track and causing the cosmic ray to pass the η cut. To help eliminate these events a looser cut is made if the muons have $|\Delta\phi|$ >177.0° and $|\eta_1 + \eta_2|$ <0.25. In this case, the data from the TDC's for the CHA or WHA must be consistent with a beam interaction and the difference between the times must be less than 14 ns (cosmic rays generally have a 28 ns time difference).

4.1.3 Isolation

In general, leptons coming from the decay of a W boson can be expected to be isolated from other particles in the event because there are only two decay products, the lepton and the neutrino and the latter does not deposit energy in the calorimeter or leave hits in the tracking chambers. Thus there should be relatively little energy in the towers surrounding those the lepton passed through and few tracks around the lepton's track from the W decay products. The only remaining sources of extra energy or tracks are jets, either from the b quarks or from initial or final state radiation, and the underlying event from the breakup of the proton and antiproton.

Two types of isolation are used, one based upon calorimeter information, I_{cal}, and the other, I_{trk}, based on tracking. I_{cal} is defined as the energy in a cone of radius 0.4 in $\eta - \phi$ space centered on the lepton direction, minus the lepton energy divided by the $E_T(p_T)$ of the electron (muon).

$$I_{cal} = \frac{E(0.4) - E(lepton)}{E_T, p_T} \tag{4.2}$$

This "normalization" of the energy around the lepton takes into account that a higher E_T particle can radiate more energy, some of which may be deposited in towers surrounding the lepton. I_{trk} is defined as the scaler sum of the momenta of all tracks within a cone of 0.4 in $\eta - \phi$ space around the lepton track, excluding the lepton track, divided by the $E_T(p_T)$ of the electron (muon).

In this analysis, both I_{cal} and I_{trk} are required to be less than 0.1 for the tight lepton. In addition, any CMIO muons are required to pass both isolation cuts even though they are not considered tight leptons. This extra requirement on CMIO muons is used to reduce the backgrounds such as mis-identifying a hadron as a lepton.

Cutting on the isolation of a lepton in both the calorimeter and the tracking chamber is very effective at removing the $b\bar{b}$ background. In direct $b\bar{b}$ production, an event could potentially get into the signal sample if both b's decayed semi-leptonically. However, this would produce leptons which are close in $\eta - \phi$ to a jet making it difficult for a $b\bar{b}$ event to pass both calorimeter and tracking isolation cuts. Figure 4.4 shows I_{trk} vs. I_{cal} for leptons in both ISAJET[23] $b\bar{b}$ and $t\bar{t}$ events with $m_{top} = 160$ GeV/c^2.

4.1.4 Charge

The events of interest in this analysis have two leptons which come from the decay of the W bosons. This implies that the two leptons should have opposite charges. By requiring the two leptons to be oppositely charged, the backgrounds can be reduced (for example the fake lepton background is reduced by a factor of two by this cut alone, see Section 5.2.2). However the signal only takes a small hit because after the lepton

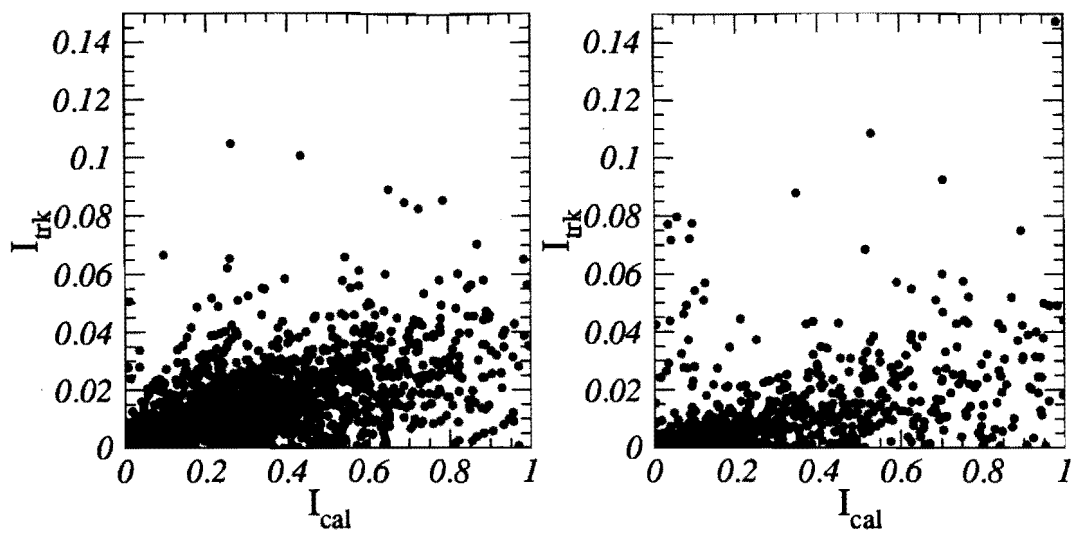

Figure 4.4: I_{trk} vs. I_{cal} for leptons from $b\bar{b}$ production (left) and from $t\bar{t}$ production (right). The $b\bar{b}$ sample was generated using a modified version of ISAJET which increases the efficiency for producing semi-leptonic decays of the b quark. The $t\bar{t}$ sample was generated using standard ISAJET with $m_{top} = 160$ GeV/c^2.

identification and isolation cuts, only about 3% of the events are removed by this cut. The events removed have one lepton from the semileptonic decay of a b quark.

4.2 Optimization of Kinematic Cuts

After selecting events with two leptons as described in the previous section, there is still a significant amount of background remaining in the sample. The background due to mis-identifying a light quark or gluon jet as a b-jet is proportional to the number of jets in the sample before tagging. If there are too many events prior to applying the b-tag requirement, the signal over background will be less than one. To improve the quality of the signal, cuts can be made on the kinematics of the events to select events with a topology more like that of $t\bar{t}$ decays than the background processes. However, because the cross section for $t\bar{t}$ production is small, the cuts must not be so hard that the acceptance for $t\bar{t}$ is reduced too much. Since the presence of a b-tag is an *a priori* requirement for this analysis, the cuts on the kinematics of the $t\bar{t}$ system are optimized

with that in mind. The procedure used to find the optimum point to cut at for several kinematic variables is described below.

4.2.1 Monte Carlo Programs Used

To calculate many of the numbers needed for this analysis, various Monte Carlo programs are used. There are three different Monte Carlo generators used to compute the hard scattering process, ISAJET, HERWIG[24], and PYTHIA[25]. ISAJET is a parton shower Monte Carlo based on the leading-order QCD matrix elements for the hard-scattering sub-process which includes incoherent gluon emission and independent fragmentation of the outgoing partons. The HERWIG Monte Carlo program is based on the leading-order QCD matrix elements for the hard-scattering sub-process with coherent parton shower evolution, cluster hadronization, and an underlying event model based upon data. PYTHIA is a Monte Carlo program which includes matrix elements for many hard-scattering sub-processes, structure functions, and initial and final state parton showers based on the string fragmentation model. Additionally, the CLEO[26] Monte Carlo is used to decay the b hadrons in the event. This is done because the CLEO Monte Carlo has been fine tuned to reproduce the properties of b hadrons seen by the CLEO collaboration. The decays of τ's are handled by TAUOLA[27].

For all Monte Carlo samples the response of the CDF detector to the final state particles in the event is simulated using a parametric detector simulation. The CDF reconstruction algorithms are used to reconstruct the leptons and jets in the event just as for the data. This makes it possible to use the same programs to analyze the data from the detector and from the Monte Carlos.

4.2.2 Variables to Optimize

Most of the variables chosen to look at were aimed at removing specific background processes. Using only the lepton identification cuts to select events will leave a very large contribution from Drell-Yan and Z production in both the ee and $\mu\mu$ channels as well as

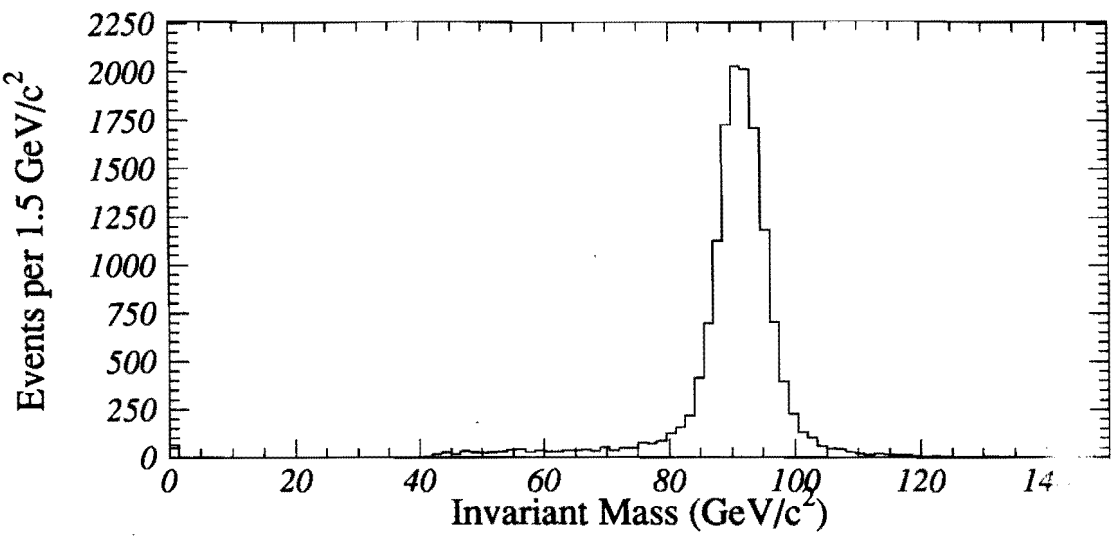

Figure 4.5: Invariant mass spectrum from a Drell-Yan and Z boson Monte Carlo sample generated using PYTHIA after only the lepton identification cuts.

W^+W^- and $Z \rightarrow \tau^+\tau^-$ in all channels. Many of the Z events can be removed by cutting on the invariant mass of the two charged leptons for ee and $\mu\mu$ events. The Drell-Yan events contain no neutrinos and thus cuts implying the presence of a neutrino signature were explored. Also, requiring the presence of two jets in the event can help reduce the background expectation from processes such as W^+W^- and Drell-Yan production.

Invariant Mass While in general the Drell-Yan mass spectrum is steeply falling, there is a large resonance in the dilepton invariant mass spectrum at the mass of the Z boson (Figure 4.5). Also, the cross section times branching ratio for $Z \rightarrow \ell^+\ell^-$ (ℓ is e or μ) at the Tevatron has been measured to be $.214 \pm .023$ nb[28], which is about 50 times the theoretical $t\bar{t}$ production cross section. This results in a potentially enormous background in the ee and $\mu\mu$ channels due to Drell-Yan and Z production. While the falling portion of the Drell-Yan spectrum can not be removed by a cut on the invariant mass without removing the entire signal, the Z peak can be easily removed. A cut around the nominal mass of the Z, $m_Z = 90$ GeV/c^2, will remove most Z events, however it will also result in a loss of acceptance for $t\bar{t}$ (see Figure 4.6). The width of the cut around m_Z on the

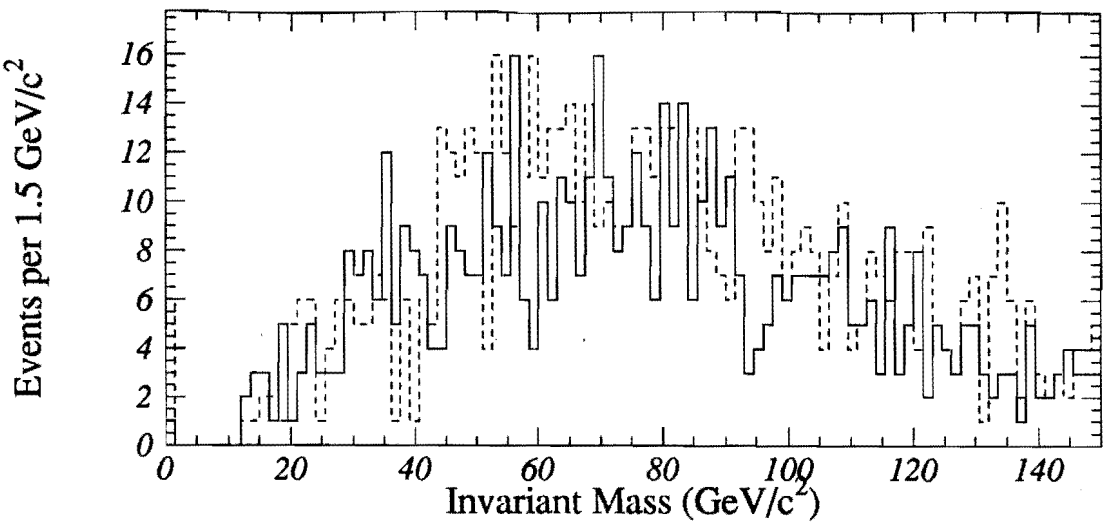

Figure 4.6: Invariant mass spectrum from a $t\bar{t}$ Monte Carlo sample generated using ISAJET after only the lepton identification cuts. The solid histogram was generated $m_{top} = 160$ GeV/c^2 while the dashed histogram used $m_{top} = 180$ GeV/c^2. This plot only shows events categorized as either ee or $\mu\mu$.

invariant mass of the dilepton pair is a variable to optimize.

Neutrino Momentum For real $t\bar{t}$ events decaying in the dilepton channel, there are two neutrinos in the final state. In general the momentum of the two neutrinos does not add to zero and large momentum imbalance, $\not{E}_T$ (see Appendix C for a definition), is expected in the event (Figure 4.7). While some background processes such as W^+W^- production also have large $\not{E}_T$, others such as Drell-Yan and Z boson production do not (Figure 4.8). The backgrounds with large $\not{E}_T$ can be reduced by cutting on the $\not{E}_T$ observed in the event.

Neutrino Direction Some of the backgrounds have a favored direction for the $\vec{\not{E}}_T$ in relation to other physics objects in the event. One way this can happen is if the energy of a jet is badly measured. In this case, the $\vec{\not{E}}_T$ will tend to point along the jet whose energy is under-measured. Another case occurs in $Z \rightarrow \tau^+\tau^-$ events where both τ's decay

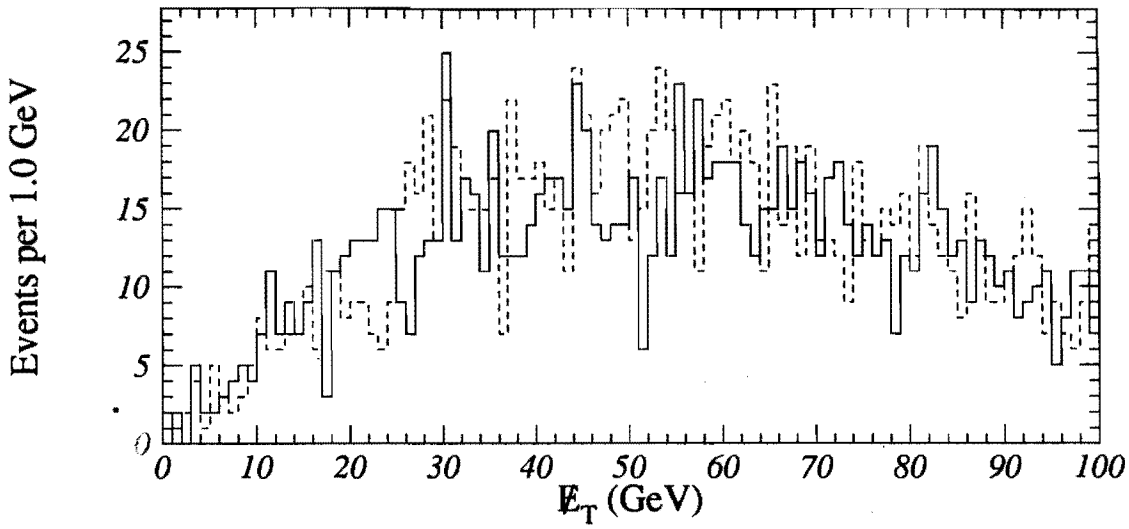

Figure 4.7: $\not{E}_T$ spectrum from ISAJET $t\bar{t}$ dilepton events with $m_{top} =$ 160 GeV/c^2 (solid) and $m_{top} = 180$ GeV/c^2 (dashed).

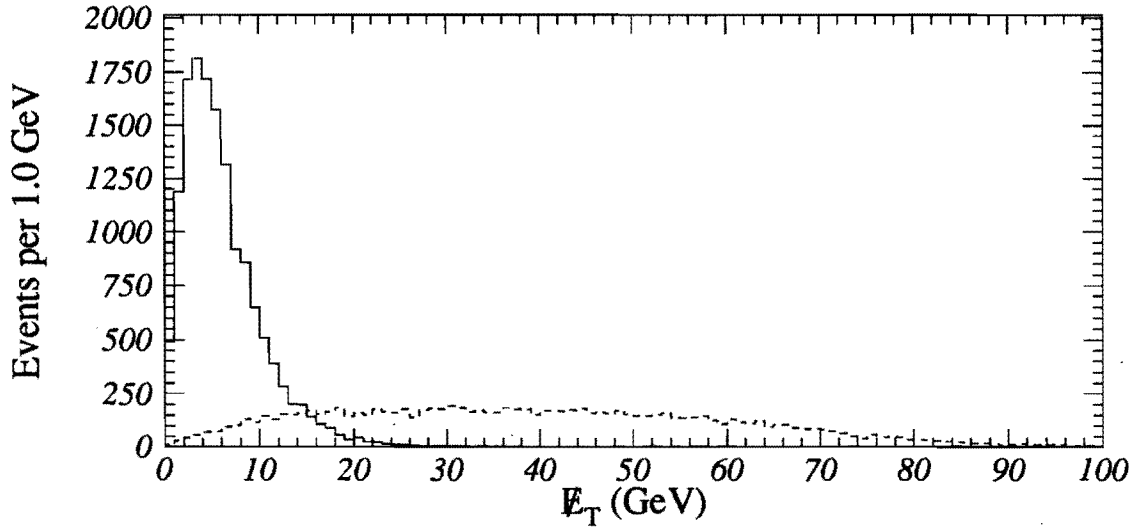

Figure 4.8: $\not{E}_T$ spectrum from W^+W^- production (dashed) and Drell-Yan plus Z production (solid), both generated using PYTHIA. The relative normalization of the plots is arbitrary, it is the shape which matters. These plots are after the lepton identification cuts have been made.

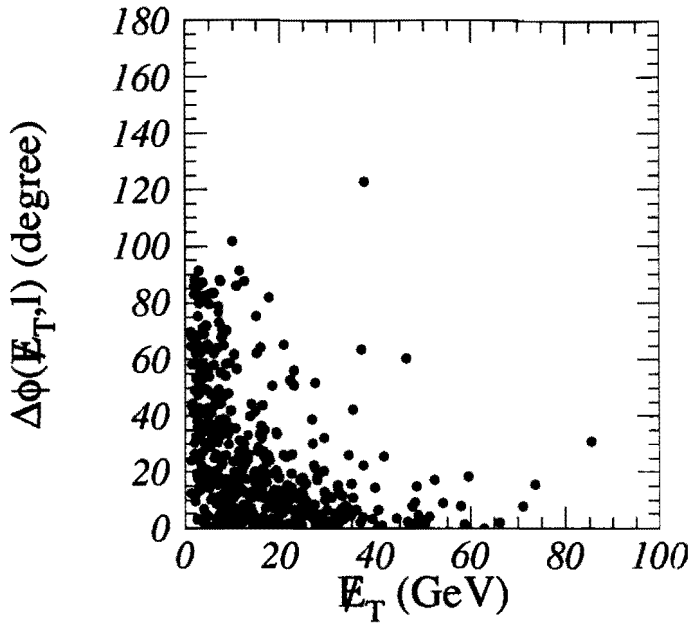

Figure 4.9: $\not{E}_T$ vs. $\Delta\phi(\not{E}_T,\ell)$ for $Z \to \tau^+\tau^-$ events.

leptonically, the $\vec{\not{E}}_T$ will often point along one of the charged leptons in ϕ because of the 2 neutrinos accompanying the lepton (see Figure 4.9). To examine the removal of these two backgrounds, a cut on the $\Delta\phi$ between the $\vec{\not{E}}_T$ and the ϕ of the nearest lepton or jet if the $\not{E}_T$< 50 GeV was explored. The distribution for $t\bar{t}$ shown in Figure 4.10, is more spread out than the $Z \to \tau^+\tau^-$.

Jet Energy Since there are two high p_T b quarks in the decay of a $t\bar{t}$ pair with a large m_{top}, there should be two high E_T jets in these events (see Figure 4.11). Only jets with an $|\eta| < 2.0$ are included in the counting because jets in this region tend to be better measured and a large value of m_{top} tends to make the partons in a $t\bar{t}$ event more central. As a result of the cuts on the η and E_T of the jets some events will have fewer than two jets because either one jet is lost or the jets overlap. On the other hand, there are also events which have more than two reconstructed jets because of initial or final state gluon radiation. Raising the E_T cut on the jets will reduce the backgrounds due to such processes as W^+W^-, Drell-Yan, and Z production whose jets are on a steeply falling

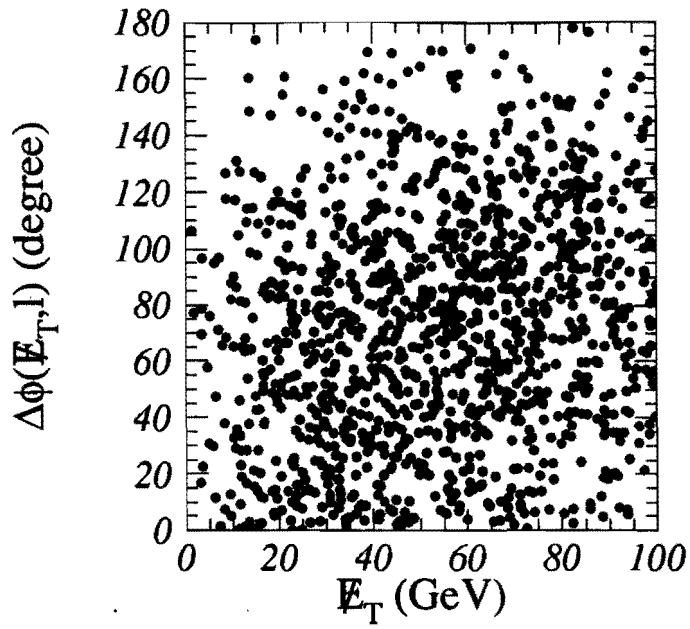

Figure 4.10: $\not{E}_T$ vs. $\Delta\phi(\not{E}_T, \ell \text{ or } jet)$ for $t\bar{t}$ events generated using ISAJET with $m_{top} = 160$ GeV/c^2.

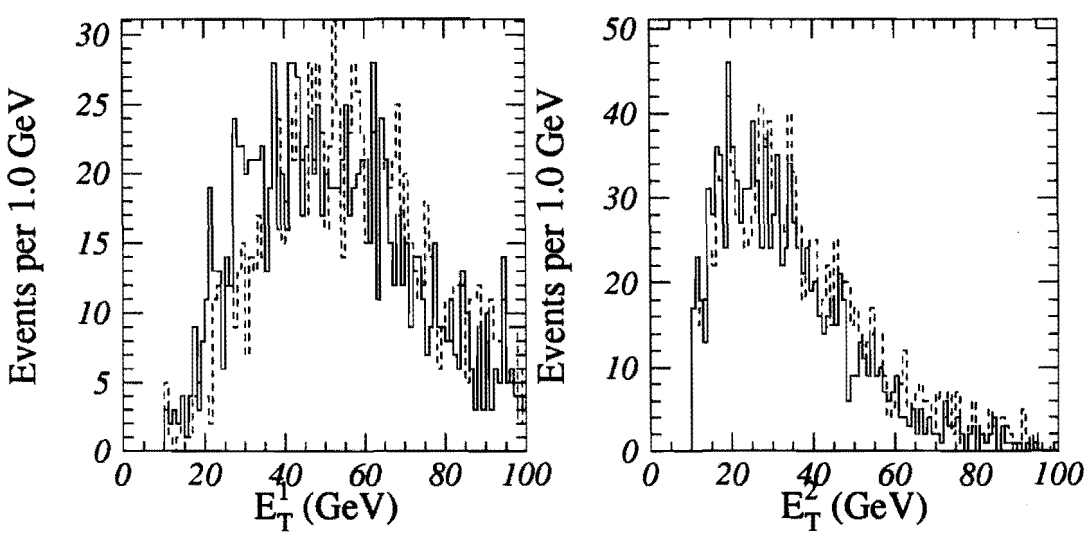

Figure 4.11: E_T spectrum of the two leading jets in $t\bar{t}$ events with $|\eta| < 2.0$. These events were generated using ISAJET with $m_{top} =$ 160 GeV/c^2 (solid) $m_{top} = 180$ GeV/c^2 (dashed).

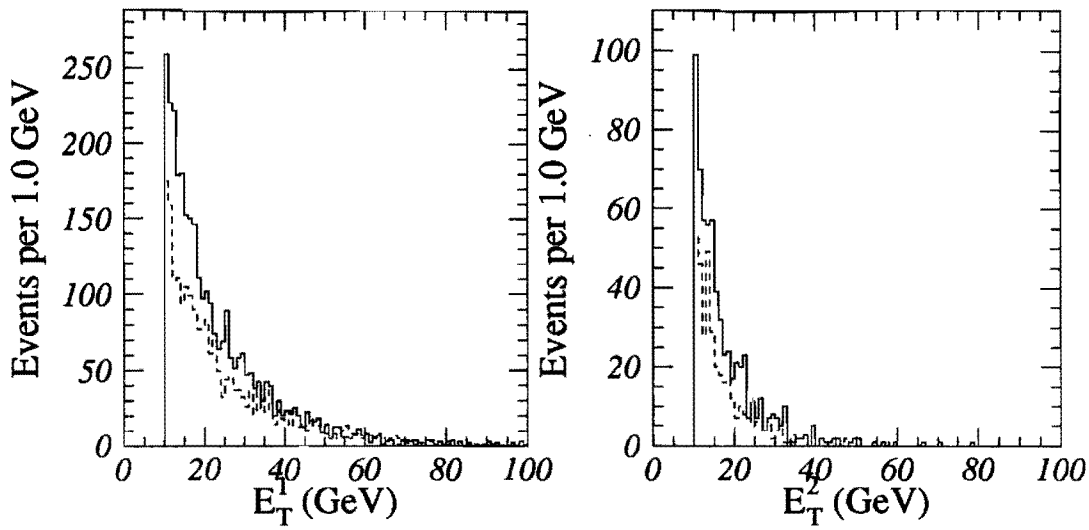

Figure 4.12: E_T spectrum of the two leading jets in Drell-Yan (solid) and W^+W^- (dashed) events generated using PYTHIA. The normalization is arbitrary, the important feature is the shape.

spectra (Figure 4.12). The E_T of the two leading jets was optimized. Their minimum value was not required to be the same.

