Search for the Top quark at $\sqrt{s} = 1.8$ TeV

by

Baber Khan Farhat

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

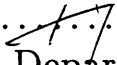Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1994

Author . . . .
Department of Physics
January 31, 1994

Certified by. . . . .
Jerome I. Friedman
Institute Professor
Thesis Supervisor

Accepted by . . . .
George F. Koster
Chairman, Departmental Committee on Graduate Students

Search for the Top quark at $\sqrt{s} = 1.8$ TeV

by

Baber Khan Farhat

Abstract

We present a search for the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV using the CDF detector at the Fermilab Tevatron Collider. We use 21.4 pb^{-1} of data collected during the 1992-93 Collider Run. In the minimal Standard Model, the top quark decays to a W boson and a b quark. To separate top events which decay into a lepton and jets from the W + multijet background, we identify b quarks via their semileptonic decays. In the W + 3 or more jet sample, we observe 7 events with an estimated background of 3.1 ± 0.3.

Thesis Supervisor: Jerome I. Friedman
Title: Institute Professor

Acknowledgments

First and foremost I would like to thank my advisor, Jerome Friedman, for granting me a free rein throughout my graduate career. He is without a doubt the nicest human being I have ever met. His constant support, especially in times of conflict, is directly reponsible for my being able to complete this work. I am also deeply grateful to Claudio Campagnari, Paris Sphicas and Avi Yagil for giving me the opportunity to work with the best minds at CDF on the hottest topic in HEP. Their insights into this analysis coupled with their patience with me, I will cherish forever.

I would also like to thank Gong Ping Yeh (GP) for his interest in my activities and constant support throughout my stay at Fermilab. My thanks to the entire MIT group at Fermilab, especially my roommates Wasiq and Troy. Special thanks to Luc for his companionship both on and off the Tennis court. From the testbeam days until the present day, Vaia Papadimitriou has always had my interests in mind. For that I am very grateful. I also appreciate the constant help I got from Colin, Jaco, Bob, Steve, Lou and the rest of the calorimetry group back in the early days.

I would like to thank the entire support staff at Fermilab without whom none of this would have been possible. Special thanks to Carol Picciolo who came through for me many a time and who was always fun to talk to. Thanks also to Peggy Berkovitz and Sandra Fowler at MIT for taking care of all the administrative matters.

I would also like to thank the MIT group at SLAC for their initial support in launching my graduate career: John, Saul, Phil, and Amit. My classmates at MIT Frank, Ethan, Ed, Ken, Ron and Robin with whom I learned a lot of physics and had a great time doing so – thanks for the memories gentlemen. Many thanks to Sanjay, Sharad, Rafey, Christian and all my other friends throughout the world for making life tolerable over the past few years.

And finally special thanks to my family whose constant support I have enjoyed throughout my life. Mom, Dad, Shahla, Rashid, Jahangir, Alamgir and Shoaib — you're the best.

This thesis is dedicated to my parents.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A search for the elusive top quark by the Collider Detector at Fermilab (CDF) collaboration, which has been investigating proton-antiproton interactions at a center of mass energy of 1.8 TeV, has recently been carried out. The search was conducted on 21.4 pb^{-1} of data collected between August of 1992 and May of 1993 and combines three separate channels. It is the purpose of this paper to provide details on one of these channels: the lepton plus jets mode with b quarks tagged using their semileptonic decays. The other two channels, the dilepton and the lepton plus jets with b tagging using displaced vertices, are discussed in detail elsewhere [1, 2].

The rest of this chapter motivates and defines the top quark search. Evidence for the existence of the top quark, results of searches to date, expected top quark properties together with signatures at hadron colliders are discussed. Our adopted search mode is defined and compared with the other two search modes being pursued at CDF. In Chapter 2 we briefly discuss aspects of the detector relevant to this analysis. The third chapter describes the event selection procedure. We describe the parameters relevant to lepton identification, jet clustering schemes and the calculation of missing transverse energy. Chapter 4 discusses the crux of this search: the low-P_T (transverse momentum) lepton identification schemes that are instrumental in reducing W+multijet backgrounds to the top signal. Chapter 5 describes the Monte Carlo datasets used in the search. Generators used to study the top production and decay at hadron colliders together with those used in determining the background

predictions are discussed. In Chapter 6 we define the signal region and present the results of the search. We compare our signal to background predictions and consider systematic effects affecting our result. We explore some kinematic properties of events in the signal region and extract a $t\bar{t}$ production cross section assuming the observed excess in the signal region is due to solely to $t\bar{t}$ production. Some consistency checks of our analysis as well as a complete list of events in the signal region are also included in the appendices.

1.1   Evidence for the top quark

The electroweak sector of the Standard Model is based on the SU(2)$_L\times$U(1) gauge group [3, 4, 5]. In this SU(2) structure, the bottom quark is assigned a weak isospin eigenvalue, T_{3L}, of $-\frac{1}{2}$ and must therefore belong to a multiplet of states. One of these states, $T_{3L} = +\frac{1}{2}$, is defined to be the top quark.

Perhaps the most compelling experimental evidence that the bottom quark belongs to a multiplet, and hence that the top quark exists, comes from the measurement of the forward-backward charge asymmetry, A$_{FB}$, in the reaction $e^+e^- \to b\bar{b}$. At e^+e^- machines, A$_{FB}$ takes the form

$$\text{A}_{FB} \propto [\text{T}_{3L} - \text{T}_{3R}][\text{T}_{3L} + \text{T}_{3R} + \frac{2}{3}\sin^2\theta_W] \tag{1.1}$$

where θ_W is the electroweak mixing angle and T_{3R} is the right handed weak isospin eigenvalue. With the assumption that $T_{3R} = 0$, a measurement of A$_{FB}$ yields T_{3L}. The world average of all such measurements [6]

$$T_{3L} = -0.54 \pm 0.13 \tag{1.2}$$

is consistent with the Standard Model.

Further indication that the b quark is not an SU(2) singlet comes from the absence of flavor changing neutral current (FCNC) decays of B mesons. If the b quark is assumed to decay through the Standard Model gauge bosons, the Ws and Z^0, and if

17

it is not a member of an SU(2) doublet, then it should exhibit FCNC decays. For a weak SU(2) singlet, theory predicts [7]

$$BR(B \to l^+l^-X) > 0.013 \qquad (1.3)$$

where l^+ and l^- are leptons from the direct decay of the B meson. This lower limit includes the possibility of additional weak singlet quarks heavier than the bottom quark and is also insensitive to the value of $\sin^2\theta_W$ within the experimental uncertainties. Looking at the electron and muon decay channels of B mesons, the CLEO collaboration [8] has published a value of

$$BR(B \to l^+l^-X) = \frac{BR(B \to \mu^+\mu^-X) + BR(B \to e^+e^-X)}{2} < 0.0012 \qquad (1.4)$$

at the 90% C.L. This is in stark contrast to the theoretical expectations and implies the existence of the top quark.

Another manifestation of the absence of flavor changing neutral current decays of bottom quarks lies in the mixing in the B meson system. Oscillations between B_d^0 and $\overline{B_d}^0$ are often measured in terms of the parameter $x_d = \Delta m/\Gamma$, where Δm is the mass difference between the B_d^0 and $\overline{B_d}^0$ states and Γ is the mean decay width. The ARGUS collaboration has measured this mixing parameter to be $x_d = 0.73 \pm 0.18$ [9]. The presence of flavor changing neutral current decays would increase this value by at least an order of magnitude [10].

From a purely theoretical perspective the top quark is a rather convenient tool for cancelling triangle anomalies. Such anomalies arise because gauge bosons couple differently to left-handed and right-handed fermions. However, it has not been established whether these anomalies would have measurable consequences or not. In fact, it is not even clear that the top quark is the only solution to eliminating these anomalies [11].

1.2 Current limits on the top quark mass

The Standard Model does not predict the mass of the top quark, m_{top}. Upper and lower bounds on m_{top} have been imposed through numerous searches, both direct and indirect, at particle accelerators worldwide and also through purely theoretical considerations. These limits have been obtained at both e^+e^- machines and hadron colliders using a variety of decay channels. Direct searches as well as inferences from precise measurements of other electroweak parameters have contributed. Here we summarize some of the more significant results.

Numerous searches have been conducted at the LEP collider in Europe. In looking for the process

$$e^+e^- \rightarrow t\bar{t} + X, \quad t\bar{t} \rightarrow bW^+\bar{b}W^- \tag{1.5}$$

where both the W bosons in the subsequent top quark decay are virtual, the ALEPH collaboration has excluded m_{top} in the range 26.0 to 45.8 GeV/c^2 by looking for spherical events with an isolated charged particle [12]. The OPAL and DELPHI collaborations have tested the top decay modes in extensions of the Standard Model. In particular, they tested the decay

$$t \rightarrow b + H^+ \tag{1.6}$$

hypothesis with 100% branching ratio and the subsequent decay of the charged Higgs to hadrons. The OPAL collaboration obtained a limit on m_{top} without any isolated charged particle requirement, by looking at the acoplanarity event shape parameter. Assuming the charged Higgs exists and the decay represented by equation (1.6) occurs 100% of the time, they eliminated m_{top} in the range $(m_{H^+} + 5.2)$ to 45.2 GeV/c^2 at the 95% C.L. where the charged Higgs mass took the values $m_{H^+} = 23, 28, 33$ or 38 GeV/c^2 [13]. DELPHI excluded m_{top} in the range 33 to 44 GeV/c^2 assuming the mass of the charged Higgs to be greater than 30 GeV/c^2 but at least 6 GeV/c^2 lighter than the top quark [14].

A decay-mode independent limit was also obtained by ALEPH. They measured

the total hadronic cross section at the Z^0 peak and extracted a lower limit of m_{top} > 45.8 GeV/c^2 at the 95% C.L. [12]. Theoretical constraints on the top quark mass can be imposed via the electroweak radiative corrections to the W and Z boson masses, the Z decay width and several weak neutral current observables (Figure 1-1). Recently, global fits to precision electroweak measurements at LEP have yielded a top mass of 164^{+17+18}_{-16-20} GeV/c^2 [15].

(a)
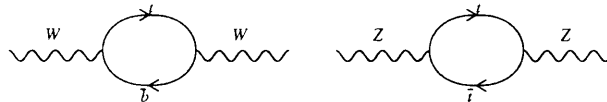
(b)
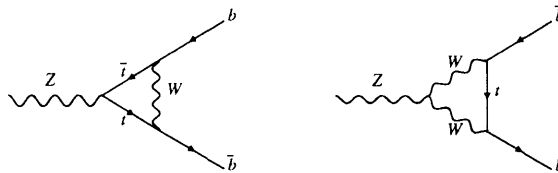
Figure 1-1: Radiative effects involving the top quark. (a) W and Z boson masses, and (b) vertex correction to the Z partial width.

Searches performed at hadron colliders have constrained the top mass further. At the CERN SPS, which operated at the center of mass energy of 630 GeV, the UA2 collaboration looked for events with a high energy electron and a neutrino. Assuming charged current decays of the top quark, they excluded m_{top} in the range 30 to 69 GeV/c^2 at the 95% C.L. [16].

Meanwhile, CDF has performed several top quark searches using data collected prior to the 1992-93 collider run. Assuming the minimal Standard Model (no charged

Higgs decays), CDF has looked for events with two high energy leptons together with considerable missing transverse energy, the so-called dilepton decay channel. From looking at the (e,μ) channel alone CDF excluded the top mass in the range 28 to 72 GeV/c^2 [17] and extending this search to all dilepton modes the lower limit has been extended to 91 GeV/c^2 at the 95% C.L. [18]. Analysis of part of the 1992-93 data has further increased this limit to 113 GeV/c^2 [2].

The CDF collaboration has also measured the ratio of cross sections of the production of W and Z^0 bosons and their subsequent decays into electrons, R, produced at $p\overline{p}$ collisions. R is related to the W decay width, $\Gamma(W)$ through [19]

$$R = \frac{\sigma(W \to e\nu)}{\sigma(Z^0 \to e^+e^-)} = \frac{\sigma(p\overline{p} \to WX)}{\sigma(p\overline{p} \to Z^0X)} \frac{\Gamma(W \to e\nu)}{\Gamma(Z^0 \to e^+e^-)} \frac{\Gamma(Z^0)}{\Gamma(W)} \qquad (1.7)$$

The ratio of cross sections is calculated using Standard Model couplings and parton distribution functions. The LEP measurement of the Z boson partial decay width, $\Gamma(Z^0) \to e^+e^-$ [20] together with the R measurement enables CDF to extract the branching ratio, $BR(W \to e\nu)$. This branching ratio is sensitive to new decay modes of the W boson. One such decay mode is the W decaying to top [1]. So precise determination of the $(W \to e\nu)$ branching ratio constrains the top mass independent of the top quark decay mode. This is an important addition to the constraints from direct searches which assume no extensions to the Standard Model. CDF has measured [21]

$$R = 10.90 \pm 0.32(stat) \pm 0.28(sys) \qquad (1.8)$$

from which a decay mode-independent lower limit on the top mass has been imposed. This limit is 62 GeV/c^2 at the 95% C.L. [21].

---

[1] if the W boson is heavier than the top quark.

1.3 Production mechanism

Top quarks can be produced directly through the QCD process

$$pp \to t\bar{t}X \tag{1.9}$$

or, if kinematically allowed, through W boson production and subsequent decay $W \to t\bar{b}$. At the Tevatron ($\sqrt{s} = 1.8$ TeV), the dominant production mechanism is through the QCD process $p\bar{p} \to t\bar{t}\ X$ for all top masses (see Figure 1-2). Henceforth we consider this direct production mechanism only.
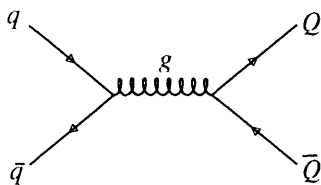
Figure 1-2: Top quark production. Shown is the Leading Order diagram for the direct QCD production process.

The theoretical calculation of the $t\bar{t}$ production cross section is based on the QCD improved parton model. The two main ingredients to the calculation are

- The calculation of the matrix elements of the interactions between the different constituent quarks and gluons, each interaction being treated as a separate subprocesses and the sum of all pertinent subprocesses yielding the total hard-scattering cross section, $\hat{\sigma}_{ij}$. Some Next-to-Leading Order Feynman diagrams used in the calculating $\hat{\sigma}_{ij}$ are shown in Figure 1-3.

22

- The distribution of each type of parton within the parent proton and anti-proton is given by the structure function for those partons, $f(x_i)$.
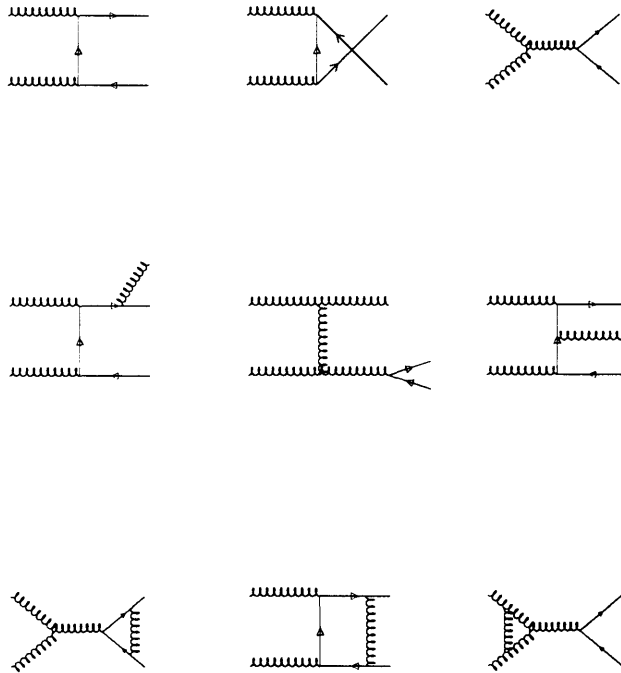
Figure 1-3: Some Next-to-Leading Order graphs for $t\bar{t}$ production at hadron colliders.

The structure functions predict the fraction of incoming hadron (proton or anti-proton) momentum carried by each constituent parton. Typically each parton carries less than 30% of the parent hadron's momentum. So the energy available for $t\bar{t}$ production is $x_p \cdot x_{\bar{p}} \cdot S$ where S is the center of mass energy at the Tevatron, 1.8 TeV, and x_p and $x_{\bar{p}}$ are the fractional momenta of the proton and antiproton carried by the partons.

Structure functions for quarks and anti-quarks are determined from deep inelastic scattering of leptons (e, μ, ν) from proton and neutron targets. Since gluons do not interact with electroweak probes, their structure functions are the hardest to determine. Gluon structure functions are determined by studying scattering from

23

"sea" quarks and also indirectly by invoking the momentum sum rule i.e., the sum of all fractional momenta of all the partons in a proton (anti-proton) should add up to one. These distribution functions have the largest uncertainties associated with them. Structure functions are determined at much lower four-momentum transfers (Q^2), typically less than 10 GeV^2, than occur at Tevatron energies. Four-momentum transfer dependence is obtained via the Alterelli-Parisi equations [22]. The top cross section used in this analysis is based on MRSD' structure functions [23].

The contributions calculated in this manner for gluon-gluon fusion (gg) and quark-antiquark annihilation $q\bar{q}$ graphs are shown in Figure 1-4. Flavor excitation graphs (qg and $\bar{q}g$) are found to be negligible in the top mass range of interest. Also, $q\bar{q}$ annihilation is expected to dominate gluon-gluon fusion at top masses greater than about 100 GeV/c^2 [24].
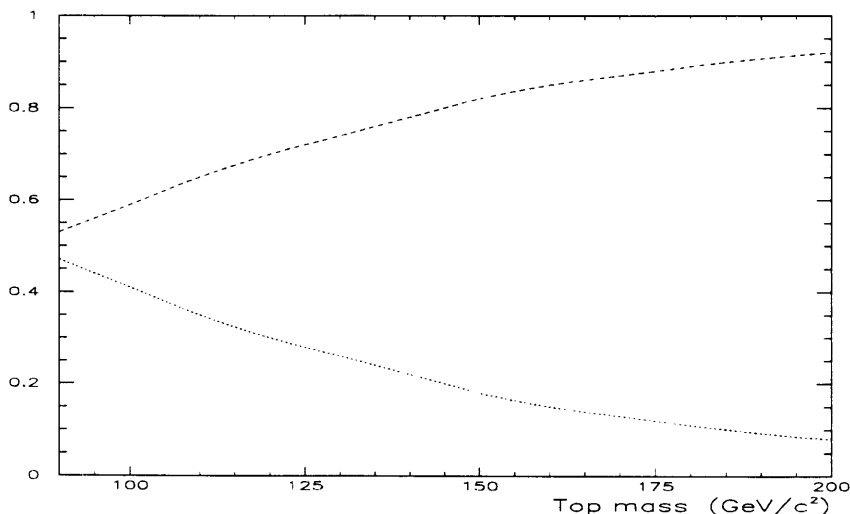
Figure 1-4: Contributions from gluon-gluon fusion gg (dotted line) and quark-antiquark annihilation $q\bar{q}$ (dashed line) processes at the Tevatron. From [24].

Putting all this together, one gets

$$\sigma(S, m_{top}, \mu_R, \mu_F) = \sum_{i,j} \int_{\tau_0}^{1.0} dx_1 \int_{\frac{\tau_0}{x_1}}^{1.0} dx_2 f_i^p(x_1, \mu_F) f_j^{\bar{p}}(x_2, \mu_F) \hat{\sigma}_{ij}(x_1 x_2 S, m_{top}, \mu_R, \mu_F)$$

(1.10)

here $\sqrt{S}$ is the $p\bar{p}$ center of mass energy, $\tau_0 = \frac{2m_{top}^2}{S}$ and $f_i^p(x, \mu_F)$ are the structure

24

functions, μ_R is the renormalization constant and $\hat{\sigma}_{ij}$ is the parton cross section for the $ij \to t\bar{t}X$ process and is evaluated by expanding in powers of $\alpha_s(\mu_R)$, the strong running coupling constant. Interactions involving more than one parton per hadron are not included; they are suppressed at high Q^2. Also, final state interactions, which cause decay and hadronization, are assumed to be independent of $\hat{\sigma}_{ij}$. The $\hat{\sigma}_{ij}$s used in this analysis have been calculated using the exact order α_s^3 corrected cross section and the resummation of the leading soft gluon corrections in all orders of perturbation theory. A theoretical prediction of the top quark cross section is shown in Figure 1-5. Details of the calculation of the cross section can be found in references [24, 25].
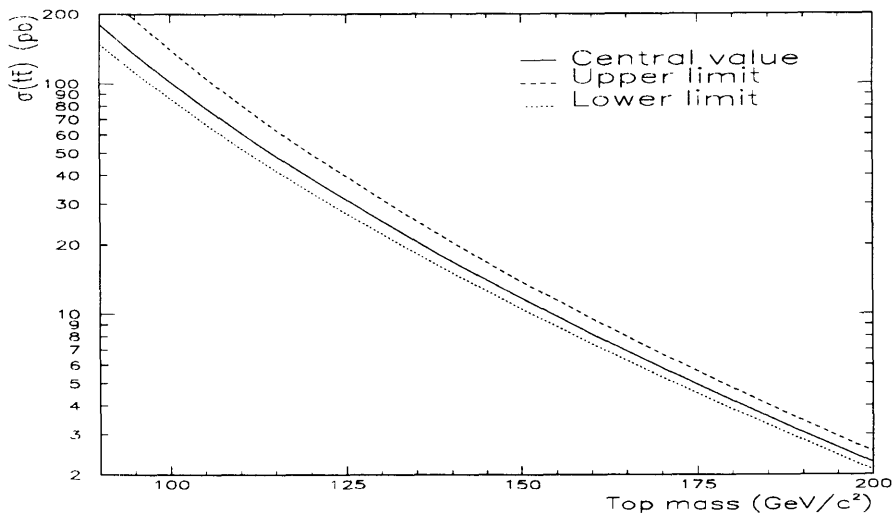
Figure 1-5: Top quark production cross section at the Tevatron. From [24].

1.4 Hadronization and decay

After creation, the top quark hadronizes and decays. There is no fundamental understanding of hadronization, it is a soft process and therefore not calculable in perturbative QCD. Two empirically successful heuristic approaches are employed in most Monte Carlo generators: the string model by the Lund group of Sweden [26] and the Feynman-Field independent fragmentation model [27]. The Feynman-Field ansatz is employed in this work. Here we summarize the basic ideas of this approach.

The top quark is always created in a color singlet state with respect to some recoiling system of quarks and/or gluons. As the two systems move apart, the color field between them creates $q\bar{q}$ pairs from the vacuum. The initial quarks then combine with one or two of these quarks to form colorless hadrons, leaving the other quark to continue the reaction. The momentum fraction, z, is considered the only relevant parameter in the model.

Hadronization is described by a set of functions $D_q{}^h(z)$, determined by phenomenological models and empirical data. The $D_q{}^h(z)$ (q-to-h fragmentation function) is defined as the probability density for finding a hadron of type h sharing a fraction between z and $z + dz$ of the energy of it's parent parton, q. The assumption that $D_q{}^h(z)$ depends only on z is plausible if the parton energy is large compared to all participating masses and P_Ts.

For heavy quark fragmentation, Peterson et al have supplied functions that agree well with data on bottom and charm quark fragmentation [28]. The Peterson model is based on the fact that heavy quarks do not have to give up much energy to pick up a light quark travelling at the same velocity. Hence heavy quarks should fragment into heavy hadrons with large z: the heavier the quark, the larger the z of the formed hadron. The fragmentation function in this model takes the form

$$D_q{}^h(z) = \frac{K}{z\left(1 - \frac{1}{z} - \frac{\epsilon_Q}{(1-z)^2}\right)} \tag{1.11}$$

where K is a normalization constant and ϵ_Q is a parameter that is proportional to $1/m_Q{}^2$.

Within the Standard Model framework, the realm of this search, the top quark very nearly always decays into a bottom quark and a W boson. The decays $t \rightarrow s + W^+$ and $t \rightarrow d + W^+$ are strongly suppressed due to the smallness of the Kobayashi-Maskawa matrix elements, V_{ts} and V_{td} compared to V_{tb}.

$$t \rightarrow b + W^+ \tag{1.12}$$

The W boson is real if the top mass is $\gtrsim 85$ GeV/c^2, the sum of the W boson mass and the b quark mass.

In extensions of the Standard Model where the Higgs sector becomes important, the decay

$$t \rightarrow b + H^+ \tag{1.13}$$

with the subsequent hadronic decay of the charged Higgs $H^+ \rightarrow c\bar{s}$, $\tau\nu_\tau$, ... becomes comparable to the Standard Model decay described earlier. The search described in this paper assumes no extensions to the Standard Model. If the Higgs sector as described by equation (1.13) is important, the results of a search in the lepton+jets mode will be significantly changed.

The W bosons subsequently decay either semileptonically or into hadrons. The Standard Model branching ratio for semileptonic decay of the W boson is $\frac{1}{9}$ for each lepton species. The branching ratio for hadronic decay (summed over quark species) is $\frac{2}{3}$. The b quarks produced in the decay hadronize into jets. Table 1.1 displays the branching ratios for all the possible Standard Model decays of the top quark.

decay mode	Branching ratio (BR)
$e\nu b + e\nu b$	1/81
$\mu\nu b + \mu\nu b$	1/81
$\tau\nu b + \tau\nu b$	1/81
$e\nu b + \mu\nu b$	2/81
$e\nu b + \tau\nu b$	2/81
$\mu\nu b + \tau\nu b$	2/81
$e\nu b + qq\bar{b}$	12/81
$\mu\nu b + qq\bar{b}$	12/81
$\tau\nu b + qq\bar{b}$	12/81
$qq\bar{b} + qq\bar{b}$	36/81

Table 1.1: Lowest order branching ratios of the top quark assuming charged current decays.

1.5 Top signature at CDF

In the Standard Model framework, the $t\bar{t}$ decays are characterized by the decay mode of the final W^+W^- pair. Each W boson decays either leptonically, into (e, ν_e), (μ, ν_μ) or (τ, ν_τ), or hadronically, into quark-antiquark pairs (see Figure 1-6). The hadronic decay products appear as jets in the detector. Also, since tau leptons are difficult to identify in the detector only electrons and muons are considered. The three distinct decay modes being considered at CDF are now described.
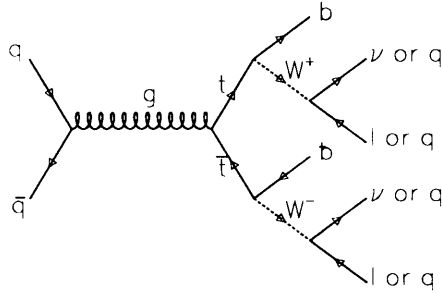
Figure 1-6: $t\bar{t}$ production and decay.

The dilepton mode

In the event that both the W bosons decay semileptonically, which happens only $\sim 5\%$ of the time, there are two high-P_T leptons together with considerable missing transverse energy[2] and some jet activity due to the b quarks in the event. This is by far the cleanest decay mode but suffers from a small braching fraction. The main backgrounds to this signature arise from $Z \rightarrow \tau\tau$ decays, $b\bar{b}$, WW and WZ production. Kinematic and topological cuts are applied to make the signal-to-background greater than 1 while maintaining reasonable efficiency for $t\bar{t}$ production.

---

[2]Due to the neutrinos which go undetected.

The hadronic mode

About 44% of the time both the Ws decay into quark-antiquark pairs resulting in events with 6 jets. This large branching fraction is swamped by the huge QCD multijet production resulting in a signal-to-background ratio at the production level of 1:16,000 [29]. Kinematic cuts can increase this ratio to 1:100. Efforts are underway to further reduce the backgrounds with b-tagging techniques discussed later. This mode could prove to be very useful in increasing the statistics if the backgrounds could be reduced to within tolerable limits.

The lepton + jets mode

When one of the Ws decays semileptonically and the other hadronically, the event signature is a high-P_T lepton with significant missing transverse energy and four jets[3]. In practice, due to detector effects such as cracks and inefficiencies in clustering algorithms, the lepton+jets search requires three or more jets in the event. This mode occurs roughly 30% of the time. By far the dominant background here is the production of W+multijets. Figure 1-7 shows one example of the higher-order production of a W recoiling against significant jet activity. The rate of W+multijet production is 2 to 10 times the rate of $t\bar{t}$ production at the Tevatron depending on the top mass and the exact jet cuts used. This signal-to-background ratio can be significantly improved with some mechanism for identifying b quarks. Two such b-tagging approaches have been adopted at CDF and are discussed below. With the b-tag the signal-to-background ratio becomes better than 1 to 1.

Table 1.2 summarizes the main features of these three decay channels.

1.5.1 b-tagging

As mentioned earlier, the background due to W+multijet production is 2 to 10 times the signal in the lepton+jets channel. The presence of the additional b quarks in a top event becomes an important discriminatory aspect of these events. Thus identifying

---

[3]Two from the decay of the W and two from the b quarks.
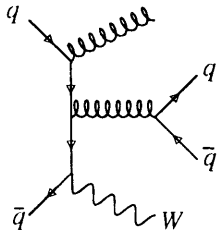
Figure 1-7: A typical W+jet production diagram.

the b quarks in top decay is an extremely powerful tool in reducing backgrounds in all decay modes. At CDF b quarks are currently identified in one of two ways:

1. SVX-TAGGING. The silicon microvertex detector is used to find displaced vertices due to the long lifetime of B mesons [2].

2. SOFT-LEPTON TAGGING (SLT). The semileptonic decays, both direct and sequential,

$$b \rightarrow e/\mu \tag{1.14}$$

$$b \rightarrow c \rightarrow e/\mu \tag{1.15}$$

of b quarks are used to tag the presence of heavy flavor in an event. These leptons tend to have a softer P_T spectrum than the leptons from direct W decay (see Figure 1-8) and are consquently often referred to as 'soft leptons.' Algorithms developed to identify these low-P_T leptons in CDF are described in detail in Chapter 4.

This paper describes the search for the top quark in the lepton+jets channel with soft-lepton b-tagging (the SLT search).

	Dilepton	Hadronic	Lepton+jets
Branching fraction	~5%	~44%	~30%
Signature	2 high-P_T leptons significant $\not{E}_T$ 2 jets	6 jets	1 high-P_T lepton significant $\not{E}_T$ 4 jets
Main background	$b\bar{b}$ $Z \to \tau\tau$ Dibosons Misidentifications	QCD multijet production	W+multijet production Misidentifications
Signal-to-background (kinematic cuts only)	1:1	1:100	1:2-10
Effects of b-tagging	would help, BR too small though	absolutely essential	turns it into powerful technique
Acceptance	~1%	~2%	~1%
Comments	cleanest signature, reasonable efficiency	high BR, but needs lots of work to reduce huge QCD backgrounds	coupled with b-tagging, very promising mode of top discovery

Table 1.2: The three decay modes being considered at CDF.
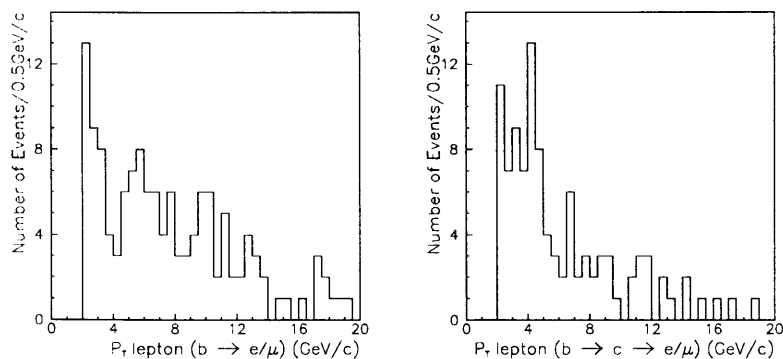
Figure 1-8: P_T spectrum of leptons from the decay of b and c quarks in Monte Carlo top events (m_{top} = 120 GeV/c^2 , a P_T cut of 2 GeV/c has been applied).

31

Chapter 2

Collider Detector at Fermilab

The Collider Detector at Fermilab (CDF) is the first of two general purpose detectors built to study $p\bar{p}$ interactions at the Fermilab Tevatron. The Tevatron is currently the world's highest energy collider; proton-antiproton collisions take place at an unprecedented center of mass energy of 1.8 TeV.

CDF is a general purpose azimuthal and forward-backward symmetric detector. It was commissioned in 1987 and took 4.3 pb^{-1}of data in the 1988-89 Collider Run. Since then CDF has undergone several upgrades, some of which are crucial to this analysis. These upgrades include the addition of a Silicon microstrip Vertex detector (SVX), a Pre-Shower Counter (CPR), extension of the muon coverage and several enhancements to the Data Acquisition system (DAQ). This paper presents analysis of 21.4 pb^{-1}of data collected with CDF in the 1992-93 Collider Run.

Fermi National Accelerator Laboratory and the Tevatron

A setup of the $p\bar{p}$ colliding facility is shown in Figure 2-1. The protons used in the collisions originate from Hydrogen gas molecules which are first ionized and the resulting ions accelerated to 750 KeV in a Cockroft-Walton electrostatic generator. The Hydrogen ions are then accelerated to 200 MeV in a linear accelerator (Linac). The ions emerge from the Linac through a carbon foil that strips their outer electrons thereby leaving only protons. Protons thus produced are stored in a Booster Ring, a synchrotron accelerator, where they are accelerated to 8 GeV. In the Booster,
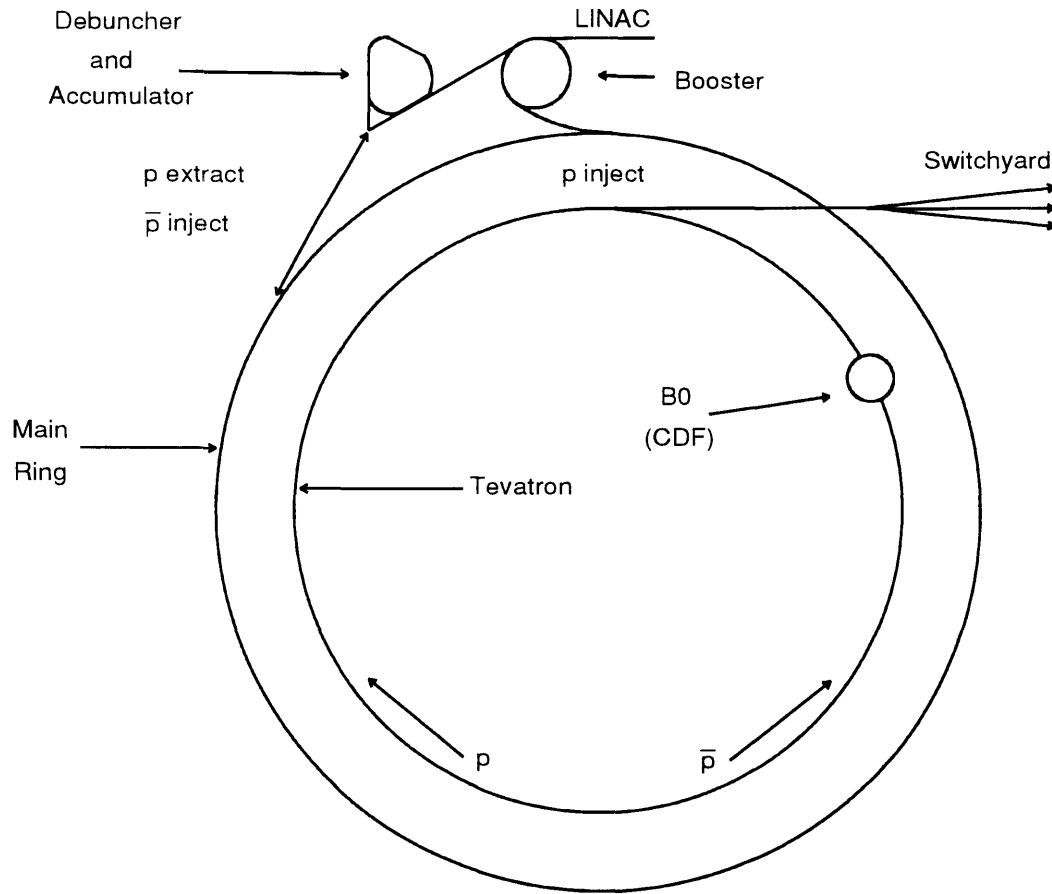
Figure 2-1: The $p\bar{p}$ Collider setup at Fermilab.

proton bunches are collected and injected into the Main Ring, also a synchrotron accelerator. The Main Ring consists of 4 miles of alternating dipole (bending) and quadupole (focusing) magnets. The Main Ring also contains RF cavities that boost the protons to 150 GeV. The Main Ring then injects these protons into the Tevatron or to Fixed Target Experimental Stations. Protons from the Main Ring also act as the source of antiprotons. These protons are removed from the Main Ring and focused on a target which produces, amongst other secondary particles, antiprotons. Antiprotons are stored in the Accumulator ring. After enough antiprotons have been accumulated, they are reinjected first into the Main Ring and then eventually into the Tevatron. The Tevatron is a ring of superconducting magnets that lies directly beneath the Main Ring. In the Tevatron protons and antiprotons are accelerated to

900 GeV and then made to collide at two interaction regions, B0 and D0, at which are housed Fermilab's two detectors, CDF and D0 respectively.

