

**Observation of $t\bar{t}$ Production using a Soft
Lepton b Tag in $p\bar{p}$ Collisions at $\sqrt{s}=1.8$ TeV**

A thesis presented

by

David Samuel Kestenbaum

to

The Department of Physics

in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in the subject of

Physics

Harvard University
Cambridge, Massachusetts
June 1996

©1996 by David Samuel Kestenbaum
All rights reserved.

To my mother and sister
and to my father
who sees muons in the backyard.

Acknowledgements

Numerous people have contributed to my education and to this work. The long list includes the many people who designed, built and worked on CDF; the top quark is their discovery. This is the short list.

My deepest thanks go to my advisor Melissa Franklin who served as mentor and friend, and who taught me the graffiti artist approach to physics. Melissa provided a wealth of ideas and inspiration. I would be lost without her.

I am equally indebted to Claudio Campagnari who lead me through this entire analysis, who encouraged me to think for myself, who taught me how to see when things were right and when they were wrong, and who demonstrated the benefits of slow walks and careful thought. Thanks also to Avi Yagil who developed the soft electron machine, and who is as insightful as he is ugly.

I have been fortunate also to work with Paolo Giromini who taught me to care about nanoseconds, and who has that rare ability to be able to understand the world from studying a single plot for five minutes. Thanks also to Phil Schlabach and Jorge Troconiz for their wisdom and humor during work on the CMX.

I acknowledge the constant support of Andrew Gordon whose intelligence, hair and footwork were a constant source of inspiration. Thanks also to George Michail, my office spouse, who made working in a 12x8x8 foot cube enjoyable. Thank you to Jaco Konigsberg for driving me to work and driving me to finish, also to Maria Spiropulu for her support and for the diagram on page 193.

My love and gratitude go to Anastasia Congdon, without whom, quarks would not mean so much.

Abstract

This thesis presents the results of a search for $t\bar{t}$ production in 110 pb^{-1} of $p\bar{p}$ collisions at $\sqrt{s}=1.8 \text{ TeV}$. The data were collected using the Collider Detector at Fermilab (CDF). The search identifies events consistent with the Standard Model decay ($t\bar{t} \rightarrow W^+bW^-\bar{b}$). W bosons are identified by the presence of high momentum electrons or muons associated with missing energy transverse to the beam. At least three jets are required to reduce the background from W bosons produced with associated jets. Finally the presence of a b-quark is inferred by searching for a soft lepton from the semileptonic decay ($b \rightarrow \ell \nu X$) or the sequential decay ($b \rightarrow c \rightarrow \ell \nu X$). Backgrounds from fake leptons or from real leptons occurring in generic jets are estimated from the data and account for 90% of the total background. The remaining backgrounds from Drell-Yan, $Z \rightarrow \tau^+\tau^-$, Wc , $b\bar{b}$ and Diboson production are estimated from a combination of Monte Carlo simulated events and data. Forty candidate events are identified containing 44 soft lepton tags with an expected background of 24.6 ± 3.0 tags. The probability of the background fluctuating upward to give the observed excess is 0.18%. The acceptance for top quark events is calculated and the $t\bar{t}$ cross section measured to be $\sigma_{t\bar{t}} = 8.5 \pm 3.4 \text{ pb}$. The results of this analysis are combined with the results of other searches and a kinematic fit for the top mass is performed yielding a top mass of $175.6 \pm 5.7 \text{ (stat)} \pm 7.1 \text{ (syst)} \text{ GeV}/c^2$. For the measured mass the cross section is in good agreement with predictions from the Standard Model.

Contents

Abstract	3
Table of Contents	5
List of Tables	10
List of Figures	16
1 The Top Quark and the Standard Model	30
1.1 The Standard Model	31
1.2 Quarks and Leptons	31
1.3 The Forces	33
1.4 The Bottom and Its Top	34
1.4.1 The Charge and Weak Isospin of the Bottom	36
1.4.2 Why the Bottom Should Have a Top	37
1.5 Previous Searches For the Top Quark	40
1.6 Top Quark Production at the Tevatron	41
1.7 The Decay of the Top Quark	47
1.8 The Lifetime of the Top Quark	47
2 The Tevatron and the Collider Detector at Fermilab	48
2.1 The Collider Detector at Fermilab	51
2.1.1 Tracking Detectors	53
2.1.2 Calorimeters	56

2.1.3	Muon detectors	60
2.1.4	The Trigger	64
2.1.5	Offline Reconstruction	68
3	Method	70
3.1	The Topology of $t\bar{t}$ Decays	70
3.2	The $t\bar{t}$ Decay Modes	71
3.3	The Lepton+Jets Channel	72
3.3.1	Identifying W's	72
3.3.2	Identifying b-hadrons	75
3.3.3	Backgrounds to Top	77
3.4	Outline of the Analysis	79
4	Data and Monte Carlo Samples	81
4.1	The W Sample	81
4.1.1	The Inclusive Electron Sample	81
4.1.2	The Inclusive Muon Sample	85
4.1.3	Final W Sample Selection	89
4.2	Monte Carlo Samples	94
4.3	The Low P_t Electron Sample	95
4.4	The Low P_t Muon Sample	96
4.5	The Jet Sample	98
4.6	The Photon Sample	100
4.7	The Z Sample	101
5	The Soft Lepton Tag (SLT) Algorithm	102
5.1	Soft Electron Tagging	103
5.1.1	Fiducial requirements	103
5.1.2	Central Strip Chamber Requirements	104
5.1.3	The Central Preradiator Requirement	107
5.1.4	The dE/dX Requirement	107

5.1.5	The E/p and E_{had}/E_{em} Requirements	109
5.1.6	Conversion Removal	111
5.1.7	Efficiency and Rejection Power of Algorithm	111
5.1.8	A Check of the Isolation Dependence of the Algorithm	114
5.2	Soft Muon Tagging	117
5.2.1	Muon Fiducial Requirements	118
5.2.2	Muon Selection Criteria	121
5.2.3	Rejection Power of Soft Muon Requirements	127
5.3	Monte Carlo Simulation of the Soft Lepton Tagger	131
6	Acceptance for $t\bar{t}$ events	133
6.1	Geometric Acceptance and Lepton Identification	135
6.2	Calculation of the Trigger Efficiency (ϵ_{trig})	136
6.3	Calculation of the Soft Lepton Tagging Efficiency (ϵ_{SLT})	137
6.4	Statistical and Systematic Uncertainties	138
7	Application of the SLT to the W+jets data sample	142
7.1	Background Removal	142
7.1.1	Drell-Yan	142
7.1.2	Low Mass Lepton Pairs	144
7.1.3	Upsilons	146
7.1.4	Z Bosons (second pass)	146
7.2	Soft Leptons in the W+Jets Sample	148
8	The SLT Backgrounds	150
8.1	The Soft Lepton Fake Rate from Jet Data	151
8.1.1	A Justification	151
8.1.2	Definition of the Fake Rate	154
8.1.3	Determination of the Muon Fake Rate	155
8.1.4	Determination of the Electron Fake Rate	159
8.2	Systematic Uncertainty on the Fake Rate	159

8.2.1	Check of Fake Rate Using Jet Samples	159
8.2.2	Check of Fake Rate Using the Photon Sample	161
8.2.3	Check of Fake Rate Using the Z Sample.	164
8.2.4	Systematic Uncertainty of the Fake Rate	164
8.3	Application of the Fake Rate to the W Sample	165
8.4	Drell-Yan background	166
8.4.1	Drell-Yan Removal Underefficiency	166
8.4.2	Drell-Yan Removal Overefficiency	168
8.4.3	Calculation of Remaining Drell-Yan	168
8.5	$b\bar{b}$ background in the W sample	169
8.5.1	Hadronic Background in the W Sample	170
8.5.2	$b\bar{b}$ background, First Method	173
8.5.3	$b\bar{b}$ Background, Second Method	174
8.5.4	Calculation of Residual $b\bar{b}$ Background	178
8.6	Wc background	179
8.7	$Z \rightarrow \tau^+ \tau^-$ background	180
8.8	Diboson background	181
8.9	Summary of the backgrounds	182
9	Analysis of the Results	184
9.1	Significance of the Excess of Tags	184
9.2	Calculation of the $t\bar{t}$ Production Cross Section	185
9.2.1	Combined Results from Run 1a and Run 1b Data	186
9.2.2	Comparison with other Analyses	186
9.3	Properties of SLT Tags in the W Sample	187
9.3.1	Stability of the Excess of Tags	187
9.3.2	Double Tagged Events	187
9.3.3	Events Tagged by the SVX and the SLT	190
9.3.4	Consistency with Jet Multiplicity Expected from Top . .	190
9.4	Determination of the Top Mass	193

9.5	Conclusions	201
A	Alternate Method of Estimating Fake, $Wb\bar{b}$ and $Wc\bar{c}$ Back- grounds	204
A.1	The True Fake Rate	204
A.2	$Wb\bar{b}$ and $Wc\bar{c}$ Backgrounds Using Monte Carlo	207
B	Calculation of the W Production Cross Section	209
C	Test of SLT on SVX Tagged b's	211
D	Soft Electron Fake Rate Composition	215

List of Tables

1.1	The Standard Model contains two families of six fermions (quarks and leptons) and four vector bosons which carry the electromagnetic, strong and weak forces. The weak isospin (I_3) shown is for the left handed fermions. The right handed components have $I_3=0$ and are weak isospin <i>singlets</i>	35
1.2	Comparison of theoretical calculations of $\sigma(p\bar{p} \rightarrow t\bar{t})$ for $M_{top}=175$ GeV/ c^2 at $\sqrt{s} = 1.8$ TeV. The various calculations use different structure functions for the central value. The uncertainties in the Laenen <i>et al.</i> and Berger and Contopanagos calculations do not include variations in the choice of structure functions.	43
3.1	Branching ratios for $t\bar{t}$ decay modes assuming Standard Model couplings. Here q stands for a u, d, c or s quark.	71
4.1	Inclusive high P_t electron selection requirements.	84
4.2	Conversion Removal	85
4.3	Level 2 trigger paths for primary muons. The Level 1 trigger identifies CMU-only, CMUP and CMX muons by reconstructing the appropriate stubs. There is no Level 1 or Level 2 stub trigger for CMP-only muons. CMP-only muons are accepted at Level 2 with a \cancel{E}_t + jets trigger (see table) and typically with a calorimeter trigger at Level 1.	87
4.4	Inclusive high P_t muon selection requirements.	89

4.5	Loose lepton quality cuts for the second (non-primary) muon in Z removal.	91
4.6	Events remaining in lepton sample after cuts.	92
5.1	Soft electron quality cuts. "P" is the track momentum in GeV/c. Local CPR and CES coordinates have $z=0$ at the $\eta = 0$ boundary and have \hat{x} in the azimuthal direction with $x=0$ at the center of the calorimeter wedge.	113
5.2	Efficiency of soft electron cuts for electrons and tracks in generic jets. The electron efficiency has been determined using conversion electrons with $P_t < 5$ GeV/c except for the E_{had}/E_{em} and E/p cuts which use b's from HERWIG generated $t\bar{t}$ events ($M_{top}=175$ GeV/c ²). Uncertainties for electrons are statistical only. The statistical uncertainties on $\epsilon_{jet\ tracks}$ numbers are less than 1%. The efficiencies are for each cut applied individually. Many cuts are correlated.	115
5.3	Selection requirements for soft muon candidates.	126
5.4	The rejection of muon candidate tracks in generic jet events at various stages of the soft muon selection. Initially there are 150k tracks which point to the muon chambers. Requiring a muon stub reduces the number of candidates to 35k, most of which are CMU-only types originating from hadronic punch-through. The final selection criteria which require that the candidate track be well matched to a muon stub provide a rejection factor of roughly 24, leaving 1466 soft muons. The average fake rate for this sample is then 0.98 %.	129

6.1	Components of the acceptance for $t\bar{t}$ events measured from HERWIG generated events with $M_{top} = 175 \text{ GeV}/c^2$. The uncertainties shown are from Monte Carlo statistics only. Each efficiency is measured after the requirements listed above it have been applied.	134
6.2	The uncertainty on the $t\bar{t}$ acceptance.	140
7.1	The number of events identified by the various removal algorithms. Since there is some overlap, the numbers given for each are after the removal algorithms above have already been applied.	144
7.2	The number of events before tagging, the number of soft lepton tags, and the number of soft lepton tagged events in the Run 1b W +jets data sample. There are four double tagged events in the $W+\geq 3$ jet bin and four in the $W+1$ jet bin.	148
8.1	The predicted and observed tags for fake rates obtained from generic jet samples collected with different triggers. In each case the fake rate is applied to an independent sample. For example the fake rate derived from the JET20 sample is applied to the combined JET50 and JET70 samples. The muon fake rate predictions agree with the observed tags to a few percent which is within the statistical uncertainty expected, the electron fake rate performs less well across samples.	160
8.2	Test of the fake rate in the photon + jet sample comparing the expected tags calculated using the fake rate to the number of observed tags.	164
8.3	Expected tags calculated using the fake rate and observed tags in the Z sample. Overall 83 tags are predicted by the fake rate and 82 are observed.	165
8.4	The underefficiency of the Drell-Yan removal algorithm in $Z \rightarrow l^+l^-$ events as a function of jet multiplicity.	166

8.5	The non-W fraction for various jet multiplicities using the \cancel{E}_t vs Isolation method described in Section 8.5.1.	173
8.6	From inclusive leptons in the the $\cancel{E}_t < 10$ GeV, $I < 0.1$ region - the number of events, number of soft lepton tags , number of expected residual Drell-Yan events (as calculated in Section 8.4), the expected number of fake tags, and the final tag rate after the backgrounds have been subtracted. The uncertainties include statistics and systematics on the background subtraction. The numbers are used to calculate the excess tag rate (due to $b\bar{b}$) in the non-W part of the signal sample.	175
8.7	The various components of the Wc background calculation. The “Real W fraction” accounts for expected contributions from QCD non-W events (measured using the \cancel{E}_t vs I method de- scribed in the text) as well as diboson production and Z events where one lepton has not been reconstructed or where the Z produced a single lepton through the decay $Z \rightarrow \tau^+ \tau^-$, $\tau \rightarrow \ell X$	179
8.8	Summary of the expected background and observed tags in the W sample. There is good agreement in the W+1 and W+2 jet bins and an excess in the W+ ≥ 3 jet bins as expected from $t\bar{t}$ production.	182
9.1	Results of the SLT analysis for Run 1a combined with Run 1b representing 109.4 ± 7.9 pb $^{-1}$ of data.	185
9.2	The $t\bar{t}$ production cross section measured in the dilepton channel and in the $\ell +$ jets channel using an SVX b-tag. Within the quoted uncertainties the measurements are in good agreement.	187

9.3	The number of observed tags compared to the expectations from background for soft electrons, soft muons and SLT's with the minimum momentum raised to $P_t > 4$ GeV/c. In each case there is good agreement in the W+1 and W+2 jet bins and an excess in the signal region as expected from $t\bar{t}$ production. . . .	188
9.4	Comparison, using Run 1a and Run 1b data, of the number of observed double tagged events with predictions from background and top.	190
9.5	Sources of systematic uncertainty on the measured top mass . .	200
A.1	The number of SVX and soft lepton tags found in a jet sample selected with a JET 50 trigger. The number of soft lepton tags found and predicted by the fake rate are shown for the subset of tracks in SVX tagged jets. A clear excess of soft lepton tags is observed as expected for real heavy flavor events.	206
A.2	The fraction of W+jets events predicted from a HERWIG Monte Carlo simulation of jet events to contain a $b\bar{b}$ or $c\bar{c}$ pair. Systematic uncertainties are estimated by comparing SVX tag rates in a jet sample to HERWIG predictions. Also shown is the soft lepton tag rate for $Wb\bar{b}$ and $Wc\bar{c}$ events.	207
A.3	Backgrounds calculated using Method II. The fake background is calculated by scaling the fake rate prediction down by 0.84 ± 0.10 . The $Wb\bar{b}$ and $Wc\bar{c}$ backgrounds are obtained from HERWIG generated Monte Carlo samples. "Other sources" includes the $Z \rightarrow \tau^+\tau^-$, $b\bar{b}$, Wc, Drell-Yan and Diboson backgrounds which are calculated as in the standard method (Table 8.8). The background calculated with the standard (fake rate) method is shown for comparison.	208

B.1	The cross section $\sigma(p\bar{p} \rightarrow WX)$ measured using the inclusive W sample from an integrated luminosity of $90.1 \pm 7.2 \text{ pb}^{-1}$. The systematic error on the luminosity dominates the final uncertainty. The uncertainty on the acceptance from Monte Carlo is taken as 5 % to cover uncertainties in the lepton selection criteria. The results are in good agreement with previous measurements of $20.4 \pm 2 \text{ nb}$	210
D.1	The number of tracks, soft electron tags found and soft electron tags predicted by the fake rate in the K_{short} and Λ mass peaks. For each number a background subtraction has been performed using the sidebands. The large uncertainties on the numbers from the Λ peak are a result of the large background present under the mass peak.	217
D.2	K_{short} and Λ selection requirements.	217

List of Figures

1-1	Example of how an isosinglet b-quark could mix with d or s quark and decay through a Z boson producing two leptons. . . .	38
1-2	Example of a triangle diagram giving rise to an anomaly. The contribution from a lepton isodoublet is exactly cancelled by a quark isodoublet.	38
1-3	Box diagrams for $B_d\bar{B}_d$ mixing. The internal lines (q) can be <i>up</i> , <i>charm</i> or <i>top</i> type quarks. The top quark gives the largest contribution because of its large mass and because it is the weak partner of the b.	39
1-4	Two examples of radiative corrections involving the top quark. The left diagram shows a one loop contribution to the Z mass. A correction to $\text{BR}(Z \rightarrow b\bar{b})$ is shown on the right.	40
1-5	The leading order diagrams for $t\bar{t}$ production at the Tevatron. .	41
1-6	The theoretical top cross section expected for $p\bar{p}$ collisions with $\sqrt{s} = 1.8$ TeV as calculated by E. Laenen <i>et al.</i> The upper and lower limits are obtained by varying the nonperturbative parameter μ_0 used in the initial state gluon bremsstrahlung resummation [23]. Cross sections for other processes are shown for comparison. At a top mass of $M_{top}=175$ GeV/c ² the $t\bar{t}$ cross section is 10 orders of magnitude below the total inelastic cross section and 3 orders of magnitude below the inclusive W cross section.	44

1-7	The $q\bar{q}$ ($u\bar{u}, d\bar{d}, s\bar{s}$) and gluon-gluon <i>parton luminosities</i> (using the MRSD0' structure functions) at the Tevatron as a function of the parton-parton center of mass energy \hat{s} . τ is defined as $\tau = \hat{s}/s$ where s is the center of mass energy of the $p\bar{p}$ collision ($\sqrt{s}=1.8$ TeV). The total $t\bar{t}$ production cross-section can be written as a sum over the parton luminosities and a matrix element for each process.	46
2-1	The layout of the Fermilab accelerator complex. The circumference of the Main Ring is approximately 6 km.	49
2-2	Isometric and side view of a quadrant of the CDF detector. . . .	52
2-3	Drawing of the CTC endplate. Sense wires are organized into cells which are positioned at 45° to radail so that the drift direction in the magnetic field and electric drift field is azimuthal.	55
2-4	A single central calorimeter wedge. Each wedge spans 15° in azimuth and contains ten towers in η each covering 0.1 units of pseudorapidity. Strip and wire proportional chambers are imbedded in the CEM at a depth roughly corresponding to shower maximum.	57
2-5	Location of the CMU in the central calorimeter wedge seen in an azimuthal slice and a polar slice of the wedge.	61
2-6	Schematic of a single CMU tower. A tower consists of 4 radial layers of 4 drift tubes. Sense wires in the first and second tubes are radially aligned as are wires in the third and fourth layers. The two pairs are offset by 2 mm (as measured at the chamber midpoint) to resolve left-right ambiguities in the reconstruction.	62
2-7	Regions of η and ϕ space covered by the CMU, CMP and CMX muon detectors. The 2.4° gap between CMU chambers (see Figure 2-5) is not shown.	63

2-8	Layout of the drift tubes in a 15° wedge of the CMX. The tubes are mounted onto four free standing arches which cover the region $0.6 > \eta > 1.0$ and span 240° in azimuth.	64
3-1	A schematic of the <i>lepton + jets</i> decay mode where one W decays to leptons and the other to quarks. The final state contains two light quark jets, two b-quark jets, a lepton and a neutrino. . . .	73
3-2	Jet multiplicity distribution for $t\bar{t}$ events generated with HERWIG (described in Section 4.2) which contain a lepton with $P_t > 20$ GeV/c and $E_t > 20$ GeV. Shown is the distribution for $M_{top}=140$ GeV/c ² (solid), $M_{top}=175$ GeV/c ² (dashed) and $M_{top}=210$ GeV/c ² (dotted).	74
3-3	A diagram for W+multijets production which is the major source of background in the lepton+jets channel. In this case the W is produced with four associated jets.	75
3-4	The expected W+jets and $t\bar{t}$ cross sections as a function of top mass and jet multiplicity. The W+jets bands reflect the estimated uncertainty in the theoretical calculation which is performed to leading order for each multiplicity. This plot is taken from Reference [59].	76
3-5	The P_t spectra of leptons from b and c decays in $t\bar{t}$ events ($M_{top}=160$ GeV/c ²) generated with the ISAJET Monte Carlo program (described in 4.2) [12].	78
3-6	Diagrams for heavy flavor production in W+multijet events. The left diagram shows Wc production where a gluon and a strange quark from the proton or anti-proton sea produce a W boson with an associated charm quark. The diagram on the right shows $Wb\bar{b}$ production where a final state gluon splits into a $b\bar{b}$ pair. $Wc\bar{c}$ events are produced in an identical manner.	79

4-1	Variables used in defining primary electrons, plotted for electrons in a $Z \rightarrow e^+e^-$ sample where the other electron in the event was required to pass tight cuts and to fire the trigger. Arrows indicate the primary lepton selection criteria listed in Table 4.1.	83
4-2	Variables used in defining primary muons, plotted for muons in a $Z \rightarrow \mu^+\mu^-$ sample where the other muon in the event was required to pass tight cuts and to fire the trigger. Arrows indicate the primary lepton selection criteria listed in Table 4.4.	88
4-3	Transverse mass of E_t and primary electron (left) and muon (right) for events in the W sample.	93
4-4	Variables used to select the conversion sample. ΔS is the distance between the two tracks in the transverse plane at the point where the tracks are tangent. $\Delta \text{Cot}(\theta)$ is the difference between the cotangents of the polar angles of the two tracks and measures their separation in the r-z plane. The bottom plot shows the radius of convergence of the tracks. The peak between 20 and 30 cm corresponds to the outer wall of the VTX and the inner wall of the CTC. The shaded histogram shows the radius of convergence for same-sign pairs which appears close to the origin as expected for tracks produced in the primary interaction. Arrows indicate the requirements for the low P_t electron sample (Section 4.3	97
4-5	Dimuon mass for muons in J/ψ sample. In the plot on the left, one muon is in the CMU,CMP or both. In the plot on the right one muon is in the CMX. The shaded areas show the peak (light) and sideband regions (dark). A distribution for real muons is obtained by plotting the desired quantity for tracks in the peak region and subtracting the same distribution for tracks in the sidebands.	98

4-6	Transverse energy of jets in events collected with a trigger that requires a calorimeter cluster with a minimum E_t of 20 GeV (top), 50 GeV (middle), and 70 GeV (bottom). The fake rate is calculated using tracks with $P_t > 2$ GeV/c, 75% of which are contained in jets with $E_t > 15$ GeV.	99
4-7	Invariant mass for lepton pairs in the Z sample described in the text. A explicit cut requiring $75 \text{ GeV}/c^2 < M_{\ell\ell} < 105 \text{ GeV}/c^2$ has been applied to the data.	100
5-1	E/p for five channel wire (left) and strip (right) clusters. Dashed histograms are from conversion electrons and are normalized to unit area. Solid histograms are from a sample of candidate tracks in jet events and have been normalized to an area of 1/3 to show shape. All samples have high statistics and so error bars have been omitted.	105
5-2	The mean E/p measured in the CES for conversion electrons as a function of the track momentum. The error bars mark one standard deviation from the mean. The superimposed line is the minimum required E/p for soft electrons. (See Table 5.1.) .	106
5-3	(Top) Preradiator response (Q_{CPR}) for conversion electrons (dashed), for candidate tracks in jet events with no id cuts applied (solid) and with all other cuts applied (dotted). All distributions have been normalized to unit area. (Bottom) Mean Q_{CPR} as a function of P/P_t . The error bars denote one standard deviation from the mean. Electrons deposit more energy in the CPR at large P/P_t where they have a longer path length through the the solenoid coil. The solid line shows the minimum CPR requirement listed in Table 5.1.	108

5-4	Q_{CTC} (dE/dX) as a function of momentum for electrons, pions, protons and muons from the data. This plot was prepared by Hovannes Keutelian [94]. Also shown is the predicted curve for kaons.	109
5-5	Q_{CTC} (dE/dX) for conversion electrons (dashed) and candidate tracks from jet events (solid). Q_{CTC} units are arbitrary. Both distributions are normalized to unit area.	110
5-6	E_{had}/E_{em} and E/p distributions from electrons from bottom or charm decay in HERWIG generated $t\bar{t}$ events with $M_{top} = 175 \text{ GeV}/c$. Roughly 55% are from bottom and 45% are from charm.	112
5-7	P_t of electrons in the conversion sample (left) and candidate tracks in the jet sample (right) with $P_t > 2 \text{ GeV}/c$. The conversion sample was collected with an $E_t > 8 \text{ GeV}$ electron trigger which produces a corresponding enhancement in the P_t distribution for the sample. In calculating the efficiencies and rejections in Table 5.2 only the tracks with $P_t < 5 \text{ GeV}/c$ are used so that the P_t spectrum of the two samples resemble that of leptons from b's in top events.	114
5-8	The efficiency of all soft electron cuts except E/p and E_{had}/E_{em} as a function of P_t as measured with conversion electrons. Uncertainties are statistical only.	116
5-9	Efficiency of the $E_{had}/E_{em} < 0.1$ cut as a function of $\Sigma_p^{0.2}$ for conversion electrons (triangles) and electrons from b or c decay in Herwig $t\bar{t}$ Monte Carlo with $M_{top}=175 \text{ GeV}/c^2$	118
5-10	Efficiency of all soft electron cuts except E/p and E_{had}/E_{em} as a function of P_t in a cone of $R=0.2$ around the electron track. The P_t dependence shown in Figure 5-8 has been removed. Within statistical uncertainties the efficiency is flat with $\Sigma_p^{0.2}$	119

5-11	Number of absorption lengths (averaged in azimuth) as a function of pseudorapidity calculated with $z = 0$	122
5-12	Expected one sigma track-stub mismatch ($\sigma_{\Delta x}$) in the x direction for muons in the CMU (square), CMX(triangle) and CMP(dots) as a function of P_t . In each case, $\sigma_{\Delta x}$ is measured at the front of the muon chamber. The CMP has the most shielding but muons in the CMX travel further after scattering before arriving at the muon chambers which gives them a larger $\sigma_{\Delta x}$. .	123
5-13	Muon matching variables from J/ψ 's. All plots have been background subtracted using the mass sidebands as described in the text. Also shown are gaussian fits, their means and sigmas. The $\Delta\phi_{CMP}$ distribution is not divided by the expected sigma since this is the quantity of interest in selecting CMP-only muons. (See Table 5.3.)	124
5-14	Efficiency of soft muon matching cuts as a function of P_t . Efficiencies are obtained from J/ψ 's and Z 's. For the J/ψ points, a background subtraction has been performed using the mass sidebands as described in the text.	128
5-15	Efficiency of muon id cuts as a function of P_t in a cone of $R=0.2$ around the muon track. The P_t dependence shown in Figure 5-14 has been taken out. The drop in efficiency for CMU-only muons is implemented in the Monte Carlo, the apparent drop in CMP is negligible. The CMUP and CMX efficiencies are flat with $\Sigma_p^{0.2}$ within statistical uncertainties.	130
6-1	Efficiency of the \cancel{E}_t and CFT trigger requirements. The logical OR of the two triggers is expected to be $> 99\%$ for electrons in the W sample.	136

6-2	(Top) The $t\bar{t}$ acceptance (in percent) as a function of top mass. (Bottom) The number of soft lepton tags expected from $t\bar{t}$ events in 90.1 pb^{-1} of data using the theoretical cross section from Laenen <i>et al.</i> (Figure 1-6) [23]. A 16 % uncertainty has been assumed on the cross section.	141
7-1	P_t spectrum of SLT's in events removed as Drell-Yan candidates from the $W+\geq 1$ jet sample. The top plot shows electron events, the bottom shows muon events. In each the histogram represents the opposite charge pairs that were actually removed and the triangles represent the same charge pairs that pass all other Drell-Yan removal cuts. The highest P_t bin contains entries for $P_t \geq 20 \text{ GeV}/c$	145
7-2	The invariant mass of same flavor primary lepton and SLT pairs in the inclusive lepton sample. Shown is the mass for opposite sign leptons with the distribution for same sign pairs subtracted. Drell-Yan pairs have been removed as described in the text. Visible are the low mass hump from the sequential decay $b \rightarrow$ $l\nu c, c \rightarrow lX$, and mass peaks from the $J/\psi, \psi'$ (for muons) and the Υ . Events with primary-SLT pair masses below $5 \text{ GeV}/c^2$ or between 9 and $11 \text{ GeV}/c^2$ are removed. Note also that there is no excess of events for larger masses. After the Isolation and \cancel{E}_t requirements have been imposed to select the W sample, only a handful of events are removed by these cuts. (See Table 7.1.) . .	147
7-3	Shown are the number of W+jets events before and after the application of the soft lepton tagger. The last multiplicity bin is an overflow bin corresponding to events with at least four jets. The data is for Run 1b only. Results from Run 1a are added in Section 9.2.1.	149

- 8-1 Diagrams for heavy flavor production in generic jet events - direct production (left), gluon splitting (middle), flavor excitation (right). In each case the $b\bar{b}$ pair may be replaced with a $c\bar{c}$ pair. 153
- 8-2 Top plot: CMU/P fake rate as a function of track P_t in jet events shown for all tracks (circles) and for tracks which were required to be away from a jet which passed the trigger (triangles). Bottom plot: The ratio of the fake rate for tracks away from trigger jets to the fake rate for all tracks. Jets containing muons are less likely to fire the trigger since they will have their energy measured systematically low. The rate for tracks well separated from a jet which would have caused the trigger to fire, is free from trigger bias. The effect is more pronounced at high P_t where the energy carried away by the muon is greater. 156
- 8-3 The fake rate for tracks extrapolating to the CMU or CMP (top) and to the CMX (bottom). The fake rate initially drops with P_t in part because the cut on matching chi-square becomes tighter with increasing P_t where multiple scattering is reduced for real muons. The selection criteria for the CMX change at $P_t=5$ GeV/c (Section 5) which causes a kink in the fake rate. 157
- 8-4 The muon fake rates as a function of $\Sigma_p^{0.2}$ for tracks extrapolating to the CMU/P (top) and CMX (middle) regions where the momentum dependence shown in Figure 8-3 has been removed. Also shown (bottom) is the electron fake rate as a function of $\Sigma_p^{0.2}$. Because of this dependence, the electron fake rate is calculated in three bins of $\Sigma_p^{0.2}/P$ 158
- 8-5 The electron fake rate as a function of track P_t for three different bins in $\Sigma_p^{0.2}/P$. The fake rate drops with increasing $\Sigma_p^{0.2}/P$ where tracks are less likely to pass the E/p and E_{had}/E_{em} cuts. . . . 162

8-6	In the photon sample, the track P_t , $\Sigma_p^{0.2}$ and E_t of jet in a cone of $\Delta R=0.4$ around the soft lepton (triangles) and as predicted by the fake rate (histogram).	163
8-7	Isolation (I) for electrons and muons from Z decay selected with the SLT algorithm. The fraction of events with $I > 0.1$ gives the overefficiency of the Drell-Yan removal. The overefficiency increases with jet multiplicity as shown in Table 8.4.	167
8-8	The regions in the \cancel{E}_t - Isolation plane used to estimate the non-W component of the W sample. The A region covers $I < 0.1$ and $\cancel{E}_t < 10$, the B region covers $I > 0.3$ and $\cancel{E}_t < 10$ GeV, the C region covers $I > 0.3$ and $\cancel{E}_t > 20$ GeV. D is the signal region and corresponds to $I < 0.1$ and $\cancel{E}_t > 20$ GeV.	171
8-9	The \cancel{E}_t distributions for inclusive leptons with Isolation (I) < 0.1 (solid) compared to leptons with $I > 0.3$ normalized to the $\cancel{E}_t < 10$ GeV region (triangles). Also shown are the (log scale) Isolation distributions for leptons with $\cancel{E}_t > 20$ GeV (solid) compared to leptons with $\cancel{E}_t < 10$ GeV normalized to the $I > 0.3$ region. The agreement of the shapes in the $\cancel{E}_t < 10$ GeV region and in the $I > 0.3$ region indicate that to a good approximation, I and \cancel{E}_t are uncorrelated for the non-W background.	172
8-10	The P_t distribution (histogram) for soft lepton tags in events with $\cancel{E}_t < 10$ GeV and $I < 0.1$ and at least one jet with $E_t > 15$ GeV. Also shown is the distribution after fakes (triangles) and the over- and under- efficiency of the Drell-Yan removal (squares) have been accounted for. The excess is assumed to come from $b\bar{b}$ events.	175

- 8-11 The invariant mass of primary and soft leptons in events with $\cancel{E}_t < 10$ GeV. The top histogram shows opposite sign primary- e and SLT- μ pairs with the same sign distribution subtracted. The bottom histogram shows opposite sign primary- μ and SLT- e pairs with the same sign distribution subtracted. An excess of events with mass below 5 GeV indicates the presence of b quarks decaying through $b \rightarrow l^{+(-)}c\nu, c \rightarrow l^{-(+)}\nu$ 177
- 8-12 The number of events, observed tags and predicted background tags in the W+jets sample from Run 1b. Boxes show the estimated uncertainty on the background prediction. There is good agreement in the W+1 and W+2 jet bins, and an excess in the W+ ≥ 3 jet bins as expected from $t\bar{t}$ production. 183
- 9-1 A transverse view of SVX tracks in a top candidate event with two jets tagged by the SVX, one of which contains a soft electron tag. The “+” denotes the origin ($x, y, z = 0$) and the cluster of tracks above it the primary vertex which is displaced by 0.05 cm in \hat{x} and .11 cm in \hat{y} . The other two clusters of tracks come from vertices which are displaced from the primary by ~ 6 mm and are tagged by the SVX b-tagger. The box shows an area 1.33 cm in \hat{y} by 3.86 in \hat{x} 191
- 9-2 The number of events, observed tags and predicted tags from background and $t\bar{t}$ production in the W+jets sample of Run 1a and Run 1b combined. A top mass of $M_{top} = 175$ GeV/ c^2 and cross section of $\sigma_{t\bar{t}} = 8.5 \pm 3.4$ pb is assumed. The data fits well to a combination of $t\bar{t}$ and background. 192

9-3	Reconstructed top mass from HERWIG generated $t\bar{t}$ samples where one jet has been b-tagged. Shown are distributions for $M_{top}=140,160,175,190$ GeV/c ² (clockwise). The full histogram shows the best fit obtained by the fitter. The dashed histogram shows the shape of reconstructed events with the correct jet-quark assignment.	195
9-4	The background template used in the top mass likelihood fit. The distribution is obtained by generating W+4 jet events with VECBOS, implemented with HERWIG fragmentation, and passing them through the SQUAW program which tries to reconstruct a top mass.	197
9-5	The reconstructed top mass for SVX and SLT tagged events with at least 4 jets (solid histogram). The inlay shows the change in the fit likelihood as a function of top mass. The dashed histogram shows the expected distribution from top ($M_{top} = 175$ GeV/c ²) and background with their relative contributions given by the fit. The shaded histogram shows the background shape normalized to 6.4 events.	198
9-6	The $t\bar{t}$ production cross section measured using the SLT plotted for the CDF measured top mass using SVX and SLT tagged top candidates. The solid line shows the theoretical prediction from Laenen <i>et al.</i> [23]. The dashed lines show the variation in the result when the nonperturbative parameter μ_o is changed. . . .	201

C-1 Distributions from the low P_t electron sample. E/p of the electron after quality requirements have been imposed (top left), and the E_t spectrum of the jet in the event which is separated by at least $\Delta R=0.4$ from the electron. Also shown is the L_{xy} distribution for those jets which are tagged by the SVX (lower left) and for those electron jets tagged by the SVX (lower right). The expected L_{xy} distributions from HERWIG $b\bar{b}$ Monte Carlo is overlayed. 213

D-1 **Top:** The reconstructed Kshort (K_s) and Λ mass peaks. **Bottom:** The radius of convergence for tracks in the sidebands used for background subtraction (left) and for a wider sideband region (right) showing a bump from photons which have converted in the outer VTX and inner CTC wall producing an e^+e^- pair. . 218

Chapter 1

The Top Quark and the Standard Model

The recently discovered top quark completes a set of twelve particles believed to be the fundamental constituents of matter. To date our best understanding of these particles and their interactions at small distance scales is expressed in a theory called the Standard Model [1]. The model contains free parameters which must be fixed by experimental measurements. Where predictions can be made from the Standard Model, agreement with experiment has been superb. Over the two decades of its development hundreds of measurements have confirmed the consequences of the Standard Model in many respects. The discovery of the bottom quark in 1977 and the subsequent measurement of its characteristics indicated the probable existence of a partner, the “top” quark. Though the Standard Model may one day be replaced by a more fundamental theory, it nevertheless provides a remarkably precise description of the subnuclear world over many orders of magnitude. The existence of the top is only one of its many triumphs. For completeness this chapter begins with a brief overview of the Standard Model.

1.1 The Standard Model

The Standard Model (SM) identifies three varieties of elementary particles - *quarks* (six in number), *leptons* (also six in number) and *gauge bosons* (four in number). Quarks and leptons may be considered the fundamental constituents of matter which interact through the four known forces, the gravitational, electromagnetic, weak and strong forces. Only the gravitational force is not incorporated into the SM.

Each gauge boson, the W, Z, photon and gluon is associated with a force. The quarks are named *up*, *down*, *charm*, *strange*, *top* and *bottom* (u,d,c,s,t and b). The leptons are the *electron*, *muon* and *tau* (e, μ , τ) and the three neutrinos ν_e , ν_μ , and ν_τ .

1.2 Quarks and Leptons

The quark and lepton “families” are further grouped into three “generations” each containing a pair of leptons or quarks. The weak force couples only to the left handed quarks and leptons. In each generation the left handed components couple to the weak force in the same manner, one particle carrying an isospin (the weak charge) of $+1/2$, the other of $-1/2$. The particle masses are unspecified in the SM. Each particle has a corresponding anti-particle with identical mass but opposite charge.

Each generation of the lepton family contains a lepton with unit charge¹ (the electron, by definition, the muon and the tau) and an electrically neutral neutrino partner (ν_e , ν_μ , and ν_τ). Experimental results constrain the neutrino masses to be quite small ($M_{\nu_e} < 5.1$ eV at the 95 % C.L. [26]) and are often taken to be zero.² Since neutrinos do not carry color (the charge of the strong force) they interact only through the weak force. The electron, discovered in

¹Electrical charge will be referred to in units of Q_e the electron charge.

²The LSND collaboration has recently announced evidence for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations indicating that neutrinos may have a small mass [2].

1897, was the first lepton and the first of the now considered elemental constituents of matter to be observed. The existence of the neutrino was inferred from the apparent momentum imbalance in beta decay in the early thirties but was not directly observed until 1958 [3]. In 1937 the muon was observed in a cosmic ray experiment [4] though it was not initially recognized as a lepton. The muon neutrino was observed directly, and determined to be distinct from the electron neutrino in fixed target experiment at Brookhaven in 1962 [5]. In 1975 the first third generation particle, the tau was discovered [6]. Its neutrino partner has not been directly observed. Measurement of the Z boson width at LEP indicates that there are only 3 light ($M_\nu < M_Z/2$) neutrino types, from which it may be inferred that there exist only three lepton generations [26].

Quarks differ most obviously from leptons in that they experience the strong force. Quarks come in three “colors” (*red* (r), *green* (g) or *blue* (b)) which is the charge of the strong force. Anti-quarks carry “anti-colors” (anti-red (\bar{r}) for instance). In nature quarks are observed in pairs (mesons) or triplets (baryons) bound together by the strong force. This phenomenon, called confinement, is a consequence of the structure of the strong force. The bulk of everyday matter is composed of two baryons, the proton (which contains two up quarks and one down quark) and the neutron (which contains one up quark and two down quarks). Free (single) quarks have never been observed. The existence of the up and down quarks was first experimentally inferred from deep inelastic scattering experiments which probed the inner structure of protons. Cosmic ray and fixed target experiments in the 1950’s revealed a plethora of particles that today are interpreted as different combinations of quarks. “V” particles [7] and their odd decays were eventually interpreted as evidence for a third quark, the “strange” quark. The absence of certain decay modes of the strange quark implied the existence of a fourth quark. In 1974 (the “November revolution”) a narrow resonance corresponding to a particle with mass ~ 3 GeV was observed in e^+e^- experiments at SLAC and Brookhaven, and interpreted as a quark-antiquark bound state of a fourth (the charm) quark [8]. In 1977 another

resonance, this time around 10 GeV pointed to a fifth (the bottom) quark [9]. As is discussed in the next section, this strongly implied the existence of a partner, the “top” quark, which is the subject of this thesis. It is a curious fact, unexplained by the SM, that the quark masses increase with each successive generation. All quarks carry fractional electric charge (of magnitude $2/3$ or $1/3$) and experience both the electromagnetic and weak forces. The SM indicates (see below) that there are the same number of quark families as lepton families, and hence that no undiscovered quarks remain.

1.3 The Forces

The mathematical language used to describe the SM is that of quantum fields. Particles appear as excitations in “matter” fields. Forces appear as other fields which “couple” to the matter fields with a constant that characterizes their strength. For each force there corresponds a particle which is said to “mediate” the interaction; these are the gauge bosons, the photon, W, Z and gluon. The photon, W and Z are associated with the electromagnetic and weak forces (and collectively described by Quantum Electrodynamics - QED in the SM). The gluon is associated with the strong force (and described by Quantum Chromodynamics - QCD in the SM). While all quarks and leptons have spins of magnitude $1/2$ (and so obey Fermi-Dirac statistics), the gauge bosons all have spin 1 (and so obey Bose-Einstein statistics).

The simplest way to introduce forces in the SM requires that the force carriers be massless. While this appears to hold for the photon and gluon, the W and Z (observed directly in 1983 at CERN) are extraordinarily heavy [11]. The SM elegantly accounts for the W and Z masses while also allowing the electromagnetic and the weak forces to be treated in a unified manner as aspects of a single “electroweak” force. This simplification comes at the expense of requiring an additional component to the SM called the Higgs sector. In its simplest form the Higgs sector may be a single particle (“*the Higgs*”). While

experimental results have placed limits on the allowed masses of various types of Higgs, the workings of this sector are still unknown.

In the SM, quark confinement arises because the strong force grows in strength with increasing distance. Baryons and mesons are said to be color neutral objects, the baryon by having a quark of each color (eg. red, green and blue), the meson by having a quark and anti-quark with the same color (eg. red and anti-red). Because the gluons themselves carry color, any attempt to remove a quark from a bound state produces a gluon “string” which acquires an energy density sufficient to split into quark-antiquark pairs. These pairs coalesce into more baryons or mesons, a process called fragmentation. Thus, quarks cannot be liberated from these bound states. In general the magnitude of the strong coupling often makes QCD predictions difficult.

An important feature of the SM (and of the world itself) is that the effective coupling strength of the forces changes with distance. In the case of electromagnetism the coupling becomes stronger at small distance scales. Unlike the photons which carry no charge, the gluons carry color and so the effect is reversed for the strong force; at very high energies (short distance scales) the strong force coupling decreases in strength. In this realm, strong interactions in the SM can be perturbatively calculated as an expansion in powers of α_s , the strong coupling constant. In this limit bound quarks can effectively be treated as free particles, a consequence called asymptotic freedom. Table 1.1 summarizes the particles in the SM.

1.4 The Bottom and Its Top

In the years following the discovery of the bottom quark its properties were extensively studied by many experiments. Given these inputs, the structure of the SM strongly implied the existence of a third generation partner, the top quark. When evidence for the top was announced in 1994 and confirmed in 1995, the surprise was not that the top existed but only that it was so heavy.