4.2.3 Optimization Method

At the time this analysis was begun, the top quark had not yet been discovered. As a result, the ultimate goal of the analysis was to discover the top quark in a sample of approximately 100 pb^{-1}. Optimizing the kinematic cuts for discovery requires estimates of both the number of signal and the number of background events expected as a function of the cut being optimized. Knowing these, a toy Monte Carlo is used to predict the probability that the expected observed number of events could be due to a statistical fluctuation of the expected background. The optimum point to cut at is where this probability is minimized.

Expectations for $t\bar{t}$ The acceptance for a dilepton analysis which did not require a b-tag had already been calculated[1][29]. The $t\bar{t}$ acceptance was calculated from Monte Carlo samples generated using ISAJET with $m_{top} = 160$ GeV/c^2 and 180 GeV/c^2. Events were selected from these samples using the cuts from the dilepton analysis of Reference [29] resulting in 899(1064) events found in the m_{top}=160(180) GeV/c^2 samples (N_{dil}). To get the expected number of events with respect to a given cut (N_{opt}), events were selected from the same sample using the same cuts except for the cut being examined. The number of $t\bar{t}$ events expected in dilepton analysis of Reference [29] is scaled by the ratio of the number of Monte Carlo events found using the cut being optimized to the number of Monte Carlo events found using the dilepton analysis cuts from Reference [29] and the ratio of integrated luminosities ($R_{LUM} = \frac{100 \text{ pb}^{-1}}{19.3 \text{ pb}^{-1}}$):

$$N_{exp} = \epsilon_{tag} N_{exp}^{dil} \frac{N_{opt}}{N_{dil}} R_{LUM}, \tag{4.3}$$

where ϵ_{tag} is the efficiency for tagging a b (for the optimization only SECVTX was used) estimated to be about 44% and N_{exp}^{dil} is the expected number of events in 19.3 ± 0.68 pb^{-1} in the Reference [29] dilepton analysis, 1.3(0.7) for m_{top}=160(180) GeV/c^2. Normalizing in this fashion made it possible to perform the optimization without having to calculate the acceptance from the ground up at each point.

Background Estimates The major sources of backgrounds in a dilepton top analysis are W^+W^- production, $Z \rightarrow \tau^+\tau^-$, Drell-Yan, and fake leptons. To estimate the contribution from each of these processes before the tagging algorithm is applied, a procedure similar to that for the $t\bar{t}$ expectation was followed. From the same samples used to calculate the backgrounds in the analysis from Reference [29], events were selected using the cuts from that analysis (N_{dil}) and with only the cut being optimized modified (N_{opt}). These are then used to scale the background expectations from the non-tagging analysis

[1]In this dilepton analysis the kinematic cuts were chosen to be: 75 GeV/c^2 < $M_{\ell\ell}$ < 105 GeV/c^2, $\not{E}_T$> 25 GeV, $\Delta\phi(\not{E}_T, \ell \text{ or } jet) > 20°$, and $E_T^1, E_T^2 > 10$ GeV.

(N_{exp}) to get the expected background before tagging:

$$N_{bkg}^{notag} = R_{LUM} \cdot (N_{exp}^{WW} \frac{N_{opt}^{WW}}{N_{dil}^{WW}} + N_{exp}^{b\bar{b}} \frac{N_{opt}^{b\bar{b}}}{N_{dil}^{b\bar{b}}} + N_{exp}^{Z\tau\tau} \frac{N_{opt}^{Z\tau\tau}}{N_{dil}^{Z\tau\tau}} + N_{exp}^{DY} \frac{N_{opt}^{DY}}{N_{dil}^{DY}} + N_{exp}^{FL} \frac{N_{opt}^{FL}}{N_{dil}^{FL}}) \tag{4.4}$$

where the superscript defines which background is being calculated and R_{LUM} is the same as before. The expected background after b-tagging is given by:

$$N_{bkg}^{tag} = (N_{exp} + N_{bkg}^{notag}) \cdot P_{mistag} \tag{4.5}$$

where N_{exp} is given by Equation 4.3, N_{bkg}^{notag} is given by Equation 4.4, and P_{mistag} is an average probability that a jet in a given event would be tagged. Since the mistag background is proportional to the total number of events the signal must also be included in Equation 4.5 to test the hypothesis that there is no $t\bar{t}$ in the sample. P_{mistag} was estimated from the tags in a sample of Z+2 jet events using the SECVTX algorithm and was found to be 0.015 tags per event.

Toy Monte Carlo A toy Monte Carlo is used to randomly select an observed number of signal events according to a Poisson distribution with the expected number of events as the mean. The number of observed background events is selected in the same fashion using the expected number of background events as the mean. The sum of these observed numbers of events is the number of observed events for that experiment (N_{obs}). The probability that the expected number of background events could have fluctuated to the number of observed events or more is calculated for this experiment using:

$$P = \frac{1}{\Gamma(N_{obs})} \int_0^{N_{bkg}^{tag}} e^{-t} t^{N_{obs}-1} dt. \tag{4.6}$$

This single experiment is repeated 20,000 times and the fraction of the experiments consistent with a $\geq 3\sigma$ result is calculated. The 3σ point was chosen since a discovery

would be made using the significance combined with the lepton + jets analysis (recall that this analysis was started before the top quark was found) although the result is not sensitive to the $n\sigma$ chosen. The optimal point to cut is where the fraction of experiments to observe a $\geq 3\sigma$ result is maximized. However, since the expected number of observed signal events is small for all cases, the acceptance must be checked to insure that it is not reduced too much. Since only a few events are expected in the best case, it is very easy to wind up cutting so hard on a variable that the expected number of $t\bar{t}$ events is $\ll 1$, which is not useful.

4.2.4 Results of Optimization

This procedure was applied to each of the cuts listed in Section 4.2.2 individually. The result of varying the individual cuts is displayed in Table 4.4 for the ISAJET sample with $m_{top} = 160$ GeV/c^2, the sample with $m_{top} = 180$ GeV/c^2 looks similar. As a result of the optimization, ee and $\mu\mu$ events with 80 GeV/c^2 $< M_{\ell\ell} <$ 100 GeV/c^2 are removed from the signal sample. The $\not\!\!E_T$ is required to be greater than 20 GeV while the cut on the direction of the $\vec{\not\!\!E}_T$ relative to the leptons and jets in the event was removed. For the jet selection the optimum cut was to require one jet with $E_T > 20.0$ GeV and a second with $E_T > 10.0$ GeV. These cuts were then checked to see that when they were applied together, they were at the optimum point.

	Signal	Background	$F(\geq 3\sigma)$
$E_T{}^1, E_T{}^2$ (GeV)			
10,10	2.96	0.19	0.777
20,10	2.92	0.18	0.780
30,10	2.57	0.15	0.740
15,15	2.68	0.16	0.759
20,20	2.28	0.13	0.706
25,25	1.86	0.10	0.631
30,30	1.46	0.08	0.548
minimum $\not{E}_T$(GeV)			
10	3.19	0.29	0.712
15	3.16	0.24	0.753
20	3.08	0.21	0.779
25	2.96	0.19	0.774
30	2.86	0.18	0.769
$\Delta\phi$			
0°	3.33	0.25	0.766
10°	3.11	0.21	0.772
20°	2.96	0.19	0.776
Z window (GeV/c^2)			
0	3.37	0.28	0.752
10	3.10	0.20	0.785
15	2.96	0.19	0.774
20	2.81	0.18	0.754

Table 4.4: Results of optimization method using the ISAJET sample with $m_{top} = 160$ GeV/c^2. The signal is the number of events expected after tagging in 100 pb^{-1}, the background is the expected number of background events expected in 100 pb^{-1}, and $F(\geq 3\sigma)$ is the fraction of experiments with a greater than or equal to 3σ discovery.

Chapter 5

Backgrounds

The backgrounds for this analysis come from a combination of physics processes which mimic the signal and from reconstruction problems. The physics processes considered include Drell-Yan, diboson, and $b\bar{b}$ production. Contributions due to reconstruction failures resulting in a hadron being identified as a lepton are also considered. In all cases, the background primarily comes from the tagging algorithms either mistagging a gluon or light quark jet or tagging b jet which is the result of a process like gluon splitting. Two different approaches for calculating the background are described below. The first method (method 1) depends on the assumption that the jets in background events will have the same heavy flavor content as inclusive jets, and involves looking only at our data samples to estimate the background. The second method (method 2) uses Monte Carlo programs to derive an estimate of the background due to physics processes but uses the data samples to estimate backgrounds due to reconstruction errors which are not well modeled by the Monte Carlo. Since the tagging algorithms are not highly correlated on the backgrounds, the backgrounds are calculated separately for each algorithm and added.

5.1 Method 1

There are two ways for a background event to get into the sample. First the event might contain a jet from the hadronization of a b quark in which case the jet may be correctly tagged as heavy flavor by the tagging algorithms. In the second case, a light quark or gluon jet might be tagged due to track reconstruction errors. Looking at a sample of inclusive jets, there will also be tags from these same two sources: real tagged b or c jets and false tags. If the fraction of jets from a b or c quark is the same in both the inclusive jets sample and in the background sources, the tagging rate per jet found in the inclusive jets sample can be applied to the jets in the pre-tag sample to predict the background.

The b jets in the inclusive jets samples come mainly from three processes: direct production ($gg \rightarrow b\bar{b}$), a final state gluon splitting into a $b\bar{b}$ pair, and flavor excitation (initial state gluon splitting). However, in the physics background events some of these b quark production mechanisms, such as direct production, are absent resulting in a smaller fraction of b jets in the background processes than what is found in the inclusive jets sample. As a result method 1 is expected to be at least a slight overestimate of the background.

5.1.1 SECVTX

For the SECVTX tagging algorithm, the tagging rate is defined as the number of tagged jets found divided by the number of "taggable" jets[1]. Looking in the inclusive jets sample, the tagging rate is found to be a function of both the E_T of the jet and the number of SVX tracks associated with the jet (N_{SVX}). To take these variations into account, the tagging rate is parameterized as a function of these two variables for both the $+L_{xy}$ and $-L_{xy}$ cases separately. The $-L_{xy}$ tags are mainly due to mistags where the impact parameter of the tracks is poorly reconstructed although about 30% come from b or c jets. Since the track reconstruction makes mistakes equally in the direction it screws

[1]A taggable jet is defined as a jet with E_T>10 GeV, and two SVX tracks within a cone 0.4 of the jet axis passing all the track selection cuts except the d/σ_d cut (see Section D.1).

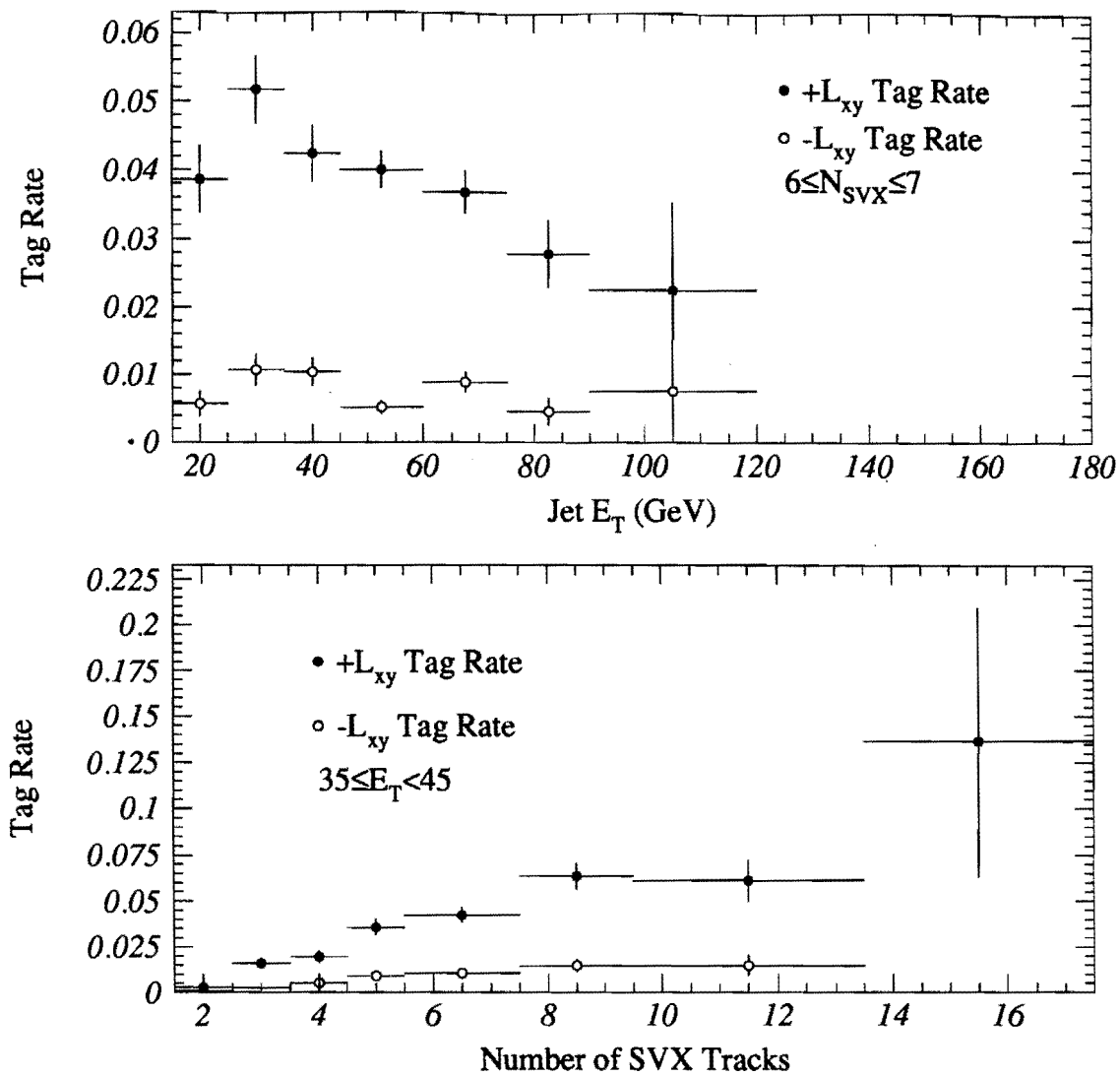

Figure 5.1: SECVTX tag rates in inclusive jets as a function of jet E_T and N_{SVX} using the parameterization. The particular E_T and N_{SVX} bins shown were arbitrarily chosen.

up an impact parameter, the L_{xy} distribution is expected to be symmetric about 0 and there will be roughly equal numbers of mistags with $-L_{xy}$ and $+L_{xy}$. The difference is that $+L_{xy}$ tags will also contain most of the real tags of b jets, less than about 10% come from mistags. The parameterizations from a sample of inclusive jets which triggered the Level 2 jet triggers are shown in Figure 5.1.

		SECVTX		SLT	
N_{jets}	$N_{taggable}$	N_{tag}	$N_{predict}$	N_{tag}	$N_{predict}$
1	956	4	10.3 ± 1.6	14	14.6 ± 1.3
2	350	4	$3.7 \pm .6$	2	$5.8 \pm .5$
≥ 3	102	1	$1.6 \pm .2$	1	$1.7 \pm .2$

Table 5.1: Number of tags predicted and found in a sample of Z's as a function of the number of jets with E_T>10 GeV and $|\eta| < 2.0$. Here the errors only take into account the error on the fake tag probability. Also shown is the number of jets in each bin before tagging.

Getting a background prediction for a given sample is simply a matter of applying the parameterization to the jets in that sample. This is done for each event by looking up the tag rate for the given E_T and N_{SVX} of each jet in the event. The probability that the event will have at least one tagged jet is given by

$$B_{SVX} = 1 - \prod_{i=1}^{N_{jets}} (1 - P(E_T, N_{SVX})) \tag{5.1}$$

where $P(E_T, N_{SVX})$ is the tag rate parameterization giving the probability of tagging each jet and N_{jets} is the number of jets in the event. The total background in the sample is just the sum over all events of the background per event.

To check that these parameterizations are doing a reasonable job of predicting the number of tags expected due to b or c jets and mistags, the parameterization is applied to the jets in a sample of Z boson events. Here a Z boson is defined as having either two electrons or two muons defined as in Section 4.1 with opposite charges and an invariant mass in the range 75 GeV/c^2 < $M_{\ell\ell}$ < 105 GeV/c^2. The jet counting is done with an E_T threshold of 10 GeV and $|\eta| < 2.0$. The number of tags found as a function of the number of jets in the event as well as the prediction based upon the tag rate parameterization are listed in Table 5.1. The Z sample is an excellent control sample for this study because most of the jets are gluon jets which is also the case for the backgrounds.

The comparison between the predicted number of tags and the number actually found looks good except in the one jet bin. In the one jet bin a smaller heavy flavor content than in inclusive jets is expected because most production mechanisms for b quarks in association with a Z result in two separate jets, one for the b and another for the $\bar{b}$. For such an event to get into the one jet bin the second jet must be lost either because of the η cut, E_T cut, or in a crack between calorimeter modules. Using tags in the Z's from Drell-Yan Monte Carlo with one reconstructed jet to get a prediction using the method of Section 5.2.3 of the number of expected tags results in prediction of 6.4 ± 1.6 tags which is in much better agreement with the observed 4 tags.

In the pre-tagged signal sample there are 16 events (see Table 7.1) to which the tag rate parameterization must be applied. This results in a predicted background of $0.584 \pm .088$ events. In addition to the statistical error, a $\pm 15\%$ systematic uncertainty has been included which takes into account how well the tag rate parameterization does at predicting the number of tags seen in a sample of events with large $\sum E_T$. The $\sum E_T$ sample is made by looking for multi-jet events where the scaler sum of the E_T of the jets is greater than 300 GeV. The ratio of the number of $+L_{xy}$ SECVTX tags found divided by the number of tags predicted for that sample is shown in Figure 5.2 for several variables. The prediction is systematically low by $\sim 30\%$. Since the $\sum E_T$ in this sample is much harder than expected for $t\bar{t}$ events, one half the difference is taken as a systematic error.

5.1.2 SLT

The method 1 background calculation for the SLT tagging algorithm follows a very similar procedure as was used for SECVTX. The biggest difference is that instead of being based upon jets, the calculation centers on the tracks near the jets. For the SLT algorithm the tag rate is defined as the number of tracks corresponding to an SLT tag divided by the number of tracks which extrapolate to a fiducial region of the detector for that type of lepton. The tag rates are parameterized as a function of the p_T of the track. The electron tag rate is also parameterized as a function of the isolation of the track in the

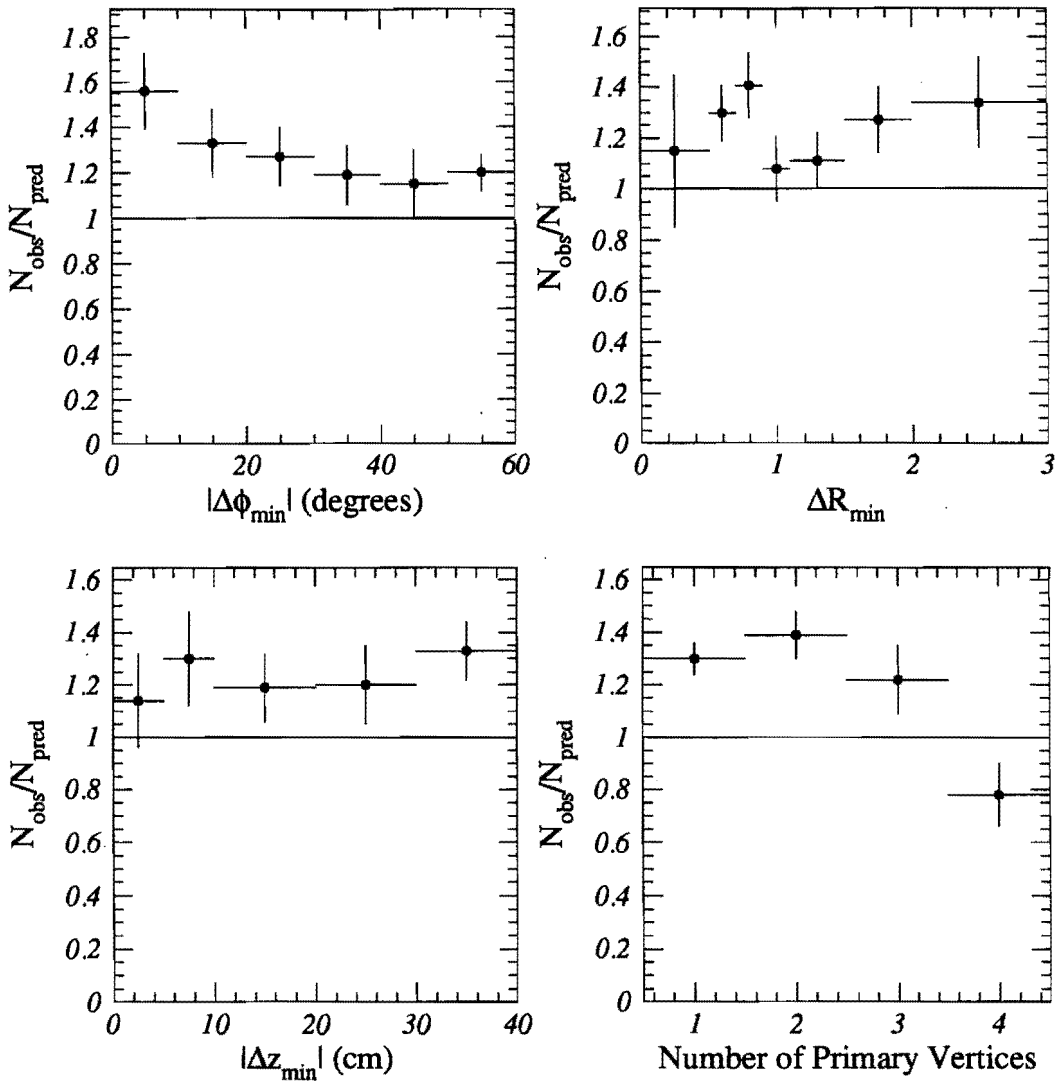

Figure 5.2: Ratio of the number of $+L_{xy}$ tags observed to the number predicted in the $\sum E_T$ sample as a function of various quantities. $|\Delta\phi_{min}|$ (ΔR_{min}) is the minimum $\Delta\phi$ (ΔR) to the closest jet, and Δz is the distance along the beamline between two good primary vertices.

tracking chamber $\sum p/p$, where $\sum p$ is the sum of the momentum of all tracks within a cone of 0.2 around the candidate track. This procedure is applied to each type of lepton separately, electron, CMUO muons in the CMX, and CMUO muons in the CMU and CMP. The tag rates are again found by looking at the tracks in a sample of inclusive jets which came in on several different jet trigger thresholds. Examples of the SLT tagging rate parameterizations are shown in Figure 5.3.

The calculation of the background for each type of lepton tag in an event is very similar to that for the SECVTX algorithm:

$$B_{SLT} = 1 - \prod_{i=1}^{N_{trk}} (1 - P(p_T, \sum p/p)) \tag{5.2}$$

where N_{trk} is the number of tracks in the event and $P(p_T, \sum p/p)$ is the tag rate parameterization. The total background per event is just the the sum over electrons and muons and the total background in a sample is the sum of the per event backgrounds.

Again the usefulness of the background prediction can be tested by looking at the same sample of Z boson events used in the previous section. The prediction as a function of the number of jets in the event compared with the actual number of tags observed is given on the right in Table 5.1. Here the agreement is better in the one jet bin than was found with the SECVTX algorithm. This is in part due to the fact that the purity of the SLT algorithm is only about 25% (see Section D.2). This means that the majority of the tags found in the Z sample are really mistags which is what the number predicted represents. Since the tags in the 1 jet bin are predominantly not due to the presence of a b or c, the agreement is expected to be better for SLT than for SECVTX.

Looking at the events selected before tagging and performing the background calculation gives an expectation of $0.772 \pm .067$ tagged events. This error includes a systematic error of $\pm 10\%$ on the muon tag rate and $\pm 15\%$ on the electron tag rates. These errors are estimated by looking at the tracks in jet events which are selected using different jet trigger thresholds.