At the B0 interaction region the beam is roughly circular in cross-section with a radius defined by one σ of $\sim 40\mu$m. The longitudinal profile of the beam is approximately gaussian with a width of 30-35 cm.

2.1 Overview
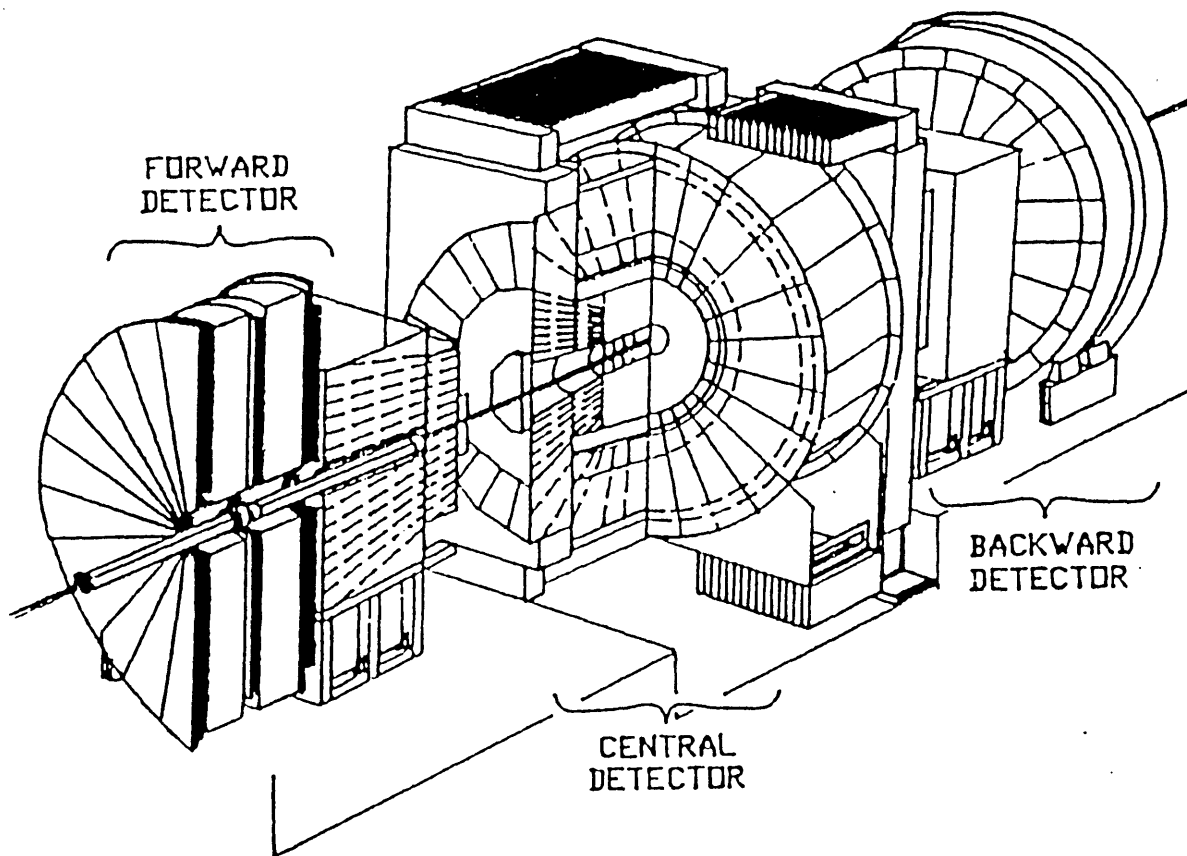
Figure 2-2: A perspective view of CDF showing the forward, central and backward sections. This figure illustrates the azimuthal and forward-backward symmetry of CDF.

The CDF detector has been described in detail elsewhere [30], in this chapter we summarize features of the detector relevant to this analysis. The CDF detector employs

34

charged particle tracking, magnetic momentum analysis and fine-grained calorimetry. It determines the energy, momentum and in some cases the identity of particles created in $p\bar{p}$ collisions. Figure 2-2 shows a perspective view of the detector and Figure 2-3 shows a side view cross section of one half of the detector. Particles produced at the vertex encounter in succession a Beryllium beam pipe, various tracking chambers, sampling calorimeters and muon chambers.
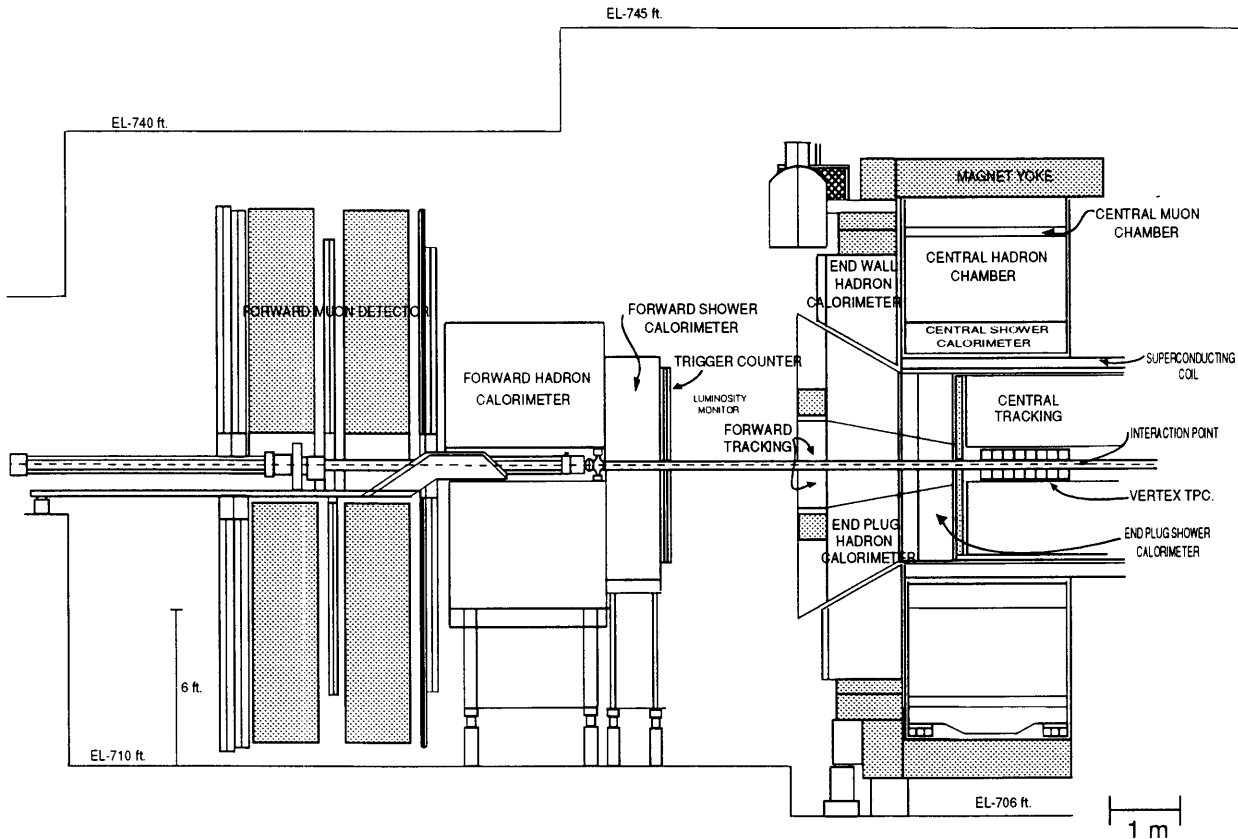
Figure 2-3: A cross section of one half of the CDF detector as configured for the 1989 Run.

Immediately surrounding the interaction point is a Silicon Vertex detector (SVX) used to measure displaced vertices. Surrounding the SVX are a series of Vertex Time Projection Chambers (VTX) which reconstruct the z coordinates of the vertices of every interaction. A Cylindrical Drift Chamber (CTC) envelops the VTX and provides tracking information. An axial magnetic field of 1.412 Tesla permeates the

CTC. Surrounding the CTC are both electromagnetic and hadronic calorimeters. Behind thick plates of steel and furthest from the beamline lie CDF's muon chambers that detect minimum ionizing particles.

Coordinate System

CDF employs a right-handed coordinate system in which the z−axis lies in the proton direction, the y−axis points upwards from the plane of the Tevatron and the x-axis points radially outwards (see Figure 2-4). The interaction point, coordinates $(0,0,0)$, is taken to be the geometric center of the detector. The polar angle, θ, is measured in the $x - z$ plane and goes to zero in the positive z−direction
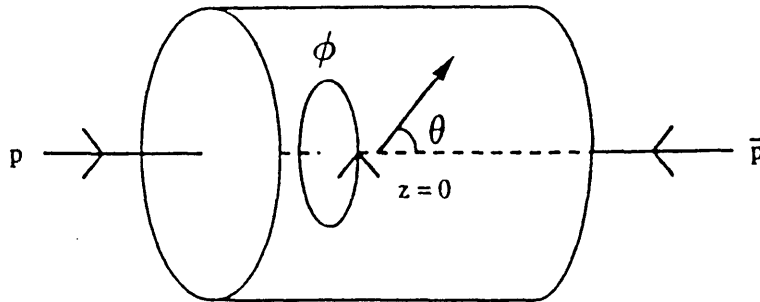
Figure 2-4: The CDF coordinate system.

It is useful to work in terms of the geometric parameters, rapidity,

$$y = \frac{1}{2}\ln\left(\frac{E + P_z}{E - P_z}\right)$$ (2.1)

Tracking system	Angular coverage	2 track resolution	Momentum resolution
SVX	$\sim 6.3° < \theta <\sim 173.7$	$-$	$\delta P_T/P_T= 0.011 \times P_T(\text{GeV/c})$
VTX	$\sim 8° < \theta <\sim 172°$	6mm/θ (Z) 6mm (R) 3 cm (ϕ)	$-$
CTC	$\sim 40° < \theta <\sim 140°$	3.5 mm	$\delta P_T/P_T= 0.002 \times P_T(\text{GeV/c})$

Table 2.1: Tracking systems at CDF.

and the closely related pseudorapidity,

$$\eta = \ln\left(\tan\frac{\theta}{2}\right) \tag{2.2}$$

Both η and y describe the direction of partons in the detector and the longitudinal boost of the parton collision frame. Compared to parton collision energies at the Tevatron (typically greater than 30 GeV), the parton masses are taken to be zero. In this limit, η and y are equivalent. Throughout this paper η will be used extensively.

2.2   Tracking

The CDF tracking system covers the angular range $\sim 8°$ to $\sim 172°$ in polar angle, with three dimensional tracking available from 40° to 140°. The tracking system achieves excellent resolution due to the 1.412 Tesla magnetic field provided by a Superconducting Solenoid. Tracking at CDF is used to provide position information of charged particles along their helical trajectories in the soledoinal magnetic field. Absence of tracks matched to electromagnetic clusters in the calorimeter (described later) are also used to identify photons.

The tracking system consists primarily of the solenoid, a silicon microvertex detector (SVX), a vertex time projection chamber (VTX), and the Central tracking chamber (CTC). These devices are summarized in Table 2.1.

2.2.1 Solenoid

A 1.412 T magnetic field is produced by a NbTi/Cu superconducting solenoid and permeates a cylindrical volume of 4.8 m in length and 3 m in diameter. An average current of 4650 Amps produces the magnetic field. The magnetic field flux is returned through a steel yoke which also functions to support the calorimeters described later in this chapter (see Figure 2-3). The magnetic field is mapped precisely and Offline corrections to tracking are preformed using this map. Some non-uniformities in the field are also corrected for using this map.

2.2.2 The SVX

One of the major upgrades of CDF for Run 1A was the installation of a silicon microstrip vertex detector (SVX), the innermost and closest to the interaction point detector of CDF [31]. The SVX surrounds the 1.9 cm radius beampipe and is itself surrounded by the VTX. It provides precise tracking in the r-ϕ plane in order to measure the impact parameter of particles traversing through its fiducial volume.

The SVX is segmented into two halves (barrels), one on either side of the nominal interaction point (see Figure 2-5). Each barrel is composed of 12 wedges, each wedge occupying 30° in ϕ. At both ends of each barrel is a beryllium support structure (bulkhead) that contains readout electronics and chilled water cooling tubes. The SVX is operated at a nominal temperature of 20°C. There are 4 layers in each wedge, numbered 0 to 3. Each layer consists of a ladder which in turn consists of 3 rectangular strip detectors laid end-to-end (3×8.5 cm). The active length of each barrel is 25.5 cm, that is, a total of 51 cm along the beam direction. The ladder charactersitics of each layer are shown in Table 2.2. Each wedge is read out independently, which leads to 46,080 channels for the SVX alone. This is a substantial fraction of all readout channels for CDF.

Because the longitudinal spread of the beam is $\sim$ 30 cm (see Figure 2-6), the SVX acceptance is 60% for $p\bar{p}$ interactions. The SVX hit information, when combined with CTC hit information improves track helix measurements. Track reconstruction in the
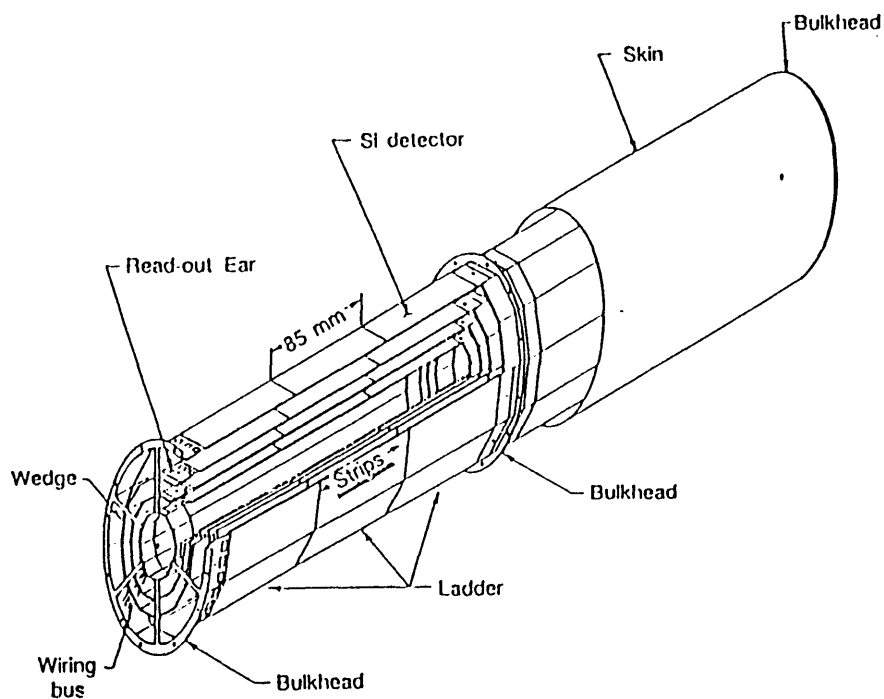
Figure 2-5: An isometric view of the CDF Silicon Vertex detector (SVX).

SVX begins with a CTC track that is required to have at least 2 hits in the SVX not associated with any other track. The SVX momentum resolution is $\delta P_T/P_T = 0.011 \times P_T(\text{GeV}/\text{c})$. The individual hit resolution is measured in the data to be $\sim 11\mu\text{m}$.

The SVX can be used to study particles with lifetimes on the order of a picosecond. Thus the SVX provides the basis for a top search [2], in which displaced vertices (believed to come from b quarks possibly arising from top quark decay) are detected in the SVX. This analysis only uses the SVX's superior beam position measurement but otherwise does not use its tracking capabilities.

Layer	Face	Radius [cm]	Tot. Width [cm]	Act. Width [cm]	Pitch [μm]	Num. of Strips	Num. of Chips
0	IN	2.9899	1.6040	1.536	60	256	2
1	OUT	4.2710	2.3720	2.304	60	384	3
2	OUT	5.7022	3.1400	3.072	60	512	4
3	IN	7.8508	4.2925	4.224	55	768	6

Table 2.2: Characteristics of each SVX layer.
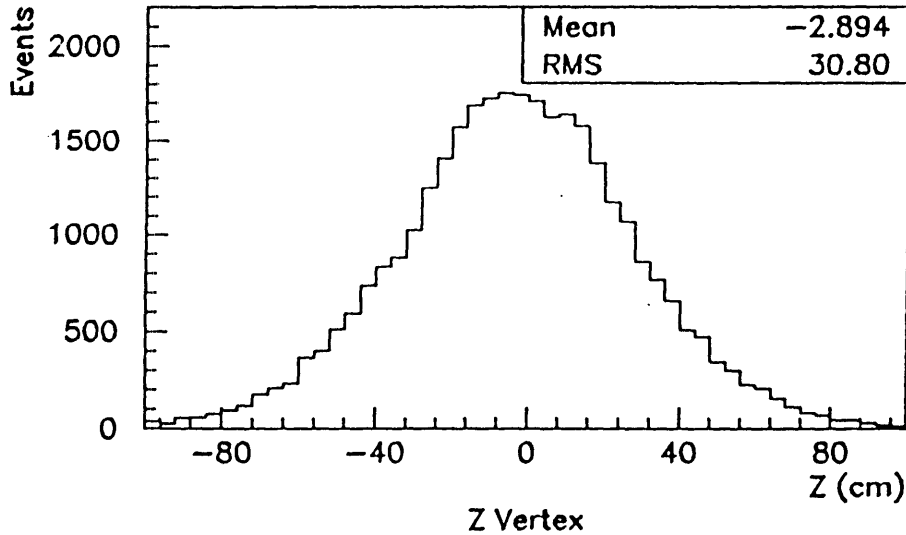
Figure 2-6: The event vertex distribution for a typical CDF Run.

2.2.3 The VTX

The Vertex Time Projection Chamber (VTX) for the 1992-93 Run is an upgrade to a similar chamber (VTPC) [30] used in previous runs. The upgrade was necessary to handle the much higher luminosity expected in Run 1A. It provides tracking information out to a radius of 22 cm ($| \eta | < 3.25$). The VTX is segmented into 2 halves, East and West, with the two halves lying on either side of the nominal interaction vertex. The construction of a VTX module is shown in Figure 2-7. As mentioned earlier, the VTX completely surrounds the SVX. Each half consists of 28 Octagonal

SENSE WIRES

SENSE WIRE READ-OUT

CATHODE PADS

CENTER H.V. GRID

CARBON FIBER OCTOGON

RADIAL BOARD

CARBON FIBER OCTOGON

INNER FIELD CAGE

PAD READ-OUT

DRIFT

DRIFT

P.W.C. CATHODE GRID

OUTER FIELD CAGE

RESISTIVE DIVIDER
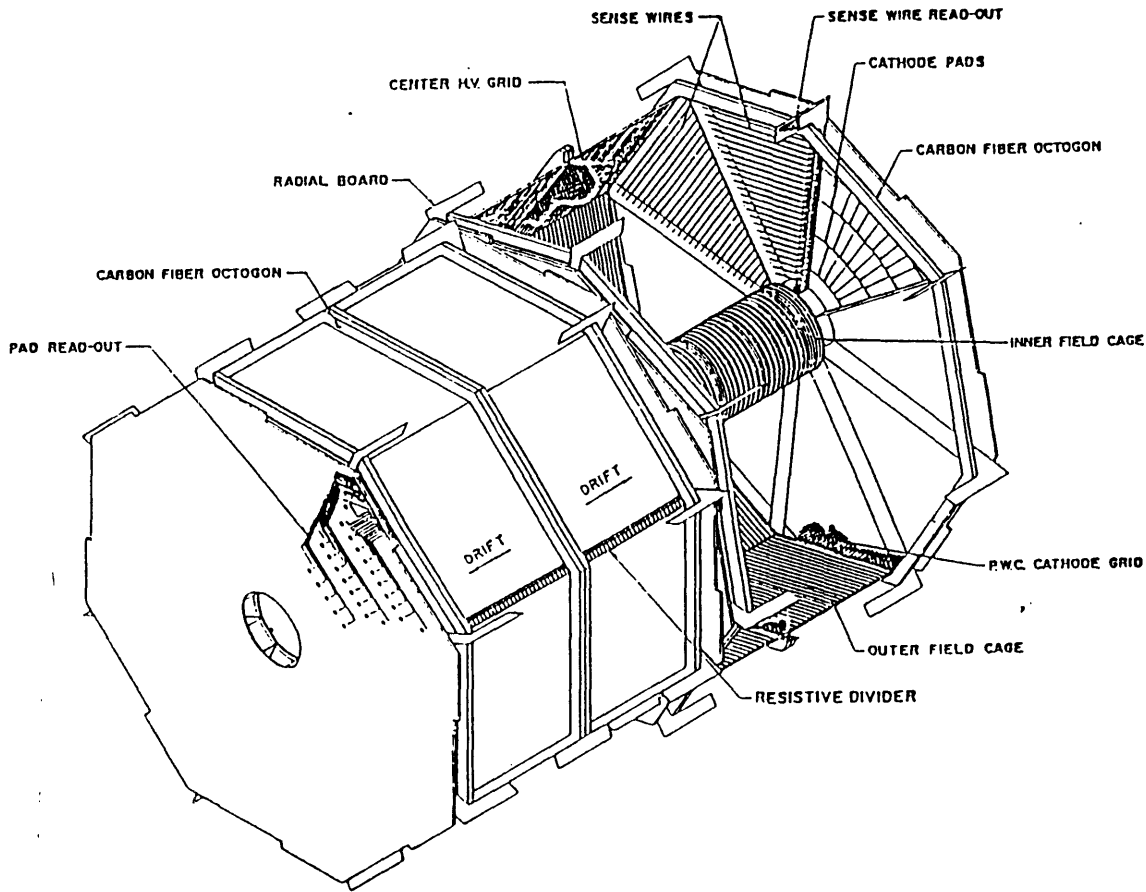
Figure 2-7: An isometric view of two of the 56 modules that make up the VTX. Shown above are modules from the Vertex Time Projection chamber used in the 1988-89 Run.

modules which extend out to 1.5 m on either side of the nominal interaction point. The inner 18 modules that surround the SVX contain 16 sense wires strung in the r-ϕ plane perpendicular to a radial line from the origin. Ten additional modules have 24 sense wires. The drift gap in the Argon-Ethane atmosphere between sense wires is 4 cm. Adjacent modules are rotated $11.3°$ in ϕ to aid better matching of the CTC r-ϕ track segments to the VTX segments.

By measuring the drift times of electrons hitting sense wires, a primary particle track can be reconstructed in the r-z plane. The azimuth of the track is obtained from charge induced on the cathode pads. The VTX provides 3-D track reconstruction for $|\eta| < 3.25$.

The main function of the VTX is to locate the primary vertex of an event, that is, the $p\bar{p}$ interaction point. This is achieved by finding the convergence of all reconstructed tracks in the event. The z resolution of primary vertices found by the VTX is 1-2 mm depending on track multiplicity in the event. Since collisions are gaussianly distributed with $\sigma \approx 35$ cm, the VTX provides ample coverage. The event vertex is used to calculate transverse quantities.

VTX materials are chosen to have low mass and long radiation lengths. This minimizes photon conversions and Multiple Coulomb scattering which degrades tracking efficiency and momentum resolution. For $| \eta | < 1.5$, the VTX presents less than 2% of a radiation length of material. The VTX is also used to identify photon conversions, as will be elaborated on in a later chapter.

2.2.4 The CTC

Surrounding the VTX is a 3.20 m long Cylindrical Drift Chamber. It fits inside a superconducting magnetic solenoid that provides a 1.412 T axial magnetic field. The chamber consists of 84 layers of sense wires grouped in 9 superlayers: 5 axial and 4 stereo (see Figure 2-8). The axial superlayers consist of 12 wires each arranged parallel to the beam direction that enable track reconstruction in the $r - \phi$ plane. The stereo superlayers, interspersed between the axial ones, have 6 wires each, at $\pm 3°$ to the axial direction and enable stereo track reconstruction i.e in the $r - z$ plane.

Table 2.3 summarizes the mechanical parameters and resolution performance of the CTC. The superlayers are further subdivided into cells. This results in a maximum drift distance of 40 mm which corresponds to a drift time of 800 ns which is much shorter than the 3.5 μs between beam crossings. Cells are tilted by 45° with respect to radial direction (see Figure 2-9). This tilt resolves the left-right ambiguity. Combined with the large number of wires per cell and cells per superlayer, the large tilt angle insures that tracks will come very close to at least one sense wire in each superlayer they cross. This condition is exploited Offline to separate closely spaced tracks. It is also used to generate a fast trigger signal when one or more high-p_T tracks are present as is done in the Central Fast Tracker (CFT). The azimuthal position resolution in

42

554.00 mm I.D.

2760.00 mm O.D.
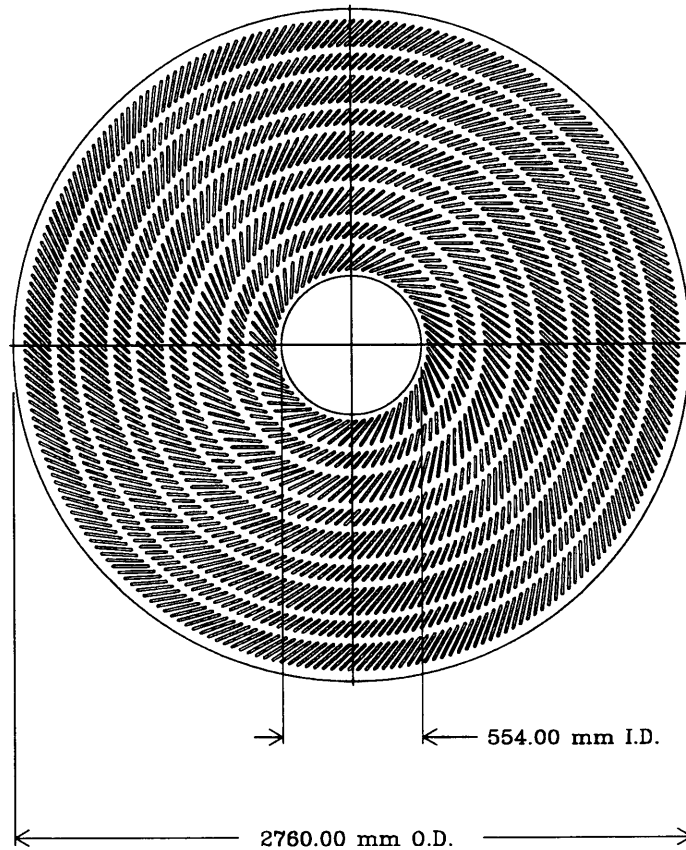
Figure 2-8: The structure of the Central Tracking Chamber (CTC).

each layer is better than 200 μm and the z resolution of a stereo wire is less than 4 mm.

The bulk of this analysis relies heavily on CTC tracking. The CTC performance is excellent, as is extensively documented elsewhere [30]. The excellent momentum resolution achieved by the CTC is largely due to the large magnetic field strength and the large tracking volume. Though beam constrained fits further improve the momentum resolution, they are not important to this analysis.

The excellent position resolution is used to identify electrons (by matching tracks to energy deposition in electromagnetic calorimeters) and muons (by matching tracks to hits in muon chambers). The CTC is also used to identify secondary vertices due to decay of long-lived particles, to study calorimetry response as a function of

Number of layers	84
Number of superlayers	9
Stereo angle for each superlayer	0°, +3°, 0°, -3°, 0°, +3°, 0°, -3°, 0°
Number of super cells per layer	30, 42, 48, 60, 72, 84, 96, 108, 120
Number of sense wires per cell	12, 6, 12, 6, 12, 6, 12, 6, 12
Sense wire spacing	0.1 cm
Tilt angle	45°
Radius at innermost sense wire	30.9 cm, inner diameter = 0.55 m
Radius at outermost sense wire	132.0 cm, outer diameter = 2.76 m
Wire length	321.4 cm
Electric Field	1350 V/cm
Magnetic Field	14.1 T
Nominal r-ϕ spatial resolution	200 μm
Nominal r-z spatial resolution	4 mm
Nominal $\delta P_T/P_T$	.002 $\times$ P_T(GeV/c)
Nominal $\delta P_T/P_T$(beam constrained)	.0011 $\times$ P_T(GeV/c)

Table 2.3: CTC mechanical parameters.

momentum and position in the calorimeter, and also to identify energy directed at cracks and holes in the calorimeter. A time-over-threshold circuit placed on the CTC sense wires provides dE/dx information. This dE/dx information is utilized in the low-P_T electron identification to discriminate against charged hadrons.

2.3 Calorimetry

Calorimetry coverage at CDF extends from $\sim 2°$ to $\sim 178°$ in θ, the polar angle, and 2π in ϕ, the azimuthal angle. The calorimeters are segmented into projective towers in η and ϕ, pointing to the nominal event vertex. CDF calorimetry is divided into three segments based on pseudorapidity coverage: the Central calorimeter provides coverage in the range $\mid \eta \mid < 1.1$, the Plug in the region $1.1 <\mid \eta \mid < 2.4$ and the Forward extends this coverage to $2.2 <\mid \eta \mid < 4.2$. Each calorimeter consists of an electromagnetic part followed by a hadronic part. Properties of the CDF calorimeters are summarized in Table 2.4. An upgrade to CDF was the installation of a pre-shower counter (CPR), which is used to sample early development of electromagnetic showers. The CPR is important in the identification of soft electrons discussed in Chapter 4.

44
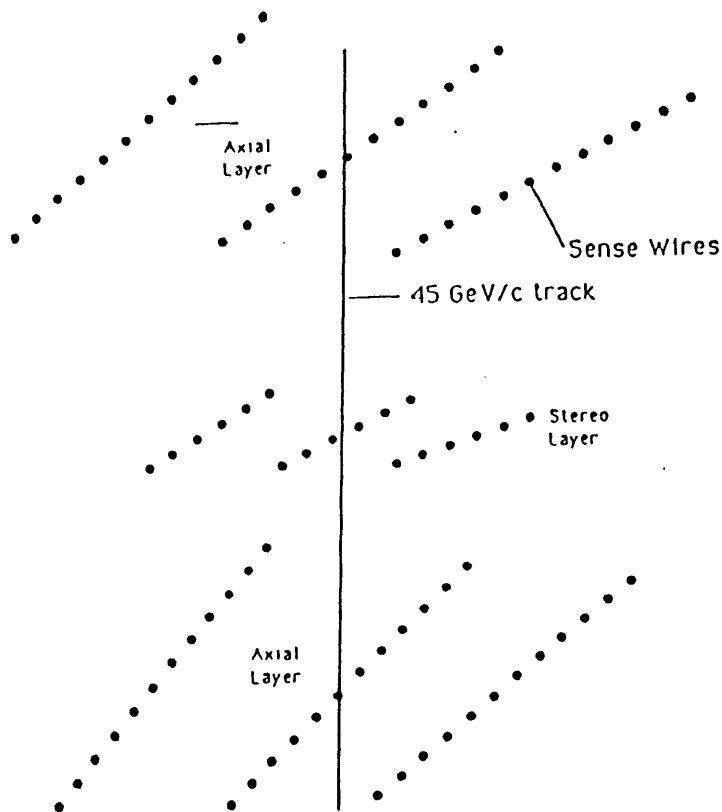
Figure 2-9: A cut away view of the CTC. Shown are two axial and one stereo super-layer together with a 45 GeV/c track through the chamber.

In this analysis we use the central calorimeter for electron identification and to some extent for muon identification as well. The Central and Plug calorimeters are used for measuring jets within $| \eta | < 2.0$. All three calorimeters are used in the determination of missing transverse energy, attributed to neutrinos from W boson decay, which is calculated as the energy imbalance in calorimeter cells out to $| \eta | < 3.6$.

Since muons are minimum ionizing particles and background from hadronic punch through is often accompanied by large amount of energy in the calorimeters, cuts on calorimeter energy deposition can reduce this background. In addition, muons from decays-in-flight (e.g $K \rightarrow \mu\nu_\mu$) are likely to have the muon direction different from

45

System	η Range	Energy Resolution $(\frac{\%}{\sqrt{E}})$+constant term	Position Resolution	Thickness
CEM	$\mid \eta \mid < 1.1$	13.5% $\oplus$ 2.0% (constant)	0.2 cm $\times$ 0.2 cm	18 X_0
PEM	$1.1 < \mid \eta \mid < 2.4$	28% $\oplus$ 2% (constant)	0.2 cm $\times$ 0.2 cm	18-21 X_0
FEM	$2.2 < \mid \eta \mid < 4.2$	25% $\oplus$ 2% (constant)	0.1 cm $\times$ 0.4 cm	25 X_0
CHA	$\mid \eta \mid < 1.1$	75% $\oplus$ 3% (constant)	0.2 cm $\times$ 0.2 cm	4.5 Λ_{abs}
PHA	$1.1 < \mid \eta \mid < 2.4$	90% $\oplus$ 4% (constant)	0.2 cm $\times$ 0.2 cm	5.7 Λ_{abs}
FHA	$2.2 < \mid \eta \mid < 4.2$	130% $\oplus$ 4% (constant)	0.2 cm $\times$ 0.2 cm	7.7 Λ_{abs}

Table 2.4: Calorimetry at CDF. $\oplus$ means add in quadrature and X_0 and Λ_{abs} are the radiation and absorption lengths respectively.

the parent particle direction and thus are likely to leave energy in towers different from the one pointing to the muon hits. Here again, cuts on energy deposition in towers pointing to the muon hits reduce backgrounds.

2.3.1 The central pre-radiator (CPR)

Another upgrade to the CDF detector for the 1992-93 Run was the addition of a pre-radiator chamber (CPR) built to study prompt gamma production at the Tevatron [32]. The pre-shower detector is located on the inner face of the Central calorimeter and is built up of MWPCs, similar to the CES chambers, that are used to sample early development of electromagnetic showers. The CPR's usefulness in separating single charged hadron backgrounds in low$-P_T$ electron identification is exploited in this analysis. Table 2.5 contains the essential parameters of the CPR and Figure 2-10 shows a side view of one of it's chambers.

2.3.2 The Central calorimeter

The Central calorimeter uses scintillators as the active medium and a phototube readout. It is segmented into two halves in the z-plane at $\eta = 0$. Each half is further segmented into 24 wedges covering $15°$ in ϕ and the η range 0 to 1.1. Figure 2-11 shows a perspective view of a Central calorimeter wedge. Each wedge is divided into 10 towers (numbered 0 to 9, with tower 0 at the $\theta = 90°$ edge) along the z-axis. One

Material seen by traversing particle:	a) 1/32" Cu-clad FRG4 (G10-like)
	b) 1/4" Hexcell (cardboard-like)
	c) 1/32" Cu-clad FR4
	d) 7/16" 95/5 ArCo2 Gas Volume
	e) Repeat of a)-c)
Total Depth	1.125"
Total Width	14.6875" (gas volume = 14")
Total Length	45.75"
Ground Field wires	33
Sense wires	32 at +1790V
Typical current draw of sense wires	35nA
Wire spacing	7/32"
Cell size after ganging	2832"
Readout channels	32 per wedge, total of 1536
Chamber gain from Fe55	32K
Gas inlets	1" from chamber corner, add 1" to depth
Mass	G10 = 1.6% X_0
	Cu = 2.1% X_0
	Total = 3.7% X_0
Source Capacitance	960 pF

Table 2.5: CPR mechanical parameters.

wedge, called the chimney, is notched to allow access to the coil, and consequently has only 8 towers.