Quarks			Q	I ₃
$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
Leptons				
$\begin{pmatrix} e \\ \nu_e \end{pmatrix}$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -1/2 \\ +1/2 \end{pmatrix}$

Gauge Bosons	Force
Photon (γ)	Electromagnetic
$W^{+/-}, Z^0$	Weak
gluon (g)	Strong

Table 1.1: The Standard Model contains two families of six fermions (quarks and leptons) and four vector bosons which carry the electromagnetic, strong and weak forces. The weak isospin (I₃) shown is for the left handed fermions. The right handed components have I₃=0 and are weak isospin *singlets*.

This section reviews some of the earlier, indirect evidence for top.

1.4.1 The Charge and Weak Isospin of the Bottom

The charge of the bottom quark was first deduced from measurements of the leptonic width of the Υ in the late seventies at the DORIS e^+e^- storage ring [71]. Because the decay $\Upsilon \rightarrow l^+l^-$ takes place primarily via the exchange of a photon, the width is proportional to the square of the charge of the b-quark. This, and later measurements of $R=\sigma(e^+e^- \rightarrow \text{hadrons})/\sigma(e^+e^- \rightarrow \mu^+\mu^-)$ established the magnitude of the b quark charge to be $1/3$ [72].

The weak isospin of the bottom quark was first inferred from the forward-backward asymmetry in the reaction $e^+e^- \rightarrow b\bar{b}$ which is mediated by the photon and Z. The Z's axial and vector couplings to fermions depend differently on isospin and so the b's are produced with an angular asymmetry which can be used to infer the isospin of the b. Defined in terms of the b-quark production cross section, the measured asymmetry is

$$A_{fb}^{b\bar{b}} = \frac{\sigma(b, \theta > 90^\circ) - \sigma(b, \theta < 90^\circ)}{\sigma(b, \theta > 90^\circ) + \sigma(b, \theta < 90^\circ)}$$

where θ is the polar angle of the b-quark relative to the e^- beam direction. The JADE collaboration at PETRA ($\sqrt{s} = 35\text{GeV}$) first measured the asymmetry in the mid-eighties and found it consistent with SM predictions for a b-quark isospin of $-1/2$ [10]. Experiments at LEP [46] [47] have since made more accurate measurements at the Z resonance, where to lowest order the asymmetry is given by

$$A_{fb}^{b\bar{b}} \approx \frac{3}{4} \frac{2a_e v_e}{a_e^2 + v_e^2} \frac{2a_b v_b}{a_b^2 + v_b^2}$$

where a_f and v_f are the tree-level axial and vector couplings of the Z to fermion f . If the b quark were an isospin singlet ($I_3=0$), one would expect $A_{fb}^{b\bar{b}}=0$. On the other hand (without taking into account higher order corrections) a b-

quark with $I_3 = -1/2$ would result in a value of $A_{fb}^{b\bar{b}} \approx 0.1$.³ The combined LEP average is 0.0960 ± 0.0043 [45], in good agreement with this expectation.

1.4.2 Why the Bottom Should Have a Top

The charge and isospin of the bottom quark suggested that it belonged to a third quark generation since it resembled a heavier version of the strange quark. Theoretical arguments, which are outlined in this section, implied that it was part of a doublet.

Flavor Changing Neutral Current B decays

If the b were a weak isospin singlet (and so did not have a top quark partner) it could still decay through a W or Z by mixing with d and s quarks. It was shown in 1982 by Kane and Peskin [50] that this would imply

$$\frac{BR(B \rightarrow X l^+ l^-)}{BR(B \rightarrow X l^+ \nu)} \geq 0.12$$

since the b quark could then also decay through a Z as shown in Figure 1-1. In 1983 the first upper limits on flavor changing neutral current B decays by the Mark II collaboration ($BR(b \rightarrow l^+ l^- X) < 0.008$ at the 90 % C.L.) ruled this out, indicating that the bottom was part of a doublet [51].

Anomalies

Without a top quark, anomalies arising from diagrams (for example the triangle diagram shown in Figure 1-2) threaten the renormalizability of the SM. For the SM to remain internally consistent, these diagrams must sum to zero when the contributions from all fermions are taken into account. Each fermion (f in the drawing) contributes a term proportional to $N_c g_a^f Q_f^2$ where

³This assumes a value of $\sin^2(\theta_W) = 0.2311$ which is taken from the LEP average forward-backward asymmetry measured using $e^+e^- \rightarrow l^+l^-$ [45].

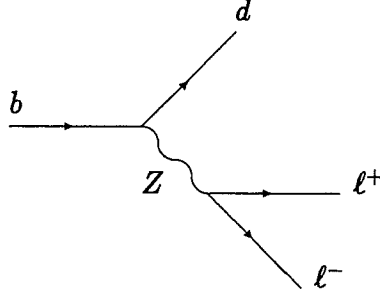


Figure 1-1: Example of how an isosinglet b-quark could mix with d or s quark and decay through a Z boson producing two leptons.

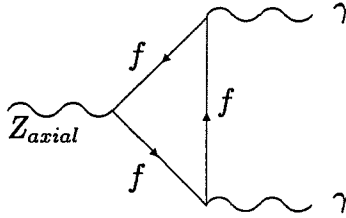


Figure 1-2: Example of a triangle diagram giving rise to an anomaly. The contribution from a lepton isodoublet is exactly cancelled by a quark isodoublet.

g_a^f is the axial coupling of the Z to fermion f and Q_f is the fermion charge, and N_c is a color factor (3 for quarks 1 for leptons). The contribution from a lepton isodoublet is exactly offset by a quark doublet. With three generations of leptons, the existence of a top quark ensures the cancellation of the anomaly [52].

B-mixing

The phenomenon of $B_d^0 - \overline{B}_d^0$ mixing observed in the late eighties [53] provided further evidence for the existence of a top quark. In the SM a B_d^0 meson can transform into a \overline{B}_d^0 meson through box diagrams as shown in Figure 1-3. The internal quark lines (q) can be up, charm or top quarks. The matrix

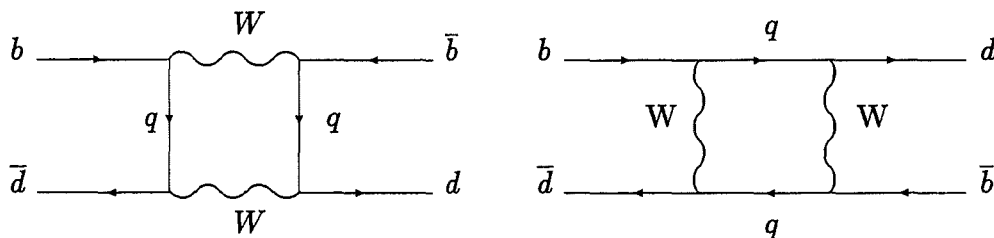


Figure 1-3: Box diagrams for $B_d \bar{B}_d$ mixing. The internal lines (q) can be *up*, *charm* or *top* type quarks. The top quark gives the largest contribution because of its large mass and because it is the weak partner of the b .

element is proportional to $M_q^2 |V_{qb} V_{qd}^*|^2$ and so the top quark gives the largest contribution. It has been shown that the observed level of mixing requires the existence of a heavy top ($M_{top} > 45 \text{ GeV}/c^2$) or physics beyond the SM [48].

Precision Electroweak Measurements and the Top

Although the top mass is a free parameter in the SM, it appears in a variety of radiative corrections for other observables. Today many high precision electroweak measurements are sensitive to these corrections. The top quark, for instance, appears in a one loop contribution to the Z mass (see Figure 1-4) although in this case similar corrections involving the Higgs dilutes what can be inferred about the top mass. The ratio, however, of $\text{BR}(Z \rightarrow b\bar{b})$ to $\text{BR}(Z \rightarrow \text{hadrons})$ (called R_b), is sensitive to the top mass (see Figure 1-4) while remaining largely independent of corrections to the propagator from the Higgs which drop out in the ratio. A fit to electroweak measurements made at LEP predicts a top mass of $177 \pm 11^{+18}_{-19} \text{ GeV}/c^2$ where the central value assumes $M_{Higgs} = 300 \text{ GeV}/c^2$ and the second uncertainty comes from varying the Higgs mass from $60 \text{ GeV}/c^2$ to $1 \text{ TeV}/c^2$ [45]. It is interesting to note that the measurement of R_b (which has the smallest dependence on Higgs mass) prefers a significantly lower top mass and disagrees with the mean of the global fit at the level of three standard deviations [49].

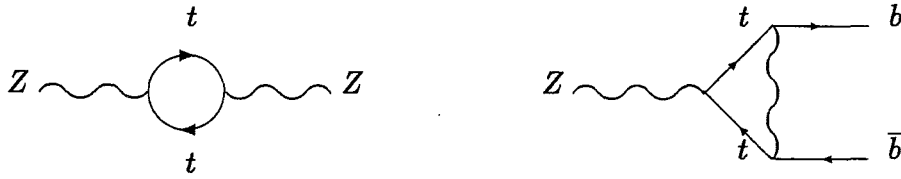


Figure 1-4: Two examples of radiative corrections involving the top quark. The left diagram shows a one loop contribution to the Z mass. A correction to $\text{BR}(Z \rightarrow b\bar{b})$ is shown on the right.

1.5 Previous Searches For the Top Quark

While early searches for the top took place at e^+e^- experiments, limits on its mass in the late eighties by the UA1 collaboration at CERN $Sp\bar{p}S$ collider ($M_{top} > 44$ GeV) [63] effectively left the task to the hadron colliders. In 1988 and 1989 the UA1 and UA2 collaborations collected more data setting a lower limit of $M_{top} > 69$ GeV/ c^2 . At the same time CDF completed its first high statistics run at the Tevatron.

Top quarks can be produced at the Tevatron either singly or in pairs. For top masses below roughly the W mass the two processes have comparable cross sections. Above the W mass and below ~ 200 GeV/ c^2 , $t\bar{t}$ pair production is expected to dominate. In 1992 a CDF search using 4.1 pb $^{-1}$ of data placed a 95 % confidence level lower limit on the top quark mass of 91 GeV/ c^2 . In 1994 the D0 collaboration using 15 pb $^{-1}$ of data, extended the limit to $M_{top} > 131$ GeV/ c^2 [73]. In March of that year the CDF collaboration, using 19.3 pb $^{-1}$ of data, published the first evidence for $t\bar{t}$ production with a cross-section of $13.9^{+6.1}_{-4.8}$ pb and measured a top mass of $174 \pm 10^{+13}_{-12}$ GeV/ c^2 where the first uncertainty is statistical and the second is systematic [12]. In April of 1995, after accumulating twice as much data, both the CDF and D0 collaborations announced unequivocal discovery of the top quark.⁴ CDF, using a combined

⁴The results of the soft lepton analysis were included in this paper.

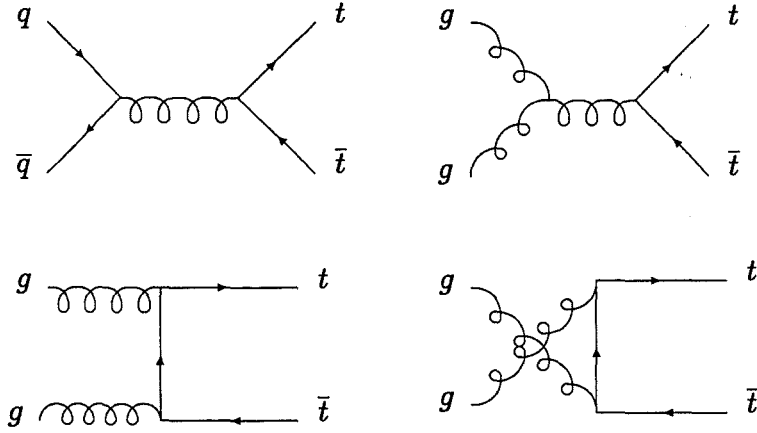


Figure 1-5: The leading order diagrams for $t\bar{t}$ production at the Tevatron.

67 pb^{-1} of integrated luminosity, measured a top quark mass of $176 \pm 8 \pm 10 \text{ GeV}/c^2$ and a $t\bar{t}$ production cross-section of $6.8^{+3.6}_{-3.4} \text{ pb}$ [14]. D0, using approximately 50 pb^{-1} measured a top quark mass of $199^{+19}_{-21} \pm 22 \text{ GeV}/c^2$ and a production cross-section of $6.4 \pm 2.2 \text{ pb}$. An excellent, comprehensive review of the search for and discovery of the top quark is given in [59].

1.6 Top Quark Production at the Tevatron

Figure 1-5 shows the leading order diagrams for $t\bar{t}$ pair production at the Tevatron. The pairs are produced either by gluon fusion or quark- anti-quark annihilation. The gluon initial state is expected to contribute $\sim 20 \%$ of the total cross-section for a top mass of $M_{top}=150 \text{ GeV}/c^2$, decreasing to 10% for a top mass of $M_{top}=200 \text{ GeV}/c^2$ [23]. Some non SM theories such as Technicolor predict possible enhancements in the $t\bar{t}$ cross-section [22].

In perturbative QCD the total inclusive cross section for $t\bar{t}$ pair production can be calculated as a sum over the contributions from the partons (quarks, anti-quarks and gluons) within the colliding proton and anti-proton:

$$\sigma(p\bar{p} \rightarrow t\bar{t}X) = \sum_{ij} \int dx_i dx_j \hat{\sigma}_{ij}(\hat{s}, M_{top}, \mu^2) F_i^p(x_i, \mu^2) F_j^{\bar{p}}(x_j, \mu^2)$$

where

- x_i (x_j) is the fraction of the proton (anti-proton) longitudinal momentum carried by parton i (j).

- \hat{s} is the center of mass energy of the two partons ($\hat{s} = x_i x_j S$ where S is the center of mass energy of the $p\bar{p}$ system).

- F_i^p and $F_j^{\bar{p}}$ are the parton number densities for the colliding proton and anti-proton. These denote the probability density that a parton of type i (j) carries a fraction x of the proton's (anti-proton's) momentum and are determined from fits to experimental data.⁵ The number densities are evaluated at a scale μ described below.

- $\hat{\sigma}_{ij}$ is the short distance matrix element for an initial state with partons i and j to produce a $t\bar{t}$ pair. The matrix element depends on the top mass and μ described below.

- μ is the factorization and renormalization scale. The scale μ has units of energy and is introduced in the renormalization procedure. There is no rigorous method for selecting a particular value of μ though it is usually taken to be on the order of the top mass.

The cross-section for producing a pair of heavy quarks was first calculated (to order α_s^2) in the late seventies [60]. In 1988 the full next-to-leading order (NLO) calculation was performed by Nason, Dawson and Ellis [56]. These were subsequently convoluted with parton distribution functions by Altarelli *et al.* [54] and Ellis [55] who found that the order α_s^3 corrections raised the cross-section by ~ 30 %. More recently Laenen, Smith and van Neerven [23] have included leading logarithm corrections to the α_s^3 calculation, corresponding to initial state gluon bremsstrahlung and report a ~ 15 % enhancement in the

⁵For example, the number of up-type valence quarks in the proton is two, so $\int_0^1 (F_u^p(x) - F_{\bar{u}}^p(x)) dx = 2$.

Calculation	Order	Cross Section ($\sigma_{t\bar{t}}$)
Altarelli <i>et al.</i>	NLO	$4.37^{+0.26}_{-1.1}$
Laenen <i>et al.</i>	NLO plus gluon resummation	$4.95^{+.70}_{-.42}$
Berger and Contopanagos	NLO plus gluon resummation	$5.52^{+.07}_{-.42}$
Catani <i>et al.</i>	NLO plus gluon resummation	$4.75^{+.63}_{-.68}$

Table 1.2: Comparison of theoretical calculations of $\sigma(p\bar{p} \rightarrow t\bar{t})$ for $M_{top}=175$ GeV/c² at $\sqrt{s} = 1.8$ TeV. The various calculations use different structure functions for the central value. The uncertainties in the Laenen *et al.* and Berger and Contopanagos calculations do not include variations in the choice of structure functions.

predicted cross section. This technique introduces a new scale μ_0 below which perturbative QCD is expected to break down. Similar calculations have been performed using a different resummation technique by Berger and Contopanagos yielding cross-sections roughly 10 % higher than those attained by Laenen, Smith and van Neerven [57]. A soon to be published paper [58] by Catani, Mangano, Nason and Trentadue re-evaluates the soft gluon resummation calculation and finds results smaller than both Laenen *et al.* and Berger *et al.*. Table 1.2 compares these predictions for a top mass of 175 GeV/c². Figure 1-6 compares the predicted $t\bar{t}$ cross-section as a function of top mass from Laenen *et al* to cross sections for other familiar processes. At a top mass of $M_{top} = 175$ GeV/c² the $t\bar{t}$ cross section is 10 orders of magnitude below the total inelastic cross section and 3 orders of magnitude below the inclusive W cross section.

Uncertainties in the theoretical predictions arise partly from uncertainties in the input quantities (the parton distribution functions for example, and Λ_{QCD} which determines how the parton densities and α_s evolve with μ) as well as intrinsic theoretical ambiguities such as the choice of the scale μ . Since μ is an arbitrary parameter, observable quantities should not depend on it. It is encouraging that the NLO top cross-section prediction is considerably more stable with variations in μ than the LO calculation which diverges for low μ

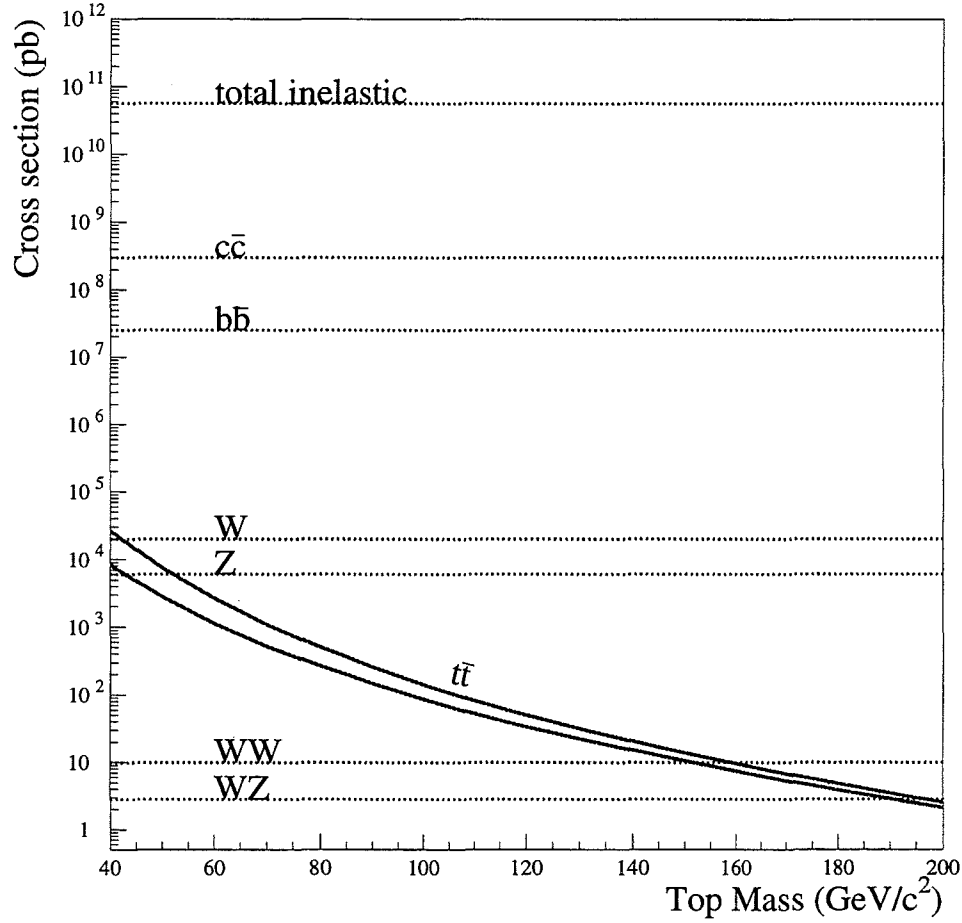


Figure 1-6: The theoretical top cross section expected for $p\bar{p}$ collisions with $\sqrt{s} = 1.8$ TeV as calculated by E. Laenen *et al.* The upper and lower limits are obtained by varying the nonperturbative parameter μ_0 used in the initial state gluon bremsstrahlung resummation [23]. Cross sections for other processes are shown for comparison. At a top mass of $M_{top}=175$ GeV/ c^2 the $t\bar{t}$ cross section is 10 orders of magnitude below the total inelastic cross section and 3 orders of magnitude below the inclusive W cross section.

[54].

Typically these quantities are varied and the change in cross-section taken as measure of the overall uncertainty on the prediction. The NLO predictions were found to be stable at the $\sim 20\%$ level when μ was varied from $2M_{top}$ to $0.5 M_{top}$ and Λ_{QCD} from 90 MeV to 250 MeV [54]. Laenen *et al.* find the cross section with gluon resummation to be stable at the 10% level when μ_o is varied from $0.05M_{top}$ ($0.2M_{top}$) to $0.1M_{top}$ ($0.25M_{top}$) for the $q\bar{q}'$ (gg) channel [23].

Some general features of $t\bar{t}$ production are evident from the leading order predictions. While matrix elements for the gluon fusion diagrams (Figure 1-5) are several times larger than the quark anti-quark diagram, the total cross-section also depends (through the number densities) on the distributions of quarks and gluons in the proton and anti-proton. To make this more explicit, the total cross-section can also be written as an integral over σ_{ij} and a quantity called the *parton luminosity* which is essentially an integral over the number densities.⁶ (See reference [64] for more details.) The effective gluon-gluon and quark-antiquark parton luminosities are shown as a function of \hat{s} in Figure 1-7. Top pairs ($M_{top} = 175 \text{ GeV}/c^2$) are produced with center of mass energies of order 400 GeV, at which point the gluon-gluon luminosity is an order of magnitude below the quark anti-quark luminosity. As a consequence, for large top masses, the $q\bar{q} \rightarrow t\bar{t}$ process dominates.

Calculated to leading order, the $q\bar{q} \rightarrow t\bar{t}$ matrix element is given by [60]

$$\hat{\sigma} = \frac{8\pi\alpha_s^2}{27\hat{s}} \sqrt{1 - \frac{4M_{top}^2}{\hat{s}}} \left(1 + \frac{2M_{top}^2}{\hat{s}} \right)$$

which is zero at threshold ($\hat{s}=4M_{top}^2$) and reaches a maximum at $\hat{s}=5.58M_{top}^2$. As a result the top quarks are produced with significant momentum. After convoluting the matrix element with the parton luminosities the most probable momentum of a top quark is $\sim 0.4M_{top}$ [59].

⁶The cross section can be written $\frac{d\sigma}{d\tau} = \sum_{ij} \frac{dL_{ij}}{d\tau} \hat{\sigma}_{ij}(\hat{s}, \mu^2, M_{top})$ where $\tau = \hat{s}/s$ and $dL_{ij}/d\tau$ is the parton luminosity.

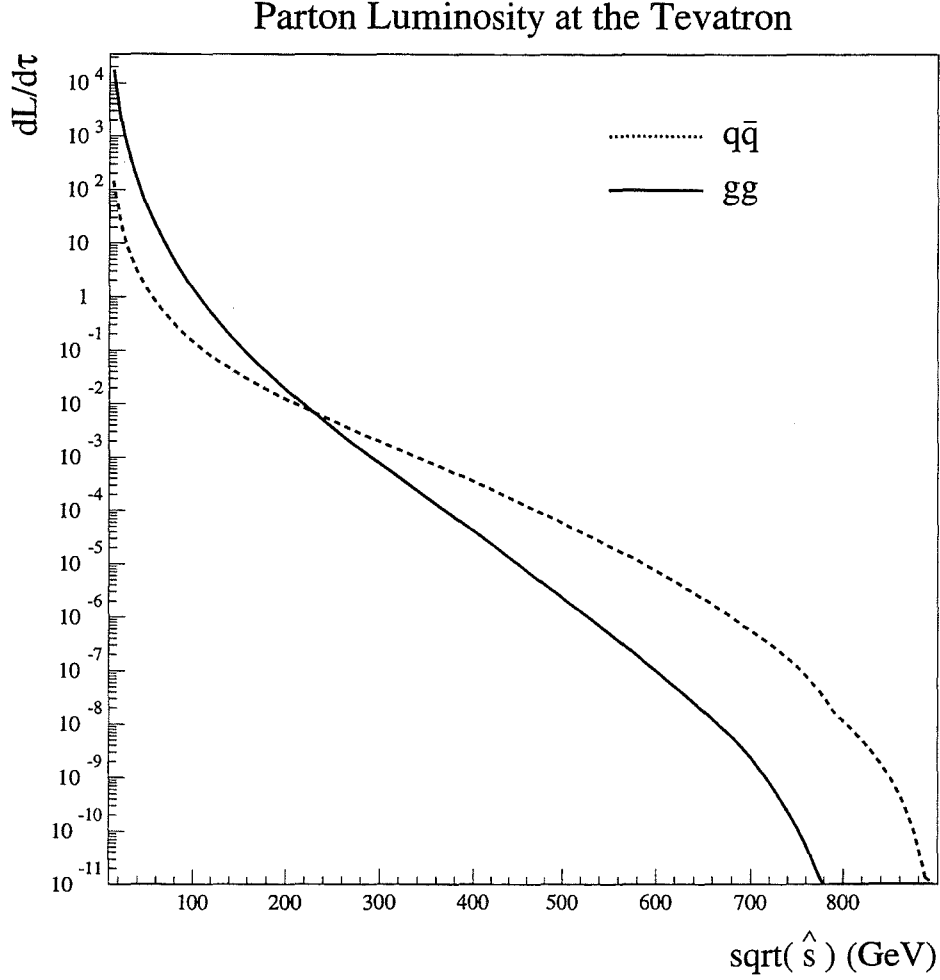


Figure 1-7: The $q\bar{q}$ ($u\bar{u}, d\bar{d}, s\bar{s}$) and gluon-gluon *parton luminosities* (using the MRSD0' structure functions) at the Tevatron as a function of the parton-parton center of mass energy \hat{s} . τ is defined as $\tau = \hat{s}/s$ where s is the center of mass energy of the $p\bar{p}$ collision ($\sqrt{s}=1.8$ TeV). The total $t\bar{t}$ production cross-section can be written as a sum over the parton luminosities and a matrix element for each process.

1.7 The Decay of the Top Quark

In the Standard Model, the top quark decays almost exclusively to a W boson and a b-quark. The decay, for instance, to a W and an s-quark, while allowed by the SM is heavily suppressed since it involves an off-diagonal element in the Cabibbo-Kobayashi-Maskawa (CKM) matrix which determines how the quarks mix in their coupling to the W. The ratio $\text{BR}(t \rightarrow Ws) / \text{BR}(t \rightarrow Wb)$ is proportional to $|V_{ts}|^2 / |V_{tb}|^2$ which is constrained to be less than 0.009 (90 % C.L. assuming unitarity and three generations) [26]. It should be noted that non SM theories like SUSY allow other sizeable decay channels for the top quark. A more complete discussion of the topology of $t\bar{t}$ decays is contained in Chapter 3 which describes the overall search strategy for this analysis.

1.8 The Lifetime of the Top Quark

In the limit $M_{top}, M_W \gg M_b$ (and taking $|V_{tb}| = 1$) the partial width for the decay ($t \rightarrow Wb$) is given by [62]

$$\Gamma(t \rightarrow Wb) \approx 175 \text{ MeV} \left(\frac{M_{top}}{M_W} \right)^3$$

so that a top quark with mass $175 \text{ GeV}/c^2$ should have a width of nearly 2 GeV and a lifetime of $\sim 4 \times 10^{-25}$ seconds. It is interesting to note that given this extremely short lifetime, the top is expected to decay before forming a hadron [62] [61]. While the hadronization processes is not particularly well understood, it has been argued that hadronization does not occur before the outgoing quarks are more than $\sim 1 \text{ fm}$ apart. At this distance, the stretched color string is expected to break producing $q\bar{q}$ pairs out of the vacuum which can combine with the top quarks to form mesons or baryons. A $175 \text{ GeV}/c^2$ top quark produced at the Tevatron travels only $\sim 0.04 \text{ fm}$ in one lifetime and so should not hadronize. However, because the top is so much heavier than the other quarks, the decay of a free top quark and a top hadron are not expected to be differentiable in current experiments [61].

Chapter 2

The Tevatron and the Collider Detector at Fermilab

The Tevatron is a large superconducting magnet accelerator that collides bunches of protons (p) and anti-protons (\bar{p}) at a center of mass energy (\sqrt{s}) of 1.8 TeV. The acceleration process begins with a series of staged accelerators which prepare the protons and anti-protons for the Tevatron. The protons and anti-protons travel in an evacuated “beampipe” which is held at a vacuum of $\sim 10^{-10}$ Torr. Figure 2-1 shows a schematic of the accelerator complex.

Proton bunches are made starting from a bottle of H_2 gas which is ionized to make H^+ ions. The ions are first accelerated through a 750 kV potential in a Cockroft-Walton electrostatic accelerator and subsequently to 400 MeV in a 145 m linear accelerator (the Linac). At the end of the Linac, the ions pass through a copper foil which removes the electrons, leaving the bare protons. The protons are transferred to a 475 m circumference synchrotron (the **Booster ring**) where they are accelerated to 8 GeV and formed into bunches. After the Booster, the protons are injected into a 6km circumference synchrotron called the **Main Ring**. The Main Ring accelerates the protons by 550 KeV per turn to an energy of 150 GeV. The Main Ring dipole magnets generate magnetic fields up to 0.7 T to keep the protons in the ring. The protons are finally coalesced into a single bunch and transferred to the Tevatron

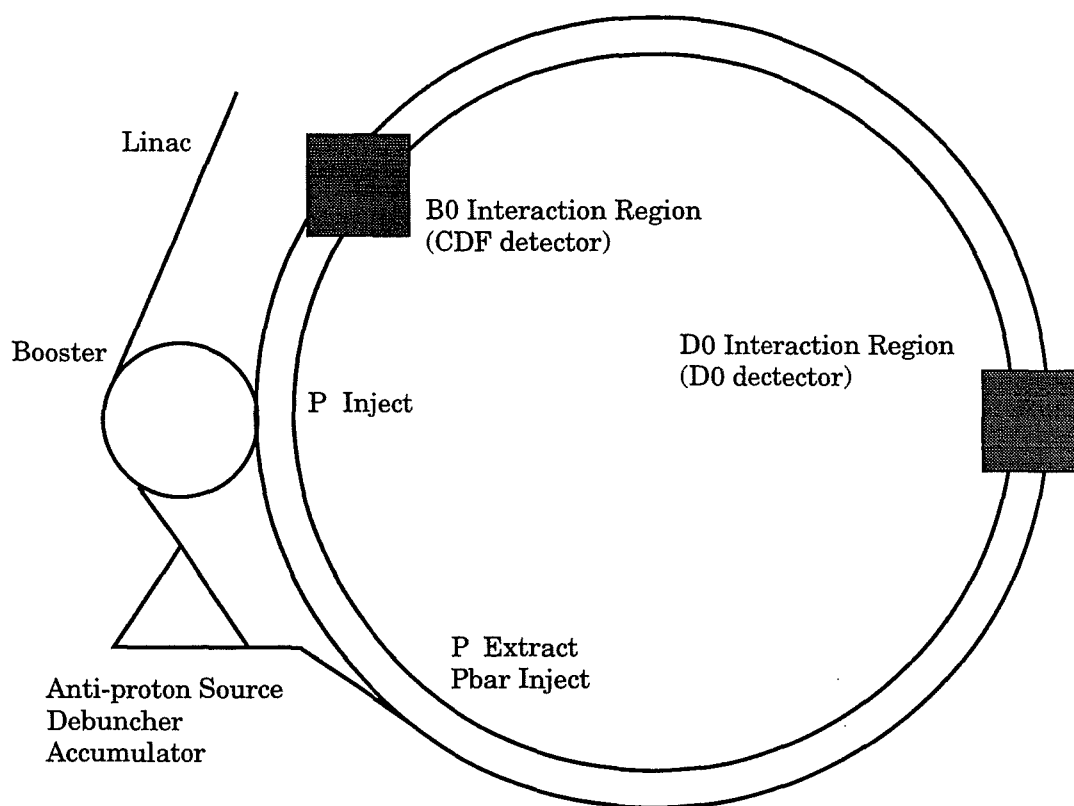


Figure 2-1: The layout of the Fermilab accelerator complex. The circumference of the Main Ring is approximately 6 km.

which is located directly below the Main Ring. The Tevatron employs superconducting magnets capable of generating fields from 0.66 to 4.4 T. Here the protons reach their final energy of 900 GeV. The entire process takes about one minute.

Anti-proton bunches require significantly more effort to create. Protons of 120 GeV from the Main Ring are shined onto a Tungsten target which produces a spray of secondary particles, some of which are anti-protons. The anti-protons are selected, focused with a Lithium lens and passed to the Antiproton Debuncher where they are stochastically cooled to reduce the phase space of the beam. The anti-protons are subsequently transferred to the **Accumulator**, a ring concentric with the Debuncher, where they are stored. Typically 4×10^{10} antiprotons are added to this “stack” every hour. When enough anti-protons are collected, six proton bunches (typically 2×10^{11} per bunch) are injected into the Tevatron. Anti-protons are then taken from the Accumulator and accelerated in the Main Ring and injected into the Tevatron in six bunches of typically 5.5×10^{10} anti-protons each. The protons and anti-protons travel in opposite directions in different orbits and are focussed to collide at two points (the *interaction regions*) around the ring where the detectors are positioned. The revolution frequency is approximately 50 kHz, and so bunch crossings occur in the interaction regions roughly every $3.5 \mu s$. The energy of each beam is ~ 900 GeV so the center of mass energy of a colliding proton and anti-proton is 1.8 TeV. The beams are typically left to collide in the machine for ~ 10 hours before the process is restarted. Anti-protons can be accumulated during this time since the Main Ring is no longer needed.

The accelerator quantity of interest for determining expected event yields is the *luminosity* (\mathcal{L}). The number of $t\bar{t}$ events produced, for instance, is given by

$$N_{t\bar{t}} = \int \mathcal{L} \cdot \sigma(p\bar{p} \rightarrow t\bar{t}X)$$

where $\int \mathcal{L}$ is the time integrated luminosity. The luminosity attainable by the

accelerator is roughly given by

$$\mathcal{L} = \frac{B f_0 N_p N_{\bar{p}}}{4\pi\sigma^2}$$

where B is the number of bunches, f_0 is the revolution frequency, N_p ($N_{\bar{p}}$) are the number of protons (anti-protons) per bunch and σ is the cross-sectional area of the bunches [69]. Typical starting luminosities for the 1995 run were $1.6 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$. The highest luminosity achieved was $2.8 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$. The luminosity falls exponentially with time due to transverse spreading of the beam and losses from collisions. Characteristic beam lifetimes are of the order 12 hours.

This analysis uses data collected from January 1994 through July 1995 (*Run 1b*) with the Collider Detector at Fermilab (CDF). Run 1b represents a (time) integrated luminosity of 90.1 pb^{-1} . Results from Run 1a data (August 1992 through May 1993) representing 19.3 pb^{-1} are discussed in Section 9.2.1.

2.1 The Collider Detector at Fermilab

CDF is a large multipurpose solenoidal detector designed to observe the particles produced in $p\bar{p}$ collisions. The perfect detector would determine the trajectory, momentum, energy and type (*id*) of every particle produced. In practice a detector design targets a few of these quantities. CDF was designed to explore physics at the high center of mass energy available at the Tevatron. This section briefly discusses the major components of the detector pertinent to this analysis. A more detailed description is given elsewhere [16].

CDF uses a right handed coordinate system centered on the interaction point with the z -axis pointing in the proton direction, the x -axis pointing horizontally toward the center of the ring and the y -axis pointing upwards. Components of momentum or energy transverse to the beam line are referred to as P_t and E_t . The pseudorapidity η is defined as a function of the polar angle (θ)

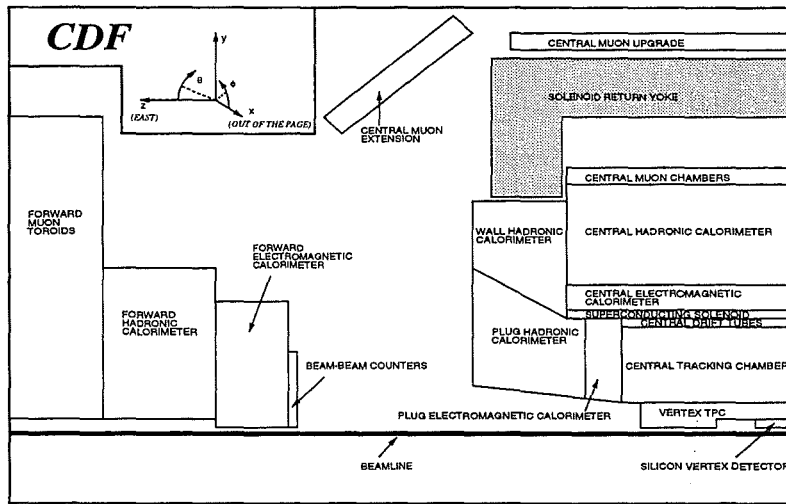


Figure 2-2: Isometric and side view of a quadrant of the CDF detector.

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

The detector is largely symmetric in z about $z=0$. Figure 2-2 shows side and perspective views of the detector.

The overall design of CDF, like most collider detectors, is dictated by the way in which particles interact with matter. CDF is composed of a variety of smaller devices. The collision point is surrounded with tracking chambers, followed by calorimeters and muon chambers.

All tracking takes place within the volume of the CDF solenoidal magnet. The superconducting solenoid measures 1.5 m in radius, 5m in length and provides an axial magnetic field of 1.4 Tesla. By measuring the curvature of (charged) tracks in the field, their momentum can be determined. The calorimeters surround the tracking chambers and measure the energy of electromagnetic showers and jets. The large amount of material present in the calorimeters absorbs most particles, leaving the muons to be observed with the outermost detectors, the muon chambers.

2.1.1 Tracking Detectors

CDF employs three layers of tracking devices, each tailored to a particular task. The Silicon Vertex Detector (SVX) [17] is closest to the beampipe, provides the highest position resolution, and is used to identify the displaced vertices of b-hadrons. The SVX consists of two identical barrels, one on each side of $z=0$. The active length of the SVX is 51 cm. The $p\bar{p}$ collisions occur with a spread of $\sigma \sim 30\text{cm}$ about $z=0$, so that roughly 60 % of all interactions take place within the SVX. Each barrel is divided into 12 azimuthal wedges of 30° , and into 4 concentric radial layers. The inner (outer) layer is 2.86 cm (7.87 cm) from the beampipe. Each layer consists twelve ladders of single sided silicon microstrip detectors. The strips are aligned parallel with the beamline and so provide 2-D tracking information in the $r-\phi$ plane. There are a total of 46080 channels in the SVX which comprise nearly a third of the total for the entire detector. Channels are read out only when they produce a hit which reduces this number by roughly a factor of 20. The SVX has an intrinsic single hit resolution of $\sim 15\mu\text{m}$.

The Vertex Time Projection Chambers (VTX) surround the SVX and are

primarily used to identify the z position of the primary interaction. There are eight separate chambers in the beam direction which cover 2.8 meters in z and extend to a radius of 22 cm. The chambers are octagonal and have wires in the end-caps which run parallel to each of the lines that define the octagon. A 50/50 mixture of Argon and Ethane gas is circulated through the modules. Charged particles passing through the VTX ionize the gas and free electrons which drift in the axial direction to the wires. The drift times give a measurement of the position of the track in the r - z plane. By combining this information for multiple tracks in the event, the z position of the primary vertex can be determined with a resolution of ~ 1 mm.

A large cylindrical open-wire drift chamber called the Central Tracking Chamber (CTC) surrounds the VTX. The CTC extends 3.2 m in the z direction, has an inner (outer) radius of 0.3 m (1.3 m) and contains over 6000 sense wires. The sense wires form 84 layers which extend the length of the chamber and which are grouped into 9 *superlayers* (numbered 0-8 increasing with r). Superlayers 0,2,4,6 and 8 each contain 12 sense wires which run in the z direction. Superlayers 1,3,5 and 7 each contain 6 wires which are tilted by $\pm 3^\circ$ to provide 3-D tracking. Field wires create a 1350 V/cm drift field. Wires in each superlayer are grouped into cells which are tilted by 45° with respect to the radial direction to compensate for the Lorentz angle from the crossed E and B fields. Figure 2-3 shows the cells seen in the r - ϕ plane. In this configuration the electrons from ionization drift in the azimuthal direction which simplifies track reconstruction. Tracks are reconstructed by fitting the hits to a helix. Residuals in the r - ϕ direction are better than $200\ \mu\text{m}$ and 6 mm in the z direction. The curvature in the transverse plane gives the momentum of the track. The momentum resolution is

$$\frac{\delta P_t}{P_t} \approx (0.002\ \text{GeV}^{-1}) P_t$$

for tracks with $|\eta| < 1.0$. When the track is required to pass through the beamline (a *beam constrained* fit) the resolution improves to $\frac{\delta P_t}{P_t} \approx (0.0011$

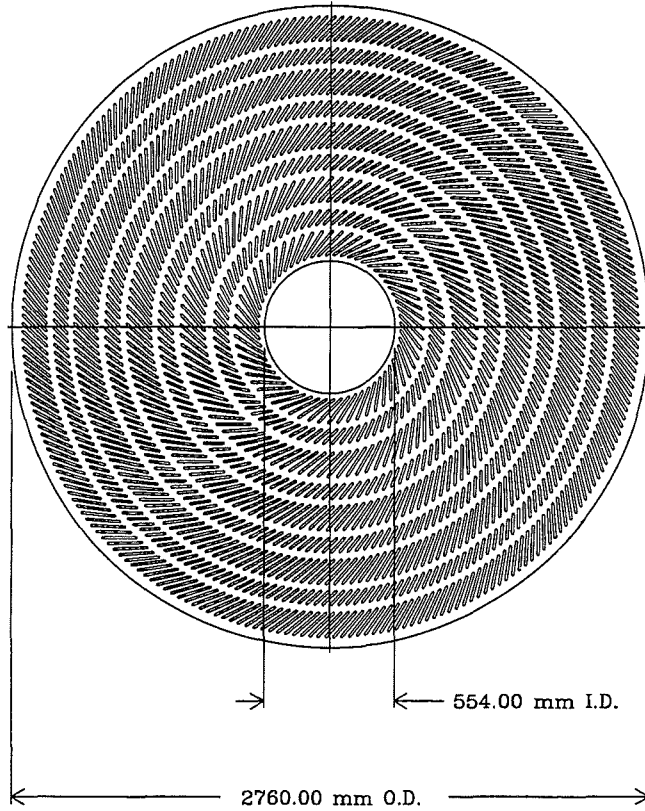


Figure 2-3: Drawing of the CTC endplate. Sense wires are organized into cells which are positioned at 45° to radail so that the drift direction in the magnetic field and electric drift field is azimuthal.

$\text{GeV}^{-1}) \times P_t$ Full 3-D tracking is not possible in the region $|\eta| > 1.6$ where particles traverse less than half of the CTC superlayers. The calibration of the CTC is checked offline by reconstructing the J/ψ mass from the measured muon momenta in the decay $J/\psi \rightarrow \mu^+ \mu^-$.

A measure of the charge deposited on wires in superlayers 3 through 8 of the CTC is obtained by integrating the ionization pulse when it is over a threshold. The information is passed on as the width of the digital TDC pulse which reports the time of the hit. Combining this information for all the hits on a track yields a measure of the ionization rate (dE/dX) of the particle which can be used to determine its identity.

2.1.2 Calorimeters

Calorimeter coverage at CDF is broken up into three eta regions, the *central* region ($|\eta| < 1.1$) the *plug* region ($1.1 < |\eta| < 2.2$), and the *forward* region ($2.2 < |\eta| < 4.2$). Because electromagnetic showers begin and finish much earlier than hadronic showers, each region has an electromagnetic calorimeter followed in depth by a hadronic calorimeter. Both are sampling calorimeters in which layers of absorber are alternated with layers of an active material. The absorber layers cause an incident particle to shower while the active material samples the energy flow in depth and gives a measure of the energy of the original particle. The calorimeters are segmented into *towers* which point back to the interaction point ($x, y, z = 0$).

Central Calorimeters

The central electromagnetic and hadronic calorimeters (the CEM and CHA) are divided into towers subtending 15° in azimuth and roughly 0.1 units of η . The calorimeters are separated physically into two sets ($\eta > 0$ and $\eta < 0$) of 24 (ϕ) wedges, each containing 10 eta towers. A cutaway view of a single calorimeter wedge is shown in Figure 2-4.