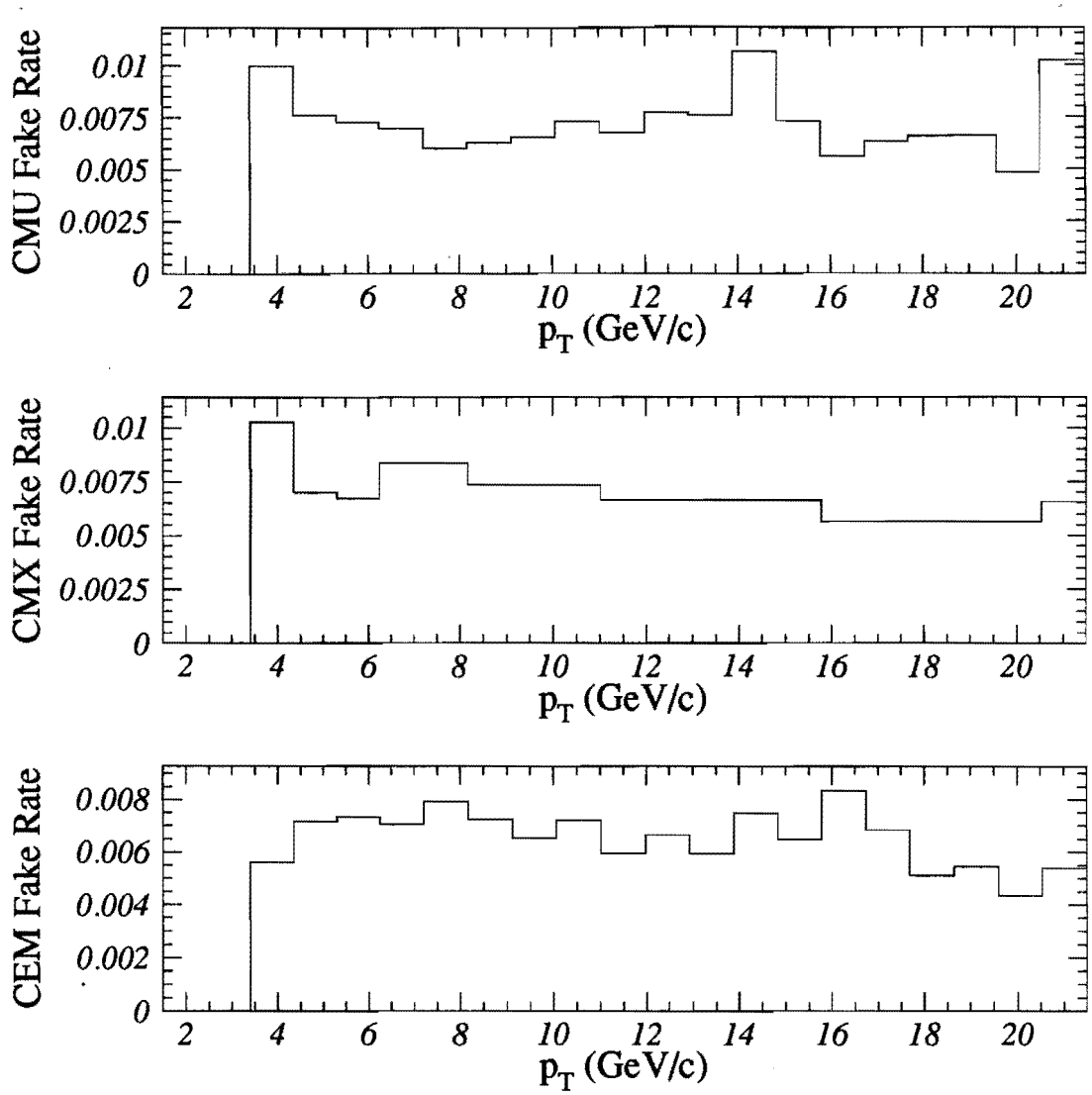

Figure 5.3: Tagging rate parameterizations for the SLT algorithm. The CMU muon, CMX muon and CEM electron rates are shown separately.

5.1.3 Combined

The two tagging algorithms are expected to have very little correlation on the backgrounds since they use very different approaches to find the b jets in an event. As a result, the background calculations for the two algorithms are not very correlated. This means that the expected background for the combination of the two algorithms is close to the sum of the separate backgrounds, $1.35 \pm .11$. Correlations between the tagging algorithms on the background would reduce the expected background further, so using the sum is conservative. This estimation method includes all of the dilepton backgrounds (Drell-Yan, $Z \rightarrow \tau^+\tau^-$, W^+W^- production, fake leptons, and $b\bar{b}$).

There are reasons to expect that the method 1 calculation of the background results in an overestimation of the background in the sample. As was mentioned above, the b content is expected to be higher in the inclusive jet samples than in the pre-tagged signal sample due to the absence of processes like direct $b\bar{b}$ production as the source of the tagged jets in the pre-tag sample. (For direct $b\bar{b}$ production, the high p_T leptons which cause the event to be selected as a dilepton + two jet event generally each come from the semi-leptonic decay of the b and $\bar{b}$ quarks. Since the tagged jets are required to be separated from the leptons by 0.4 in $\eta - \phi$ space, the jets from the b hadronization and decay are generally not eligible for tagging and the tagged jet must be a light quark or gluon jet.)

The second reason this is expected to result in an overestimation of the background is that the pre-tag sample has been assumed to be all background. In fact, following the hypothesis that there is top in the sample, then to get the proper background, a correction must be made to remove the expected mistags in signal events from the background calculation (see Equation 4.5). To show that there is an excess of tagged events, the conservative path is taken and this correction is not made. When the production cross section is calculated, this effect is taken into account.

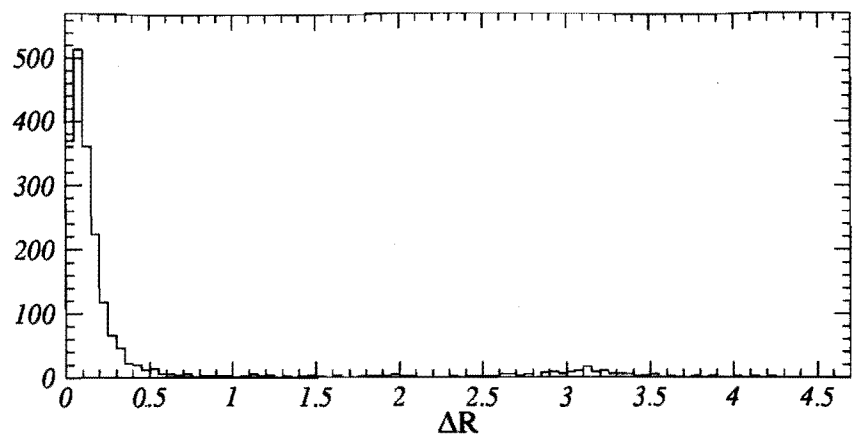

Figure 5.4: ΔR distribution between a lepton and the nearest jet in $\eta - \phi$ for a $b\bar{b}$ Monte Carlo sample generated using ISAJET. The leptons are required to have $p_T > 15$ GeV/c.

5.2 Method 2

The second technique used to estimate the background for this analysis involves a combination of estimates based upon Monte Carlo techniques and calculations very similar to those of method 1. For backgrounds such as W^+W^- and Drell-Yan production, Monte Carlo calculations are used to include the theoretical knowledge available. For the mistags and fake leptons, the data alone is used since these are not modeled well in simulations.

The background contribution due to $b\bar{b}$ production is extremely difficult to estimate accurately due to its small size. An estimate based upon the calculation for Reference [2] is made here and shown to be quite small. In that paper the total background from $b\bar{b}$ was estimated to be $0.03 \pm .02$ events in a 67 pb^{-1} sample. The tagged jet in a $b\bar{b}$ event is required to be outside a cone of radius 0.4 in $\eta - \phi$ space centered on the lepton meaning that in general it will not be from a b quark. Figure 5.4 shows the ΔR distribution between the leptons and the nearest jet in $b\bar{b}$ events for leptons with $p_T > 15$ GeV/c. The fraction of leptons which have a minimum $\Delta R < 0.4$ is about 0.9 which means the jets which are taggable are dominated by gluon jets to which the mistag probability, conservatively about 10% for the sum of the two algorithms, should be applied further

reducing the estimate to about 0.003. Taking into account the increased integrated luminosity, the estimate becomes ~ 0.005 events which is small compared to the other processes considered.

5.2.1 Mistags

For the method 2 background calculation, the background due to real b jets is calculated using various Monte Carlo programs. As a result the estimation of the mistag background should include only the true mistags without the contribution due to b jets. For the SECVTX tagging algorithm the method to estimate the mistag background is essentially the same as in method 1. The difference is the use of the $-L_{xy}$ parameterization instead of the $+L_{xy}$ parameterization and the assumption that the mistags are symmetric about $L_{xy} = 0$. Applying the $-L_{xy}$ parameterization to the jets in the events in the pre-tag signal sample results in an estimated background of $0.155 \pm .031$ events due to mistags alone. The error here includes a $\pm 20\%$ uncertainty which is taken from the number of tags found and the number predicted in the $\sum E_T$ sample. The ratio of the number of $-L_{xy}$ SECVTX tags found divided by the number of tags predicted for that sample is shown in Figure 5.5 for several variables. The prediction is systematically low by $\sim 40\%$. Since the $\sum E_T$ in this sample is much harder than expected for $t\bar{t}$ events, one half the difference is taken as the systematic error.

The SLT mistag background calculation makes use of the purity of the SLT algorithm as shown in Section D.2. It was found that the purity of the SLT tags is only $25 \pm 12\%$. This implies that only 25% of the method 1 background for the SLT algorithm ($0.772 \pm .067$) was due to the presence of real b quarks and 75% was due to mistags. Using this argument, the SLT mistag background in the pre-tag sample is estimated to be $0.57 \pm .11$ events.

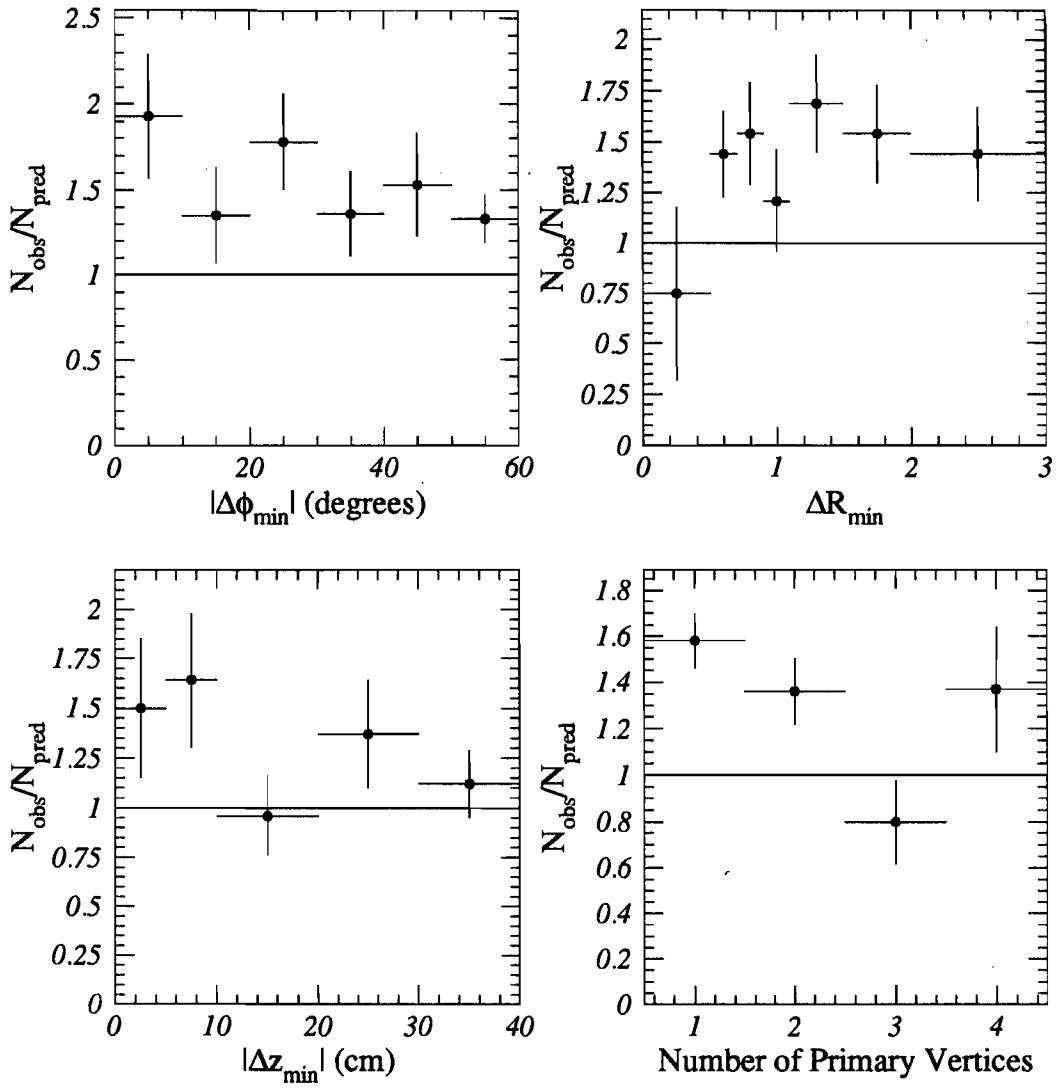

Figure 5.5: Ratio of the number of $-L_{xy}$ tags observed to the number predicted in the $\sum E_T$ sample as a function of various quantities. $|\Delta\phi_{min}|$ (ΔR_{min}) is the minimum $\Delta\phi$ (ΔR) to the closest jet, and Δz is the distance along the beamline between two good primary vertices.

5.2.2 Fake Primary Leptons

The "fake lepton" background contribution is the result of events which pass the signal selection cuts, but one of the primary leptons found is really either from either a parton which either fragmented in some unusual way, was poorly measured and ended up being reconstructed as a lepton, or leptons coming from the decay of particles such as K and π mesons or b or c hadrons. For an event of this type to get into the sample, the rest of the event must still pass all the cuts imposed. This implies that events with a fake lepton which gets into the signal sample are in fact probably W+jets events where the $\not\!\!E_T$ and real lepton come from the leptonic decay of the W, the fake lepton is either a jet or track which looks similar to an electron or muon, and the jets are mostly gluon jets which may contain a b quark from gluon splitting and can be tagged.

To calculate this background, a fake lepton rate is derived from inclusive jets and applied to events which contain a lepton, $\not\!\!E_T$, two jets, and a fake lepton candidate. The fake lepton rates are derived from an inclusive jets sample where cuts are made to suppress real electron from leptonic W and Z decays to prevent overestimating the fake lepton rates and thus the background. This technique assumes that the jets in an inclusive jet sample will contain leptons at the same rate as events which might mimic the signal with a fake lepton. The calculation is designed to minimize the overlap between the sample of events containing fake lepton candidates and the signal sample. If the signal events are allowed into the fake lepton sample, the background would be overestimated.

To calculate the fake rates, a new class of lepton, called "denominator lepton", is defined. These denominator leptons are selected using cuts which have no overlap with the lepton identification cuts defined in Section 4.1. This exclusion is done so that when the fake rates are applied to the W+jets events, the two primary leptons in the signal events will not contribute to the background estimation resulting in an overestimate of the background. A denominator central electron is defined as an energy cluster in the CEM calorimeter having:

- $E_T > 20.0$ GeV

- passing fiducial cuts
- not consistent with being a conversion
- $L_{shr} > 0.2$ or $E_{HAD}/E_{EM} > 0.055 + 0.045E_T/100$.

The L_{shr} and E_{HAD}/E_{EM} requirements here are the inverse of what is required for a loose central electron in the analysis. Denominator CMUO muons are defined by:

- $p_T > 20$ GeV/c
- $E_{HAD} + E_{EM} > 0.1$ GeV
- impact parameter and vertex cuts same as for CMUO muons
- $E_{EM} > 2.0$ GeV or $E_{HAD} > 6.0$ GeV.

Here the E_{EM} and E_{HAD} cuts are used to insure that muons passing the signal selection criteria are not included in the denominator selection. The CMIO muon denominator cuts are the same as the CMUO denominator cuts except that the CMIO fiducial cuts are also required to insure the energies are well measured and a cut $E_{HAD}/E_{EM} > 0.2$ is added to remove electrons. (Isolated electrons will pass the CMIO denominator cuts if the E_{HAD}/E_{EM} cut is not applied. Since an electron has no chance of passing the CMIO lepton identification cuts, including it in the CMIO denominator is wrong.)

The fake lepton rates are defined as the number of leptons passing the analysis cuts divided by the number of denominator leptons found in a sample of inclusive jets which passed a 20 GeV jet trigger. To remove real leptons from W decays from both the numerator and the denominator, events with $\not{E}_T >$ 15 GeV are removed from the sample used to measure the fake lepton rates. Events containing a $Z \rightarrow e^+e^-$ decay are removed by requiring that there be no other energy cluster in the electromagnetic calorimeter with $E_T > 15$ GeV in the event. To remove $Z \rightarrow \mu^+\mu^-$ decays there must be no other CMUO or CMIO muon with $p_T > 10$ GeV/c which has an invariant mass in the range 75 GeV/$c^2 < M_{\ell\ell} <$ 105 GeV/c^2 in the event. To prevent any bias from the trigger at

Category	Isolated	Non-isolated
Tight CE	$0.0153 \pm .0055$	$0.0069 \pm .0016$
Loose CE	$0.0363 \pm .0085$	$0.0165 \pm .0025$
CMU/P	$0.1053 \pm .0573$	$0.0136 \pm .0048$
CMX	$0.2857 \pm .2645$	$0.0649 \pm .0305$
CMIO	$0.0370 \pm .0222$	

Table 5.2: Fake rates for isolated and non-isolated leptons derived for this analysis. Note that the CMIO category always requires isolation.

least one jet, which is outside a cone of 0.4 in $\eta - \phi$ space around the fake lepton candidate, must pass the trigger threshold.

The resulting fake rates are summarized in Table 5.2 for both isolated and non-isolated lepton candidates. The fact that the fake rates for isolated leptons is higher than for the non-isolated leptons is a reflection of the definition used. The quantity which actually matters is the product of the fake lepton rate and the number of events with a denominator lepton. Since the number of isolated denominator leptons is much smaller than the number of non-isolated denominator leptons, the background from non-isolated fake leptons actually dominates.

These fake lepton rates are then applied to a sample of events selected with one lepton passing the analysis cuts and one "lepton" passing the denominator cuts. The dilepton invariant mass must pass the Z mass window cut, $\not\!\!E_T$>20 GeV, and there must be at least two jets with $E_T^1 > 20$ GeV and $E_T^2 > 10$ GeV. There are two different ways of applying the b-tag to calculate the final background contribution. The first is to require that one jet is tagged, the second is to apply the fake tagging parameterizations to the jets/tracks in the events. Requiring a tag in the event results in a very statistically limited sample (there is only 1 event passing these cuts which contains a tag). The fake tagging parameterization can be used to get a statistically more robust answer. The background for a given event is

$$B^{FL} = P(fake\,lepton) \times (B_{SVX} + B_{SLT}) \tag{5.3}$$

Category	SECVTX	SLT
ee	$0.0^{+.0001}_{-.000}$	$0.0^{+.0002}_{-.000}$
$\mu\mu$	$0.0008 \pm .0010$	$0.0029 \pm .0026$
$e\mu$	$0.0269 \pm .0086$	$0.039 \pm .014$
Total	$0.0267 \pm .0086$	$0.040 \pm .014$

Table 5.3: Expected background due to fake primary leptons for each b-tagging algorithm separately.

where B_{SVX} and B_{SLT} are given by Equations 5.1 and 5.2 respectively and $P(fake\ lepton)$ is the appropriate fake lepton rate from Table 5.2. In calculating the fake tagging probabilities, the jets used must be outside a cone of radius 0.4 around the fake lepton candidate.

The background broken down by dilepton categories is shown in Table 5.3 with the total background being $.067 \pm .017$. Obviously the background calculation is still limited by the number of events with a fake lepton candidate, especially in the e^+e^- case. The background is entirely dominated by the $e\mu$ category.

5.2.3 Drell-Yan Background

Drell-Yan events[2] which have two or more jets can mimic the $t\bar{t}$ signal if there is some $\not{E}_T$ reconstructed in the event. If one of the jets is mismeasured the $\not{E}_T$ reconstruction algorithm, which makes use of momentum balance transverse to the beamline, will reconstruct some $\not{E}_T$ in the event. Also $Z \rightarrow \tau^+\tau^-$ events can fall into the signal region when both τ's decay leptonically. In this case the $\not{E}_T$ is due to the four neutrinos in the event. The background in the signal region due to these processes will be calculated simultaneously by considering the complete Drell-Yan spectrum.

To estimate the size of this background a sample of 6,000,000 Drell-Yan events were generated using the PYTHIA generator with both the Drell-Yan and Drell-Yan+jets

[2]In this section "Drell-Yan" will be used to refer to the entire invariant mass spectrum, while "Z" will be used to refer to only events with an invariant mass close to the mass of the Z boson.

processes enabled. Events with at least one electron or muon at generator level with $p_T > 15$ GeV/c and $|\eta| < 3.0$ were passed through a CDF detector simulation.

To verify that the Monte Carlo is doing a resonable job of modeling the data, the E_T of the two leading jets, the reconstructed $\not\!\!E_T$ spectrum, and the jet multiplicity distributions were compared to the data. Figure 5.6 shows this comparison after the same lepton identification and opposite charge cuts are applied to both the data and the Monte Carlo. The Monte Carlo sample is normalized to the number of events in the data. In this case the jet counting includes all jets with $E_T > 10$ GeV with $|\eta| < 2.0$. The invariant mass distribution for the lepton pairs is shown in Figure 5.7 for the sample as a whole and for the 0, 1, and ≥ 2 jet bins separately. Again the Monte Carlo is rescaled to the number of events seen in the data. There is some disagreement between the data and the Monte Carlo as to the relative numbers of events inside and outside the Z peak. Since the Z peak, which dominates the normalization, is cut away in the analysis, what is important is the shape outside the Z peak. To check how well the Monte Carlo does at reproducing the shape outside the Z window, the same plots are shown in Figure 5.8 except the normalization is now done based only on the bins in the region 40 GeV/$c^2 < M_{\ell\ell} <$ 80 GeV/c^2. Now the shape of the invariant mass spectra appears to match reasonably well above the Z peak. As a result this is the method used to normalize the background prediction.

To get a background prediction from this Monte Carlo sample, the normal analysis cuts are applied with the exception of the $\not\!\!E_T$ cut. The $\not\!\!E_T$ cut is not applied since it is not modeled particularly well in the Monte Carlo (see Figure 5.6). Instead, the $\not\!\!E_T$ cut is modeled by computing a rejection factor, defined below, from the Drell-Yan+2 jet data. The predicted background is given by:

$$B = N^{MC}_{tag,2jet} * R_{\not\!\!E_T} * \frac{N^{data}_{0jet}}{N^{MC}_{0jet}} \tag{5.4}$$

where $N^{MC}_{tag,2jet}$ is the number of tagged events with $\geq$2 jets found in the PYTHIA sample without the $\not\!\!E_T$ cut; $N^{data}_{0jet}/N^{MC}_{0jet}$ normalizes the Monte Carlo to the data in the 0 jet

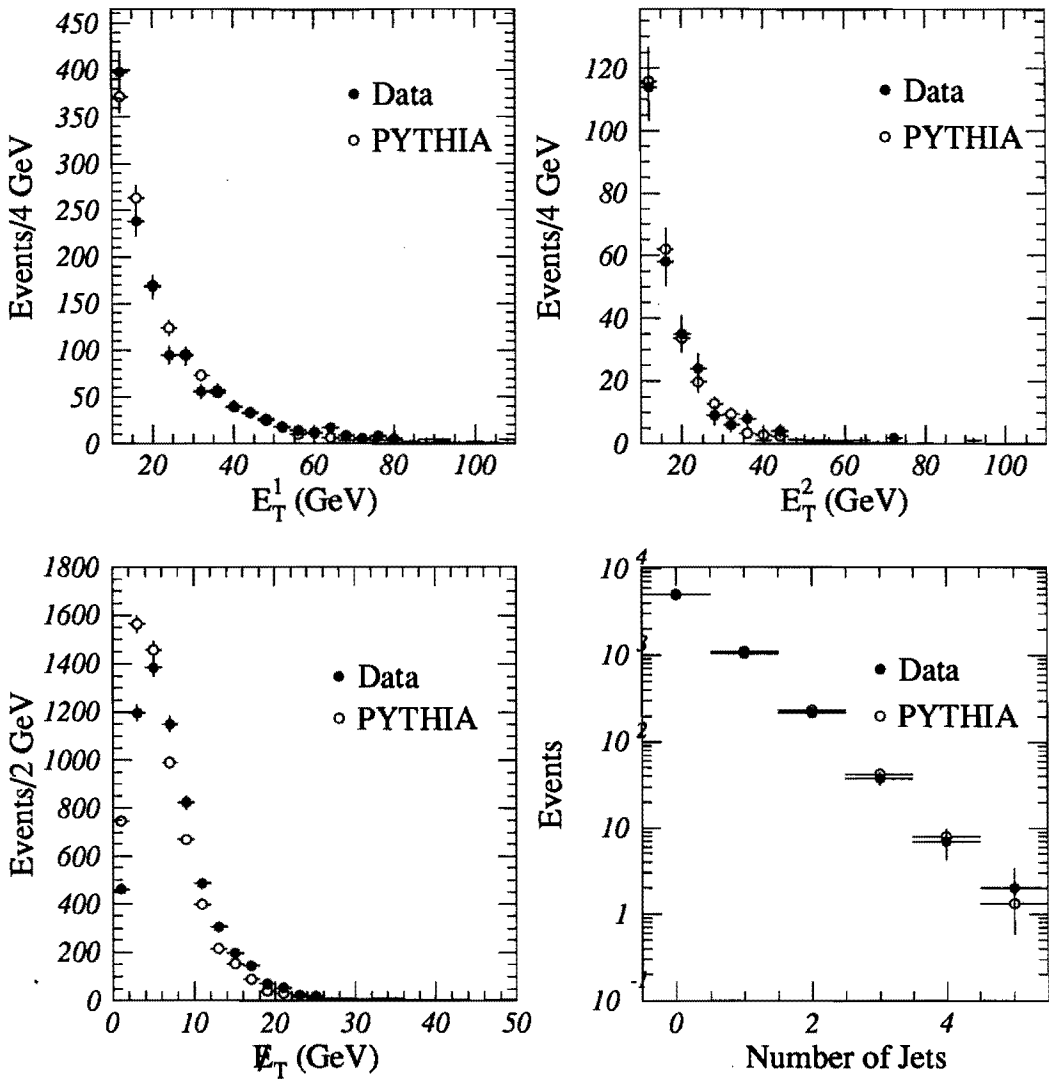

Figure 5.6: Comparison of the E_T^1, E_T^2, $\not{E}_T$, and N_{jets} spectra between PYTHIA Drell-Yan Monte Carlo and the data. Only the $\not{E}_T$ spectrum show much disagreement.