The electromagnetic part (CEM) uses lead sheets interspersed with polysterene scintillator. It consists of 31 layers of 0.5 cm thick scintillator interspersed with 30 sheets of 0.32 cm thick lead absorber. It is 18 radiation lengths deep and also presents 1 absorption length to hadrons. At approximately 6 radiation lengths (shower maximum for electromagnetic showers) inside the CEM lie proportional chambers, the CES. The CES consists of strips perpendicular to the beam and wires along beam direction. Figure 2-12 shows the structure of the CES. The CES measures the shower position and transverse extent in z and $R - \phi$ with a position resolution of 0.2 cm by 0.2 cm. The shower centroid measured in the CES is then matched with tracks from the CTC to select electron candidates and is crucial to electron identification at CDF. The energy resolution of the CEM is $13.5\%/\sqrt{E} + 2.0\%$. It was calibrated with a 50 GeV/c electron beam and uses Cesium sources to monitor long term gain variations with an accuracy of better than 2%.
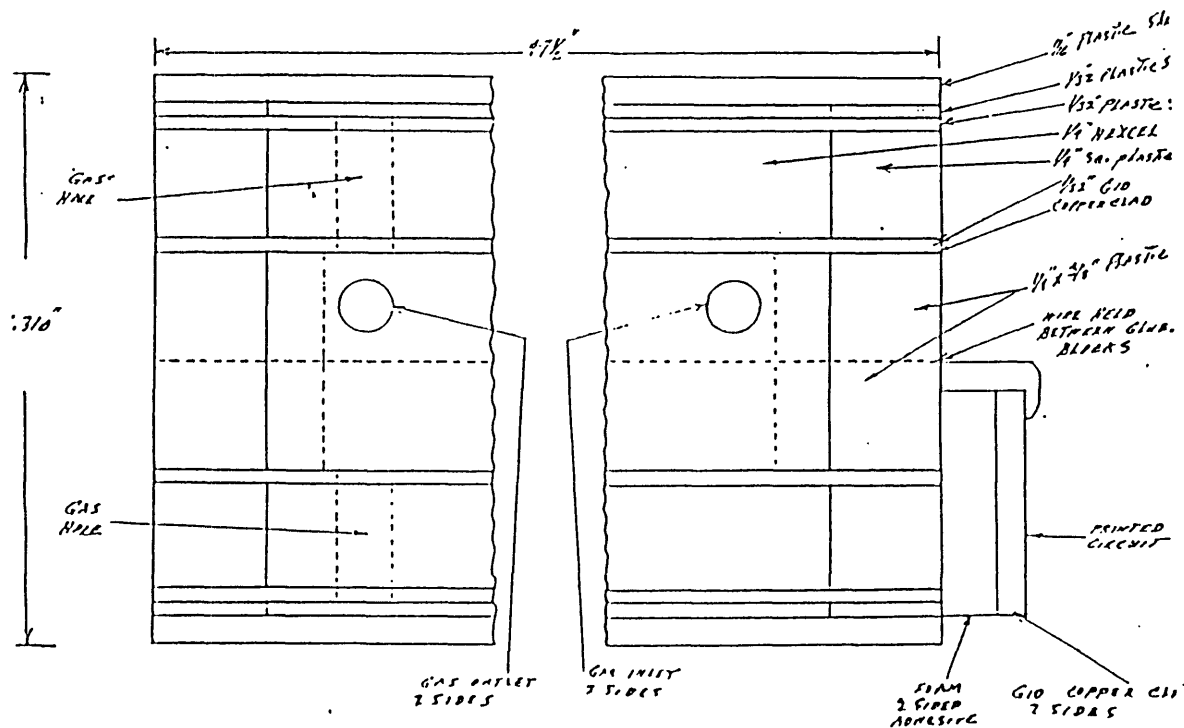
Figure 2-10: Side view of a CPR chamber.

The hadronic part of the central calorimeter (CHA) is mounted around the solenoid and consists of steel plates and acrylic scintillator. It has the same tower segmentation as its electromagnetic counterpart. The CHA covers the pseudorapidity range $|\eta| < 0.9$, with the endwall hadron calorimeter (WHA) extending this coverage out to $|\eta| < 1.3$. The hadron calorimeter has the same segmentation as its electromagnetic counterpart. The CHA is made up of 32 layers of 1 cm thick scintillator interleaved with layers of 2.5 cm thick steel. The WHA is made up of 15 layers of 5 cm thick steel followed by 1 cm thick scintillator. It presents 4.5 absorption lengths and has an energy resolution of $75\%/\sqrt{E} + 3\%$ for isolated pions.

Arrival times of signals from the CHA and WHA phototubes are measured and used to reject out-of-time backgrounds such as cosmic rays.

48
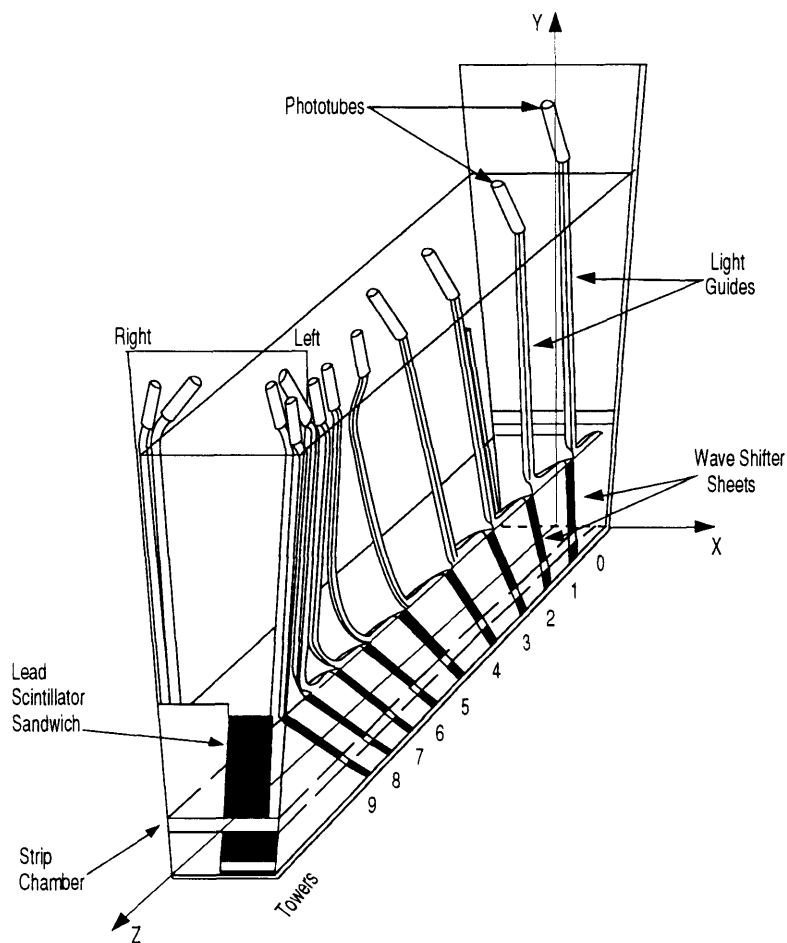
Figure 2-11: Structure of a central calorimeter wedge.

2.3.3 The Plug calorimeter

The Plug calorimeters consist of gas proportional chambers with cathode pad read-out. The electromagnetic part (PEM) uses lead absorber and conductive plastic proportional tube arrays. It presents 18-21 radiation lengths (1 absorption length). It is segmented into towers of 0.09 in η by 5° in ϕ. It has an energy resolution of $28\%/\sqrt{E} + 2.0\%$. Shower position is obtained from strips with a resolution of 0.2 cm by 0.2 cm. The PEM is divided into 4 quadrants. A gas mixture of 50-50 Argone-Ethane together with alcohol quencher is used.

The hadronic portion of the Plug calorimeter (PHA) consists of steel plates and conductive plastic proportional tube arrays. It is 5.7 absorption lengths thick, has the same tower segmentation as the PEM and has an energy resolution of $130\%/\sqrt{E} + 4\%$

49

Strip Spacing = **1.67 cm in Towers 0–4**
2.01 cm in Towers 5–9

Wire Spacing = **1.45 cm Throughout**
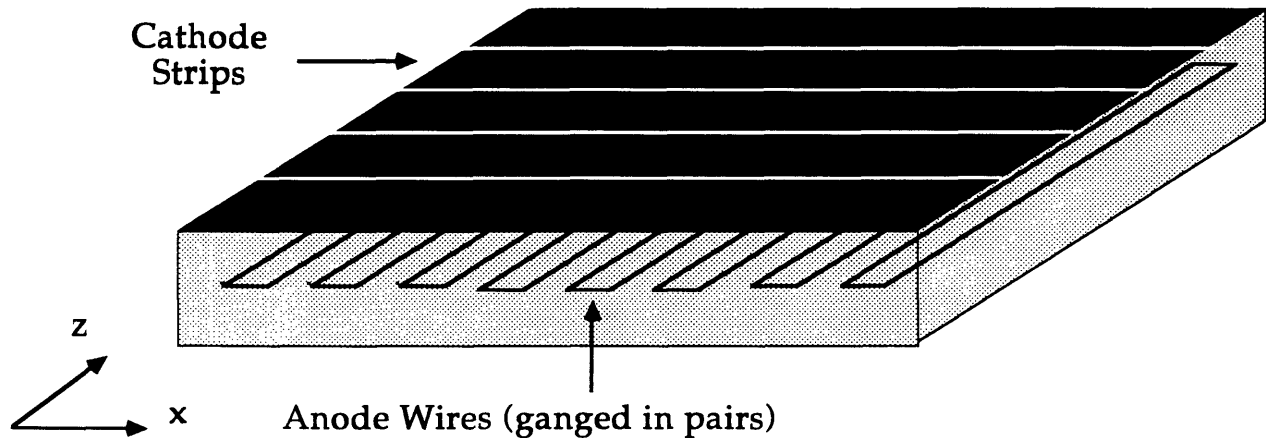
Figure 2-12: The central Strip chambers (CES).

for isolated pions.

In this analysis, the Plug calorimeter is used only for jet measurements and missing transverse energy calculation.

2.3.4 The Forward calorimeter

The Forward calorimeter at CDF is also based on gas proportional chambers with cathode pad readout. The electromagnetic component (FEM) uses lead absorber and conductive plastic proportional tube arrays. It is 25.5 radiation lengths deep (1 absorption length), has a tower segmentation of 0.1 in η by 5° in ϕ and has an energy resolution of $25\%/\sqrt{E} + 2\%$. The θ and ϕ strips measure the shower position with a resolution of 0.1 cm to 0.4 cm depending on the location in the calorimeter. The same gas mixture as for the Plug calorimeter flows through the Forward calorimeter.

The Forward hadronic calorimeter (FHA) consists of steel plates and conductive plastic proportional tube arrays. It is 7.7 absorption lengths long, has the same

tower segmentation as the FEM and has a measured resolution of $130\%/\sqrt{E} + 4\%$ for isolated pions.

2.4 The muon chambers

Muon identification capability in CDF has undergone improvements in two important directions. More steel has been added to reduce "punch through" background and the coverage has been extended.

Muons are as important to this analysis as electrons. It is important to be able to trigger on prompt muons with high efficiency and high background rejection. The muon chambers at CDF are designed with this in mind.

The muon chambers at CDF are divided into three regions: the Central muon system (CMU), the Central muon upgrade (CMP) and the Central muon extension (CMX). The CMP and CMX are upgrades to the CDF detector for the 1992-93 Run.

2.4.1 The central muon system (CMU)

The Central muon system provides muon detection up to $\mid \eta \mid < 0.63$. A muon/hadron must traverse the 4.9 absorption lengths of the central calorimeter before it encounters the CMU, 3.47 m away from the beam line. There is a CMU chamber behind the central hadron calorimeter in each wedge (see Figure 2-13). Each chamber subtends $12.61°$ in azimuth, leaving $\sim 2.4°$ gaps between calorimeter wedges. Thus the total coverage in ϕ is about 85%. The chambers extend to 226.0 cm in z on either side of the nominal interaction point. Due to the thickness of the calorimeters, only muons above $P_T \sim 1.4$ GeV/c make it to the muon chambers.

Each muon chamber is divided into three modules, each of which in turn consists of four layers of four rectangular drift cells parallel to the beam axis (see Figure 2-14). There is a sense wire at the center of each cell. Particles traversing the muon chamber radially will pass through all four cells. To resolve the left-right ambiguity in track azimuth measurement, the sense wires in the outer two cells are offset by 2 mm with respect to the the two inner wires. Sense wires in alternate cells of the same layer are
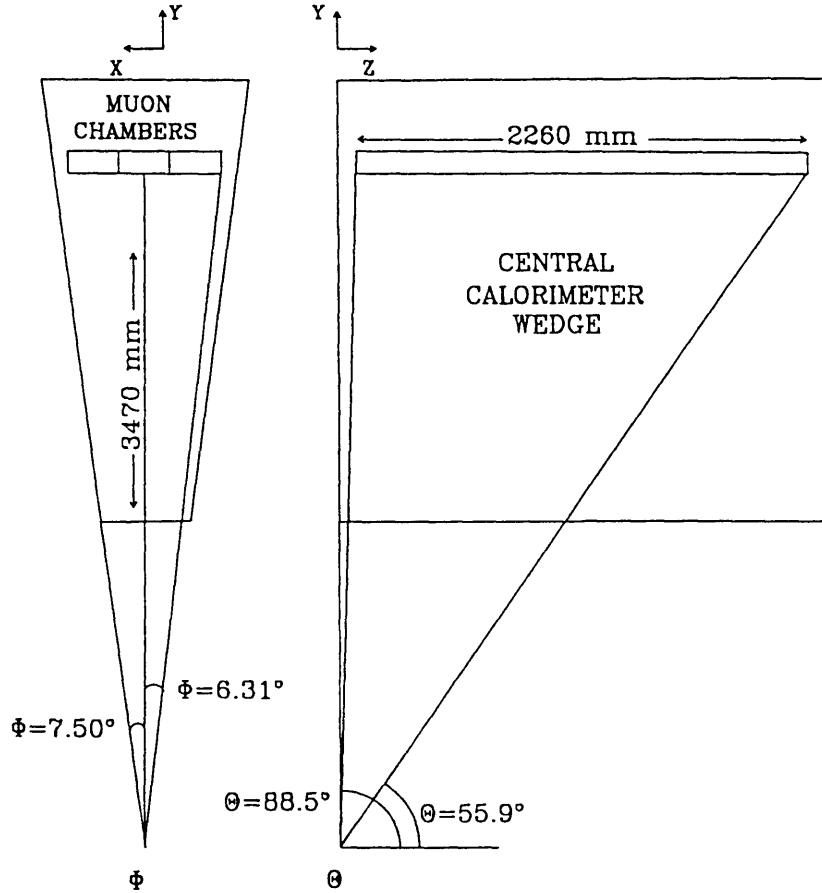
Figure 2-13: The central muon system (CMU) layout.

connected at the $\phi = 90°$ end of the chamber and are read out separately at the other end. The z position of a track is obtained by charge division. The angle α between a trajectory in a muon chamber and a reference plane containing the beam axis is related to the track's P_T by:

$$\text{Sin}\alpha = \frac{qBL^2}{2DP_T} \quad (2.3)$$

where q is the charge of the particle, B the magnitude of the magnetic field, L is the distance of the chamber from the beam line, D is the curvature of the track and P_T its transverse momentum. The angle α is determined by measuring the difference in arrival times of drift electrons at the four sense wires crossed by a given muon, i.e the greater of $t_4 - t_2$ and $t_3 - t_1$ (see Figure 2-14). The muon P_T is then inferred using

52

the above formula and used in the trigger. The momentum resolution is dominated by multiple Coulomb scattering of the muon in the calorimeter steel.
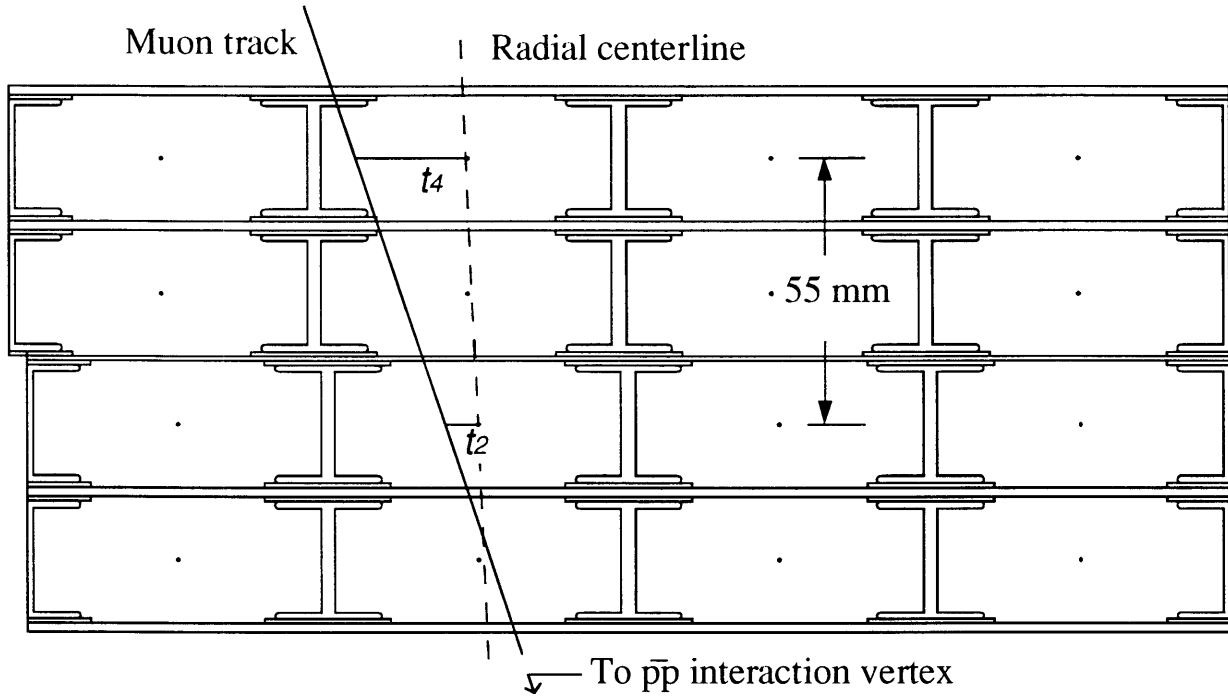
Figure 2-14: Cross-section of a CMU module showing the position of the sense wires.

2.4.2 The central muon upgrade (CMP)

The 5 absorption lengths of the central calorimeter material permit some hadrons to "punch through" and leave hits in the muon chambers. To minimize this particular background, the CMP was conceived and built. The CMP chambers form a four sided box around the CDF detector (see Figure 2-15). On two sides of the detector are steel walls $\sim$ 535 cm from the interaction point. The CMP provides additional muon chambers behind 0.6 m of additional steel behind the CMU. Thus 8 absorption lengths have to be penetrated by a hadron/muon before it hits the CMP chambers. Consequently muons with P_T less than $\sim$ 2.5 GeV/c cannot make it to the CMP. This is an important constraint in our low P_T muon search as is elaborated on in Chapter 4.

53

Design parameters of the CMP are tabulated in Table 2.6. Requiring hits in the CMP in addition to hits in the CMU increases the purity of real muons. The CMP covers an η range of upto $\sim \pm 0.6$ and has $\sim 80\%$ coverage in azimuth.

Pseudo-rapidity coverage	$\mid \eta \mid \leq 0.7$
Cell size	5 cm $\times$ 15 cm
Max drift distance	7.5 cm
Max drift time	1.4 μm
Number of X chambers (286 cm long)	344
Number of Y chambers (511 cm long)	344
Number of Z chambers (640 cm long)	424
Total number of channels	1112
Total additional iron	630 tons
Pion interaction lengths at $\theta = 90°$	7.8
Minimum detectable muon P_T	1.8 GeV/c
Multiple scattering at $\theta = 90°$	15 cm/P

Table 2.6: CMP mechanical parameters.

2.4.3   The central muon extension (CMX)

The second thrust of the muon upgrade in CDF prior to the 1992-93 Run was extending the muon coverage. The central muon extension (CMX) was built and installed to alleviate this problem. The CMX consists of 4 free-standing arches and extends muon coverage from 0.6 to 1.0 in pseudorapidity and has 80% coverage in ϕ. The chambers are located behind the 6 absorption lengths of the calorimeters. During the early part of the Run the CMX trigger was inefficient due to noise hits from particles scattering off the beam pipe and Forward calorimeters. The CMX trigger is discussed fully in a later chapter. Design parameters of the CMX are summarized in Table 2.7.

2.5   Trigger

Collisions take place at an interval of 3.5 μs resulting in a staggering event rate of 300,000 Hz. However events can only be written to tape at a few hertz ($\sim$8 Hz). To achieve this drastic reduction in event rate and still remain efficient to events of interest, CDF employs a 3-tiered trigger system. This trigger system employs
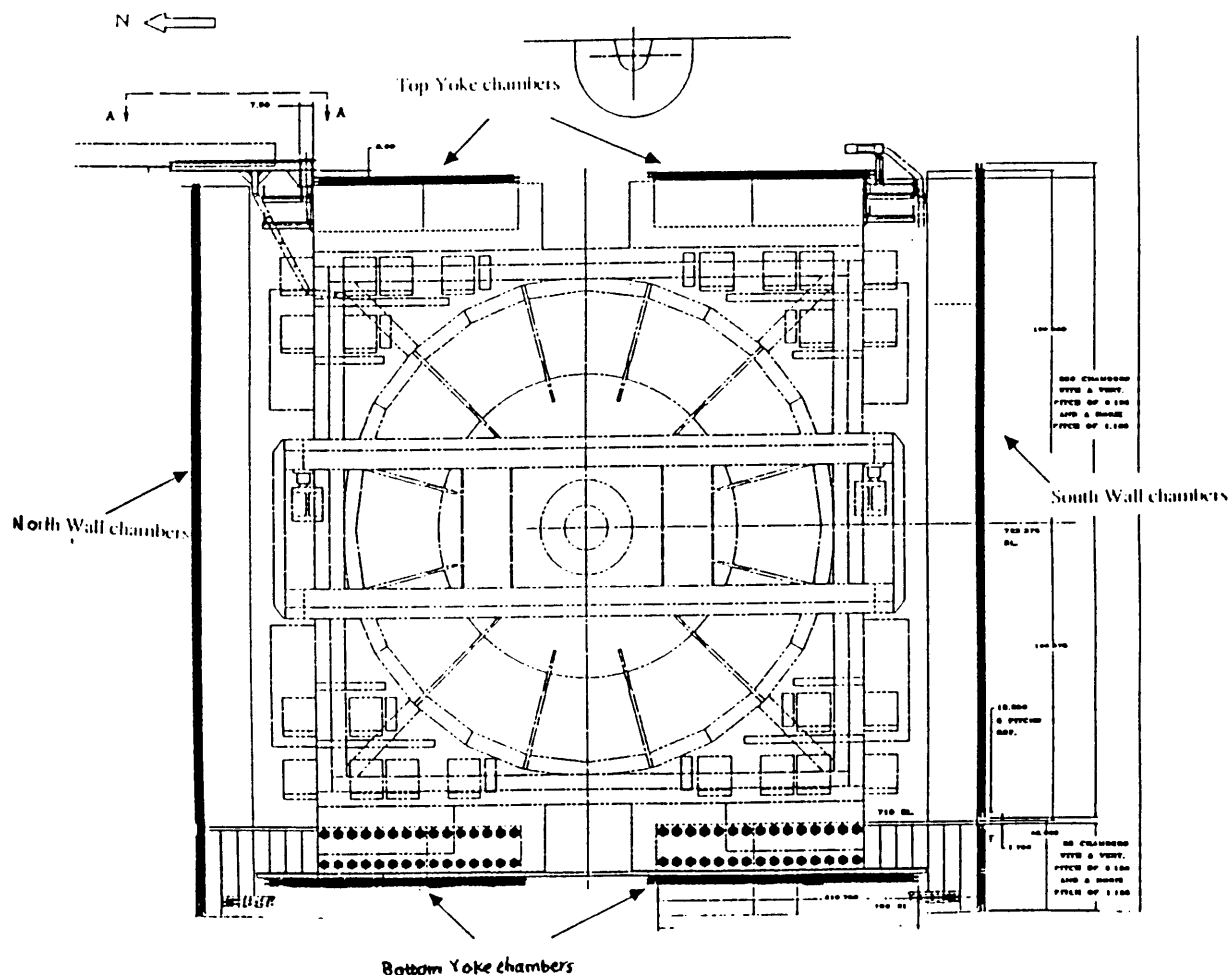
Figure 2-15: The central muon upgrade (CMP) chambers.

both hardware and software decisions that are designed to minimize deadtime and be efficient and robust. Ascension in the trigger level implies higher complexity.

All detector elements at CDF have front-end electronics which send analog signals that are digitized and stored as a CDF event. The Front-end electronics includes FASTBUS TDCs (time-to-digital converters) to readout fast tracking and prompt muon data to be used in the trigger decision, and RABBIT cards for the calorimeter and muon drift chamber readout . These front end cards are connected to SSPs and MX scanners which provide some rudimentary formatting. These digital signals are then passed to a hardware event builder which builds the event structure at a rate of approximately 30 Hz (average size of event is 200 KBytes). Events that ultimately pass the Level 3 trigger (described below) are logged onto 8 mm tapes at a rate of

Pseudo-rapidity coverage	$0.7 \leq \mid \eta \mid \leq 1.0$
Cell size	3 cm × 15 cm
Max drift distance	7.5 cm
Max drift time	$1.5\mu m$
Total number of channels	2304
Pion interaction lengths at $\theta = 55°$	6.2
Minimum detectable muon P_T	1.4 GeV/c
Multiple scattering at $\theta = 55°$	13 cm/P

Table 2.7: CMX mechanical parameters.

~8 Hz. A schematic view of the CDF trigger and data acquisition system is shown in Figure 2-16.
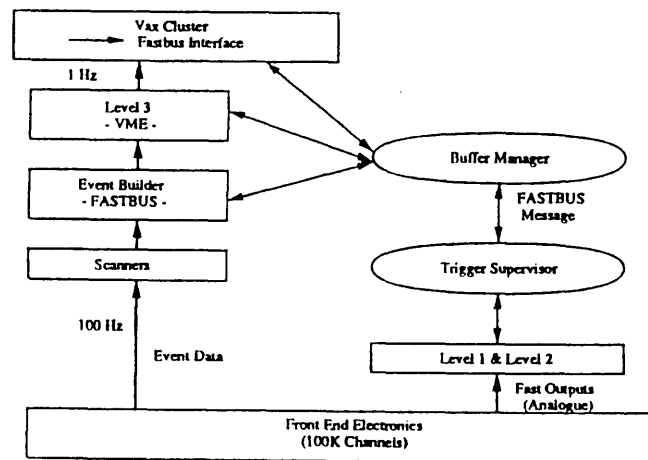
Figure 2-16: A schematic view of the CDF trigger and data acquisition system.

Each level of the trigger consists of a logical OR of a number of triggers designed to select events with physics objects of interest: electrons, muons, jets or significant missing transverse energy (neutrinos).

2.5.1 Level 1

This is the first step in reducing the 300 kHz event rate to a more manageable one. The Level 1 trigger essentially requires either a large amount of energy in a calorimeter tower or a large P_T inferred from Fastouts from the muon chambers. Typical single trigger tower thresholds for Level 1 triggers are shown below. It should be noted that trigger towers in the calorimeter are twice the size of conventional trigger towers described earlier: $\Delta\phi \times \Delta\eta = 15° \times 0.2$.

Detector	Energy threshold in GeV
CEM	6
CHA	8
PEM	8
PHA	25
FEM	8
FHA	25

Table 2.8: Typical trigger tower thresholds.

Provisions exist for passing events with two or more trigger towers above a lower threshold to allow efficient triggering of dilepton events. The output from Level 1 is roughly 2.4 kHz.

2.5.2 Level 2

Level 2 attempts to further decrease the 2.4 kHz rate by adding a level of sophistication to the events passing Level 1 triggers. Instead of single trigger tower thresholds, a nearest-neighbor hardware cluster finder is used to find clusters of energy. For each cluster, the E_T, $< \phi >$, and $< \eta >$ are determined. These clusters are then matched with tracks from the a hardware track processor that uses fast timing information from the CTC as input (CFT). The CFT resolution is $\delta P_T/P_T = 0.035 \times P_T$(GeV/c). The CFT efficiency for isolated tracks with $P_T > 10$ GeV/c has been measured to be $93.5 \pm 0.3\%$.

In the case of electrons and photons, the cluster energy is required to be predominantly electromagnetic. Muon triggers require a match between the "stubs" (track

segments) in the muon chambers (CMU, CMP and CMX) and tracks from the CFT. Requirements are also made on missing transverse energy and other neural network algorithms. The event rate out of Level 2 is ~24 Hz.

2.5.3 Level 3

Here a full blown event reconstruction is performed on a farm of dedicated CPUs. In practice, Silicon Graphics Multi-CPU Power Servers running Fortran code are used. Most of the CPU time is spent on the three dimensional track reconstruction for the CTC. Physics quantities are defined as they would be Offline and tight identification cuts are imposed. Clustering algorithms are employed to measure energy in jets and electromagnetic clusters; CTC tracking is performed and matching between CTC tracks and energy clusters is imposed in the case of electrons. For muons, CTC tracks are matched to stubs in the muon chambers and tight cuts are applied. Isolation requirements are also imposed where applicable.

Events that pass Level 3 are then written out to tape at a rate of about 8 Hz. There are 2 output streams: Stream 1, which writes out events passed by all Level 3 triggers, and Stream 2, which consists of a subset of prioritized triggers (e.g. top) and has roughly 10% the bandwidth of Stream 1. Stream 2 events are immediately processed and available for analysis. They also provide the basis for Online monitoring of data quality. The analysis described here uses Stream 2 events exclusively. However, Stream 1 events were used for other studies such as tracking efficiency determination.

2.6 Luminosity measurement

CDF uses sets of scintillator planes, known as beam-beam counters (BBCs), located 5.8 m from the nominal interaction point on either side of the detector. The BBCs cover an angular range of 0.32° to 4.47° ($3.24 <| \eta |< 5.90$). They have excellent timing properties ($\sigma < 200$ ps) and provide the measurement of the time of interaction.

Chapter 3

Event selection

The event signature of our top search comprises of a high P_T lepton, considerable missing transverse energy ($\not{E}_T$) and several jets. We then look for additional leptons that might arise from direct and sequential b quark decays. Event selection begins with the making of inclusive electron and muon samples. The creation of these samples is described in this chapter, together with a short description of the calculation and meaning of missing transverse energy and jet clustering algorithms applied at CDF. Section 3.1 discusses the high energy inclusive electron sample, Section 3.2 the corresponding muon case. Section 3.3 explains the calculation and meaning of missing transverse energy and we end the chapter with a brief discussion of jet clustering schemes employed at CDF in section 3.4. The next chapter deals with the identification of extra leptons in these events, the so-called 'soft leptons.'

3.1 Inclusive electron sample

We only consider electrons detected by the Central calorimeter ($| \eta | < 1.1$), Central electrons, in this search. We describe the trigger path, the Offline identification parameters and their cut values, and the efficiency of our electron selection scheme.

3.1.1 Trigger

As mentioned in an earlier chapter, the Level 1 CEM trigger requires a trigger cell with at least 6 GeV of transverse energy. Using the geometric center of the detector as the event vertex, angles to all the trigger cells are determined and used in computing the transverse energy deposited in these cells. Level 2 then clusters these trigger cells via a hardware processor. The clustering scheme first identifies all trigger cells with electromagnetic energy above 9 GeV, dubbed trigger seeds. The four adjacent cells to each trigger seed are then added to form a cluster if their transverse electromagnetic energy is above 7 GeV. This clustering is repeated until no more cells can be added to the cluster. The total cluster transverse electromagnetic energy is then simply the sum over all cells in the cluster. Level 2 further requires a match in the azimuthal angle with a stiff track from the Central Fast Tracker (CFT). The Level 2 trigger then cuts on the following cluster parameters:

- $E_T >$ 9 GeV

- $\frac{HAD}{EM} < 0.125$, here HAD and EM are the energy depositions in the hadronic and electromagnetic compartments of the calorimeter respectively

- cluster matched to CFT track with $P_T > 9.2$ GeV/c

The Level 3 trigger employs the same reconstruction algorithm as the Offline production package. Here clustering is done in towers, finer segmentation than trigger cells. Tracks are reconstructed with a momentum resolution of $\delta(1/P_T) = 0.007 (\mathrm{GeV}/c)^{-1}$. Level 3 cuts are summarized below.

- $E_T{}^{EM} > 18$ GeV

- $\frac{HAD}{EM} < 0.125$

- Require CTC track with $P_T > 13$ GeV/c

Here $E_T{}^{EM}$ is the tranverse electromagnetic energy of the cluster, HAD and EM the total hadronic and electromagnetic energies associated with the cluster respectively

and P_T the transverse momentum of the track matched to the electromagnetic cluster. The efficiency of the 9 GeV trigger is measured by looking at lower threshold triggers (6 GeV) and also by looking at W and Z boson events selected with independent triggers (e.g $\not{E}_T$) [33]. The efficiency for these trigger requirements is measured to be $92.8 \pm 0.2\%$ for electrons with $20 < E_T < 150$ GeV.

3.1.2 Offline electron identification parameters

Offline reconstruction looks for clusters of electromagnetic energy and then attempts to distinguish electron-like objects from hadrons and photons by applying some quality cuts. The clustering scheme together with the parameters used Offline are now discussed.

Clustering scheme

Clustering begins with seeding towers above 3 GeV of transverse electromagnetic energy, $E_T \overset{\text{def}}{=} E \sin\theta$, where E is the energy deposisted in the tower and θ is the polar angle between the center of the tower and the true event vertex. The center of the tower is defined to be at shower maximum in depth and halfway in detector azimuth between the tower boundaries. Surrounding towers are then added if their E_T is greater than 0.1 GeV. Electron showers are small compared to the size of CEM towers, therefore clusters consist of only 1 or 2 towers. Since CEM towers are separated by approximately 1 cm of inactive dense material in azimuth, showers far from azimuthal boundaries[1] will not leak across boundaries. Consequently CEM clusters are restricted to 3 or fewer towers in the same wedge i.e three towers in η and one tower in ϕ.

All clusters formed in the aforementioned manner are required to have electromagnetic energy greater than 5.0 GeV and $HAD/EM < 0.125$. The candidate clusters having survived thus far are further subjected to stringent requirements to reject hadronic backgrounds and distinguish electrons from photons. We first describe the

---

[1]Full response is obtained at the center of tower, away from edges.

parameters that quantify the likelihood of these candidates being electrons, then summarize their cut values.

Energy

The raw energy in all the towers in the cluster is summed up and three corrections are applied to the sum:

1. Within a tower, the response is corrected as a function of shower position determined by the CES. The Response map used for the correction is obtained from Testbeam electron data. An example of a response map is shown in Figure 3-1.

2. Tower-to-tower response variations are then normalized by E/P in a sample of inclusive electrons with $E_T > 15$ GeV.

3. The overall scale is then determined by comparing E/P of electrons from W decay to that of a radiative Monte Carlo simulation. In this simulation, electrons from the decay of W bosons are allowed to radiate in the detector.
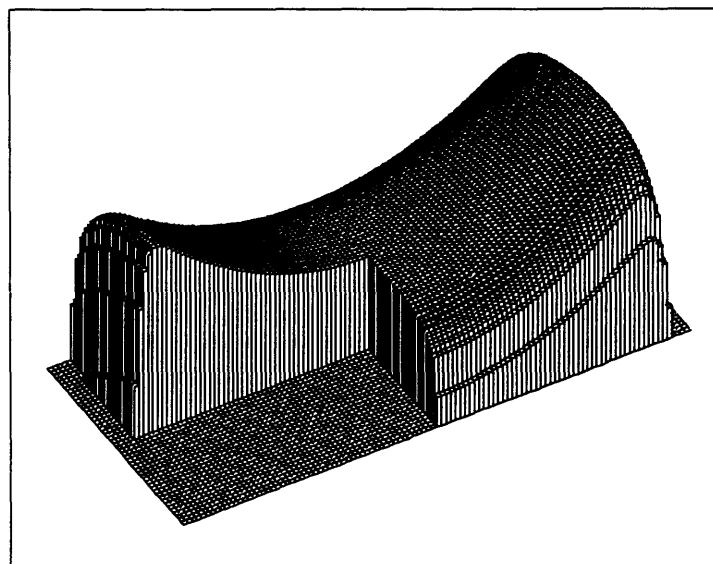
Figure 3-1: A typical CEM response map: relative response to electrons in the central calorimeter tower.

E_T is then calculated as $E_T = E \sin \theta$, where E is the corrected calorimeter energy and θ is the polar angle of the beam constrained track.