The CEM contains 31 layers of 5mm thick polystyrene scintillator interleaved with 30 layers of $\frac{1}{8}$ inch thick lead sheets. The CEM begins at an inner radius of 173 cm and is 35 cm deep. A particle traversing the CEM sees 18 radiation lengths of material. Particles in the shower produce light in the scintillator which is collected through wavelength shifters on both sides of the tower in ϕ and transmitted to light guides which are attached to photomultipliers located at the back of each wedge. The measured energy resolution for electromagnetic showers is

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{13.7\%}{\sqrt{E_t}}\right)^2 + (2\%)^2$$

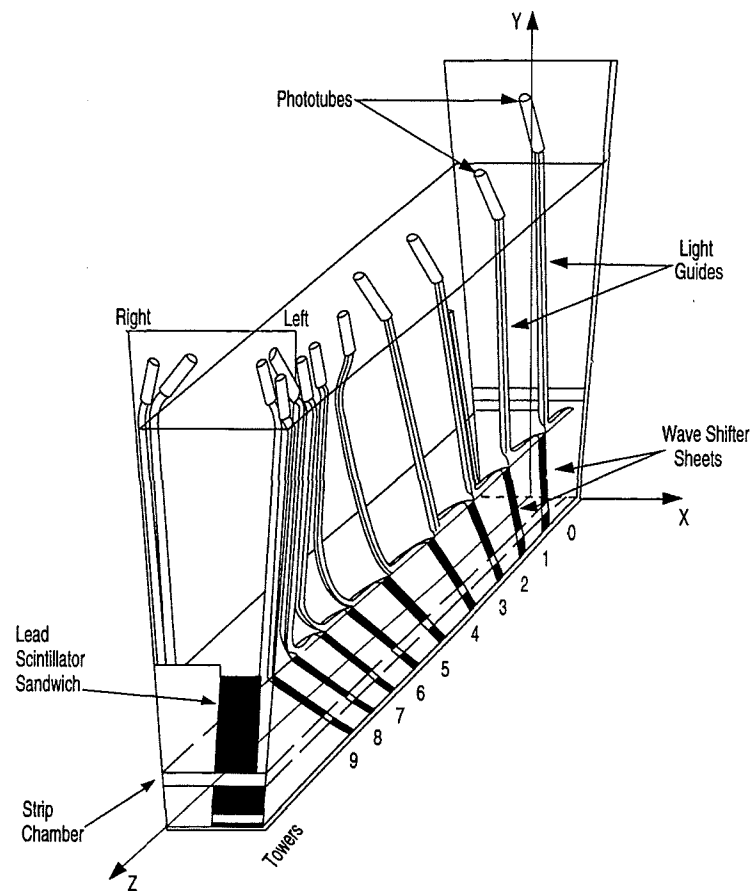


Figure 2-4: A single central calorimeter wedge. Each wedge spans 15° in azimuth and contains ten towers in η each covering 0.1 units of pseudorapidity. Strip and wire proportional chambers are imbedded in the CEM at a depth roughly corresponding to shower maximum.

with E in GeV. The CEM was originally calibrated using testbeam electrons and is checked periodically using ^{137}Cs sources.

The CHA has 32 absorber-sampling layers. The absorber layers are 2.5 cm of steel, the sampling layers are 1.0 cm of plastic scintillator. The scintillation light passes through a waveshifter and lightguide to photomultiplier tubes. The hadronic calorimeter alone contains ≈ 4.7 absorption lengths of material. The measured resolution for isolated pions is

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{50\%}{\sqrt{E_t}}\right)^2 + (3\%)^2$$

where E_t is in GeV. In the region $0.6 < |\eta| < 1.1$ particles do not pass through all layers of the CHA and so an additional hadronic calorimeter (the WHA) is necessary. The construction of the WHA is similar to the CHA except that the steel layers are 5 cm thick.

To enable a more precise measurement of the transverse profile of an electromagnetic shower, a proportional strip and wire chamber is instrumented in the CEM at a depth of approximately 6 radiation lengths (184 cm from the beamline). The strips are roughly 2 cm in width, run in the azimuthal direction, and are read out every tower in ϕ . The wires are separated by 1.45 cm, run in the z direction and are read out every 5 towers in η . The position resolution for 50 GeV electrons is roughly 2 mm in each direction.

To aid in distinguishing electrons from hadrons, a set of proportional tubes have been instrumented in the central region between the solenoid coil and the CEM to serve as a preradiator (the CPR) [93]. For each 15° wedge of the CEM there are two modules in η , each with sense wires read out every 2.2 cm. Electrons are likely to interact in the solenoid coil (which provides approximately .86 radiation lengths of material at 90°) and so will result in several particles which can deposit energy in the CPR. Hadrons, on the other hand, are less likely to interact and should deposit little or no energy in the CPR.

Plug and Forward Calorimeters

The plug and forward calorimeters are used primarily in this analysis for measuring jet energies and for determining the missing transverse energy (E_t) for an event. The calorimeters are subdivided into towers which subtend 5° in azimuth and 0.1 units of pseudorapidity.

The plug electromagnetic and hadronic calorimeters (PEM and PHA) cover the angular region from 10° to 30° with respect to the beamline ($1.1 < |\eta| < 2.4$) and the full 2π in azimuth. The calorimeters are physically divided into four quadrants of $\Delta\phi = 90^\circ$.

The PEM consists of 34 layers of proportional tubes interleaved with 2.7 mm thick sheets of lead. The proportional tubes are constructed from resistive plastic and are epoxied to sheets of copper-clad G10 which form the cathode. The cathode is etched into *pads* which form the basis of the towers. The pads at different depths are ganged together into three depth segments. When a particle showers in the calorimeter, the particles in the shower ionize the gas in the proportional tubes. The electrons from ionization drift to the wire in the tube, and the ions induce a charge on the cathode pads. The charge, which is amplified and read out, gives a measure of the energy of the shower. The energy resolution of the PEM is measured from testbeam electrons to be

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{22\%}{\sqrt{E}}\right)^2 + (2\%)^2.$$

The PHA is composed of 20 layers of proportional tubes interleaved with 5cm thick plates of steel. The cathode pads for the PHA are ganged together to form towers. The energy resolution determined from pions in a testbeam is

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{90\%}{\sqrt{E}}\right)^2 + (4\%)^2.$$

The forward electromagnetic and hadronic calorimeters (FEM and FHA) have constructions similar to the plug calorimeters. The FEM and FHA cover the region from 2° to 10° from the beampipe ($2.2 \leq |\eta| \leq 4.2$) and are divided into quadrants. The FEM consists of 30 layers of proportional tubes interleaved

with 4.5 mm thick lead sheets. The energy resolution, measured from electrons in testbeam is

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{26\%}{\sqrt{E}}\right)^2 + (2\%)^2.$$

The FHA is composed of 27 layers of proportional tubes interleaved with 5 cm thick steel plates. The measured energy resolution of the FHA is approximately

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{137\%}{\sqrt{E}}\right)^2 + (4\%)^2.$$

The gain of a proportional tube is a function of the gas composition, pressure, temperature and the chamber voltage. The plug and forward calorimeters use a 50/50 mix of argon and ethane gas. The gain of a monitor tube is kept stable to $\sim 1\%$ [68] by monitoring the temperature and pressure in the collision hall and adjusting the chamber voltage. The gain is also monitored with a set of proportional tubes with ^{55}Fe sources attached. The gain of the FEM and PEM is also checked offline by reconstructing the Z mass from a sample of electrons from $Z \rightarrow e^+e^-$ decay where one electron is detected in the CEM and the other in the PEM or FEM.

2.1.3 Muon detectors

Muons are detected at CDF with arrays of drift tubes placed outside of the calorimeters. The central muon chambers (CMU) and the central muon upgrade (CMP) cover the region $|\eta| < 0.6$. The central muon extension (CMX) covers the region $0.6 < |\eta| < 1.0$. The CMU, CMP and CMX are all single wire, rectangular drift tubes. A muon leaves hits in the drift tubes which are reconstructed to form a *stub*. Toroidal muon spectrometers (the FMU) are also located behind the forward calorimeters but are not used in this analysis.

The CMU and CMP

The central muon chambers (CMU) are located behind the CHA in the back of the calorimeter wedges. The CMU has an average of 5.4 pion interaction

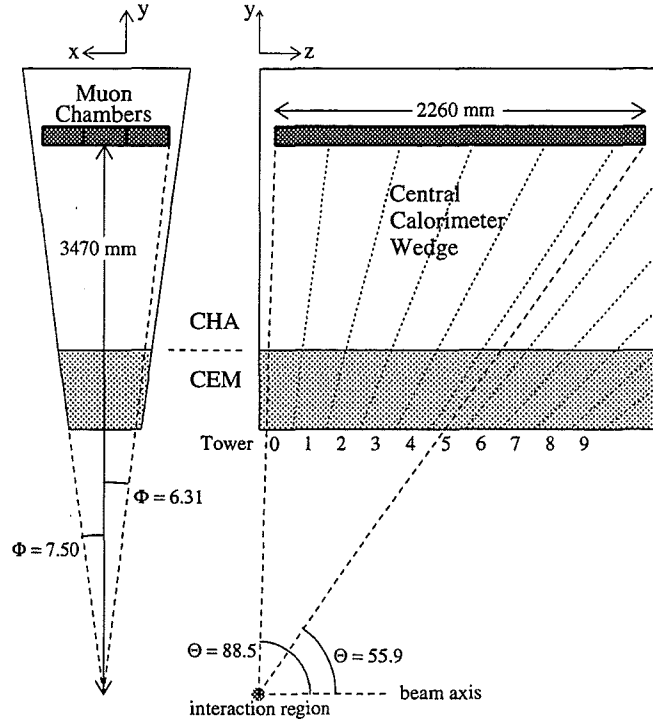


Figure 2-5: Location of the CMU in the central calorimeter wedge seen in an azimuthal slice and a polar slice of the wedge.

lengths between it and the event vertex. A significant number of hadrons (roughly 0.5 %) will traverse the CHA without interacting, and reach the CMU. [15]. Hadrons can also interact toward the back of the CHA, so that the shower leaks into the CMU. To reduce these backgrounds (collectively called *punch through*) additional steel (on average 3 pion interaction lengths) and muon chambers were added behind the CMU (called the CMP). The steel reduces the background from hadronic punch through by a factor of 20.

Each central calorimeter wedge contains three CMU *towers* with an inner radius of 347 cm. (See Figure 2-5) There is a gap of 2.4° in the coverage between each wedge. Each tower contains four radial layers of four drift tubes. (See Figure 2-6.) The tubes are 6.4 cm wide \times 2.7 cm high and run the length of the wedge. Wires from alternating cells are ganged together and read out at the $\phi = 90^\circ$ edge. Sense wires in alternating layers are radially aligned and are offset from the other pair by 2mm (at the midpoint of the chamber) so

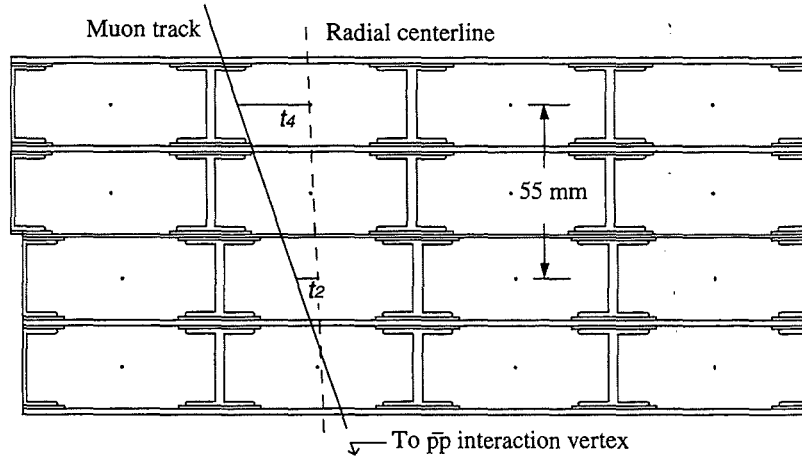


Figure 2-6: Schematic of a single CMU tower. A tower consists of 4 radial layers of 4 drift tubes. Sense wires in the first and second tubes are radially aligned as are wires in the third and fourth layers. The two pairs are offset by 2 mm (as measured at the chamber midpoint) to resolve left-right ambiguities in the reconstruction.

that the ambiguity as to which side of the wires (in ϕ) the particle has passed can be resolved. The sense wires are read out with ADC's at both ends which provides a measurement of the z position of the stub through charge division.

The CMP chambers consist of four layers of drift tubes which are stacked to form a box around central part of the detector. The CMP is not implemented in some regions of ϕ . (Figure 2-7) The CMP tubes are 2.5 cm wide \times 15 cm high and typically 640 cm long. Tubes in alternate layers half cell staggered to avoid gaps in coverage and left-right ambiguities. Since the chambers are a fixed length in z , the pseudorapidity coverage varies with azimuth. The overlap of the CMU and CMP coverage is shown in Figure 2-7. Approximately 84 % of the region $|\eta| < 0.6$ is covered by the CMU, 63 % by the CMP and 53 % by both.

CMX

The CMX detector covers the region $.65 < |\eta| < 1.0$, spans 240° in azimuth and consists of 8 layers of rectangular drift chambers organized into 4 groups

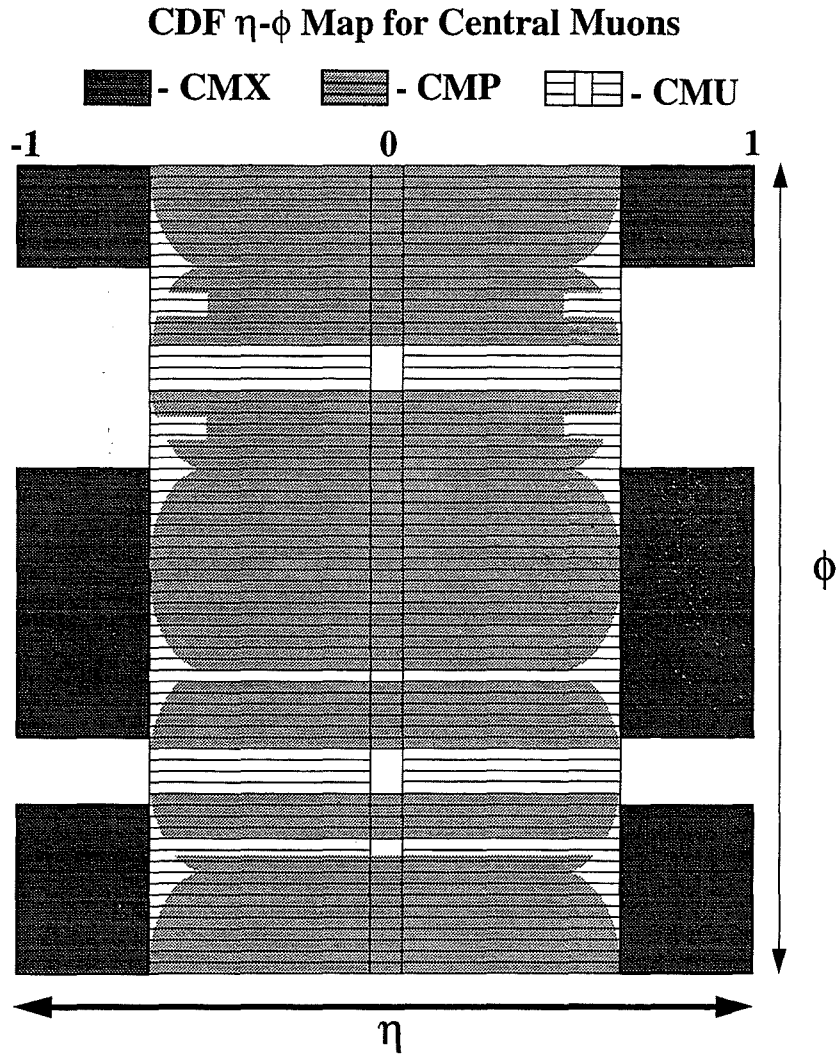


Figure 2-7: Regions of η and ϕ space covered by the CMU, CMP and CMX muon detectors. The 2.4° gap between CMU chambers (see Figure 2-5) is not shown.

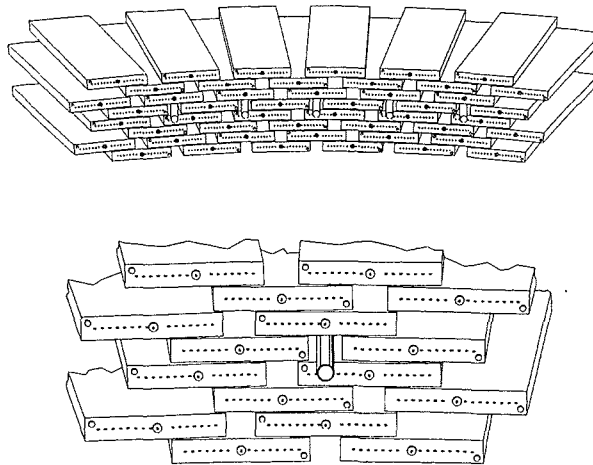


Figure 2-8: Layout of the drift tubes in a 15° wedge of the CMX. The tubes are mounted onto four free standing arches which cover the region $0.6 > |\eta| > 1.0$ and span 240° in azimuth.

of radially aligned pairs. The tubes themselves are 2.5 cm x 15 cm x 180 cm and are mounted to form four arch-like structures. Each arch lies on the surface of a virtual cone with an opening angle of 41.4° and a vertex 10 meters from the interaction point. Figure 2-2 shows the overall placement, Figure 2-8 shows the positioning of tubes in a 15° wedge. Drift tubes are grouped into radially aligned pairs, and adjacent layers are offset by a half cell. Most muons traversing the CMX intersect between four and six tubes. The CMX is sandwiched between two layers of scintillators (the CSX) which aid in identifying real muons.

2.1.4 The Trigger

In current operation, Tevatron beam crossings at CDF occur almost 286,000 times per second, roughly one every $3.5 \mu s$. With a total inelastic cross-section of roughly 50 mb at a luminosity of $1.6 \times 10^{31} \text{ cm}^{-2} s^{-1}$ an interaction rate of nearly 1 MHz is expected, an average of 3.5 interactions per crossing. Unfortunately CDF can record events to 8mm tape at a rate of only $\sim 10 \text{ Hz}$ and

so only 0.001 % of all interactions can be recorded. To accomplish this, CDF employs a three level *trigger*. Each level of the trigger is a logical OR of many separate triggers designed to select events of interest to different analyses. Ideally the trigger would be able to select or discard an event fast enough so that other beam crossings would not be missed, though in practice there is some dead time at high luminosities. The lowest level of the trigger is the fastest and simplest, the highest level is the slowest, but the most sophisticated. The Level 1 and Level 2 triggers are made by hardware processors, the Level 3 trigger is implemented in software running on a farm of Silicon-Graphics computers. The following discussion focuses on the lepton triggers used for this analysis.

Level 1

The Level 1 trigger is based on identification of energy clusters in the calorimeter or muon stubs in the muon chambers and reduces the event rate by roughly a factor of 1000. The decision time is less than $3.5 \mu s$ and so incurs no dead time. Electrons and jets are selected at Level 1 with a calorimetry trigger that requires a single *trigger tower* (defined as $\Delta\eta \times \Delta\phi = 0.2 \times 15^\circ$ sections of the calorimeter) to have energy over a threshold. The thresholds vary for each calorimeter and are 8 GeV for the CEM, 12 GeV for the CHA, 11 GeV for the PEM and 51 GeV for the PHA, FEM and FHA.

The Level 1 muon triggers require a pair of hits (called a Level 1 stub) in two parallel muon drift tubes with radially aligned wires. The difference in the drift times between the tubes (Δt) is a measure of the P_t of the muon. High P_t tracks are roughly radial and so have small values of Δt . Low P_t tracks undergo larger deflections by the solenoidal field, and so arrive at the muon chambers at a small angle to a radial line, resulting in a larger Δt . The Level 1 single muon trigger requires values of Δt corresponding to $P_t > 6$ GeV for the CMU and $P_t > 10$ GeV for the CMX.

The Level 1 muon trigger also requires a minimum of energy (~ 300 MeV) in a hadron calorimeter tower associated with the stub which reduced back-

grounds from accidental overlap of unrelated stubs and tracks. To reduce backgrounds from hadronic punchthrough, stubs in the CMU must also have a corresponding stub in the CMP except in regions of ϕ where the CMP coverage is not complete. There is no Level 1 trigger path for muons which leave a stub in the CMP but miss the CMU as may happen in the CMU gaps at the ϕ edges of the calorimeter wedges. Because of the cylindrical shape of the CMU and the rectangular shape of the CMP, some muon trajectories which strike the CMU at large η will miss the CMP. These muons, called “eta gap” muons also do not have a Level 1 trigger path.

The CMX detector, unlike the CMU, is completely unshielded from the Tevatron beam pipe which runs through the detector. The Level 1 CMX trigger suffers large backgrounds from particles from the collision which scatter in the beampipe (or subsequently the FEM) and into CMX. This is partially reduced by the hadronic energy requirement. To gain further rejection, the Level 1 trigger calculates an arrival time for a CMX stub using the accompanying scintillator array (the CSX) and requires it to be consistent with a particle coming from the $p\bar{p}$ interaction.

Level 2

The Level 2 trigger decision takes approximately $20\ \mu s$ and incurs a dead time of a few percent since the detector electronics must be read-out or cleared before processing the next event. The Level 2 trigger identifies calorimeter clusters, stiff tracks and muon candidates. Calorimeter clusters are formed by searching for a seed tower over a threshold and adding in neighboring towers which are over a lower (*shoulder*) threshold. The E_t and mean η and ϕ are calculated for each cluster. Stiff tracks are identified with a hardware processor called the CFT (Central Fast Tracker) which uses hit pattern masks to determine possible high momentum tracks. The CFT is essentially capable of measuring P_t with a resolution $\delta P_t/P_t = 3.5\% (\text{GeV}^{-1})P_t$.

The Level 2 electron trigger requires that a CFT track with $P_t > 12\ \text{GeV}/c$

be matched to a CEM cluster with $E_t > 16$ GeV and that the energy in the corresponding CHA towers must be less than 12.5% of the CEM energy. The Level 2 muon trigger requires a CFT track corresponding to a minimum P_t of 12 GeV matched to a Level 1 stub. Without further requirements, the rates for CMX, and CMU where the CMP was not required, are unacceptably large at high luminosities. To keep the rate manageable, these triggers are *prescaled* so that only 1 out of every n triggers (where n is lowered with decreasing luminosity) is allowed to pass. The inclusive CMX trigger, for example is prescaled by a factor of 80 at high luminosities, and a factor of 5 at low luminosities. Averaged over the whole run the effective prescale factor is 2.7. To retain good efficiency for the top analysis, a Level 2 trigger which also requires a calorimeter cluster with $E_t > 15$ is used. The rate for this trigger is low enough that prescaling is not necessary.

The Level 2 trigger also performs a rough calculation of the E_t in the event from the vector sum of the transverse energy in the calorimeter towers.

The event rate out of Level 2 is of order 20-35 Hz. A Level 2 trigger pass initiates a full detector readout which takes of order 30 ms which results in ~ 10 % deadtime overall.

Level 3

The final decision as to whether data is written to tape is made by a software trigger written in FORTRAN, running on a farm of Silicon Graphics processors.¹ The Level 3 trigger runs a scaled down version of the full offline code which reconstructs the event in detail and applies more stringent electron and muon identification requirements which are discussed in Chapter 4. Level 3 rejects 60-80 % of events which pass Level 2.

¹The farm has a combined CPU power of 3200 Vups. Vups is a Vax 11/780 unit of processing.

2.1.5 Offline Reconstruction

Events which pass Level 3 are written to disk and spooled to 8 mm tape. The tapes are processed offline with the full version of the CDF event reconstruction code. The code performs full 3-D tracking in the CTC and identifies jet, electron and muon candidates.

Electron reconstruction

Electron identification begins with a calorimeter based clustering algorithm to identify electron showers. An electron cluster consists of a seed tower with at least 3 GeV of electromagnetic transverse energy (E_t^{EM}) and adjacent (shoulder) towers with $E_t^{EM} > 0.1$ GeV. Beginning with each seed tower, a cluster is formed by adding shoulder towers and repeating until no more towers are added or until a maximum cluster size is reached ($\delta\eta \approx 0.3$). Finally the total E_t^{EM} of the cluster is required to be at least 5 GeV and the ratio of the hadronic to electromagnetic energy in the cluster is required to be less than 0.125.

Muon Reconstruction

A muon candidate consists of a CTC track matched to a track segment in a muon detector. Hits in the CMU, CMP and CMX are first fit to form track segments called *stubs*. Tracks in the CTC are then extrapolated to the muon chambers. Each stub is linked with the nearest CTC track in $r-\phi$ to form a muon candidate.

Jet reconstruction

Jets are identified using a clustering algorithm similar to the one described above for electron candidates. Plug and forward towers are first grouped into sets of three in ϕ to mimic the 15° segmentation in the central region and seed clusters formed by grouping all adjacent towers with $E_t > 1$ GeV. The energy weighted centroid is calculated for each group, and a jet defined as the sum of

all towers (now without the 15° grouping in the plug and forward) within a cone in $\eta - \phi$ space of radius $R=0.4$ from this centroid.² Overlapping jets are merged or separated depending on how the energy is shared [74]. The jet energies used in this analysis are the *raw* energies obtained from the calorimeter and are not corrected for effects such as reduced response near calorimeter boundaries or energy from the underlying event. (See Section 9.4.) These corrections can improve the resolution of the reconstructed top mass, but are unnecessary for this analysis since the same corrections apply for top and background events.

² ΔR is defined as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Chapter 3

Method

This chapter describes the overall search strategy for the analysis. The expected backgrounds and acceptances are discussed qualitatively to motivate its development.

3.1 The Topology of $t\bar{t}$ Decays

In the SM, the top quark decays almost exclusively to a W boson and a b-quark. The b-quark hadronizes and eventually decays. The W boson decays weakly to two leptons or quarks. The b and the quarks from the W fragment into hadrons and appear as collimated “jets” of particles in the detector. Only of order one in one billion neutrinos is expected to interact in the detector volume, and so the presence of a neutrino is inferred from an imbalance in calorimeter energy transverse to the beam (E_t).

Since the W couples to each weak isospin doublet with equal strength (“universality”) its branching fractions are essentially determined by the number of available decay channels (with the obvious exclusion of the top quark which is heavier than the W). The branching fraction to leptons, $BR(W \rightarrow l\nu)$ is approximately 1/9. The branching fraction to “light” quarks (ie. up, down, strange or charm) $BR(W \rightarrow q\bar{q}')$ is three times larger (approximately 1/3) since quarks come in three colors. Table 3.1 shows the decay modes for a $t\bar{t}$

Decay Mode	Branching Ratio
$t\bar{t} \rightarrow (q\bar{q}'b)(q\bar{q}'\bar{b})$	36/81
$t\bar{t} \rightarrow (q\bar{q}'b)(e\nu\bar{b})$	12/81
$t\bar{t} \rightarrow (q\bar{q}'b)(\mu\nu\bar{b})$	12/81
$t\bar{t} \rightarrow (q\bar{q}'b)(\tau\nu\bar{b})$	12/81
$t\bar{t} \rightarrow (e\nu b)(\mu\nu\bar{b})$	2/81
$t\bar{t} \rightarrow (e\nu b)(\tau\nu\bar{b})$	2/81
$t\bar{t} \rightarrow (\mu\nu b)(\tau\nu\bar{b})$	2/81
$t\bar{t} \rightarrow (e\nu b)(e\nu\bar{b})$	1/81
$t\bar{t} \rightarrow (\mu\nu b)(\mu\nu\bar{b})$	1/81
$t\bar{t} \rightarrow (\tau\nu b)(\tau\nu\bar{b})$	1/81

Table 3.1: Branching ratios for $t\bar{t}$ decay modes assuming Standard Model couplings. Here q stands for a u, d, c or s quark.

pair with their approximate branching ratios. Searches for the top quark are categorized by the decays of the two W 's in the event (one from the t , one from the \bar{t}). While observation of an excess of events in one of these channel might suffice to establish the existence of the top, it is useful to identify $t\bar{t}$ events in all channels. Because the majority of the hadronic decays of the τ are not easily distinguishable from quark or gluon jets, in what follows, a “lepton” will refer to a (more easily identified) electron or muon. Searches for top with hadronic tau decays are underway, but in general have very small acceptances ($\approx 0.1\%$) [65].

3.2 The $t\bar{t}$ Decay Modes

- The “dilepton” mode where both W bosons decay leptonically offers a clean signature (two high momentum leptons, two jets and large E_t) but has a small branching fraction (5%). In the current analyses, the largest backgrounds in

this channel come from $Z \rightarrow \tau^+ \tau^-$ decays and Drell-Yan events produced with associated jets [31] [12] [14].

- The “**all hadronic**” mode where both W’s decay to quarks has the largest branching fraction (44 %). Searches for top in this channel must contend with large backgrounds from multijet production and typically require at least one of the jets to be identified as coming from a b-quark [66].

- The “**lepton + jets**” mode, where one W decays to leptons and the other to quarks is the subject of this analysis. This is a particularly attractive channel for observing $t\bar{t}$ production because it has a sizeable branching fraction (~ 30 %) and smaller backgrounds than the all-hadronic mode. A typical lepton+jets event will have a well isolated high momentum lepton, large E_t (from the neutrino) and four jets (two from the second W and two from the b’s). The decay is shown schematically in Figure 3-1.¹

3.3 The Lepton+Jets Channel

3.3.1 Identifying W’s

The lepton+jets analysis begins with a sample of leptonic-decay W events. The selection of the sample is described in detail in the next chapter, but to aid the following discussion the main features are summarized here: The W sample contains events with an isolated electron or muon with $P_t > 20$ GeV/c, and $E_t > 20$ GeV. For the purposes of this analysis, jets are required to have $E_t > 15$ GeV and $|\eta| < 2.0$. While four jets should be present in $t\bar{t}$ events (Figure 3-1)

¹Another nice feature of the W+jet analyses are that they identify all the main features of a top decay, namely the presence of a W and b. Because, in principle all the $t\bar{t}$ decay products are identified in W+jets events, they can also be used to directly reconstruct the top mass.

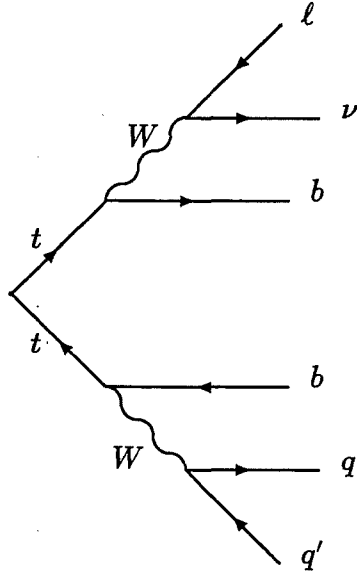


Figure 3-1: A schematic of the *lepton + jets* decay mode where one W decays to leptons and the other to quarks. The final state contains two light quark jets, two b-quark jets, a lepton and a neutrino.

in practice jets may fall outside of the detector acceptance or may have too low E_t to be counted. Initial or final state gluon radiation can contribute extra jets. Figure 3-2 shows the expected jet multiplicity in $t\bar{t}$ events.

The largest background to this analysis is the production of W bosons with associated jets (“W+multijets”). The inclusive W cross section is ~ 20 nb, more than three orders of magnitude higher than the $t\bar{t}$ cross section for $M_{top} = 175 \text{ GeV}/c^2$. The cross section for $W + N$ jets, however, is roughly suppressed by a factor of α_s^N . Backgrounds from W production can therefore be greatly reduced by requiring additional jets in the event. Figure 3-4 shows the expected $t\bar{t}$ and W+jets cross sections (calculated at tree-level) as a function of the number of reconstructed jets in the event. After requiring four jets in addition to the leptonic W, there are expected to be 3-5 times more W+jets events than top events [59]. (Figure 3-3 shows a diagram for production of a W and four jets.) To retain good efficiency for top, the final event sample is chosen to be those W events which also contain at least three jets, and is termed the $W + \geq 3 \text{ jet}$ sample.

Unfortunately there are large theoretical uncertainties in the predicted W+N jets cross-section due to variations with the choice of μ^2 , the scale at

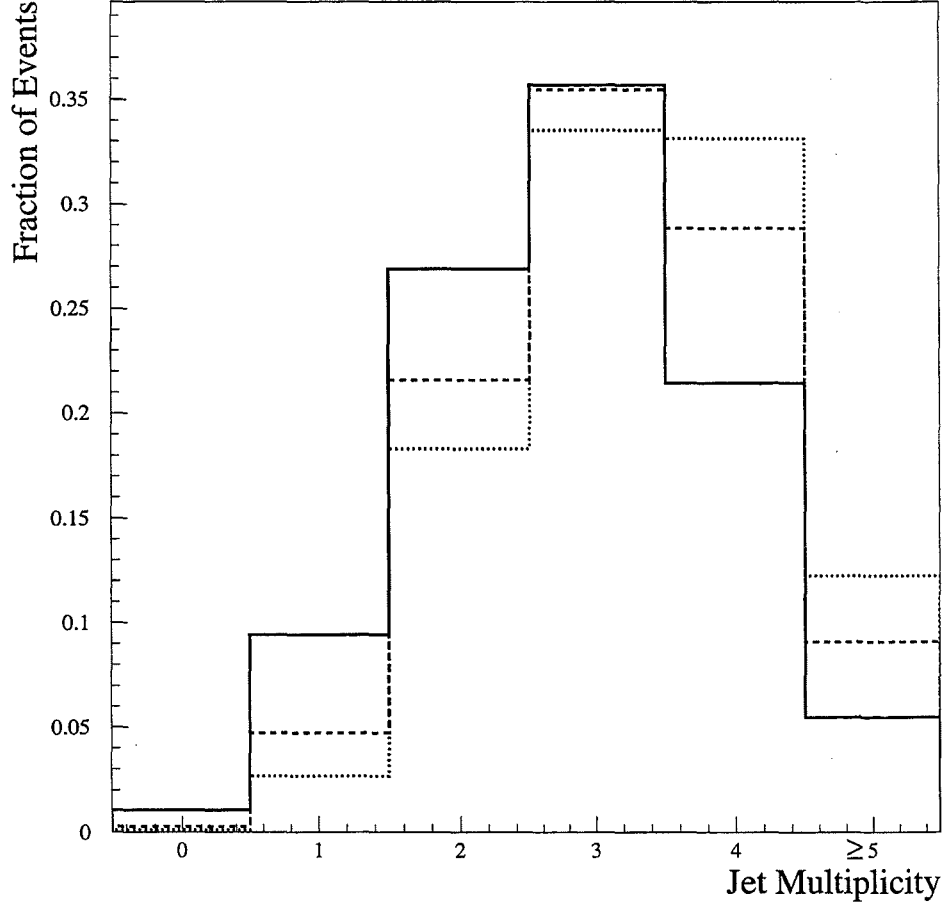


Figure 3-2: Jet multiplicity distribution for $t\bar{t}$ events generated with HERWIG (described in Section 4.2) which contain a lepton with $P_t > 20$ GeV/c and $\cancel{E}_t > 20$ GeV. Shown is the distribution for $M_{top}=140$ GeV/c² (solid), $M_{top}=175$ GeV/c² (dashed) and $M_{top}=210$ GeV/c² (dotted).

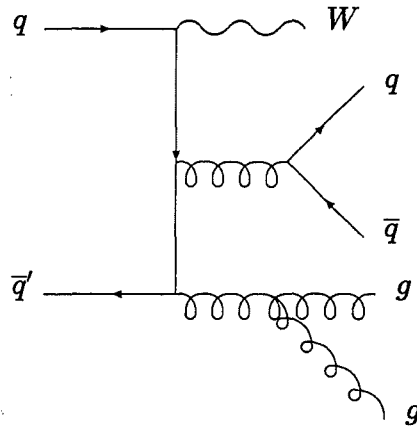


Figure 3-3: A diagram for W+multijets production which is the major source of background in the lepton+jets channel. In this case the W is produced with four associated jets.

which α_s and the parton distribution functions are evaluated. Because of this uncertainty (which grows with N) the existence of the top cannot be firmly established by an excess of W+jets events over the theoretical expectation. The signal to background, however, can be further improved by requiring that one of the jets come from a b-quark.

3.3.2 Identifying b-hadrons

Each $t\bar{t}$ event will contain two b-quarks which form b-hadrons. CDF utilizes two complementary techniques for identifying b-hadrons (called b-tagging). The first uses the SVX to associate tracks in a jet with a vertex which is displaced in the transverse direction from the primary vertex. The b-hadrons have a long lifetime (~ 1.5 ps) and receive a large boost in a top decay. In a $t\bar{t}$ event ($M_{top}=160$ GeV/ c^2) the b's travel an average of 3.4 mm in the radial direction before decaying. The SVX can resolve the displaced vertex to roughly $130 \mu m$ and can be used to select a high-purity sample of b's. The SVX analysis is detailed in references [67] [12] [13].

The second technique, which is the thrust of this analysis, tags b-quarks

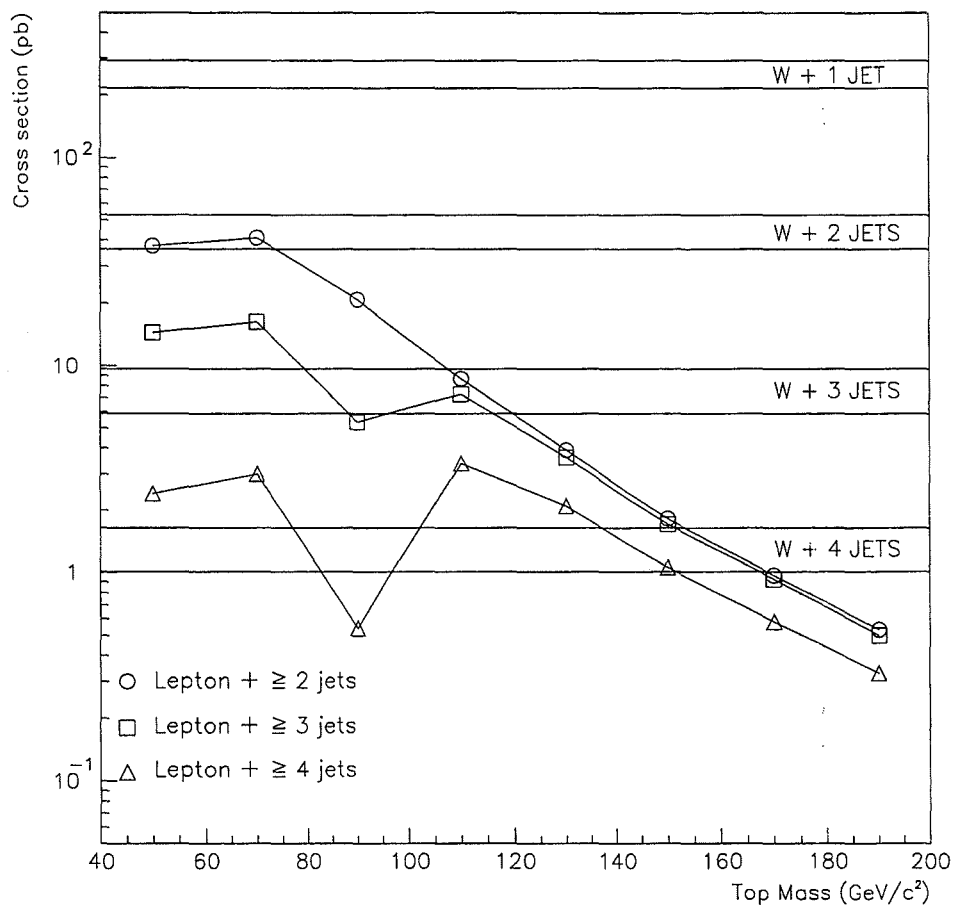


Figure 3-4: The expected W +jets and $t\bar{t}$ cross sections as a function of top mass and jet multiplicity. The W +jets bands reflect the estimated uncertainty in the theoretical calculation which is performed to leading order for each multiplicity. This plot is taken from Reference [59].

by searching for an electron or muon from their decay. Since leptons from b's are produced with low momenta compared to the lepton from the W (the *primary* lepton), this is called a *Soft Lepton Tag* (SLT). At first glance one might expect only a few $t\bar{t}$ events to contain such a lepton, but this is not the case. Consider first that the branching ratio for $b \rightarrow (e, \mu) + X$ is approximately 20% [26] [99] [100]. Soft leptons can also result in b hadron decays through the semileptonic decay of a sequential charm quark. The branching ratio $\text{BR}(b \rightarrow c \rightarrow e, \mu)$ is also approximately 20% [26]. Since there are two b's in each $t\bar{t}$ event, nearly 60% of top events will contain at least one soft e or mu from a b decay, and $\sim 20\%$ will contain two. Leptons from b decay typically have momentum of a few GeV/c, and are contained in jets. Figure 3-5 shows the expected P_t distribution from bottom and charm decays from a Monte Carlo with $M_{top} = 180 \text{ GeV}/c^2$.

A small number of soft lepton tags are also expected from the second W in the event, either where it decays to a charm quark or a tau lepton, producing a soft lepton, or where it decays leptonically and the event is not classified as a dilepton event. (The dilepton analysis selection is described in Section 4.1.3.) After adding in contributions from $W \rightarrow (c, \tau); (c, \tau) \rightarrow l$, roughly one soft lepton is expected in each $t\bar{t}$ event. It is also worth mentioning that roughly 2% of primary leptons are expected to come from the decay chain $W \rightarrow \tau \nu_\tau; \tau \rightarrow l \nu_l \bar{\nu}_\tau$.

3.3.3 Backgrounds to Top

After identifying SLT's in the $W + \geq 3 \text{ jet}$ sample, the number of expected tags from non-top (background) sources is calculated. $t\bar{t}$ production should evince itself as an excess of events over the expected background. The main background for this analysis arises from "fake" soft lepton tags. Fake tags can be defined as particles identified as leptons whose origin is not a heavy flavor

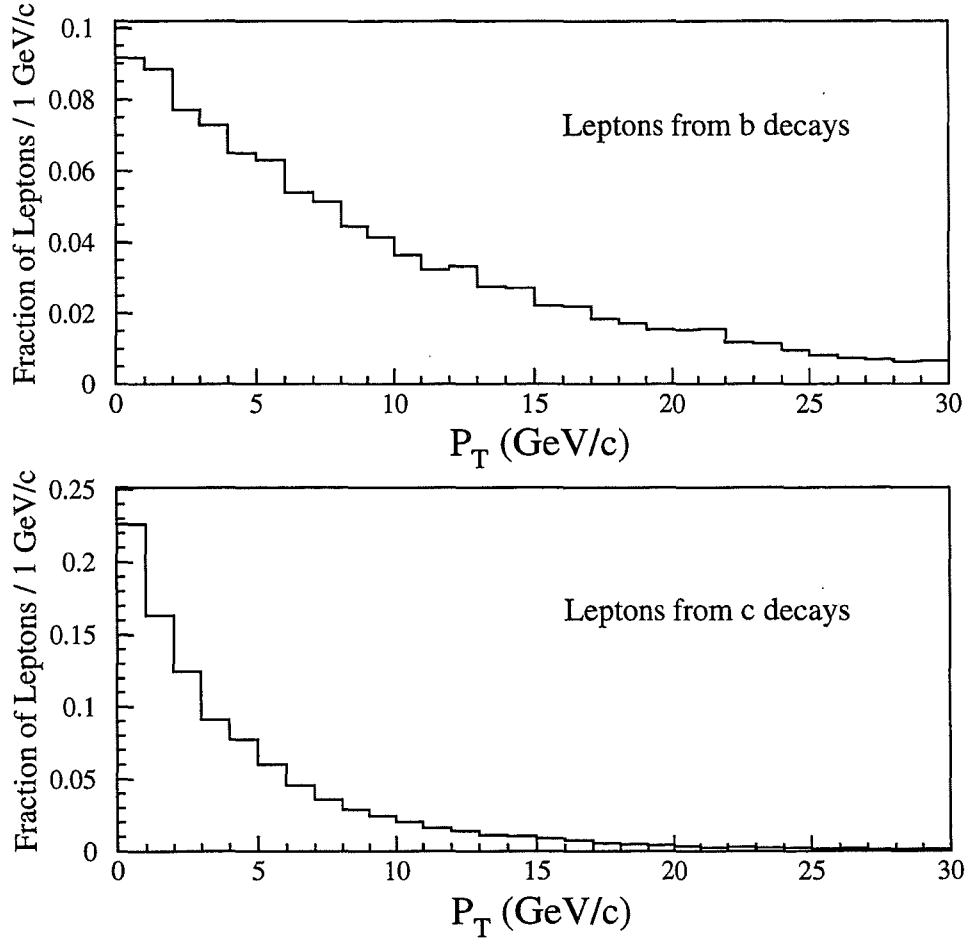


Figure 3-5: The P_t spectra of leptons from b and c decays in $t\bar{t}$ events ($M_{top}=160$ GeV/ c^2) generated with the ISAJET Monte Carlo program (described in 4.2) [12].