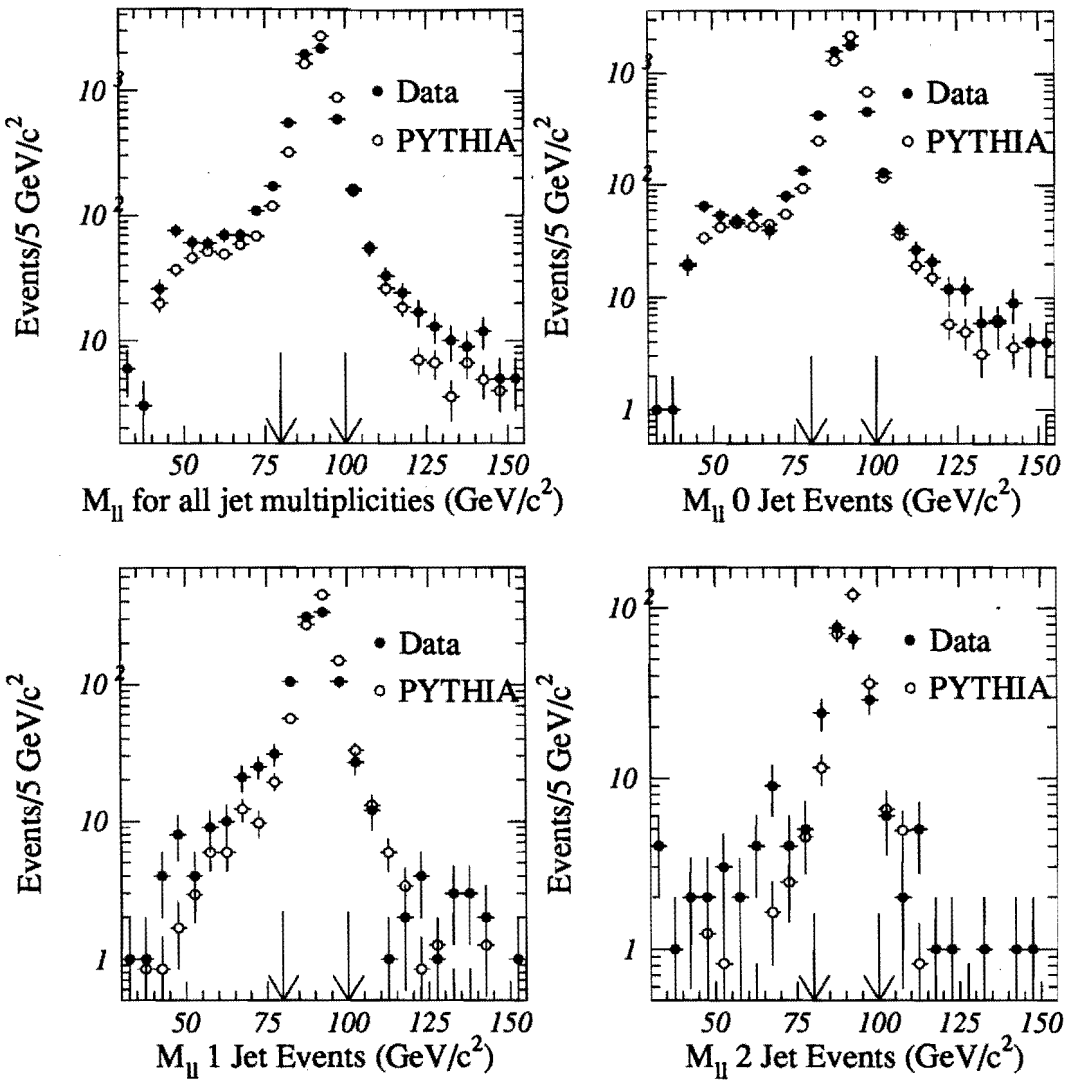

Figure 5.7: Comparison of invariant mass of two leptons between data and PYTHIA Drell-Yan Monte Carlo as a function of the number of jets with $E_T > 10$ GeV and $|\eta| < 2.0$. The arrows mark the region cut by the invariant mass cut.

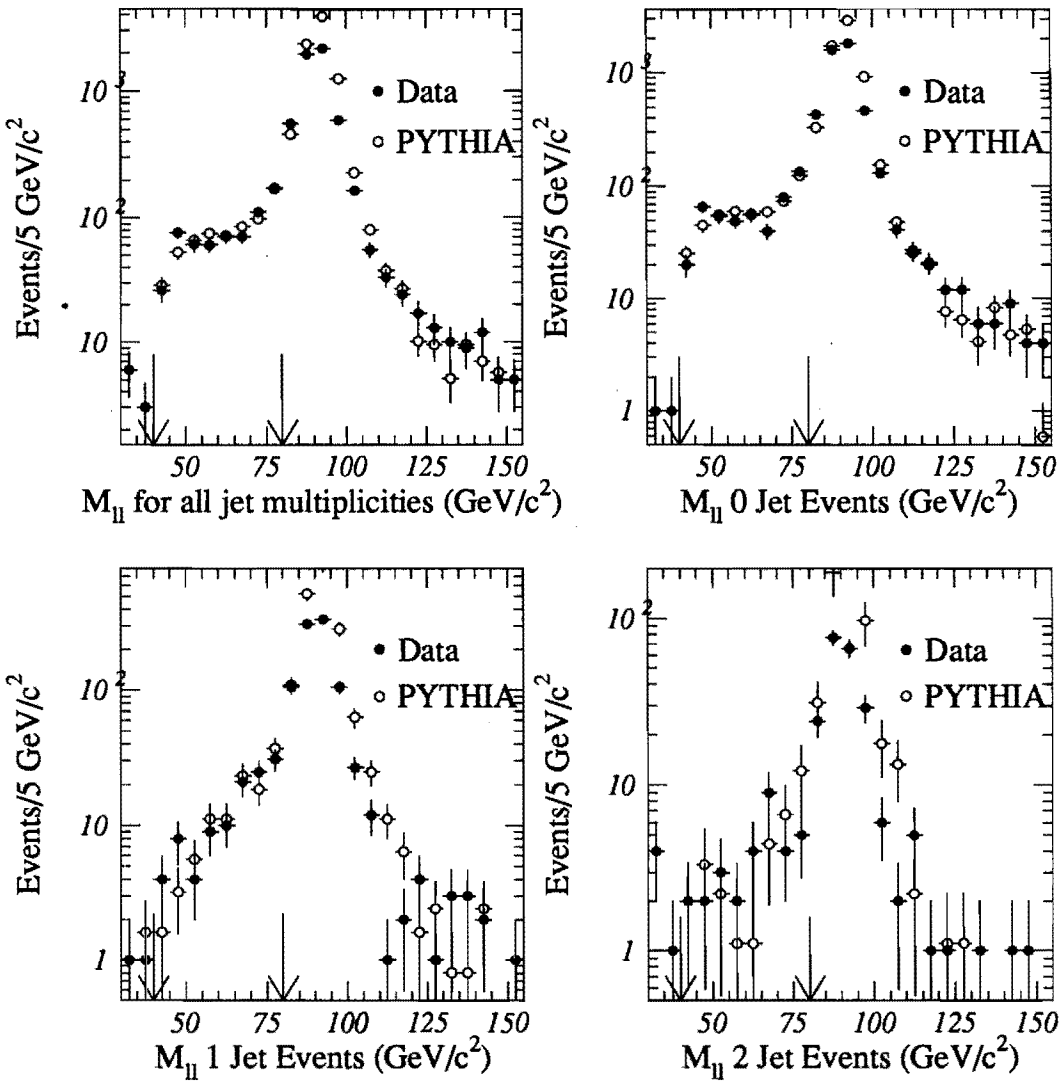

Figure 5.8: Comparison of invariant mass of two leptons between data and PYTHIA Drell-Yan Monte Carlo as a function of the number of jets with $E_T > 10$ GeV and $|\eta| < 2.0$. For this figure, the Monte Carlo is normalized in the mass range 40 GeV/c^2 $< M_{\ell\ell} <$ 80 GeV/c^2 which is between the arrows.

bin with no $\not\!\!E_T$ cut applied and using the low mass region; and $R_{\not\!\!E_T}$ is the $\not\!\!E_T$ rejection factor.

In the Monte Carlo sample there are 4.47 ± 2.05 SECVTX tags and $0.71 \pm .71$ SLT tags found. The number of tags has been corrected for several effects resulting in a non-integer number of tags. For the SECVTX algorithm, the expected mistag contribution calculated by applying the $-L_{xy}$ parameterization to the jets is subtracted since mistags are already accounted for by the calculation in Section 5.2.1 and the Monte Carlo tagging scale factor ($0.87 \pm .07$) (see Section D.1) has been applied. For the SLT algorithm, the number of tags in the Monte Carlo is overestimated because the detector simulation does not model some cuts well. This is taken into account by applying a known efficiency correction. There is no mistag correction for the SLT algorithm because on Monte Carlo events it only tags leptons from b or c decays. In both cases the detector simulation overestimates the efficiency of the primary lepton identification cuts (see Section 6.2) and this is taken into account here as well by using the lepton efficiency scale factors from Section 6.2. The normalization factor, $N_{0jet}^{data}/N_{0jet}^{MC}$, is given by 544/848. The $\not\!\!E_T$ rejection factor, $R_{\not\!\!E_T} = 39/188$, is found by looking at the fraction of $Z + 2$ jet data events which have $\not\!\!E_T > 20$ GeV. This results in an expected background of $0.60 \pm .30$ SECVTX tags and $0.094 \pm .094$ SLT tags. This calculation is clearly dominated by the statistics of the Monte Carlo samples.

5.2.4 W^+W^- Background

The W^+W^-+jets process can also mimic the $t\bar{t}$ signal. To get into the signal region both W bosons must decay leptonically and one of the jets must be tagged, the $\not\!\!E_T$ comes naturally from the leptonic W decays. The total W^+W^- production cross section has been calculated to be 9.53 pb using the HMRSB structure functions[30] and a $\pm 30\%$ systematic error which includes the choice of structure functions is taken on this calculation. Since little is known experimentally about W^+W^- production, two different Monte Carlo generators were used to estimate the background. The results of the two calculations are averaged to get the final expectation.

Using the PYTHIA Monte Carlo, 1,000,000 events were generated with the W^+W^- and W^+W^-+jets processes. Similarly a sample of 400,000 events was generated using the ISAJET Monte Carlo. In both cases, events with at least one electron or muon with $p_T > 15$ GeV/c and $|\eta| < 3.0$ were put through a detector simulation. To check that the results of the two Monte Carlo programs are in reasonable agreement, comparisons of some of the relevant kinematic quantities were done. The E_T of each of the two leading jets, the $\not{E}_T$ and the N_{jets} spectra for each of the two Monte Carlos is shown in Figure 5.9. The two Monte Carlos show some disagreement especially in the jet multiplicity. This could greatly affect the estimation of the background, however, only more data will determine which Monte Carlo program, if any, is giving the right jet multiplicity distribution.

Using the number of tags found in the Monte Carlo samples when the full analysis is applied (N_{tags}), the expected background can be calculated using the formula

$$B = N_{tags} * \frac{\sigma_{th} \int \mathcal{L} dt}{N_{gen}} \tag{5.5}$$

where the number of events generated is N_{gen}, σ_{th} is the theoretical cross section for inclusive W^+W^- production, and $\int \mathcal{L} dt$ is the total integrated luminosity of the data sample. In the PYTHIA sample there are 22.8 ± 4.9 SECVTX tags and $1.54 \pm .77$ SLT tags found after corrections. These result in background estimates of $0.024 \pm .009$ tags for SECVTX and $0.002 \pm .001$ for SLT. The ISAJET sample contains 16.1 ± 4.1 SECVTX tags which results in a background estimate of $0.042 \pm .014$ tags and an estimated background of $0.002 \pm .001$ SLT tags based upon the $0.62 \pm .44$ SLT tags found in the Monte Carlo sample. The difference between the predictions of the two Monte Carlos is due to differences in the fraction of W^+W^- events with two or more jets. To account for this the background prediction is taken to be the average and the difference between the average and the individual Monte Carlo results is taken as a systematic error. Averaging results in an estimate of $0.033 \pm .015$ SECVTX tags and $0.002 \pm .001$ SLT tags expected

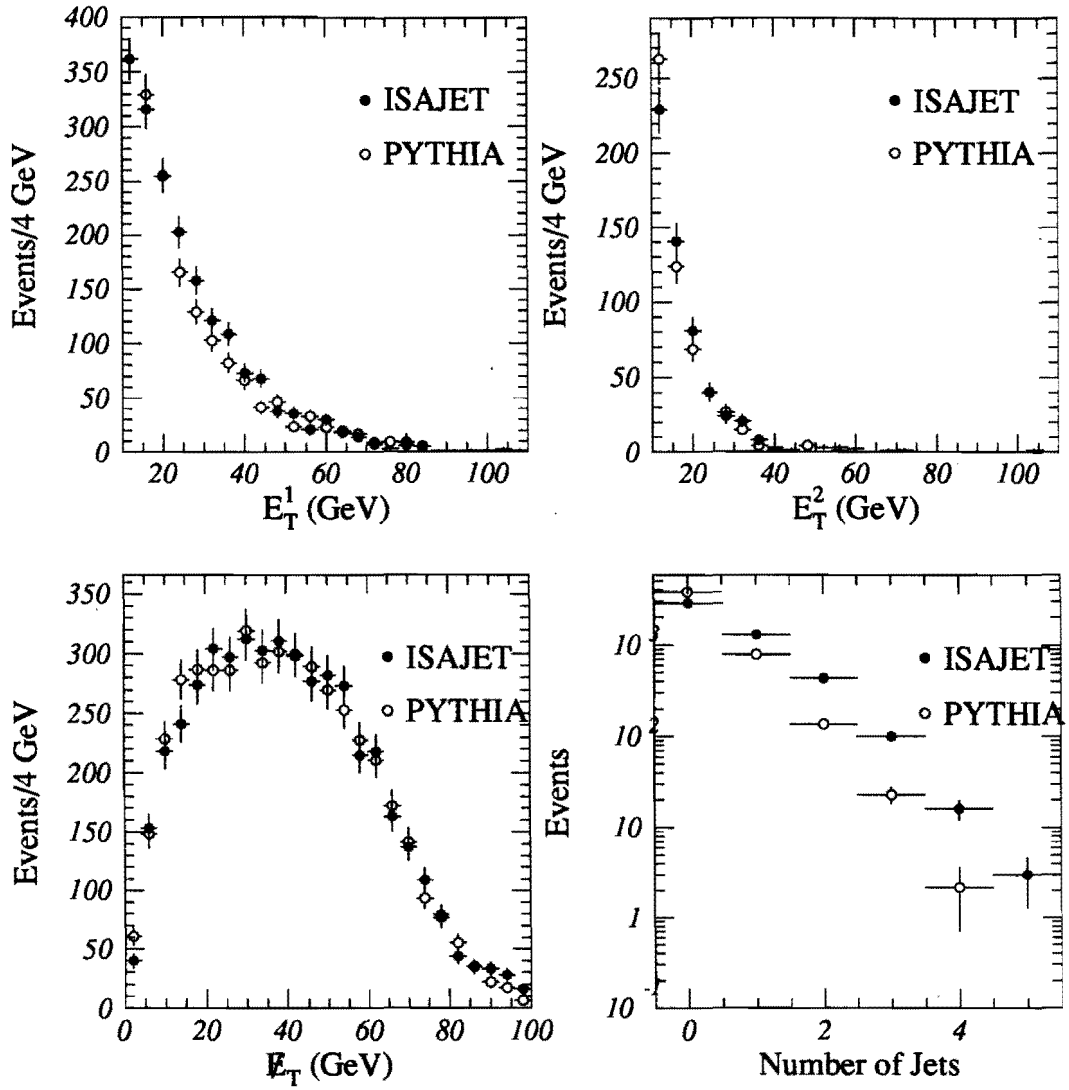

Figure 5.9: Comparison of the E_T^1, E_T^2, $\not{E}_T$, and N_{jets} spectra for the ISAJET and PYTHIA W^+W^- Monte Carlos.

	SECVTX	SLT	Combined
Mistags	$0.155 \pm .031$	$0.57 \pm .11$	$0.725 \pm .11$
Drell-Yan	$0.60 \pm .29$	$0.094 \pm .094$	$0.69 \pm .31$
Fake Leptons	$.027 \pm .009$	$.040 \pm .014$	$0.067 \pm .017$
W^+W^-	$.033 \pm .015$	$0.002 \pm .001$	$0.035 \pm .015$
Total	$0.82 \pm .29$	$0.71 \pm .15$	$1.53 \pm .33$

Table 5.4: Method 2 backgrounds for the dilepton+b-tag search shown for each tagging algorithm separately and for the combination.

in the signal region due to W^+W^- production. The errors include the $\pm 30\%$ uncertainty in the theoretical cross section.

5.2.5 Method 2 Total

Since the background estimates for each portion of the background are independent, the total background for each tagging algorithm is just the sum of all the parts. There would be an overlap if the mistag component had not been removed from the SECVTX calculations based upon the Monte Carlos. The method 2 backgrounds are summarized in Table 5.4 for each tagging algorithm individually and for the combination.

There has been no attempt to model the trigger in the Monte Carlo based calculations. However the trigger efficiencies should be essentially the same for the backgrounds as for $t\bar{t}$. The trigger efficiencies for $t\bar{t}$ are almost 100% (see Table 6.6) and here they are assumed to be one.

5.3 Summary

The background estimation for the search for $t\bar{t}$ pairs in the dilepton+b-tag channel has been done with two different methods. The first method which relies only upon $\bar{p}p$ data, results in an estimate of $1.35 \pm .11$ tags. The second method which attempts to make use of our knowledge of the theory, results in an estimate of $1.53 \pm .33$ tags. Since the

expected overlap between tags is small, there is little difference between the number of background tags expected and the number of background events expected and at this point the two are used interchangeably.

Chapter 6

Acceptance

The acceptance calculation is factorized into the following form:

$$A = A_{geom} \times \epsilon_{ID} \times \epsilon_{kin} \times \epsilon_{trig} \times \epsilon_{tag}. \tag{6.1}$$

The geometric acceptance for pairs of reconstructed leptons with $p_T >$ 20 GeV/c is given by A_{geom}, the lepton identification and isolation efficiencies combined for the two leptons is given by ϵ_{ID}, the efficiency of the opposite charge and the kinematic cuts is given by ϵ_{kin}, the trigger efficiency for the two leptons is given by ϵ_{trig}, and the efficiency for tagging at least one of the b's by either the SLT or the SECVTX algorithm is given by ϵ_{tag}. The acceptance calculation is done for each combination of lepton pairs (CEMU, MUCE, CEMI, CECE, MUMU, and MUMI where CE is a central electron, MU is a CMUO type muon, and MI is a CMIO type muon) independently. In this nomenclature, the first lepton must pass the tight lepton identification and isolation cuts, while the second lepton must pass the loose lepton identification cuts. The total acceptance is just the sum over the lepton combinations.

The acceptance for $t\bar{t}$ events is measured as a function of the mass of the top quark using three samples of $t\bar{t}$ Monte Carlo generated using PYTHIA/CLEO/TAUOLA. Samples of about 40,000 events were generated using m_{top} = 160, 175, and 180 GeV/c^2. These masses are centered around the previously published mass measurement from CDF[2].

Category	$A_{geom}(160)$	$A_{geom}(175)$	$A_{geom}(190)$
CECE	$.0079 \pm .0004$	$.0089 \pm .0005$	$.0088 \pm .0005$
CEMU	$.0164 \pm .0006$	$.0182 \pm .0007$	$.0204 \pm .0007$
MUCE	$.0164 \pm .0006$	$.0182 \pm .0007$	$.0204 \pm .0007$
CEMI	$.0041 \pm .0003$	$.0038 \pm .0003$	$.0042 \pm .0004$
MUMU	$.0074 \pm .0004$	$.0088 \pm .0005$	$.0097 \pm .0005$
MUMI	$.0040 \pm .0003$	$.0058 \pm .0004$	$.0073 \pm .0005$

Table 6.1: Geometrical acceptance for two leptons with $p_T > 20$ GeV/c determined using the PYTHIA sample as a function of m_{top}. Note that the CEMU and MUCE categories are equivalent at this point.

Some parts of the calculation are not modeled particularly well by the detector simulation and information from data samples will be used to correct these efficiencies.

6.1 Geometric Acceptance

The geometric acceptance represents the fraction of $t\bar{t}$ events which are expected to have two leptons each having $p_T > 20$ GeV/c and passing fiducial cuts. For the electron and CMIO muon cases, the fiducial cuts mean the lepton does not hit the calorimeter face too close to the edge of a tower. For the muon case, the fiducial cut means that a track is reconstructed in a muon system with a CTC track pointing to it. The geometric acceptance calculation begins by counting the number of events in the Monte Carlo samples which have two fiducial leptons with $p_T > 20$ GeV/c (N_{geom}). Dividing N_{geom} by the number of events generated in the sample gives the geometric acceptance. The acceptance calculated in this fashion includes the assumed Standard Model branching ratio for $t\bar{t}$ into e, μ pairs, 4/81. Table 6.1 lists the geometric acceptance for the PYTHIA Monte Carlo as a function of m_{top}.

Category	ϵ_{ID}^{MC}	$R(\ell_1)R(\ell_2)$	ϵ_{ID}
CECE	$0.525^{+.027}_{-.028}$	$0.910 \pm .007$	$0.478^{+.025}_{-.026}$
CEMU	$0.368^{+.019}_{-.018}$	$0.917 \pm .006$	$0.338^{+.018}_{-.017}$
MUCE	$0.080^{+.011}_{-.010}$	$0.929 \pm .005$	$0.074^{+.010}_{-.009}$
CEMI	$0.212^{+.038}_{-.034}$	$1.040^{+.036}_{-.035}$	$0.221^{+.041}_{-.036}$
MUMU	$0.496 \pm .028$	$0.936 \pm .004$	$0.464 \pm .026$
MUMI	$0.247^{+.031}_{-.029}$	$1.062^{+.036}_{-.035}$	$0.262^{+.034}_{-.032}$

Table 6.2: Lepton identification efficiency for $m_{top} = 175$ GeV/c^2. Note that there is no overlap between the CEMU and MUCE categories After the lepton identification cuts are made. Most event which contain a tight CEM electron and a tight CMUO muon are classified in the CEMU category for historical reasons.

6.2 Lepton Identification Efficiency

The lepton identification efficiency, ϵ_{ID}, is the efficiency for a pair of fiducial leptons with $p_T >$ 20 GeV/c to pass the lepton identification cuts including the isolation cuts but not including the opposite charge requirement. A first approximation (ϵ_{ID}^{MC}) can be made by just counting the number of events in the $t\bar{t}$ Monte Carlo which pass the lepton identification and isolation cuts, N_{ID}, and then dividing by N_{geom}. The result of this approach is shown on the left in Table 6.2 for $m_{top} = 175$ GeV/c^2. However, using this number does not take into account the fact that the detector simulation does not model some of the variables used for lepton identification particularly well. Figure 6.1 shows $(N^i_{data} - N^i_{MC})/N_{data}$, where N^i_{MC} (N^i_{data}) in the number of data (Monte Carlo) events in the ith bin and N_{data} is the number of data events in the plot, for some of the lepton identification variables. The discrepancies in the cut regions contribute to the differences in lepton identification efficiency seen between the data and the Monte Carlo.

The $t\bar{t}$ efficiency derived from the Monte Carlo is corrected by two quantities

$$\epsilon_{ID} = \epsilon_{ID}^{MC} R(\ell_1)R(\ell_2). \tag{6.2}$$

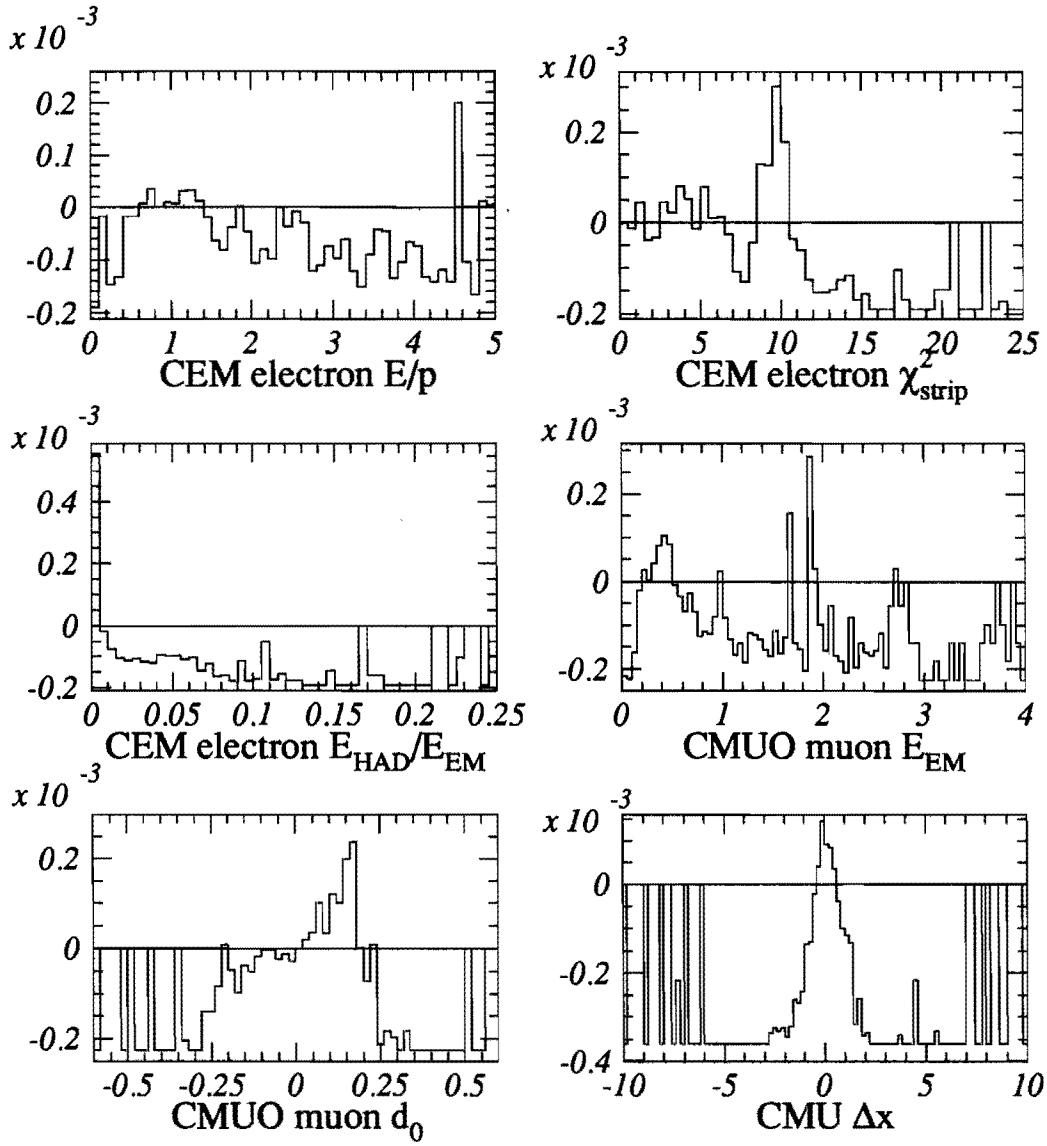

Figure 6.1: Plots showing $(N^i_{MC} - N^i_{data})/N_{data}$ which emphasize the discrepancy between data and Monte Carlo for some of the lepton identification variables for CEM electrons and CMUO muons.