Track momentum

To distinguish electrons from photons, a three dimensional CTC track pointing to the calorimetry cluster is required. For high energy electrons, such as the ones from W decay, the momentum resolution in the Central tracking chamber is worse than the energy resolution of the CEM. Consequently we only use the track to yield the direction of the electron-like object.

Strip chamber variables

The CES is used to measure the shower location (center) and to quantify 'cleanliness' of electron candidates. The strip and wire views yield z and x information respectively.

The shower profiles in each view are fitted to 50 GeV Testbeam data [34]. In the case of strips, the z-coordinate of the shower center, Z_{CES}, and the cluster energy E_S are obtained by minimizing

$$\chi^2(z, E) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \frac{(E_i^{meas} - E \cdot q_i^{pred}(z))^2}{\sigma_i^2(z)} \tag{3.1}$$

where $n{=}11$ is the number of channels, E_i^{meas} is the measured channel energy and $q_i^{pred}(z)$ are predicted energies normalized to 1 and corresponding to a given z-coordinate of the shower center. Fluctuations in a single channel response are given by $\sigma_i^2(z) = (0.026)^2 + (0.096)^2 q_i^{pred}(z)$. This equation is obtained from 10 GeV/c Testbeam electron data.

A single electron/photon hypothesis is then tested using

$$\chi_{strips}^2 \stackrel{\text{def}}{=} \frac{1}{4} \left(\frac{E_{CEM}}{10} \right)^{0.747} \sum_{i=1}^{n} \frac{(q_i^{meas} - q_i^{pred}(Z_{CES}))^2}{\sigma_i^2(Z_{CES})} \tag{3.2}$$

where $\{q_i^{meas}\}_{i=1}^{n}$ is the measured strip profile normalized to 1.

The wires are treated analogously and the x-coordinate, X_{CES}, of the shower center and χ_{wires}^2 are obtained. These χ^2s are then cut upon.

Finally, matching between the shower position in the CES and tracks extrapolated

63

from the CTC is imposed:

$$\Delta X = X_{extrap}^{CTC} - X_{CES} \tag{3.3}$$

$$\Delta Z = Z_{extrap}^{CTC} - Z_{CES} \tag{3.4}$$

Here X_{extrap}^{CTC} is the $x-$coordinate at the CES of the extrapolated CTC track and Z_{extrap}^{CTC} the analogous $z-$coordinate. These variables help reject fake electrons caused by a charged pion track which overlaps with a neutral pion showering in the electromagnetic calorimeter.

Ratio of hadronic and electromagnetic energies

The CDF calorimeter's ability to measure the hadronic leakage of showers is exploited in rejecting QCD backgrounds from 'overlap' events involving charged hadrons accompanied by extra photons. The variable HAD/EM, where HAD is the hadronic energy associated with a cluster and EM its counterpart in the electromagnetic compartment of the calorimeter, provides a powerful discriminant between charged hadrons and electrons.

Transverse shower development

The transverse shower development of electromagnetic showers is monitored through the variable L_{shr} defined by

$$L_{shr} \overset{\text{def}}{=} 0.14 \sum_{i} \frac{E_i^{dep} - E_i^{exp}}{\sqrt{(\Delta E)^2 + (\Delta E_i^{exp})^2}} \tag{3.5}$$

Here the sum is over the two towers adjacent to the seed tower in the same wedge. The quantity E_i^{dep} is the energy deposited in tower i, E_i^{exp} is the energy expected in tower i. Expected energy is calculated from profiles of transverse shower development at the Testbeam. It depends on the seed energy and the direction of shower impact point in strip chamber relative to the event vertex. The denominator normalizes the expression taking into account the finite resolution of energy measurement: $\Delta E = 0.14\sqrt{E}$ is the uncertainty on cluster energy E, and ΔE_i^{exp} is the error in E_i^{exp} associated with

a 1 cm error in shower impact point measurement.

E/P

This parameter is used to verify matching between the CEM and CTC measurements of electron energy, E and P respectively. High energy electrons radiate in the detector. The CTC only measures charged particle track momenta whereas the calorimetry captures most of the radiated energy, consequently the mean value of E/P for tracks associated with electrons is slightly above 1 (see Figure 3-2).
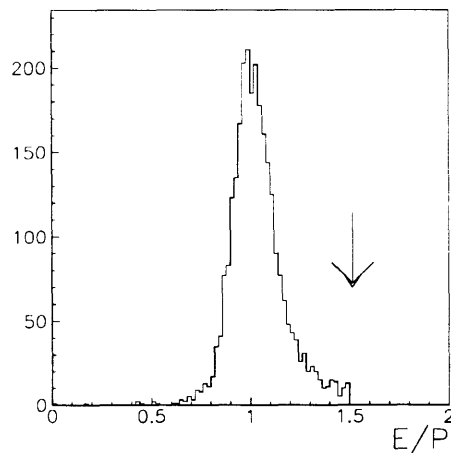
Figure 3-2: E/P for W electrons showing the radiative tail. The arrow indicates the cut value used.

Isolation

This variable detects the presence of energetic particles near an electron. Since electrons from W boson decay are expected to be highly isolated and backgrounds are not, this variable has high background rejection capability. Two isolation variables are commonly used at CDF. One measures the excess energy in a fixed cone around the candidate cluster, $Isol(R)$,

$$Isol(R) \stackrel{\text{def}}{=} \frac{E_T^R - E_T^{cluster}}{E_T^{cluster}} \tag{3.6}$$

65

where E_T^R is the total transverse energy, hadronic plus electromagnetic, in a cone centered on the electron cluster and with radius $R \stackrel{\text{def}}{=} \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$; $E_T^{cluster}$ is the uncorrected electron cluster transverse energy. A cone size of $R = 0.4$ is used in this analysis.

Another commonly used isolation variable at CDF is Border Tower Energy, BTE. It is simply the sum of total, electromagnetic as well as hadronic, energy of all towers adjacent to an electron cluster. We do not use BTE in this analysis.

Fiducial cuts

Fiducial requirements are applied to avoid cracks in the detector and insure reliable energy measurement. The CEM fiducial cuts are:

1. Seed towers must not be one of the outermost (large $\mid \eta \mid$) towers of the calorimeter.

2. At the CES ($R = 184$ cm), the extrapolated CTC track must be at least 2.5 cm away from the azimuthal boundaries between the wedges.

3. Shower position in strip chambers must be at least 9 cm away from the Z=0 plane to avoid the 90° crack.

These fiducial cuts cover 84% of the solid angle in $\mid \eta \mid < 1.0$ as can be seen from Figure 3-3.

3.1.3 Cut values

Table 3.1 lists all the cuts used in the creation of the inclusive electron sample. Figure 3-4 shows all the cut parameters together with their cut values. Each plot has all other cuts applied to it except for the variable it displays.
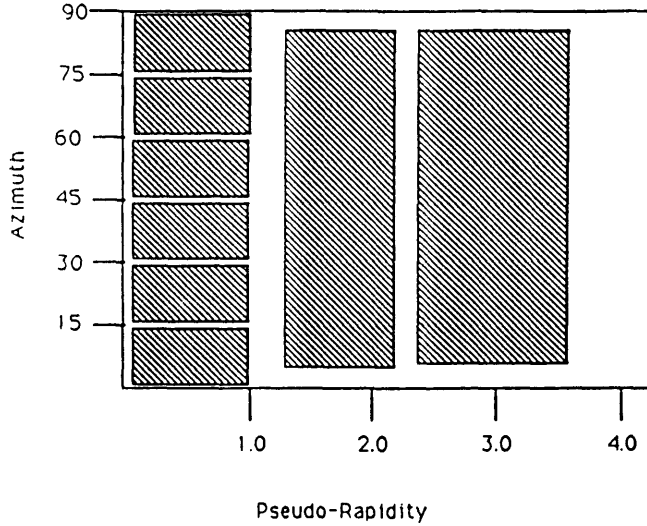
Figure 3-3: The fiducial region used in central electron identification. The shaded region shows the area covered by the fiducial cuts.

3.1.4 Electron identification efficiency

Method

For high P_T electrons, we measure the efficiency of our identification cuts from $Z^0 \rightarrow e^+e^-$ decays. The Z boson sample is selected by requiring one leg to pass all the electron identification cuts listed in Table 3.1 and the second leg to form an invariant mass with the first in a narrow window around the Z° resonance. In this manner we obtain a source of unbiased electrons from which we can determine the efficiency of each of our cuts as well as all of them combined.

The efficiency calculation is straightforward and we only summarize it here, see reference [35] for details. We define the following parameters:

N = number of $Z^0 \rightarrow e^+e^-$ candidates

N_1 = number of Z^0s for which at least one leg passes the identification cuts

N_2 = number of Z^0s for which both legs pass the cuts

N_c = number of Z^0s where one leg passes all cuts and the other cut c

ϵ = efficiency of all cuts

ϵ_c = efficiency of cut c

Parameter		Cut value
E_T	$>$	20 GeV
E/P	$<$	1.5
$HAD/EM(3 \times 3)$	$<$	0.05
$\mid \Delta X \mid$	$<$	1.5 cm
$\mid \Delta Z \mid$	$<$	3.0 cm
L_{shr}(3-tower)	$<$	0.2
χ^2(strip)	$<$	10.
$\mid z_v - z_0^e \mid$	$<$	5 cm (Vertex class $\geq$ 10)
$\mid z_v \mid$	$<$	60 cm
Fiducial cuts		Applied
Coversion removal		Applied

Table 3.1: Central electron selection criteria.

then the following relations can be seen to hold,

$$N_1 = \epsilon(2 - \epsilon)N \qquad (3.7)$$

$$N_2 = \epsilon^2 N \qquad (3.8)$$

$$N_c = \epsilon(2\epsilon_c - \epsilon)N \qquad (3.9)$$

We can solve for N, ϵ and ϵ_c from these constraints,

$$N = \frac{(N_1 + N_2)^2}{4N_2}, \quad \epsilon = \frac{2N_2}{N_1 + N_2}, \quad \epsilon_c = \frac{N_2 + N_c}{N_1 + N_2} \qquad (3.10)$$

Since N_1, N_2 and N_c are measured from the data, ϵ and ϵ_c can be extracted.

To compute the uncertainty on ϵ and ϵ_c, we consider N_1 and N_2 to be binomial variables with respective efficiencies $\epsilon_1 \stackrel{\text{def}}{=} \epsilon(2 - \epsilon)$ and $\epsilon_2 \stackrel{\text{def}}{=} \epsilon^2$, on a sample space of size N. We then have

$$\sigma_{N_1}^2 = N\epsilon_1(1 - \epsilon_1) = N\epsilon(2 - \epsilon)(1 - \epsilon)^2 \qquad (3.11)$$

$$\sigma_{N_2}^2 = N\epsilon_2(1 - \epsilon_2) = N\epsilon^2(1 - \epsilon^2) \qquad (3.12)$$

$$\sigma_{N_1 N_2}^2 = N\epsilon^2(1 - \epsilon)^2 \qquad (3.13)$$
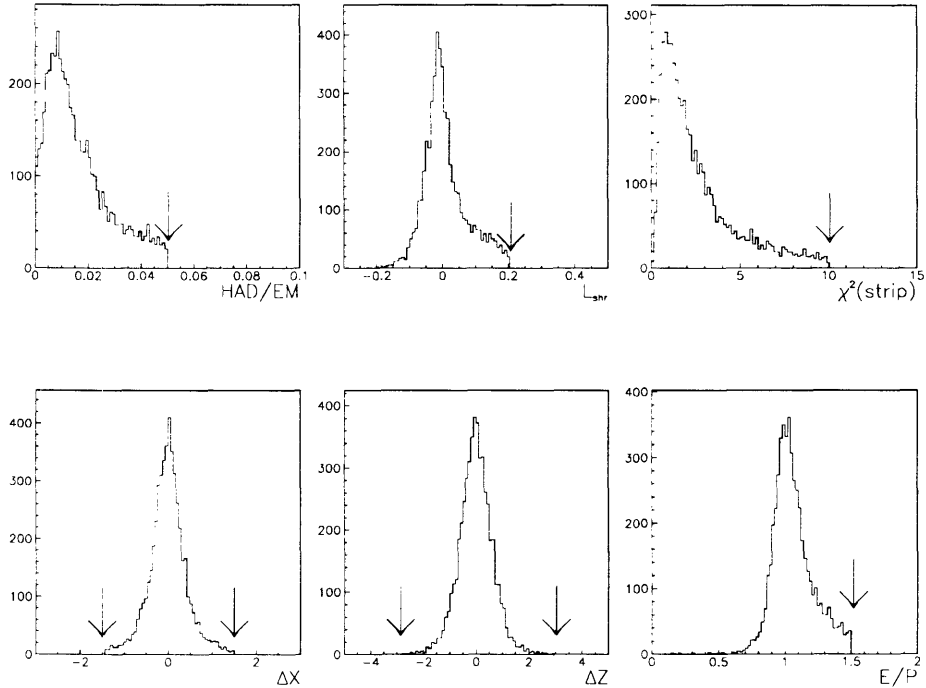
Figure 3-4: Central electron selection parameters. (a)HAD/EM, (b) L_{shr}, (c) χ^2(strip) (d) ΔX in cm, (e) ΔZ in cm, and (f) E/P. Arrows indicate cut values used in this analysis.

and propagating these uncertainties onto ϵ yields

$$\sigma_\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)(2-\epsilon)}{N_1 + N_2}} \qquad (3.14)$$

Similar reasoning yields

$$\sigma_{\epsilon_c} = \sqrt{\frac{\epsilon_c(1-\epsilon)(1+\epsilon/\epsilon_c - \epsilon)}{N_1 + N_2}} \qquad (3.15)$$

Results

We have 518 events in our Z sample (see Figure 3-5). Applying the above procedure to this dataset gives us the efficiencies listed in Table 3.2. The total efficiency in this table does not include the 5% inefficiency due to the conversion removal algorithm (discussed later) and the 13% due to isolation requirements imposed on high-P_T

69

Cut	N_c	ϵ_c
HAD/EM	285	0.969 ± 0.008
L_{shr}	293	0.985 ± 0.005
$\chi^2(\text{strips})$	275	0.950 ± 0.010
ΔX	288	0.975 ± 0.007
ΔZ	284	0.967 ± 0.008
E/P	265	0.931 ± 0.012
All cuts	217	0.84 ± 0.02

Table 3.2: Electron identification efficiencies.

electrons in the top quark search.
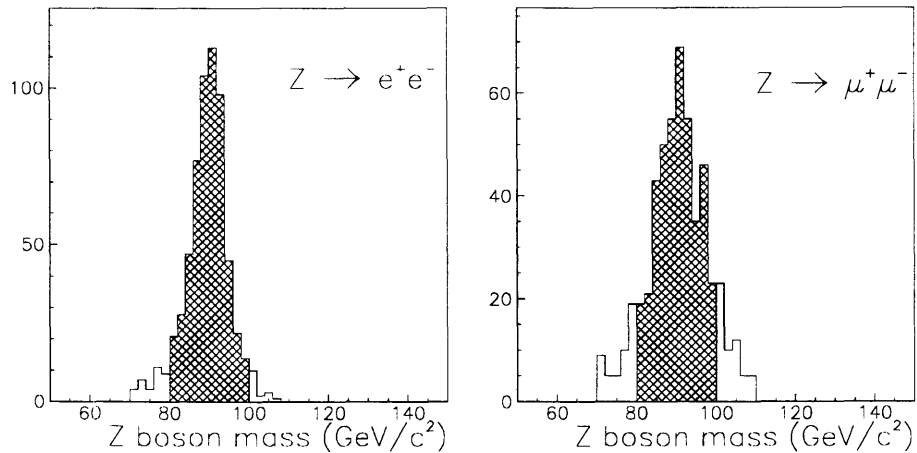
Figure 3-5: The invariant mass of (a) $Z \rightarrow e^+e^-$, and (b) $Z \rightarrow \mu^+\mu^-$ used in the efficiency determination of high$-P_T$ leptons. Events in the shaded area were used.

3.1.5 Conversion removal

A significant source of electrons in CDF are Dalitz decays of π^0s $(\pi^0 \rightarrow \gamma e^+e^-)$ and photons interacting with material in the detector to produce e^+e^- pairs. Both these processes are referred to as 'conversions.'

Conversion electrons can be identified by looking for low mass opposite-charge track pairs. In practice the following cuts are imposed on the track pair

70

1. $\mid \Delta(\cot \theta) \mid < 0.06$

2. $\mid \Delta S \mid < 0.3$ cm

3. $M_{ee} < 500$ MeV/c^2

4. $R < 50$ cm

where θ is the polar angle, ΔS is the distance of closest approach between the initial track and its conversion partner, M_{ee} is the mass of the track pair and R is the radius of conversion.

A detailed description of the conversion sample selection is given in chapter 4, where such electrons constitute a control sample for studying low-P_T electron cuts. Here it is only relevant to remove high-P_T electrons from the inclusive electron sample if they are identified as conversions. The conversion removal procedure employed is $\sim 3 - 5\%$ over-efficient. The efficiency of the conversion removal procedure has been determined to be $88 \pm 4\%$.

3.2   Inclusive muon sample

Muons are detectable in three separate detectors at CDF: the central muon chambers (CMU), the central muon upgrade (CMP) and the central muon extension (CMX). Muons can also be detected in $\eta - \phi$ space not covered by one or more of these muon chambers by looking for minimum ionizing tracks in the calorimeter (CMIOs). Because of the large backgrounds associated with CMIOs, muons in this analysis must leave hits in one or more of CDF's three muon chambers. This section details the trigger path, the Offline reconstruction parameters and their cut values as well as the efficiency of our selection scheme. Removal of Cosmic ray background is also discussed.

3.2.1 Trigger

The Level 1 CMU (CMX) trigger looks for tracks crossing a CMU (CMX) chamber leaving at least 2 hits. By requiring the time difference between hits in two alternate layers of the chamber to be lower than a threshold value, it imposes a minimum P_T cut on candidate muons. In the 1992-93 Run a cut of 6 GeV/c was used for the CMU in addition to expected hits being present in the CMP. In the CMX a P_T cut of 10 GeV/c in coincidence with hits on both sides of the chambers was imposed.

The CMU Level 1 efficiency has been determined to be $94.7^{+0.82}_{-0.98}\%$ for muons with $P_T > 15$ GeV/c [36]. The CMX inclusive muon trigger was fully functional for only $\sim 30\%$ of the run because of a rate-dependent prescale. The CMX trigger also required a Level 1 single tower threshold in addition to the CMX muon candidate in the event. This trigger was present for 83% of the total integrated luminosity of the run.

At Level 2 the presence of a central fast tracker (CFT) track of $P_T > 9.2$ GeV/c that matched in azimuth the stubs found in the muon chambers which triggered at Level 1 was required. The matching was performed after the CFT track was extrapolated to the CMU chambers taking care of any multiple scattering and possible error in CFT azimuth measurement. The Level 2 efficiency for isolated muons with $P_T > 15$ GeV/c was measured to be $93.7^{+1.27}_{-1.52}\%$ [36].

As mentioned earlier, at Level 3 full blown event reconstruction takes place. The CMU Level 3 trigger requires a reconstructed CTC track of $P_T > 18$ GeV/c that matches CMU chamber hits to within ± 10 cm in the azimuthal direction. In addition, the CHA energy of the tower traversed by the muon is required to be less than 6 GeV. The CMU Level 3 efficiency was measured to be $86.8 \pm 1.9\%$ for muons with $P_T > 20$ GeV/c . The corresponding CMX efficiency was measured to be $54.4 \pm 5.5\%$ for isolated muons with $P_T > 15$ GeV/c [37]. There were no CMP-only triggers in the Run.

3.2.2 Muon identification parameters

This section describes parameters employed to separate muons from hadrons that interact in the calorimeter.

Momentum

The momentum of a muon is the momentum of the CTC track that best matches the CMU/CMP/CMX track segment. Matching cuts are described below.

Impact parameter

To reduce backgrounds from Cosmic rays, decay-in-flight muons from kaons and pions, the CTC track is required to pass close to the event vertex. The impact parameter, b, and the z-coodinate of the track, z_{track}, are cut upon. The impact parameter is the distance of closest approach of the reconstructed track to the event vertex.

Track matching

The CTC track is extrapolated to the muon chambers (CMU/CMP/CMX) where it undergoes survey corrections. Using a local coordinate system the CTC track and muon segment are fit to straight lines in the $x-y$ and $z-y$ planes. The difference in slope and intercept in each plane are used for matching purposes. Because of the better CTC resolution in the transverse plane, only the difference in the intercept in the $x-y$ plane, ΔX, is used in this analysis.

Energy in the calorimeters

Since muons are minimum ionizing particles, requiring the energy deposited in calorimeter towers traversed by muon candidates to be significantly less than the momentum of the muons rejects interacting punch-through hadron background. Testbeam muons of $P_T = 57$ GeV/c leave on the average $EM_\mu = 0.3$ GeV and $HAD_\mu = 2$ GeV as shown in Figure 3-6. Also shown in this figure is the $EM + HAD$ energy deposited by pions of 57 GeV/c .
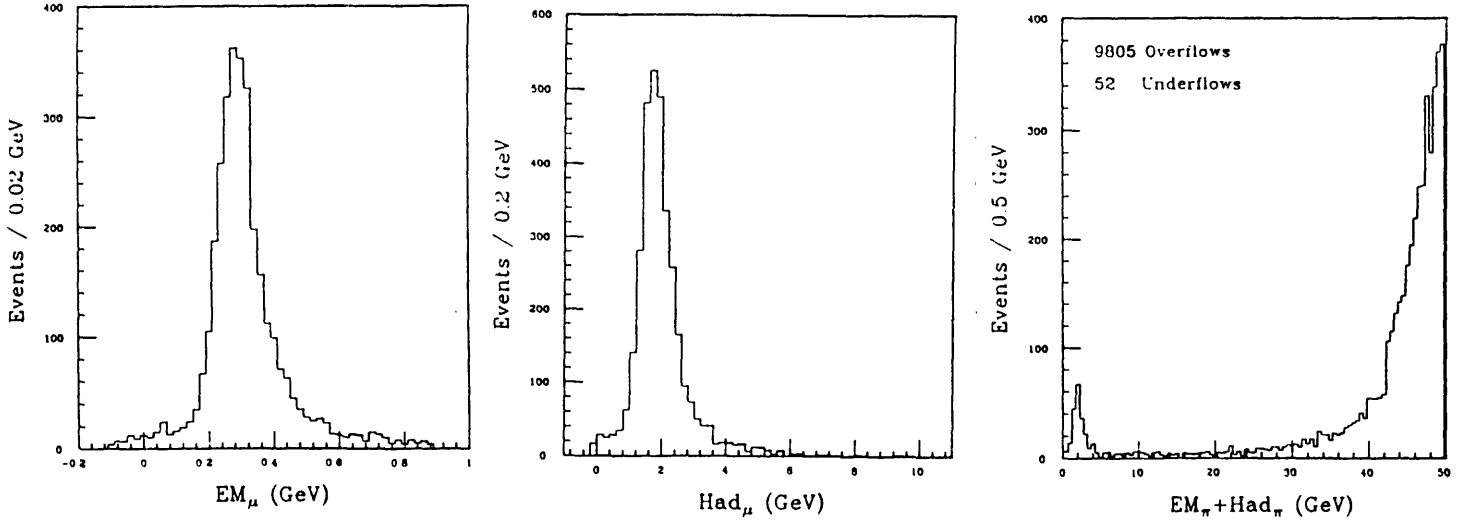
Figure 3-6: Energy deposited in the calorimeters by 57 GeV/c muons and pions at the testbeam. (a) Energy deposited in the CEM by muons, (b) energy deposited in the CHA by muons and (c) total energy deposited by pions in the calorimeter (CEM+CHA). The low energy peak in (c) is due to non-interacting pions or unidentified muons. (a) and (b) are from [38] and (c) from [39].

Isolation

As in the electron case, we define an isolation variable

$$Iso(R) \stackrel{\text{def}}{=} E_T^{0.4} - E_T^{0.13} \tag{3.16}$$

where $E_T^{0.4}$ is the electromagnetic plus hadronic transverse energy in the calorimeter in a cone of radius $R = 0.4$ centered about the μ track and $E_T^{0.13}$ is the same quantity for a cone of 0.13. We expect muons from W decay to be highly isolated and thus this variable is very useful in reducing punch-through background which is not isolated.

3.2.3 Muon identification cuts

Table 3.3 displays all the cuts used in the inclusive muon sample.

Parameter		Cut value
P_T	$>$	20 GeV/c
EM_μ	$<$	2 GeV
HAD_μ	$<$	6 GeV
$\mid Z_v - Z_0^\mu \mid$	$<$	5.0 cm
$\mid Z_v \mid$	$<$	60.0 cm
b	$<$	0.2 cm
Track-Stub matching:		
CMU: $\mid \Delta x \mid$	$<$	2.0 cm
CMP: $\mid \Delta x \mid$	$<$	5.0 cm
CMX: $\mid \Delta x \mid$	$<$	5.0 cm
Fiducial cuts		Applied
Remove cosmic rays		Applied

Table 3.3: Central muon selection criteria.

Fiducial cuts

At the edges of CMU chambers the detection efficiency falls off due to field distortions. The CMU fiducial volume cuts are applied to minimize this efficiency degredation. The CTC track is extrapolated to the CMU and the following cuts imposed:

1. The track must be at least 1.5° away from the azimuthal boundaries of CMU wedge it hits.

2. In local wedge pseudo-rapidity, η_{wedge}, the track must satisfy the condition $0.040 < \eta_{wedge} < 0.61$.

3. In the 'chimney' wedge the above condition becomes $0.040 < \eta_{wedge} < 0.50$.

3.2.4 Muon identification efficiency

An identical method to the one employed in the electron case is applied here. A sample of $Z^0 \rightarrow \mu^+ \mu^-$ decays are selected from which the efficiencies are extracted. There are 394 events in our sample. The results are summarized in Table 3.4. The total efficiency listed in the table does not include the 19% inefficiency due to the isolation requirement.

75

Cut	N_c	ϵ_c
EM_μ	368	0.965 ± 0.007
HAD_μ	385	0.993 ± 0.005
Δx	393	0.999 ± 0.001
All cuts	365	0.906 ± 0.010

Table 3.4: Muon identification efficiencies.

3.2.5 Cosmic ray removal

Cosmic ray muons could be confused with products of a real $p\bar{p}$ collision. Cosmic ray muons have broader impact parameter and $|\ z_{track} - z_{event}\ |$ distributions than prompt muons. Also, incoming legs of Cosmic ray tracks are more difficult to reconstruct due to timing problems. A Cosmic ray filter based on these criteria has been used.

The Cosmic ray filter rejects muons whose vertex is greater than 60 cm from the nominal interaction point, impact parameter is greater than 0.5 cm and $z-$coordinate of the track is less than 5 cm from the $z-$coordinate of the event vertex. Muons within $2°$ back to back in azimuth with a 'bad' track of $P_T > 10$ GeV/c are removed. 'Bad' tracks are defined as those which are not three-dimensional, have too few hits or track segments in the CTC, have impact parameter greater than 0.5 cm or whose $z-$coordinate is more than 5 cm away from the $z-$coordinate of the event vertex.

The Cosmic ray filter is more than 99.8% efficient for W and Z events, and the filtered samples contain less than 0.37% cosmic background.

3.3 Missing transverse energy

The $\not{E}_T$ variable is used to estimate the total transverse momentum of all the neutrinos in the event. It is computed as

$$\vec{\not{E}}_T \overset{\text{def}}{=} -\sum_{|\eta|<3.6} \vec{E}_T^{tower} \tag{3.17}$$

76

where the sum is over all towers within the pseudo-rapidity range $| \eta | < 3.6$. The quantity $\vec{E}_T^{tower}$ is a two-dimensional vector pointing from the event vertex to the center of the tower. Towers are included only if their energy is greater than a certain threshold: CEM and CHA > 0.1 GeV, PEM > 0.3 GeV, PHA and FEM > 0.5 GeV, and FHA > 0.8 GeV. The full calorimetry range is not used in the calculation of $\not{E}_T$ because of the Tevatron low$-\beta$ quads that obscure part of the FHA.

The above definition is inadequate in the inclusive muon sample because muons are minimum ionizing. So we correct the $\not{E}_T$ in the following manner

$$\vec{E}_T^{\nu} \stackrel{\text{def}}{=} \vec{\not{E}}_T + \sum (\vec{E}_T^{\mu-tower} - \vec{P}_T^{\mu}) \tag{3.18}$$

where the sum runs over all CTC tracks with $P_T > 15$ GeV/c that satisfy muon identification requirements listed in an earlier section and any other minimum ionizing muons with $P_T > 10$ GeV/c as identified by the soft muon algorithm described in the next chapter. The quantity $\vec{E}_T^{\mu-tower}$ is calculated from the track direction and the electromagnetic and hadronic energy deposited in the calorimeter tower traversed by this track.

3.4   Jets

Due to the confining properties of strong interactions, quarks and gluons manifest themselves as jets rather than isolated free particles. In $(\eta - \phi)$ space jets have a circular cone shape covering several CDF calorimetry towers. To get back the original parton energy from the final state jet, we need (1) a clustering algorithm and (2) a jet energy scale. In this search we do not attempt to reconstruct the energy of the original partons, consequently we select jets based on their observed properties only.

3.4.1   Trigger

The QCD jet sample will be used extensively in the next chapter in the context of determining fake backgrounds for our $b-$tagging algorithms. Here we briefly describe

the four jet triggers that will be used in that chapter.

All jet triggers require a Level 1 single trigger tower above the energy thresholds listed in Table 2.8. A localized cluster of electromagnetic and hadronic transverse energy in the calorimeter is then required at Level 2. Clustering proceeds in a manner similar to the one described for electron identification earlier. Towers with E_T above 3 GeV form seeds and all contiguous towers with E_T greater than 1 GeV are added to the seed to form a cluster. At Level 3 four separate triggers, each with a different transverse energy threshold, are used to write out events. The Jet-20, 50, 70, and 100 triggers require the presence of a cluster with E_T greater than 20, 50, 70, and 100 GeV respectively.

3.4.2 Clustering algorithm

A fixed-cone iterative clustering scheme is used Offline to reconstruct jets at CDF. The Plug and Forward calorimetry towers are first combined to form towers of the central region's dimensions: $\Delta\phi \times \Delta\eta = 15° \times 0.1$. The electromagnetic and hadronic transverse energies in these towers are then summed up to form the transverse energy,

$$E_T = E_T^{em} + E_T^{had} \tag{3.19}$$

$E_T^{em} \stackrel{\text{def}}{=} E_{em} \mathrm{Sin}\theta_{em}$ where E_{em} is the electromagnetic energy in the tower and θ_{em} is the polar angle between the true event vertex and the center of the tower. E_T^{Had} is defined similarly.

Jet clustering involves seeding towers, forming pre-clusters, extending pre-clusters to clusters and resolving overlap region issues between clusters. First all towers with $E_T > 1$ GeV are seeded. Seed towers adjacent to each other, on a corner or side, are grouped into pre-clusters. Pre-clusters are expanded to clusters using a fixed-cone iterative algorithm.

The E_T-weighted $(\eta - \phi)$ center of the pre-cluster is computed. A cluster is then

defined as the set of all towers with $E_T > 100$ MeV and within $R < 0.4$ [2] from the centroid of the pre-cluster. The cluster centroid is recomputed and its set of towers redefined accordingly. This procedure is repeated until the set of towers in a cluster does not change. Initial pre-clusters are always kept in a cluster regardless of their distance to the centroid. This prevents the centroid from shifting too far away in pathological situations.

If a cluster is a subset of another cluster, only the larger cluster is kept. If two clusters share towers, they are merged if the total E_T in common towers is $> 50\%$ of the E_T of the smaller cluster. Otherwise towers in overlap region are divided according to their proximity to cluster centroids. The centroids are recomputed until a stable configuration is reached.

3.4.3 Jet parameters

The four-momenta of jets are computed and used extensively in the search. The four-momenta are constructed from the energy deposited in the calorimeters and the event vertex measured by the VTX.

$$P_x = \sum_i (E^i_{em}\mathrm{Sin}\theta^i_{em} + \mathrm{E}^i_{had}\mathrm{Sin}\theta^i_{had})\mathrm{Cos}\phi^i \qquad (3.20)$$

$$P_y = \sum_i (E^i_{em}\mathrm{Sin}\theta^i_{em} + \mathrm{E}^i_{had}\mathrm{Sin}\theta^i_{had})\mathrm{Sin}\phi^i \qquad (3.21)$$

$$P_z = \sum_i (E^i_{em}\mathrm{Cos}\theta^i_{em} + \mathrm{E}^i_{had}\mathrm{Cos}\theta^i_{had}) \qquad (3.22)$$

$$E = \sum_i (E^i_{em} + E^i_{had}) \qquad (3.23)$$

The cluster momentum is approximately equal to the corresponding fragmenting parton momentum, in the limit where the particles forming the jet have small masses compared to their energies.

From the four-momenta we then form variables analogous to variables used in describing electrons and muons

---

[2] $R \overset{\mathrm{def}}{=} \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$

$$P_T = \sqrt{P_x^2 + P_y^2} \qquad (3.24)$$

$$P = \sqrt{P_x^2 + P_y^2 + P_z^2} \qquad (3.25)$$

$$\eta \stackrel{\text{def}}{=} \frac{1}{2}\ln\left(\frac{\mathrm{P} + \mathrm{P}_z}{\mathrm{P} - \mathrm{P}_z}\right) \qquad (3.26)$$

$$E_T \stackrel{\text{def}}{=} \mathrm{E}\frac{\mathrm{P_T}}{\mathrm{P}} \qquad (3.27)$$

We also define the detector pseudo-rapidity, η_d, analogous to the η mentioned above but using the nominal interaction point as the event vertex instead of the measured event vertex. Jet selection is then based on η_d and E_T.

Chapter 4

Low$-P_T$ lepton identification

It was mentioned in Chapter 1 that the direct production of W bosons in conjunction with jets (W+multijets) is known to be the dominant background for a top search in the lepton+jets channel. This background can be greatly suppressed by tagging the b quarks from the top quark decay. This search attempts to identify these b quarks via their semileptonic decays. Since these leptons are expected to have a much softer P_T spectrum and be surrounded by considerable jet activity, we must devise algorithms different from the primary lepton algorithms described in Chapter 3. In this chapter we describe the low$-P_T$ electron and muon algorithms used to tag b quarks in the top quark search. The two defining characteristics of such algorithms, fake rates and efficiencies, are also discussed here. We conclude the chapter with a description of the tests performed to check the validity of the algorithms developed.

4.1 Low$-P_T$ electron identification

The standard CDF electron reconstruction algorithm was designed for high P_T electrons. The numerous implicit isolation requirements in this algorithm render it inefficient for non-isolated, low$-P_T$ electrons that are expected from decays of b quarks coming from top decays (see Figure 1-8). Consequently we have adopted a track-based algorithm that utilizes the central strip chambers (CES), the central pre-radiator (CPR) and CTC tracking.

81

4.1.1 Electron reconstruction algorithm

To determine the detector response to soft electrons, an unbiased sample of low energy electrons had to be found. A good source of these exist in CDF data as the second leg of asymmetric photon conversions. In this section, we describe the selection criteria for this sample, demonstrate the quality of the resulting electron sample, and use it to define the electron identification cuts.

The Conversion sample

Conversions are selected solely by tracking requirements and the kinematics of the decay. We start with a low P_T inclusive electron sample and select conversions in the following manner:

- impose track quality cuts: consider only tracks that have hits in at least 2 axial and 2 stereo superlayers, and an impact parameter $<$ 3 mm

- since photons are massless, the two daughter tracks will appear parallel in the $R - Z$ view. Therefore, loop over all other tracks in the event and select the one that minimizes $\Delta(\cot\theta)$, requiring it to be less than 0.03 (see Figure 4-1). Call this track the 'conversion partner'

- now switch over to the $R - \phi$ plane and compute ΔS, the distance of closest approach between the initial track and it's conversion partner. Cut at $\mid \Delta S \mid <$ 0.2 cm

- require that the conversion partner propagates to a different wedge (CEM/CES/CPR) than the primary track

- plot $\Delta(\cot\theta)$, ΔS, radius of conversion point (R), and Z of conversion point (for the VTX active region). These are shown in Figure 4-1.