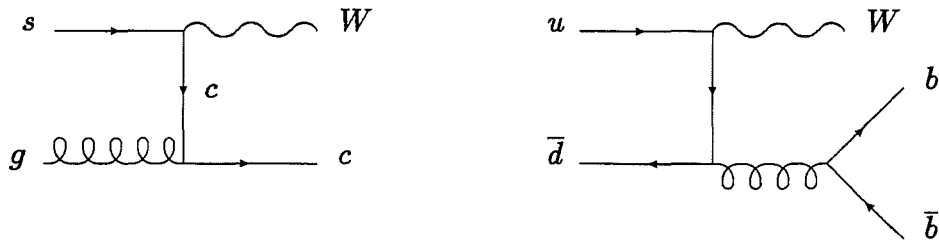


Figure 3-6: Diagrams for heavy flavor production in W +multijet events. The left diagram shows Wc production where a gluon and a strange quark from the proton or anti-proton sea produce a W boson with an associated charm quark. The diagram on the right shows $Wb\bar{b}$ production where a final state gluon splits into a $b\bar{b}$ pair. $Wc\bar{c}$ events are produced in an identical manner.

decay.² This includes non-leptons which pass the lepton selection requirements (such as a pion faking an electron or a muon) as well as electrons from conversions, or muons from pions or kaons which decay in flight. Backgrounds also arise from real heavy flavor tags in W +jets events. Gluon splitting, for instance, can produce $Wb\bar{b}$ or $Wc\bar{c}$ events, and Wc events can be produced from an initial state strange quark as shown in Figure 3-6. Other, relatively small backgrounds from $b\bar{b}$, Drell-Yan, $Z \rightarrow \tau^+\tau^-$ and Diboson (WW, WZ) production (which can appear in the W sample) are also estimated.

3.4 Outline of the Analysis

Chapter 4 describes the data and Monte Carlo samples used in this analysis. Chapter 5 outlines the development of the soft lepton tagging algorithm. Chapter 6 presents the calculation of the $t\bar{t}$ acceptance and the expected yields. Chapter 7 presents the results from applying the soft lepton tagger to the W +

²or the decay of a τ .

≥ 3 jet sample. Chapter 8 contains the calculation of the backgrounds. The significance of the observed excess of events and the $t\bar{t}$ production cross section is calculated in Chapter 9. The properties of the soft lepton tagged events and the calculation of the top mass using soft lepton tagged events are also discussed in Chapter 9. The Appendix contains a variety of tests of the soft lepton tagger and an alternate method for calculating the background.

Chapter 4

Data and Monte Carlo Samples

This analysis uses a variety of data samples which are described together here for simplicity. A trigger path is specified for each sample and quality selection criteria imposed. In general any selection requirement is referred to as a *cut*. Many of the variables used are discussed in detail in Chapter 5 which describes the soft lepton tag algorithm. Monte Carlo samples and generators used for background calculations and other studies are also described here.

4.1 The W Sample

The search for $t\bar{t}$ production begins with a W sample selected by requiring an isolated high P_t electron or muon accompanied by large \cancel{E}_t . The W sample is a subset of the inclusive electron and muon samples which have no \cancel{E}_t or isolation cuts applied. The efficiency of the cuts used to select the inclusive sample are measured using electrons and muons from Z boson decay which are expected to be quite similar to their counterparts from W 's in top events.

4.1.1 The Inclusive Electron Sample

High E_t electrons are identified in the Level 1 trigger by requiring a CEM tower with at least 8 GeV of E_t . The Level 2 trigger requires a CEM cluster with

$E_t > 16$ matched to a CFT track with $P_t > 12$. Because the CFT requirement is only $\sim 90\%$ efficient for electrons with $E_t > 20$ GeV, a second Level 2 trigger path is allowed for this analysis which instead requires 20 GeV of E_t and a CEM cluster with $E_t > 16$. A Level 3 trigger applies loose cuts on some of the variables used for the final selection which are described next.

To increase the purity of the electron sample, a set of quality cuts are applied to the events passing the trigger selection. The cuts are listed in table 4.1 and have a combined efficiency of $81.9 \pm 0.7 \%$. The variables are shown for electrons from a Z sample in Figure 4-2. The requirements are designed to eliminate non-electrons, electrons in jets, and electrons from photon conversions.

The electron track must first be well matched to wire and cathode strip clusters located at shower maximum in the CEM. The quantity Δx (Δz) denotes the distance between the extrapolated track position and the wire (strip) cluster in the $r - \phi$ (z) direction. The profile of the strip clusters (χ^2_{strip}) is required to be consistent with that measured from electrons in testbeam. The ratio of the electromagnetic calorimeter energy to the momentum of the track measured in the CTC (E/p) is required to be close to one. The ratio of energy in the hadronic calorimeter to the electromagnetic calorimeter energy associated with the electron (E_{had}/E_{em}) is required to be less than 5%. The E_{had}/E_{em} and E/p cuts can eliminate events where a charged pion and a nearby pi-zero result in an electromagnetic cluster with an nearby track, or where a charged pion *charge exchange's*, in the coil or the calorimeter and produces a neutral pion ($\pi^+ n \rightarrow \pi^0 p$) resulting in a single track, well matched to an electromagnetic cluster.

Using the z position of the track origin and the z position as measured in the CES, the energy sharing between neighboring towers in η can be predicted based on studies at a testbeam. The LSHR variable is an indicator of how well the cluster matches this prediction and is defined as (all energies in GeV):

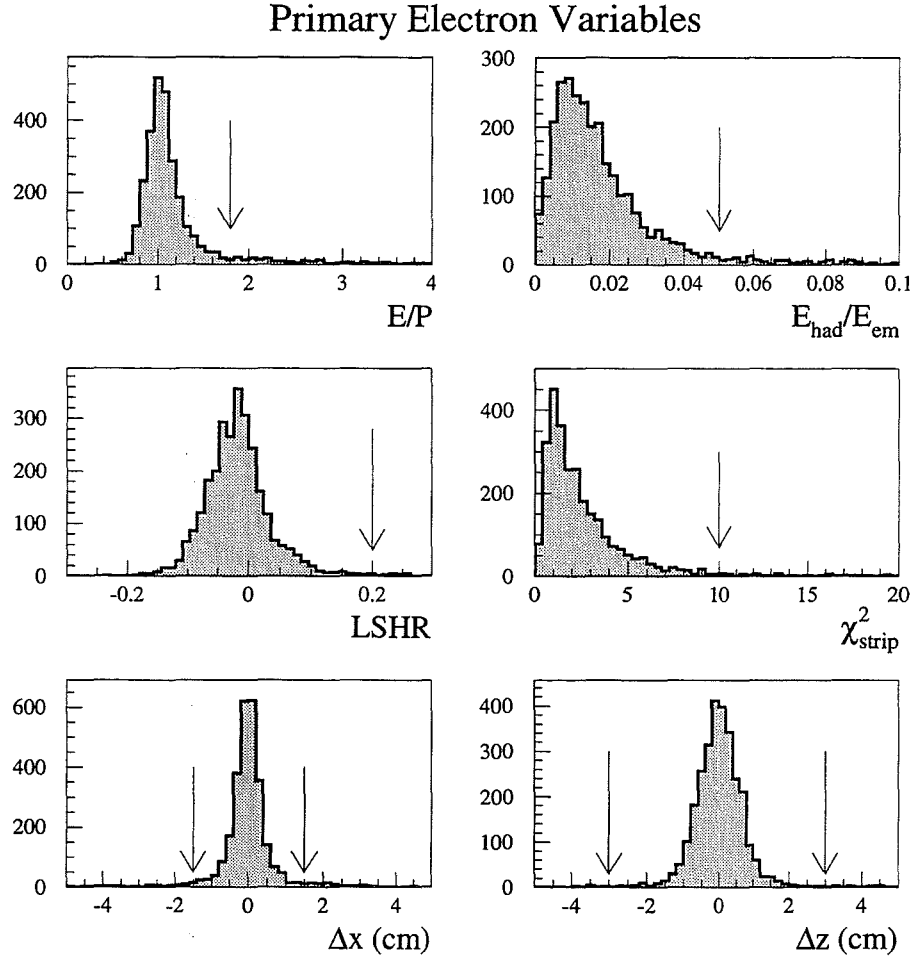


Figure 4-1: Variables used in defining primary electrons, plotted for electrons in a $Z \rightarrow e^+e^-$ sample where the other electron in the event was required to pass tight cuts and to fire the trigger. Arrows indicate the primary lepton selection criteria listed in Table 4.1.

$E_t > 20 \text{ GeV}$
$E/p < 1.8$
$E_{had}/E_{em} < 0.05$
$LSHR < 0.2$
Track Strip/Wire Matching:
$\Delta x < 1.5$
$\Delta z < 3.0$
$\chi^2_{strip} < 10$
$ Z_{electron} - Z_{vertex} < 5 \text{ cm}$
$ Z_{vertex} < 60 \text{ cm}$
<i>Fiducial Requirements</i>

Table 4.1: Inclusive high P_t electron selection requirements.

$$L_{shr} = 0.14 \sum_i \frac{E_i^{obs} - E_i^{pred}}{\sqrt{(0.14\sqrt{E})^2 + \sigma_{pred}^2}}$$

where the sum is over all adjacent towers, E_i^{obs} is the observed energy in tower i , E_i^{pred} is the predicted energy in the tower, σ_{pred} is the uncertainty on the predicted value, and $0.14\sqrt{E}$ is the uncertainty on the measured energy.

Finally, to ensure that the electron energy is well measured, the cluster is required to be in the fiducial region of the CEM, away from tower boundaries in ϕ and uncovered areas such as the $|\eta| = 0$ boundary. The z position of the electron track origin is also required to be within 5 centimeters of a well measured event vertex in the VTX. Because of the projective geometry of the detector, the event vertex is also required to have $|z| < 60 \text{ cm}$.

Conversion Removal

The inclusive sample contains some real electrons which are not from W decay. Photons which interact with material in the detector can convert to electron

$$\begin{array}{c}
\overline{\overline{\left(\begin{array}{l} |\Delta R\phi| < 0.3cm \\ |\Delta Cot(\theta)| < 0.06 \end{array} \right)}} \\
\text{OR} \\
\overline{\overline{\left(VTXoccupancy < 0.2 \right)}}
\end{array}$$

Table 4.2: Conversion Removal

positron pairs. Roughly two percent of π^0 's Dalitz decay at the origin also producing electron-positron pairs ($\pi^0 \rightarrow e^+e^-\gamma$). Fortunately most of these can be identified using the tracking chambers and removed. The cuts, listed in Table 4.2, remove electrons that have a nearby partner track of opposite sign. Since the photon is massless the tracks are required to be close in θ , and to pass within 0.3 cm of each other in the $r - \phi$ plane at the point where the tracks are tangent to each other. To remove conversions where the partner track is not reconstructed (usually because it has too low P_t) electron candidates with fewer than 20% of the expected hits in the VTX are also removed. The efficiency of the conversion removal cuts are estimated to be $90.7 \pm 3.8\%$ for this sample [24]. The *overefficiency* of the removal (the fraction of non-conversions which are removed) is determined by applying the algorithm to a sample of $Z \rightarrow e^+e^-$ events. Of 781 $Z \rightarrow e^+e^-$ candidates, 17 are identified as conversions, yielding an overefficiency of $2.2 \pm 0.6\%$ [24]. Before removal, conversion electrons comprise roughly 30 - 40 % of the inclusive electron sample.

4.1.2 The Inclusive Muon Sample

Muons at CDF are categorized by detector region. Those reconstructed in both the CMU and CMP are called CMUP muons. Muons reconstructed in only one are called CMU-only or CMP-only. Muons in the CMX are sensibly dubbed CMX muons.

Muons are selected for the inclusive sample if they pass a Level 2 trigger

which requires a CFT track with $P_t > 12 \text{ GeV}/c$ pointing to within 5° of a Level 1 muon stub. Some of these triggers had large rates and so were prescaled. To retain good efficiency for top, alternate Level 2 triggers which require a Level 2 calorimeter cluster with $E_t > 15$ in addition to the muon are also employed. CMP-only muons are accepted for this analysis if they pass a Level 2 trigger which requires 35 GeV of \cancel{E}_t (calculated using calorimeter energies only), accompanied by at least two jets. The Level 2 muon triggers used for this analysis are summarized in Table 4.3.

The Level 3 muon trigger runs full offline reconstruction of stubs in the muon chambers, and a pared down version of the offline track reconstruction. The distance between the extrapolated track and the stub in the $r - \phi$ plane (Δx) is required to be less than 10 cm for CMU-only or CMUP muons, 25 cm for CMX muons, and 40 cm for CMP-only muons.

After trigger selection, a final set of muon quality cuts (Table 4.4) is applied to the sample. The variables are shown in Figure 4-2 for muons from a Z sample. The matching cuts are looser for muons in the CMP and CMX because they typically traverse more material than for the CMU (see Figure 5-11) and so incur greater deflections through multiple scattering. High energy muons are minimum ionizing and so the electromagnetic (hadronic) energy in the calorimeter tower associated with the muon is required to contain less than 2 GeV (6 GeV) of energy. Finally the z position of the muon track at its point of closest approach to the beam line is required to be within 5 cm of a well measured event vertex in the VTX, and this event vertex is required to be within 60 cm of the center of the detector. To improve the measurement of the muon momentum, the track is refit with the requirement that it pass through the beam line.

The efficiency of the combined cuts, measured with muons from Z boson decay, is $91.4 \pm 1\%$ for CMX muons, $90 \pm 2\%$ for CMU-only muons, $88 \pm 2\%$ for CMP-only muons and $93.6 \pm 0.7\%$ for CMUP muons.

Level 2 Trigger Paths for Muons	
Triggers	Prescaled
CMU-only muons	
1) $\cancel{E}_t > 35$ GeV, Two Jets with $E_t > 3$ GeV	NO
2) CFT track $P_t > 12$ GeV/c matched to CMU stub	YES
3) CFT track $P_t > 12$ GeV/c matched to CMU stub <i>and</i> One Jet $E_t > 15$ GeV	NO
CMUP muons	
1) $\cancel{E}_t > 35$ GeV, Two Jets with $E_t > 3$ GeV	NO
2) CFT track $P_t > 12$ GeV/c matched to CMU and CMP stubs	NO
3) CFT track $P_t > 12$ GeV/c matched to CMU and CMP stubs <i>and</i> One Jet with $E_t > 15$ GeV	NO
CMP-only muons	
1) $\cancel{E}_t > 35$ GeV, Two Jets with $E_t > 3$ GeV	NO
CMX muons	
1) $\cancel{E}_t > 35$ GeV, Two Jets with $E_t > 3$ GeV	NO
2) CFT track $P_t > 12$ GeV/c matched to CMX stub	YES
3) CFT track $P_t > 12$ GeV/c matched to CMX stub <i>and</i> One Jet with $E_t > 15$ GeV	YES

Table 4.3: Level 2 trigger paths for primary muons. The Level 1 trigger identifies CMU-only, CMUP and CMX muons by reconstructing the appropriate stubs. There is no Level 1 or Level 2 stub trigger for CMP-only muons. CMP-only muons are accepted at Level 2 with a \cancel{E}_t + jets trigger (see table) and typically with a calorimeter trigger at Level 1.

Primary Muon Variables

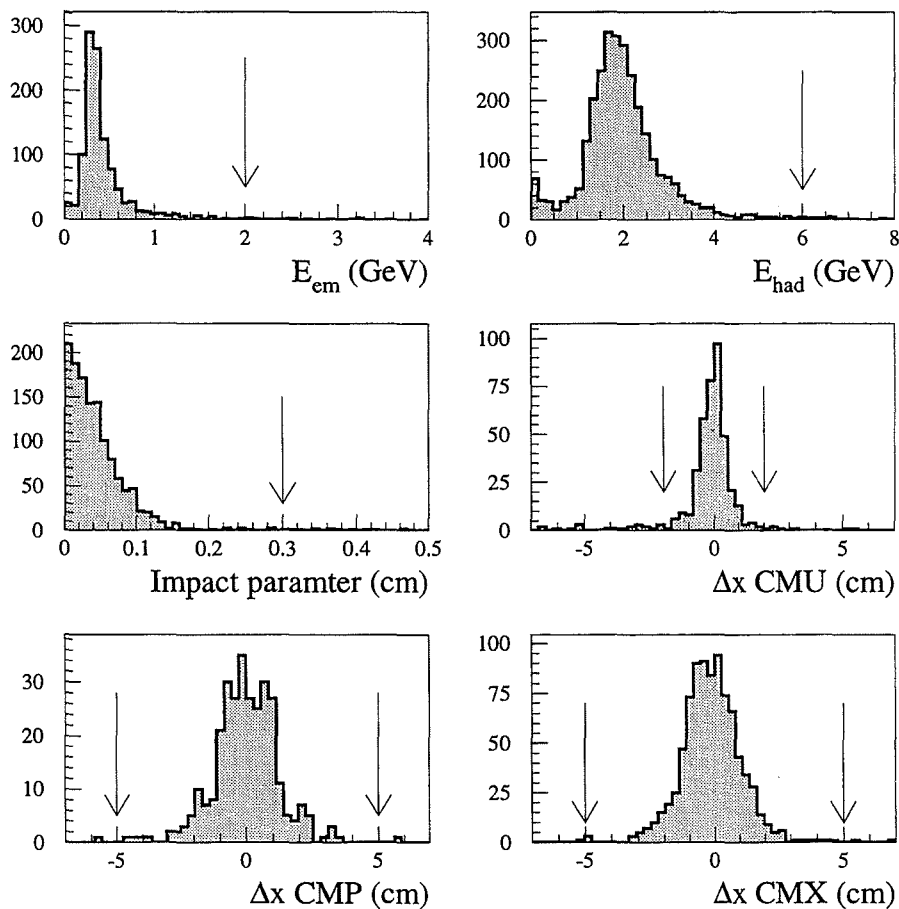


Figure 4-2: Variables used in defining primary muons, plotted for muons in a $Z \rightarrow \mu^+ \mu^-$ sample where the other muon in the event was required to pass tight cuts and to fire the trigger. Arrows indicate the primary lepton selection criteria listed in Table 4.4.

$P_t > 20 \text{ GeV}/c$
$E_{em} \text{ energy in tower} < 2.0 \text{ GeV}$
$E_{had} \text{ energy in tower} < 6.0 \text{ GeV}$
$Impact \text{ parameter} < 3mm$
$ Z_{muon} - Z_{vertex} < 5cm.$
$ Z_{vertex} < 60cm.$
<i>Track Stub matching:</i>
$ \Delta X _{CMU} < 2.0 \text{ cm}$
OR
$ \Delta X _{CMP} < 5.0 \text{ cm}$
OR
$ \Delta X _{CMX} < 5.0 \text{ cm}$

Table 4.4: Inclusive high P_t muon selection requirements.

4.1.3 Final W Sample Selection

From the inclusive lepton samples the final W sample is made by requiring that the primary lepton be isolated and accompanied by significant imbalance of transverse energy (E_t) indicating the presence of a neutrino.

Isolation

For electrons, a quantity called *isolation* is defined as

$$Iso^{ele} = \frac{E_t^{cone} - E_t^{ele}}{E_t^{ele}}$$

where E_t^{cone} is the calorimeter energy contained in a cone of radius $R = 0.4$ centered on the electron cluster centroid, and E_t is the calorimeter energy of the electron. For muons a similar quantity is defined

$$Iso^{muon} = \frac{E_t^{cone} - E_t^{tower}}{P_t^{muon}}$$

where E_t^{tower} is the amount of energy found in the tower associated with the muon track, and P_t^{muon} is the transverse momentum of the muon track. Leptons are required to have $Isolation < 0.1$. The efficiency is measured using a PYTHIA generated $t\bar{t}$ Monte Carlo sample to be $88.8 \pm 0.9\%$ ($84.8 \pm 0.7\%$) efficient for electrons (muons) with $P_t > 20$ GeV/c. The quoted errors are statistical; systematic uncertainties related to the overall acceptance are treated in Section 6.4.

Z Boson Removal

The cuts listed in Table 4.5 are next applied to the data to remove events where the lepton originates from the decay of a Z boson. The cuts identify events with a second, same flavor lepton candidate passing loose quality cuts. Tracks without stubs are considered muon candidates if they are minimum ionizing (see Table 4.5). Events where this second lepton makes an invariant mass between 75 and 105 GeV/ c^2 with the primary are removed as consistent with Z boson decay. Where there is a track associated with the second lepton (plug electrons do not require one) it must have the opposite charge of the primary lepton. Some Z 's will remain in the sample, where the second lepton is not identified by this procedure. This is taken into account in the calculation of the W purity (Section 8.6) and as a background source for soft lepton tags (Section 7.1.4).

Missing Transverse Energy (\cancel{E}_t)

Since the neutrino from the decay of the W is not expected to interact in the detector, its presence is inferred indirectly from a large imbalance in observed energy transverse to the beam direction (\cancel{E}_t or *missing E_t*). First a raw \cancel{E}_t for the event is calculated as the negative of the vector sum of the transverse energy in the calorimeter towers. The raw \cancel{E}_t is then corrected for primary muons by vectorially adding the P_t of the muon track, and subtracting the E_t in the calorimeter tower associated with the muon. Similar corrections

Electron cuts:
$E_t > 10 \text{ GeV}$
$E_{had}/E_{em} < 0.12$
$Isolation < 0.2$
$E/p < 2.0$ (if in CEM)
Muon cuts:
$P_t > 10 \text{ GeV}/c$
If associated with a stub
$E_{em} \text{ energy in tower} < 5.0 \text{ GeV}$
$E_{had} \text{ energy in tower} < 10.0 \text{ GeV}$
$ \Delta X _{CMU,CMP \text{ or } CMX} < 5.0 \text{ cm}$
$Isolation < 0.1$
If no stub
$E_{em} \text{ energy in tower} < 2.0 \text{ GeV}$
$E_{had} \text{ energy in tower} < 6.0 \text{ GeV}$
$ \eta < 1.1$
$Isolation < 0.2$

Table 4.5: Loose lepton quality cuts for the second (non-primary) muon in Z removal.

	Muons	Electrons
Quality Cuts	87892	121123
Badrun Removal	84251	115699
Trigger Requirement	79955	111895
$Isolation < 0.1$	51102	76791
$\cancel{E}_t > 20 GeV$	38602	57675

Table 4.6: Events remaining in lepton sample after cuts.

are applied for minimum ionizing tracks with $P_t > 10$ GeV/c that pass loose matching requirements with a muon stub, or extrapolate to regions not covered by the CMU, CMP or CMX detectors [13]. Finally events are selected with this corrected $\cancel{E}_t > 20$ GeV. This requirement is $83 \pm 0.1(\text{stat})$ % efficient for top events with a primary lepton (systematic uncertainties are treated in Section 6.4) and reduces the number of isolated lepton candidates with three or more jets by 35%.

Bad Run Removal

To ensure the good quality of the dataset, runs where there were known detector or readout problems are removed. This eliminates 4 % of the events in the inclusive sample.

Events in the W Sample

Table 4.6 shows the number of events surviving each of the W selection criteria. Because the transverse momentum of the neutrino is unmeasured, the full invariant mass of the W cannot be reconstructed. Using only the momentum and energies measured in the transverse direction gives a related quantity, the *transverse mass*

$$M_t = \sqrt{(|P_t^{lep}| + \cancel{E}_t)^2 - (\vec{P}_t^{lep} + \vec{\cancel{E}}_t)^2}$$

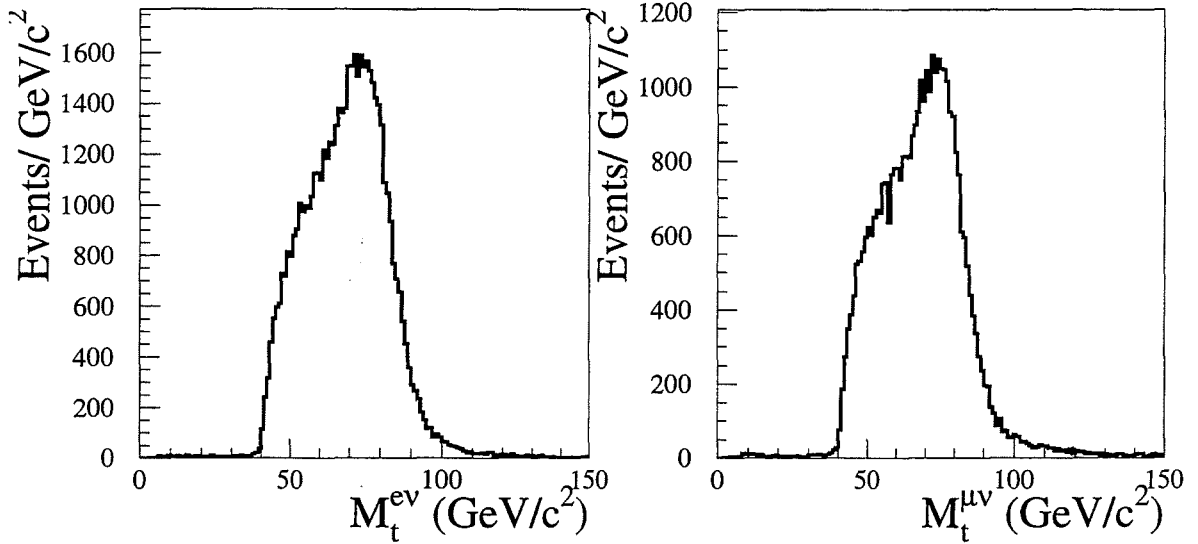


Figure 4-3: Transverse mass of \cancel{E}_t and primary electron (left) and muon (right) for events in the W sample.

where P_t^{lep} is the transverse energy of the electron measured in the calorimeter or the transverse momentum of the muon measured from the muon track. This is the quantity used in the measurement of the W mass [88]. Figure 4-3 shows the transverse mass of the lepton and \cancel{E}_t for events in the W sample which exhibits the expected smeared Jacobian peak at roughly the W mass (80 GeV/c²).

Dilepton Removal

To avoid overlap with the dilepton channel search for top, events passing the cuts used in that analysis [12] [31] are removed from the sample. The dilepton analysis begins with the inclusive lepton samples described here, and then looks for an additional, opposite charge electron or muon with $E_t > 20\text{GeV}$. The analysis also requires $\cancel{E}_t > 25\text{ GeV}$ and two jets with $E_t > 10\text{ GeV}$. If the magnitude of the \cancel{E}_t is less than 50 GeV, the direction of the \cancel{E}_t is required to be at least 20° away from the nearest jet or lepton in azimuth. Monte Carlo studies indicate that removing these events reduces the SLT acceptance by 15

%. In most of these cases the soft lepton is from the decay of the second W , and so the event is correctly categorized as a dilepton event.

4.2 Monte Carlo Samples

This analysis employs three different Monte Carlo generators called ISAJET [75], HERWIG [76] and PYTHIA [77]. All generators begin by convoluting parton distribution functions with a tree level matrix element for the desired process. In ISAJET, initial and final state radiation are incorporated in the leading log approximation and the final state quarks and gluons are fragmented into hadrons independently. HERWIG and PYTHIA employ more sophisticated treatments of jet production which take into account color correlations between the initial and final state partons. HERWIG has been shown to reproduce well the observed properties of multijet events in the data [83]. Each of the Monte Carlo generators output a list of four-vectors of stable particles. A CDF detector simulation package (QFL) takes the generated particle list as input and simulates the event. The simulation uses parameterizations of detector response based on testbeam measurements. This method is significantly faster than the GEANT style simulation which models particle showers and interactions in material in detail. In QFL, all short-lived particles except for B 's are decayed according to the branching fractions compiled by the Particle Data Group [26]. The B hadron decay branching fractions and decay kinematics are taken from the measurements of the CLEO collaboration. This is implemented with a software module called CLEOMC [86].

Top quark pair events are generated for this analysis using the HERWIG generator and checked using PYTHIA. Where needed, background processes are generated with ISAJET or HERWIG.

VECBOS is a matrix element Monte Carlo for generating W +jets events and contains the leading order matrix elements for $W+0,1,2,3$ and 4 jets. VECBOS produces the four vectors of the outgoing partons, and a weight for the

event. In this analysis, HERWIG fragmentation (called HERPRT) is used to turn the VECBOS output into a list of stable particles which is passed on to QFL. VECBOS reproduces well the observed kinematics and rates of W +jets events in the data [95].

Unless otherwise stated all Monte Carlo samples are generated using the MRSD0' parton distribution functions which describe well the observed W asymmetry at CDF [98].

4.3 The Low P_t Electron Sample

The development of a soft lepton tagger requires a good understanding of the detector response to low P_t leptons. A high purity sample of low P_t electrons can be easily obtained by identifying electron-positron pairs produced from photon conversions in a manner similar to that described in Section 4.1.1. First a sample of electron candidates is selected with a Level 2 trigger that requires an electromagnetic cluster with $E_t > 8$ GeV associated with a CFT track of $P_t > 7.5$ GeV/c. To avoid biasing the strip and wire information, a trigger that additionally requires the presence of a cluster in the CES is not used. As in Section 4.1.1 an opposite sign partner track is searched for with $|\Delta C_{ot}(\theta)| < 0.03$ and $\Delta S < 0.2$ cm. Figure 4-4 shows these variables and the radius where the two tracks converge. The ΔS variable is defined as

$$\Delta S = D - \rho_1 - \rho_2$$

where D is the distance between the center of the circles which describe the tracks in the $r - \phi$ plane, and ρ_1 and ρ_2 are the radii of the circles. The ΔS distribution is peaked slightly below zero ($\approx -500\mu m$) because the CTC track reconstruction tends to pull tracks to pass near the beamline, inflating ρ_1 and ρ_2 . This is noticeable only because many of the conversion tracks originate many centimeters from the beamline [89].

Since photons can only convert in matter, the tracks are expected to orig-

inate from a location where there is considerable material in the detector. The plot of the radius of convergence shows a large peak around 25 cm corresponding to the inner wall of the CTC and the outer wall of the VTX. Most fake conversions due to random overlapping tracks should have a conversion radius near the interaction point since this is where the majority of tracks originate. To insure a clean sample of conversions, only events in this region ($20 < R_{conversion} < 30$ cm) are selected. The two conversion tracks are required to extrapolate to different calorimeter towers in the CEM so that each can be examined independently. A study repeating this procedure with same sign pairs indicates that the sample is over 99 % pure. The radius of convergence for same sign tracks passing the ΔS and $\Delta Cot(\theta)$ requirements are shown in Figure 4-4.

4.4 The Low P_t Muon Sample

A pure sample of low P_t muons is obtained from the data by reconstructing the decay $J/\psi \rightarrow \mu^+ \mu^-$. The Level 1 trigger accepts J/ψ candidates by requiring two low P_t muon stubs (the trigger rate for a single low P_t stub is unmanageably large). At Level 2 the stubs are matched to CFT tracks of at least 2.2 GeV. At Level 3, opposite sign pairs making an invariant mass consistent with a J/ψ are retained. The reconstructed mass of the two muon candidates, after offline tracking is run, is shown for CMU, CMP, CMUP and CMX muons in Figure 4-5. Assuming a flat background, the $\pm 3\sigma$ area around the center of the peak is $\sim 75\%$ pure $J/\psi \rightarrow \mu^+ \mu^-$. In practice it is a simple matter to do a background subtraction using the sidebands to measure the efficiency for real muons to pass a particular cut.

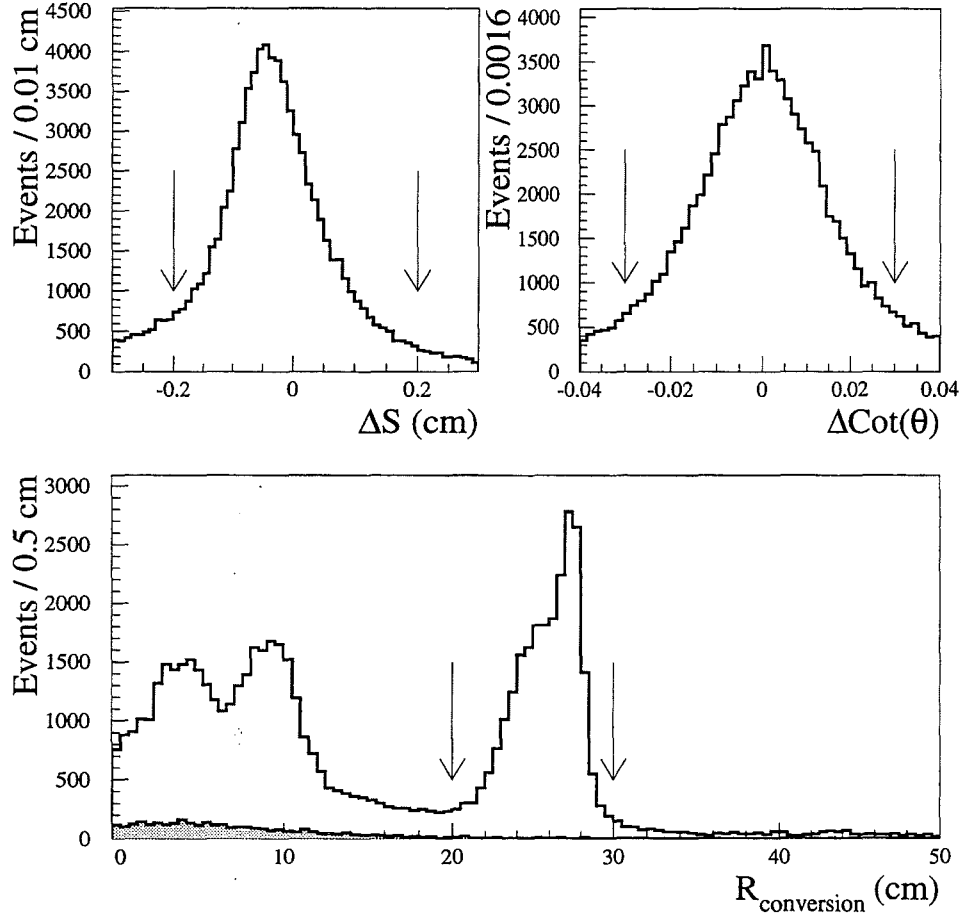


Figure 4-4: Variables used to select the conversion sample. ΔS is the distance between the two tracks in the transverse plane at the point where the tracks are tangent. $\Delta \text{Cot}(\theta)$ is the difference between the cotangents of the polar angles of the two tracks and measures their separation in the r-z plane. The bottom plot shows the radius of convergence of the tracks. The peak between 20 and 30 cm corresponds to the outer wall of the VTX and the inner wall of the CTC. The shaded histogram shows the radius of convergence for same-sign pairs which appears close to the origin as expected for tracks produced in the primary interaction. Arrows indicate the requirements for the low P_t electron sample (Section 4.3

).

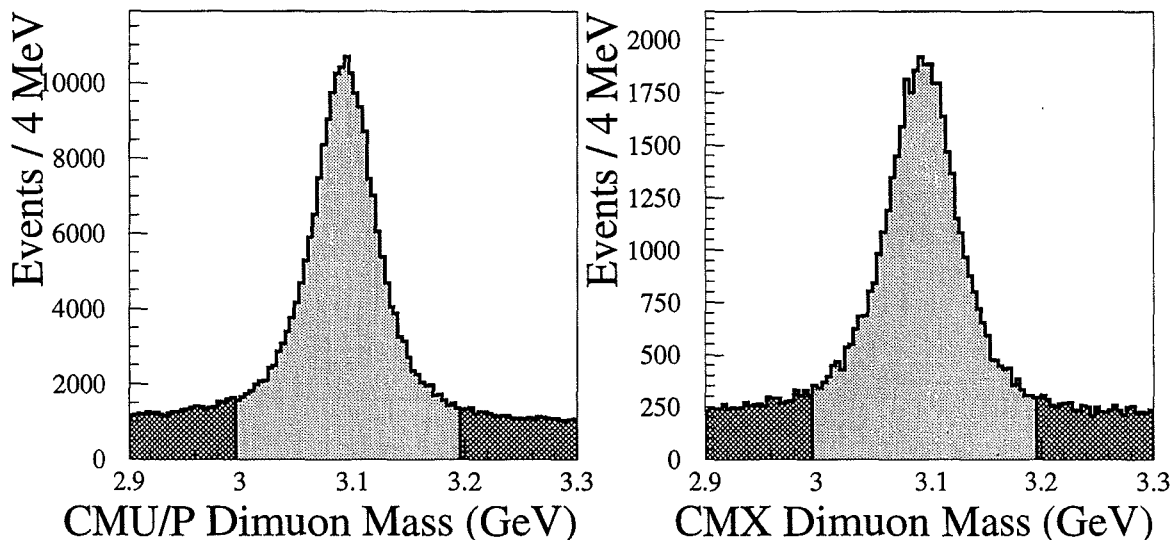


Figure 4-5: Dimuon mass for muons in J/ψ sample. In the plot on the left, one muon is in the CMU,CMP or both. In the plot on the right one muon is in the CMX. The shaded areas show the peak (light) and sideband regions (dark). A distribution for real muons is obtained by plotting the desired quantity for tracks in the peak region and subtracting the same distribution for tracks in the sidebands.

4.5 The Jet Sample

The primary method for estimating the main background for this analysis (consisting of fake soft lepton tags and tags from $Wb\bar{b}$ and $Wc\bar{c}$ production) involves a study of soft lepton tags in a sample of jets. The Level 2 trigger selects jet events by searching for calorimeter clusters with E_t greater than 20, 50 or 70 GeV. Due to the large cross section for multijet production, these triggers are all prescaled and typically have at least one jet recoiling opposite the trigger jet in ϕ . Since these events should not be enhanced in heavy flavor jets, they are termed *generic* jets. The E_t spectrum of jets in these events are shown in Figure 4-6.

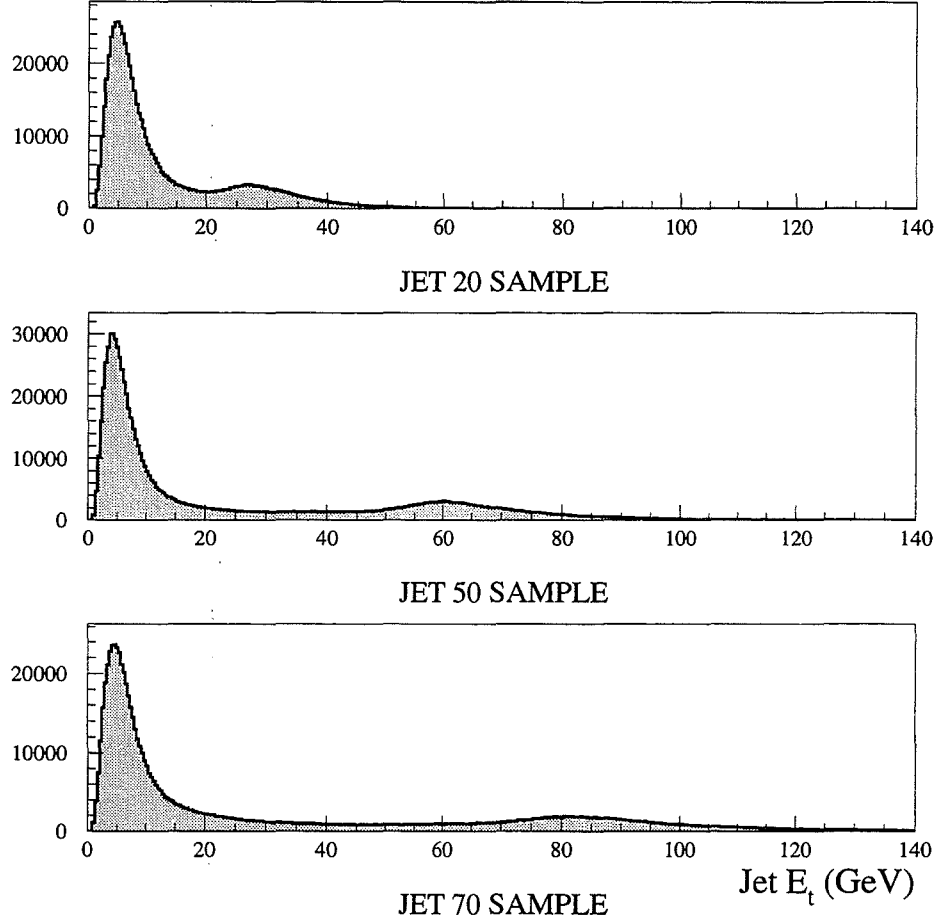


Figure 4-6: Transverse energy of jets in events collected with a trigger that requires a calorimeter cluster with a minimum E_t of 20 GeV (top), 50 GeV (middle), and 70 GeV (bottom). The fake rate is calculated using tracks with $P_t > 2$ GeV/c, 75% of which are contained in jets with $E_t > 15$ GeV.

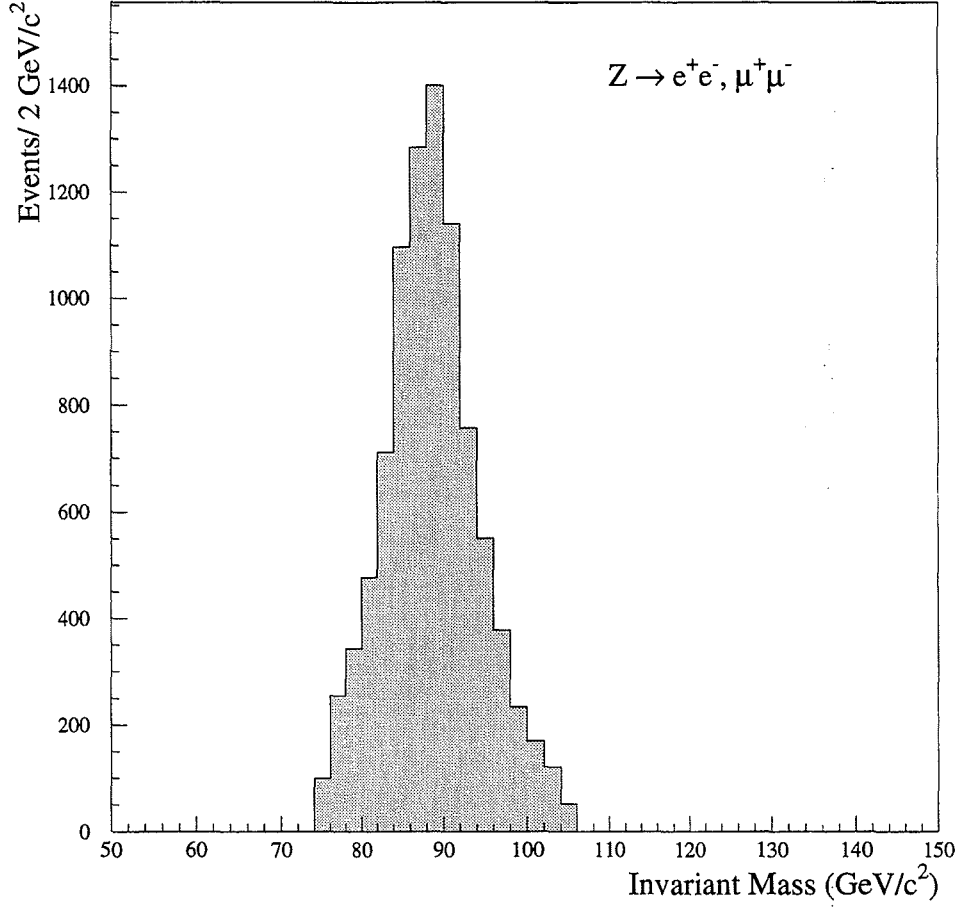


Figure 4-7: Invariant mass for lepton pairs in the Z sample described in the text. A explicit cut requiring $75 \text{ GeV}/c^2 < M_{\ell\ell} < 105 \text{ GeV}/c^2$ has been applied to the data.

4.6 The Photon Sample

A sample of events containing a photon and at least one jet provides an independent sample in which parts of the background calculation can be tested. Events are selected with a Level 2 photon trigger which requires a CEM cluster with $E_t > 23 \text{ GeV}$. Offline photons are selected by requiring that no 3-D tracks extrapolate to the cluster. The final sample selection requires an additional jet with at least 15 GeV of transverse energy.