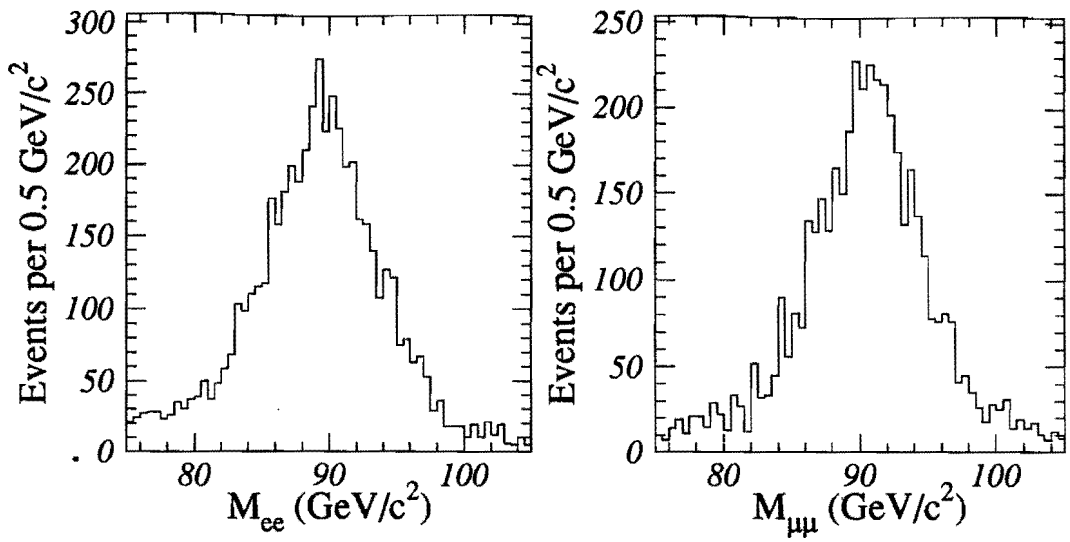

Figure 6.2: Invariant mass distributions for ee pairs (left) and $\mu\mu$ pairs (right) used to calculate the single lepton identification efficiency. One lepton passes the tight lepton identification cuts while the second leg has only fiducial, an opposite charge, and p_T>20 GeV/c or E_T<20 GeV cuts applied.

The $R(\ell)$ defined as the ratio of the single lepton identification efficiency measured in Z data to the single lepton identification efficiency measured in Z Monte Carlo and depends only on the lepton type. The Z events used to derive the correction factor were selected by looking for events with one lepton which was required to pass all tight lepton identification cuts except the isolation cut, the second was required to pass only the fiducial and p_T (E_T) cuts, and the pair was required to have opposite charge and an invariant mass in the range 75 GeV/c^2 < $M_{\ell\ell}$ < 105 GeV/c^2. The invariant mass distributions for $e-e$ and $\mu-\mu$ pairs passing these cuts is shown in Figure 6.2. The Monte Carlo is the PYTHIA sample used in Section 5.2.3 with the same cuts as applied to the data. Since one of the leptons is required to pass the tight lepton identification cuts, the other lepton is used to determine the efficiency of the identification cuts. When calculating the efficiency of tight lepton cuts where isolation is required, both leptons are required to pass the isolation requirements.

To determine the single tight CEM electron and both tight and loose CMUO muon cut efficiencies, the following procedure is used. There are four combinations of passing the cuts that a pair of leptons can have: PP, PF, FP, FF where P implies passing the cuts and F implies failing the cuts. The first three combinations have at least one lepton passing the tight identification cuts and will be in the Z sample, but no events of the last type (FF) will be in the Z sample. The ratio of the number of events with two tight leptons to the number of events with at least one tight lepton is given by:

$$R_T = \frac{\epsilon_T^2}{\epsilon_T^2 + \epsilon_T(1-\epsilon_T) + (1-\epsilon_T)\epsilon_T} \tag{6.3}$$

where ϵ_T is the efficiency for a single isolated, fiducial lepton with $p_T >$ 20 GeV/c to pass the tight lepton identification cuts. Solving for ϵ_T gives:

$$\epsilon_T = \frac{2R_T}{1+R_T}. \tag{6.4}$$

The single lepton identification efficiency for tight and loose CMUO muons and tight CEM electrons is given in Table 6.3 along with the ratio $R(\ell)$.

Using this same form to measure the loose single CEM identification efficiency could result in the sample having a large background component. Since the leptons in background events could have a different efficiency than those from Z decays, this would need to be corrected for and is difficult. Instead an approach similar to that followed for the tight leptons is taken. Here there are three possible outcomes for each lepton: passes the tight cuts, passes the loose cuts but fails the tight cuts, and fails the loose cuts. The ratio of number of events with one tight lepton and a second loose lepton (which can also pass the tight identification cuts since the tight cuts are a subset of the loose cuts) divided by the number of events with at least one tight lepton can be written as

$$R_L = \frac{\epsilon_T^2 + \epsilon_T(\epsilon_L - \epsilon_T) + (\epsilon_L - \epsilon_T)\epsilon_T}{\epsilon_T^2 + \epsilon_T(1-\epsilon_T) + (1-\epsilon_T)\epsilon_T} \tag{6.5}$$

Category	ϵ_{ID}^{Data}	ϵ_{ID}^{MC}	$R(\ell)$
Tight CEM	$0.850 \pm .004$	$0.894 \pm .002$	$0.951 \pm .006$
Loose CEM	$0.873 \pm .004$	$0.911 \pm .002$	$0.958 \pm .004$
Tight MU	$0.959 \pm .002$	$0.987 \pm .001$	$0.986 \pm .003$
Loose MU	$0.952 \pm .003$	$0.987 \pm .001$	$0.985 \pm .003$
MI	$0.447 \pm .013$	$0.409 \pm .006$	$1.094 \pm .039$

Table 6.3: Single lepton identification efficiency measured using data and Monte Carlo Z events. The ratio of efficiencies for Data/Monte Carlo, $R(\ell)$, is also given.

where ϵ_L is the efficiency of the loose cuts and ϵ_T is the efficiency of the tight lepton cuts. Solving this for ϵ_L gives

$$\epsilon_L = \frac{(R_L(2 - \epsilon_T) + \epsilon_T)}{2}. \tag{6.6}$$

The single loose CEM lepton identification efficiency is given in Table 6.3.

The single lepton identification efficiency for CMIO muons, which are always considered a loose lepton, is straight forward. The efficiency is simply the ratio of the number of events with one CMUO muon passing the tight lepton identification cuts and one CMIO muon passing its identification cuts to the number of events with one tight CMUO muon and a track passing the CMIO p_T and fiducial requirements.

The single lepton scale factors combined in the proper combinations to apply to the $t\bar{t}$ Monte Carlo are shown in the middle column of Table 6.2. The final dilepton identification and isolation efficiency for $m_{top} = 175$ GeV/c^2 are shown on the right. The final lepton identification and isolation efficiency as a function of the mass of the top quark is summarized in Table 6.4.

6.3 Kinematic Cuts

This section describes the efficiency of the combination of the the kinematic cuts from Section 4.2 and the opposite charge requirement. Again this measurement is done using the PYTHIA Monte Carlo samples. The efficiency is calculated by dividing the number

Category	$\epsilon_{ID}(160)$	$\epsilon_{ID}(175)$	$\epsilon_{ID}(190)$
CECE	$0.571^{+.026}_{-.026}$	$0.478^{+.025}_{-.026}$	$0.459^{+.027}_{-.027}$
CEMU	$0.346^{+.018}_{-.018}$	$0.338^{+.018}_{-.017}$	$0.306^{+.017}_{-.016}$
MUCE	$0.082^{+.011}_{-.010}$	$0.074^{+.010}_{-.009}$	$0.068^{+.010}_{-.009}$
CEMI	$0.251^{+.039}_{-.036}$	$0.221^{+.041}_{-.036}$	$0.239^{+.040}_{-.036}$
MUMU	$0.481^{+.028}_{-.028}$	$0.464^{+.026}_{-.026}$	$0.416^{+.025}_{-.025}$
MUMI	$0.299^{+.043}_{-.040}$	$0.262^{+.034}_{-.032}$	$0.199^{+.028}_{-.026}$

Table 6.4: Lepton identification and isolation efficiency as a function of m_{top}.

of events passing all cuts except the b-tag (N_{kin}) by N_{ID} defined in the previous section. The efficiency is summarized in Table 6.5 as a function of the top quark mass.

6.4 Trigger Efficiencies

The data used in this analysis is derived primarily from the single electron and muon triggers. Which lepton in a $t\bar{t}$ event actually caused the trigger is not important so that the trigger efficiency for the event is just

$$1 - (1 - \epsilon_1)(1 - \epsilon_2) \tag{6.7}$$

Category	$\epsilon_{kin}(160)$	$\epsilon_{kin}(175)$	$\epsilon_{kin}(190)$
CECE	$0.375^{+.037}_{-.035}$	$0.418^{+.039}_{-.038}$	$0.415^{+.041}_{-.040}$
CEMU	$0.697^{+.030}_{-.031}$	$0.810^{+.024}_{-.027}$	$0.793^{+.026}_{-.028}$
MUCE	$0.607^{+.068}_{-.072}$	$0.525^{+.072}_{-.073}$	$0.828^{+.051}_{-.064}$
CEMI	$0.905^{+.045}_{-.069}$	$0.758^{+.079}_{-.098}$	$0.919^{+.044}_{-.073}$
MUMU	$0.538^{+.042}_{-.043}$	$0.522^{+.040}_{-.040}$	$0.569^{+.041}_{-.042}$
MUMI	$0.646^{+.075}_{-.082}$	$0.569^{+.072}_{-.074}$	$0.604^{+.074}_{-.078}$

Table 6.5: Kinematic acceptance for PYTHIA $t\bar{t}$ Monte Carlo as a function of the top quark mass.

Category	ϵ_{trig}
CECE	$0.9998 \pm .0001$
CEMU	$0.9981 \pm .0005$
MUCE	$0.9981 \pm .0005$
CEMI	$0.987 \pm .003$
MUMU	$0.979 \pm .003$
MUMI	$0.854 \pm .009$

Table 6.6: Trigger efficiencies for $t\bar{t}$.

where ϵ_1 and ϵ_2 are the two trigger efficiencies relevant for that event. Note that the CMIO type muons, which consist of a stiff track pointing to very little energy in the calorimeter, are not triggered upon. Thus the trigger efficiencies for the CEMI and MUMI categories are just the electron and muon trigger efficiencies respectively.

The single lepton trigger efficiencies have been determined by looking at samples of events which pass at least one trigger other than the trigger being examined. For example to measure the muon trigger efficiency, a sample of events passing a jet trigger could be used. The electron trigger was measured to be $98.7 \pm .3\%$ efficient while the single muon trigger efficiency is $85.4 \pm .9\%$. These triggers efficiencies are then combined to form the dilepton trigger efficiencies which are summarized in Table 6.6. Note that the triggers for single leptons are all designed to have a constant efficiency for $p_T > 20$ GeV/c. This implies that the trigger efficiency for each dilepton pair should not be dependent upon the mass of the top quark.

6.5 b-tagging

To measure the efficiency of the tagging algorithms a sample of $t\bar{t}$ Monte Carlo events generated using PYTHIA with $m_{top} = 175$ GeV/c^2 is used. After processing with a detector simulation, the number of events which pass all selection cuts except the tag is counted (481). In this sample there are 189 SECVTX and 93.7 SLT tagged events. The tagging efficiency for each algorithm to tag a $t\bar{t}$ event is the number of tagged events divided

Category	$A_{tot}(160)$	$A_{tot}(175)$	$A_{tot}(190)$
CECE	$0.00079 \pm .00012$	$0.00083 \pm .00012$	$0.00078 \pm .00012$
CEMU	$0.00186 \pm .00021$	$0.00234 \pm .00026$	$0.00232 \pm .00026$
MUCE	$0.00038 \pm .00008$	$0.00033 \pm .00007$	$0.00054 \pm .00010$
CEMI	$0.00043 \pm .00009$	$0.00030 \pm .00007$	$0.00042 \pm .00009$
MUMU	$0.00088 \pm .00013$	$0.00098 \pm .00014$	$0.00106 \pm .00015$
MUMI	$0.00031 \pm .00007$	$0.00035 \pm .00007$	$0.00035 \pm .00008$
Total	$0.00465 \pm .00031$	$0.00513 \pm .00034$	$0.00547 \pm .00036$
Systematic error	$\pm .00065$	$\pm .00072$	$\pm .00077$

Table 6.7: Total acceptance for $t\bar{t}$ events from PYTHIA $t\bar{t}$ Monte Carlo as a function of top quark mass.

by the number of events. Note that for the SECVTX algorithm, this must be corrected to account for the over-efficiency of the algorithm in the Monte Carlo $0.87 \pm .07$ (see section D.1). The individual tagging efficiencies per event are found to be $0.342 \pm .034$ for SECVTX and $0.195 \pm .20$ for SLT. To calculate the total tagging efficiency per event, the efficiencies are summed and the overlap is removed:

$$\epsilon_{tag} = \epsilon_{SVX} + \epsilon_{SLT} - \epsilon_{SVX}\epsilon_{SLT} \tag{6.8}$$

This results in an efficiency of $0.470 \pm .040$ for one of the two algorithms to tag a $t\bar{t}$ event.

6.6 Total Acceptance

The acceptance is first calculated on a channel by channel basis according to Equation 6.1. The total acceptance is just the sum of all channels. Table 6.7 summarizes the total efficiency for $t\bar{t}$ as a function of top mass. Several additional studies have been done to estimate possible systematic effects on the acceptance.

To calculate the acceptance, a specific generator has been chosen. To get an idea of how the generator selection affects the total acceptance, a sample of 210,526 events generated with HERWIG and a sample of 36,000 events generated with ISAJET using $m_{top} = 175$ GeV/c^2 were examined. After applying all analysis cuts except the b-tag,

there are 2297 (392) events remaining in the HERWIG (ISAJET) sample. The PYTHIA sample has 482 out of 40688 events remaining after all cuts are applied except the b-tag. The maximum difference between the fraction of events passing is taken as a systematic error on the total acceptance ($\pm 8\%$).

Additionally, there is no direct evidence which tells how much initial and final state gluon radiation there should be in an event. To look at these effects, three samples with $m_{top} = 175$ GeV/c^2 were generated using the PYTHIA Monte Carlo. A sample of 40,688 events were generated just as the default sample was except the mass, 482 passed all cuts except the b-tag. A second sample of 38,114 events were generated with the final state gluon radiation process turned off, 440 passed the cuts. The third sample made with the initial state radiation disabled consisted of 42123 events, 431 of which passed the cuts. In each of the two cases the difference from the default radiation settings is taken as a systematic error. This results in $\pm 3\%$ for final state radiation and $\pm 14\%$ for initial state radiation.

Taking the largest of these as a systematic error on the total acceptance results in an estimated $\pm 14\%$ uncertainty. The systematic uncertainty is shown on the final line of Table 6.7.

Chapter 7

Results

The selection cuts described in Chapter 4 were applied to the inclusive central electron and inclusive central muon datasets to select top candidate events. The number of events remaining after each of the cuts is summarized in Table 7.1. The bottom line is that there are four events which contain six b-tags. These 4 events are shown in Figures 7.1-7.4 in both an end view and in a lego display of the E_T measured in the calorimeter.

From the method 2 background prediction from Section 5.2, $1.53 \pm .33$ tagged events are expected in this sample due to backgrounds. When compared to the observed 4 events, a small excess is observed. The statistical significance of this excess can be calculated by looking at how often the predicted background fluctuates to greater than or equal to the observed signal. In doing this calculation, no correction is made to the

Selection Criteria	Number of Events Remaining
Lepton Identification and Isolation	8873
Opposite charge	8856
Invariant Mass	1171
$\not\!\!E_T$	77
2 jets	16
b-tag	4

Table 7.1: Number of event remaining after each of the cuts in the analysis.

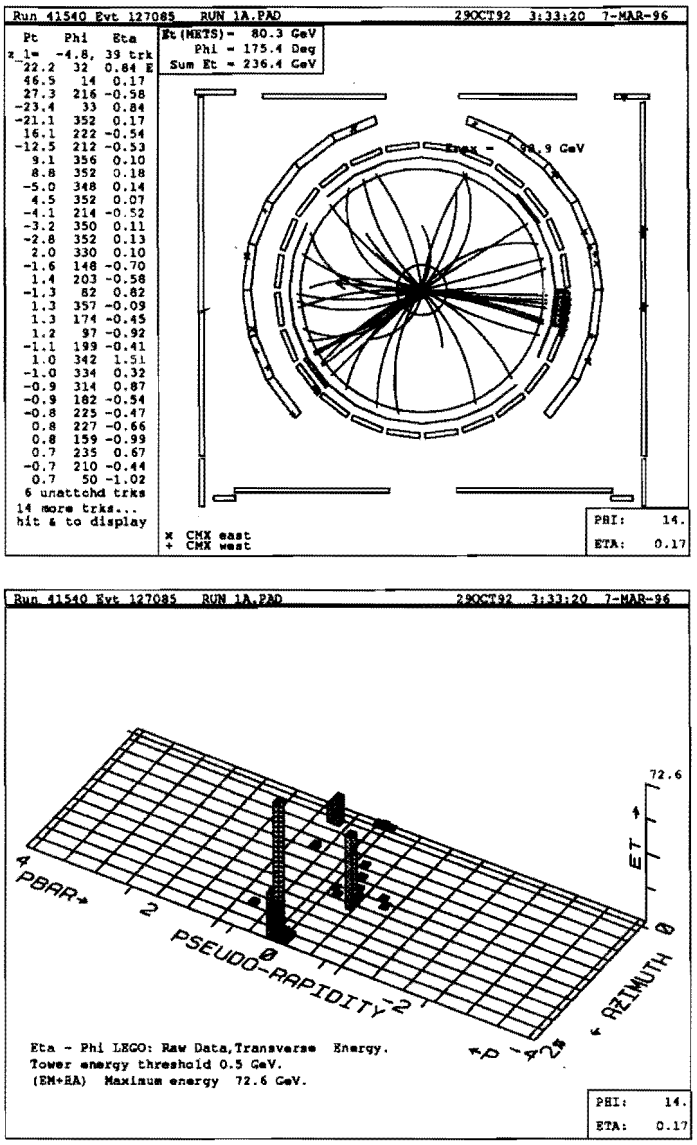

Figure 7.1: CDF event displays showing end view and lego plot of the calorimeter for Run 41540 Event 127085.

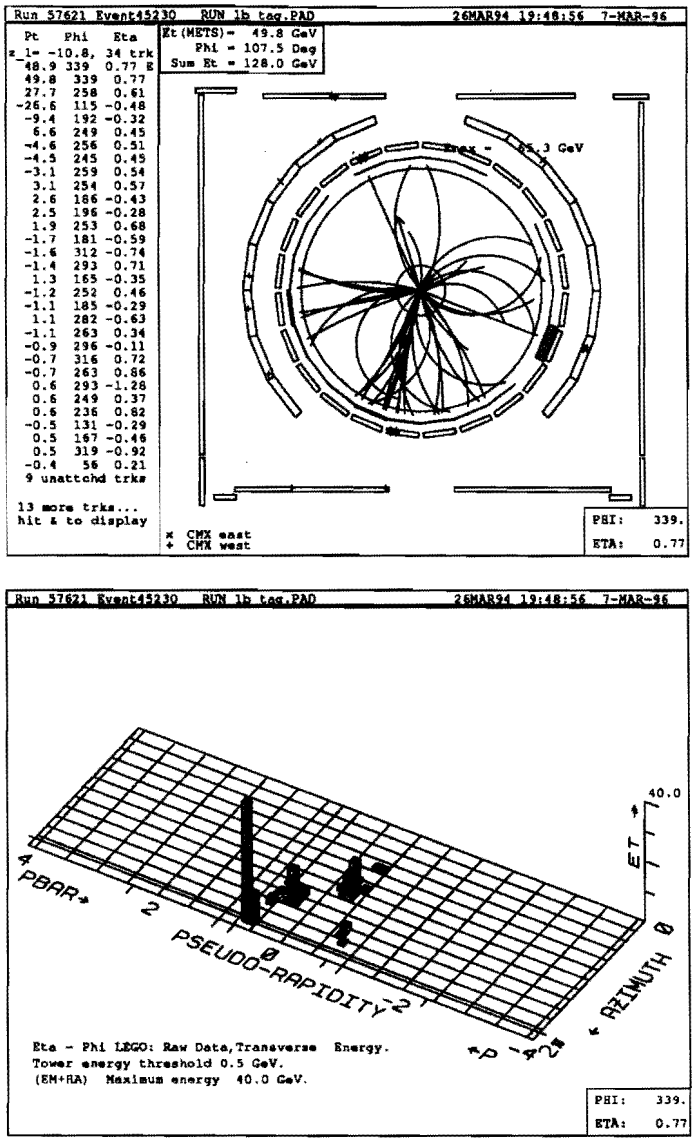

Figure 7.2: CDF event displays showing end view and lego plot of the calorimeter for Run 57621 Event 45230.

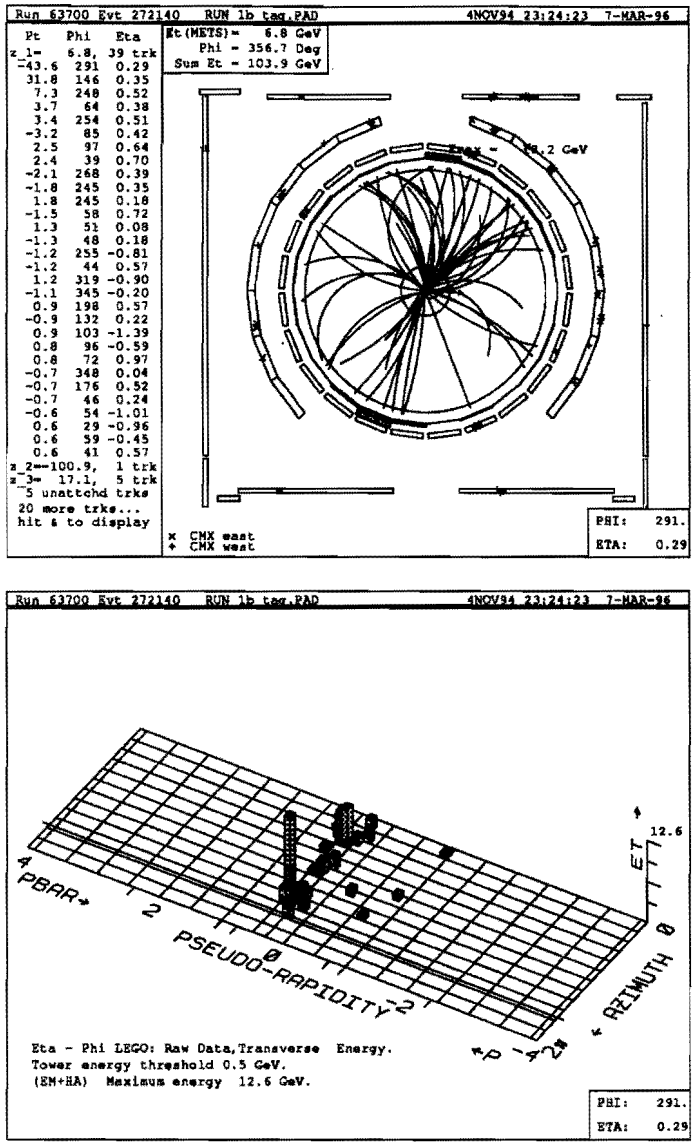

Figure 7.3: CDF event displays showing end view and lego plot of the calorimeter for Run 63700 Event 272140.