- to further clean the sample, only conversions occurring at radii consistent with the VTX outer wall or the CTC inner tube will be used (conversion radius in the range 22-30 cm).
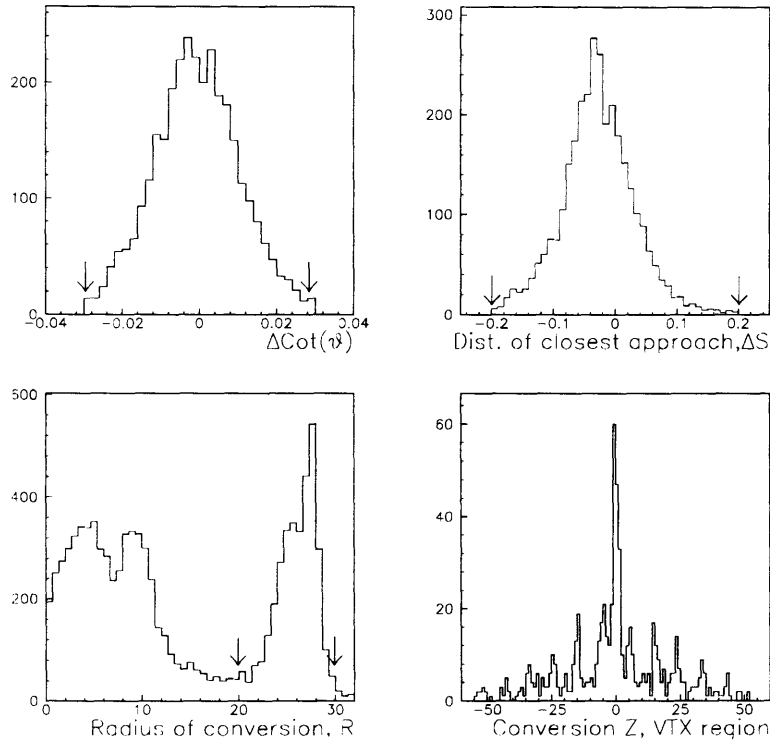
Figure 4-1: Characteristic plots for conversions. The arrows indicate cut values.

Up to this point, only tracking information has been used in the reconstruction. Taking the energy deposited in the electromagnetic compartment of the calorimeter in the tower which is struck by the conversion partner, we plot the ratio E/P (Figure 4-2a) for tracks with momenta in excess of 1 GeV/c. We observe a nice signal centered at 1, as expected for electrons, and a cluster of events with low energies.

Next we define our fiducial volume, requiring the track to satisfy:

- Local CES radial position $| X_{CES} | < 22$ cm

- Local CES Z position 6.22 cm $< | Z_{CES} | < 237.45$ cm

- Exclude chimney region

- Require CPR fiducial volume. This results in a little loss in efficiency since the CPR fiducial volume is in general smaller than the CES cuts described above.

and generate the same distribution. Figure 4-2b shows the E/P after fiducial cuts. A fit to the distribution yields a width of 15.3±0.4%, consistent with the electromagnetic

83

calorimeter resolution. Figure 4-3 shows the matching between the track and the strip clusters in the x and z views respectively.
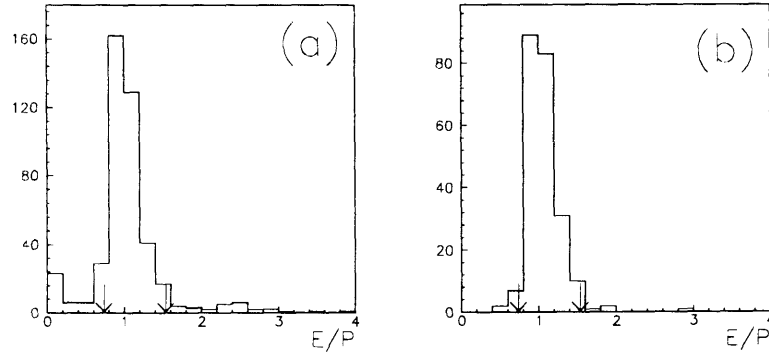
Figure 4-2: E/P for conversion electrons (a) without and (b) with fiducial cuts. The arrows indicate cut values.

Based on this sample, we define our electron selection criteria.
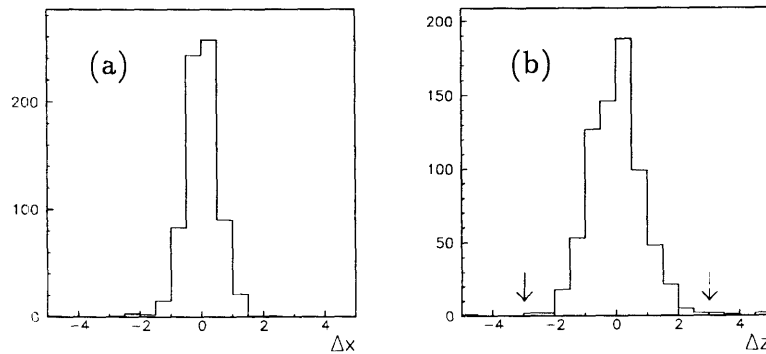
Figure 4-3: Track-CES matching for conversion electrons: (a) Wire view, (b) Strip view. Cut values are indicated by arrows.

Soft electron selection criteria

The details of the algorithm are finalized by balancing efficiencies and fake rates. They are listed below:

1. Extrapolate each track with $P_T > 2$ GeV/c that passes track quality cuts, mentioned in the previous section, to the CPR and CES.

2. Apply CES and CPR fiducial cuts (as defined above).

84

3. Require minimum pulse-height in the CES: Sum up the energy in 5 strips (E_S^5) and 5 wires (E_W^5) around the extrapolated track position and apply the momentum-dependent cuts listed below. The momentum dependence reflects shower fluctuations in the calorimeter.

$$
\begin{array}{llll}
E_S^5 & > 0.6 & \text{GeV} & \text{for P} > 12 \text{ GeV/c} \\
E_S^5 & > 0.24 + 0.03 \times \text{P} & \text{GeV} & \text{for P} \leq 12 \text{ GeV/c} \\
E_W^5 & > 0.6 & \text{GeV} & \text{for P} > 12 \text{ GeV/c} \\
E_W^5 & > 0.24 + 0.03 \times \text{P} & \text{GeV} & \text{for P} \leq 12 \text{ GeV/c}
\end{array}
$$

4. Matching cuts between the extrapolated track and CES strip and wire clusters are then imposed. The CES cluster positions are taken as the energy-weighted mean positions of 3-channel wire and strip clusters around the extrapolated track position. The wires and strips give us x and z information respectively. The cut values are:

   - $| \delta x | < 0.7$ cm for P > 6 GeV/c

   - $| \delta x | < 1.82 - 0.1867 \times$ P cm for P< 6 GeV/c

   - $| \delta z | < 2.0$ cm

   Once again the momentum dependence of these cuts reflects shower fluctuations in the calorimeter.

5. Require that the shower shape be consistent with that of an electron. This is imposed via a cut on the 7-channel CES cluster χ^2:

   - $\chi^2(\text{wire})/6 < 16$

   - $\chi^2(\text{strip})/6 < 16$

6. The charge deposition in the CPR is then considered. The sum charge of 3 CPR wires around the extrapolated track position (Q_{CPR}) is calculated and cut upon. Figure 4-4(a) plots the Q_{CPR} distributions for conversion electrons and generic tracks that fail the E/P < 0.5 cut. Figure 4-4(b) shows the CPR

85

response to electrons and pions passing the same exact cuts except for the CPR requirement. The cut value adopted

- $Q_{CPR} > 2000$ fC

is also illustrated in the figure. This CPR energy deposition cut corresponds to at least 4 minimum ionizing particles.

7. Finally, the ionization loss in the CTC is required to be consistent with that of an electron [40]. Figure 4-5 shows the dE/dx distribution for a sample of conversion electrons and pions. A cut of

- $Q_{CTC} > 14.55$ ns

is imposed. This value is chosen to make the cut 90% efficient for conversion electrons. The units are nanoseconds since the integrated charge information is stored in the digital pulse width [1] of a signal. It is this digital pulse width that is used to calculate average dE/dx .

4.1.2   Additional cuts for top quark search

After imposing the CES-CPR cuts listed above, we apply additional cuts to the electron candidates in our top search:

- We first remove conversions via the following loose criteria

    - $\mid \Delta(\cot\theta) \mid < 0.06$

    - $\mid \Delta S \mid < 0.3$ cm

    - M(ee) < 500 MeV/c^2

    - R < 50 cm.

---

[1]The digital pulse width is the difference between the trailing and leading edge times
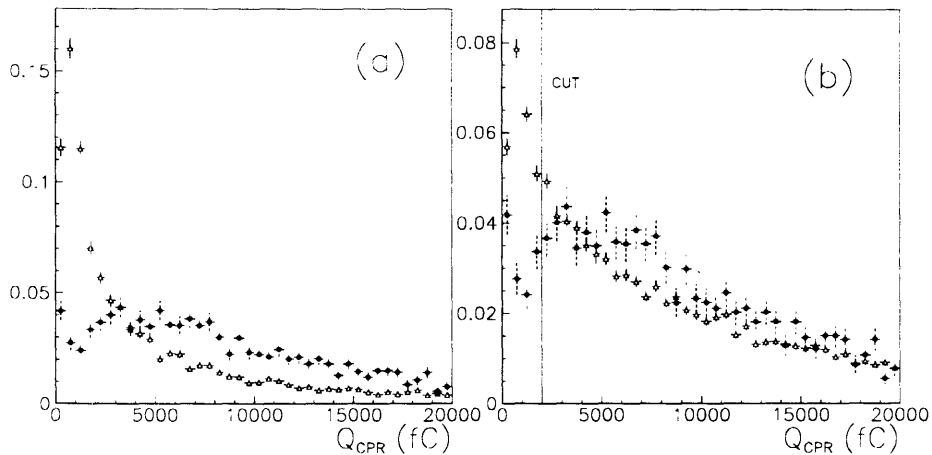
Figure 4-4: The CPR charge deposition, Q_{CPR}, for (a) conversion electrons (dots) and generic tracks that fail the E/P< 0.5 requirement (triangles), and (b) electrons and pions passing the same cuts except for the CPR cut. The line indicates the cut value adopted.

Here M(ee) is the mass of the track pair and all the other variables are as described earlier. From a study of random tracks in jet events, we estimate that 3% of prompt electrons will also be removed by these requirements.

The calorimeter energy (E), in both electromagnetic (EM) and hadronic (HAD) compartments of the detector, is defined as the energy in the tower that the track points to. If the track is within 2 cm in the $z-$direction of the nearest tower boundary, we sum up the energy of the two towers. We then impose E/P and HAD/EM cuts on the electron candidates:

- $0.7 < E/P < 1.5$

- $HAD/EM < 0.1$

These isolation cuts, E/P and HAD/EM, are motivated by the distributions in Figure 4-6 and the fact that they decrease the fake rates by a factor of two. It is clear that these cuts are more physics dependent than the CES-CPR cuts described above, and hence the determination of their efficiencies will have a larger systematic uncertainty.

87
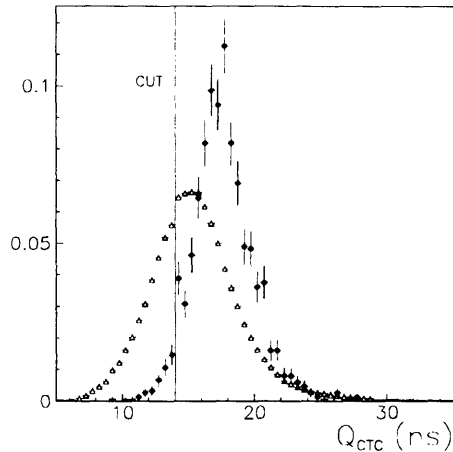
Figure 4-5: The dE/dx distribution for electrons (circles) and pions (triangles). The line indicates the cut value of 14.55 ns.

Cut	$b \to e$	$b \to c \to e$
E/P	0.82±0.06	0.50±0.05
HAD/EM	0.58±0.05	0.38±0.04
E/P and HAD/EM	0.51±0.04	0.27±0.03

Table 4.1: The E/P and HAD/EM efficiencies from Monte Carlo generated top events. This table was made for a top mass of 140 GeV/c^2. Shown separately are efficiencies for electrons coming from direct b decay and sequential $b \to c \to e$ decay. The errors are statistical only.

4.1.3 Efficiency

The P_T spectra of all fiducial conversions and those passing the CES-CPR cuts are shown in Figures 4-7(a) and (b). The total efficiency obtained from these histograms, as a function of the electron transverse momentum, is then shown in Figure 4-7(c). This plot does not include the 10% inefficiency from the dE/dx cut discussed above.

The efficiencies of the E/P and HAD/EM cuts are extracted from Monte Carlo generated top events and are summarized in table 4.1. The details of this extraction are relegated to Chapter 5. A treatment of the systematics associated with these isolation cuts will also be discussed in that chapter.
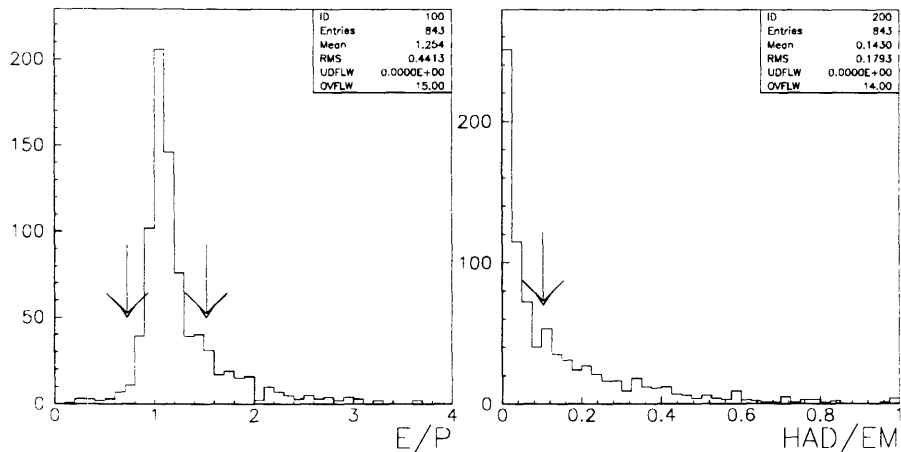
Figure 4-6: E/P and HAD/EM distributions of electrons from b and c decays in Monte Carlo generated top events. Shown here are plots for $M_{top} = 140$ GeV/c^2 (The arrows indicate cut values).

4.1.4 Fake rate

The probability for a hadron to pass the electron selection requirements is obtained from a study of tracks in QCD jet events. In this study we assume that all tracks in this sample are hadrons. In reality any such sample will also contain a small fraction of conversion electrons that were not removed by our algorithms, as well as prompt electrons from bottom and charm decays. For these reasons a fake rate obtained from the QCD sample will always be an over-estimate of the true hadron fake rate. This issue will be discussed further in chapter 6.

The fake rate is then used to predict the number of expected background electron candidates in the lepton + jets sample. Such usage of the QCD fake rate is equivalent to making the assumption that the heavy flavor content of jets in the W sample is the same as that of jets in the QCD sample.

Our electron selection algorithm has a number of implicit isolation requirements (e.g. the E/P and HAD/EM cuts). An isolated, early-interacting hadron has a higher probability to be reconstructed as an electron than a non-isolated one, since the energy deposition due to nearby particles can cause it to fail the HAD/EM

89
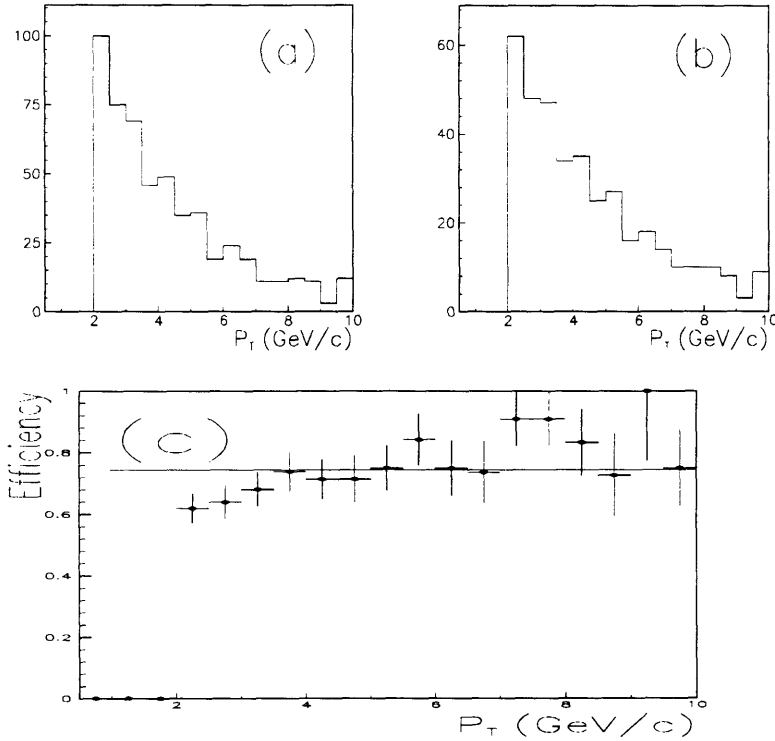
Figure 4-7: P_T spectra for (a) all fiducial conversions, (b) conversions passing CES-CPR cuts mentioned in the text. (c) shows the soft electron efficiency from conversions (no dE/dx , E/P or HAD/EM cut has been applied).

or E/P requirements. Conversely, a hadron that deposits little energy in the EM compartment would fail the electron requirements, but it is possible that nearby π^0s would deposit enough energy in the same cell to make the E/P value consistent with that of an electron. Therefore we expect the probability for a hadron to fake an electron to depend on isolation.

Method

For each track of momentum P, we define an isolation variable SUMP2 as the scalar sum of the momenta of all tracks within a cone of 0.2 centered around the track. The cone is defined at the face of the calorimeter, and the size of the cone is chosen to be not much larger than the typical hadronic shower. In Figure 4-8 we show the expected SUMP2/P distribution for electrons from bottom and charm decays in Monte Carlo top events that pass the HAD/EM and E/P requirements. We extract fake rates as

90

a function of SUMP2/P from a sample of jets.
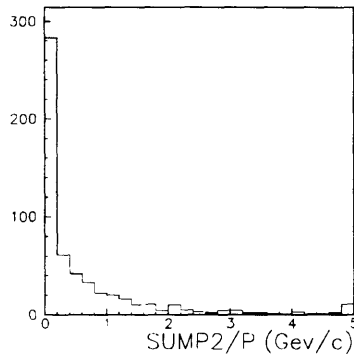
Figure 4-8: The SUMP2/P distribution from Monte Carlo top events of mass 140 GeV/c^2. The overflow bin is also shown.

We divide the QCD sample into three bins of roughly equal statistics:

- SUMP2/P < 0.2

- 0.2 < SUMP2/P < 5.0

- SUMP2/P > 5.0

Figure 4-9 shows the fake rates as a function of transverse momentum for the three isolation bins. These fake rates are obtained by taking the ratio of the P_T spectra of all reconstructed tracks passing all the electron requirements to the P_T spectra of all tracks in the fiducial region described in Section 4.1.1.

To calculate the expected fake rates in a given sample, we then apply the prescription:

- Convolute the P_T spectra of all tracks in the fiducial volume in each of the three SUMP2/P bins with the corresponding fake rate plot from Figure 4-9 to obtain the expected number of fake electrons in each bin.

- Add the three fake expectations to obtain the total number of expected fake electrons.
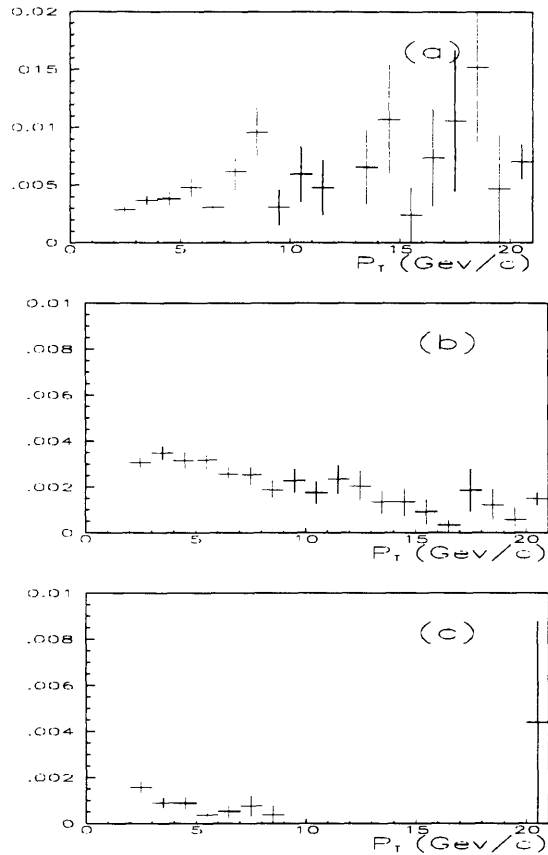
Figure 4-9: Fake rates as a function of isolation. The three plots represent the three isolation bins:(a) SUMP2/P < 0.2, (b) 0.2 < SUMP2/P < 5.0, and (c) SUMP2/P > 5.0. The isolation variable SUMP2/P is described in the text.

Systematics

It is important to check the stability of the fake rate calculation under different conditions. The heavy flavor content of different samples may be different, and as indicated in Figure 4-9 there is a clear isolation dependence which we have attempted to model. The jet sample we used was a mixture of Jet−20, 50, 70, and 100 triggers. We have computed the fake expectations for the four triggers individually. These expectations together with the number of observed events are tabulated in Table 4.2. We observe a ±15% spread on the fake rate as a function of jet trigger. We take this spread to be the systematic uncertainty on our fake rate prediction.

Trigger	Expected tags	Observed tags
Jet−100	242 ± 36	185
Jet−70	267 ± 40	260
Jet−50	239 ± 36	239
Jet−20	436 ± 65	499

Table 4.2: Fake rate expectations versus the number of observed events as a function of jet triggers for the HAD/EM cut. Errors shown are the 15% systematic errors for expectations from fake events.

Leading jet dependence		
Trigger	Expected tags	Observed tags
With Leading jet	467 ±70	438
Without Leading jet	707 ±106	745

Table 4.3: Fake rate predictions versus observed events with and without the leading jet in the event. The normalization is arbitrary and the 15% systematic errors are shown.

Trigger bias in the Jet sample

We also investigate any possible trigger bias in our fake rate determination from the QCD sample. We compare fake rates and numbers of observed events with and without the leading jet in the event. The results are summarized in Table 4.3. With the leading jet we expect 467 events and we observe 438, and without it we predict 707 and see 745. We conclude that the trigger dependence is of order $\sim 5\%$ or less.

4.2 Low−P_T muon identification

Unlike the low−P_T electron identification scheme, where a reconstruction algorithm different from the standard one for high−P_T electrons had to be devised, the low−P_T muon identification uses the standard muon reconstruction code. This is possible because the muon reconstruction code does not have implicit isolation biases inherent to it − all objects that pass the Level 3 muon trigger form Offline data storage struc-

tures (CMUO banks). Thus any muon search in CDF begins with these CMUO banks which contain all the relevant information about the candidate, such as momentum, position, calorimetery information and much more.

As mentioned in Chapter 2, muons can be detected in the three central muon systems: CMU, CMP and CMX. In this search we do not include low$-P_T$ CMX muons since their addition to the acceptance is small while their associated backgrounds are large. Depending on the muon's trajectory and the actual stub reconstruction in the muon code, CMUO banks will have CMU$-$only, CMP$-$only or both CMU$-$CMP stubs. The first step in identifying soft muons is then the requirement that the appropriate stubs be present. Candidates tracks are extrapolated to the muon chambers taking into consideration multiple scattering and ionization loss effects. If a track extrapolates to the fiducial volume of the CMU (CMP) and is at least 3σ of multiple scattering or 5 cm away from the nearest chamber edge, we require the CMU (CMP) stub to be present. The x and z positions of a CMU stub, if present, are also used to decide if a CMP stub is expected. Further, a study of J/ψ muons shows that tracks with $P_T < 3$ GeV/c will not make it to the CMP chambers. Consequently CMP stubs are not required for muon candidates with $P_T < 3$ GeV/c .

The stub requirement classifies our soft muon candidates into three distinct categories: CMU$-$only , CMP$-$only, and CMU$-$CMP muons. Each of these three categories is treated slightly differently as is detailed below. There is still the question of how to treat cases where the track extrapolated to more than 3σ outside the fiducial volume of the CMP but a CMP stub is present. This scenario could arise either when we're not dealing with a muon or there is some error in the muon stub reconstruction code (CMLNK). In the latter case, a good muon would be lost if CMP matching cuts were applied (discussed below). Evidence of pattern recognition failure has been obtained from samples of J/ψ muons. To avoid this inefficiency we ignore the CMP stub in such cases at the price of a modest increase in the background.

The efficiency of the muon algorithm is extracted from J/ψ and Z boson data. Fake rates are obtained analogously to the electron case: a fake probability per track is obtained from QCD data and then convoluted with a track spectrum in the data

CMU−only	Number of TDC hits ≥ 3	
	Number of ADC hits ≥ 3	
	CMCLUS < 6	(see text)
	$\mid \delta z \mid < $ Max$(3\sigma, 8\text{cm})$	
	$\chi^2(\text{x}, \text{slope}) < 15$	$(P_T < 20 \text{ GeV}/\text{c})$
	$\mid \delta x \mid < $ Max$(3\sigma, 2\text{cm})$	$(P_T > 20 \text{ GeV}/\text{c})$
	Minimum ionizing (see text)	$(P_T > 6 \text{ GeV}/\text{c})$
CMU−CMP	Number of TDC hits ≥ 3	
	Number of ADC hits ≥ 3	(CMU)
	$\mid \delta z \mid < $ Max$(3\sigma, 8\text{cm})$	(CMU)
	$\mid \delta x \mid < $ Max$(3\sigma, 2\text{cm})$	(CMU)
	No requirement on CMP stub	
	Minimum ionizing (see text)	$(P_T > 6 \text{ GeV}/\text{c})$
CMP−only	$\chi^2(\text{x}, \text{slope}) < 10$	$(P_T < 10 \text{ GeV}/\text{c})$
	δslope < 0.1	$(P_T > 10 \text{ GeV}/\text{c})$
	$\mid \delta x \mid < $ Max$(3\sigma, 5\text{cm})$	$(P_T > 10 \text{ GeV}/\text{c})$
	Minimum ionizing (see text)	$(P_T > 6 \text{ GeV}/\text{c})$

Table 4.4: Soft muon selection criteria.

sample of interest. This fake probability is inherently an over-estimate of the total fake background since there are some real muons in the QCD sample from charm or bottom decay.

4.2.1 Cuts

We start by making simple quality cuts on the tracks: the impact parameter is required to be less than 3 mm, and hits in at least two axial and two stereo superlayers in the CTC are required. The impact parameter cut is designed to remove cosmic ray muons and some decays in flight. Objects that survive the track quality cuts undergo matching and minimum ionizing cuts. These cuts are different for the three categories of low P_T muons mentioned earlier. Table 4.4 summarizes the cuts.

The slope measurements in CMU and CMP have non-gaussian tails that become important at high P_T, at which multiple scattering effects become quite small. Thus, in order to mantain good efficiency, all slope cuts are relaxed or removed at high P_T .

The variable CMCLUS is defined for CMU stubs as the number of TDC hits clustered around the stub, including hits used in constructing the stub. A perfect muon will have CMCLUS = 4; one can get CMCLUS = 5 if there are delta rays and/or the muon crosses chamber cells. On the other hand, an interacting hadron that punches through the calorimeter often results in a shower of particles emerging at the back of the calorimeter, causing a splash of hits in the CMU chambers. Figure 4-10 shows the CMCLUS distribution for muons from J/ψ decays (after sideband subtraction) and for CMUOs in jet events that pass all other cuts. A modest ($\approx$ 8%) additional background rejection can be achieved with essentially no loss in efficiency by demanding CMCLUS < 6.
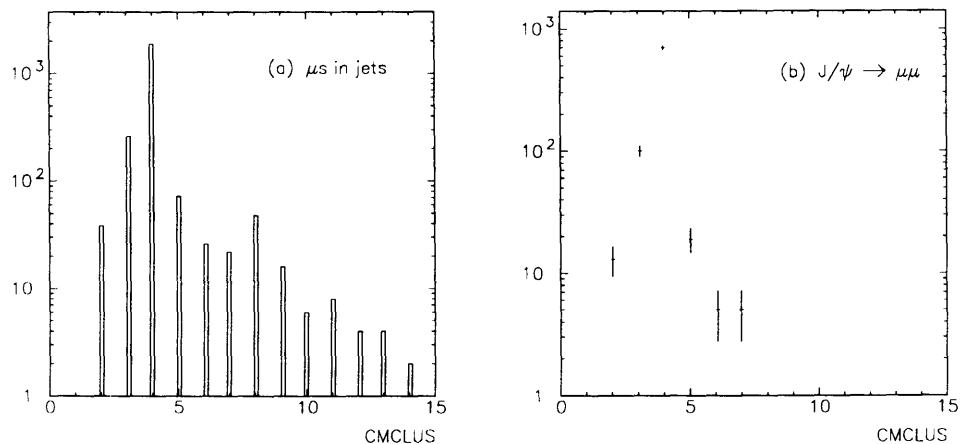
Figure 4-10: The CMCLUS distribution for muons (a) in jets, and (b) from J/ψ decays.

A minimum ionizing cut is applied to muon candidates of P_T > 6 GeV/c, since at low momentum calorimeter information cannot be used to separate muons from interacting hadrons. We propagate the track to the middle of the CHA and take the energy of the tower to which the track points (E_{HAD}). In order to remain efficient for muons inside jets, we do not cut directly on E_{HAD}. Instead we use the variable

96

SUMP2 as previously defined. We then require

$$E_{HAD} - SUMP2 \;\; < 6 \;\; \text{GeV} \qquad\qquad (4.1)$$

This cut rejects muon candidates that (i) have $E_{HAD} > 6$ GeV, inconsistent with monimum ionizing alone, and (ii) such that the extra energy in the tower cannot be attributed to nearby hadrons. Monte Carlo studies show that this cut is approximately 99% efficient for muons from b-decays. Approximately 10% of high P_T muon candidates from QCD jet triggers, passing all other requirements, are rejected by this cut.

The TDC and ADC cuts on CMU are very efficient, and reject punch-through background. The $\chi^2(x, \text{slope})$ cut for CMU is slightly inefficient, however if this cut is replaced by a δx only requirement, the background rates for the CMU-only muons increase by almost a factor of two. In the CMU-CMP case, the punch-through background is greatly reduced, and the slope requirement is removed. In the CMP-only case, we impose requirements on both position and slope matching, since the background rates appears a little high probably due to hadrons going through cracks in the calorimeter.

4.2.2 Efficiency

The efficiency of the muon selection requirements can be broken up into three parts:

1. the stub efficiency, i.e the efficiency for a muon passing through the muon chambers to leave a reconstructable stub, and for the muon reconstruction code to associate it with the CTC track, ignoring pattern recognition problems,

2. The (CMLNK) pattern recognition efficiency,

3. The efficiency of the cuts listed in the previous section.

Pattern recognition inefficiencies are physics dependent, and are modelled by the Monte Carlo. All other efficiencies are measured directly from J/ψ data.

Muon quality cuts efficiency

The J/ψs used in this study were selected from the J/ψ Stream 2 datasets reconstructed with Offline version 6.15; a minimal set of quality cuts had already been applied to these samples, however the cuts were loose and close to 100% efficient.

The efficiency for muons was statistically corrected for the small non-muon contamination using the events in the J/ψ sidebands. The efficiency is shown in Figure 4-11 together with a simple parametrization derived from J/ψ data.
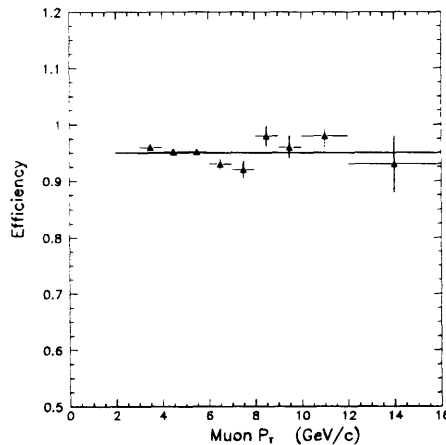
Figure 4-11: The soft muon efficiency as a function of P_T as measured from a sample of $J/\psi \rightarrow \mu^+\mu^-$ decays.

Efficiency of stub requirement

As was mentioned earlier, muon candidates in the CMU(CMP) fiducial volumes are considered only if the CTC track is matched to the relevant stub. The stub may be missing for a number of reasons: chamber inefficiencies (there are a small number of dead CMP chambers), CMLNK pattern recognition problems, and problems with track propagation.

The efficiency of the CMP stub reconstruction is measured using muons from J/ψ and Z° decays that are within the CMP fiducial volume. To avoid trigger biases, the Z° events used in this study were required to have fired the dimuon trigger,

which does not require the CMP. To separate out the stub-finding efficiency from pattern recognition inefficiencies, we required events to have no additional CMUO banks containing a CMU or CMP stub (this rejects the class of events where the muon left a stub in both the CMU and the CMP and the CMP stub was matched to a different CTC track). The efficiencies using muons from J/ψ and Z decays were found to be 96.1 ± 0.2% and 96 ± 2% respectively.

Muons from J/ψ decays cannot be used to extract the CMU stub efficiency since the presence of a CMU stub is required at the trigger and event selection level. To measure the CMU stub efficiency, we used the second muon from Z decays (the first muon was required to pass the CMU-CMP quality requirements). We find the efficiency to be 98 ± 2%.

4.2.3 Fake rate

The Fake rate is defined as the probability (as a function of transverse momentum) that a hadron satisfies the muon selection requirements described earlier. In order to measure the fake rate, we turn to a large sample of jet events (Jet-20, 50, 70 and 100 triggers) and compute the probability, as a function of transverse momentum, that a track within the fiducial volume of the muon chambers passes the muon selection requirements. As was mentioned earlier, this is an overestimate of the true fake rate, since there are a small number of prompt muons from bottom and charm decays in any jet sample.

A jet containing a hadron that decays in flight to a muon or that does not interact in the calorimeter will have its calorimeter energy measured systematically low. Because the trigger requires a minimum energy deposistion, and because the jet E_T spectrum falls rather steeply, one must be careful to remove trigger bias from the data used in the fake rate determination. For this reason we only consider tracks and muons candidates in jets that are unbiased by the trigger. Specifically, we require (i) that there be at least one Level 1 trigger tower above threshold which is at least 0.7 units of R away from the track/muon-candidate, and (ii) that there be at least one Level 2 trigger jet cluster above threshold which is also 0.7 units of R away from the

track/muon-candidate.

The Fake rates for the different types of muons (CMU−CMP, CMU−only, CMP−only and the total) are shown in Figure 4-12. There are three different fake rates, for the three different fiducial regions, and the total fake rate is the weighted average of the three fake rates. When trying to predict the number of expected background muons in a large enough sample of tracks, it is sufficient to use the total fake rate, and this is what is done in this search.
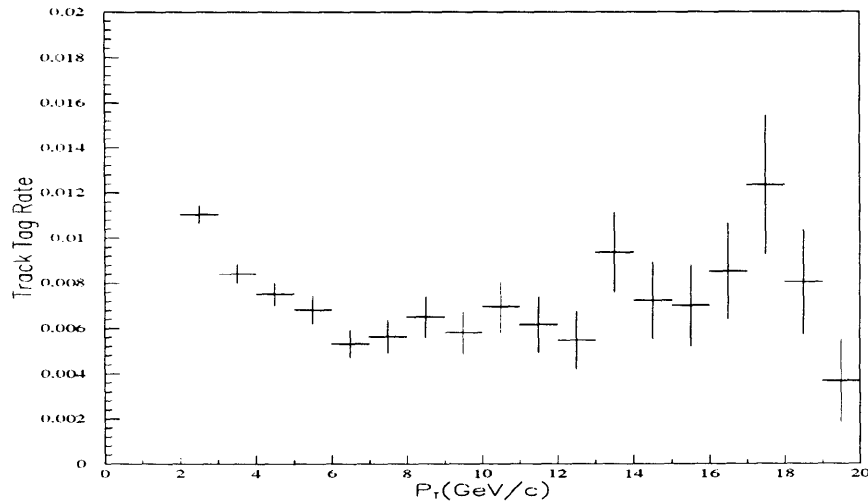
Figure 4-12: The soft muon fake rate.