4.7 The Z Sample

A sample of Z bosons is selected from the inclusive electron and muon samples by looking for a second isolated, same flavor, opposite charge lepton that makes an invariant mass between 75 and 105 GeV with the primary. The invariant mass for pairs in this sample is shown in Figure 4-7.

Chapter 5

The Soft Lepton Tag (SLT) Algorithm

This chapter describes in detail the SLT algorithm. The overall philosophy has been to develop a track based b-tagger. The cluster based electron finding algorithm used to identify primary electrons has many implicit isolation requirements, and hence an unnecessarily low efficiency for finding electrons from b's which are expected to lie in jets. The standard CDF algorithm already defines muons in a track based manner.

The soft lepton algorithm begins by extrapolating all tracks which pass a loose set of quality cuts to the CEM or the muon chambers. If the track extrapolates to a fiducial region of the detector it is termed a "candidate" track. A set of selection criteria are then applied to the candidate tracks which define an SLT electron or muon.

As will be discussed in Chapter 8 most of the SLT background for the top search uses a fake rate derived from tracks in "generic" jet events which are not expected to be enriched in heavy flavor. The fake rate is defined as the fraction of candidate tracks in these events which are soft lepton tagged. The goal has been to achieve fake rates per track of less than 1% while retaining high efficiency for electrons and muons from b decay. At a fake rate of 1% per track, background levels in the W sample are sufficiently low for an excess of tagged

events from top to be observed.

The efficiency for a set of requirements to tag an electron or muon from a b decay can be broken down into two parts. First, there are variables which are expected to be isolation independent, that is, largely unaffected by nearby particles. The quantity dE/dX , calculated with the CTC and used in electron identification, is a perfect example of this; since the dE/dX value comes directly from the hits on the track, it should depend minimally on the surrounding environment. In the second category are variables which are quite sensitive to nearby particles; the quantity E_{had}/E_{em} , for instance, depends quite strongly on the surrounding jet activity.

Fortunately, samples of low P_t electrons (from photon conversions) and muons (from J/ψ 's) are available in the data which can be used to measure the efficiency of the isolation independent cuts. For muons, all cuts are determined using J/ψ muons. Most electron cuts are isolation independent and can be derived from the conversion sample except for the E_{had}/E_{em} and E/p cuts which must be studied using Monte Carlo simulated events. These assumptions are critical for the calculation of the $t\bar{t}$ cross section and are evaluated in more detail in Section 5.1.8.

5.1 Soft Electron Tagging

5.1.1 Fiducial requirements

All candidate soft electron tracks are required to have at least two axial and two stereo superlayers used in their reconstruction, and an impact parameter less than 3 mm. This guarantees a well measured 3D track, and removes some background from photon conversion tracks with large impact parameters. Tracks are further required to have $P_t > 2$ GeV/c, which retains good efficiency for leptons from b 's from top decay (see Figure 3-5), while rejecting tracks from generic jet events which fall more quickly with P_t . The search is further

restricted to electrons in the CEM detector, as 3D tracking is unavailable for $|\eta| > 1.6$ and the preradiator does cover $|\eta| > 1$.

Tracks passing these requirements are extrapolated to the central wire and strip chambers (CES) which are positioned at shower maximum ($R=184$ cm) in the CEM. Fiducial tracks are defined as those that extrapolate to a point at least 2 cm away from the phi edge of a tower (towers are 48 cm wide) and ~ 6 cm from the z boundary of the arch (an arch is 249 cm long). The “chimney” region of the CEM where there is a gap in the calorimeter coverage that allows for access to the solenoid is also excluded from the fiducial region. Finally, the track is extrapolated to the Central Preradiator (CPR) and is required to be away from the edges of the chambers. The fiducial regions were determined by studying the response of the detectors to electrons from conversions as a function of position.

5.1.2 Central Strip Chamber Requirements

Electrons from b decay are typically imbedded in jets. Ideally, soft electrons could be selected in an isolation independent manner without having to resort to the gross calorimeter energies which are likely to be effected by the other nearby particles. Electron clusters are typically much smaller than a calorimeter tower and so electron identification begins with the CES which has finer segmentation. In the following, the number of strips or wires used to determine each variable has been chosen somewhat arbitrarily, though the general idea was to keep the number to a minimum. In calculating the electron energy, five strips (or wires) were necessary to span the electron shower. For measurement of the shower position, however, it was found necessary to use only three.

The CES electron energy is obtained by summing up the energy in the 5 strips (Es_5) and 5 wires (EW_5) around the extrapolated track position. Figure 5-1 shows this energy divided by the momentum of the track for electrons from the conversion sample (Section 4.3). The plots exhibit peaks around 1 indicating that this is a good measure of the electron energy. The large

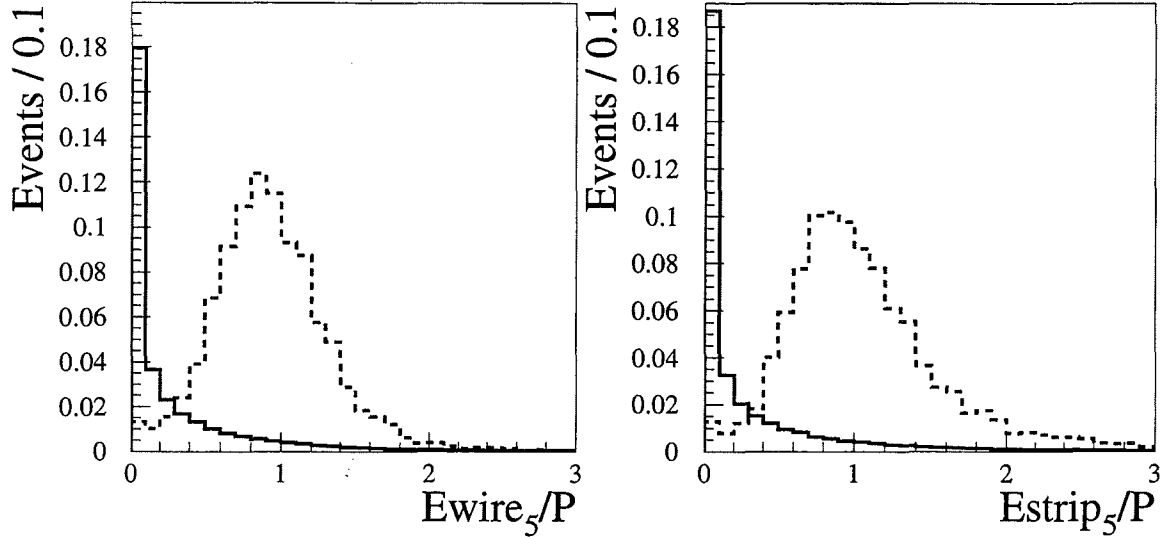


Figure 5-1: E/p for five channel wire (left) and strip (right) clusters. Dashed histograms are from conversion electrons and are normalized to unit area. Solid histograms are from a sample of candidate tracks in jet events and have been normalized to an area of $1/3$ to show shape. All samples have high statistics and so error bars have been omitted.

width of the distribution arises because the CES amounts to a single sampling layer calorimeter, and so has limited resolution. There is also some variation with momentum due to shower fluctuations in the calorimeter. Figure 5-1 also shows Es_5/P and EW_5/P for candidate (background) tracks in jets normalized to area $1/3$ to show the shape of the distributions. Hadrons shower later in the calorimeter than do electrons and so deposit less energy in the CES. The CES measured E/p is required to be larger than a minimum value (see Table 5.1) which varies with momentum from 0.5 to 0.24 to account for the effect mentioned above. Figure 5-2 shows E/p as a function of the track momentum and the minimum value imposed by the cut. The efficiency of this requirement for conversion electrons and for tracks in generic jets is summarized in Table 5.2.

Next, the CES cluster position is calculated by taking the energy-weighted mean of the 3 wires and 3 strips around the extrapolated track position, which give azimuthal and z position information respectively. Using electron tracks

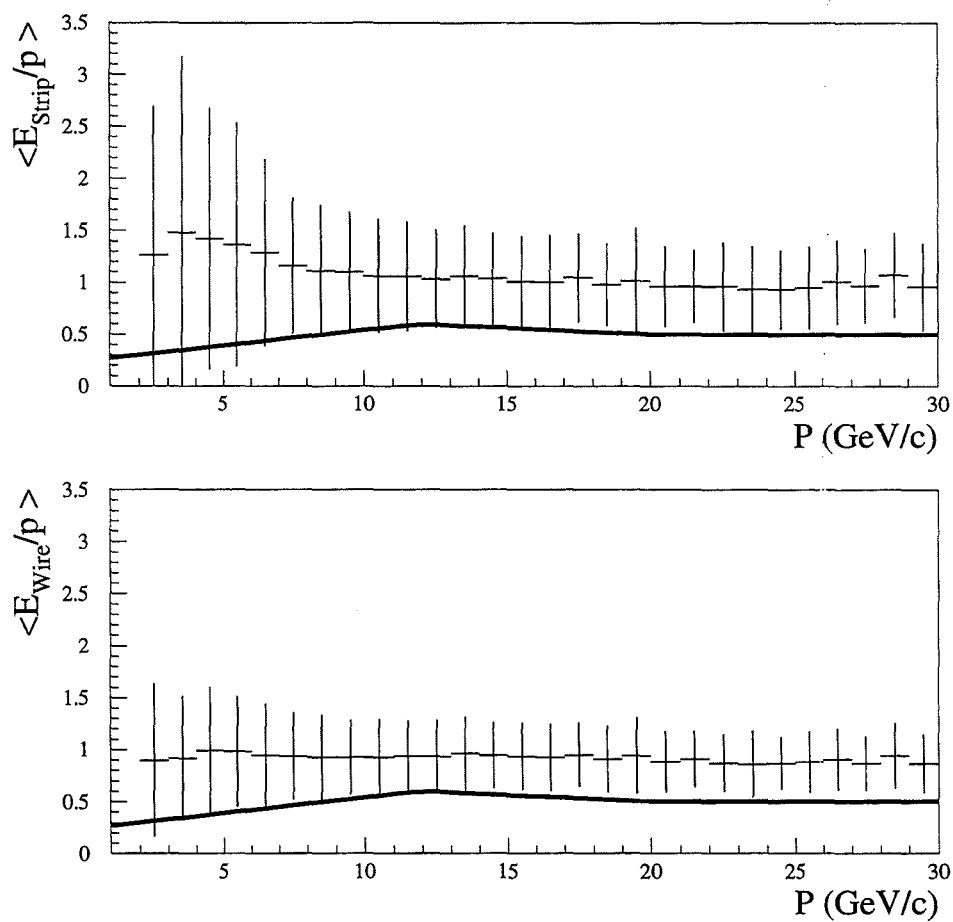


Figure 5-2: The mean E/p measured in the CES for conversion electrons as a function of the track momentum. The error bars mark one standard deviation from the mean. The superimposed line is the minimum required E/p for soft electrons. (See Table 5.1.)

from conversions, alignment corrections are calculated on a wedge-by-wedge basis, and the distance between the extrapolated track position and the cluster means ($\delta x, \delta z$) are required to be less than values determined from the photon conversions. To account for shower fluctuations in the calorimeter, these cuts are also momentum dependent ($|\delta z| < 2$ cm, and $|\Delta x| \leq \text{Max}(0.7 \text{ cm}, 1.82 \text{ cm} - 0.1867P)$ with P in GeV/c). Finally the shape of 7 channel CES clusters ($\chi^2_{wire}, \chi^2_{strip}$) are required to be consistent with that of an electron.

5.1.3 The Central Preradiator Requirement

The preradiator energy is determined by extrapolating the track to the CPR and summing the charge of the 3 surrounding CPR wires. Figure 5-3 shows this quantity (Q_{CPR}) for electrons from conversions and candidate tracks in jet events. Electrons at large values of $|\eta|$ have a longer path length through the solenoid coil than do electrons traveling perpendicular to the beam line ($\eta=0$), causing the CPR response to vary with polar angle. (See Figure 5-3.) Tracks at small $\text{Sin}(\theta)$ interact with more material and produce more secondary particles which deposit additional charge in the CPR. The CPR cut value used (Table 5.1) contains an angular dependence (P/P_t) to correct for this effect. The minimum cut value increases with P/P_t and corresponds to four minimum ionizing particles.

5.1.4 The dE/dX Requirement

The ionization rate for a track (dE/dX) is calculated offline from the charge associated with the CTC hits used in the track. Quality cuts are applied to the charge pulses and geometric effects are corrected for. The resulting quantity (Q_{CTC}), plotted for electrons, pions, muons and protons, exhibits the expected dE/dX behavior (Figure 5-4). Figure 5-5 shows the Q_{CTC} distribution for electrons from conversions and candidate tracks from jet events. The Q_{CTC} distribution is essentially independent of momentum over the range of interest

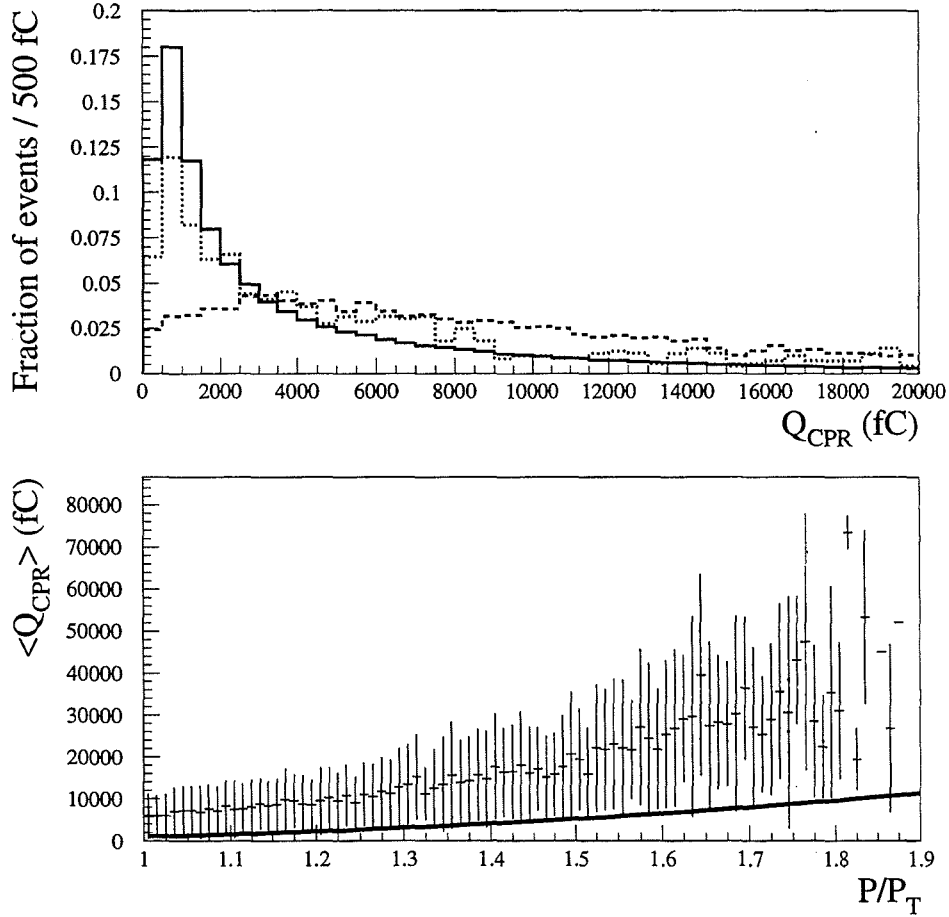


Figure 5-3: (Top) Preradiator response (Q_{CPR}) for conversion electrons (dashed), for candidate tracks in jet events with no id cuts applied (solid) and with all other cuts applied (dotted). All distributions have been normalized to unit area. (Bottom) Mean Q_{CPR} as a function of P/P_t . The error bars denote one standard deviation from the mean. Electrons deposit more energy in the CPR at large P/P_t where they have a longer path length through the the solenoid coil. The solid line shows the minimum CPR requirement listed in Table 5.1.

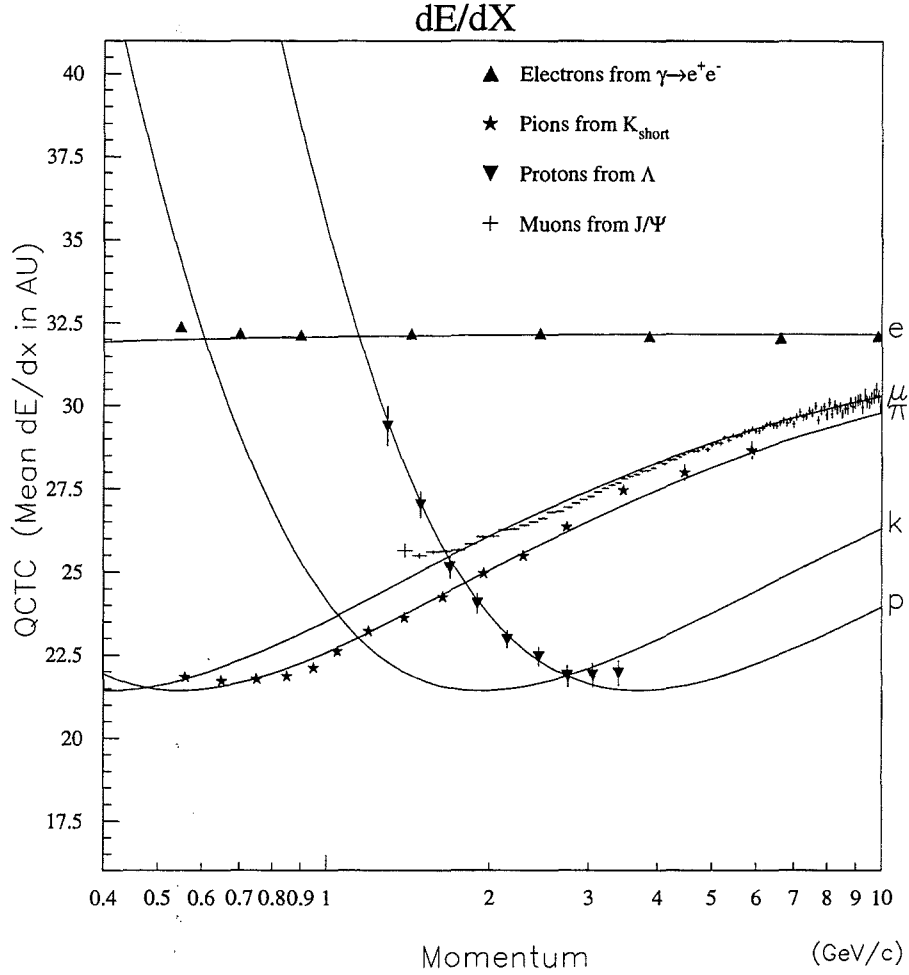


Figure 5-4: Q_{CTC} (dE/dX) as a function of momentum for electrons, pions, protons and muons from the data. This plot was prepared by Hovannes Keutelian [94]. Also shown is the predicted curve for kaons.

for electrons, but depends strongly on momentum for tracks in generic jets. For $P > 12$ GeV/c there is little separation between the two and so no cut is applied.

5.1.5 The E/p and E_{had}/E_{em} Requirements

Without further cuts, the electron fake rate would be 2.5 % per track (on average). Additional rejection can be achieved by constructing calorimeter based E/p and E_{had}/E_{em} quantities similar to those used for primary lepton selection. For both these variables, the candidate track is extrapolated to a tower in the central calorimeter. The electron energy is taken as the energy

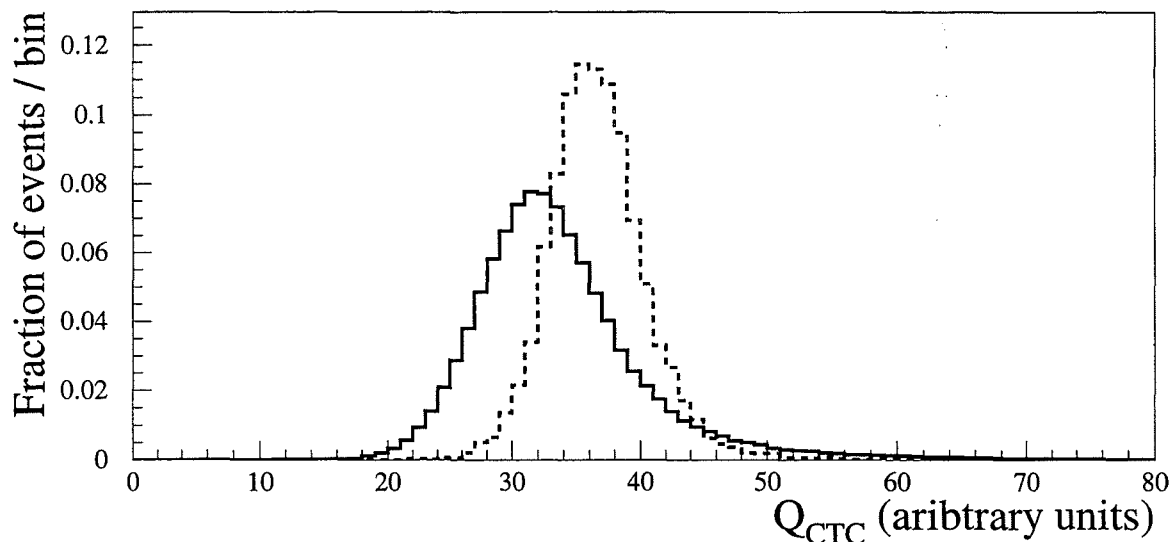


Figure 5-5: Q_{CTC} (dE/dX) for conversion electrons (dashed) and candidate tracks from jet events (solid). Q_{CTC} units are arbitrary. Both distributions are normalized to unit area.

in that tower plus the energy in the adjacent η tower if the track extrapolates to within 2 cm of the boundary between the two. This is in contrast to the primary electron clustering which uses 3 towers by default. The quantity E/p is defined as the electromagnetic component of the energy divided by the track momentum, and E_{had}/E_{em} as the ratio of the hadronic to electromagnetic components in the tower or towers. The efficiency of these cuts must be determined from Monte Carlo since they will certainly be effected by nearby particles in a b-jet. Figure 5-6 shows these distributions for soft electrons from b's in HERWIG $t\bar{t}$ Monte Carlo events with $M_{top} = 175 \text{ GeV}/c^2$. The efficiency of requiring $E_{had}/E_{em} < 0.1$ and $0.7 \leq E/P \leq 1.5$ measured from this Monte Carlo sample are listed in Table 5.2 . Without any other id requirements, the E/p (E_{had}/E_{em}) cut is $72 \pm 1\%$ ($56 \pm 2\%$) efficient for electrons from $b \rightarrow e$ and only $48 \pm 2\%$ ($35 \pm 2\%$) efficient for electrons from $b \rightarrow c \rightarrow e$, considerably lower than the for the previously described cuts. Applied together the E_{had}/E_{em} and E/p cuts are $46 \pm 2\%$ ($23 \pm 2\%$) for electrons from b (c) decay. Together, they reduce the fake rate by a factor of 8.5 to 0.3 % per track which

is substantial enough to merit their application.

5.1.6 Conversion Removal

While photon conversions provide an excellent source of low P_t electrons for selecting the id requirements, they also present a serious source of background for tagging b's. Photons in jets arise largely from the decay of neutral pions and can convert in the detector, producing e^+e^- pairs which may appear as soft lepton tags. Roughly 2 % of neutral pions also produce electrons directly by decaying to a photon and an e^+e^- pair (the “Dalitz” decay, $\pi^0 \rightarrow \gamma e^+e^-$).

To minimize the number of SLT's which arise from these sources, a *conversion veto* is applied. The veto removes all tracks from consideration which pass a looser version of the standard conversion cuts (Section 4.3). The cuts on ΔS and $\Delta C \otimes (\theta)$ remain unchanged but now all conversion radii less than 50 cm are allowed. A loose requirement that the invariant mass of the two tracks be less than 500 MeV/c² is also applied. The overefficiency of these cuts is expected to be small, on the order of a few percent. The conversion removal is included in the Monte Carlo version of the soft lepton tagger so this effect should be modeled when calculating $t\bar{t}$ acceptance.

As will be discussed in Appendix D, roughly 22% of soft electron tags in generic jets are expected to be real electrons that do not originate from heavy flavor decay. Presumably these are largely residual electrons from photon conversions or Dalitz decays.

5.1.7 Efficiency and Rejection Power of Algorithm

Table 5.2 lists the efficiency of the soft electron identification cuts for conversion electrons and candidate tracks in jet events. The conversion tracks are required to have $P_t < 5$ GeV/c to avoid the momentum bias of the trigger so that the two samples have similar P_t spectra. Figure 5-7 shows the P_t spectrum of tracks in the conversion sample and the jet sample. This table is meant only

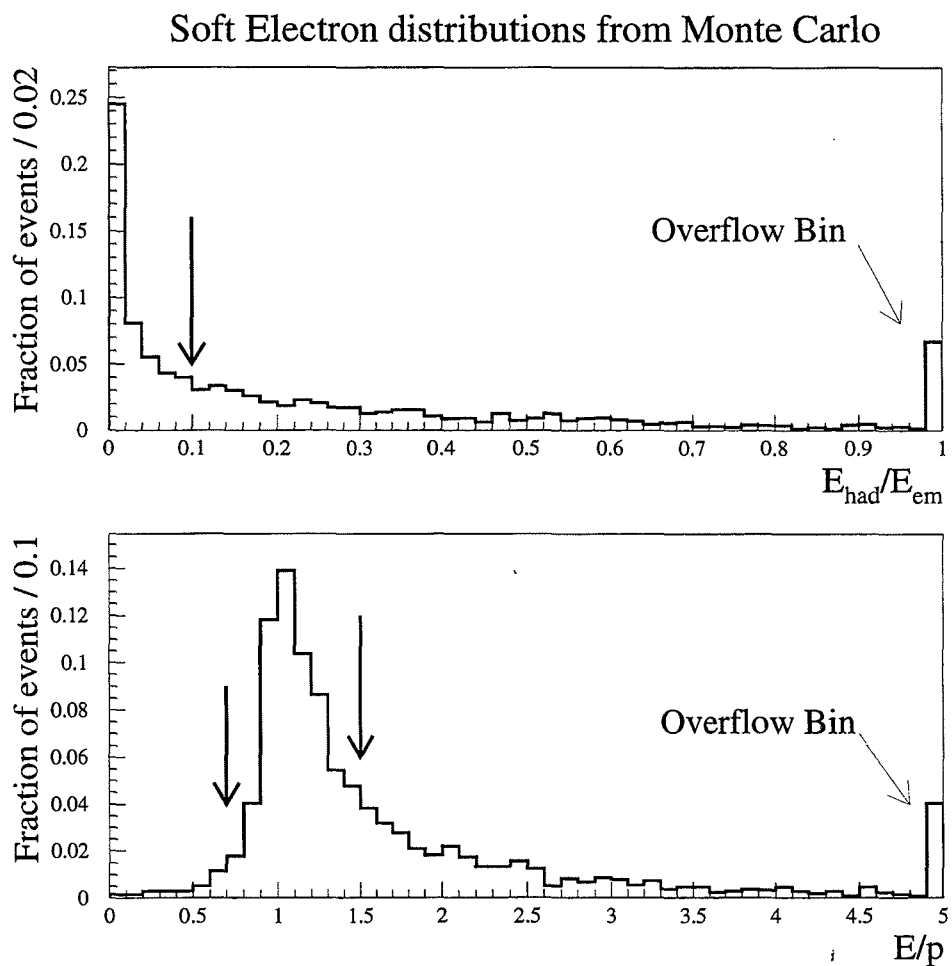


Figure 5-6: E_{had}/E_{em} and E/p distributions from electrons from bottom or charm decay in HERWIG generated $t\bar{t}$ events with $M_{top} = 175\text{GeV}/c$. Roughly 55% are from bottom and 45% are from charm.

Soft Electron Cuts

CPR and CES fiducial cuts

Local CES x position $|X_{CES}| \leq 22cm$

Local CES z position $6.22cm \leq |Z_{CES}| \leq 237.45cm$

Exclude chimney region

Local CPR x position $|X_{CPR}| \leq 17.78cm$

Local CPR z position $9.0cm \leq |Z_{CPR}| \leq 118.0cm$

Local CPR z position $125.0cm \leq |Z_{CPR}| \leq 235.26cm$

CES cuts

$(Es_5/P, Ew_5/P) \geq 0.24 + 0.03P$ if $P < 12GeV/c$

$(Es_5/P, Ew_5/P) \geq 0.6 - (P - 12).0125$ if $12 < P < 20GeV/c$

$(Es_5/P, Ew_5/P) \geq 0.5$ if $P > 20$

$|\Delta x| \leq \text{Max}(0.7 \text{ cm}, 1.82 \text{ cm} - 0.1867 P)$

$|\Delta z| \leq 2 \text{ cm}$

$(\chi^2_{wire}/6, \chi^2_{strip}/6) \leq 16$

Other cuts

$Q_{CPR} \geq 4744 - 11592(P/P_t) + 7923(P/P_t)^2$

$Q_{CTC} \geq 29.15 + e^{1.671 - 0.08P}$ for $P \leq 15 \text{ GeV/c}$

$0.7 \leq E/P \leq 1.5$

$E_{had}/E_{em} < 0.1$

Table 5.1: Soft electron quality cuts. “P” is the track momentum in GeV/c. Local CPR and CES coordinates have z=0 at the $\eta = 0$ boundary and have \hat{x} in the azimuthal direction with x=0 at the center of the calorimeter wedge.

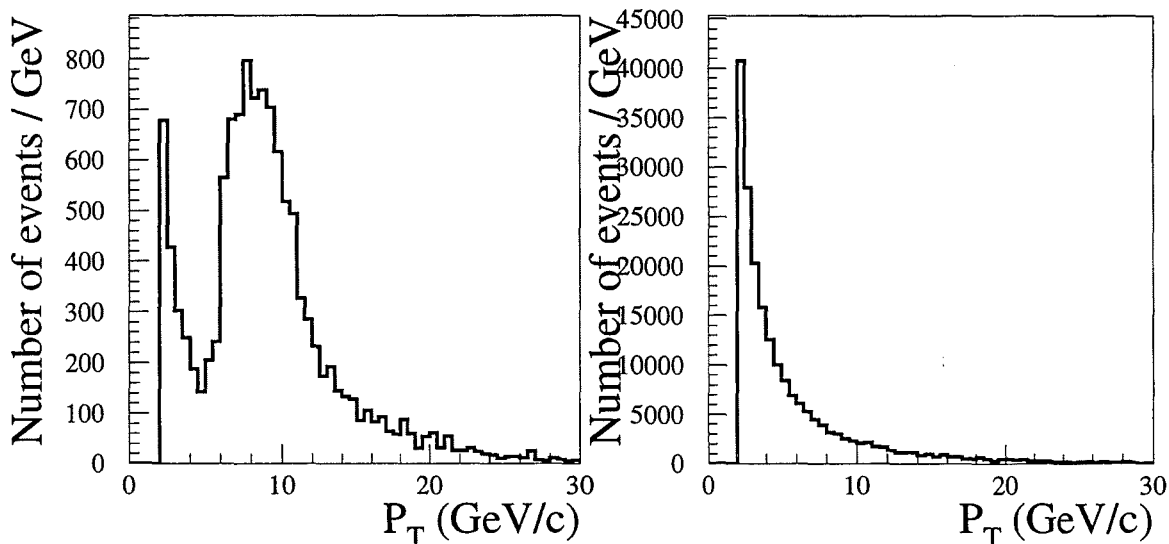


Figure 5-7: P_t of electrons in the conversion sample (left) and candidate tracks in the jet sample (right) with $P_t > 2$ GeV/c. The conversion sample was collected with an $E_t > 8$ GeV electron trigger which produces a corresponding enhancement in the P_t distribution for the sample. In calculating the efficiencies and rejections in Table 5.2 only the tracks with $P_t < 5$ GeV/c are used so that the P_t spectrum of the two samples resemble that of leptons from b's in top events.

as a general guide to the efficiencies and rejections gained by the individual cuts.

The efficiencies shown in the table are for the cut applied alone. Obviously many of these variables are correlated; the Es_5/P and Ew_5/P cuts for example essentially measure the same quantity. The average fake rate after all cuts, (Chapter 8) is around 0.3 % per track as desired. Figure 5-8 shows the efficiency of the isolation independent cuts measured from conversion electrons as a function of P_t .

5.1.8 A Check of the Isolation Dependence of the Algorithm

The assumption that the CES, CPR and dE/dX requirements are not effected by nearby particles can be tested by plotting the efficiency as a function of the

Cut	$\epsilon_{\text{electrons}}$	$\epsilon_{\text{jet tracks}}$
$ \Delta x $	$92.7 \pm 0.6\%$	59 %
$ \Delta z $	$94.8 \pm 0.5\%$	47 %
Q_{CPR}	$84.0 \pm 0.8\%$	53 %
dE/dX	$85.4 \pm 0.8\%$	51 %
Es_5/p	$89.7 \pm 0.7\%$	24 %
EW_5/p	$87.5 \pm 0.7\%$	23 %
E/p	$62 \pm 1\%$	22 %
χ^2_{strip}	$98.9 \pm 0.2\%$	92 %
χ^2_{wire}	$98.2 \pm 0.3\%$	90 %
$E_{\text{had}}/E_{\text{em}}$	$46 \pm 1\%$	23 %

Table 5.2: Efficiency of soft electron cuts for electrons and tracks in generic jets. The electron efficiency has been determined using conversion electrons with $P_t < 5$ GeV/c except for the $E_{\text{had}}/E_{\text{em}}$ and E/p cuts which use b's from HERWIG generated $t\bar{t}$ events ($M_{\text{top}}=175$ GeV/ c^2). Uncertainties for electrons are statistical only. The statistical uncertainties on $\epsilon_{\text{jet tracks}}$ numbers are less than 1%. The efficiencies are for each cut applied individually. Many cuts are correlated.

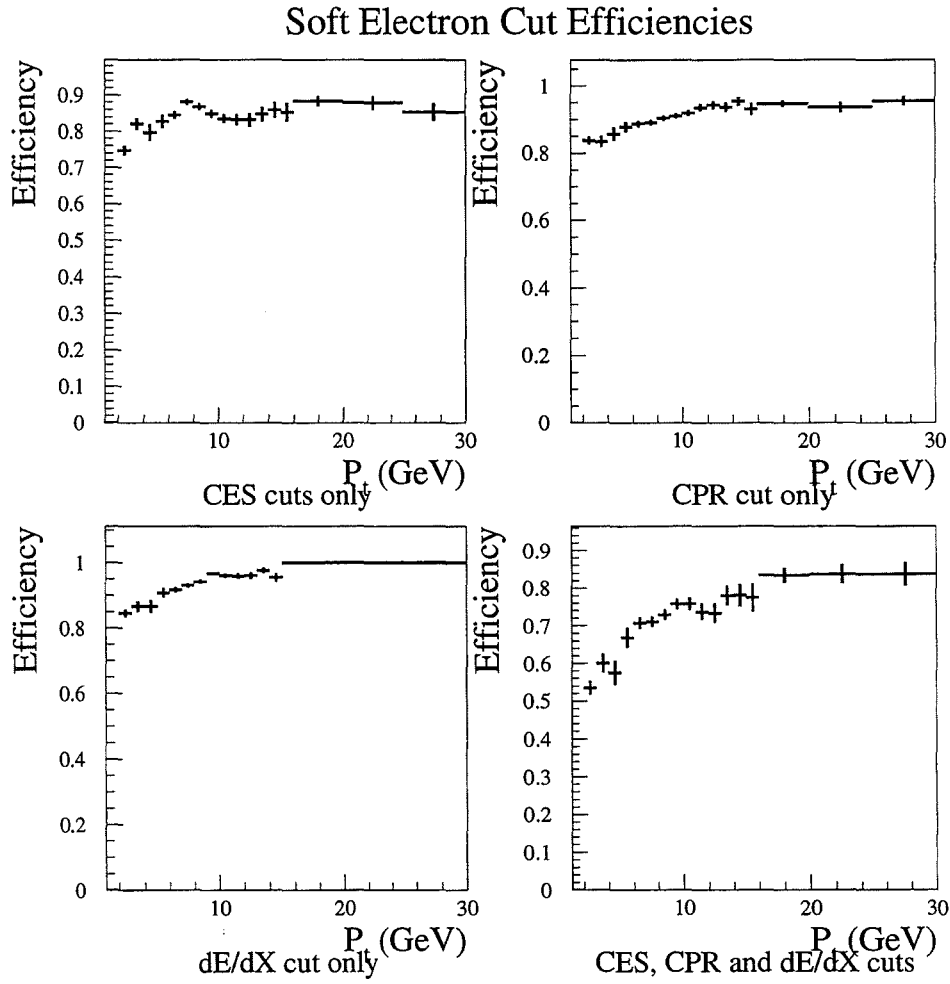


Figure 5-8: The efficiency of all soft electron cuts except E/p and E_{had}/E_{em} as a function of P_t as measured with conversion electrons. Uncertainties are statistical only.

total momentum of nearby tracks. The quantity $\Sigma_p^{0.2}$ is defined as the sum of the momentum of tracks in a cone around the candidate track of radius $R=0.2$ drawn at the calorimeter. Figure 5-10 shows the efficiency of the CES, CPR and dE/dX cuts as a function of $\Sigma_p^{0.2}$ for conversion electrons after correcting for the P_t dependence shown in Figure 5-8. No significant dependence on $\Sigma_p^{0.2}$ is observed, indicating that the efficiencies may be accurately applied to electrons in b jets.

Since the E/p and E_{had}/E_{em} efficiencies are taken from Monte Carlo they also bear further study. Electrons are expected to fail the E/p cut if a nearby photon (from a pi-zero decay for instance) deposits an extra photon in the electron tower. Since photons are relatively simple to simulate, this should be well modeled by the Monte Carlo. The E_{had}/E_{em} quantity however, is likely to be sensitive to the fragmentation model and the simulation of the hadronic calorimeter. Since soft electrons have energies of a few GeV, a variation of a few hundred MeV of hadronic energy can cause the electron to fail the cut. Conversion electrons can be used to test how well the simulation models these effects. Figure 5-9 shows the efficiency of the E_{had}/E_{em} cut as a function of $\Sigma_p^{0.2}$ for electrons from conversions and electrons from b's in HERWIG $t\bar{t}$ Monte Carlo. To avoid a trigger bias (the electron trigger requires $E_{had}/E_{em} < 0.125$) the event is required to pass a trigger unrelated to the electron. The momentum of the conversion partner is removed from the $\Sigma_p^{0.2}$ sum. Within statistical uncertainties, there is good agreement between the Monte Carlo predicted efficiency as a function of $\Sigma_p^{0.2}$ and that observed in the data.

In Appendix C these assumptions are tested together by running the tagger on a sample of high purity b's selected with the SVX.

5.2 Soft Muon Tagging

Soft muon selection begins by applying the same track quality requirements that are imposed on electron candidates; muon candidate tracks must have at

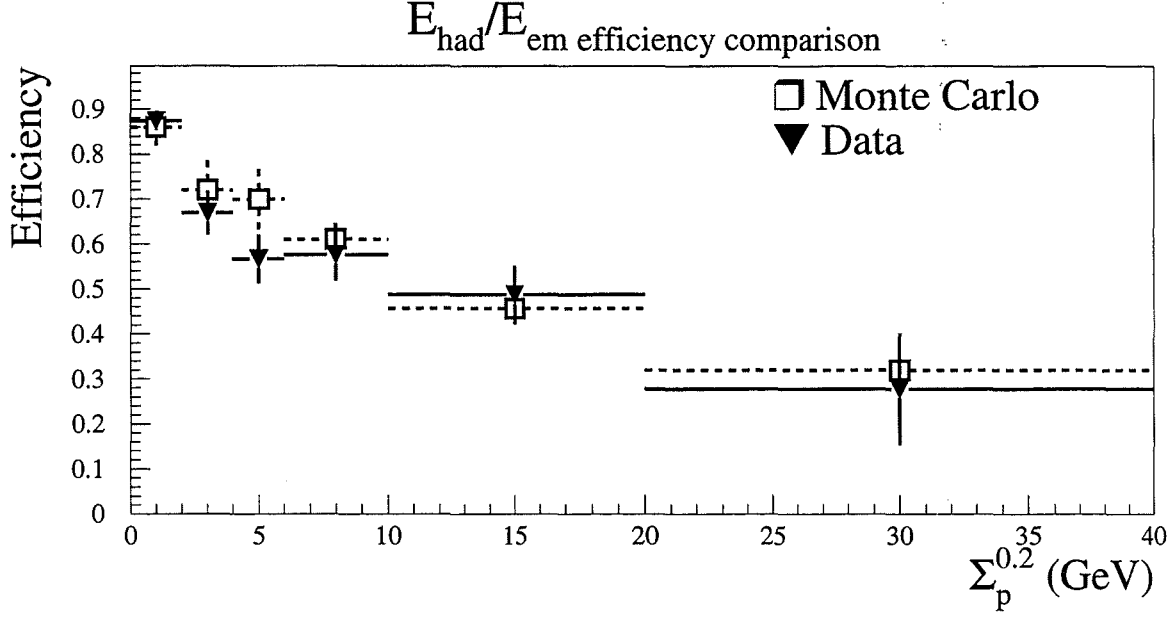


Figure 5-9: Efficiency of the $E_{had}/E_{em} < 0.1$ cut as a function of $\Sigma_p^{0.2}$ for conversion electrons (triangles) and electrons from b or c decay in Herwig $t\bar{t}$ Monte Carlo with $M_{top}=175$ GeV/c².

least two axial and at least two stereo superlayers used in their reconstruction, and an impact parameter less than 3mm. This guarantees a well measured 3D track and cosmic rays and rejects some muons from pions or kaons which decay in flight. As for the soft electrons, tracks must have $P_t > 2$ GeV/c. The amount of material in the detector essentially places this lower limit independently as muons with $P_t < 1.4$ GeV/c will range out before reaching the CMU and must typically have $P_t > 2$ GeV/c to reach the CMP [15]. Figure 5-11 shows the absorption lengths traversed by a particle as a function of $|\eta|$.

5.2.1 Muon Fiducial Requirements

Tracks passing the quality cuts are extrapolated to the muon chambers. For the CMX the simulation must take into account the return field of the solenoid which is non-uniform at large $|\eta|$. The extrapolation routine also calculates the distance to the closest edge of each detector in the x and z directions¹ and the

¹As before, x refers to the tranverse ($r - \phi$) plane.

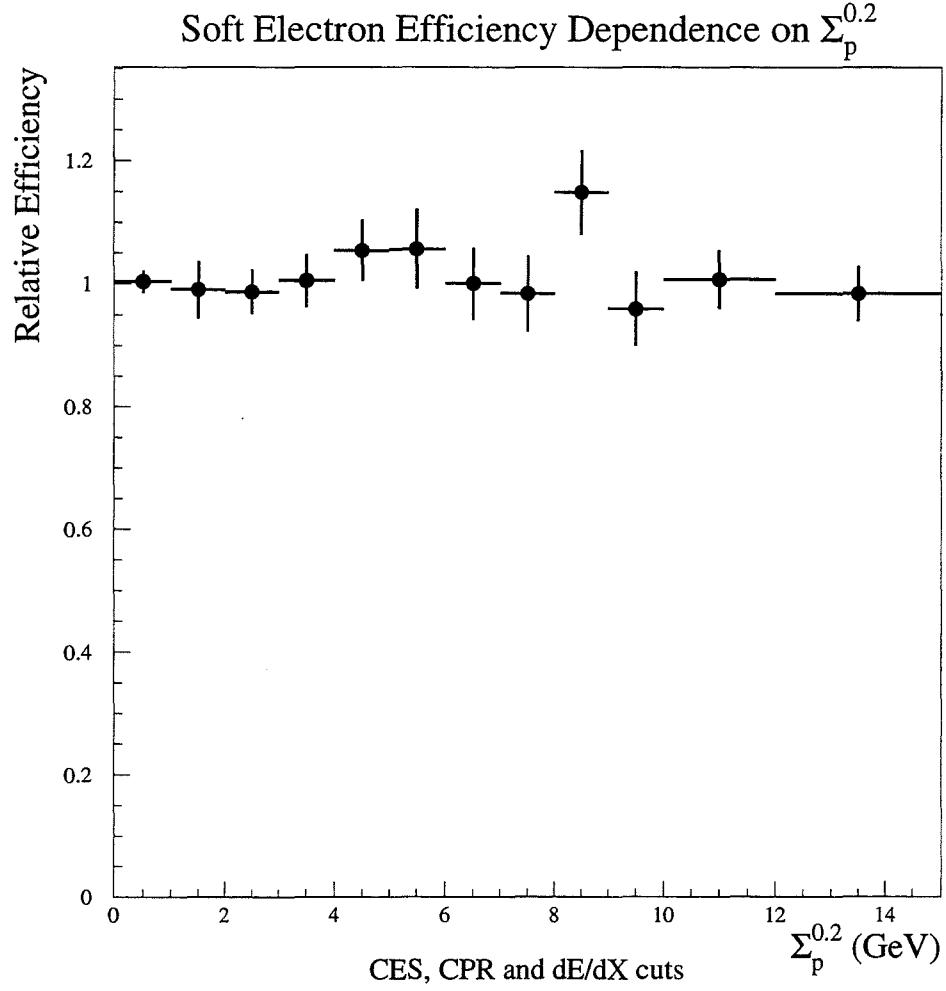


Figure 5-10: Efficiency of all soft electron cuts except E/p and E_{had}/E_{em} as a function of P_t in a cone of $R=0.2$ around the electron track. The P_t dependence shown in Figure 5-8 has been removed. Within statistical uncertainties the efficiency is flat with $\Sigma_p^{0.2}$.

expected multiple scattering distance σ_{MS} . The tracks are then further divided into four fiducial types. The CMU and CMP region is subdivided based on which stubs are expected for the track.