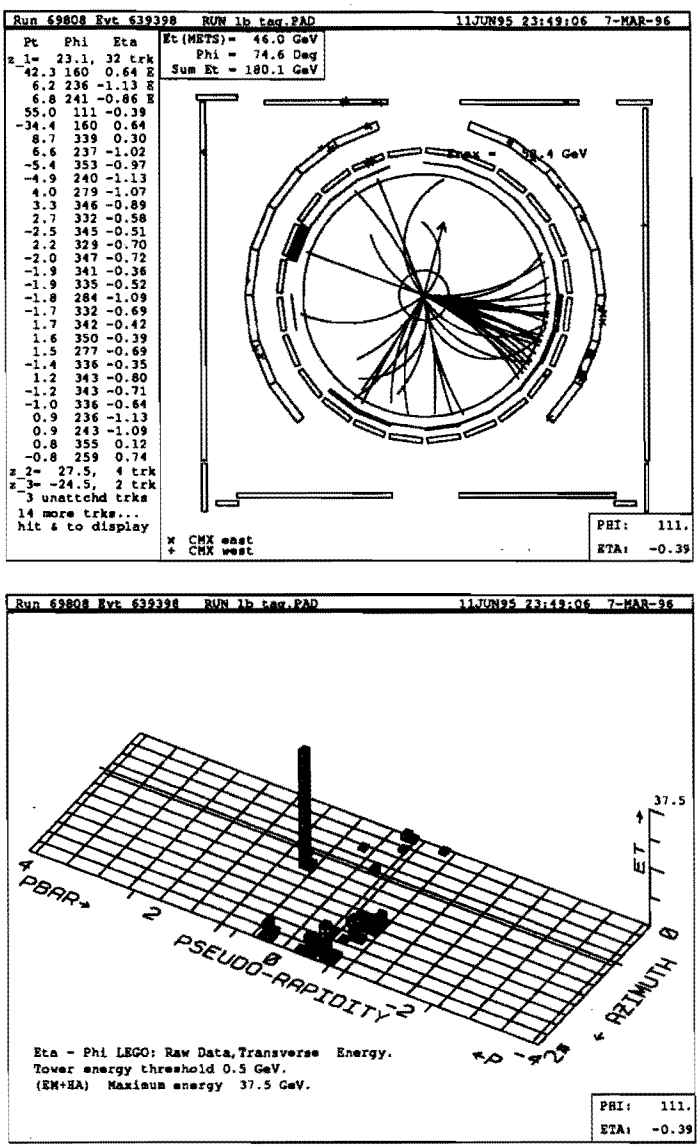

Figure 7.4: CDF event displays showing end view and lego plot of the calorimeter for Run 69808 Event 639398.

expected background to account for the presence of top in the sample as this is the most conservative approach. A toy Monte Carlo is used to calculate the probability by counting the fraction of events with a poisson distribution with a mean equal to the background which is greater than or equal to the observed number of events. For the method 2 background prediction the statistical significance of the observed excess is found to be .076.

Calculating the significance using the number of events found is expected to be an underestimate of the real statistical significance because there is additional information contained in the double tagged events. Calculating the significance using the 6 tags found and assuming that the number of tags in the background and the number of background events is the same results in a significance of 6.8×10^{-3}. This is in fact expected to be a slight overestimate because it does not take into account tag correlations in the background. (If the number of Drell-Yan events with a $b\bar{b}$ pair fluctuates high, this will affect both the SECVTX and SLT background calculations in the same way.) Check on the effect of this correlation indicate that it is indeed a small effect.

These events all contain two leptons and $\not{E}_T$ which are consistent with coming from W^+W^- pairs. There are also two jets, at least one of which is consistent with having come from a b quark, in the event. The most natural interpretation of these events is as a $t\bar{t}$ pair which decayed into $W^+\bar{b}W^-b$ with both W's decaying to electrons and muons. Proceeding under the assumption that these events are from top, the cross production cross section can be calculated. Before calculating the $t\bar{t}$ production cross section implied by the excess, the background contribution from SECVTX mistags must be corrected since the calculation was done under the assumption that there is no $t\bar{t}$ in the sample. The presence of $t\bar{t}$ will tend to cause an overestimation of the background contribution from the mistags. There should also be a small correction to the the fake primary lepton background, however since this background is only about 4% of the total the correction would be negligible and is not considered further.

The correction to the mistag background is made using a simple iteration procedure:

$$N_{t\bar{t}} = (N_{obs} - B_{tot})/\epsilon_{tag} \tag{7.1}$$

$$B_{cor} = (16 - N_{t\bar{t}})\frac{B_{cor}}{16} \tag{7.2}$$

$$B_{tot} = B_{cor} + B_{uncor} \tag{7.3}$$

where N_{obs} is the number of tagged events seen, B_{cor} is the portion of the background which must be corrected and B_{uncor} is the fixed portion from Monte Carlo studies. The 16 in the above equations is the number of events found before the tagging algorithms are applied. Only the portions of the background which are derived from the events before tagging are corrected, the others are absolute predictions. The iteration procedure loops around the above equations until the background changes by less than 0.01 event.

For the SECVTX algorithm, B_{cor} is initially 0.155 (the mistag contribution) and B_{uncor} is 0.66. This results in a corrected background of $0.72 \pm .30$ events where the error is a combination of the fractional error before correction plus a 10% error from the tagging efficiency. Adding in the SLT background results in a total estimated background of $1.43 \pm .33$ events. The SLT mistag background does not need a correction applied to it because the acceptance is calculated using only leptons coming from the decay of either a b or c quark.

After showing that there is a signal and assuming that it is $t\bar{t}$, the cross section can be calculated. The cross section and its associated uncertainties are calculated using the following likelihood function:

$$L = e^{-\frac{(\int \mathcal{L}dt - \overline{\int \mathcal{L}dt})^2}{2\sigma_{\mathcal{L}}^2}} e^{-\frac{(A-\overline{A})^2}{2\sigma_A^2}} e^{-\frac{(B-\overline{B})^2}{2\sigma_B^2}} \frac{(A \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt + B)^N}{N!} e^{-(A \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt + B)} \tag{7.4}$$

where A is the total acceptance, B is the expected background, N is the number of observed candidate events, $\sigma_{t\bar{t}}$ is the $t\bar{t}$ production cross section, and $\overline{\int \mathcal{L}dt} = 109.4 \pm 7.2$ pb^{-1} is the total integrated luminosity. The production cross section is found by minimizing $-\ln L$. The parameters A, B, and $\int \mathcal{L}dt$ are initially set to their expected values, $\overline{A}$,

m_{top}	160 GeV/c^2	175 GeV/c^2	190 GeV/c^2
$\sigma_{t\bar{t}}$	$5.0^{+4.9}_{-3.4}$ pb	$4.6^{+4.4}_{-3.1}$ pb	$4.3^{+4.2}_{-2.9}$ pb

Table 7.2: Measured cross section for $t\bar{t}$ production as a function of m_{top}.

$\overline{B}$ and $\overline{\int \mathcal{L}dt}$ and allowed to vary according to their uncertainties. The uncertainties on the cross section are the points where the $-\ln L$ has changed by 1/2 from its minimum value. The minimization is done using the MINUIT[31] package. This procedure reduces to the form:

$$\sigma_{t\bar{t}} = \frac{N - \overline{B}}{\overline{A} \int \mathcal{L}dt} \tag{7.5}$$

for the central value, however the formal expression is necessary to get the proper uncertainties.

Using this formalism the cross section is calculated as a function of m_{top}. The calculation uses the total integrated luminosity 109.4 ± 7.2 pb^{-1} the total corrected background $1.43 \pm .33$, and the acceptance as a function of m_{top} from Table 6.7 to give $\sigma_{t\bar{t}} = 4.6^{+4.4}_{-3.1}$ pb for an assumed $m_{top} = 175$ GeV/c^2. This value of m_{top} is chosen because analyses to measure the mass of the top quark find $m_{top} = 176 \pm 9$ GeV/c^2[32]. The result is summarized in Table 7.2 an shown in comparison to the theory predictions in Figure 7.5 for top masses around the expected mass.

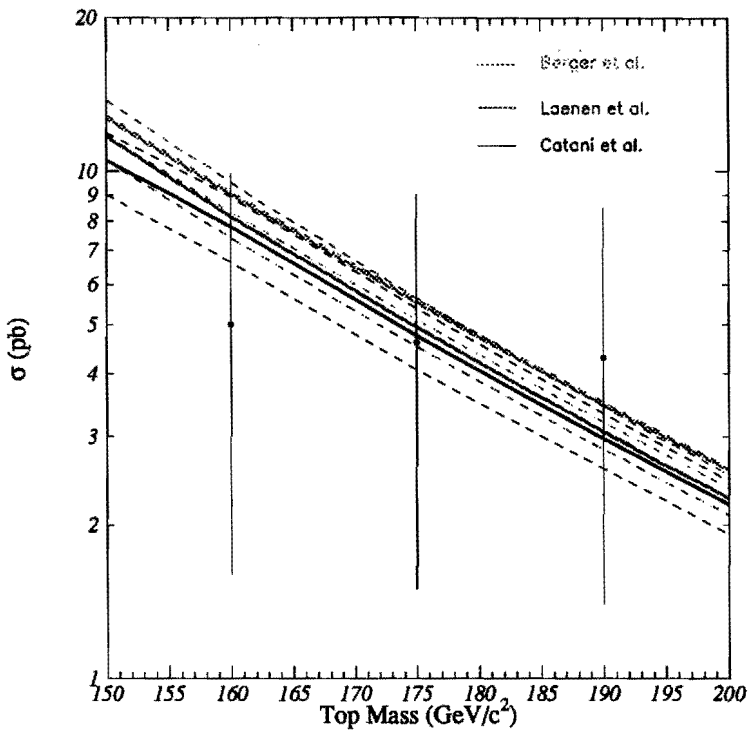

Figure 7.5: Comparison of measured $t\bar{t}$ production cross section vs theory. Note that the errors are highly correlated.

Chapter 8

Conclusion

The analysis presented here describes a method for searching for events containing a $t\bar{t}$ pair which decays to give two leptons, $\not\!\!E_T$, and two jets, one of which is b-tagged, in the final state. Looking in the CDF Run 1 data sample of 109.4 ± 7.2 pb^{-1}, 6 tags were found in four events where only $1.53 \pm .33$ are expected from background sources. Assuming the observed excess is due to $t\bar{t}$ production, the production cross section is measured to be $4.6^{+4.4}_{-3.1}$ pb.

This analysis is just one part of the overall strategy for searching for the top quark at CDF. Results from the lepton+jets and dilepton searches are already published and have firmly established the existence of the top quark. Other analyses pursuing the signature in the all hadronic mode and extending the dilepton channel to include hadronic τ decays are also underway. Figure 8.1 shows a comparison between the cross section derived from this analysis and the CDF[34] combined and D$\emptyset$[33] measurements. These measurements are all consistent with one another.

At this point data collection has stopped for both CDF and D$\emptyset$ and will resume in 1999. During this shutdown both the accelerator and the detectors will undergo improvements in the hardware. The accelerator will be upgraded with the addition of the Main Injector which will allow the Tevatron to operate at instantaneous luminosities over 10^{32} cm^{-1}s^{-1}. The center of mass energy will be increased from 1.8 TeV to 2.0 TeV which should increase the $t\bar{t}$ production cross section by about 37%[7].

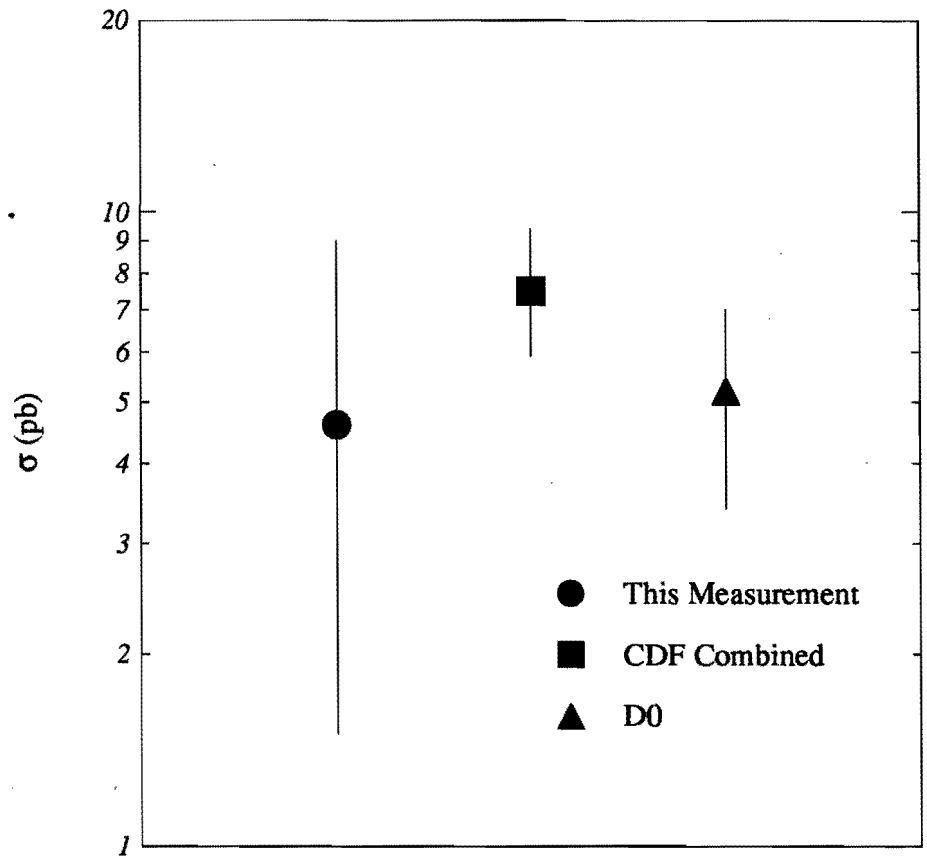

Figure 8.1: Comparison of the cross section from this analysis with those from other measurements. The CDF combined cross section does not include this analysis which is highly correlated with other CDF analyses. Also the DØ cross section is calculated at $m_{top} = 180$ GeV/c^2.

The CDF detector will undergo several upgrades which will result in improved performance for measuring the top quark system. The SVX detector will be replaced with a longer version which will increase the acceptance for b jets and will improve the tagging efficiency per jet to approximately 60% for vertex tagging algorithms alone[35]. The CTC and VTX will be replaced as will with detectors which are better able to withstand the density of tracks in the RunII environment and improve the tracking efficiency at higher $|\eta|$. The plug and forward calorimeters are being replaced with a scintillating tile calorimeter which should improve the energy resolution in those regions and increase the acceptance for electrons from W decays by about 36%. The gaps in ϕ in the CMP and CMX muon systems are being covered with chambers which should improve the acceptance for muons by as much as 25%.

With an expected 2 fb^{-1} of data, there should be about 700 events in the lepton+jets channel and over 100 dilepton candidates. The resolution on the mass of the top quark which is now about 9 GeV/c^2, should be about 4 GeV/c^2. Knowing the top quark mass to within a few GeV/c^2 combined with the mass of the W boson may make it possible to make predictions about the Higgs boson (see Figure 8.2). Additionally the large sample of $t\bar{t}$ events will make it possible to measure such quantities as $|v_{tb}|$, the branching ratios $Br(t \rightarrow Wb)/Br(t \rightarrow Wq)$ and $Br(t \rightarrow W_0 b)$ where W_0 is a longitudinally polarized W. Searches for single top production and for the rare decays such as $t \rightarrow Zc, \gamma c$ will also be possible.

One has to wonder if there is any significance to the top mass being 175 ± 9 GeV/c^2 when all of the other quarks are very light in comparison (the next heaviest quark is the b with a mass of about 5 GeV/c^2). Is this giving us a probe into the higgs sector, or perhaps an indication of new physics? These are questions whose answers may come from looking at more data.

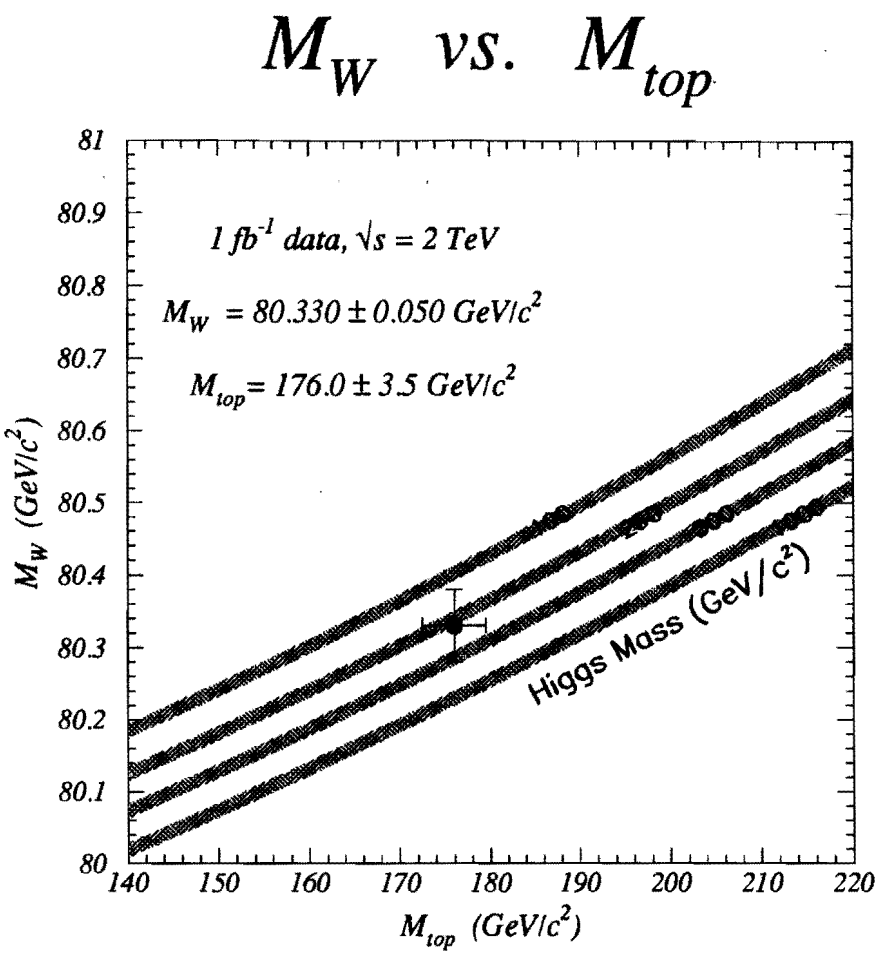

Figure 8.2: The point is located with at the current values of m_W and m_{top} with the uncertainties being those expected in Run II. The curves are from a calculation of the dependence of the mass of the W boson on the mass of the top quark and Higgs boson[36]. The uncertainty bands on the curves are obtained from the uncertainty on $\Delta\alpha(M_Z^2)$ ($\pm.0004$).

Appendix A

Design and Performance of the CMP

In the 1988-1989 run the CDF detector relied on the CMU detector to identify muons with $|\eta| < 0.6$. In this region there is an average of 5.4 pion interaction lengths between the vertex and the CMU chambers. Because of this, the trigger rate for CMU muons is dominated by hadronic "punch through", particles passing out of the calorimeter. During the 1988-1989 run the average instantaneous luminosities were about 10^{30} cm^{-2}s^{-1} and the trigger rates were manageable. In the 1992-1993 run the expected instantaneous luminosities were of order 5×10^{30} cm^{-2}s^{-1} which would have caused an excessive trigger rate. This hadronic punch through also made the identification of muons in jets difficult limiting the ability to study semi-leptonic decays of heavy quarks. By building the Central Muon Upgrade (CMP) and installing it outside the CMU behind an additional 60 cm of steel, the identification of muons both at the trigger level and offline can be improved by requiring the presence of a track in the CMP system. This makes it possible for CDF to better exploit the physics potential of the Tevatron collider.

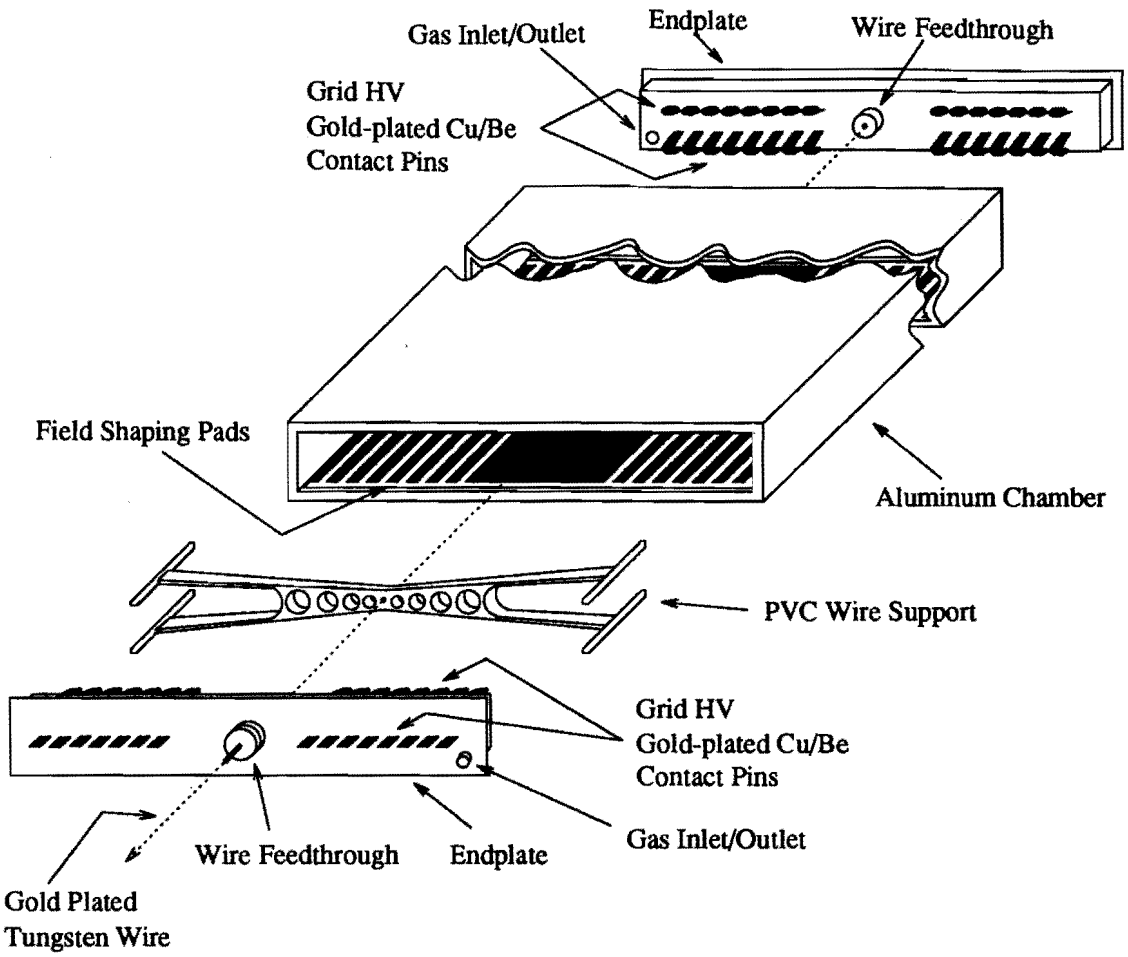

Figure A.1: CMP stack in an exploded view showing the endplate and wire support. The scale is not exactly correct in this figure.

A.1 Chamber Design

The CMP chambers were constructed using extruded rectangular tubes with an outside dimension of 2.54 cm × 15.24 cm and length of 6.4 m. The shape of the tubes was constrained by pre-existing obstructions around the detector and by the requirement of four layers of chambers. The chamber walls are 2.6 mm thick (Figure A.1). Four of these chambers were glued together with half the chambers offset by half their width to form a "stack". This staggering removes the left-right ambiguity of a reconstructed track. In addition, the time needed to make a trigger from the chambers is cut in half because hits in only two of the four chambers are required for a valid trigger.

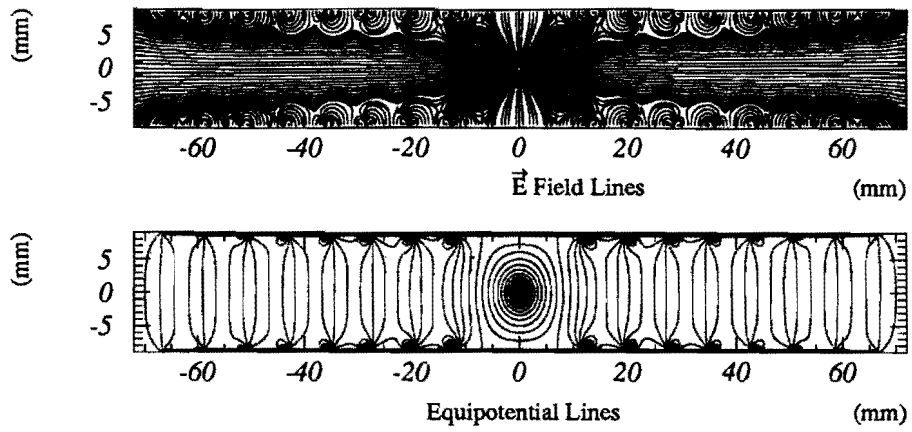

Figure A.2: Electric field lines (top) and equipotential lines (bottom) inside a CMP chamber.

The chambers have a 50 μm diameter gold-plated tungsten wire down the center which serves as the anode at +5400 V. There are cathode pads on the top and bottom of the chamber to shape the electric field (see Figure A.2). The center pad is set to +2800 V and has eight shaping pads on either side which have voltages decreasing in steps of 375 V towards the outer edge. These voltages are passed through the endplate by gold-plated Cu/Be contacts which were put into the endplate when it was molded. The division of the field shaping voltages is done by a 20 M$\Omega\times$8 divider resister which is mounted on a printed circuit board soldered to the contacts outside the endplate. The outer most shaping strips are electrically connected to the chamber wall using a conductive epoxy.

The chambers are run in proportional mode with an active volume filled with a 50%/50% Ar-C_2H_6 mixture bubbled through isopropyl alcohol at a temperature of -7° C. Setting the pads at 2800 V and the wire at 5400 V gives a nearly saturated drift velocity. At these voltages the chambers have a measured gain of about 6×10^4 yet remain several hundred volts from streamer mode. It is important to keep the chambers out of streamer mode because of the running conditions at the Tevatron where some of the chambers are hit by a spray of particles from the Fermilab Main Ring as it negotiates the overhead

bypass. This spray of particles produces large currents in streamer chambers which will cause the high voltage supplies to either trip or reduce the output voltage.