It is important to check the stability of the fake rate under different conditions. In general the fake rate will depend on

1. The heavy flavor content of the sample.

2. The particle composition of the sample. This is because different particles have different interaction lengths. Furthermore, the decay in flight probabilities depend on the relative abundance of pions and kaons in the sample.

3. The shape of the parent P_T spectrum.

In addition, the fake rate may also depend on the isolation properties of the sample, since there is a finite probability that CMLNK will link stubs to the wrong

100

track, and this probability depends on the track density. All of these are rather subtle issues that are very difficult to quantify individually. We can try to estimate their combined effect by breaking up our jet sample into several subsamples with different characteristics, and compare the fake rates obtained from these subsamples.
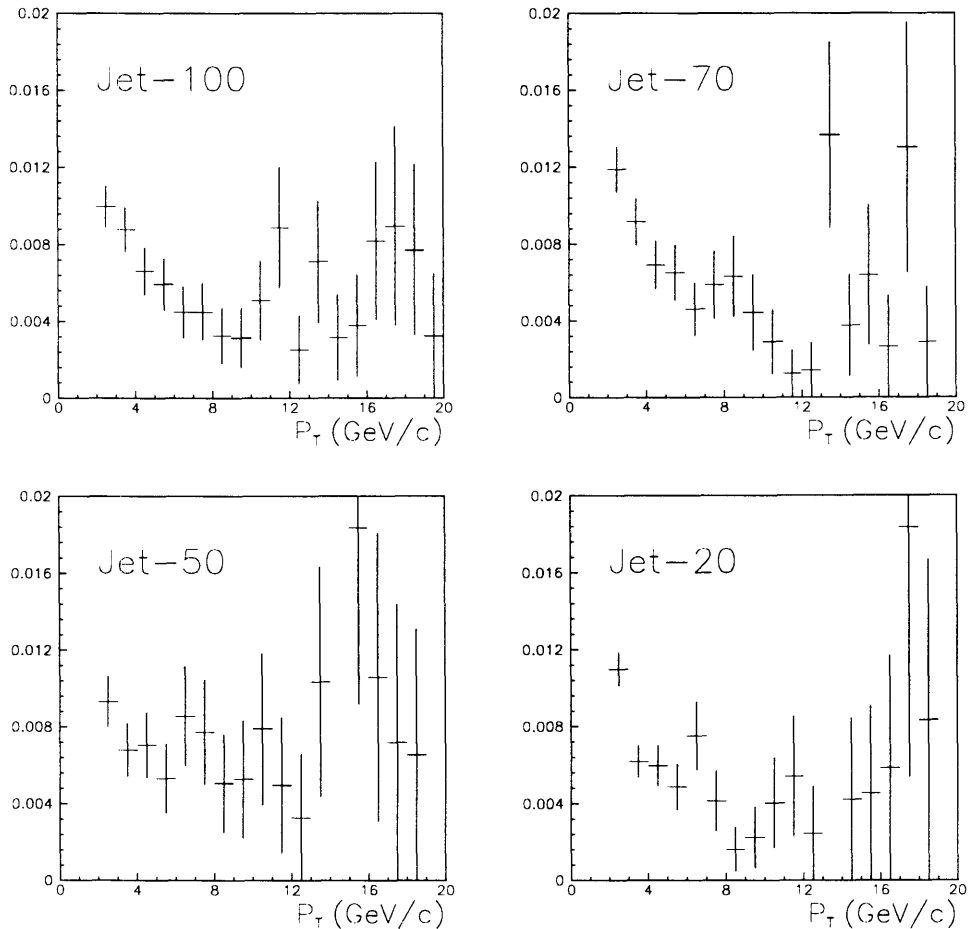
Figure 4-13: The soft muon fake rate as a function of jet trigger.

In Figure 4-13 we show the total fake rate for the four jet triggers in the data sample. The particle composition and heavy flavor content in jets of different energies may be different; in addition tracks in lower E_T jets tend to be more isolated, and the shape of the P_T spectra in the different samples also depends on the jet energies (see Figure 4-14). However, the resulting fake rates are very similar. The most relevant way of quantifying possible differences is to convolute these four fake rate histograms with the same track spectrum. (The track spectrum that is used is the sum of the

Trigger	Fake Rate (arbitrary units)
Jet−100 trigger	97 ± 4
Jet−70 trigger	107 ± 5
Jet−50 trigger	95 ± 5
Jet−20 trigger	102 ± 5
All triggers	100 ± 2

Table 4.5: Fake rate predictions vs. number of observed events as a function of jet trigger in the QCD sample. Systematic errors for the expectations are shown. The normalization is arbitrary.

track spectra for the four trigger samples). The resulting fake rates (normalized so that the average fake rate is 100) are shown in Table 4.5. The fake rates are statistically consistent with being the same for the four jet samples to 10% or better.
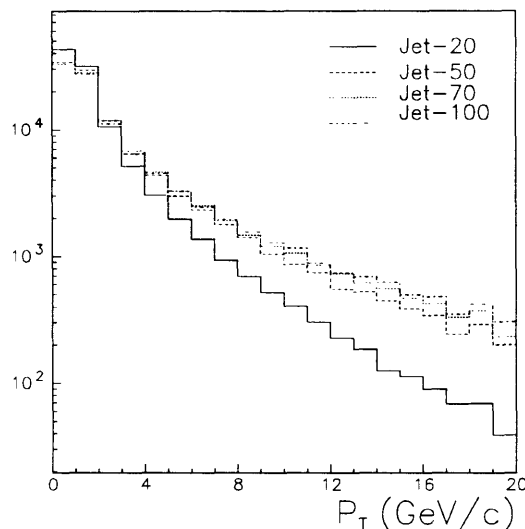
Figure 4-14: The track P_T spectra in the four jet trigger samples.

We have also computed fake rates as a function of isolation (SUMP2). In Figure 4-15(a) we show the SUMP2 distribution for all tracks in our jet sample. We divide the jet sample into three subsamples of approximately equal size, with SUMP2 < 2 GeV/c , 2 GeV/c < SUMP2 < 20 GeV/c , and SUMP2 > 20 GeV/c , and we calculate the fake rates for the three samples. Note that the particle content of the

isolated and non-isolated samples may be very different, e.g prompt leptons will tend to be somewhat more isolated, and the K/π ratios in the core and periphery of a jet may be different. Furthermore, the P_T spectra for the isolated and non-isolated sub-samples are very different (see Figure 4-15(b)).
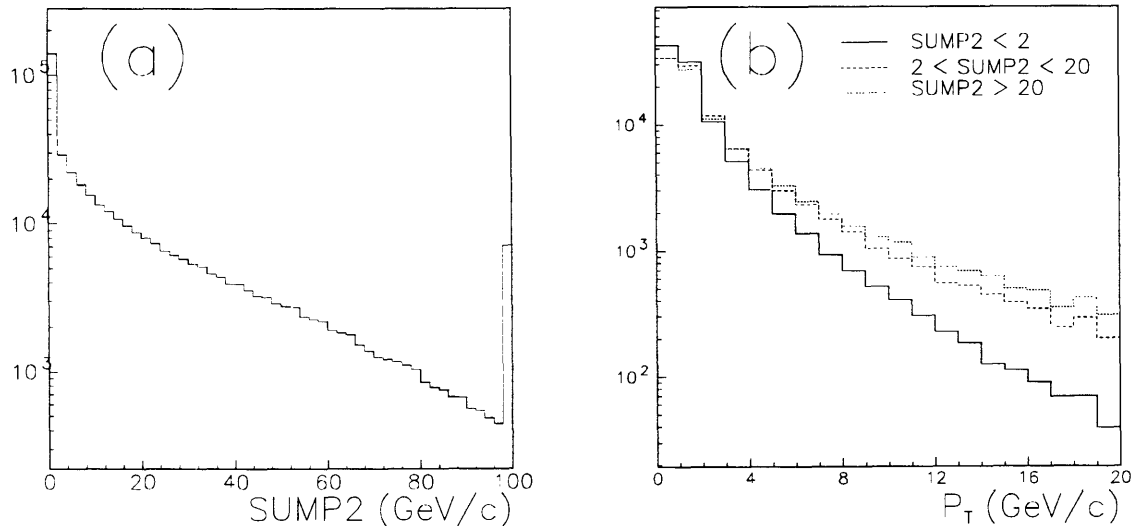
Figure 4-15: The SUMP2 isolation variable. (a) The SUMP2 distribution in the jet sample, and (b) the P_T distribution for the different SUMP2 isolation requirements.

The resulting fake rates are shown in Figure 4-16. It is immediately apparent that the fake rate for the highly non-isolated sample (SUMP2 > 20 GeV/c) is lower than for the other two samples. To quantify the difference, we again convolute the fake rates obtained from the three different samples with the same track spectrum used in the comparison of the trigger samples. The resulting fake rates are shown in Table 4.6.

All the fake rates as a function of SUMP2 are statistically consistent with each other, although there may be a trend towards higher fake rates for the more isolated tracks.

Systematics

As described in the previous sections, we have studied systematic variations of the fake rate as a function of jet trigger (i.e jet energy), isolation and data sample (i.e
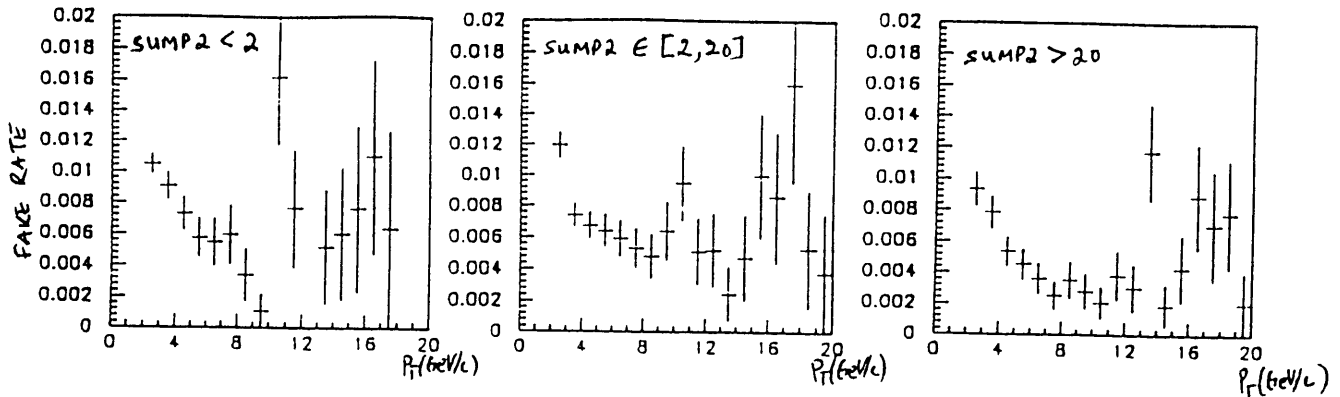
Figure 4-16: The total fake rates for the three Fake rate bins.

SUMP2	Fake rate (arbitrary units)
< 2 GeV/c	106 ± 5
> 2 GeV/cand < 20 GeV/c	103 ± 4
> 20 GeV/c	96 ± 5
All	100 ± 2

Table 4.6: Fake rate predictions vs. isolation.

generic jets vs. jets in direct photon events). The fake rate appears to vary by less than 10% as a function of the isolation properties of the sample. We take the systematic uncertainty in the fake rate calculation to be 10%.

4.3 Checks on the soft lepton algorithms

It is important to check the stability of our background prediction scheme. Here we present the result of testing our fake rate calculation in three different physics datasets: the Z+jets sample, the photon sample and an inclusive electron sample enriched with b quarks. We find our fake prediction to be remarkably accurate even in the wide variety of physics processes these three samples represent.

4.3.1 Z+jets study

Although the statistics of the Z sample are rather limited, it still provides a good check of the background prediction, since QCD Z+jets and W+jets production are very similar.

Events for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ were selected by requiring one e/μ to satisfy the standard e/μ selection. The second lepton was selected with the following requirements.

Electrons:

- $E_T > 10$ GeV

- $Isol(0.4) < 0.1$

- $HAD/EM < 0.1$

- E/P < 2 (if Central)

- $\chi^2(3X3) < 3$ (if Plug)

- VTX occupancy > 0.5 (if Plug)

Muons:

- $P_T > 10$ GeV/c

- $Isol(0.4) < 0.1$

- $EM_\mu < 2$ GeV

- $HAD_\mu < 6$ GeV

- Impact Parameter < 3 mm

- $\mid \delta x \mid < 5$ cm if CMUO

- $\mid \eta \mid < 1.1$ if CMIO

Jet Multiplicity	Number of Events	Expected Fake tags	Observed tags
≥ 1 jet	186	4.1±0.4	6
1 jet	163	2.9±0.3	6
2 jets	20	0.8±0.1	0
≥ 3 jet	3	0.44±0.04	0

Table 4.7: Background test in Z+jets sample.

The invariant mass of the lepton pair is then required to be within 15 GeV/c^2 of the Z mass. There are 466 $Z \rightarrow \mu^+\mu^-$ and 1314 $Z \rightarrow e^+e^-$ candidates in our sample. In the following table we summarize the results of our search for lepton tags in the Z+jets sample. Although the statistics are very limited, the number of tags in the Z+jets data agrees very well with our fake prediction.

4.3.2   Electron sample

We also apply our algorithm on the inclusive electron sample to test its validity. Two sub-samples are selected: a 'gold' conversion sample, i.e. conversions selected with very tight cuts, and a 'gold' electron sample. In the conversion sample the number of tags should be given by our predicted background (this serves as a cross-check of the fake rate). In the 'gold' electron sample, which is enriched with b quarks, we look for soft leptons opposite in ϕ to the trigger electron. Here we expect an excess over the calculated background due the presence of b quarks.

The conversion sample is selected with the tight cuts listed in section 4.1.1:

- $\mid \Delta(\cot \theta) \mid < 0.03$

- $\mid \Delta S \mid < 0.2$ cm

- 22 cm $< R <$ 30 cm.

The 'gold' electron sample is selected with the following cuts:

- $E_T > 12$ GeV

106

	Expected Background	electrons found
Gold Conversions	35 ± 5	32
Gold Electrons	99 ± 15	186

Table 4.8: Number of electron tags vs. expected background in the 'gold' samples.

- $E/P < 1.4$

- $HAD/EM < 0.02$

- $L_{shr} < 0.2$

- $| \delta x | < 2$ cm

- $| \delta z | < 3$ cm

- χ^2 (strip) < 4

- χ^2 (wire) < 4

- | impact parameter $| < 2$ mm

- one and only one 3-D track in the ELES cluster

A background estimate is obtained from each of these sub-samples and this estimate is compared to the number of tagged events. The background is calculated by multiplying the fake rate distribution obtained earlier with the P_T spectrum of all tracks passing fiducial cuts. Tables 4.8 and 4.9 summarize the results for electrons and muons respectively. From these tables we see that in the 'gold' conversion sample, where little or no b enrichment is expected, we tag 32 electron events which is in excellent agreement with the background calculation of 35 ± 5 events expected. And similarly, we tag 52 muon events when 45 ± 5 are expected. In addition, both the soft electron and soft muon algorithms tag a number of events well in excess of the expected background in the 'gold' electron sample. We attribute this excess to tagging genuine leptons from the second b in the event.

107

	Expected Background	μs found
Gold Conversions	45 ± 5	52
Gold Electrons	123 ± 12	181

Table 4.9: Number of muon tags vs. expected background in the 'gold' samples.

4.3.3 Photon sample

As a further check on the fake rate calculation, we have looked for electrons and muons in the photon sample. The heavy flavor content of jets in this sample is expected to be different from that of jets in QCD for a number of reasons. Most noticeably, there is no direct $b\bar{b}$ or $c\bar{c}$ production and the gluon-to-quark ratio is different. Direct photons thus provide us with a different physics sample to test our fake rate calculation.

We have used the Stream 2 photon-70 sample as well as a subset of the Stream 1 photon-16 sample. Event selection requires the presence of one and only one electromagnetic cluster with the requirements:

1. $E_T > 10$ GeV

2. there is at least one CES cluster associated with the EM cluster

3. no 3-D tracks point to the cluster

4. isolation: the energy in a cone of 0.7 around the EM cluster < 4 GeV

5. fiducial cuts: $\mid X_{CES} \mid < 17.5$ cm, 14 cm $< \mid Z_{CES} \mid < 217$ cm.

The sample is then further broken down into "π^0-like" and "γ-like" clusters. Clusters are considered "γ-like" if in addition to the cuts listed above they satisfy:

1. $< \chi^2 > < 4$

2. one and only one CES cluster is associated with the EM cluster.

We then look for electrons and muons that are at least 90° away in azimuth from the electromagnetic cluster, and compare the number of found electrons with the fake

108

Trigger	Expected e	Observed e
"γ-like"	30 ± 5	37
"π^0-like"	73 ± 11	92
All	103 ± 15	129

Table 4.10: Fake rate predictions vs. number of observed events in the photon sample. Systematic errors for the expectations are shown. The normalization is arbitrary.

Sample	Found μ	Expected μ
$\gamma - 70$		
"γ-like"	11	10
"π^0-like"	10	9
$\gamma - 16$		
"γ-like"	128	129
"π^0-like"	232	235

Table 4.11: Same as Table 4.10 except for the muon case.

rate prediction. The results for the photon-16 sample are summarized in Tables 4.10 and 4.11. The statistics in the photon-70 sample were very limited and are not presented here. From this table we note that the photons are systematically high, $\sim 1.5-2\sigma$ with a 15% systematic uncertainty. This systematic uncertainty is deemed to cover the difference.

The agreement between the observed and the expected number of muons is very good everywhere.

Chapter 5

Monte Carlo datasets

Monte Carlo datasets are used in this search to calculate acceptances and estimate certain backgrounds. A brief description of each generator used is followed by details of the actual datasets created. Normalization issues involved in calculating top acceptances as well as systematic effects associated with these predictions are also described.

We have used ISAJET [41] for top quark generation and VECBOS [42] for the kinematics of W+multijet background. The VECBOS Monte Carlo samples are not used in the background calculation at all — most of the background is determined from the data as is explained in Chapter 6. They are used solely for comparison of kinematics of events found in our signal region. The CLEO Monte Carlo [43] decay model is used to describe the decay of b quarks. The HERWIG Monte Carlo generator [44] is also used to calculate systematics of the top acceptance. All Monte Carlo events generated are run through a full detector simulation, QFL [45].

5.1 $t\bar{t}$ generation

We have used ISAJET to generate samples of $p\bar{p} \rightarrow t\bar{t} + X$. The B meson decay in ISAJET is superceded by the superior CLEO Monte Carlo model. In this section we first briefly describe the ISAJET program and then present the details of the datasets generated.

5.1.1 ISAJET

The ISAJET program generates $t\bar{t}$ pairs in four steps:

1. Parton level hard scattering is generated according to the leading order perturbative QCD two-jet cross section, convoluted with structure functions evolved to an appropriate momentum scale, Q^2. Generation efficiency is enhanced by constructing an envelope for the two-jet differential cross section

$$\frac{d\sigma}{dp_T^2 dy_1 dy_2} < F(p_T) \stackrel{\text{def}}{=} Ap_T^{-b} \qquad (5.1)$$

where p_T is the transverse momentum of either jet, and y_1 and y_2 are their rapidities. For each event, $F(p_T)$ is used to generate the jet transverse momentum. Unweighted events are produced by accepting an event if its QCD cross section is greater than a uniform random number between 0.0 and 1.0 times $F(p_T)$.

2. The Fox and Wolfram [46] branching approximation algorithm is then used to develop cascades of initial and final state partons. For the initial state, the Sjöstrand approach [47] is employed — scaling violations in jet fragmentation are implemented, inducing jet broadening and possibly producing additional, resolvable jets. This method only approximately reproduces the matrix elements for widely separated jets.

3. Final state partons undergo Field and Feynman independent jet fragmentation and form hadrons. This ansatz correctly describes the fast hadrons in a jet, but it does not conserve energy-momentum or flavor. ISAJET conserves the center of mass energy, $\sqrt{\hat{s}}$, by boosting all hadrons to the rest frame of the fragmented jets, rescaling the three-momenta and recalculating the energies. Heavy quarks (charm, beauty and top) are fragmented via the Peterson Model (see chapter

111

1), where the Peterson fragmentation parameter, ϵ_Q takes on the values

$$\epsilon_Q = \begin{cases} \frac{0.8 GeV^2}{m_Q^2} & \text{for charm} \\[2ex] \frac{0.5 GeV^2}{m_Q^2} & \text{for bottom and top} \end{cases} \tag{5.2}$$

Here m_Q is the mass of the heavy quark, Q.

4. Beam jets from spectator parton interactions are finally added. A simplified version of a scheme proposed by Abramovskii, Kancheli and Gribov [48] is used to describe minimum bias data. The scheme is modified to account for the experimental observation that spectators interact more strongly in hard scattering events.

5.1.2 $t\bar{t}$ datasets

ISAJET version 6.36 was used to generate $t\bar{t}$ samples for top quark masses of 100, 120, 140, 160 and 180 GeV/c^2. No decays were forced but only final states with at least one lepton above 12 GeV/c were kept. As mentioned earlier, the CLEO Monte Carlo was used to describe the fragmentation and decay of all b quarks. The transverse momentum of the top quarks generated was constrained to be less than 500 GeV/c . The integrated luminosities of the samples are calculated using cross sections calculated to full next-to-leading order in QCD, including corrections dominated by the emission of multiple soft gluons, tabulated in reference [24]. Normalization issues will be revisited in a later section in this chapter. Table 5.1 summarizes the top datasets generated.

5.2 W+jets generation

We have used the VECBOS Monte Carlo to generate datasets of W+1, 2, 3, and 4 jets. As in the previous section, we first describe the basic characteristics of VECBOS and then the datasets generated.

M_{top} GeV/c^2	Number of events generated	Number of events with lepton above 12 GeV/c	$\sigma_{t\bar{t}}$ (pb)	$\int \mathcal{L}dt$ pb^{-1}
100	30,000	12389	102	294
120	30,000	13812	38.9	771
140	30,000	14611	16.9	1775
160	30,000	15492	8.2	3659
180	30,000	16164	4.2	7143

Table 5.1: Monte Carlo generated top data sets.

5.2.1 VECBOS

VECBOS is a parton level Monte Carlo generator that uses exact tree-level matrix elements for the processes it simulates. In the W+jets generation mode VECBOS only produces exclusive final states with a W boson and a fixed number of partons. We have used it to generate 1, 2, 3, and 4 partons in addition to the W boson. This calculation does not cancel the divergences arising from soft and collinear partons since loop diagrams are not included in calculating the matrix elements thereby necessitating some regulatory cuts at the generation level. This fixed-order perturbation calculation also renders the overall normalization to be Q^2−dependent. This results in a rather large uncertainty, $\sim 50\%$, in the cross section normalization. With appropriate regulatory cuts however, the physical distributions can be made independent of loop diagrams. This can be achieved by choosing the regulatory cuts such that they are hidden by the experimental resolution on the transverse momentum and mutual separation of jets. The implementation of VECBOS used in this analysis hadronizes final state partons and adds the underlying event in a manner identical to that for ISAJET described earlier.

VECBOS uses the VEGAS algorithm [49, 50] to calculate a hadronic cross section σ as the convolution integral of the partonic cross section $\hat{\sigma}$ with structure functions F_a, F_b:

$$\sigma(p\bar{p} \rightarrow Wj) = \int dx_1 dx_2 \sum_{a,b} F_a(x_1, Q^2) F_b(x_2, Q^2) \hat{\sigma}(ab \rightarrow Wj) \qquad (5.3)$$

113

Min. P_T of outgoing primary parton	8 GeV
η separation of jets	0.40
Max. η of outgoing primary	3.5
Lepton from W boson decay:	
η cut	1.5
P_T cut	12 GeV/c
QCD scale	$\alpha_s = <P_T>$
Structure functions	MRSB

Table 5.2: Generation level cuts for the VECBOS W+jets datasets.

here j is the number of final state partons generated in addition to the W boson.

In practice an initial run samples the integral and creates a grid of function values to be used as a probability density for subsequent iterations. The net result is the production of weighted events. Before using these events we eliminate their weights using a straightforward procedure. The maximum weight from the initial run is obtained and in subsequent runs events are kept if their weights are greater than a uniform deviate multiplied by the maximum weight.

5.2.2 W+jets datasets

VECBOS was used to generate samples of W+1, 2, 3, and 4 partons in the final state. The generation level cuts employed are listed in Table 5.2. Due to fragmentation and jet clustering and resolution effects the number of jets in the final state may be different from the number of partons generated. Only W decays into electrons were generated. An identical dataset was then made with the electrons being replaced by muons.

In this analysis these datasets only serve the purpose of comparing kinematic variables between data, Monte Carlo top events and the W+jet background. The creation of $W + 4$ parton processes is very CPU intensive and consequently relatively small datasets were created. The characteristics of these datasets are summarized in Table 5.3.

114

$W + N$jets	Number of events generated	Cross section (pb)	$\int \mathcal{L}dt$ pb^{-1}
$W + 1$jet	3478	700	4.97
$W + 2$jets	5050	300	16.8
$W + 3$jets	3231	130	24.9
$W + 4$jets	1020	60	17.0

Table 5.3: W+jets Monte Carlo datasets created using VECBOS. Two identical datasets were made, in one the W was forced to decay to an electron and in the other to a muon.

5.3 Acceptance calculation

The acceptance to top events is calculated using the ISAJET top Monte Carlo datasets. In order to compare theory with the data, this acceptance needs to be corrected for trigger efficiencies and differences between data and Monte Carlo lepton detection efficiencies. We describe the normalization of the e+jets and the μ+jets datasets and also corrections necessary to account for soft leptons in these events.

5.3.1 Normalization of the e+jets Monte Carlo dataset

The same cuts as applied to the data are applied to the Monte Carlo electrons except for the χ^2(strip) cut and the conversion removal. The efficiency of the electron quality cuts for isolated Monte Carlo electrons, obtained from shooting single high-P_T electrons into QFL, is found to be $80 \pm 4\%$. In Chapter 3 a study of $Z \rightarrow e^+e^-$ events resulted in an isolated electron efficiency of $84 \pm 2\%$ excluding the conversion removal procedure. The conversion veto removes roughly 5% of prompt electrons. The Level 2 trigger efficiency was found to be $92.8 \pm 0.3\%$ for high P_T electrons. Therefore, the e+jets Monte Carlo acceptance needs to be corrected by a factor of

$$0.84 \times 0.95 \times 0.928/0.80 = 0.93 \pm 0.046 \tag{5.4}$$

115

This correction factor is appropriate for isolated electrons. Any additional inefficiencies of the implicit (HAD/EM, E/P, L_{shr}) and explicit ($Isol$) isolation requirements for electrons in top events are modeled by the Monte Carlo.

5.3.2 Normalization of the μ+jets Monte Carlo dataset

CMU

The Monte Carlo muons in the region of the CMU/CMP covered by the trigger are subjected to the same quality cuts as muons in the data. The Monte Carlo is then corrected for the muon trigger efficiency, 86.8 ± 1.9%, and the efficiency of the $| \Delta x |< 2$ cm requirement which is 100% efficient in QFL and is found to be 95.2 ± 1.5% efficient for high-P_T muons in CDF. The net correction factor applied to CMU muons in the Monte Carlo is then

$$0.868 \times 0.952 = 0.83 \pm 0.024 \tag{5.5}$$

As in the electron case, any additional inefficiency (e.g due to minimum ionizing requirements) for muons in top events is modeled with the Monte Carlo.

CMX

Since the CMX trigger required a single calorimeter trigger tower above threshold at Level 1, muons detected by the CMX in the Monte Carlo are required to satisfy the simulation of the Level 1 calorimetry trigger. In addition, the muon track is required to pass through Superlayer 8 of the CTC to ensure CFT acceptance. The luminosity weighted Level 2 trigger efficiency is 60.4%, and only 82% of the luminosity is usable. In addition, the $| \Delta x |< 5$ cm matching requirement is 98 ± 0.2% efficient on high-P_T CMX muons in the data and 99 ± 0.3% in QFL. Thus the CMX μ+jets acceptance, after the Level 1 simulation and superlayer requirements, is degraded by

$$0.604 \times 0.82 \times 0.98/0.99 = 0.495 \pm 0.004 \tag{5.6}$$

116

5.3.3 Soft leptons in the Monte Carlo datasets

The top Monte Carlo is also used to calculate the probability that an additional lepton is found in a lepton+jets event.

Low$-P_T$ muons

Additional muons are obtained as follows:

- loop over all CMUO banks in the event

- require that the corresponding track be matched to a generator-level (GENP) muon from bottom, charm, tau, or W decays

- require that all the expected stubs (CMU and/or CMP) be present in the CMUO bank using the algorithm outlined in chapter 4

- degrade the acceptance by the stub-finding efficiencies: 96% for CMP stubs and 98% for CMU stubs

- degrade the acceptance by the efficiency of the quality cuts on the appropriate stubs

This procedure leads to a slight underestimate of the acceptance. If a stub from a real muon is linked by the muon reconstruction code to a wrong track, then the muon is lost in the Monte Carlo, whereas in the data it is still possible that such a badly reconstructed CMUO bank will pass the quality cuts and be counted as a good muon. Based on the efficiency measurements, the systematic uncertainty in the total efficiency for the reconstructing muons in the lepton+jets events is taken to be 5%.

Low$-P_T$ electrons

Additional electrons in the Monte Carlo are treated in the following manner:

- loop over all tracks in an event within the fiducial volume described in chapter 4

- require the track to be associated with a GENP-electron originating from bottom, charm, tau or W decay

- impose the E/P and HAD/EM cuts

- degrade the acceptance by the CES-CPR cut efficiencies obtained in chapter 4

5.3.4 Systematics on the top quark acceptance

Here we discuss the systematic uncertainties on the top acceptance.

Lepton efficiencies

The uncertainty in the high P_T lepton efficiency is estimated to be roughly 5%. The efficiency of the soft muon requirements is measured from J/ψ and Z data, and is found to be 5%. The soft electron efficiency can be broken down into two pieces: (i) the instrumental efficiency of the CES-CPR-dE/dx requirements, which is obtained from a study of conversion electrons and is determined with a statistical accuracy of order 5%, and (ii) the physics dependent efficiency of the implicit isolation requirements (i.e E/P and HAD/EM).

The E/P and HAD/EM efficiencies are taken from Monte Carlo top events. They depend on the physics assumptions (e.g fragmentation , discussed later), and on the simulation of the central calorimeter in QFL. The E/P requirement is $0.7 <$ E/P $<$ 1.5; an electron will fail this requirement if there are high energy photons (e.g from π^0 decay) that deposit energy in the same tower. Since photons are straight-forward to simulate, we expect this inefficiency to be well modeled by QFL. The simulation of HAD/EM in the Monte Carlo has also been verified to be in very good agreement with a study of conversion electrons [51] and a systematic uncertainty of 10% is assigned to it.

Energy scale

To estimate the systematic uncertainty in the top acceptance due to the uncertainty in the jet energy scale, we repeat the Monte Carlo acceptance calculation rescaling

118

Jet energy	Acceptance change				
scale variation	Top 100	Top 120	top 140	Top 160	Top 180
+10 %	−20 %	−10 %	−9 %	−4 %	−3 %
−10 %	+15 %	+11 %	+8 %	+7 %	+3 %

Table 5.4: Systematic uncertainty due to the Jet energy scale.

all the jets in the Monte Carlo by ± 10%. The results are summarized in Table 5.4.

Initial state radiation

The gluon radiation model in ISAJET affects the expected jet multiplicity of top events, and therefore affects the top acceptance. To study the sensitivity of our analysis to the ISAJET modeling of gluon radiation, we have generated and simulated top events (for $M_{top} = 120$ and 140 GeV/c^2) with a modified version of ISAJET in which gluon radiation from the initial state (but not from the final state) was disabled. The resulting change in acceptance, for both top masses, was roughly 15%. We take half of this deviation, 7%, as the systematic uncertainty on the top acceptance due to the modeling of gluon radiation in ISAJET.

b-quark fragmentation

The acceptance to leptons from b-decays may depend on the b-quark fragmentation model used in ISAJET. The default value of the Peterson fragmentation parameter is $\epsilon_b = 0.5/M_b^2$ in ISAJET as mentioned earlier. We have also generated Monte Carlo top events with different values of ϵ_b: $\epsilon_b = 0.65/M_b^2$ and $\epsilon_b = 0.35/M_b^2$, which correspond to 1σ bounds on the Peterson fragmentation parameter as measured at LEP. We did not find any significant variation in the top acceptance within the limited statistics of our Monte Carlo samples (10%).

b and c quark semileptonic branching ratios

The acceptance to top events in this analysis depends on the branching ratio for semileptonic b and c decays (in the Monte Carlo we find that about 50% of leptons

Source		Comments
High P_T lepton efficiency	5%	
Soft μ efficiency	2.5%	5%, but muons contribute 1/2
CES-CPR-dE/dx efficiency	2.5%	5%, but electrons contribute 1/2
HAD/EM cut for soft electrons	5%	10%, but electrons contribute 1/2
Gluon radiation	7%	1/2 of difference ON/OFF
Energy scale	10%	m_{top} dependent ($\sim 3\%$ to 20%)
$b \to e$, $c \to e$ branching ratio	10%	
$b-$fragmentation	10%	
Luminosity	10%	
Total	21%	

Table 5.5: Summary of systematic uncertainties on top acceptance.

come from $b-$decays, about 40% from $c-$decays and the remaining 10% from W and $\tau-$decays). The semileptonic b and c branching ratios are known with an accuracy of about 10% [52], and therefore we assign a systematic uncertainty of 10% to the acceptance calculation.

B decay modeling

In the Monte Carlo, fragmentation and decay of B mesons produced by ISAJET are described by the CLEO Monte Carlo model. We have also calculated the acceptance without using CLEO, by using the ISAJET fragmentation and decay model to describe the decay of B mesons. The two models of b-decays are slightly different in the relative branching ratios and in the charge multiplicity for b-decays. However, when the same semileptonic branching ratios are used, we have verified that the two decay models result in the same top acceptance within the statistics of our Monte Carlo samples, $\pm 10\%$.

Summary of systematic uncertainties on top acceptance

In Table 5.5 we summarize our study of the systematic uncertainties associated with our top acceptance calculation.

120

Chapter 6

Top quark search

In this chapter we describe the results of our search. We begin by defining the final event samples and the signal region. After a discussion of the backgrounds to a lepton+jets top quark signature, we compare our results with these background predictions and expectations from Monte Carlo generated top events. Assuming the excess in the signal region is due solely to $t\bar{t}$ production, we estimate a production cross section for $t\bar{t}$. We end the chapter with a qualitative examination of some kinematic features of events in the signal region.

6.1   The W+jets data sample

Starting with the inclusive lepton samples described in chapter 3, we apply a $\not{E}_T$ cut and remove events consistent with Z boson decay. We count jets with observed E_T > 15 GeV and $\mid \eta_{det} \mid$ < 2.0 clustered with a cone size of 0.4. The $\not{E}_T$ cut removes Drell-Yan, $b\bar{b}$ and fake backgrounds and is ~90% efficient for top events as can be seen from Figure 6-1. We then remove Z° bosons from the data as described in the following section.

Z boson removal

One leg of the Z° decay is required to pass all the high-P_T inclusive lepton sample cuts and a second leg that forms an invariant mass with the first within 20 GeV/c^2
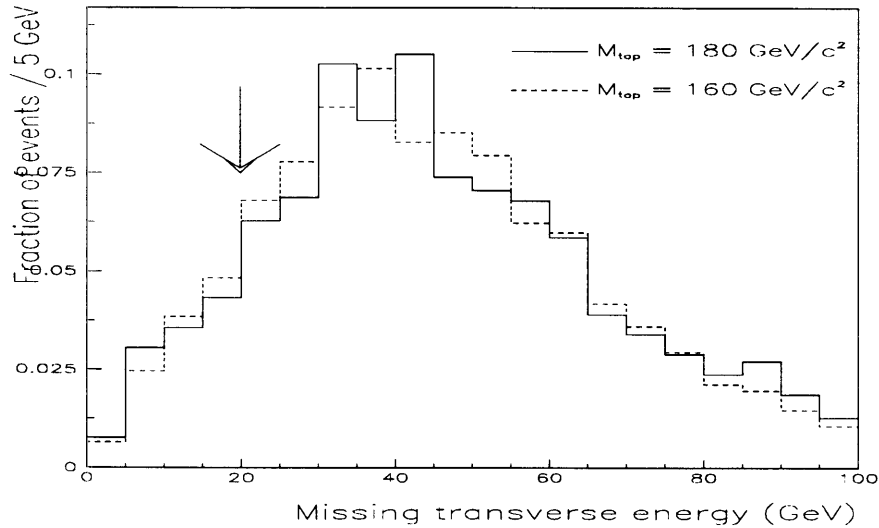
Figure 6-1: The missing transverse energy distribution for the top Monte Carlo. The solid histogram is for a top mass of 180 GeV/c² and the dashed for 160 GeV/c² . The cut value is indicated by an arrow.

of the Z boson mass is also required for removal. Events with leptons of the same species and opposite charge that pass the above requirement are rejected.