- **Type 1 (CMUP):** The track must have $P_t > 3 \text{ GeV}/c$, and extrapolate to within the CMU and be more than $3\sigma_{MS}$ or 2 cm from the nearest CMU boundary. The track must also extrapolate to within the CMP and be more than $3\sigma_{MS}$ or 5 cm from the nearest CMP boundary. The momentum requirement insures that the muon should reach the CMP, the other cuts insure that both CMU and CMP stubs should be present.

- **Type 2 (CMP-only):** The track must have $P_t > 3 \text{ GeV}/c$ and extrapolate to within the CMP and be more than $3\sigma_{MS}$ or 5 cm from the nearest boundary. The track must not qualify as Type 1 (ie. the CMU is not required).

- **Type 3 (CMU-only):** The track must extrapolate to within the CMU boundaries, and not be of Type 1 or 2.

- **Type 4 (CMX):** The track must extrapolate to within the boundaries of the CMX detector.

In the next step, the tracks are matched to stubs and matching requirements applied. There is some subtleness about which combinations of CMU and CMP stubs should be required for each of the above types; where possible a CMP stub should be required to backup a CMU stub since this greatly reduces backgrounds from punch through, however, this should not be done at the expense of a loss in acceptance. As described in Section 4.1.2, each detector has gaps in coverage which make this complicated:

- For tracks of **Type 1** both a CMU and a CMP stub are required, and the muon must pass the cuts described below for a CMUP muon.

- For a track of **Type 2** a CMP stub is required. If, however, there is a CMU stub associated with the track, and the track is not more than $3\sigma_{MS}$ or 2 cm outside the CMU, then it qualifies as a CMUP muon for the cuts listed below. If there is only the CMP stub, then it must instead pass the (more stringent) CMP-only cuts in the next Section.

- For a track of **Type 3** a CMU stub is looked for. If the CMU stub exists, it is used to extrapolate the track to the CMP. If it extrapolates to within the CMP and is at least $3\sigma_{MS}$ from the nearest boundary a CMP stub is also required. If it has the CMP stub also, then it is treated as a CMUP type for the cuts. If no CMP stub is required and there is a CMU stub, it is treated as a CMU-only type. If there is no CMU stub, but has $P_t > 3 \text{ GeV}/c$ and there is a CMP stub, and the track is no more than $3\sigma_{MS}$ outside the CMP, then it is categorized as a CMP-only type for the cuts in the next Section.

- In the rare case that a track is not of Type 1, 2 or 3, but has a CMU or a CMP stub associated with it, it follows the same categorization logic as for a Type 3 track described above.

- **Type 4** muons are unambiguous. They must have a CMX stub, and pass the CMX cuts described in the next Section.

Note that these requirements carry small inefficiencies. For instance, CMP-only muons with P_t below $3 \text{ GeV}/c$ are not accepted. Requiring a minimum of 3 GeV avoids the CMUP turn-on region which occurs at around $P_t=2.7 \text{ GeV}/c$, which may not be well modeled by the Monte Carlo. Other inefficiencies arise because the muon reconstruction algorithm will occasionally match a CMU or CMP stub to a track other than the muon.

5.2.2 Muon Selection Criteria

Most of the muon quality cuts relate to the quality of the track-stub match of the muon candidate which are described here. The cuts are summarized in Table 5.3.

Description of Variables

A good CMU muon typically leaves hits in all four layers of the CMU which can be reconstructed to measure the stub position in the x direction and angle in azimuth (ϕ). Residuals in this direction are on the order of $250 \mu m$ [15].

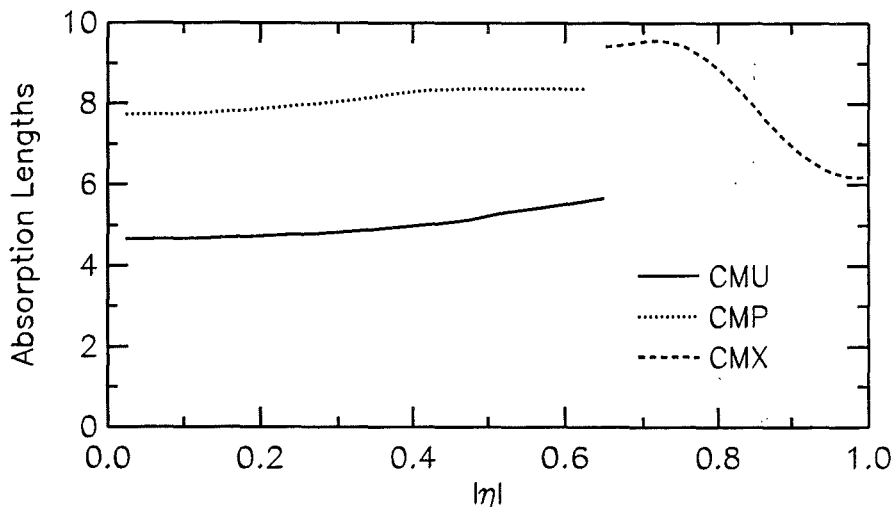


Figure 5-11: Number of absorption lengths (averaged in azimuth) as a function of pseudorapidity calculated with $z = 0$.

The z position is measured using charge division with an accuracy of 1.2 mm [79]. The θ angle of the stub is too poorly measured to be of use. Stubs in the CMP are reconstructed only in two dimensions, and so yield only x and ϕ information.

A typical muon path through the CMX intercepts between four and six drift tubes. Because of the conical geometry, there is a stereo angle (3.6 mrad) between adjacent chambers. Stubs must consequently be fit in three dimensions and so yield position information in the z , x and ϕ directions. The z position of the stub is measured with a resolution of roughly 2 cm. The θ angle of the stub cannot be well measured, and in fact, the stub fitter constrains the stub to point to $z = 0$ in the $r - z$ plane.

When the muon reconstruction routine extrapolates a CTC track to the muon chambers, it calculates the expected mismatch (σ) with the stub by adding the expected multiple scattering in quadrature to the detector resolu-

Expected track/stub mismatch

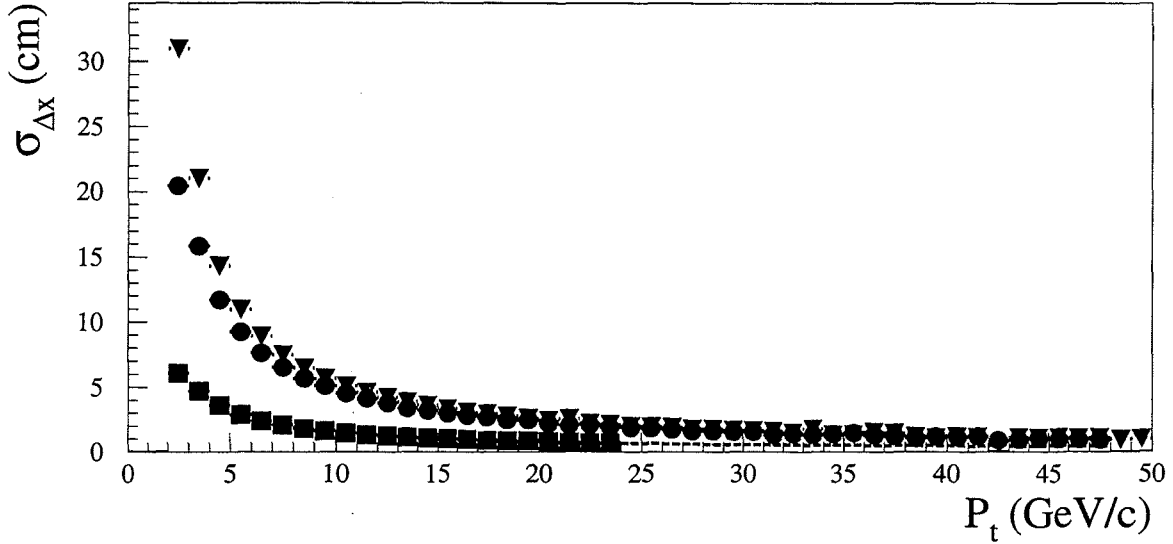


Figure 5-12: Expected one sigma track-stub mismatch ($\sigma_{\Delta x}$) in the x direction for muons in the CMU (square), CMX (triangle) and CMP (dots) as a function of P_t . In each case, $\sigma_{\Delta x}$ is measured at the front of the muon chamber. The CMP has the most shielding but muons in the CMX travel further after scattering before arriving at the muon chambers which gives them a larger $\sigma_{\Delta x}$.

tion. Because multiple scattering depends on the amount of material encountered, and because the angular deflection from multiple scattering is proportional to $1/P$ [26] [70], the expected mismatch falls roughly as $1/P_t$ as shown in Figure 5-12, and is larger for the CMP than the CMU which has less shielding. The CMX mismatch is the largest in part because of the long distance the particles travel after multiple scattering in the calorimeter before reaching the muon chambers.

From the expected mismatch, chi-square quantities are constructed for each matching variable. In the following, the Δx variable refers to the distance between the extrapolated track and the stub in the transverse plane evaluated at the inner radius of the muon detector. The Δz quantity refers to the mismatch in the z direction. The $\chi^2_{\Delta\phi\Delta x}$ chi-square takes into account the expected correlation between Δx and $\Delta\phi$. Figure 5-13 shows the muon matching quantities divided by the expected sigma for muons from J/ψ 's. In each case a background subtraction has been performed using the sidebands; each variable is

Muon matching variables

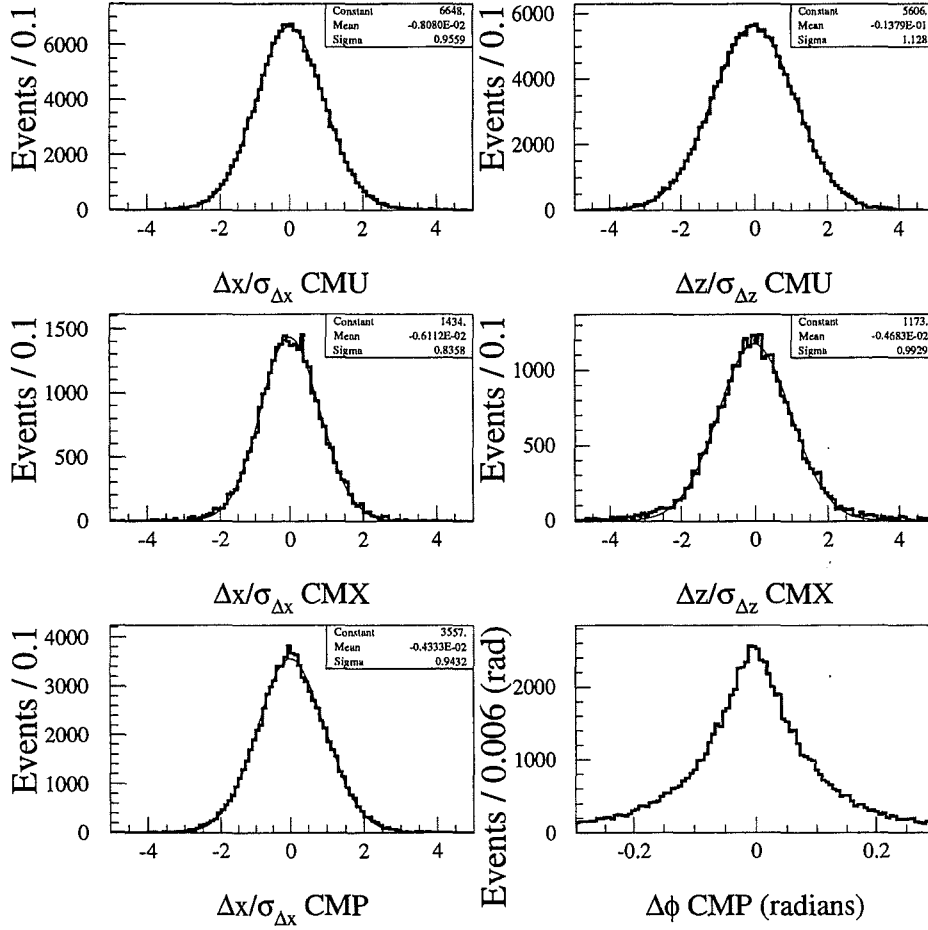


Figure 5-13: Muon matching variables from J/ψ 's. All plots have been background subtracted using the mass sidebands as described in the text. Also shown are gaussian fits, their means and sigmas. The $\Delta \phi_{CMP}$ distribution is not divided by the expected sigma since this is the quantity of interest in selecting CMP-only muons. (See Table 5.3.)

plotted for events with mass $\pm 3\sigma$ from the J/ψ mean, and the distribution for an equivalent mass width from the sidebands is subtracted. (The J/ψ mass distribution is shown in Figure 4-5.) The matching variables are each well described by a gaussian centered at zero with near unit width indicating that the expected mismatch is well modeled.

Soft Muon Quality Requirements

Table 5.3 lists the soft muon quality cuts for each muon category described above. For CMUP types, matching cuts are only applied to the CMU stub, since the requirement of a CMP stub has already increased the purity. A

minimum of 3 TDC and ADC hits are required for the CMU stub which insures that it is well measured in x and z . Note that the tightest matching cut applied is 8 cm in z and 2 cm in x . This cuts are not tightened beyond this because there are tails to the matching distributions not accounted for by σ_{MS} . The tails are apparent only at high P_t when the multiple scattering contribution is small. As shown in Figure 5-12, a cut at 2 cm corresponds to a 3σ cut only at very high P_t . Because most of the candidate tracks in jets have low P_t this does not significantly increase the background.

For CMU-only and CMP-only muons, more stringent cuts are necessary to keep the background at a level comparable to the CMUP category. CMP-only type muons can arise when a muon passes near the phi edge of a calorimeter wedge where there is a gap in the CMU coverage. Since pions also pass through these cracks, the background from hadronic punch-through is larger here than for CMUP muons.

For CMU-only types, a quantity $CMCLUS$ is used which is the number muon tubes adjacent to and including the stub which contain at least one TDC hit. A typical CMU muon will have $CMCLUS = 4$. A hadron (from a hadron shower), however, which punches through the calorimeter will often be accompanied by a number of other particles which deposit hits in the CMU. Requiring $CMCLUS < 6$ decreases the background by $\sim 8\%$ with essentially no loss in efficiency.

For all non-CMX muons a loose minimum ionizing cut is also applied to candidates with $P_t > 6\text{GeV}/c$. Below this point, muons and hadrons deposit similar energy in the calorimeter. The hadronic energy associated with the candidate is determined by extrapolating the candidate track to the middle in depth of the CHA ($R=279$ cm) and using the energy in the corresponding tower. Since muons from b 's are typically in jets, the total momentum all tracks in a cone of radius $R=0.2$ ($\Sigma_p^{0.2}$) is subtracted from the calorimeter energy to account for the effect of nearby particles. The minimum ionizing cut is then,

- **CMUP category**

Number of CMU TDC hits ≥ 3

Number of CMU ADC hits ≥ 3

$|\Delta z|_{CMU} < MAX(3\sigma, 8cm)$

$|\Delta x|_{CMU} < MAX(3\sigma, 2cm)$

$E_{had} - \Sigma_p^{0.2} < 6 \text{ GeV}$ if $P_t > 6 \text{ GeV}/c$

- **CMU-only category**

Number of CMU TDC hits ≥ 3

Number of CMU ADC hits ≥ 3

$|\Delta z|_{CMU} < MAX(3\sigma, 8cm)$

CMCLUS ≤ 5

$\chi^2_{\Delta\phi\Delta x CMU} < 10$ if $P_t < 20 \text{ GeV}/c$

$|\Delta x|_{CMU} < MAX(3\sigma, 2cm)$ if $P_t \geq 20 \text{ GeV}/c$

$E_{had} - \Sigma_p^{0.2} < 6 \text{ GeV}$ if $P_t > 6 \text{ GeV}/c$

- **CMP-only category**

$\chi^2_{\Delta\phi\Delta x CMP} < 10$ if $P_t < 10 \text{ GeV}/c$

$|\Delta x|_{CMP} < MAX(3\sigma, 5cm)$ if $P_t \geq 10 \text{ GeV}/c$

$|\Delta\phi|_{CMP} < 0.1$ if $P_t \geq 10 \text{ GeV}/c$

$E_{had} - \Sigma_p^{0.2} < 6 \text{ GeV}$ if $P_t > 6 \text{ GeV}/c$

- **CMX category**

$\chi^2_{\Delta z CMX} < 9$

$\chi^2_{\Delta x CMX} < 9$

$\chi^2_{\Delta\phi\Delta x CMX} < 9$ if $P_t \leq 5 \text{ GeV}/c$

$|\Delta\phi|_{CMX} < 9$ if $P_t > 5 \text{ GeV}/c$

Table 5.3: Selection requirements for soft muon candidates.

$$E_{had} - \Sigma_p^{0.2} < 6 \text{ GeV}/c.$$

Monte Carlo studies show that this requirement is $98 \pm 2\%$ efficient for muons from b decays with $P_t > 6 \text{ GeV}/c$. The minimum ionizing cut reduces the muon fake rate for tracks from generic jets with $P_t > 6 \text{ GeV}/c$ by $\sim 6\%$ of itself.

For both the CMU-only and CMP-only categories, the $\chi^2_{\Delta\phi\Delta x}$ cut is slightly inefficient at high P_t and so is replaced with a straight Δx matching cut. For the same reason, the $\chi^2_{\Delta\phi\Delta x}$ requirement for CMX muons is replaced with a straight cut on $|\phi|$ for tracks with $P_t > 5 \text{ GeV}/c$.

5.2.3 Rejection Power of Soft Muon Requirements

To get an idea of the background rejection gained by each of the soft muon requirements, they are applied to tracks in generic jet events. As summarized in Table 5.4, roughly 150k tracks are found pointing to the muon chambers, of which $\sim 35\text{k}$ have associated muon stubs. The cuts described above provide an overall rejection factor of 24, leaving 1466 SLT muons for an overall fake rate of 0.98 % per track. There are fewer tracks in the **type 1** and **type 2** categories because they are required to have $P_t > 3 \text{ GeV}/c$.

Efficiency of Soft Muon Requirements

The efficiency of the muon cuts measured with muons from J/ψ decay is shown as a function of P_t in Figure 5-14. To measure the efficiency at high P_t a sample of Z's is used for the $P_t > 10 \text{ GeV}/c$ bins. To test the assumption of isolation independence, the efficiency is plotted as a function of $\Sigma_p^{0.2}$ after correcting for the momentum dependence in Figure 5-15. For CMUP and CMX muons there is no observable dependence on $\Sigma_p^{0.2}$. For CMU-only muons, however, the efficiency drops by $\sim 10\%$ for values of $\Sigma_p^{0.2} > 3 \text{ GeV}/c$.² As will be discussed

²One can imagine that, where there is nearby punch-through from other particles, the stub may fail the CMCLUS cut or be poorly reconstructed; this has been found to be the

Soft muon cut efficiencies

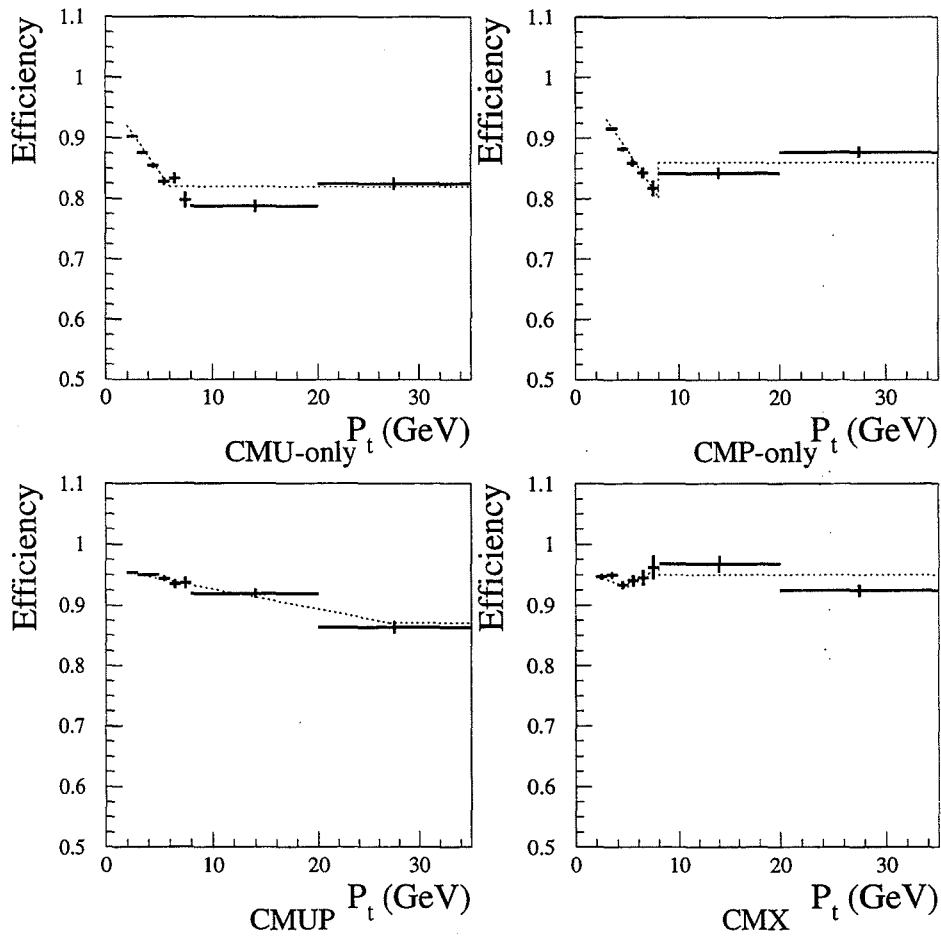


Figure 5-14: Efficiency of soft muon matching cuts as a function of P_t . Efficiencies are obtained from J/ψ 's and Z 's. For the J/ψ points, a background subtraction has been performed using the mass sidebands as described in the text.

# Track type 1 (cmup region)	38097
# Track type 2 (cmp-only region)	28346
# Track type 3 (cmu-only region)	82406
# Track type 4 (cmx region)	55440
# with CMU and CMP stubs	2600
# with only a CMP stub	1913
# with only a CMU stub	26709
# with a CMX stub	4004
# CMUP passing cuts	323
# CMP-only passing cuts	163
# CMU-only passing cuts	574
# CMX passing cuts	406

Table 5.4: The rejection of muon candidate tracks in generic jet events at various stages of the soft muon selection. Initially there are 150k tracks which point to the muon chambers. Requiring a muon stub reduces the number of candidates to 35k, most of which are CMU-only types originating from hadronic punch-through. The final selection criteria which require that the candidate track be well matched to a muon stub provide a rejection factor of roughly 24, leaving 1466 soft muons. The average fake rate for this sample is then 0.98 %.

in the following chapter, only 14 % of the SLT acceptance is expected to come from CMU-only muons, of which roughly 2/3 should have $\Sigma_p^{0.2} > 3 \text{ GeV}/c$. The final effect on the acceptance is less than 1%. For completeness, the dependence of the CMU-only efficiency on $\Sigma_p^{0.2}$ is implemented in the Monte Carlo. Although there appears to be a possible $\Sigma_p^{0.2}$ dependence in the CMP-only muons, this is an even smaller correction and so is ignored.

In summary, an algorithm for identifying muons in top events with $P_t > 2 \text{ GeV}/c$ has been developed. The fake rate, as described in Section 8.1.3 is less than one percent per track. The id cuts are over 90% efficient for real muons.
 case with an early version of the CMX stub reconstruction algorithm which would incorrectly measure the stub ϕ or z because of surrounding drift tube hits [18].

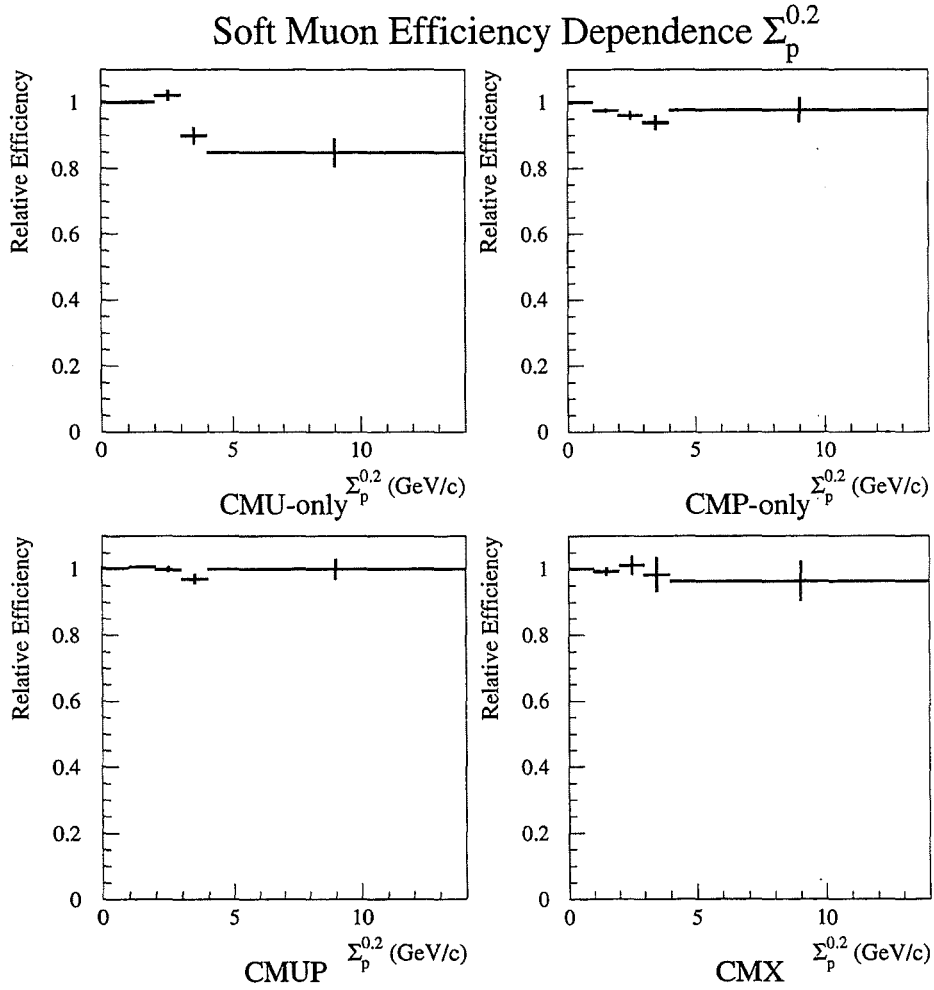


Figure 5-15: Efficiency of muon id cuts as a function of P_t in a cone of $R=0.2$ around the muon track. The P_t dependence shown in Figure 5-14 has been taken out. The drop in efficiency for CMU-only muons is implemented in the Monte Carlo, the apparent drop in CMP is negligible. The CMUP and CMX efficiencies are flat with $\Sigma_p^{0.2}$ within statistical uncertainties.

5.3 Monte Carlo Simulation of the Soft Lepton Tagger

A nice feature of the soft lepton tagger as developed in this chapter is that it has been largely developed from studying real soft leptons from the data. To avoid relying on the Monte Carlo's ability to model the response of the detector for the SLT (dE/dX for instance is not modeled at all in QFL) only the E_{had}/E_{em} and E/p quantities (which are expected to depend on the particular event topology) are taken from the Monte Carlo. For all other requirements the efficiencies measured in the data are used. The procedure is summarized as follows:

- Tracks produced in the simulation are first matched to electrons and muons produced at the generator level. The electron or muon is required to be a “real” tag, i.e. one from the decay of a b , c , W or τ .
- Electrons are extrapolated to the CPR and CES in the same manner as for the data, and required to pass the fiducial cuts and conversion removal cuts.
- Muons which are reconstructed in the muon chambers are extrapolated as in the data and classified as CMU-only, CMP-only, CMUP or CMX as described in Section 5.2.
- Finally, the efficiencies as a function of P_t determined from the data (see Figures 5-8 and 5-14) are applied to the tracks. For electrons, the E_{had}/E_{em} and E/P quantities are obtained from the Monte Carlo and required to pass the selection criteria.

Stub finding efficiency

For the muons a small correction factor is also applied in the Monte Carlo to account for the fact that a muon may pass through the drift tubes, but leave too few hits to be reconstructed into a stub. Because each drift tube has high efficiency ($99.93 \pm 0.03\%$ for CMX as measured with cosmic rays [19]) this should be a small effect. For the CMU and CMP chambers the stub finding

efficiency has been measured using muons from $Z \rightarrow \mu^+ \mu^-$ to be $98.3 \pm 0.3\%$ and $99.76 \pm 0.03\%$ respectively [20]. For the CMX the efficiency has been measured with muons from $Z \rightarrow \mu^+ \mu^-$ to be $98.9^{+0.6}_{-1.0}\%$.

Tracking efficiency

The tag efficiency is also corrected for a tracking inefficiency present in data, but which is not modeled in QFL. While isolated tracks in data are reconstructed almost 100 % of the time, it has been shown that tracks in jets can be missed if enough hits are merged with nearby hits, or are incorrectly assigned to a different track. The inefficiency has been parameterized as a function of the density of hits surrounding a track by embedding an isolated track from one event in a jet from another and retracking the event [38]. This correction effectively degrades ϵ_{SLT} by a factor of 0.96. Preliminary studies of Run 1b data indicate that an additional correction of ~ 0.95 may be necessary, presumably due to ageing of the CTC [39]. This is treated as a systematic uncertainty in Section 6.4.

Chapter 6

Acceptance for $t\bar{t}$ events

This chapter describes the calculation of the acceptance of this analysis for $t\bar{t}$ events. The acceptance is calculated using HERWIG generated $t\bar{t}$ events and defined as the expected number of “real” soft lepton tags observed divided by the number of $t\bar{t}$ events produced. The soft leptons originate mainly from the semileptonic decay of b or c hadrons (90%) and occasionally from the decay of a W or tau (10%). Note that this is the “tag acceptance” in the sense that it uses the number of *tags* and not the number of tagged *events*. Counting tags instead of tagged events reduces the statistical uncertainty¹ and greatly simplifies the calculation of $\sigma_{t\bar{t}}$. The acceptance can be factored as

$$A_{t\bar{t}} = A_{lepton} \cdot \epsilon_{lepton\ id} \cdot \epsilon_{\cancel{E}_t} \cdot \epsilon_{jet} \cdot \epsilon_{trig} \cdot \epsilon_{SLT\ tag}$$

where

- $A_{lepton} \cdot \epsilon_{lepton\ id}$ is the fraction of $t\bar{t}$ events which have an electron with $E_t > 20$ GeV or a muon with $P_t > 20$ GeV/c which passes the id cuts described in Section 4.1.
- $\epsilon_{\cancel{E}_t}$ is the fraction of these good lepton events which also have $\cancel{E}_t > 20$ GeV.

¹Roughly 10 % of tagged $t\bar{t}$ events ($M_{top}=175$ GeV/c²) are expected to contain a second soft lepton tag.

$A_{lepton} \cdot \epsilon_{lepton\ id}$	$14.9 \pm 0.1\%$
$\epsilon_{\cancel{E}_t}$	$88.9 \pm 0.3\%$
ϵ_{jet}	$84.5 \pm 0.3\%$
ϵ_{trig}	$92.4 \pm 0.2\%$
$\epsilon_{SLT\ tag}$	$20.3 \pm 0.4\%$
Final acceptance	$2.10 \pm 0.04\%$

Table 6.1: Components of the acceptance for $t\bar{t}$ events measured from HERWIG generated events with $M_{top} = 175\text{ GeV}/c^2$. The uncertainties shown are from Monte Carlo statistics only. Each efficiency is measured after the requirements listed above it have been applied.

- ϵ_{jet} is the fraction of events with a good lepton and $\cancel{E}_t > 20\text{ GeV}$ which also have at least three jets.
- ϵ_{trig} is the fraction of events passing all above cuts that are expected to pass the trigger.
- $\epsilon_{SLT\ tag}$ is the fraction of events passing the above cuts which also contain a real soft lepton tag.

Table 6.1 lists the various components of the acceptance as measured using $t\bar{t}$ events ($M_{top} = 175\text{ GeV}/c^2$). After applying the $W + \geq 3\text{ jet}$ selection, the primary leptons are expected to be half muons, and half electrons. In $t\bar{t}$ events, roughly 75 % real soft lepton tags are expected to be muons, and 25 % electrons. Of the muon tags, 44 % are CMUP type, 16 % are CMP-only, 19 % are CMU-only, 21 % are CMX.

6.1 Geometric Acceptance and Lepton Identification

$A_{lepton} \cdot \epsilon_{lepton\ id}$ is the fraction of $t\bar{t}$ events which result in a lepton that passes the primary lepton id cuts described in Section 4.1. The lepton is not required to come from the decay of a W boson. Because the QFL simulation does not exactly reproduce all of the variables used in the primary lepton selection, the number of events passing cuts in the Monte Carlo is corrected by a small factor. The factor is determined by comparing the efficiency of the id cuts measured using Z 's in the data to that obtained from a Monte Carlo sample of Z events. The electron (muon) id cuts are found to be a factor of 1.08 (1.06) more efficient in Monte Carlo than in data, and so the Monte Carlo $t\bar{t}$ efficiency is degraded accordingly. The electron efficiency is also degraded by 0.98, the measured overefficiency of the conversion removal cuts.

As will be discussed in 7.1 several event cuts will be applied to reduce backgrounds from non- W processes. These apply only to events with a soft lepton tag and remove events where the primary lepton and the SLT are consistent with coming from an Υ , J/ψ , or the sequential decay of a b . Events where the pair appear to be from Drell-Yan production or an unremoved Z boson decay are also removed. These cuts are discussed in detail later and are only mentioned here to point out that they are also applied to the Monte Carlo events in calculating the acceptance. Events used in the dilepton analysis are also removed (Section 4.1.3). In the rare case where an event contains two leptons which pass the primary lepton cuts, the lepton with the highest P_t is taken as the primary.

For a top mass of $M_{top}=175\text{ GeV}/c^2$, $A_{lepton} \cdot \epsilon_{lepton\ id}$ is determined to be $14.9 \pm 0.1\%$.

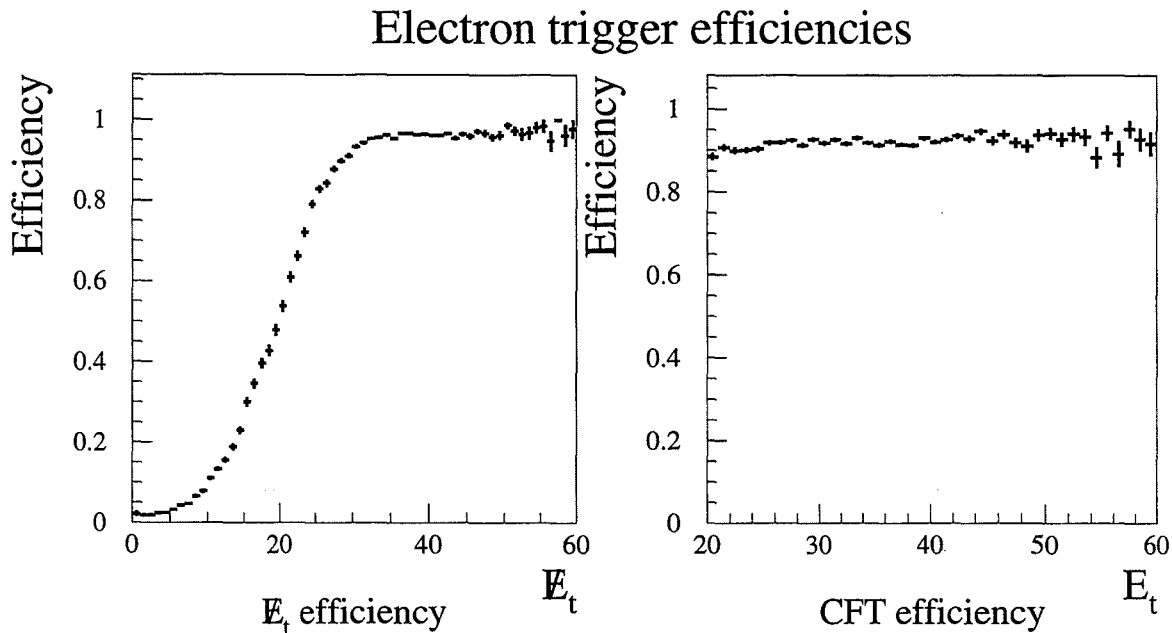


Figure 6-1: Efficiency of the E_t and CFT trigger requirements. The logical OR of the two triggers is expected to be $> 99\%$ for electrons in the W sample.

6.2 Calculation of the Trigger Efficiency (ϵ_{trig})

As discussed in Section 4.1 the W sample is selected using a variety of triggers. Electron events must come in on one of two triggers. The first requires a Level 2 CFT track corresponding to a P_t of at least 12 GeV/c, and an electromagnetic cluster with $E_t > 16$. The other trigger requires $E_t > 20$ GeV in place of the CFT requirement. Figure 6-1 shows the efficiency of the CFT and the E_t requirements obtained by plotting the efficiency of one trigger for events which have passed the other. Assuming the Level 2 CEM cluster requirement is 100 % efficient, the combined efficiency can be measured by taking a sample of Monte Carlo top events and convoluting their E_t and P_t spectrums with these efficiency curves. It is found that the OR of these two triggers should be over 99% efficient for top events with a primary electron, and is taken to be 100 %.

There are 16 different muon triggers allowed for this analysis. (See Section 4.1.2.) Prohibitively large trigger rates necessitated that the inclusive CMU-only and CMX triggers be “prescaled” so that only a fixed fraction of the events

were accepted. To maintain high efficiency for the top analysis, additional (non-prescaled) triggers are included which also require a jet. There is no Level 2 path for CMP-only muons, and so they are required to pass a \cancel{E}_t trigger. The efficiency of these triggers as a function of P_t , jet energy and \cancel{E}_t taking into account prescale correction factors has been measured by others [82]. A software trigger simulation which implements this has been incorporated into the Monte Carlo simulation [21]. The efficiency of the trigger is $84 \pm 0.9 \%$ for events with a primary muon passing the lepton, \cancel{E}_t and jet requirements. Averaged over primary leptons and muons the trigger efficiency is $\epsilon_{trig}=92.4 \pm 0.2\%$. The trigger simulation is checked in Section B by calculating the W cross section.

6.3 Calculation of the Soft Lepton Tagging Efficiency (ϵ_{SLT})

The soft lepton tagging efficiency is defined for $t\bar{t}$ events as the number of real soft lepton *tags* in events which pass the $W+\geq 3$ jet selection divided by the number of events which pass the $W+\geq 3$ jet selection. (This is 10 % larger than the *event tagging rate* defined as the fraction of events in the signal region which are tagged.) The efficiency of the majority of SLT cuts are determined from conversion electrons and muons from J/ψ 's and Z boson decay. To a good approximation, these efficiencies are isolation independent and can be applied to leptons in b jets in Monte Carlo events. The SLT algorithm tags only those electrons or muons that come from the weak decay of a b or c quark, a W or a τ . These are the “real SLT” tags and will represent the source of excess tags. While it is true, for instance, that in a b decay, a genuine muon may arise from a pion which decays in flight, this is treated as background, since it occurs also in non heavy flavor jets. Particles in b jets may also be misidentified as an electron or muon, but for similar reasons this is also considered a background

process, and does not contribute to ϵ_{SLT} .

For a top mass of $M_{top}=175 \text{ GeV}/c^2$, ϵ_{SLT} is determined to be $20.3 \pm 0.4\%$.

6.4 Statistical and Systematic Uncertainties

Many sources of systematic uncertainty are particular to a single component of the acceptance as factored above. In other cases correlations (increased gluon radiation for instance may lower the lepton id efficiency but raise the efficiency of the three jet requirement) require that the uncertainty be calculated for the acceptance as a whole. The following is a list of the major sources of uncertainty for the acceptance.

- **Lepton id statistics :** The uncertainty due to the limited number of Z events used to measure the lepton id efficiency is less than 1%.
- **Monte Carlo statistics :** Uncertainties from Monte Carlo statistics are small, in all cases less than 2 %.
- **Trigger Efficiency:** The systematic uncertainty on the trigger simulation is determined by varying the Level 1 and Level 2 efficiencies within the measured errors. The effect on the acceptance is less than 5 % which is taken as the systematic uncertainty [25]. The trigger simulation is tested by calculating the W production cross section in Section B which is found to be in good agreement with previous measurements.
- **Jet Energy Scale :** The uncertainty on the overall jet energy scale is estimated using E_t balancing in Z+jet and photon plus + jet events to be less than 3%. Scaling the jet energies in the Monte Carlo by $+(-)10\%$ changes the acceptance by $+4 (-3) \%$ for a top mass of 175 GeV. To be conservative, a systematic uncertainty of 5% is assigned based on these studies.

- **Structure functions** : By default, the Monte Carlo generated samples use MRSD0' [96] structure functions, since these give a good fit to the observed W asymmetry at CDF [98]. Recalculating the acceptance with another recent parameterization, CTEQ2L [97] has no observable effect.
- **Initial and Final State Radiation** : Although measurements of jet multiplicities and other jet quantities at CDF compare quite well with HERWIG predictions [83] , uncertainties associated with gluon radiation can have large effects on the acceptance. The sensitivity of the acceptance to initial and final state radiation is estimated by turning them off in the Monte Carlo program. A systematic uncertainty of 5% and 2% is assigned to the modeling of initial and final state radiation, which corresponds to half of the observed variation in the acceptance when they are turned on and off [78].
- **Model Dependence** : To estimate the uncertainty due to jet fragmentation and other effects, the acceptance is also calculated using a PYTHIA generated sample and found to agree with HERWIG predictions at the level of 10 % which is taken as a systematic.
- **Soft Lepton Tagging Efficiency** : The statistical uncertainties on the isolation independent cuts, (Figures 5-8 and 5-14) are small, on the order of 1 %. The E_{had}/E_{em} requirement, however, carries a greater systematic uncertainty. Figure 5-9 shows the ability of the Monte Carlo to model this variable. The efficiency as a function of $\Sigma_p^{0.2}$ measured from the conversion sample is convoluted with the distribution for electrons from b's in top Monte Carlo events and found to change the (soft electron) acceptance by 5%. The limited knowledge of the $BR(b \rightarrow \ell X)$ and $BR(b \rightarrow c \rightarrow \ell X)$ branching fractions contribute an overall uncertainty of $\sim 5\%$.² There is also a 5% uncertainty in in the tracking efficiency for

²The averaged measured semileptonic branching fraction of b-hadrons produced at LEP is $BR(b \rightarrow \ell^- \bar{\nu} X) = 0.112 \pm .004$ [101]. The branching ratio of the subsequent charm decay

Statistical Uncertainties	2 %
Trigger Efficiency	5 %
Jet Energy Scale	5 %
Initial state radiation	5 %
Final state radiation	2 %
Model Dependence	10 %
Soft lepton tag efficiency	10 %
Total uncertainty	17 %

Table 6.2: The uncertainty on the $t\bar{t}$ acceptance.

Run 1b.

The overall uncertainty on $\epsilon_{SLT\ tag}$ is checked in Appendix C by searching for soft lepton tags in a sample of b's selected with the SVX tagger. To cover these various effects, a 10% uncertainty is assigned to the combined soft electron and soft muon tagging efficiency in top events.

Table 6.2 summarizes the uncertainties in $A_{t\bar{t}}$. Added in quadrature they give a total uncertainty of 17 %.