Because the chambers are 6.4 m long, they would be electrostatically unstable if a wire support was not put at the middle of the chamber. This support consists of a PVC body glued to 5 cm long PVC legs. The legs are made from rods sorted by diameter with variations up to 200 μm. The variation in diameter allows the wire support to be customized for each chamber to insure that the friction between the legs and the chamber walls will keep the support in place.

The signal from the wires comes out through a 2200 pF blocking capacitor onto a preamp board. The preamp board is mounted directly to the stack by four delrin standoffs which are screwed into the endplates. Each preamp board contains the preamplifiers for all four channels in a stack as well as a calibration pulse (Figure A.3). The preamp boards are based on a Radeka 10354 hybrid preamps. From the preamp the signal is sent to a 48 channel Amplifier-Shaper-Discriminator (ASD) card which sends a digital signal to a LeCroy 1879 Fastbus TDC.

A.2 Chamber Performance

After each stack was constructed, it was put through a series of tests to insure it was of good quality before being shipped to FermiLab for installation.

A.2.1 Gas Tightness

The chambers were checked for leaks by pressurizing them to about 0.13 atm above atmospheric pressure with N_2 gas. A reference vessel of approximately 10 ℓ was pressurized to the same initial pressure as the chamber provided a stable reference pressure. Once the system is isolated, the pressure of the chamber, $P(t)$, can be measured as a function of time:

$$P(t) = P_0 + (P_{atm} - P_0)e^{at/V}, \tag{A.1}$$

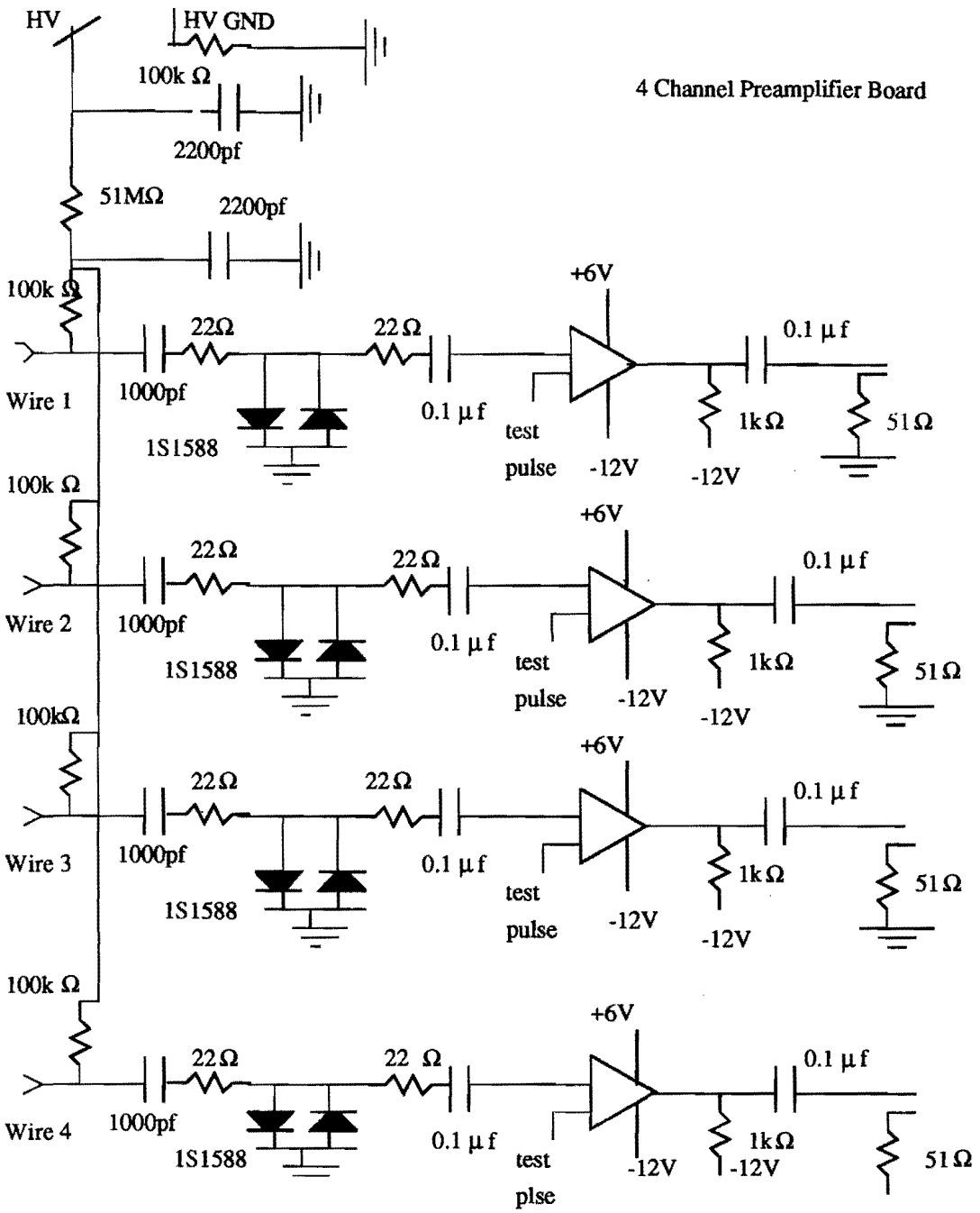

Figure A.3: CMP preamp schematic.

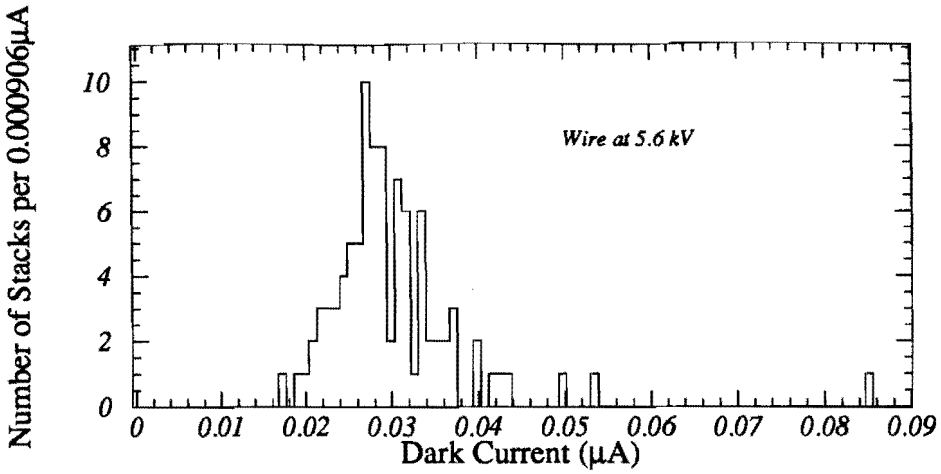

Figure A.4: CMP dark current distribution at anode voltage of 5600 V measured using good stacks.

where P_0 and P_{atm} are the initial chamber pressure and atmospheric pressure respectively; V is the volume of the chamber ($\sim$19 ℓ); and a is the leak rate for that chamber. The average measured leak rate for these chambers was about 5 cc/min. Larger leaks around the endplates were found in some chambers and were plugged with RTV.

A.2.2 Dark Currents

The dark current on the anode wires for a stack were measured as a function of anode voltage. For good chambers, the dark current was flat up to about 5800 V at which point the current draw begins to rise sharply. Chambers with defects on the wire have an increase in current draw at a lower voltage. The distribution of dark currents measured at 5600 V is shown in Figure A.4. The average dark current at an anode voltage of 5600 V is about 30 nA. Chambers were re-strung if the observed dark current was significantly larger than normal.

A.2.3 Wire Tension

The tension of the wire was measured by taking advantage of the resonance properties of a periodic signal propagating down the wire when it is placed in a magnetic field. A

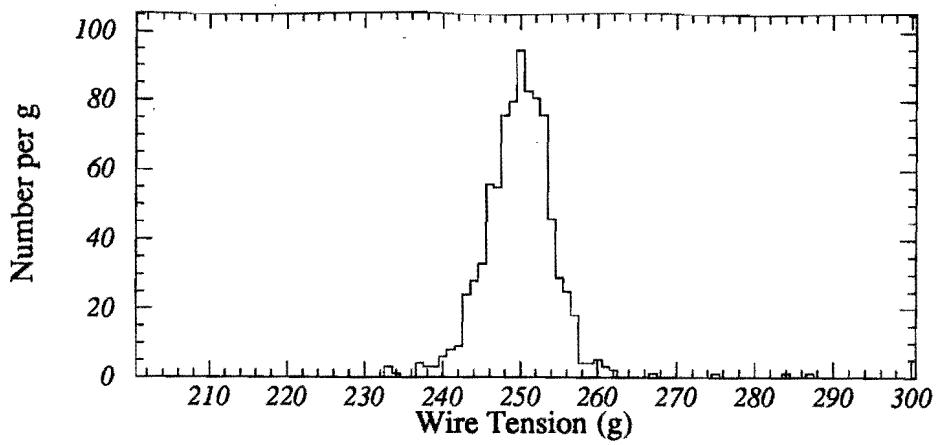

Figure A.5: Measured CMP wire tensions. The nominal tension is 250 g.

stack was place in the gap of a large horseshoe magnet. A sine-wave signal from a pulse generator was input at one end was compared to the signal measured at the other end of the chamber. By measuring the phase difference between these two signals, the resonant frequency of the wire can be found when the input and output signals are in phase. The wire tension can then be derived from the fundamental resonant frequency f_0 as

$$T = \frac{4L^2 f_0^2 \mu}{g}, \tag{A.2}$$

where L is half the chamber length (because of the wire support); μ is the mass per unit length of the wire (3.79×10^{-5} kg/m for tungsten wire with a radius of 25 μm); and g is the acceleration due to gravity. The tension measurements are shown in Figure A.5 have a mean of 249 g with a width of 4 g. The tails of the distribution are the result of friction in the pulley used to set the wire tension. A wire tension much lower than nominal was taken as a sign that the wire was not properly soldered at one end and those chambers were re-strung.

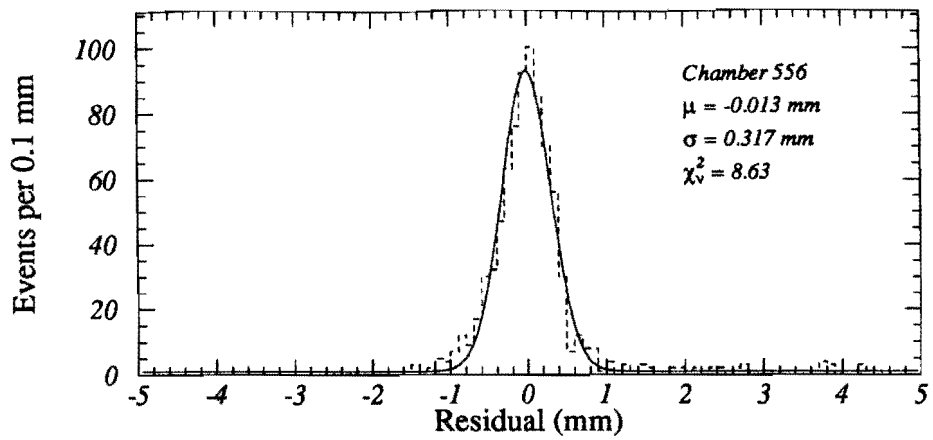

Figure A.6: CMP hit residual distribution for a single chamber.

A.2.4 Cosmic Ray Testing

General chamber performance was checked by reconstructing tracks from cosmic rays passing through the stacks. From these reconstructed tracks, the hit residuals, hit efficiencies and drift velocities can be measured.

Tracks were reconstructed using three of the four chambers in a stack and extrapolating or interpolating to the fourth chamber. The difference between the measured hit position and the predicted position is called the residual. A typical residual distribution is shown in Figure A.6 for good chamber. A Gaussian fit to the data give an average chamber resolution of about 300 μm. The distribution of resolutions for the chambers is shown in Figure A.7. The same technique of fitting three chambers and looking at the fourth chamber is used to measure the efficiency as a function of hit position across the width of the chamber. In this case the fourth chamber is checked for a valid hit. Figure A.8 shows the typical efficiency of a typical chamber as a function of the hit position within the chamber. The efficiency is measured to be ~99% out to the edge of the drift region. Inefficiencies also occur around the wire support and the end of the chambers, these distributions are shown in Figures A.9 and A.10.

The drift velocity for a given chamber can be calculated by using the drift times of the other three chambers in a stack and the stack geometry under the assumption that the

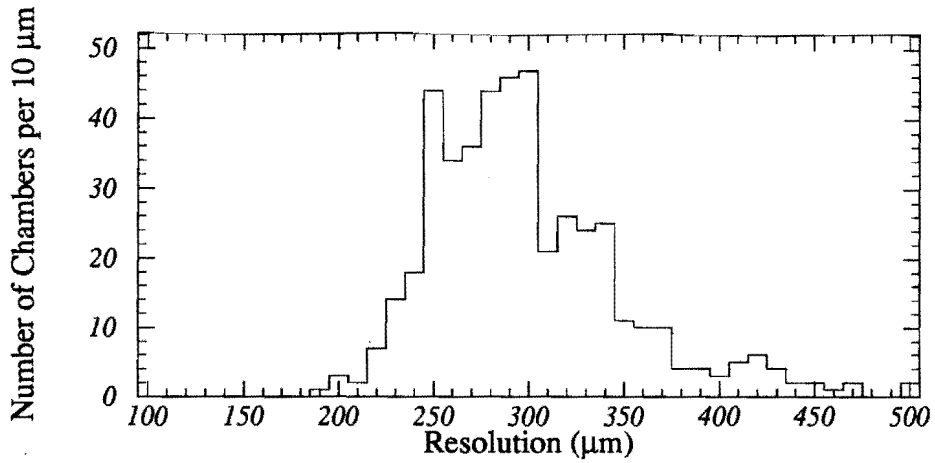

Figure A.7: CMP hit resolutions measured using good stacks.

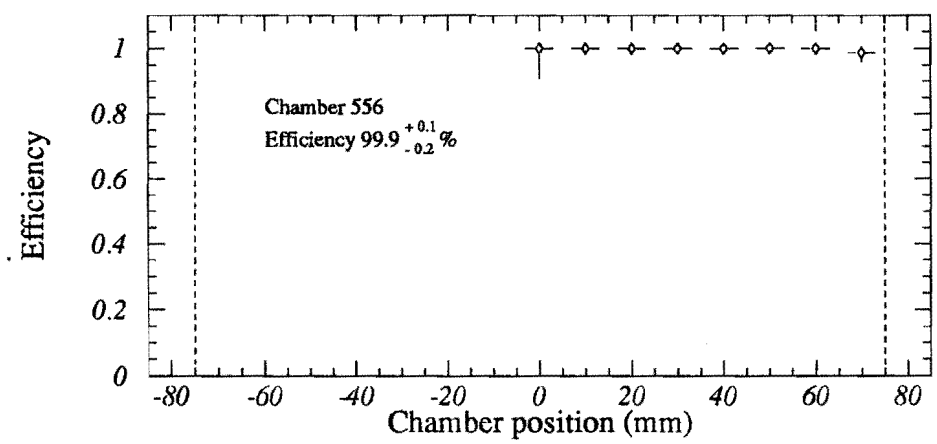

Figure A.8: CMP single hit efficiency measured for a typical chamber. The chamber position is the distance between the wire and extrapolated/interpolated track.

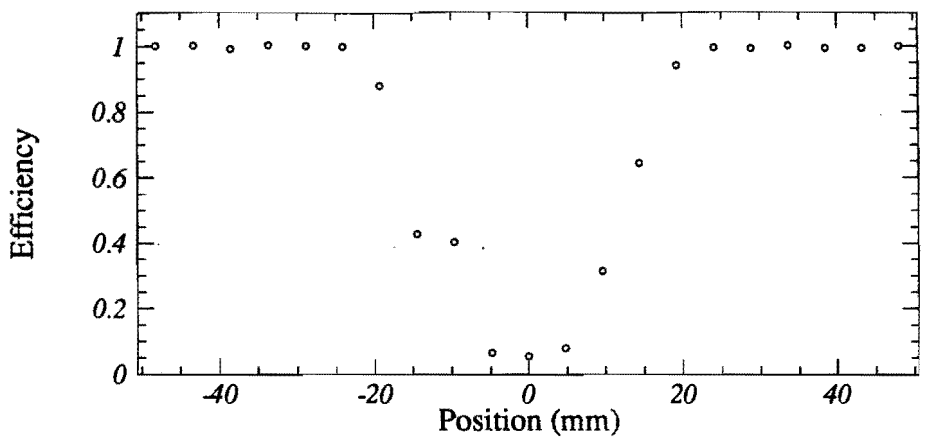

Figure A.9: CMP hit efficiency measured near the wire support. The position is the distance along the wire from the wire support to the extrapolated/interpolated track.

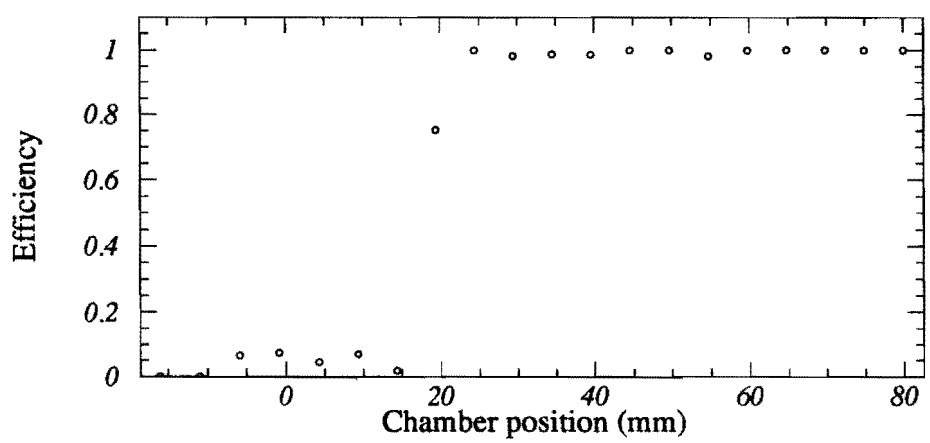

Figure A.10: CMP hit efficiency measured near the end of the chamber. The chamber position is the distance from the end of the chamber to the extrapolated/interpolated track position.

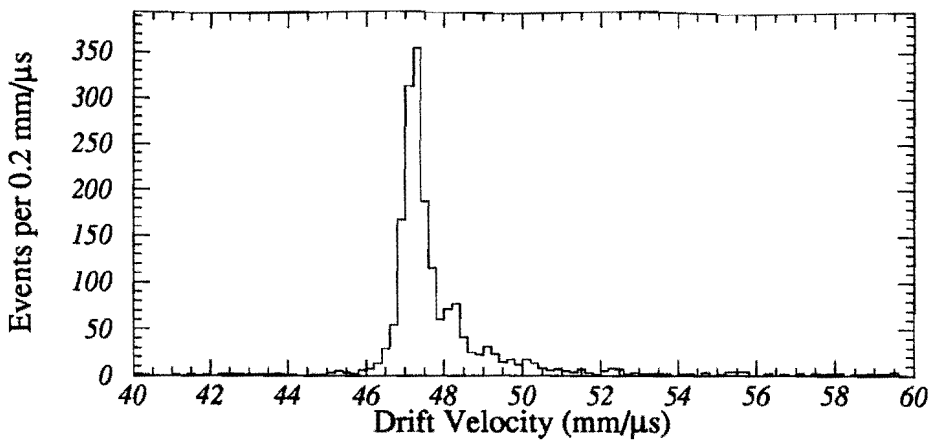

Figure A.11: CMP drift velocity distribution measured in a typical chamber. The mean is expected to fall between 45 and 50 mm/μs.

drift velocity is constant for all chambers in a give stack. In that case the drift velocity is given by:

$$v_{drift} = \frac{a}{bt_1 + ct_2 + ct_3}, \tag{A.3}$$

where a, b, c, and d depend on the relative positions of the wires and t_1, t_2, and t_3 are the three drift times. The drift velocity distribution for a typical stack is shown in Figure A.11. The tails are caused by the inclusion of tracks which were not fit well.

Appendix B

The CDF Muon Trigger Scheme

The CDF central muon trigger has three levels, two implemented in hardware and a third in software. This appendix describes the procedures used by the Level 1 and Level 2 central muon trigger to arrive at a decision.

The central muon trigger in CDF is organized into 5° segments, called "trigger towers", three for each wedge in the central detector. A decision is made based only upon information pertinent to a single 5° slice of the detector. Note that the CMP portion of the trigger was designed to be used only as a confirmation of the CMU trigger, it does not provide an independent trigger. The general organization of the trigger is shown in Figure B.1 and is described below.

B.1 Level 1

The Level 1 muon trigger consists of several different modules designed to extract information from a specific detector and generate a "primitive" and a set of modules to combine those primitives. Many of the modules are capable of generating primitives for two programmable thresholds making it possible to trigger separately on high p_T and low p_T muons. The trigger rate for single low p_T muons is very high and would cause to much dead time. However, accepting pairs of low p_T muons does not incur unreasonable dead time and yet still allows CDF to trigger on muon pairs from sources such as $J/\psi \rightarrow \mu\mu$,

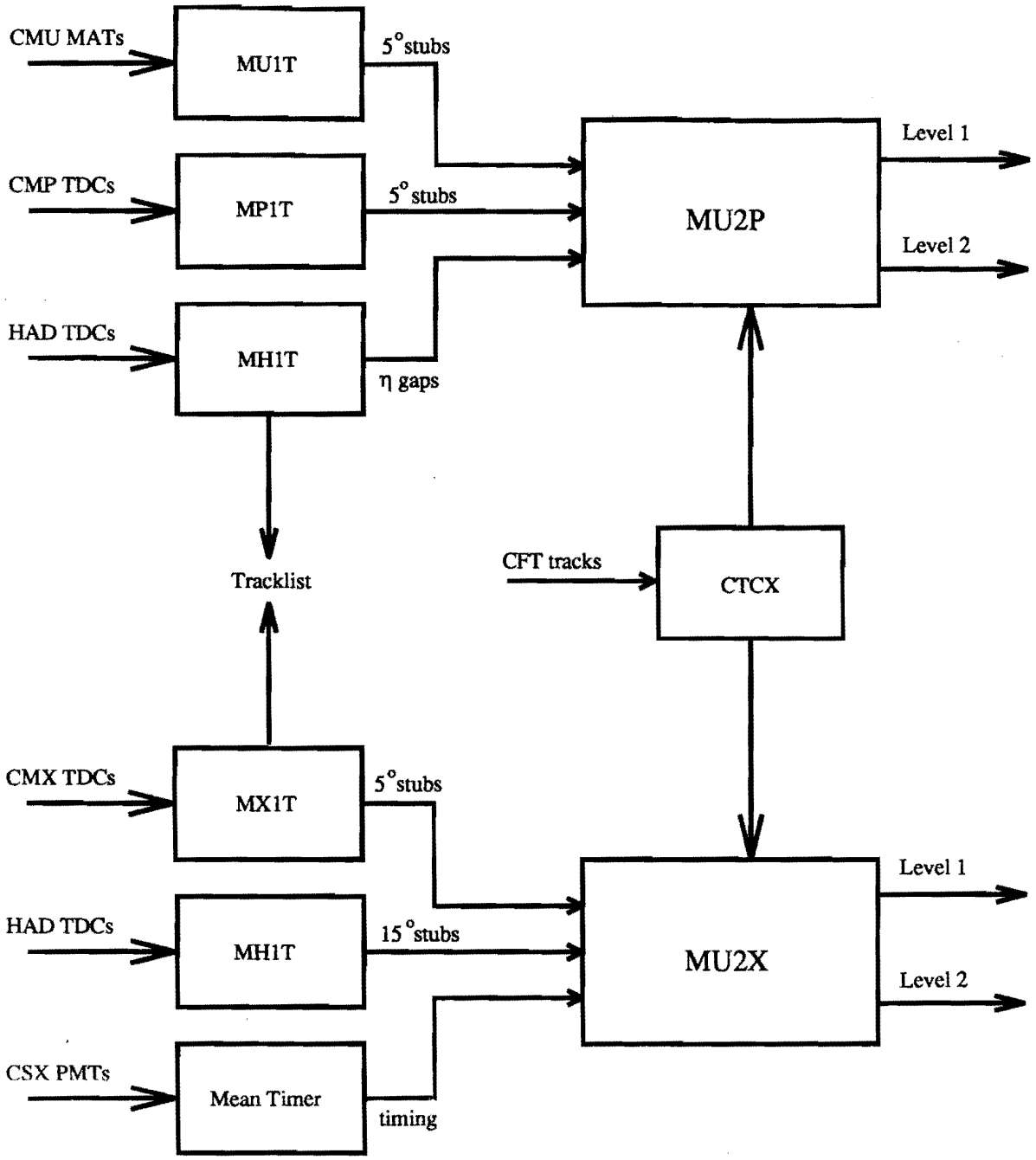

Figure B.1: Organization of the CDF central muon trigger system.

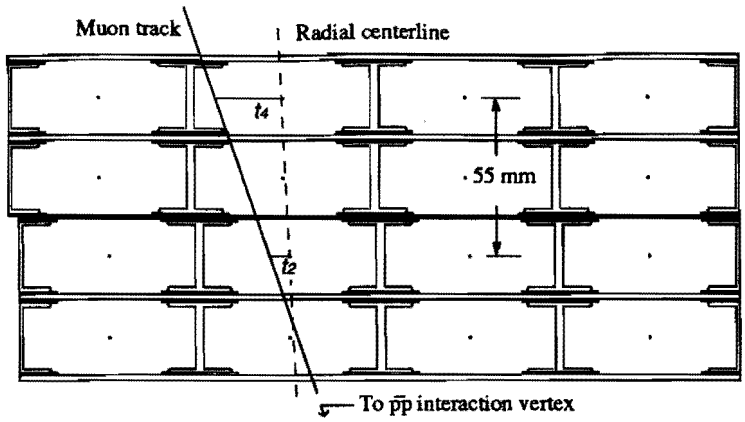

Figure B.2: Cross section of a CMU module showing the different drift times for chambers in alternate layers for a non-radial track.

a powerful tool for doing b physics. As a result, the CDF muon trigger can select events with at least one high p_T muon or at least two low p_T muons.