In the case of electrons, events in which the second leg passes the cuts

1. $E_T > 10$ GeV

2. E/P < 2.0

3. $HAD/EM < 0.12$

4. $Isol(0.4) < 0.1$

are removed. Similarly, muon events in which the second leg satisfies the criteria

1. $P_T > 15$ GeV/c

2. $| \eta | < 1.1$

3. If CMUO bank present then $| \Delta x | < 5$ cm

are removed. None of the events removed had a jet multiplicity greater than 2 and consequently no events from the signal region (defined later) were lost.

122

The final W+jets selection cuts are summarized in Table 6.1. The transverse mass of the resulting W+jets data is shown in Figure 6-2; a Jacobian peak at the W mass is clearly visible. Table 6.2 shows the W+jets data binned in jet multiplicity together with the corresponding Monte Carlo predictions.

$\not{E}_T$	$>$	20 GeV
P_T of leptons	$>$	20 GeV/c
Z bosons		REMOVED
Jet counting:		
Conesize	$=$	0.4
Observed E_T	$>$	15 GeV
Pseudo-rapidity range	$<$	2.0

Table 6.1: Final cuts for the W+jets sample.
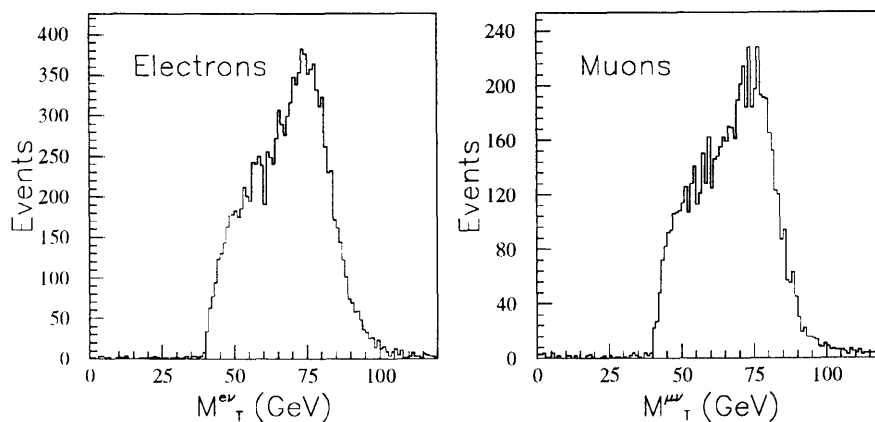
Figure 6-2: The transverse mass distribution in the W sample. Shown here are plots for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ individually. A Jacobian peak at the W mass is clearly visible.

Dilepton removal

In order not to have any overlap with the dilepton search, any events found in our search that satisfy the dilepton selection criteria are removed. These requirements

Jet Multiplicity	Number of data events
0 jets	17219
1 jet	1748
2 jets	283
3 jets	43
≥ 4 jets	8

Table 6.2: Jet multiplicity in the W boson data.

are summarized in Table 6.3

Two opposite charge leptons (e, μ) or (e, e) or (μ, μ) present
Two jets with observed E_T above 10 GeV and $| \eta_d | < 2.4$

Missing transverse energy :
$\not{E}_T > 25$ GeV
$\not{E}_T < 50$ GeV **AND** $(\Delta\phi(\not{E}_T, \text{lepton}) < 20° \ \textbf{OR} \ \Delta\phi(\not{E}_T, \text{jet}) < 20°)$

Table 6.3: The dilepton channel cuts. Events passing cuts listed in this table are explicitly removed from our search.

The signal region definition

At this point we have a final W+jets sample that will provide the basis of our top search. We will look for extra leptons in these events as a sign of the presence of b quarks that are expected from Standard Model decays of the top quark, namely

$$t \rightarrow Wb \tag{6.1}$$

with $\sim 100\%$ branching ratio. Before applying our tagging requirements however, we can further reduce the W+jets background by demanding considerable jet activity in our signal region. Figure 6-3 shows the jet multiplicity in Monte Carlo generated top events. Using this figure and the fact that the W+jets cross section falls rapidly with observed jet multiplicity [53], we can enhance the signal-to-background by requiring at least 3 jets in the event. This requirement is $\sim 75\%$ efficient for top while only $\sim 0.5\%$ of all events in our W+jets sample pass this criterion. Thus we define our

124

signal region to be events in our W+jets sample with at least 3 jets, with jet counting in accordance with the cuts listed in Table 6.1. We then look for extra leptons in these events.
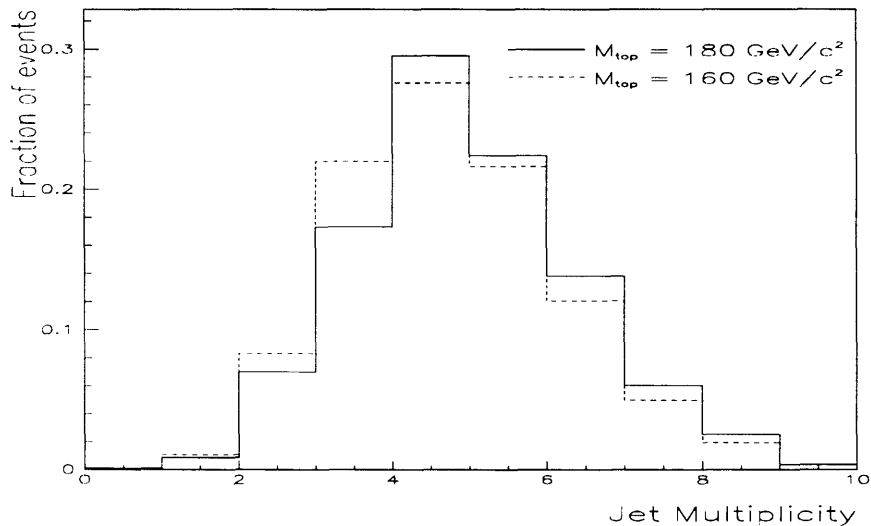
Figure 6-3: The jet multiplicity distribution from Monte Carlo generated top events. The solid histogram is for a top mass of 180 GeV/c² and the dashed for 160 GeV/c²
.

6.2 Extra leptons in signal region

We look for extra leptons as identified by the algorithms detailed in chapter 4. We apply a P_T cut of 2 GeV/c on all soft leptons except for the CMP, which requires muons to have at least a P_T of 3 GeV/c to be detected. We also apply a low invariant mass cut to remove residual Drell-Yan events (explicit Drell-Yan background removal is dicussed later) and cascade decays. Events with two leptons of the same species, primary or soft, whose invariant mass is less than 5 GeV/c² are removed.

Before we present the results of our search, we discuss the various backgrounds to this signal. The background predictions in the next section are after b−tagging algorithms have been applied.

6.3 Backgrounds to the top quark signal

An essential part of any search is the accurate determination of backgrounds to the signal. In this search we look for possible physics sources and misidentifications that would present a signature similar to that of top. Our philosophy is to utilize the data as much as possible, reverting to Monte Carlo only where data cannot be used. The largest background in our channel turns out to be the fake background which is entirely determined from data. A discussion of all the relevant backgrounds together with a summary of the total backgrounds follows.

6.3.1 Fake backgrounds

The fake background determination procedure has been detailed in an earlier chapter. It has also been emphasized that due to the presence of some prompt leptons in the QCD samples, this estimate will necessarily be an over-estimate. It has further been stressed that applying the fake rate probability obtained from jet data to W+jets data assumes that the heavy flavor content of jets in QCD events is the same as it is in W+jets events.

Heavy quarks in jet events are produced through gluon splitting and flavor excitation. In W events, heavy quarks can be produced in the $q \rightarrow W + c$ process, as well as through $W+$ gluon followed by gluon splitting into a pair of heavy quarks. The theoretical expectations for $W + c$ are presented in a later section, and this background is explicitly included in the total background estimate. The contribution due to gluon splitting is automatically included in the fake rate calculation, and it can be argued that it represents an overestimate of the real $W+b\bar{b}$ and $W+c\bar{c}$ background since:

1. The lepton-from-heavy-quark contribution in the fake probability obtained from QCD events includes leptons from the direct and flavor-excitation processes, which are not present in W events

2. Jets in the QCD sample originate mostly from gluons, whereas in the W+jets

126

sample the fraction of the light-quark jets is large. Therefore the gluon-splitting probability per jet is expected to be higher in the QCD sample.

The results of the fake calculation, as a function of jet multiplicity, are shown in Table 6.4. The systematic uncertainties are estimated to be 15% for the fake electron background, and 10% for the fake μ background as described in chapter 4. The two uncertainties are largely uncorrelated.

Jet Multiplicity	Background
1 jet	29.5 ± 3.0
2 jets	9.0 ± 0.9
3 jets	1.93 ± 0.19
≥ 3 jets	2.80 ± 0.28

Table 6.4: Fake backgrounds as a function of jet multiplicity.

6.3.2 Multijet and $b\bar{b}$ backgrounds

Assuming isolation and missing transverse energy are not correlated for QCD and $b\bar{b}$ events, one can extract the fraction (F_{bg}) of non-W events in the W sample as a function of jet multiplicity by extrapolating from the low to the high $\not{E}_T$ regimes. These events will be a mixture of fake lepton + $\not{E}_T$ events, and semileptonic bottom (or charm) decays with high $\not{E}_T$.

Lepton tags in these events will arise from (i) misidentified hadrons, and (ii) real leptons from semileptonic decays of the second b or c quarks in these events. The first type of background is automatically accounted for in the fake rate prediction, but the second needs to be estimated. In order to estimate this background, we first make the assumption that the mixture of QCD and $b\bar{b}$ events is independent of missing energy. We then study the events with low $\not{E}_T$ and low isolation, and for these events we compute the quantity R_{tag}, which is defined as the tagging rate (in excess of fake background prediction) for these events. Then in the sample of W candidates (N_W), the number of lepton-tagged events due to this process will be given by

$$L_b = N_W F_{bg} R_{tag} \qquad (6.2)$$

127

In the W+jets sample, we find F_{bg} roughly 10% independent of jet multiplicity. The statistics of the sample are such that we cannot measure R_{tag} in the ≥ 3 jet bin. In the 1 and 2 jet bins, we find R_{tag} roughly 1%; we assume R_{tag} is independent of the number of jets in the event.

In Table 6.5 we present the expected number of background events due to the $b\bar{b}$ contamination of the W+jets sample as a function of jet multiplicity. The errors shown are our estimate on the systematic uncertainties involved in this background calculation i.e 50% on F_{bg} and 50% on R_{tag} [54].

Jet Multiplicity	$b\bar{b}$ background
1 jet	1.71 ± 1.21
2 jets	0.28 ± 0.20
3 jets	0.04 ± 0.03
≥ 3 jets	0.05 ± 0.04

Table 6.5: The $b\bar{b}$ background to the top quark signal.

6.3.3 Drell-Yan background

Drell-Yan pairs are explicitly removed from the data. We remove any e^+e^- and $\mu^+\mu^-$ pairs where one lepton passes the standard high-P_T lepton identification cuts and the other is identified as a soft lepton by our algorithms and passes the following isolation requirements

$$E_T \ (\text{cone} \ = \ 0.4) \quad < 2 \ \text{GeV} \quad \text{for} \ P_T < 20 \ \text{GeV/c}$$

$$\frac{E_T \ (\text{cone} = 0.4)}{P_T(\text{lepton})} \quad < 0.1 \quad \text{for} \ P_T > 20 \ \text{GeV/c}$$

This requirement removes most Drell-Yan events, as well as a small number of W+fake lepton events, and lepton pairs from $Z \rightarrow \tau\tau$, and diboson production and decay. However, based on the number of same sign pairs, and of $e\mu$ pairs that satisfy these cuts, we conclude that the majority of events removed by this cut are indeed Drell-Yan.

From measuring the isolation efficiency for $Z^\circ \to ee$ (97.5 $\pm$ 0.5%) and $Z^\circ to \mu\mu$ (91.5 $\pm$ 0.8 %), we estimate that the isolation requirement is roughly 95$\pm$3% efficient for leptons from the Drell-Yan process. Therefore, the number of residual Drell-Yan background pairs in our sample can be estimated, as a function of jet multiplicity, by multiplying the number of removed Drell-Yan candidates by 0.05$\pm$0.03. The results are summarized in Table 6.6.

Jet Multiplicity	Number of Drell-Yan pairs removed	Estimated Residual Drell-Yan pairs
1 jet	6	0.30 $\pm$ 0.2
2 jets	1	0.05 $\pm$ 0.05
3 jets	0	< 0.12 90% C.L.
$\geq$ 3 jets	0	< 0.12 90% C.L.

Table 6.6: Explicit Drell-Yan removal and the residual background in data.

6.3.4 Z→ $\tau\tau$ and diboson backgrounds

These backgrounds are calculated using the ISAJET+QFL Monte Carlo described in chapter 5. The diboson backgrounds are normalized to next-to-leading order (NLO) predictions for the production cross sections [55]

$$\sigma(WW) = 9.90 \text{ pb} \tag{6.3}$$

$$\sigma(WZ) = 2.84 \text{ pb} \tag{6.4}$$

$$\sigma(ZZ) = 1.08 \text{ pb} \tag{6.5}$$

The number of expected background tags from these processes, as a function of jet multiplicity, is shown in Table 6.7. Errors are statistical plus systematic, where the statistical error comes from the size of the Monte Carlo sample we have generated, and the systematic errors are as follows:

- $Z \to \tau\tau$: 30% for the ISAJET modeling of the jet multiplicity for intermediate vector boson production [51].

- Dibosons: 30% for the cross section, roughly the difference between the leading-order and the NLO calculation, 30% for the ISAJET modeling of the jet multiplicity in these events, and 10% for the overall luminosity normalization.

Jet Requirement	$Z \rightarrow \tau\tau$	Dibosons
1 jet	$0.67 \pm 0.14 \pm 0.20$	$0.53 \pm 0.10 \pm 0.23$
2 jet	$0.18 \pm 0.07 \pm 0.05$	$0.14 \pm 0.05 \pm 0.06$
3 jet	$0.09 \pm 0.05 \pm 0.03$	$0.02 \pm 0.02 \pm 0.01$
$\geq$ 3 jets	$0.14 \pm 0.04 \pm 0.04$	$0.04 \pm 0.03 \pm 0.02$

Table 6.7: $Z \rightarrow \tau\tau$ and diboson backgrounds.

6.3.5 $W + c$ backgrounds

An additional source of lepton tags in W events arises from semileptonic decays of charm-quarks produced in the process $g + s \rightarrow W + c$, as well as the Cabibbo suppressed equivalent process $g + d \rightarrow W + c$. This background is estimated as follows. The fraction of W events, as a function of jet multiplicity, containing a charm quark (F_{wc}), is obtained by running W Monte Carlos (HERWIG and VECBOS). The uncertainty due to the strange structure function is obtained by varying the input PDF. Then a sample of $W + c$ events is generated using ISAJET + QFL to determine the tagging efficiency (ϵ_c) for this class of events (the tagging efficiency is defined as the number of tagged $W + c$ events divided by the number of $W + c$ events). The number of tags (N_{wc}) in the W sample due to $W + c$ is then given by

$$N_{wc} = 0.9 N_w F_{wc} \epsilon_{wc} (1.11 \pm 0.03) \qquad (6.6)$$

where N_w is the number of W candidates in the data, the factor of 0.9 accounts for the fact that of order 10% of W candidates are really QCD or $b\bar{b}$ events, and the factor of 1.11 is a correction factor to account for the fact that F_{wc} was calculated at the parton level [56]. From the Monte Carlo we find

$$\epsilon_c = 2.0 \pm 0.4\% \qquad (6.7)$$

130

where the errors are statistical only. The results are summarized in Table 6.8.

Jet Multiplicity	N_w	F_{wc}	N_{wc}
1 jet	1748	5.3±1.3%	1.81±0.37±0.45
2 jets	283	7.5±1.5%	0.42±0.08±0.08
3 jets	43	8.0±1.5%	0.07±0.01±0.01
≥3 jets	52	8.0±1.5%	0.08±0.02±0.02

Table 6.8: $W + c$ background summary.

6.3.6 $W+c\bar{c}$ and $W+b\bar{b}$ backgrounds

We have already argued that these backgrounds are already included in the fake calculation. As a check, we also obtain explicit expectations for these backgrounds based on a theoretical model.

The fraction of W+jets events that contain a heavy-quark pair (F_{wbb} and F_{wcc}) is obtained using the HERWIG Monte Carlo. A modified version of HERWIG, forcing $b\bar{b}$ or $c\bar{c}$ pairs in the final state is used to generate this class of events; and the events are then simulated with QFL to determine tagging efficiencies (ϵ_{wbb} and ϵ_{wcc}). The number of expected tagged $W+b\bar{b}$ and $W+c\bar{c}$ events (denoted by N_{wbb} and N_{wcc}) is then given as a function of jet multiplicity by

$$N_{wbb} = 0.9 \times 1.4 \times N_w \times F_{Wbb} \times \epsilon_{wbb} \qquad (6.8)$$

$$N_{wcc} = 0.9 \times 1.4 \times N_w \times F_{Wcc} \times \epsilon_{wcc} \qquad (6.9)$$

where the factor of 0.9 accounts for the non$-W$ contamination of the sample, N_w is the number of W candidates, and the factor of 1.4 is a correction factor [56]. From the Monte Carlo sample we measure:

$\epsilon_{wbb} = 7.5 \pm 1.6\%$

$\epsilon_{wcc} = 5.2 \pm 1.2\%$

where the errors here are statistical only. It should be emphasized that these tagging efficiencies are for the $W+ \geq 1$ jet sample. We do not have enough statistics in our Monte Carlo samples to measure these efficiencies for the higher jet multi-

131

plicity events. Using constant efficiencies as a function of jet multiplicity, we obtain predictions for N_{wbb} and N_{wcc} tabulated in Table 6.9. The errors are statistical only. Although the systematic uncertainties are large, roughly about 50% [57], the number of predicted background events is only a small fraction of the total background.

Jet Multiplicity	N_w	F_{wbb}	F_{wcc}	N_{wbb}	N_{wcc}
1 jet	1748	0.53±0.02%	1.19±0.04%	1.06	1.15
2 jets	283	1.04±0.08%	2.45±0.13%	0.51	0.42
3 jets	43	2.11±0.30%	3.45±0.48%	0.16	0.09
≥ 3 jets	52	2.11±0.30%	3.45±0.48%	0.20	0.11

Table 6.9: $W + b\bar{b}$ and $W + c\bar{c}$ backgrounds. Errors are statistical only.

6.3.7 Summary of backgrounds

A summary of the the backgrounds discussed is presented in Table 6.10. As can be seen from this table, fake tags are the dominant source of background to the signal.

Jet Multiplicity	Fakes	$b\bar{b}$	$Z \to \tau\tau$	dibosons	Drell-Yan	W+c	Total
1 jet	29.5±3.0	1.7±1.2	0.67±0.24	0.53±0.25	0.30±0.20	1.8±0.6	34.5±3.3
2 jets	9.0±0.9	.28±.20	0.18±0.09	0.14±0.08	0.05±0.05	.42±.11	10.1±1.0
3 jets	1.93±.18	.04±.03	0.09±0.06	0.02±0.02	0.06±0.06	.07±.02	2.2 ±0.2
≥ 3 jets	2.80±.28	.05±.03	0.14±0.06	0.04±0.03	0.06±0.06	.08±.03	3.1 ±0.3

Table 6.10: Summary of backgrounds.

6.4 Results

We now present the results of applying our soft lepton tagging requirements to the W+jets data. The results together with the expectations from Monte Carlo generated top events as well as the total backgrounds from the above section are displayed in Table 6.11. The top acceptances were determined in the previous chapter.

We observe 7 events in the signal region with a background expectation of 3.1 ± 0.3 events. This excess is too small to claim with sufficient confidence the existence

132

$P_T > 2$ GeV/c			Top mass in GeV/c^2				
Jet multiplicity	Data	Background prediction	100	120	140	160	180
1 jet	33	34.5±3.3	8.14±2.60	2.06±0.78	0.49±0.19	0.17±0.06	0.07±0.02
2 jets	12	10.1±1.0	8.56±1.88	4.04±0.89	1.59±0.35	0.70±0.17	0.29±0.07
3 jets	5	2.2±0.2	6.00±1.32	4.26±0.81	2.19±0.42	1.14±0.22	0.48±0.10
≥ 3 jets	7	3.1±0.3	9.66±2.70	7.57±1.67	4.14±0.91	2.24±0.45	1.24±0.24

Table 6.11: Results of the search. Tags, background predictions and top expectations are tabulated as a function of jet multiplicity.

of a new source of b quarks in W events beyond expectations from generic heavy flavor production. However, we examine these events to see whether their kinematic behavior is suggestive of the onset of top quark production.

6.4.1 $t\bar{t}$ production cross section calculation

We first calculate the $t\bar{t}$ production cross section assuming all the excess in the signal region is due to $t\bar{t}$ production alone. We then examine the kinematics of these events to see how well the data fits the hypotheses that either the signal region is composed of W+jets background only, or that a fraction of it consists of top quark signal and the rest W+jets background.

Method

The $t\bar{t}$ production cross section is calculated using a likelihood function. Using the number of expected events, N_{exp}, defined by

$$N_{exp} = \epsilon \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L} dt + b \qquad (6.10)$$

and the gaussian formulation

$$G(x, \bar{x}, \sigma_x) = e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}} \qquad (6.11)$$

133

	Top mass in GeV/c^2				
	100	120	140	160	180
acceptance, ϵ (%)	0.44±0.12	0.91±0.18	1.14±0.22	1.28±0.22	1.38±0.23
cross section, $\sigma_{t\bar{t}}$, in pb	$41.42^{+27.51}_{-23.49}$	$20.03^{+18.04}_{-11.96}$	$15.99^{+16.30}_{-9.52}$	$14.24^{+14.77}_{-8.57}$	$13.21^{+13.48}_{-7.90}$

Table 6.12: Results of the $t\bar{t}$ cross section calculation. The number of observed events in the signal region is 7 with an expected background of 3.1 ± 0.3. The integrated luminosity is 21.4 pb^{-1} . Details on the top acceptance calculation can be found in chapter 5.

we can write the likelihood function as

$$L = G(\mathcal{L}, \overline{\mathcal{L}}, \sigma_{\mathcal{L}}) \cdot G(\epsilon, \bar{\epsilon}, \sigma_\epsilon) \cdot G(b, \bar{b}, \sigma_b) \cdot \left[\frac{N_{exp}^n}{n!} e^{-N_{exp}} \right] \qquad (6.12)$$

where $\bar{\epsilon}$ is the total acceptance, $\bar{b}$ is the expected background, n is the number of observed events, $\sigma_{t\bar{t}}$ is the $t\bar{t}$ production cross section and $\overline{\int \mathcal{L} dt} = 21.4$ pb^{-1} is the integrated luminosity. The $t\bar{t}$ cross section is obtained by maximizing the natural logarithm of this likelihood function. The central value obtained from maximizing $\ln L$ is simply given by

$$\sigma_{t\bar{t}} = \frac{n - \bar{b}}{\bar{\epsilon} \cdot \overline{\int \mathcal{L} dt}} \qquad (6.13)$$

as is to be expected. The uncertainties on the cross section are obtained from the condition $\Delta(\ln L) = \frac{1}{2}$. The MINUIT package [58] is used to calculate these uncertainties.

Results

Table 6.12 lists all the ingredients in the calculation as well as the final results. The acceptance calculation was described in chapter 5 and only the results are presented here. In Figure 6-4 we superimpose our cross section calculation on the theoretical predictions from reference [24]. From this plot we note that the observed excess in the signal region, if attributed to exclusively to $t\bar{t}$ production, corresponds to a top quark mass in the range 125−195 GeV/c^2 at the level of ± one standard deviation, with a central value of about 150 GeV/c^2 .
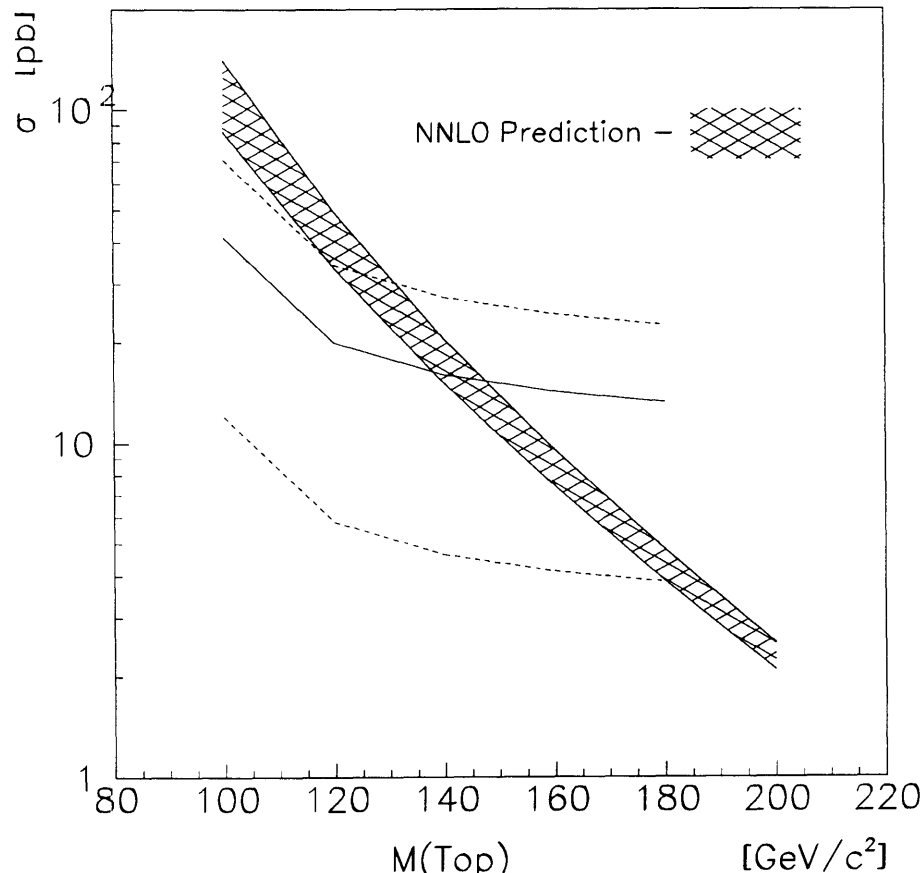
Figure 6-4: The $t\bar{t}$ production cross section. The band represents the Next-to-Next-to-Leading-Order (NNLO) theoretical prediction from [25]. The points joined by a solid line correspond to the cross section derived from the excess of lepton tags in the signal region assuming the excess is solely from $t\bar{t}$ production. The dashed lines represent the 1σ uncertainty on this measurement.

6.4.2 Kinematic analysis

We now compare simple kinematic features of the events in the signal region with expectations from Monte Carlo top events and the dominant background Monte Carlo, W+jets. Numerous studies performed at CDF have tried to extract a top quark signal based on kinematic comparisions of data with W+jets Monte Carlo alone. Due to the uncertainties in the Monte Carlos for W+multijet production currently available, primarily in normalization, strong conclusions could not be drawn from any of these studies. The intention of the comparison presented here is then not to draw any

135

quantitative conclusions from these plots, but rather to get some understanding of the nature of these events. The quantitative conclusions from this analysis have already been drawn and are summarized in a later chapter.

We first show that the W+jets Monte Carlo does a reasonable job in predicting the kinematics of these events. Figure 6-5 shows some basic plots comparing the VECBOS predictions with W+jets data from the current run. In order to minimize possible top quark contribution in the data, we have explicitly removed all events that either had a $b-$tag or passed the dilepton requirements. From these plots it can be seen that the Monte Carlo does indeed do a reasonable job in predicting the general kinematic features of W+jets events. We then turn to our signal region and compare kinematic quantities that might discriminate signal from background. It should be emphasised that $\sim 10\%$ of the background is from sources other than W+multijet production and is hence not included in the comparisons presented.

We begin our comparison with a quantity that measures the jet activity in these events. Figure 6-6(a) shows the $SUMET$ distribution defined as

$$SUMET \stackrel{\text{def}}{=} \sum_j E_T(j) \tag{6.14}$$

where the sum is over all jets of cone 0.4, with E_T greater than 10 GeV and $\mid \eta \mid <$ 2.0. In this as well as all plots to follow, arrows represent the tagged sample, dots with error bars the *untagged* data, dashed-line histograms the Monte Carlo generated W+jets events, and solid-line histograms Monte Carlo generated top events of mass 160 GeV/c^2 .

Figure 6-6(b) shows the transverse mass, M_{trans} defined as

$$M^2_{trans} \stackrel{\text{def}}{=} [E_T(e) + E_T(\not{E}_T) + \sum E_T(j)]^2 - [P_x(e) + P_x(\not{E}_T) + \sum P_x(j)]^2 - [P_y(e) + P_y(\not{E}_T) + \sum P_y(j)]^2 \tag{6.15}$$

With the use of these two kinematic distributions we proceed to test the two hypotheses mentioned earlier:

1. **Hypothesis I:** The data is composed entirely of W+jets background.

136
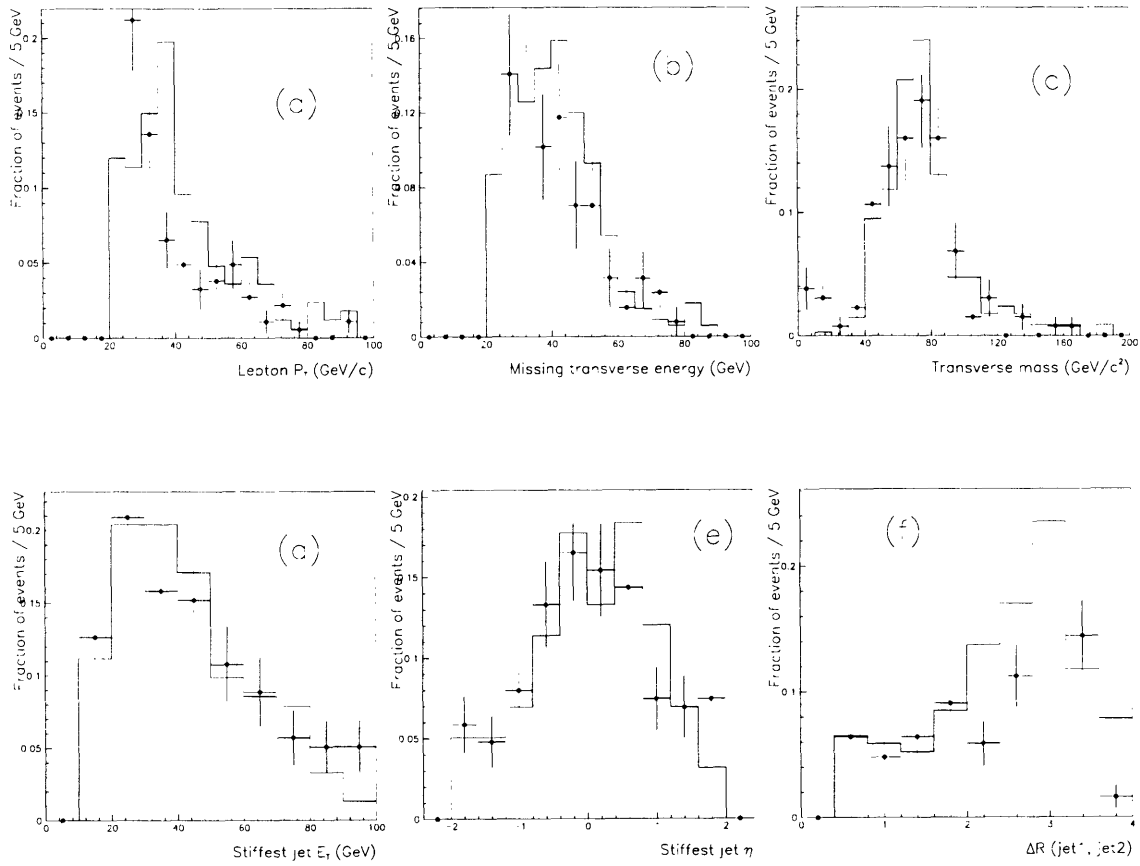
Figure 6-5: A comparison of simple kinematic distributions for W+jets data (circles) and VECBOS Monte Carlo predictions (histogram). (a) the lepton P_T , (b) the missing transverse energy, (c) the transverse mass of the lepton and neutrino, (d) E_T of the highest E_T jet, (e) η for the highest E_T jet and (f) ΔR between the highest two E_T jets.

2. **Hypothesis II**: The data consists of top quark signal together with some W+jets background contamination.

We construct likelihood functions for the two hypotheses and compute the ratio of this likelihood for each of the kinematic variables, $SUMET$ and M_{trans}, for each hypothesis. We use a likelihood function of the form

$$F(\alpha, x) = [\alpha f_B(x) + (1 - \alpha) f_t(x)] \qquad (6.16)$$

where α is the background fraction, x is the kinematic variable of interest, f_B is the
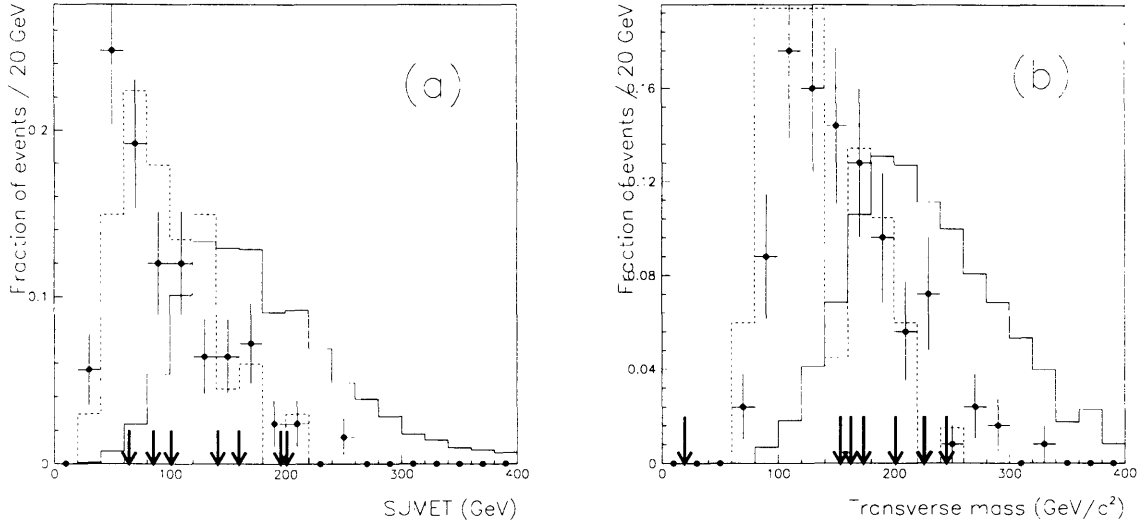
Figure 6-6: Some kinematic distributions for events in the signal region (arrows). Also shown are the expectations from W+jets Monte Carlo (dashed histograms) Monte Carlo top events of mass 160 GeV/c^2 (solid histograms) and the W+jets *untagged* sample (dots). (a) $SUMET$, and (b) the transverse mass, M_{trans}. Note that the *untagged* data is in very good agreement with VECBOS predictions. In (b) the data point at ~ 20 GeV/c^2 is not included in the likelihood calculation and could be due to fluctuations in the W+jets background or some other background (e.g. $b\bar{b}$) not included in these plots.