Figure 6-2 shows the acceptance and expected yield in 90.1 pb^{-1} as a function of top mass.

to leptons is measured to be $\text{BR}(b \rightarrow c \rightarrow \ell^+ \nu X) = 0.08 \pm .01$ (OPAL and ALEPH average) [100] [99]. For comparison, in the $t\bar{t}$ Monte Carlo events used for this analysis the branching ratios are $\text{BR}(b \rightarrow \ell^- \bar{\nu} X) = 0.114 \pm 0.002$ (stat) and $\text{BR}(b \rightarrow c \rightarrow \ell^+ \nu X) = 0.098 \pm .002$.

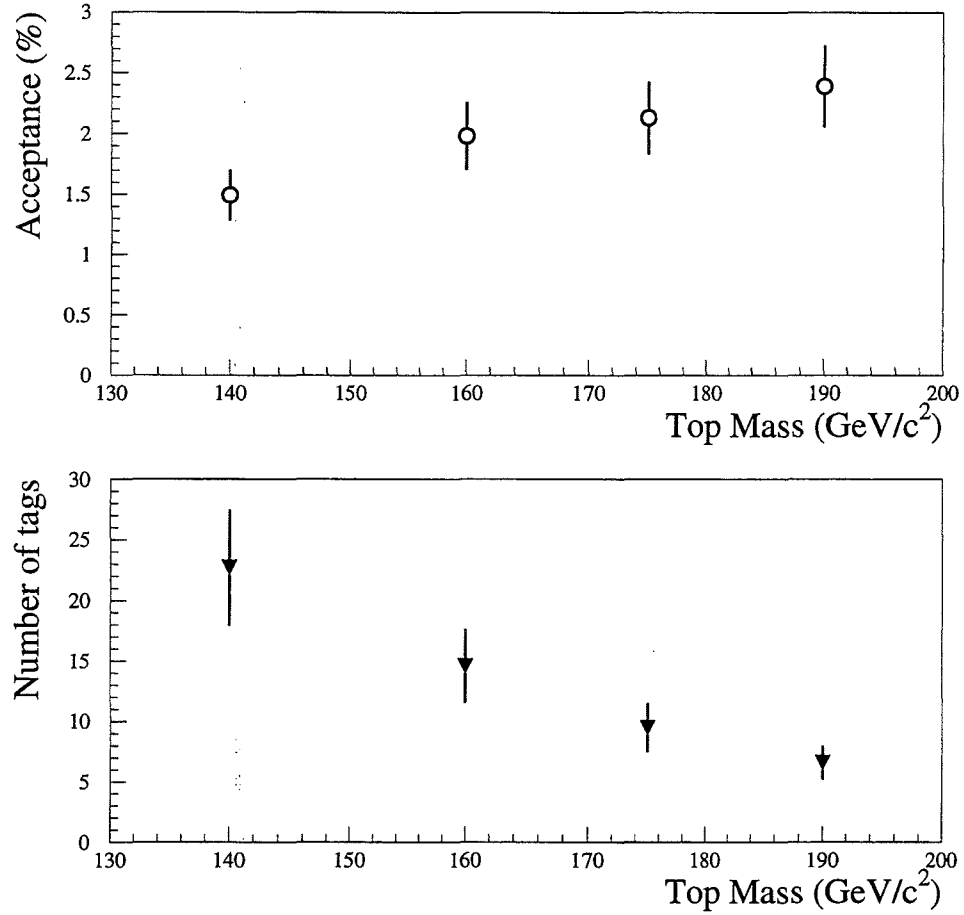


Figure 6-2: (Top) The $t\bar{t}$ acceptance (in percent) as a function of top mass. (Bottom) The number of soft lepton tags expected from $t\bar{t}$ events in 90.1 pb^{-1} of data using the theoretical cross section from Laenen *et al.* (Figure 1-6) [23]. A 16 % uncertainty has been assumed on the cross section.

Chapter 7

Application of the SLT to the W+jets data sample

This chapter describes the application of the SLT algorithm to the W+jets sample described in Chapter 4.

7.1 Background Removal

The W + jets sample contains some background from non-W sources which can generate a soft lepton tag. Examples are Drell-Yan, $b\bar{b}$, J/ψ or Υ production of leptons. Fortunately many of these events can be removed with simple cuts which do not significantly reduce the top acceptance. Residual backgrounds from these processes are discussed in the next chapter.

7.1.1 Drell-Yan

The process $p\bar{p} \rightarrow l^+l^-X$ can produce a lepton which passes the primary cuts, and a second which is soft lepton tagged. Although there are generally no neutrinos associated with Drell-Yan production, mismeasured jet energies can result in \cancel{E}_t and place the event in the W sample. For Drell-Yan pairs, the primary lepton and the SLT will have the same flavor and opposite charges,

and will typically be isolated. Isolation for a soft electron is defined as

$$Iso^{(soft\ e)} = \frac{E_t^{cone} - E_t^{ele}}{E_t^{ele}}$$

where E_t is the transverse energy of the soft electron. (As described in Section 5.1 this includes only the electromagnetic energy in the tower or towers associated with the electron) and E_t^{cone} is the transverse energy (hadronic and electromagnetic) in a cone of $R = 0.4$ around the electron. For muons isolation is similarly defined as

$$Iso^{(soft\ \mu)} = \frac{E_t^{cone} - E_t^{tower}}{P_t}$$

where E_t^{tower} is the energy in the calorimeter tower behind the muon, and P_t is the momentum of the muon track. For the purposes of Drell-Yan removal, soft leptons with $P_t > 20$ GeV/c are considered isolated if they have Isolation < 0.1 . If the soft lepton has $P_t < 20$ GeV/c, the requirement is loosened to $E_t^{cone} - E_t < 2$ GeV (for electrons) or $E_t^{cone} - E_t^{tower} < 2$ GeV for muons which is equivalent to the isolation cut at $P_t = 20$ GeV/c. Events where the primary lepton and the SLT are the same flavor, have opposite charge and are isolated are defined as Drell-Yan candidates and removed. For Drell-Yan pairs, the SLT leg is expected to have large P_t since the other has been selected as a primary lepton and so has $P_t > 20$ GeV.

To estimate the efficiency of this procedure for removing Drell-Yan pairs, a study is performed using $Z \rightarrow l^+l^-$ pairs where one lepton passes the primary cuts, and the other is tagged with the SLT algorithm. The removal procedure is found to be $87 \pm 1\%$ efficient for electron pairs, and $96.4 \pm 0.3\%$ of muon pairs. Although in these samples, both legs are higher P_t than the continuum Drell-Yan events which will appear in the W sample, the cuts are expected to be at least this efficient when the second lepton is softer and so need only pass the loosened isolation requirement.

Table 7.1 lists the number of Drell-Yan candidate events removed from the W+jets sample as a function of jet multiplicity. Also listed are the number of

	W + 1 jet	W + 2 jets	W+3 jets	W + ≥ 4 jets
Drell-Yan removal	49	15	1	0
Drell-Yan (same charge)	10	2	1	0
Low Mass removal	10	1	0	1
Upsilon removal	5	0	1	0
Residual $Z \rightarrow \mu^+ \mu^-$ removal	4	5	0	0

Table 7.1: The number of events identified by the various removal algorithms. Since there is some overlap, the numbers given for each are after the removal algorithms above have already been applied.

like charge pairs which *would* be removed, were they instead of opposite charge. As discussed in Section 8.4 this will be used to estimate the *overefficiency* of the Drell-Yan removal.

Figure 7-1 shows the momentum spectrum of the SLT in the Drell-Yan events. Also shown is the momentum spectrum for like charge pairs, that would otherwise be removed. As expected, many of the candidates have an SLT with large P_t .

7.1.2 Low Mass Lepton Pairs

Several background processes will result in events where the invariant mass of the SLT and the primary lepton is less than $5 \text{ GeV}/c^2$. Leptons pairs from sequential b decay ($b \rightarrow l\nu c, c \rightarrow lX$) are expected to have low mass. Also a high P_t J/ψ or ψ' may decay to an e^+e^- or $\mu^+\mu^-$ pair where one leg appears as a primary lepton, and the other as a soft lepton. All events in which the primary and soft lepton are of opposite sign and have an invariant mass of less than $5 \text{ GeV}/c^2$ are removed from the sample. Figure 7-2 shows the mass distributions for same flavor pairs in the *inclusive* electron and muon samples (no Isolation or E_t cuts applied) where the mass for opposite sign pairs has

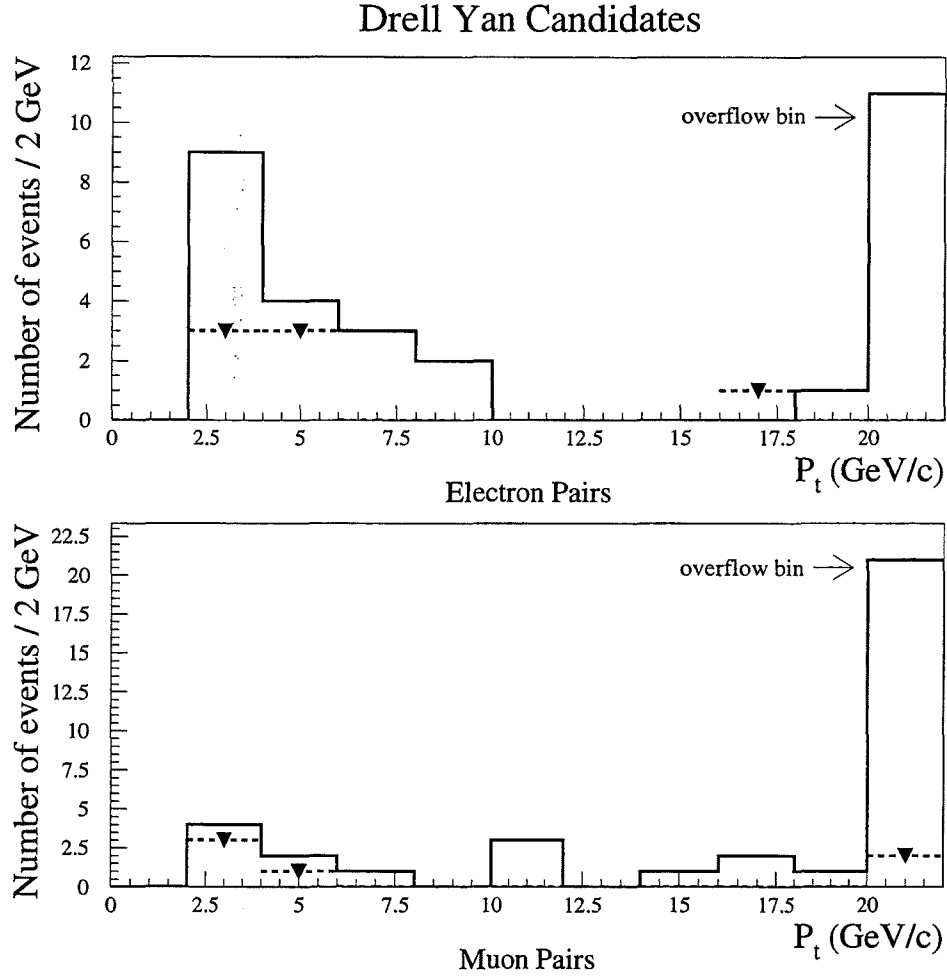


Figure 7-1: P_t spectrum of SLT's in events removed as Drell-Yan candidates from the $W + \geq 1$ jet sample. The top plot shows electron events, the bottom shows muon events. In each the histogram represents the opposite charge pairs that were actually removed and the triangles represent the same charge pairs that pass all other Drell-Yan removal cuts. The highest P_t bin contains entries for $P_t \geq 20$ GeV/c.

been plotted, and the same distribution for same sign pairs subtracted. The plot shows a falling distribution at low mass from sequential decays, and a mass peak for the J/ψ . In the muon plot a peak due to ψ' is also visible. Also apparent is a bump around $9.5 \text{ GeV}/c^2$ from the upilon which is discussed next. There are fewer electron pairs than muon pairs because the primary electron selection criteria contain implicit isolation requirements which reduce the probability of finding a second lepton nearby.

The plot demonstrates the level of these backgrounds in the inclusive lepton samples. After the E_t and isolation cuts have been applied to select the W sample, only a handful of events are removed with the low mass cuts. Table 7.1 lists the number of events removed as a function of jet multiplicity.

7.1.3 Upsilon

As shown in Figure 7-2 a mass bump corresponding to the Υ is visible around 9.5 GeV . Although the background from Υ 's is expected to be extremely small in the W sample, events are removed where the primary lepton and the SLT are of opposite sign, have the same flavor and form an invariant mass $8 \text{ GeV}/c^2 < M_{\text{primary}+SLT} < 11 \text{ GeV}/c^2$, since it would be hard to justify a top candidate event where the SLT and primary lepton formed an invariant mass near the Υ mass. There is, in fact, one such event removed from the W+3 jet bin. Table 7.1 lists the number of events removed as Υ candidates in the W sample.

7.1.4 Z Bosons (second pass)

The cuts described in 4.1.3 were the standard Z removal cuts used for CDF top analyses. However, since this analysis is particularly susceptible to background processes which produce two leptons, further cuts are required. In particular, since the Z removal discussed in 4.1.3 makes isolation and minimum ionizing cuts on the second muon leg, it may miss a lepton if there is jet activity

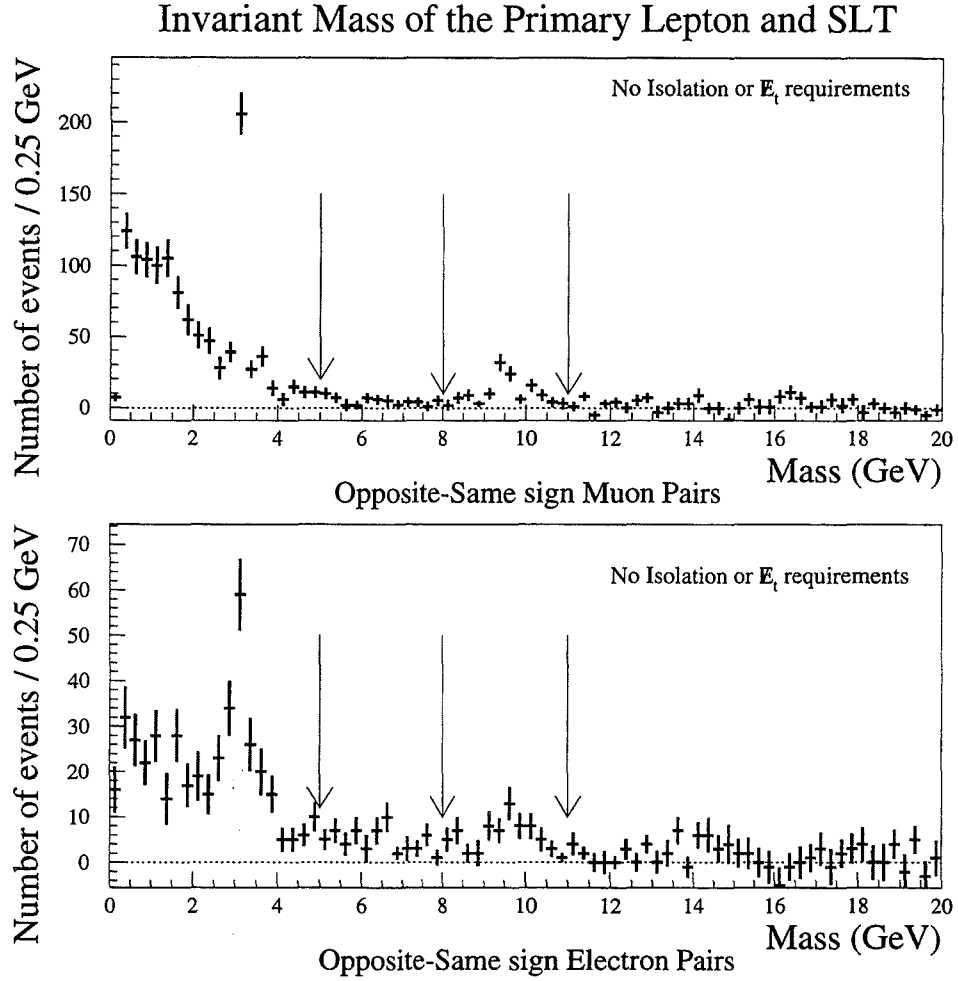


Figure 7-2: The invariant mass of same flavor primary lepton and SLT pairs in the inclusive lepton sample. Shown is the mass for opposite sign leptons with the distribution for same sign pairs subtracted. Drell-Yan pairs have been removed as described in the text. Visible are the low mass hump from the sequential decay $b \rightarrow l\nu c, c \rightarrow lX$, and mass peaks from the J/ψ , ψ' (for muons) and the Υ . Events with primary-SLT pair masses below $5 \text{ GeV}/c^2$ or between 9 and $11 \text{ GeV}/c^2$ are removed. Note also that there is no excess of events for larger masses. After the Isolation and E_t requirements have been imposed to select the W sample, only a handful of events are removed by these cuts. (See Table 7.1.)

Jet multiplicity	Events	Soft lepton tags	Soft lepton tagged events	Double tags
W+1 jet	9001	201	197	4
W+2 jets	1376	66	66	0
W+3 jets	212	23	21	2
W+ ≥ 4 jets	61	15	13	2

Table 7.2: The number of events before tagging, the number of soft lepton tags, and the number of soft lepton tagged events in the Run 1b W+jets data sample. There are four double tagged events in the $W+\geq 3$ jet bin and four in the W+1 jet bin.

in the event which overlaps with the lepton, and since the SLT algorithm is designed to identify leptons in jets, it may find these. Events where the primary and SLT have opposite charge, are of the same flavor and make a mass $70 \text{ GeV}/c^2 < M_{\text{primary}+SLT} < 110 \text{ GeV}/c^2$ are removed from the sample. Table 7.1 lists the number of events removed as $Z \rightarrow \mu^+\mu^-$ candidates.

7.2 Soft Leptons in the W+Jets Sample

Figure 7-3 shows the number of soft lepton tags found in the W+jets sample as a function of jet multiplicity. The background removal cuts described in this Chapter have been applied. Table 7.2 lists the number of events, tags and tagged events as a function of jet multiplicity. There are four double tagged (two soft lepton tags) events in the W+1 jet bin, none in the W+2 jet bin, two in the W+3 jet bin and two in the W+ ≥ 4 jet bin. Requiring a soft lepton tag has reduced the number of events in the W+ ≥ 3 jet bin by a factor of seven.

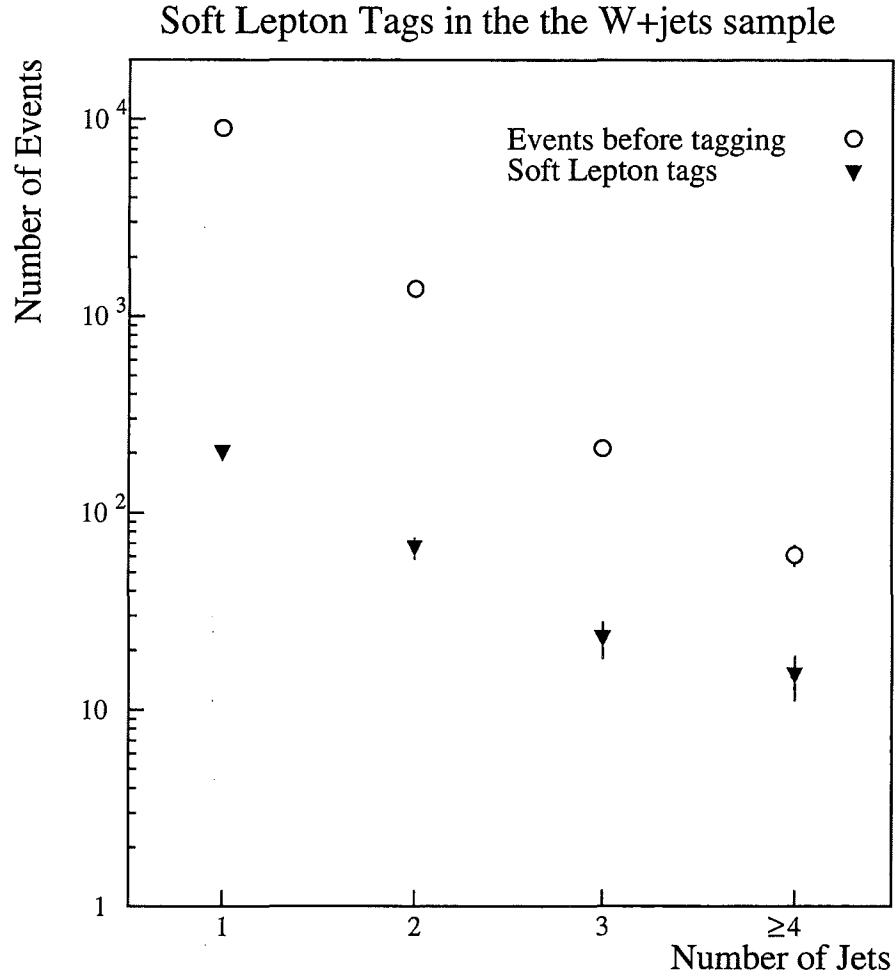


Figure 7-3: Shown are the number of W+jets events before and after the application of the soft lepton tagger. The last multiplicity bin is an overflow bin corresponding to events with at least four jets. The data is for Run 1b only. Results from Run 1a are added in Section 9.2.1.

Chapter 8

The SLT Backgrounds

The major background to the top signal in this analysis comes from “fake” soft lepton tags. Fake tags can be defined as particles identified as leptons whose origin is not a heavy flavor decay. This includes non-leptons which pass the lepton selection cuts (such as a pion faking an electron or a muon) as well as electrons from conversions or muons from pions or kaons which decay in flight. Other backgrounds can come from W bosons produced with real heavy flavor pairs ($Wb\bar{b}$, $Wc\bar{c}$). As will be discussed, backgrounds from these three processes (fakes, $Wb\bar{b}$ and $Wc\bar{c}$) are estimated with a track based fake rate derived from a sample of “generic jets” which are not expected to be enhanced in heavy flavor.

There are other background processes that can produce a W and a SLT that are not accounted for by the fake rate. These include the production of a W with an associated charm quark (Wc , shown in Figure 3-6) and diboson (WW, WZ, ZZ) production.

Non- W backgrounds can persist in the sample and contribute real soft lepton tags that have not yet been accounted for. These include backgrounds from $b\bar{b}$, $Z \rightarrow \tau^+\tau^-$ production, and residual Drell-Yan events not removed by the event cuts (Section 7.1).

Because of the complication of accurately modeling fake soft leptons, the majority of the backgrounds are estimated using the data itself (and some

theoretical assumptions). Because this analysis was targeted at observing the as yet unobserved, the philosophy has been to minimize uncertainties from theoretical cross sections and Monte Carlo.

Along the way various tests of the background calculations are performed. The reliability of the fake rate is tested by applying it to a photon sample and, most appropriately, a Z +jets sample where the heavy flavor content is expected to be similar to the W sample (excluding top). To check the residual $b\bar{b}$ background two methods are presented, one which relies entirely on the data, and another which uses a combination of data and Monte Carlo. In the Appendix, an alternative, Monte Carlo based method of measuring the $Wb\bar{b}$ and $Wc\bar{c}$ backgrounds is presented which attempts to calculate their contributions explicitly. In the end, the best check of the background estimates comes from comparing soft lepton tags and backgrounds in the $W+1$ and $W+2$ jet bins, since these are events most like those in the signal region of $W+ \geq 3$ jets.

8.1 The Soft Lepton Fake Rate from Jet Data

8.1.1 A Justification

The bulk of the background ($\sim 90\%$) is estimated with a track based fake rate derived from a sample of “generic” jets collected with triggers that require jet clusters over 20, 50, or 70 GeV as described in Section 4.5. The fake rate applied to the W data, should predict the number of fake tags as well as some real soft leptons from $Wb\bar{b}$ and $Wc\bar{c}$ production since these are expected to arise from gluon splitting which should be present also in the jet sample. The “fake” rate could perhaps be more accurately characterized as a “jet-rate” since it incorporates both fakes and some real heavy flavor tags.

The assumption inherent in applying this fake rate to the signal sample is that the jets in W +jets events are similar to the ones in the jet sample. (“A jet is a jet” the aphorism goes.) “Similar”, here refers partly to the heavy flavor

content of the jets, but also to the *fake* component which may also be different in the generic jet samples compared to the W+jets sample. Biases due to the jet trigger and the variations in the fake rate due to track isolation and P_t are expected to be the largest of these effects and are taken into account.

The Heavy Flavor Content in the Generic Jets and the W+jets Samples

Although studies measuring correlations between SVX (displaced vertex) tags and soft lepton tags estimate that only $\sim 15\%$ of soft lepton tags in generic jets are due to real heavy flavor (see Appendix A) and hence that the background is likely to be dominated by fakes, it is still necessary to compare the heavy flavor content of the generic jet sample to the W+jets where in principle the heavy flavor content could be larger. The heavy flavor content in the two samples will depend on the underlying physics processes which can produce it. A variety of theoretical arguments indicate that there should be more heavy flavor in the generic jet sample.

First, heavy flavor pairs in W+jets events (recall that Wc is treated separately) is expected to be produced entirely by gluon splitting [27]. In the generic jet sample however, heavy flavor is expected to come from several processes (including gluon splitting) which at tree level can be placed in three categories. Examples of all three are shown in Figure 8-1. They are

- **Direct Production:** Direct Production covers processes like $q\bar{q}' \rightarrow (b\bar{b}, c\bar{c})$ and $gg \rightarrow (b\bar{b}, c\bar{c})$.
- **Gluon Splitting:** This is when a final-state gluon splits into a heavy quark pair.
- **Flavor Excitation:** In this process an initial-state gluon produces a heavy flavor pair.

NLO calculations [28] predict that gluon splitting should account for approximately 65 % of $c\bar{c}$ pairs and 75 % of $b\bar{b}$ pairs produced in the jet sample.

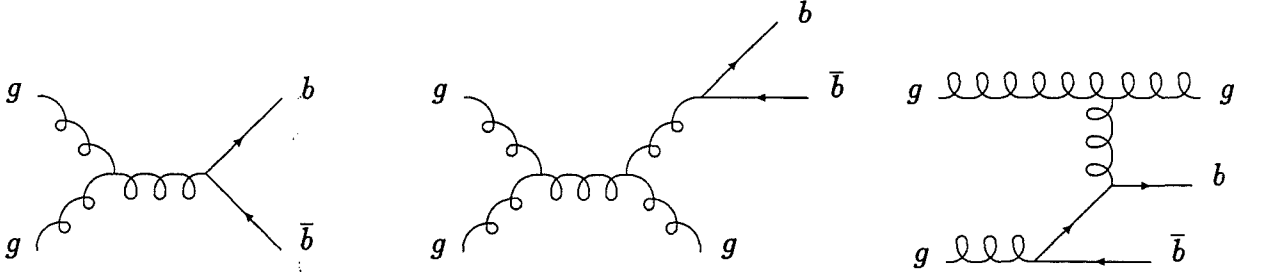


Figure 8-1: Diagrams for heavy flavor production in generic jet events - direct production (left), gluon splitting (middle), flavor excitation (right). In each case the $b\bar{b}$ pair may be replaced with a $c\bar{c}$ pair.

Since there are other processes which can produce heavy flavor in generic jets, it is reasonable to assume that the heavy flavor content is higher there.

To complete the argument, however, one needs to know what fraction of the jets in each sample come from gluons since more gluons in W+jets could lead to a larger $b\bar{b}$ component). One estimate of the gluon content comes from Monte Carlo studies which use HERWIG and VECBOS and predict a fraction of gluon jets in W+jets events smaller than in generic-jet events by a factor of 1.4 to 2 for jet multiplicities between 1 and 4. Recent studies comparing HERWIG predictions to multijet data at CDF [30], show good agreement in jet rates, mass and P_t , which gives some confidence as to its ability to model gluon radiation.

A HERWIG study which incorporates all of these effects, by calculating the ratio of actual heavy flavor production in generic jets to W+jets, obtains a similar factor of ~ 1.3 . HERWIG predictions of the heavy flavor content in generic jets have been shown to give reasonable agreement [34] to CDF data and to a full NLO calculation [36] of $b\bar{b}$ production. HERWIG has also been compared to a Monte Carlo based on a full NLO calculation of $Wb\bar{b} + Wc\bar{c}$

production [37] and also gives reasonable agreement [34].

For these reasons the fake rate obtained from generic jet events is expected to overestimate the background in W +jets events. This would lead to a conservative estimate of the significance of an observed excess of events, but would also lead to an underestimate of the $t\bar{t}$ cross section. For better or worse, the fake rate turns out to be dominated not by real heavy flavor tags, but by fake lepton tags, and so serves as a good predictor for other samples. The reliability of these assumptions will be tested by applying the fake rate to a variety of samples which should have different heavy flavor content.

8.1.2 Definition of the Fake Rate

The fake rate is defined as the fraction of fiducial tracks which pass the soft lepton id cuts described in Chapter 5. Since the fiducial regions for CMU-only, CMP-only and CMUP muons all overlap, a combined “CMU/P” muon fake rate is used which includes all these types and is defined as the number of SLT CMU-only muons, CMP-only muons and CMUP muons divided by the number of tracks of Type 1, 2, or 3 as defined in Section 5.2. This is effectively the number of SLT muons per track which extrapolates to the combined CMU and CMP regions. The CMX fake rate is defined as the number of good CMX muons divided by the number of tracks that extrapolate to within the CMX boundary (also described in Section 5.2). For soft electrons the fake rate is defined as the number of good electrons divided by the number of fiducial electron tracks as described in Section 5.1. Unless otherwise stated, the fake rates are obtained from a combination of Jet 20, 50, and 70 triggers. In these samples, 75 % of fiducial lepton tracks are within a cone of 0.4 of a jet with $E_t > 15$ GeV. Fake rates are calculated with the identical algorithm used to calculate the efficiencies from J/ψ ’s, photon conversions, and Z ’s.

8.1.3 Determination of the Muon Fake Rate

Since the muon matching cuts are dependent on P_t , and since the background may have a P_t dependence, the muon fake rate is parameterized as a function of the track P_t .

There is a trigger bias to the fake rate which can be seen by plotting the fake rate for all tracks in the event, and for tracks which are well separated from a trigger jet. As shown in Figure 8-2, the rate for tracks which are well separated from a trigger jet have a fake rate which is $\sim 15\%$ larger than the rate for all tracks. This can be understood as a consequence of the trigger requirement that the jet be over a particular threshold; jets which contain a muon (from decay in flight for instance) will have their energy measured systematically low in the calorimeter, and are hence less likely to pass the trigger. Because the underlying jet E_t distribution is steeply falling, lost energy from muons can have a large effect. By the same argument, non-trigger jets are *more* likely to have a muon since they by definition have less energy than the trigger jet. The bias, however, is entirely removed when the fake rate is calculated from tracks which are well separated from a trigger jet ($\Delta R > 0.7$). This means that if both jets fire the trigger, all tracks in the event are used. Figure 8-2 also shows the unbiased rate divided by the biased rate. As expected, the size of the effect increases with P_t where the energy carried away by the muon is larger.

Figure 8-3 shows the (unbiased) muon fake rates as a function of P_t obtained from ~ 3 million tracks. To test for a possible isolation dependence, the fake rate is plotted as a function of $\Sigma_p^{0.2}$ after correcting for the P_t dependence in Figure 8-4. Within statistical uncertainties no correlation is observed. The fake rate is around 1% for tracks with $P_t = 2$ GeV/c and falls for higher P_t . This behavior is in part a consequence of the matching chi-square requirements which effectively apply a tighter matching cut for higher P_t tracks.

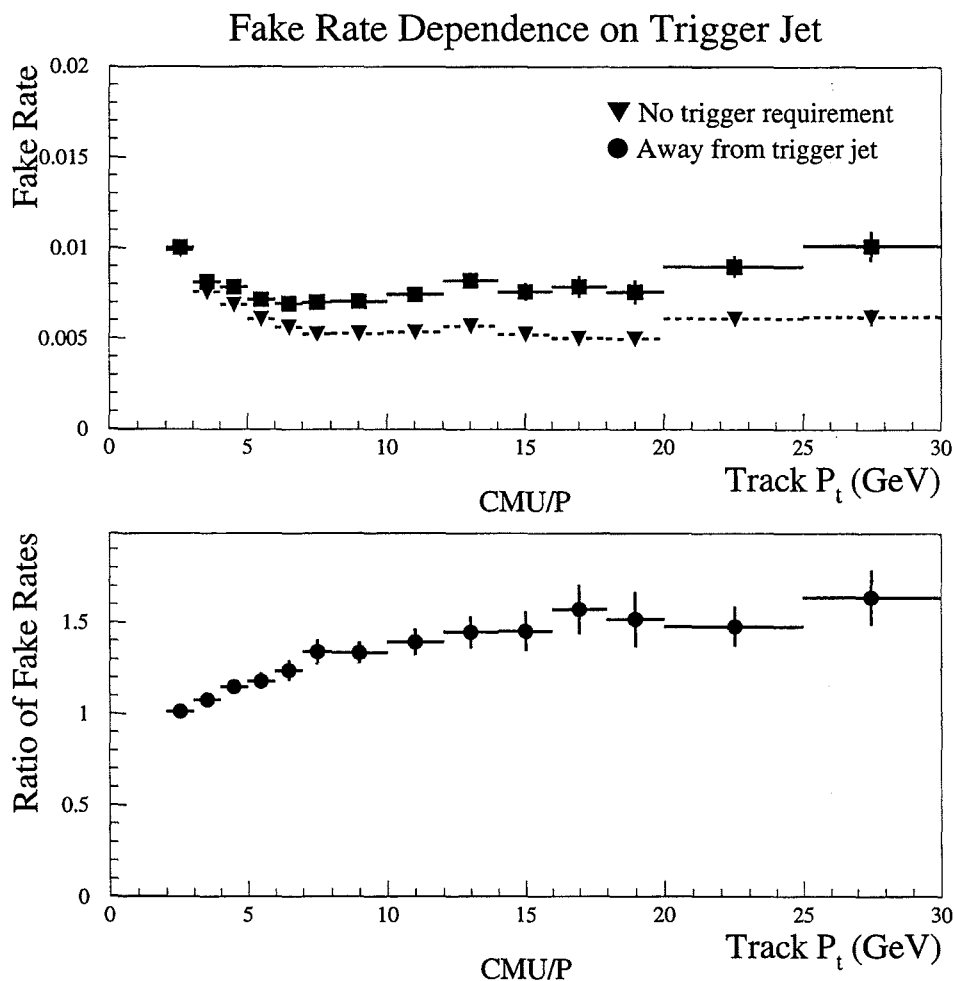


Figure 8-2: Top plot: CMU/P fake rate as a function of track P_t in jet events shown for all tracks (circles) and for tracks which were required to be away from a jet which passed the trigger (triangles). Bottom plot: The ratio of the fake rate for tracks away from trigger jets to the fake rate for all tracks. Jets containing muons are less likely to fire the trigger since they will have their energy measured systematically low. The rate for tracks well separated from a jet which would have caused the trigger to fire, is free from trigger bias. The effect is more pronounced at high P_t where the energy carried away by the muon is greater.

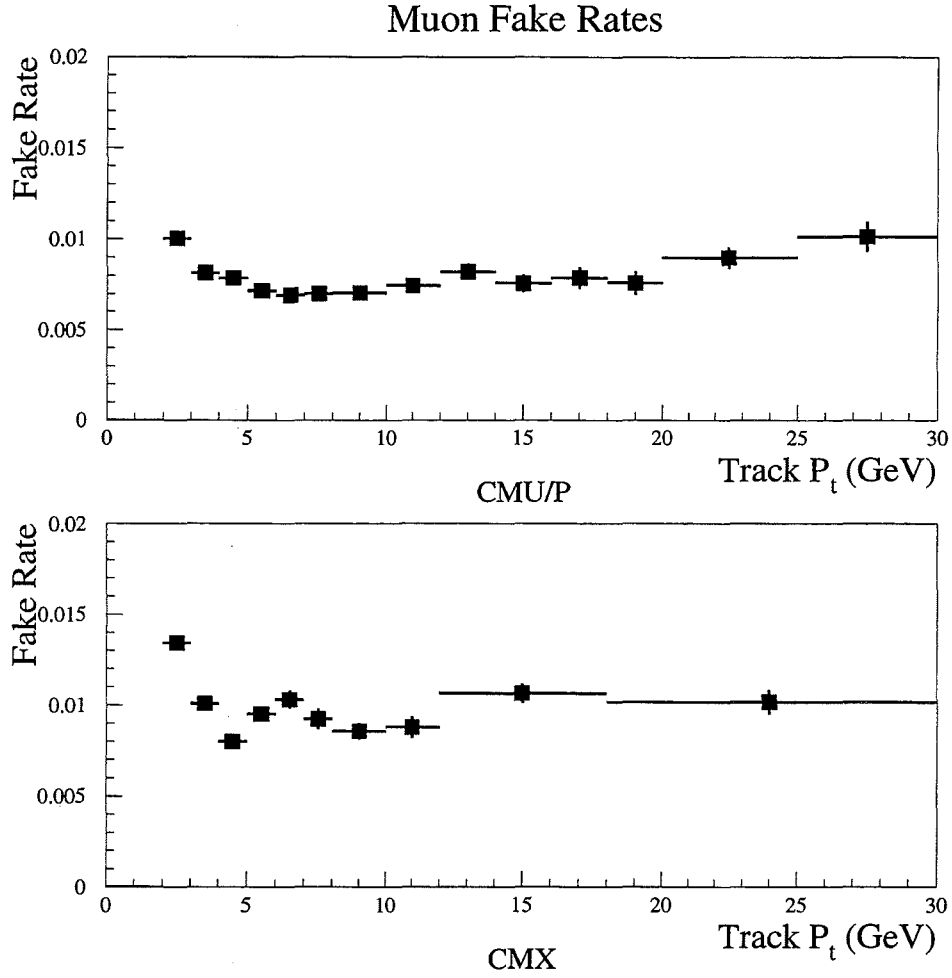


Figure 8-3: The fake rate for tracks extrapolating to the CMU or CMP (top) and to the CMX (bottom). The fake rate initially drops with P_t in part because the cut on matching chi-square becomes tighter with increasing P_t where multiple scattering is reduced for real muons. The selection criteria for the CMX change at $P_t=5$ GeV/c (Section 5) which causes a kink in the fake rate.

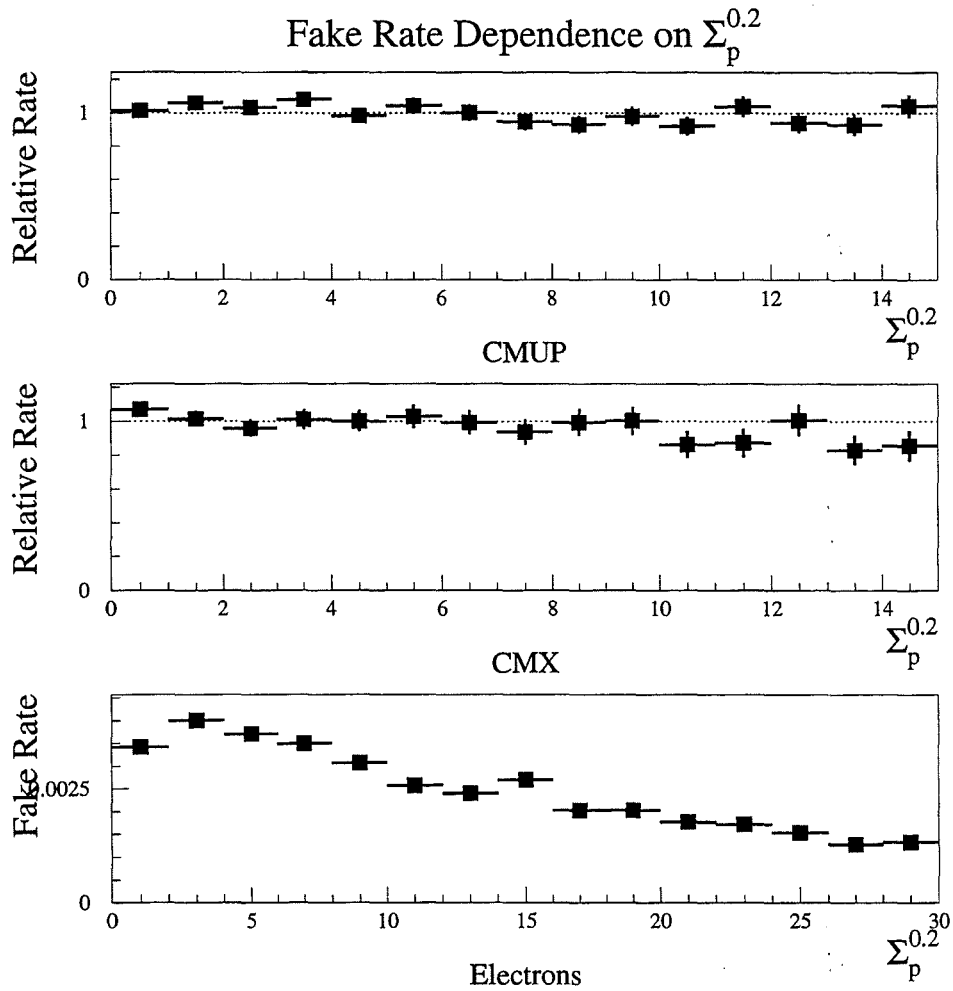


Figure 8-4: The muon fake rates as a function of $\Sigma_p^{0.2}$ for tracks extrapolating to the CMU/P (top) and CMX (middle) regions where the momentum dependence shown in Figure 8-3 has been removed. Also shown (bottom) is the electron fake rate as a function of $\Sigma_p^{0.2}$. Because of this dependence, the electron fake rate is calculated in three bins of $\Sigma_p^{0.2}/P$.

8.1.4 Determination of the Electron Fake Rate

The soft electron fake rate is expected to be dependent on both isolation and track P_t . The isolation of a track can be characterized by a quantity $\Sigma_p^{0.2}/P$ where $\Sigma_p^{0.2}$ is the sum of track momentum in a cone of $R=0.2$ around the track, and P is the track momentum. Figure 8-4 shows the electron fake rate as a function of $\Sigma_p^{0.2}/P$. The final fake rate is broken up into three bins, $\Sigma_p^{0.2}/P < 0.2$, $0.2 < \Sigma_p^{0.2}/P < 5$, and $\Sigma_p^{0.2}/P > 5$. $\Sigma_p^{0.2}/P$ has been chosen instead of just $\Sigma_p^{0.2}$ because isolation dependent cuts (E_{had}/E_{em} and E/p) effectively divide by the track momentum. The final rates obtained from ~ 4 million tracks are shown in Figure 8-5. The fake rate is lowest in the $\Sigma_p^{0.2}/P > 5$ bin because the track is less likely to pass the E_{had}/E_{em} and E/p cuts.

8.2 Systematic Uncertainty on the Fake Rate

The uncertainty of the Fake Rate can be determined by applying it to a variety of samples and comparing its predictions with the observed number of tags. This covers uncertainties due to the parameterization and variations in the heavy flavor content of the samples.

8.2.1 Check of Fake Rate Using Jet Samples

The simplest way to assess the reliability of the fake rate is to compare it across jet samples. A fake rate, for instance, can be calculated with the JET20 sample, and then applied to the combined JET50 and JET70 samples. Table 8.1 shows the predicted and observed tags using fake rates calculated using a single jet sample, and applying it to the remaining two. The uncertainties listed are statistical from the fake rate itself.