MU1T The Level 1 trigger modules for the CMU system were implemented as cards in the front end electronics crates for the 1988-1989 run[37]. The idea is to look at the time difference Δt between the arrival of hits from a muon passing through the chamber on the wires for alternate layers (see Figure B.2). Because of the magnetic field in the solenoid, a muon will follow a helical trajectory until it passes into the solenoid at which point its trajectory becomes a straight line. As a result, the muon's track in CMU will make an angle relative to the radial line which is inversely proportional to the p_T of the track. Measuring Δt determines the relative angle between the track and the radial line and also the p_T. If the time difference is less than a programmed threshold which is determined by the desired minimum p_T, the module generates a CMU primitive. Each wedge in CMU is instrumented with two of these modules making it possible to have two independent thresholds. The entire system has one output for each of the trigger towers on each half of the detector, at each threshold for a total of 288 bits.

MP1T The CMP trigger modules, MP1T, are custom FASTBUS modules which are designed to look for hit combinations in the CMP. Because of the rectangular geometry of the CMP a simple relationship between the track angle and the p_T does not exist. As a result the CMP is used only as confirmation of a CMU trigger. This philosophy means that the trigger only needs to look for the presence of hits in CMP chambers which are consistent with having come from a muon passing through a given CMU trigger tower. To do this Monte Carlo simulations of muons were used to determine which pairs of chambers could be hit for each CMU tower by muons with $p_T > 3.3$ GeV/c. The output is a single bit for each of the 72 CMU trigger towers in ϕ.

MX1T The MX1T cards are also custom FASTBUS modules which are similar to the MU1T modules in function. The MX1T modules measure the time difference between the arrival of hits from radially aligned pairs of wires to make a p_T measurement. There are two different programmable p_T thresholds for each module. Each module corresponds to 30° in ϕ for one side of the detector and has 12 outputs, giving a total of 288 bits.

CSX The CSX meantimers are designed to pass along information from the CSX for use in the CMX trigger. These modules were designed to make an OR of pairs of scintillators, one on the inner surface of CMX and the other on the outer, and to tighten the timing of the hits relative to the beam crossing. By gating the meantimers around the beam crossing time, much of background not associated with muons from the $\overline{p}p$ interaction can be removed. These backgrounds are a problem due to the location of the CMX and CSX systems.

MH1T The MH1T modules are custom designed to add information from the TDC's on the CHA and WHA calorimeters into the muon trigger. The idea is that a muon passing through a calorimeter module should have deposited some energy. The TDC's provide a fast signal which simply says that calorimeter had some energy deposited in it. The OR of the CHA TDC bits is available for use in the CMU/CMP trigger while the OR of four towers in the CHA and four in the WHA which are covered by the CMX chambers

are available for use in the CMX trigger. This can help remove backgrounds coming from fake CMU or CMX primitives which have no energy deposited in the appropriate calorimeter towers.

MU2X The MU2X cards are FASTBUS boards built here which combines the information provided by all the above systems to form the actual Level 1 trigger. These cards were designed to have a great deal of flexibility built into them so that triggers from the new systems could be integrated smoothly. There are 24 of these cards for the CMU/CMP trigger and 24 for the CMX trigger, each card make the decision for all three towers in a 15° wedge. In a given 5° trigger tower, the CMU, CMP, and hadron TDC primitives can be combined by requiring a programmable combination to form a Level 1 CMU trigger. The CMX, CSX and hadron TDC primitives can be combined in a similar fashion to provide a Level 1 trigger for the CMX system.

B.2 Level 2

The Level 2 muon trigger decision is based upon whether or not a track from the CFT extrapolates to the trigger tower where a Level 1 trigger occurred. The CFT provides information about tracks in the CTC to a FASTBUS module called the CTCX in the form of a wire number in the outer-most superlayer of the CTC and a crude p_T measurement. The CTCX uses these two pieces of information as inputs into a lookup table to determine which 5° trigger towers the track might have hit. The CTCX passes along three bits to each MU2X corresponding to a track matching to each of the trigger towers. The CTCX also passes along the track information from the CFT to the rest of the trigger system to be used in other trigger decisions such as the electron trigger.

Appendix C

Neutrino momentum

Neutrinos only interact with the other quarks and leptons through the weak interaction resulting in a very small cross section for a neutrino to be observed. As a result, the momentum of a neutrino cannot be directly measured in the CDF detector. However, the total momentum transverse to the beamline of non-interacting particles can be inferred by requiring conservation of momentum transverse to the beamline for all observed particles. A quantity called "missing transverse energy", $\not{E}_T$, is defined which represents the total transverse momentum of all the unobserved particles and is calculated as follows.

The transverse energy in the calorimeter tower i is defined to be E_T^i, given by $E^i sin\theta_i$, where E^i is the measured energy deposited in tower i and θ_i is the polar angle from the event vertex to the calorimeter tower. The total vector transverse energy, $\vec{E}_T$, measured in the detector can then be defined by:

$$\vec{E}_T = \sum_i E_T^i \hat{\mathbf{n}}_i, \tag{C.1}$$

where $\hat{\mathbf{n}}_i$ is a unit vector transverse to the beam line pointing to the ith calorimeter tower and the sum is over all towers with $|\eta| < 3.6$. The η range is restricted because the FHA is partially obscured at high $|\eta|$ by the quadrapole magnets which focus the beams. In addition only towers with an energy above a threshold of 100 MeV in the CEM, CHA, and WHA, 300 MeV in the PEM, 500 MeV in the PHA and FEM, and 800 MeV in the

FHA are included in the sum. The $\vec{E}_T$ must then be corrected for muons which deposit little of their energy in the calorimeter. This is done if the muon had $p_T > 10$ GeV/c by subtracting out the $\vec{E}_T$ in the towers the muon passed through, and adding back in the $\vec{p}_T$ of the muon. In addition, since jets are corrected back to the parton energies (see Section 3.1) these corrections are taken into account in the calculation of the $\vec{E}_T$.

If the momenta of all the particles in the event have now been accounted for, the $|\vec{E}_T|$ would be 0 GeV. However, if there is a neutrino in the event, the $|\vec{E}_T|$ will in be non-zero. The total missing transverse energy, $\not{E}_T$, is defined to be $-|\vec{E}_T|$. This technique only works in the plane transverse to the beam line because particles with $|\eta| > 4.6$ can escape from the detector without having been seen. Also particles with $|\eta| > 3.6$, which include many of the fragments of the proton and anti-proton which collided, would not contribute much to the $\not{E}_T$ since most of their momentum is along the beamline.

Appendix D

Tagging Algorithms

This Appendix servers to describe some of the details of the b-tagging algorithms used in this analysis. These algorithms were developed by other people and are presented here for completeness sake.

D.1 SECVTX

The SECVTX algorithm is used by CDF to look for displaced vertices from the decay of b hadrons. The idea is to look at tracks which are inconsistent with having originated at the primary vertex and to see if any appear to have come from a common point. The decay products of a top quark decay include b quarks which will have a typical momentum of $\sim$60 GeV/c for a top mass of 175 GeV/c^2 (Figure D.1). CDF has measured the average lifetime of b hadrons to be $1.46 \pm .08$ ps[38] by looking at inclusive J/ψ production. This implies that b hadrons from the decay of a top quark will have a flight distance of a few mm between the primary vertex and the b decay vertex.

The SECVTX algorithm first makes an association between a jet and the tracks in the event by finding all tracks within a cone of 0.4 in $\eta - \phi$ of the jet axis. Tracks which do not have a significant impact parameter transverse to the beamline are removed. For each jet, the SECVTX algorithm makes two passes through the remaining tracks looking for a common vertex associated with that jet. The first pass attempts to make a vertex with 3

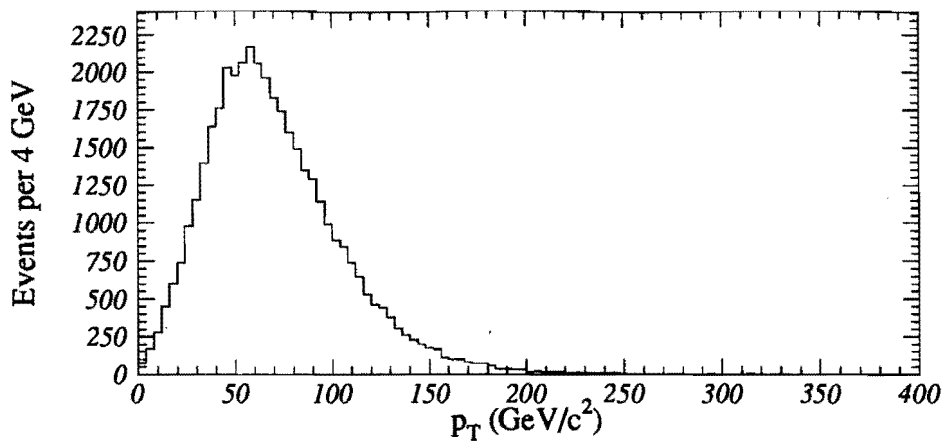

Figure D.1: Transeverse momentum spectrum of b quarks from top decays in $t\bar{t}$ Monte Carlo generated with PYTHIA at $m_{top} = 175$ GeV/c^2.

tracks, while the second pass only requires two tracks but tightens the track quality and p_T requirements on those tracks. Tracks which are consistent with being from a K_s or a Λ decay are excluded from both passes of the algorithm. (These are identified as two oppositely charged tracks with an invariant mass within 10(6) MeV/c^2 of the nominal $K_s(\Lambda)$ mass, the displacement from the primary vertex along the momentum vector of the reconstructed particle is greater than 10σ, and the displacement perpendicular to the momentum vector is less than 3σ.)

For the first pass, the tracks in a jet with $p_T > 0.5$ GeV/c and unsigned impact parameter significance ($|\frac{d}{\sigma_d}|$) greater than 2.5 are ranked according to their p_T, impact parameter significance, and number of good SVX hits used in the track's fit. The two highest ranked tracks are constrained to a single point. This point is used as a seed to look for an additional displaced track which has an impact parameter significance < 3 with respect to the seed vertex. If a vertex is found, a tag is declared and the algorithm ends, otherwise the next highest ranked pair of tracks is used as a seed. If no vertex is found when all seed combinations have been tried, then the algorithm goes to the second pass.

The second pass of the SECVTX algorithm searches for displaced vertices made of tracks which pass more stringent track quality cuts. All tracks in the jet with $p_T >$

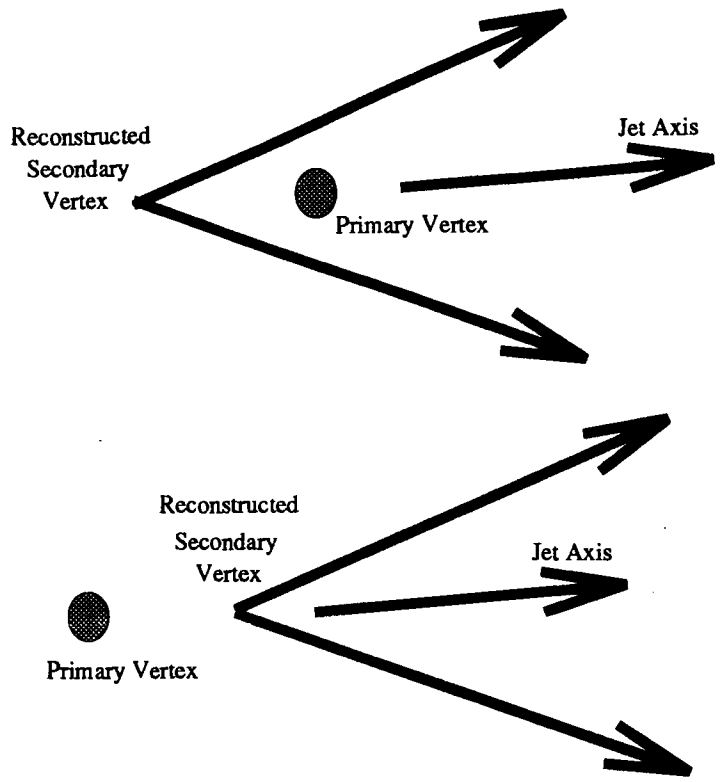

Figure D.2: Diagram showing both negative (top) and positive (bottom) L_{xy} vertices.

1.5 GeV/c and unsigned impact parameter significance > 4.0 with respect to the primary vertex are considered. The fitting routine forms a vertex using all tracks passing these cuts. All tracks contributing more than 50 to the χ^2 of the vertex fit are removed and a new fit is performed. This procedure is iterated until the tracks in the vertex remain the same. If no vertex is found during this pass, the algorithm ends.

After a vertex candidate is found in either pass, the 2D decay length, L_{xy}, and its error, $\sigma_{L_{xy}}$, are computed. The sign of L_{xy} is defined as the sign of the dot product between a vector from the primary vertex to the secondary decay vertex and a vector in the direction of the jet axis (See Fig. D.2). The L_{xy} distribution for tagged jets in a sample of inclusive jets with a 50 GeV trigger is shown in Figure D.3.

To be considered a tag, a secondary vertex must pass the following cuts: $|L_{xy}| <$ 5.0 cm and $\frac{|L_{xy}|}{\sigma_{L_{xy}}} > 3.0$. Tags with $L_{xy} < 0$ are expected to be dominated by "mistags" which are due to mismeasured tracks and detector resolution effects. The tags with

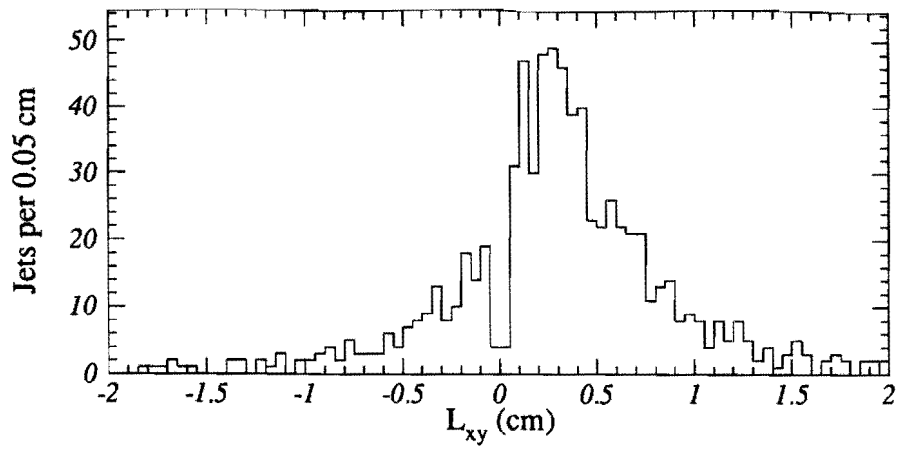

Figure D.3: L_{xy} distribution for tagged jets selected with a 50 GeV trigger threshold.

$L_{xy} > 0$ will contain some mistags but are dominated by real $b\bar{b}$ and $c\bar{c}$. The L_{xy} distribution for mistags is expected to be symmetric about $L_{xy} = 0$ since reconstruction errors and resolution effects have no preference for a sign on L_{xy}. This symmetry allows the tags with $L_{xy} < 0$ to be used to estimate the number of mistags which are contained in the $L_{xy} > 0$ sample.

A difference is seen in the tagging efficiency measured in inclusive electron data and the Monte Carlo for the SECVTX algorithm[2] and a scale factor is computed to apply to Monte Carlo samples to correct for this effect. Monte Carlo tracks are imbedded in jets with $E_T > 15$ GeV, $|\eta| < 1.0$ and $z_{vert} < 30$ cm in events which passed the 50 GeV jet trigger and the track reconstruction code is re-run on the event. The efficiency for finding a track in a jet is the fraction of the events where the added Monte Carlo track is found. Following the same procedure with a sample of inclusive jets events generated using PYTHIA results in a tracking efficiency for the Monte Carlo. The ratio of data/Monte Carlo tracking efficiencies is a measure of how many fewer tracks are found in the data than in the Monte Carlo. Applying this ratio of tracking efficiencies to tracks in Monte Carlo jets effectively reduces the track finding efficiency to be what is observed in the data.

In the $t\bar{t}$ Monte Carlo, the effect of this tracking efficiency on the SECVTX tagging efficiency can be seen by applying this tracking efficiency effect to the tracks in b jets. This is done by discarding the individual tracks in the jets according to the expected inefficiency before applying the SECVTX algorithm. The efficiency for tagging a lepton+jets event before this procedure was $47.5 \pm .4\%$, while afterward, the efficiency is $41.2 \pm .4\%$ (the errors quoted are statistical only). The ratio of the data tagging efficiency to Monte Carlo tagging efficiency is $0.87 \pm .07$ after systematic effects such as the b lifetime, the SVX resolution, b hadronization and decay effects, and track multiplicities in the jets are included.

D.2 SLT

The second method used to look for b hadrons is to look for the presence of low p_T muons and electrons near jets. This method takes advantage of the fact that approximately 20% of the b decays have an electron or muon in them. The idea in this algorithm is to look at the tracks in jets to see if they are consistent with being from either a muon or an electron. The algorithm used in this analysis is a slight variation from what has been used by CDF in the past[2, 29] in that here the soft lepton is required to be within a cone of 0.4 in $\eta - \phi$ of a jet axis.

The soft electron finding algorithm starts by locating the tracks within 0.4 of a jet axis with a $p_T > 2.0$ GeV/c, an impact parameter relative to the beamline of less than 3 mm, two axial CTC superlayers with hits, and two CTC stereo superlayers with hits. Tracks from photon conversions to e^+e^- pairs are removed if a second, oppositely charged track which makes a small invariant mass and comes within 3 mm in the r-ϕ plane. The shower shape and charge in the CPR are required to be consistent with having come from an electron. For tracks with $P \leq 15$ GeV/c, a cut is made on the dE/dX measured in the CTC. The algorithm also requires $0.7 \leq E/P \leq 1.5$ and $E_{Had}/E_{EM} \leq 0.1$. Also the track is required to extrapolate to a good fiducial region of the CES and CPR detectors.

The soft muon identification algorithm begins by looking for CTC tracks which extrapolate to the muon detectors. If the track extrapolates to inside one of the detectors, hits in that detector are required. Tracks with $p_T < 3$ GeV/c and $|\eta| < 0.6$ are only required to point to hits in the CMU detector because muons with $p_T < 2.8$ GeV/c are expected to stop before getting to the CMP. The CTC tracks are required to have hits in at least two axial and two stereo CTC superlayers as well as have an impact parameter relative to the beamline less than 3 mm. Cuts are then made on the number of TDC and ADC hits seen in the detectors and how well the muon stub matches the extrapolated track in both position and in the phi of the segments. There is also a minimum ionizing cut placed on the muon candidates, $E_{Had} - \sum P(0.2) < 6.0$ where $\sum P(0.2)$ is the scalar sum of momenta of all the tracks in a cone of radius 0.2 around the muon candidate.

The SLT purity can be measured in a sample of inclusive jets by looking for jets which are tagged by both the SLT and SECVTX algorithms. The purity can be written as

$$P_{SLT} = \frac{N_{SLT-SVX}}{N_{SLT}\,\epsilon_{SVX}} \tag{D.1}$$

where $N_{SLT-SVX}$ is the number of jets tagged with both algorithms, N_{SLT} is the number of jets tagged by the SLT algorithm which are taggable by the SECVTX algorithm, and ϵ_{SVX} is the efficiency for tagging a taggable b jet by the SECVTX algorithm. Looking in a sample of inclusive jets with a 50 GeV trigger jet, the purity is found to be $25 \pm 12\%$.

Bibliography

[1] S. Glashow. Partial-Symmetries of Weak Interactions. Nucl. Phys. **22**, 579 (1961). S. Weinberg. A Model of Leptons. Phys. Rev. Lett. **19**, 1264 (1967). A. Salam. *Elementary Particle Theory,* page 367. Almqvist & Wiksell, Sweden, 1968.

[2] F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[3] S. Abachi *et al.*, Phys. Rev. Lett. **74**, ? (1995).

[4] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).

[5] G. Altarelli, M. Diemoz, G. Martinelli, and P. Nason, Nucl. Phys. **B308**, 724 (1988). R. K. Ellis, Phys. Lett. **B259**, 492 (1991).

[6] E. Laenen, J. Smith and W. L. van Neerven, Phys. Lett. **b321**, 254 (1994).

[7] E. Berger, H. Contopanagos, Phys. Lett. **b361**, 115 (1995).

[8] S. Catani, M. Mangano, P. Nason, and L.Trentadue, CERN-TH/96-21.

[9] J. M. Benlloch, N. Wainer, and W. T. Giele, Phys. Rev. D **48**, 5226 (1993).

[10] F. Abe *et al.*, Nucl. Instr. and Meth. **A271**, 387 (1988).

[11] D. Amidei *et al.*, Nucl. Instr. and Meth. **A350**, 73 (1994).

[12] F. Bedeschi *et al.*, Nucl. Instr. and Meth. **A268**, 50 (1988).

[13] L. Balka *et al.*, Nucl. Instr. and Meth. **A267**, 272 (1988).

[14] S. Bertolucci *et al.*, Nucl. Instr. and Meth. **A267**, 301 (1988).

[15] Y. Fukui *et al.*, Nucl. Instr. and Meth. **A267**, 280 (1988); G. Brandenburg *et al.*, Nucl. Instr. and Meth. **A267**, 257 (1988); S. Cihangir *et al.*, Nucl. Instr. and Meth. **A267**, 249 (1988).

[16] G. Ascoli *et al.*, Nucl. Inst. and Meth. **A268**, 33 (1988).

[17] J. Lewis *et al.*, To be published in Nucl. Inst. and Meth. A.

[18] D. Amidei *et al.*, Nucl. Inst. and Meth. **A269**, 51 (1988); E. Barsotti *et al.*, Nucl. Inst. and Meth. **A269**, 82 (1988).

[19] G. W. Foster *et al.*, Nucl. Inst. and Meth. **A269**, 93 (1988).

[20] F. Abe *et al.*, Phys. Rev. D **50**, 5518 (1994). F. Abe *et al.*, Phys. Rev. D **50**, 5535 (1994). F. Abe *et al.*, Phys. Rev. D **50**, 5550 (1994).

[21] F. Abe *et al.*, Phys. Rev. D **45**, 1448 (1992), F. Abe *et al.*, Phys. Rev. D **47**, 4857 (1993).

[22] L. Montanet *et al.*, Phys. Rev. D **50**, 1173 (1994), and 1995 off-year partial update for the 1996 edition of the *Particle Data Book* available at http://pdg.lbl.gov/.

[23] F. Paige and S. D. Protopopescu, ISAJET Version 7.06 BNL Report No. 38034, 1986 (unpublished).

[24] G. Marchesini and B. R. Webber, HERWIG Version 5.6 Nucl. Phys. **B310**, 461 (1988).
G. Marchesini, *et al.*, Comput. Phys. Comm. **67**, 465 (1992).

[25] T. Sjöstrand, PYTHIA Version 5.6 Computer Physics Commun. **39**, 347 (1986).
T. Sjöstrand and M.Bengtsson, Computer Physics Commun. **43**, 367 (1987).
H.-U. Bengtsson and T. Sjöstrand, Computer Physics Commun. **46**, 43 (1987).

[26] P. Avery, K. Read, G. Trahern, Cornell Internal Note CSN-212, 1985 (unpublished).

[27] S. Jadach *et al.*, TAUOLA Version 2.5 (June 1994); CERN-TH-5856 (1990), CERN-TH-6195 (1991), CERN-TH-6793 (1992).

[28] F. Abe *et al.*, Phys. Rev. Lett. **69**, 29 (1992).

[29] F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994).

[30] J. Ohnemus *et al.*, Phys. Rev. D **44**, 1403 (1991). S. Frixione, Nucl. Phys. **B410**, 280 (1993).

[31] F. James, *Minuit - Function Minimization and Error Analysis*, 1994. CERN Program Library Documentation.

[32] L Galtieri for the CDF Collaboration, XXXI Rencontres de Moriond, March 1996, to bc published.

[33] R. Partridge, DØ Collaboration, XXXI Rencontres de Moriond, March 1996, to be published.

[34] P. Azzi for the CDF Collaboration, XXXI Rencontres de Moriond, March 1996, to be published.

[35] CDF Collaboration, "Physics with CDF in Run II", CDF Internal Note 3172, unpublished.

[36] The theory curves are calculated using a FORTRAN program from F. Halzen and B. A. Kniehl (private communication), described in nucl. Phys. **B353**, 567 (1991). The program has been modified to include higher-order m_{top} dependant QCD corrections to $\Delta\rho$ (See e.g. B. H. Smith and M. B. Voloshin, Phys. Rev. D **51**, 5251 (1995) and A. Sirlin, HEPPH-9403282). The values used to calculate the curves are $\alpha_{em} = 1/(137.0359895 \pm 0.0000061)$, $G_F = 1.16639 \pm 0.00002 \times 10^{-5}$ GeV^{-2}, $M_Z = 91.1888 \pm 0.0044$ GeV/c^2, $\alpha_s(M_Z^2) = 0.118 \pm 0.005$, and $\Delta\alpha_{had} = 0.02800 \pm 0.00040$.

[37] G. Ascoli, *et al.*, Nucl. Inst. and Meth. **A269**, 63 (1988).

[38] F. Abe *et al.*, Phys. Rev. Lett. **71**, 3421 (1993).

Vita

Andrew James Martin was born in Quincy, Illinois in mid-April 1966. Andrew graduated from Quincy Senior High in 1984 having learned a little about both music and physics. Andrew then proceeded to attend Carnegie-Mellon University where he took all the physics courses he could. After earning a Bachelor of Science from the Physics Department, he decided to come to the University of Illinois to do particle physics. He decided that helping Tony Liss build a thousand muon chambers sounded like fun, so he joined CDF. After a long effort he finally found a top quark signature and wrote this thesis. When not trying to find the top quark, he could often be found attempting play the bass.