W+jets background distribution for the kinematic variable x, and f_t is the top quark distribution for the kinematic variable x. We then calculate $\ln L$ as a sum over all the events in the signal region

$$\ln L = \sum_{i=1}^{7} \ln(F(\alpha, x)) \tag{6.17}$$

Finally, for each kinematic variable we compute the ratio of the likelihoods, R_t, for the two hypotheses

$$R_t = \frac{L(x, \text{hypothesis II})}{L(x, \text{hypothesis I})} \tag{6.18}$$

In this formalism, hypothesis I corresponds to $\alpha = 1.0$ and in hypothesis II $\alpha = \frac{3.1}{7}$ where we have used the background estimate from our SLT search to normalize the background distribution function. In computing f_B we have added the distributions from both $W + 3$jets and $W + 4$jets VECBOS Monte Carlo samples in the ratio of

138

Kinematic variable	R_t
$SUMET$	7.3
M_{trans}	33.3

Table 6.13: Ratio of the top signal plus background hypothesis and the entirely background hypothesis, in the two kinematic variables investigated. R_t is defined in the text.

their measured production cross sections [53].

In Table 6.13 we present the results of our determination of R_t for the $SUMET$ and M_{trans} kinematic variables. From this table it can be seen that hypothesis II, corresponding to a mixture of top quark signal and W+jets background, is more probable than the hypothesis that the events in our signal region consist entirely of background. Whereas this result suggests a top quark contribution to the signal region, it does not provide strong enough evidence to constitute top quark discovery. Clearly the statistics are insufficient and much more data will have to be collected before any firm quantitative conclusions can be drawn.

6.5   Summary

We have searched for the processes

$$p\bar{p} \rightarrow t\bar{t}X \rightarrow e\nu_e b\bar{b}q\bar{q} \tag{6.19}$$

$$p\bar{p} \rightarrow t\bar{t}X \rightarrow \mu\nu_\mu b\bar{b}q\bar{q} \tag{6.20}$$

in 21.4 pb^{-1} of data collected by the CDF detector during the 1992-93 Tevatron Collider Run. Events were selected by requiring an isolated high P_T electron or muon, significant missing energy and three or more jets. The b—quarks in the final state were further tagged through their semileptonic decays. An extensive background investigation was performed including both physics backgrounds ($b\bar{b}$, Z$\rightarrow \tau\tau$, Drell-Yan, $Wb\bar{b}$, $Wc\bar{c}$, Wc) and misidentifications.

The search yields 7 events with an expected background of 3.1 $\pm$ 0.3 events. While

the kinematic behavior of these events is suggestive of a top quark signal, the limited statistics of this experiment do not permit any definitive conclusion. However, this result combined with the searches conducted by the CDF collaboration in the other channels could well result in evidence for $t\bar{t}$ production.

With a three-fold increase in statistics as is expected from the 1993−94 Collider Run, this channel would be well suited to discovering a moderately heavy top quark.

Appendix A

Consistency checks

A.1 e+jets versus μ+jets

To verify that the e+jets data are consistent with the μ+jets data we perform the analysis separately for the two data samples. In Table A.1 we present the fake prediction and the number of found tags as a function of jet multiplicity in the e+jets and μ+jets sample respectively. From these tables we note that the e+jets data is consistent with the μ+jets data.

Soft lepton $P_T > 2$ GeV/c						
	Lepton+jets		Electron+jets		Muon+jets	
Jet multiplicity	Fake prediction	Tags	Fake prediction	Tags	Fake prediction	Tags
1 jet	29.7±3.0	33	17.7±1.8	21	12.1±1.2	12
2 jets	9.2±0.9	12	6.4±0.6	10	2.9±0.3	2
3 jets	1.83±.18	5	1.29±.13	3	0.53±.24	2
≥ 3 jets	2.55±.26	7	1.85±.19	3	0.70±.34	4

Table A.1: The lepton+jets data broken down into electron+jets and muon+jets components.

Soft lepton $P_T > 2$ GeV/c						
Jet multiplicity	Fake prediction ($e + \mu$)	Tags ($e + \mu$)	Fake prediction (e only)	Tags (e only)	Fake prediction (μ only)	Tags (μ only)
1 jet	29.7±3.0	33	10.0±1.5	16	19.8±2.0	17
2 jets	9.2±0.9	12	3.0±0.4	3	6.3±0.6	9
3 jets	1.83±.18	5	0.64±.10	2	1.19±.12	3
$\geq$ 3 jets	2.55±.26	7	0.82±.12	4	1.73±.17	3

Table A.2: The lepton+jets data split up into electron tags and muon tags.

Soft lepton $P_T > 2$ GeV/c				
	Runs > 45000		Runs < 45000	
Jet multiplicity	Fake prediction	Tags	Fake prediction	Tags
1 jet	15.8±1.6	23	13.9±1.4	10
2 jets	4.7±0.5	5	4.5±0.5	7
3 jets	0.8 ±0.1	2	1.0 ±.10	3
$\geq$ 3 jets	1.2 ±0.1	3	1.4 ±.1	4

Table A.3: Run number dependence of the lepton+jets data.

A.2 Electron tags versus muon tags

Next we split the tags into electron and muon tags. The results are shown in Table A.2. In the 1 jet bin there seems to be a slight excess (at the 1.5σ level) of electron tags.

A.3 Run dependence

Finally we split the data by run-number to check for any time dependent effects. The results of the analysis for runs > 45000 and runs < 45000 are shown in Table A.3. We observe no statistically significant run dependence.

142

Soft lepton $P_T > 4$ GeV/c						
Jet multiplicity	Lepton+jets		Electron+jets		Muon+jets	
	Fake prediction	Tags	Fake prediction	Tags	Fake prediction	Tags
1 jet	11.9±1.2	12	7.1±0.7	8	4.8±0.5	4
2 jets	4.1±0.4	7	2.9±0.3	6	1.8±0.2	1
3 jets	0.88±.09	2	0.64±.06	1	0.24±.03	1
≥ 3 jets	1.26±.13	3	0.92±.09	1	0.34±.03	2

Table A.4: The electron and muon components of the lepton+jets data with a 4 GeV/c P_T cut on the soft lepton tags.

Soft lepton $P_T > 4$ GeV/c						
Jet multiplicity	Fake prediction ($e + \mu$)	Tags ($e + \mu$)	Fake prediction (e only)	Tags (e only)	Fake prediction (μ only)	Tags (μ only)
1 jet	11.9±1.2	12	4.1±0.6	7	7.8±1.2	5
2 jets	4.1±0.4	7	1.3±0.2	1	2.8±0.4	6
3 jets	0.88±.09	2	0.30±.05	2	0.59±.09	0
≥ 3 jets	1.26±.13	3	0.40±.06	2	0.86±.13	1

Table A.5: The eletron and muon tags in the lepton+jets data. A P_T cut of 4 GeV/c has been applied on the data.

A.4 Higher P_T cut on the soft lepton

Finally we repeat all three checks just described with a P_T cut of 4 GeV/c on the soft lepton. From Tables A.4, A.5, and A.6 we conclude that the 2 GeV/c cut on the soft leptons does not bias the results.

Soft lepton $P_T > 4$ GeV/c				
	Runs > 45000		Runs < 45000	
Jet multiplicity	Fake prediction	Tags	Fake prediction	Tags
1 jet	6.4±0.6	8	5.5±0.5	4
2 jets	2.0±0.2	4	2.1±0.1	3
3 jets	0.38±.04	1	0.50±.05	1
≥ 3 jets	0.64±.06	2	0.62±.06	1

Table A.6: Time dependence of the lepton+jets data with a 4 GeV/c cut on the soft lepton.

Appendix B

Description of tagged events in the signal region

B.1   List of events

Table B.1 lists all the tagged events in the signal region together with their characterising features. This table also includes events that did not fire the appropriate trigger or for some other reason were not included in the analysis as is explained in the next section.

Run, Event	Primary lepton	Lepton tag, P_T (GeV/c)	Jet multiplicity	Comments
42548, 143286	Electron	μ, 2.2	3	
45705, 54765	Electron	e, 11.0	3	
45880, 31838	Electron	e, 2.2	4	
42517, 44047	Muon	e, 4.2	3	
43351, 266423	Muon	μ, 2.1	3	
45047, 104393	Muon	e, 22.6	3	
45879, 123158	Muon	μ, 13.5	4	
46818, 221912	Muon	μ, 10.5	3	CMX muon, did not fire CMX trigger
45178, 382599	Muon	e, 2.6	3	CMP−only muon, not triggered on

Table B.1: Characteristics of tagged events in the signal region.

B.1.1 Relaxing the trigger requirement

The analysis described only considered events which were explicitly required to pass the Level 2 lepton triggers. This requirement enabled proper normalization of the Monte Carlo top event samples. Another reason for requiring a well-defined trigger path was that events which failed the inclusive lepton trigger and contained an additional lepton, whether real or fake, are likely to have a higher probability of firing a different trigger[1]. Hence the fake calculation for this class of events may have been underestimated due to this trigger bias.

Except for the CMX trigger, all the inclusive muon and electron triggers are very efficient for high P_T leptons (see chapter 3). So requiring the trigger path does not result in a significant loss of acceptance. There are, however, a small number of events that make it into our lepton+jets sample through a different trigger path (mostly some variation of the $\not{E}_T$ trigger path). Table B.2 summarizes the jet multiplicity and the number of tags in each jet bin that do not come in on the required trigger path.

Jet Multiplicity	Number of events	Tags,P_T >2 GeV/c	Tags,P_T >4 GeV/c
1 jet	353	10	7
2 jets	44	1	0
3 jets	8	1	1
$\geq$ 3 jets	10	2	1

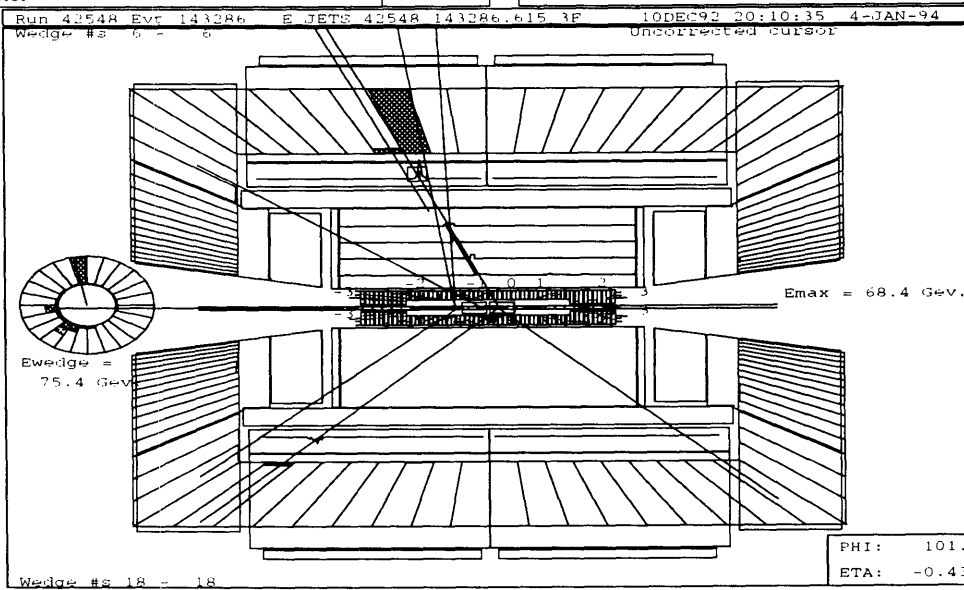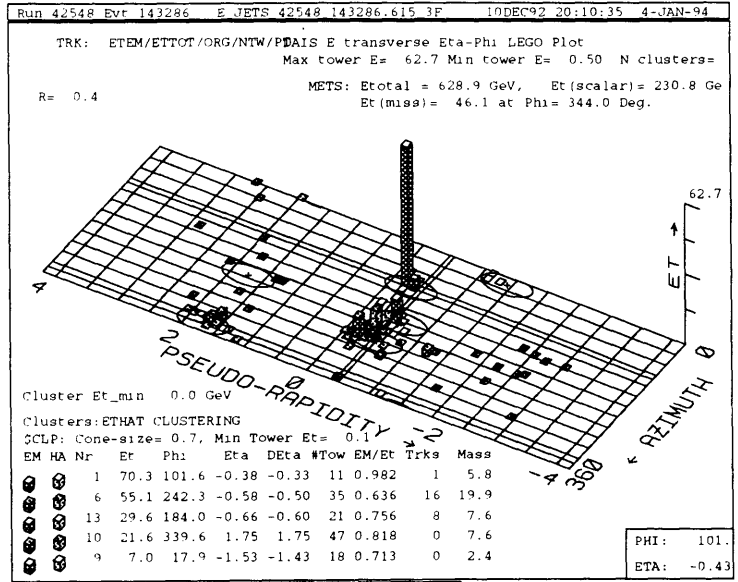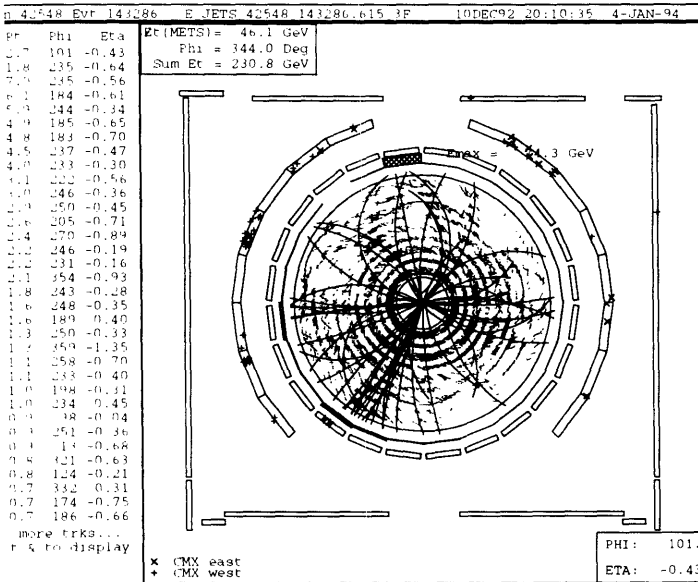Table B.2: Events that came in on alternate trigger paths.

The two events in the signal region, a CMX muon event and a CMP−only one, are included in the table in the previous section.

---

[1]e.g the dilepton trigger

146

B.2 CDF Event Display pictures of candidates

Following is a set of Event Display (DF) pictures showing the candidate events in CDF. In the top left hand corner is the CTC display together with energy deposited in the calorimeters and hits in the muon chambers (CMU/CMP/CMX). On the top right hand corner is a lego display of the transverse energy deposited in the calorimeters. At the bottom is a side view through CDF with reconstructed stiff tracks, energy deposition in calorimeter towers and also the event vertex(ices) as reconstructed in the VTX.

Run=45705, Event=54765

Run=45880, Event=31838

150

Run 42517 Event44047   9DEC92 8:43:31 4-JAN-94

Et(METS)= 64.0 GeV
Phi = 31.3 Deg
Sum Et = 122.5 GeV

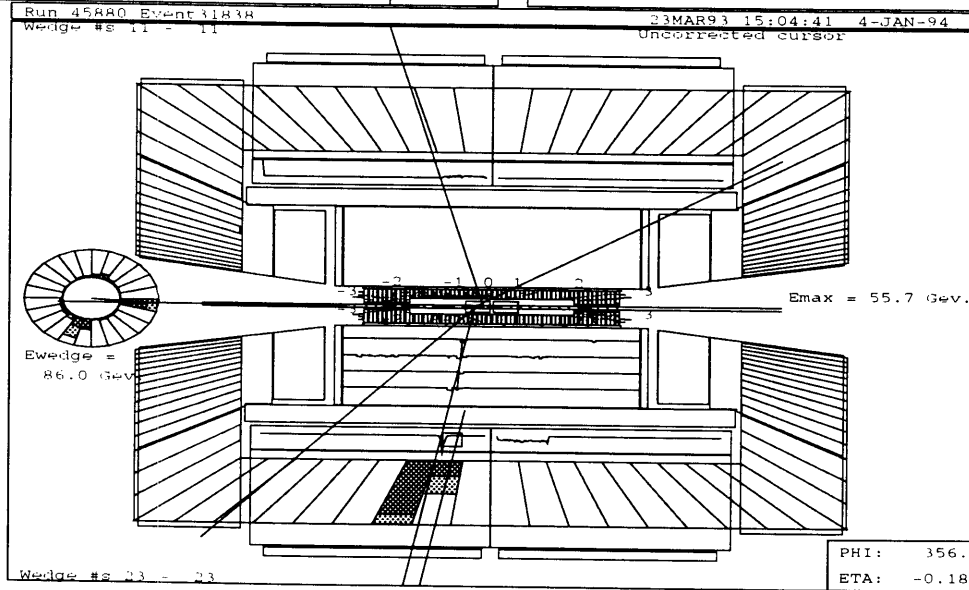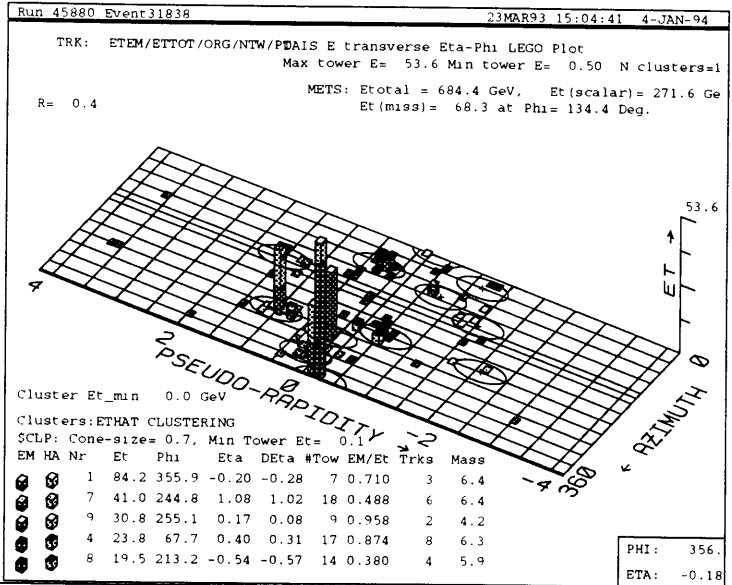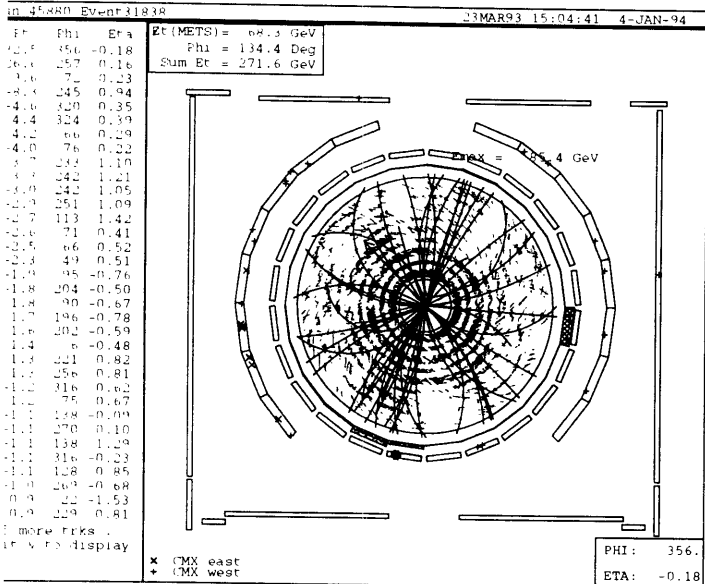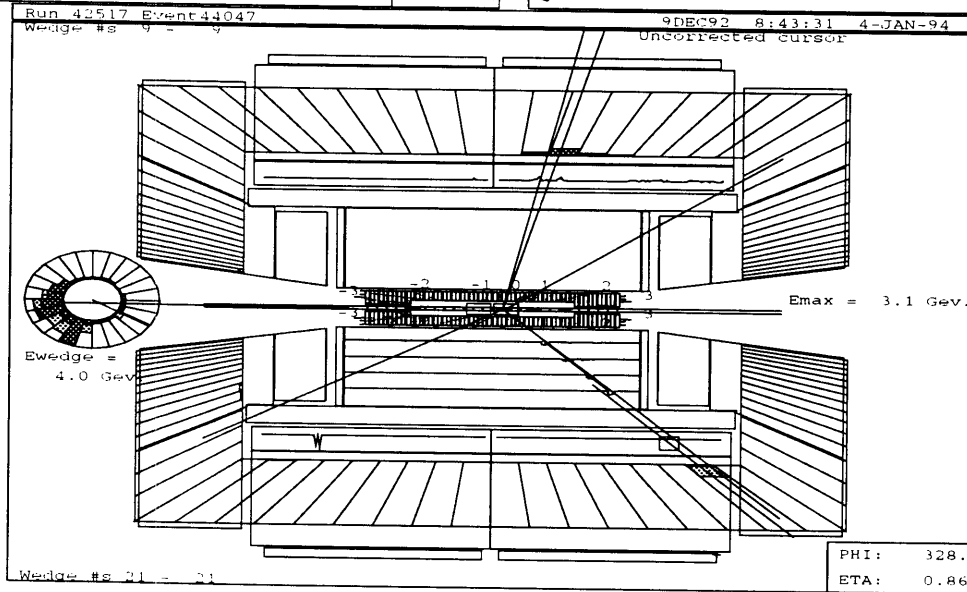| Pt	Phi	Eta
45.8	328	0.86
5.0	227	0.02
-4.2	209	0.32
-4.0	224	0.10
3.4	223	0.12
3.7	180	0.39
3.7	172	0.45
3.4	182	0.39
3.4	257	-0.57
-2.5	177	0.36
2.4	173	0.82
-2.3	253	-0.65
-2.2	227	-0.95
-2.1	168	-1.18
2.0	249	-0.69
-1.8	163	0.49
1.7	240	-0.69
1.7	135	0.20
1.6	220	0.18
1.2	255	0.19
-1.1	213	-0.09
1.0	213	-0.52
0.9	301	0.90
-0.9	140	0.25
0.9	262	0.33
-0.9	19	0.25
-0.8	175	0.57
-0.7	238	-0.92
-0.7	388	-1.00
0.6	9	-1.64
0.6	101	0.93
0.6	21	-0.41
-0.6	276	0.90

1 more trks...
It & to display

× CMX east
+ CMX west

Emax = 6.7 GeV

PHI: 328.
ETA: 0.86

Run 42517 Event44047   9DEC92 8:43:31 4-JAN-94

TRK: ETEM/ETTOT/ORG/NTW/PDAIS E transverse Eta-Phi LEGO Plot

Max tower E= 6.6 Min tower E= 0.50 N clusters=

METS: Etotal = 469.9 GeV, Et(scalar)= 122.5 Ge
Et(miss)= 64.0 at Phi= 31.3 Deg.

R= 0.4

ET 6.6

Cluster Et_min 0.0 GeV
Clusters:ETHAT CLUSTERING
$CLP: Cone-size= 0.7, Min Tower Et= 0.1

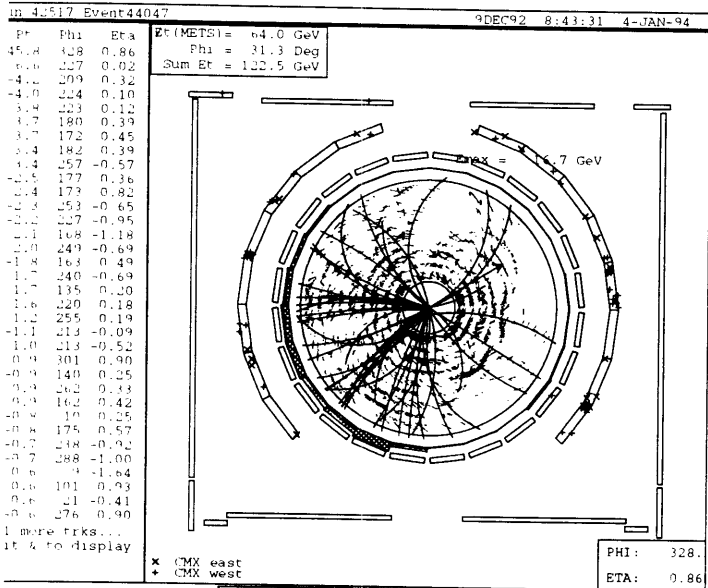| EM	HA	Nr	Et	Phi	Eta	DEta	#Tow	EM/Et	Trks	Mass
		8	50.6	197.0	0.36	0.37	31	0.685	15	23.5
		10	22.2	246.9	-0.73	-0.66	19	0.724	9	6.3
		5	6.0	168.7	-2.41	-2.33	17	0.294	0	1.7
		11	4.4	357.8	-1.85	-1.74	20	0.765	0	1.7
		7	4.6	337.9	0.67	0.68	5	0.378	2	1.5

PHI: 328.
ETA: 0.86

Run 42517 Event44047   9DEC92 8:43:31 4-JAN-94

Uncorrected cursor

Emax = 3.1 Gev.

Ewedge = 4.0 Gev

PHI: 328.
ETA: 0.86

Run=46818, Event=221912

Run 45178 Evt 382599   3MAR93 3:02:54 4-JAN-94

TRK: ETEM/ETTOT/ORG/NTW/PDAIS E transverse Eta-Phi LEGO Plot

Max tower E= 26.2 Min tower E= 0.50 N clusters=

METS: Etotal = 424.1 GeV, Et(scalar)= 162.8 Ge
Et(miss)= 17.0 at Phi= 76.3 Deg.

R= 0.4

Cluster Et_min   0.0 GeV

Clusters:ETHAT CLUSTERING

$CLP: Cone-size= 0.7, Min Tower Et= 0.1

EM	HA	Nr	Et	Phi	Eta	DEta	#Tow	EM/Et	Trks	Mass
		9	56.3	192.9	1.26	1.25	38	0.868	2	13.8
		8	43.7	354.3	0.58	0.57	14	0.637	7	11.8
		3	21.7	312.3	-0.96	-0.94	19	0.592	10	4.6
		4	14.1	120.9	1.10	1.05	20	0.507	3	5.6
		7	6.2	120.7	-1.14	-1.10	8	0.558	2	1.9

PHI: 29.
ETA: -0.16

Et(METS)=   17.0 GeV
Phi =   76.3 Deg
Sum Et = 162.8 GeV

Pt	Phi	Eta
4.5	29	-0.16
8.1	347	0.71
8.0	192	0.49
6.5	316	-1.00
6.3	308	-1.06
5.3	310	-1.02
5.0	312	-0.90
4.9	193	1.30
4.5	357	0.45
4.4	112	-1.28
3.5	118	0.86
3.5	341	-0.79
3.2	136	1.30
3.2	99	1.15
3.1	354	0.66
2.9	310	-0.96
2.8	350	0.68
1.6	336	-0.48
1.6	206	1.34
1.6	355	0.62
1.5	190	-0.67
1.5	352	0.43
1.4	405	-0.95
1.4	140	-1.09
1.2	312	-1.04
1.2	216	1.14
1.0	215	1.43
0.9	221	1.02
0.9	353	0.28
0.8	131	0.78
0.8	159	-0.77
0.8	170	-0.60
0.7	334	0.54
0.7	316	0.53

more trks...
F & to display

× CMX east
+ CMX west

PHI: 29.
ETA: -0.16

Run 45178 Evt 382599   3MAR93 3:02:54 4-JAN-94

Wedge #s 1 - 1   Uncorrected cursor

Emax = 25.5 GeV.

Ewedge =
3.0 GeV

Wedge #s 13 - 13

PHI: 29.
ETA: -0.16

Bibliography

[1] See for example, F. Abe *et al*, "Top dilepton search at CDF," Fermilab Preprint FERMILAB-CONF-93-212-E, Aug 1993. Submitted to the XVI International Symposium on Lepton and Photon Interactions, Ithaca, NY, 10-15 Aug 1993

[2] P. Tipton *et al*, CDF collaboration, Proceedings of the XVI International Symposium on Lepton−Photon Interactions, Ithaca, New York, 1993

[3] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967)

[4] S.L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberger, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in *Elementary Particle Theory*, edited by N. Svartholm (Almqvist and Wiksell, Sweden, 1968), p. 367

[5] S.L. Glashow, J. Illiopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970); M. Kobayashi and M. Maskawa, Prog. Theor. Phys. **49**, 652 (1973)

[6] R. Marshall," A determination of $\sin^2\theta_W$ and the lepton and quark electroweak couplings from e^+e^- data", Z. Phys. **C43**, 607 (1989)

[7] G. L. Kane, M.E. Peskin, "A constraint from B decay on models with no t quark," Nucl. Phys. **B195**, 29 (1982)

[8] A. Bean *et al*, "Improved upper limit on flavor-changing neutral current decays of the b quark," Phys. Rev. **D35**, 3533 (1987)

[9] H. Albretch *et al*, "Observation of $B^0 - \overline{B}^0$ mixing," Phys. Lett. **B192**, 245 (1987)

[10] D.P. Roy and S.U. Sankar, "Upper limit on $B_d{}^0 - \overline{B_d}{}^0$ mixing as evidence for the existence of the top quark," Phys. Lett. **B243**, 296 (1990)

[11] G. L. Kane, "Top quark topics" in *Proceedings of the eighteenth SLAC Summer Institute on Particle Physics, July 1990*, Edited by J. Hawthorne, SLAC-report-278, p. 123

[12] D. Decamp *et al*, "A search for new quarks and leptons from Z^0 decay at LEP," Phys. Lett. **B236**, 511 (1990)

[13] M.Z. Akrawy *et al*, "A search for the top and b' quarks in hadronic Z^0 decays," Phys. Lett. **B236**, 364 (1990)

[14] M.Z. Abreu *et al*, "Search for the t and b' quarks in hadronic decays of the Z^0 boson," Phys. Lett. **B242**, 536 (1990)

[15] W. Hollick, Proceedings of the XVI International Symposium on Lepton-Photon Interactions, 10-15 August 1993, to be published

[16] T. Åkesson *et al*, "Search for top quark production at the CERN $p\bar{p}$ collider," Z. Phys. **C46**, 179 (1990)

[17] F. Abe *et al*, Phys. Rev. Lett. **64**, 147 (1990)

[18] F. Abe *et al*, Phys. Rev. Lett. **68**, 147 (1992); F. Abe *et al*, Phys. Rev. **D45**, 3921 (1992)

[19] N. Cabibbo, III Topical Conference on Proton-Antiproton Collider Physics, Rome, Jan. 1983; F. Halzen and M. Marsula, Phys. Rev. Lett. **51**, 857 (1983); K. Hikasa, Phys. Rev. **D29**, 1939 (1984); N.G. Deshpande *et al*, Phys. Rev. Lett. **54**, 1757 (1985); A. D. Martin, R.G. Roberts, and W.J. Stirling, Phys. Lett. **189B**, 220 (1987); E.L. Berger, F. Halzen, C.S. Kim, and S. Willenbrock, Phys. Rev. **D40**, 83 (1989)

[20] R. Tanaka, XXVI International Conference on High Energy Physics, Dallas, TX, Aug 1992

[21] F. Abe *et al*, Fermilab Preprint FERMILAB-Conf-93/210-E, 1993

[22] V. Barger and R. Phillips, *collider physics*, Addison-Wesley Pub. Co. (1987), p. 215

[23] A.D. Martin, R.G. Roberts, W.J. Stirling, Phys. Lett **B306** , 145 (1993)

[24] E. Leanen, J. Smith and W.L. Van Neerven, "Top Quark Production Cross Section," Fermilab Preprint FERMILAB-Pub-93/270-T, ITP-SB-93-55 (1993)

[25] R.K. Ellis, "The theory of heavy flavor production," Fermilab Preprint FERMILAB-Conf-89/168-T (1989)

[26] B.Anderson *et al*, Phys. Rep. **97**, 31 (1983); T. Sjöstrand, Comp. Phys. Comm. **27**, 243 (1982); T. Sjöstrand, Comp. Phys. Comm. **28**, 229 (1983); T. Sjöstrand, Comp. Phys. Comm. **39**, 347 (1986)

[27] R.P. Feynman and R.D. Field, Nucl. Phys. B **136**, 1 (1978)

[28] C. Peterson, D. Schlatter, I. Schmitt, P.M Zerwas, "Scaling Violations in Inclusive e^+e^- Annihilation Spectra," Phys. Rev. **D27**, 105 (1983)

[29] B. Denby *et al*, the CDF collaboration, Proceedings of the 9th Topical Workshop on Proton-Antiproton Collider Physics, University of Tsukuba, Tsukuba, Japan, October 18-22, 1993

[30] For a detailed description of the CDF detector as configured for the 1988-89 Collider Run, see F. Abe *et al*, Nucl. Instrum. Methods Phys. Res., Sect. **A** **271**, 387 (1988), and references therein. For recent upgrades to the detector see A. Yagil "Highlights from CDF," edited by P.Kasper *et al*, in 7th Meeting of the American Physical Society, Division of Particles and Fields, 10-14 November 1992

[31] The SVX of the Collider Detector at Fermilab, Nucl. Instrum. Method Phys. Res., to be published.

[32] R. Blair *et al*, CDF Internal note 760 (1988)

[33] F. Abe *et al*, Phys. Rev. **D43**, 664 (1991)

[34] R.M. Harris, "Definition of CES χ^2," CDF Internal note 1329 (1990)

[35] L. Demortier, Ph.D. thesis, Brandeis University, 1991

[36] T. Liss, CDF Internal note 2367 (1993)

[37] P. Schlabach *et al*, CDF Internal note 2186 (1993)

[38] T.K Westhusing, "Observation of $W \rightarrow e\nu$ decays in proton-antiproton collisions at $\sqrt{s} = 1.8$ TeV," Ph.D. thesis, University of Illinois at Urbana-Champaign (1989)

[39] D.A. Smith, "Study of muons associated with jets in proton- antiproton collisions at $\sqrt{s} = 1.8$ TeV," Ph.D. thesis, University of Illinois at Urbana-Champaing (1989)

[40] H. Keutelian *et al*, CDF Internal note 2223 (1993)

[41] F.Paige and S.D.Protopopescu, ISAJET Monte Carlo version 6.21, BNL Report No. BNL38034, 1986 (unpublished). Throughout the analysis presented in this paper, we have used version 6.36

[42] F.A.Behrends *et al*, Fermilab-Pub-92/196-T, (1992), to be published

[43] P. Avery, K. Read, G. Trahern, Cornell Internal Note CSN-212, March 25, 1985, unpublished

[44] G. Marchesini and B. R. Webber, *Nucl.Phys.* **B238** (1984), 1; G. Marchesini and B. R. Webber, *Nucl.Phys.* **B310** (1988), 461

[45] M. Shapiro, D. Brown and S. Kannappan, CDF Internal note 753 (1988). M. Shapiro *et al*, CDF Internal note 1810 (1992)

[46] G.C. Fox and S. Wolfram, "A model for parton showers in QCD," Nucl. Phys. **B168**, 285 (1980)

[47] T. Sjöstrand, "A model for initial state parton showers," Phys. Lett. **B157**, 321 (1985)

[48] V.V.Abramovskii, O.V. Kancheli and V.N. Gribov, "Structure of inclusive spectra and fluctuations in inelastic processes caused by multiple-pomeron exchange", in *Proceedings of the XVI International Conference on High Energy Physics, National Accelerator Laboratory, Batavia, Illinois, September 1972*, Edited by J.D. Jackson, A. Roberts, and R. Donaldson, (National Accelerator Laboratory, 1973), Vol. 1, p. 389

[49] G.P. Lepage, "A new algorithm for adaptive multidimensional integration," J. of Comp. Physics **27**, 192 (1978)

[50] G.P. Lepage, "VEGAS − An adaptive multidimensional integration program," Newman Laboratory of Nuclear Studies, Cornell University Report No. CLNS-80/447 (1980)

[51] C. Campagnari, B. Farhat, P. Sphicas and A. Yagil, CDF Internal note 2150 (1993)

[52] See for example, OPAL collaboration CERN-PPE-93-106, submitted to Zeit. der Phys. C

[53] F.Abe *et al,* Phys. Rev. Lett. **70**, 4042 (1993)

[54] F.Abe *et al,* Phys. Rev. Lett. **64**, 152 (1990). B. Winer, Ph.D. thesis, University of California at Berkeley, 1990. C. Campagnari, B. Farhat, P. Sphicas and A. Yagil, CDF Internal note 2245 (1993). C. Campagnari, CDF Internal note 1025 (1989)

[55] J. Ohnemus, Phys. Rev. **D44**, 1403 (1991); J. Ohnemus and J.F. Owens, Phys. Rev. **D43**, 3626 (1991); J. Ohnemus, FSU-HEP-910621 and Fermilab-PUB-91/163-T, to be published.

[56] N.M.Shaw *et al*, CDF Internal note 2206 (1993)

[57] M. Mangano, preprint IFUP-TH 23/92

[58] F.E. James and M. Roos, "MINUIT - a system for function minimization and analysis of the parameter errors and correlations," Comp. Phys. Comm. **10**, 343 (1975)