The fake rate performs least well for soft electrons. In the case where the rate was derived from the JET20 sample and applied to the combined JET50 and JET70 samples, 27 % more tags are predicted than observed. The rate

Electron fake rate

Fake Rate from	Test Sample	Predicted	Observed	Difference/Observed
JET 20	JET50+JET70	8814 \pm 201	6939	+ 27 %
JET 50	JET20+JET70	6596 \pm 114	6790	- 3 %
JET 70	JET20+JET50	5912 \pm 101	6459	- 9 %

Muon fake rate

<i>CMU/P</i>				
Fake Rate from	Test Sample	Predicted	Observed	Difference/Observed
JET 20	JET50+JET70	12019 \pm 219	11782	- 2 %
JET 50	JET20+JET70	10128 \pm 138	10015	-1 %
JET 70	JET20+JET50	8886 \pm 112	8939	-0.6 %
<i>CMX</i>				
JET 20	JET50+JET70	6091 \pm 147	6283	- 3 %
JET 50	JET20+JET70	5459 \pm 101	5428	+ 1 %
JET 70	JET20+JET50	5004 \pm 87	4879	+3 %

Table 8.1: The predicted and observed tags for fake rates obtained from generic jet samples collected with different triggers. In each case the fake rate is applied to an independent sample. For example the fake rate derived from the JET20 sample is applied to the combined JET50 and JET70 samples. The muon fake rate predictions agree with the observed tags to a few percent which is within the statistical uncertainty expected, the electron fake rate performs less well across samples.

derived in the JET70 sample, applied to the combined JET20 and JET50 sample underestimates the number of tags by 9%. It might seem that given the variation in fake rates between the three isolation bins, the situation could be improved by better accounting for the isolation dependence. A few alternative techniques have been investigated which include finer $\Sigma_p^{0.2}/P$ binning, smoothing the entire distribution in the $\Sigma_p^{0.2}/P$ vs P_t plane to obtain a continuous fake rate, and applying a straight line extrapolation between the three bins to estimate the fake rate for intermediate values of $\Sigma_p^{0.2}/P$. In each case no significant improvement in the ability of the electron fake rate to predict the number of tags in other samples was observed. Since the fake rate drops for higher Jet E_t triggers, it seems likely that there remains an isolation dependence which is not characterized by $\Sigma_p^{0.2}/P$. Since the accuracy of this method of calculating the background is eventually limited by the uncertainties on the heavy flavor content of generic jets compared to the W sample, and since electrons comprise less than a third of the SLT acceptance, this is not pursued further. The agreement between the muon predictions and the observed tags is excellent, and is at the level of the statistical uncertainty expected from the fake rate.

8.2.2 Check of Fake Rate Using the Photon Sample

While the test of the fake rate using the generic jet samples says something about its predictive power in a sample selected in a similar manner to the one it was derived from, it should also be tested on completely unrelated events. In particular it is useful to look in samples where the heavy flavor content is expected to be different from the generic jet sample. To this end, the fake rate is applied to a “photon plus jet” sample of 44053 events which has been described in 4.6. Here, additional heavy flavor is expected from charm quarks in the proton sea which scatter off a gluon and produced a photon [42]. Table 8.2 compares the predictions from the fake rate with the found tags. The agreement for the CMU/P muons and for the electrons are both within the expected statistical uncertainty from the number of found tags. The prediction

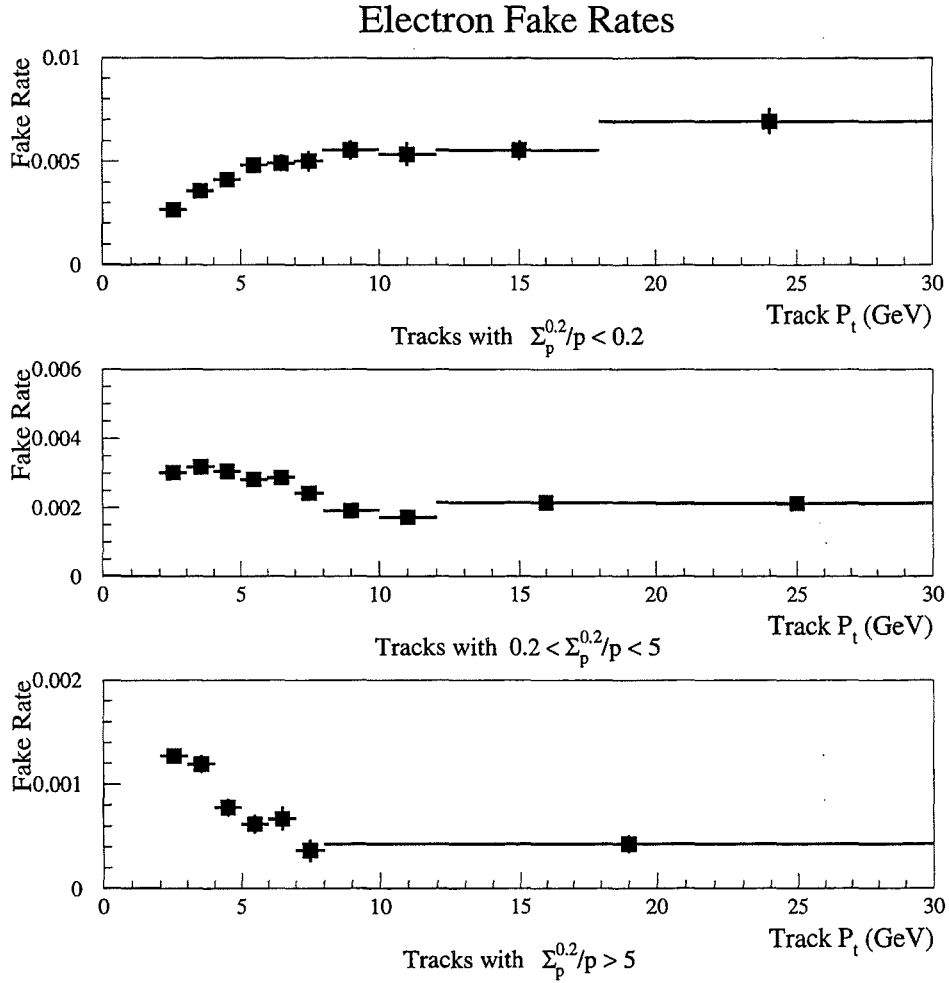


Figure 8-5: The electron fake rate as a function of track P_t for three different bins in $\Sigma_p^{0.2}/P$. The fake rate drops with increasing $\Sigma_p^{0.2}/P$ where tracks are less likely to pass the E/p and E_{had}/E_{em} cuts.

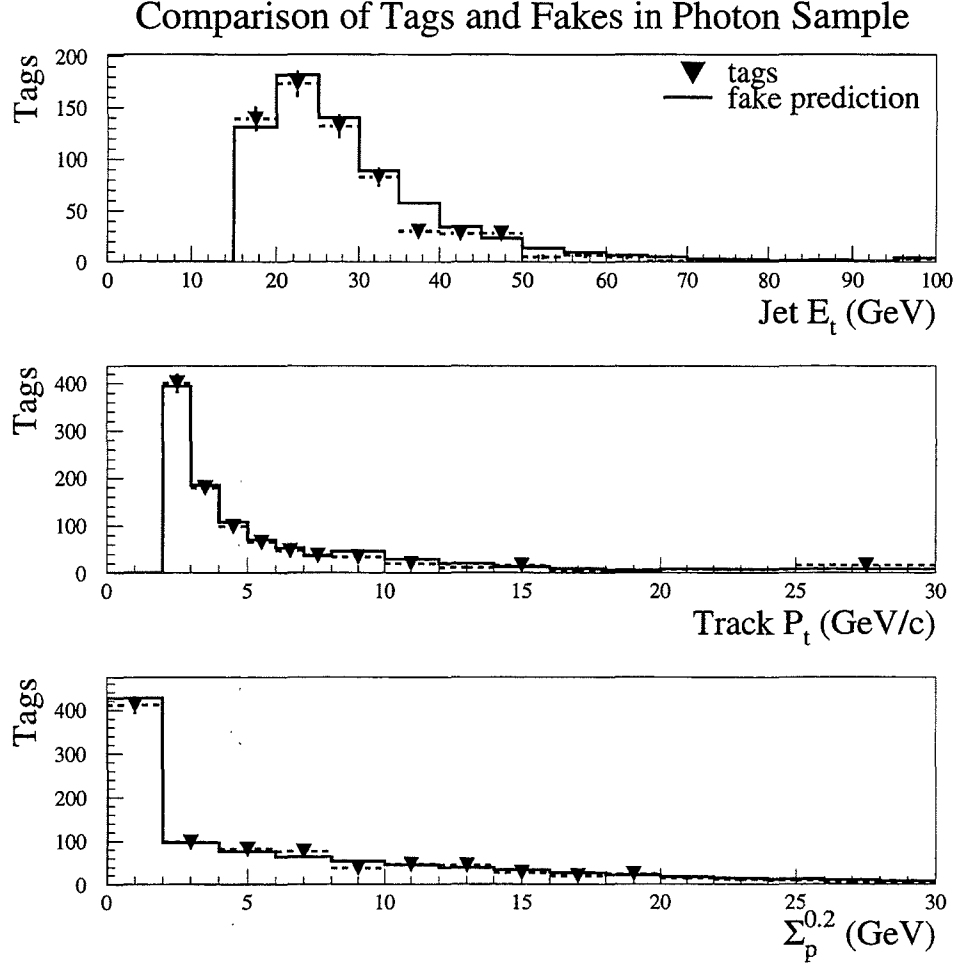


Figure 8-6: In the photon sample, the track P_t , $\Sigma_p^{0.2}$ and E_t of jet in a cone of $\Delta R=0.4$ around the soft lepton (triangles) and as predicted by the fake rate (histogram).

for CMX muons is 16 % high, although there is an 8% statistical uncertainty from the number of observed tags. Combining all SLT types, 985 tags are predicted and 935 observed. Overall the prediction is high by 5%.

It is useful to compare also the distributions for the tags to the predictions from fake rates. In Figure 8-6 the expected P_t , $\Sigma_p^{0.2}$ and Jet E_t distributions from the fake rates is compared to the distributions for the actual tags. Here Jet E_t refers to the jet within a cone of $\Delta R=0.4$ of the SLT. (If the SLT is not near a jet, it does not appear in this plot.) For all three variables the agreement is good, indicating that the fake rates have been parameterized in a sensible manner, and that there are no unexpected biases such as an excess

<i>SLT type</i>	<i>fake prediction</i>	<i>observed tags</i>	<i>Difference/Observed</i>
CMU/P muons	498.9	478	+ 4 \pm 5 %
CMX muons	246.4	211	+ 16 \pm 8 %
Electrons	239.4	246	- 3 \pm 6 %
TOTAL	984.7	935	+5 \pm 3 %

Table 8.2: Test of the fake rate in the photon + jet sample comparing the expected tags calculated using the fake rate to the number of observed tags.

of tags in high energy jets.

8.2.3 Check of Fake Rate Using the Z Sample.

The Z sample (Section 4.7) provides a good check of the fake rate while also testing the validity of the assumption that it can be applied to the W sample. If there were an anomalous source of heavy flavor in W events (not due to top) this might also be apparent in the Z sample. In this sense the Z's provide the perfect control sample - a selection of high-purity vector bosons with no expected top component.

Unfortunately there are an order of magnitude fewer Z's than W's produced which decay to charged leptons, and so statistics are limited. The predicted and observed tags as a function of jet multiplicity are listed in Table 8.3. Within the limited statistics available, the agreement is excellent. Overall 83 tags are expected and 82 observed.

8.2.4 Systematic Uncertainty of the Fake Rate

Based on the above three studies a 20% uncertainty is assigned to the electron fake rate and a 10% uncertainty to the muon fake rate. Since these uncertainties are expected to be uncorrelated, the systematics are added in quadrature when

	Z+0 jets	Z+1	Z+2	Z+ ≥ 3	Z+any
Number of events	8586	861	129	18	9594
CMU/P fake prediction	24.9	11.2	3.0	0.7	39.8
CMU/P tags observed	24	9	1	1	36
CMX fake prediction	16.1	5.4	1.6	0.2	23.3
CMX tags observed	19	6	0	0	25
Electron fake prediction	12.7	5.2	1.4	0.3	19.6
Electron tags observed	15	5	1	1	21
TOTAL predicted	53.7	21.8	6	1.2	82.8
TOTAL observed	58	20	2	2	82

Table 8.3: Expected tags calculated using the fake rate and observed tags in the Z sample. Overall 83 tags are predicted by the fake rate and 82 are observed.

predictions from the two are combined. The statistical error on the fake rate is ignored since this is much smaller in all cases ($\sim 3\%$).

8.3 Application of the Fake Rate to the W Sample

The fake rate applied to the W sample predicts 202 ± 18 tags in the W+1 jet bin, 57 ± 5 tags in the W+2 jet bin, 13.2 ± 1.2 in the W+3 jet bin and 6.1 ± 0.6 in the W+ ≥ 4 jet bin. The results are summarized in Table 8.8.

Jet Multiplicity	Electrons	Muons
Z+any	$13.2 \pm 0.8 \%$	$3.4 \pm 0.3 \%$
$Z + \geq 1$ jet	$18 \pm 3 \%$	$10 \pm 2 \%$
$Z + \geq 2$	$36 \pm 10 \%$	$18 \pm 6 \%$

Table 8.4: The underefficiency of the Drell-Yan removal algorithm in $Z \rightarrow l^+l^-$ events as a function of jet multiplicity.

8.4 Drell-Yan background

8.4.1 Drell-Yan Removal Underefficiency

Although Drell-Yan candidates where both legs are isolated have already been removed (Section 7.1), some may remain, especially if there is jet activity in the event which can overlap with one of the leptons. The underefficiency of the removal criteria is estimated by applying them to a sample of $Z + \text{jets}$ events, where the Z decays to leptons, one which passes the primary lepton cuts, and the other which is tagged as an SLT. It is found that the Drell-Yan removal identifies 1598 out of 1842 $Z \rightarrow e^+e^-$ events, and 2760 out of 2857 $Z \rightarrow \mu^+\mu^-$ events. The underefficiency of the removal then, is $13 \pm 1\%$ for electrons and $3.4 \pm 0.3\%$ for muons. Figure 8-7 shows the isolation for electrons and muons from Z 's.

The underefficiency is likely to be effected by jet activity in the event, and so is calculated separately for $Z + \geq 1$ jet and $Z + \geq 2$ jets. The results are shown in Table 8.4. Although the statistics are low, the removal is clearly less efficient when there are other jets in the event. The numbers in the $Z + \geq 2$ jet bin compare well with the isolation efficiency for leptons from W 's in top Monte Carlo which is on the order of 85 %.

One might worry that the P_t of leptons from Z decay will be larger than the P_t of the leptons from Drell-Yan production which may end up in the sample. The isolation variable used to identify Drell-Yan candidates in particular, could

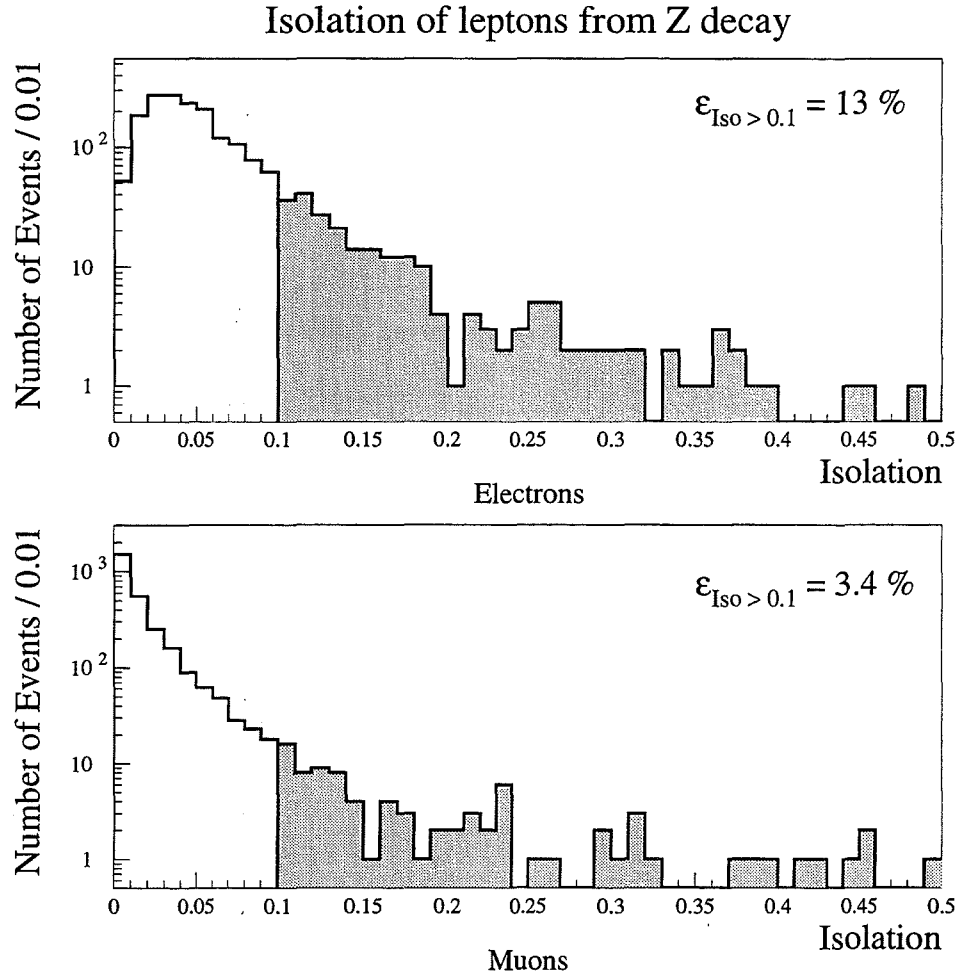


Figure 8-7: Isolation (I) for electrons and muons from Z decay selected with the SLT algorithm. The fraction of events with $I > 0.1$ gives the overefficiency of the Drell-Yan removal. The overefficiency increases with jet multiplicity as shown in Table 8.4.

become less efficient at low P_t . However, as discussed in Section 7.1, this is not expected to be a problem since SLT's with $P_t < 20$ GeV/c are removed if they have $E_t^{cone} - E_t < 2$ GeV. This is less stringent than requiring Isolation < 0.1 and in fact is equivalent to the isolation cut at $P_t = 20$ GeV/c. Consequently it should be an accurate estimate of the removal under efficiency.

8.4.2 Drell-Yan Removal Overefficiency

The loose Drell-Yan removal cuts at low P_t (see above) mean that some non Drell-Yan events will also be taken out. This must be corrected for this since these are events that are accounted for by other parts of the background estimate. If, for instance, a “fake” tag is removed as a Drell-Yan, then the estimate of the fake background will come up one event too high (if it were perfect). In a sense this is just bookkeeping.

A measure of the *overefficiency* of the Drell-Yan removal is obtained by counting the number of *same sign* pairs that pass all other Drell-Yan cuts, and would otherwise be removed. Assuming that there is no correlation between the sign of the primary and SLT in these events, the number of non-Drell-Yan's incorrectly removed should equal the number of these same-sign pairs. Table 7.1 lists the number of same-sign and opposite-sign pairs in the W sample.

8.4.3 Calculation of Remaining Drell-Yan

The number of residual Drell-Yan's is given by

$$Drell - Yan \ background = \frac{N_{opp. \ sign} - N_{same \ sign}}{\epsilon_{under \ efficiency}} - N_{same \ sign}$$

where $\epsilon_{under \ efficiency}$ is the under efficiency as measured from Z 's, $N_{opp. \ sign}$ is the number of opposite sign pairs removed and $N_{same \ sign}$ is the number

of same sign pairs that would have been removed, had they not been same sign. Subtracting $N_{same\ sign}$ compensates for the overefficiency of the removal. Where there has been more over-removal than under-removal, the resulting “background” is negative.

This calculation is performed for electrons and muons separately, with the underefficiencies from Table 8.4 applied as a function of jet multiplicity. For events with three or more jets with $E_t > 15$ GeV, an estimated underefficiency of 20 ± 10 % is used. The final results are listed in Table 8.8.

8.5 $b\bar{b}$ background in the W sample

Another source of non-W background in the W sample are $b\bar{b}$ events where one b decays leptonically ($b \rightarrow \ell \nu c$) producing a high P_t lepton and where neutrinos in the b decay or mismeasured jet energies result in \cancel{E}_t . These events can produce additional soft lepton tags if the decay of the other b in the event produces a lepton. Recall that events have already been removed from the W sample if the primary and the soft lepton form a low mass and consistent with the *sequential* decay of a single b (Section 7.1) so only leptons from the other b in the event are of concern.

The expected number of soft lepton tags from $b\bar{b}$ events in the W sample is calculated from two quantities.

- F_{non-W} : This is the fraction of non-W events in the W sample and is estimated by fitting a background shape to the tails of the \cancel{E}_t and *Isolation* distributions.
- $\epsilon_{b\bar{b}}$: This is the fraction of non-W events which contain a soft lepton from the decay of a b.

Multiplying these two numbers with the observed number of W candidates in

the data yields the expected number of soft lepton tags from $b\bar{b}$.

Two techniques are employed to measure $\epsilon_{b\bar{b}}$. Both methods use the low \cancel{E}_t lepton sample which is assumed to be representative of the non-W background. The first method looks for an excess of soft lepton tags over the prediction from the fake rate, the second method identifies lepton pairs from sequential decays and uses Monte Carlo to estimate how many $b\bar{b}$ events are in the sample, and how many soft lepton tags are expected from the decay of the other b .

8.5.1 Hadronic Background in the W Sample

Backgrounds in the W sample can be placed in two categories. The first are Z bosons which produce a primary lepton, either directly or through the decay of a τ , but are not removed with the Z removal (Section 4.1.3). The second source, is *hadronic* background, where a generic jet-type event produces something which is identified as a primary lepton. This hadronic component (which shall be referred to as “non-W” for simplicity) contains leptons from b decay and so is the quantity of interest here. In general these leptons from hadronic background (which can be real or fake) are not isolated and so the non-W component of the W sample can be estimated by extrapolating a background distribution into the signal region from the low \cancel{E}_t and the high lepton isolation (I) regions. Specifically, the \cancel{E}_t vs I plane is divided into four regions shown in Figure 8.5.1. Region D is the signal region and should contain mostly W’s. Under the assumption that \cancel{E}_t and I are uncorrelated for the non-W background, the non-W component of region D is given by

$$D = \frac{A}{B} \times C$$

Note that while it may seem reasonable that \cancel{E}_t and I are uncorrelated for a particular background (eg. light quark jets or $b\bar{b}$) in general the *sum* will have some correlation. (Consider the simple case where one distribution has $I < 0.1$ and is flat in \cancel{E}_t added to another distribution which has $\cancel{E}_t < 10$

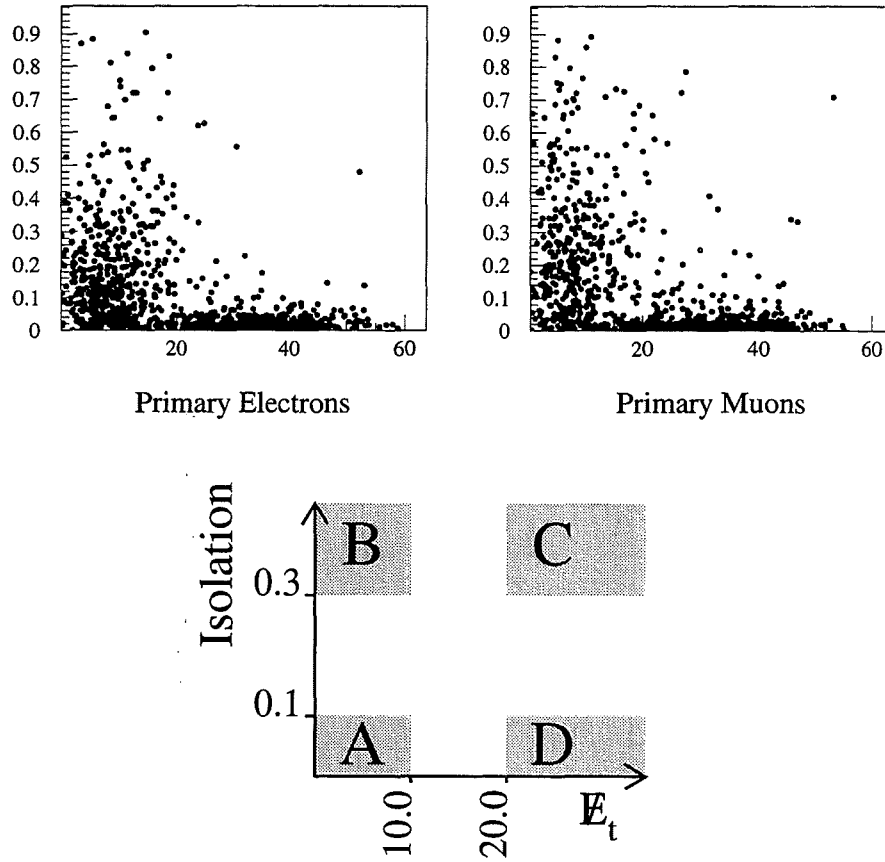


Figure 8-8: The regions in the E_t - Isolation plane used to estimate the non-W component of the W sample. The A region covers $I < 0.1$ and $E_t < 10$, the B region covers $I > 0.3$ and $E_t < 10$ GeV, the C region covers $I > 0.3$ and $E_t > 20$ GeV. D is the signal region and corresponds to $I < 0.1$ and $E_t > 20$ GeV.

GeV and is flat in I . While there is no correlation between E_t and I for the individual distributions, there are large correlations for the sum.)

This method is essentially a poor man's fit of the E_t and I distributions as seen in Figure 8-9 which shows the I distribution for the inclusive sample with $E_t > 20$ GeV overlapped with the distribution for events with $E_t < 10$ GeV normalized to $C/B = 0.15$. Also shown is the analogous plot of E_t for $I < 0.1$ and $I > 0.3$ scaled by $A/B = 1.49$. The quality of the fit in the $E_t < 10$ and $I < 0.3$ regions indicates that to a good approximation, E_t and *Isolation* are uncorrelated for the non-W background.

Table 8.5 lists the measured non-W fraction as a function of jet multiplicity.

Non-W component in Signal Region

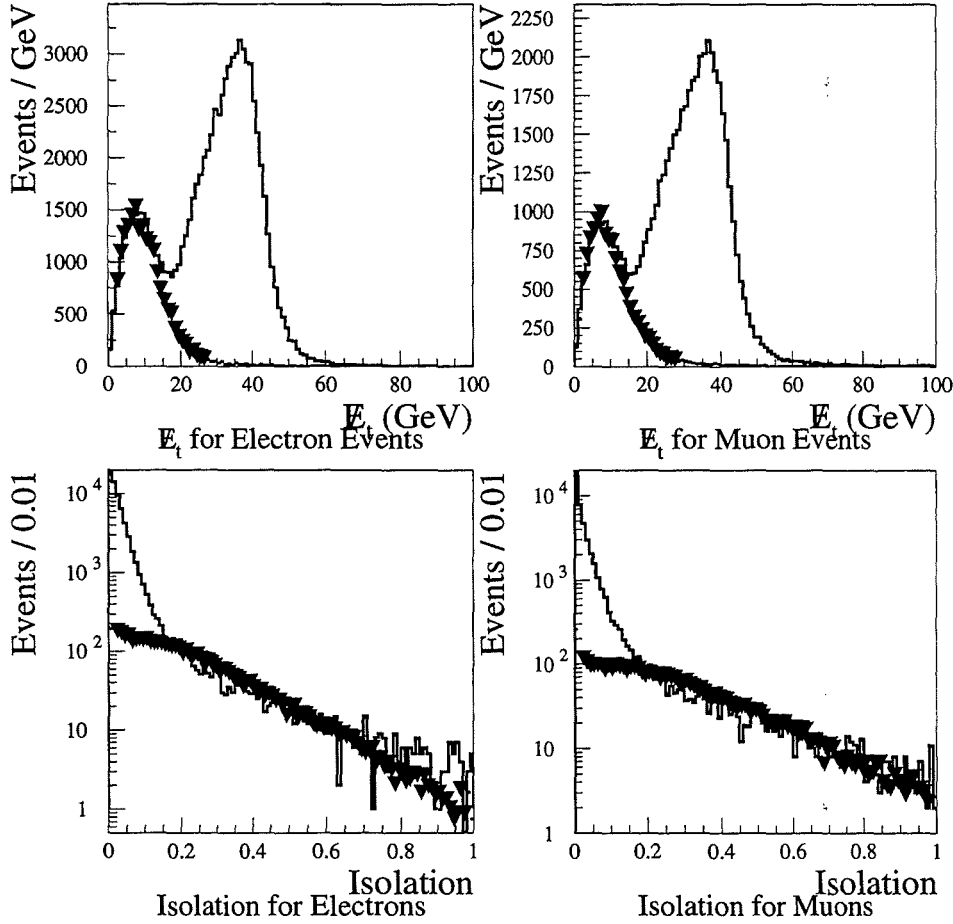


Figure 8-9: The E_t distributions for inclusive leptons with Isolation (I) < 0.1 (solid) compared to leptons with $I > 0.3$ normalized to the $E_t < 10$ GeV region (triangles). Also shown are the (log scale) Isolation distributions for leptons with $E_t > 20$ GeV (solid) compared to leptons with $E_t < 10$ GeV normalized to the $I > 0.3$ region. The agreement of the shapes in the $E_t < 10$ GeV region and in the $I > 0.3$ region indicate that to a good approximation, I and E_t are uncorrelated for the non-W background.

	Electrons	Muons
$W+\geq 0$	$2.5 \pm 0.1 \%$	$2.7 \pm 0.1\%$
$W+1$	$9.7 \pm 0.5 \%$	$8.7 \pm 0.5\%$
$W+2$	$8 \pm 1 \%$	$5.3 \pm 0.5\%$
$W+\geq 3$	$13 \pm 4 \%$	$7 \pm 2\%$

Table 8.5: The non-W fraction for various jet multiplicities using the \cancel{E}_t vs Isolation method described in Section 8.5.1.

Systematic Uncertainties

In general, W 's which “leak” into regions A and C, as well as un removed Z 's and Drell-Yan events will limit the validity of the assumptions. Fortunately only the excess of tags due to $b\bar{b}$ are of concern here, and so this method should provide a reasonable estimate for these purposes. Other analyses have checked this method by estimating the various background components explicitly and show agreement at the 20 % level for inclusive W 's [32]. To cover this, a 50% systematic uncertainty is assigned to the non-W fraction.

8.5.2 $b\bar{b}$ background, First Method

Having calculated the fraction of events in the W sample from non- W sources, the next step is to determine what fraction of the non- W events come from b 's and how often they result in a soft lepton tag ($\epsilon_{b\bar{b}}$). As in the previous section, the low \cancel{E}_t and high *Isolation* portions of the inclusive lepton sample are assumed to be identical to the non- W events in the W sample. The number of tags from $b\bar{b}$ in these events is calculated by looking for an excess of tags over the expectation from the fake rate.

Although it has already been implicitly assumed that the fractional $b\bar{b}$ content in non- W events is independent of \cancel{E}_t and I , one could imagine that the fraction of non- W events from $b\bar{b}$ could be dependent on the *Isolation* of the

primary lepton. Since region A has the same *Isolation* as the signal sample, it is used for the calculation. Events consistent with Drell-Yan production are removed as in Section 7.1.

Table 8.6 shows the number of tags found in lepton+1 jet events in region A, the number predicted by applying the fake rate, and the number expected from Drell-Yan events. Note that the Drell-Yan contribution is significant since unlike the W sample, region A has been selected by requiring $E_t < 10$ GeV. Figure 8-10 shows the P_t distribution for soft lepton tags in region A. The number of tags with high P_t is significantly reduced after subtracting the expected Drell-Yan background, indicating that the Drell-Yan background calculation is reasonable.

Ideally these results could be broken up into the lepton+1,2,3 and ≥ 4 jet samples to check for a variation with jet multiplicity, but unfortunately statistics are limited and so the results from the inclusive lepton+ ≥ 1 jet sample are taken to be valid for events with larger jet multiplicity. Figure 8-10 shows the P_t distribution of tags after backgrounds from fakes and Drell-Yan events have been subtracted. There are nearly twice as many tags as expected from fakes and Drell-Yan events. For primary electron (muon) events an excess of 97 ± 6 (73 ± 5) tags are observed in 5570 (3032) events. Thus 1.7 ± 0.2 % (24 ± 0.4 %) of the non-W events with a primary electron (muon) are expected to contain an additional soft lepton. This is the quantity $\epsilon_{b\bar{b}}$. Note that this method covers, quite generally, all possible enhancements of soft lepton tags in the non-W part of the sample, and does not explicitly assume that it is entirely due to $b\bar{b}$. The quoted errors include statistics and the systematic uncertainties on the fakes and Drell-Yan backgrounds.

8.5.3 $b\bar{b}$ Background, Second Method

An alternate approach to estimating $\epsilon_{b\bar{b}}$ identifies b's directly by searching in the low E_t sample for an SLT near the primary lepton from the sequential decay ($b \rightarrow \ell_{primary} \nu c \rightarrow \ell_{SLT} X$). A $b\bar{b}$ Monte Carlo sample where one b produces

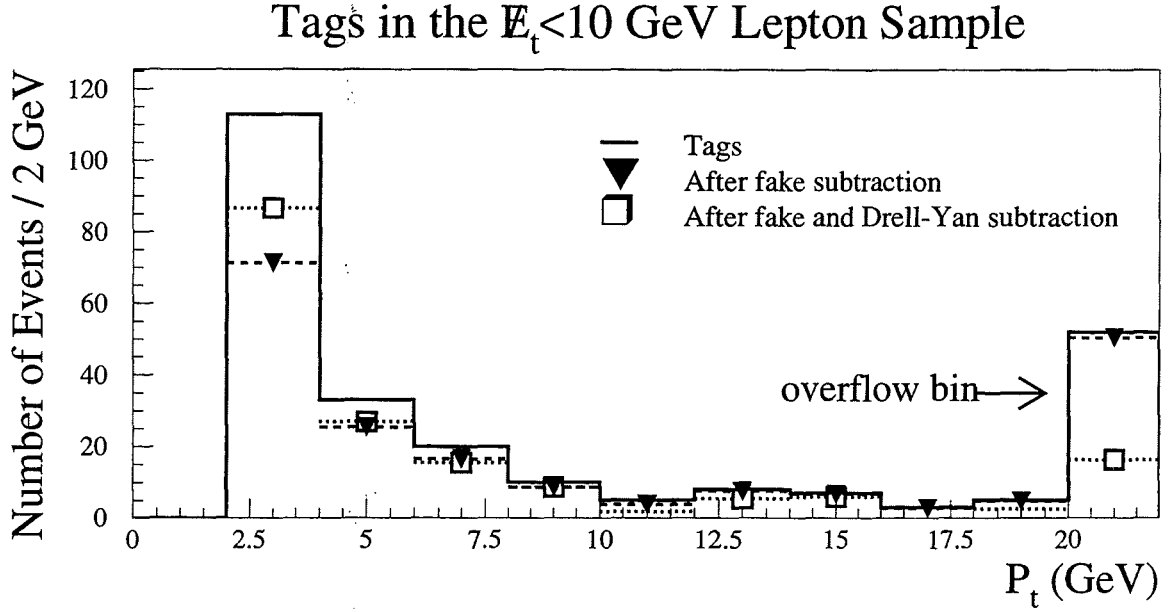


Figure 8-10: The P_t distribution (histogram) for soft lepton tags in events with $E_t < 10$ GeV and $I < 0.1$ and at least one jet with $E_t > 15$ GeV. Also shown is the distribution after fakes (triangles) and the over- and under- efficiency of the Drell-Yan removal (squares) have been accounted for. The excess is assumed to come from $b\bar{b}$ events.

Primary lepton	Number of events	Number of tags	Number of DY remaining	fake prediction	excess tag rate
Electrons	5570	169	35 ± 6	38 ± 4	$1.7 \pm 0.2 \%$
Muons	3032	87	-6 ± 3	20 ± 2	$2.4 \pm 0.4 \%$

Table 8.6: From inclusive leptons in the the $E_t < 10$ GeV, $I < 0.1$ region - the number of events, number of soft lepton tags , number of expected residual Drell-Yan events (as calculated in Section 8.4), the expected number of fake tags, and the final tag rate after the backgrounds have been subtracted. The uncertainties include statistics and systematics on the background subtraction. The numbers are used to calculate the excess tag rate (due to $b\bar{b}$) in the non-W part of the signal sample.

a primary lepton is then used to find the probability for finding a SLT nearby. From this, the total number of $b\bar{b}$ events in the low E_t sample can be estimated. The same Monte Carlo sample can then be used to estimate how often the other b in these events should produce an SLT.

Sequential decays are searched for in the inclusive lepton sample with $E_t < 10$ GeV by selecting events which have a primary electron and a soft muon or a primary muon and a soft electron. Lepton pairs from sequential decay will have opposite signs, but should be uncorrelated for other processes. Figure 8-11 shows the invariant mass distribution (opposite sign minus same sign) of the primary and SLT in these events. A clear excess of 197 (99) events is observed from primary- e SLT- μ (primary- μ SLT- e) pairs in the mass region below 5 GeV/ c^2 , where b 's are expected. Same flavor pairs have not been considered because of the backgrounds from Drell-Yan and other sources, and because of the difficulty of finding a soft electron near a primary electron.

The remaining steps use a HERWIG generated $b\bar{b}$ sample. The Monte Carlo sample is selected by requiring a primary electron with $E_t > 15$ GeV. The soft lepton tagger (discussed in Section 6), is run on the Monte Carlo events and finds a sequential SLT- μ in 73 out of 3084 of the events, or 2.4 ± 0.3 (stat) % of the time. At the expense of statistics, the electron E_t cut can be raised to $E_t > 20$ GeV (as is the case for the inclusive sample) which leaves 105 events containing 2.5 sequential muons¹ which gives an efficiency for finding a sequential muon of 2.4 ± 1.4 (stat) %. Dividing the number of sequential decays observed in the data by this number should give the number of $b\bar{b}$ in the entire sample.

The final number needed is the probability to find an additional SLT from the other b in the event. Using the same $b\bar{b}$ Monte Carlo, events are selected with a primary electron with $E_t > 20$ GeV and a jet with $E_t > 15$ GeV. (This is similar to the region A, $W + \geq 1$ jet study above.) In 538 events, 36.5 soft

¹In the Monte Carlo simulation, soft leptons are weighted by their tagging efficiency, so the number of tags is not an integer.

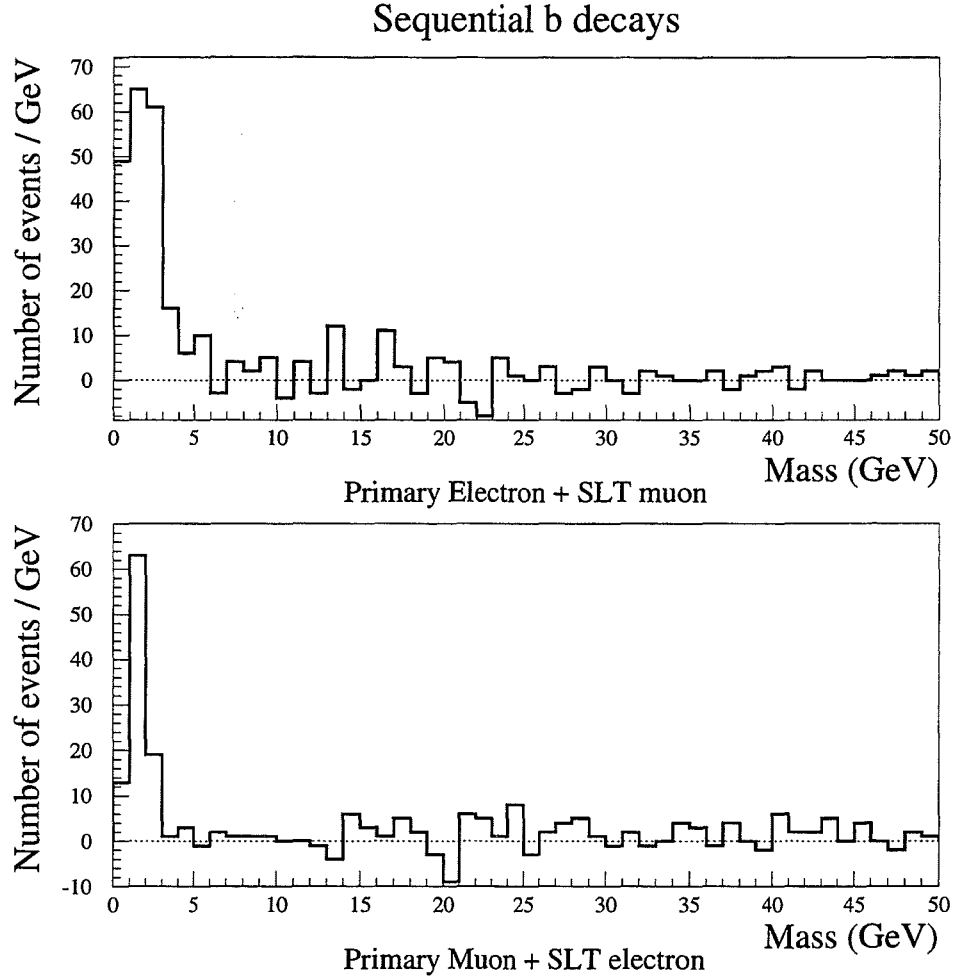


Figure 8-11: The invariant mass of primary and soft leptons in events with $E_t < 10$ GeV. The top histogram shows opposite sign primary- e and SLT- μ pairs with the same sign distribution subtracted. The bottom histogram shows opposite sign primary- μ and SLT- e pairs with the same sign distribution subtracted. An excess of events with mass below 5 GeV indicates the presence of b quarks decaying through $b \rightarrow l^{+(-)}c\nu, c \rightarrow l^{-(+)}\nu$.

leptons are found which are well separated from the primary electron, or $7 \pm 1(\text{stat})$ % of the time.

Turning back now to the inclusive lepton sample with $E_t < 10$ GeV, these numbers can be used to estimate the number of tags expected from the second b in $b\bar{b}$ events. First, from the sequential decay excess of 197 primary- e SLT- μ pairs and the 2.4 % efficiency for finding the sequential muon, it can be inferred that there are ~ 8200 $b\bar{b}$ events in the sample. If the other b produces an additional SLT 7 % of the time, this would predict ~ 574 tagged events. There are 32195 in this sample, so this corresponds to an excess fake rate of ~ 1.8 %, in good agreement with the rates of 2-3 % for the first method. This also provides a rudimentary check of the Monte Carlo treatment of the soft lepton tagger.

8.5.4 Calculation of Residual $b\bar{b}$ Background

Because of the uncertainties associated with the $b\bar{b}$ Monte Carlo, and because the first method incorporates more generally any excess tag rate in non-W events ($c\bar{c}$ events for example), the first method is used to calculate the $b\bar{b}$ background and a conservative systematic uncertainty of 50 % assigned. The number of tags in the sample from $b\bar{b}$ production is given by

$$N_{b\bar{b} \text{ tags}}^j = N_{\text{events}}^j \times F_{\text{non-W}}^j \times \epsilon_{b\bar{b}}$$

where $N_{b\bar{b} \text{ tags}}^j$ is the number of W events with j jets, $F_{\text{non-W}}^j$ is the non-W fraction for the W+j jet events, and $\epsilon_{b\bar{b}}$ is the fraction of these events expected to arise from \bar{b} and produce a real soft lepton², taken from Table 8.6. The final results are listed in Table 8.8. The uncertainties include the 50 % systematic on the non-W fraction and the 50% systematic on the $b\bar{b}$ tag rate.

²It should be noted that fake tags in non-W events have already been accounted for as part of the overall fake background since the fake rate was applied to the whole sample without subtracting the non-W component.

Jet Multiplicity	# of events	Real W fraction	F_{Wc}	$(\epsilon_{SLT} Wc \text{ tag } eff)$
W+1 jet	9001	$86 \pm 2 \%$	$4.8 \pm 1.3 \%$	$2.2 \pm 0.4 \%$
W+2 jet	1376	$84 \pm 2 \%$	$7.2 \pm 2.2 \%$	$2.2 \pm 0.4 \%$
W+3 jet	212	$81 \pm 4 \%$	$7.5 \pm 2.3 \%$	$2.2 \pm 0.4 \%$
W+ ≥ 4 jet	61	$80 \pm 4 \%$	$7.5 \pm 2.3 \%$	$2.2 \pm 0.4 \%$

Table 8.7: The various components of the Wc background calculation. The “Real W fraction” accounts for expected contributions from QCD non-W events (measured using the \cancel{E}_t vs I method described in the text) as well as diboson production and Z events where one lepton has not been reconstructed or where the Z produced a single lepton through the decay $Z \rightarrow \tau^+ \tau^-$, $\tau \rightarrow \ell X$.

8.6 Wc background

Since the production of a W boson and a charm quark (through the “excitation” of a strange sea quark, as shown in Figure 3-6) is not accounted for by the fake rate, its contribution must be calculated separately. Unlike the $b\bar{b}$ background, the number of Wc events in the data cannot currently be measured and must be taken from theoretical predictions. To minimize the uncertainty on the overall Wc cross section, the *fraction* of W +jets events from Wc (F_{Wc}) is calculated from Monte Carlo. This number is then convoluted with the number of *real* W +jets events in the data, to calculate the number of Wc events in the data. To get the *real* number of W +jets events one must subtract the hadronic component calculated in 8.5.1 as well as the expected number of events in the sample from diboson production and unremoved Z ’s either where one leg was missed or where the Z decayed to τ ’s one of which decayed to a lepton. The expected number of events from these sources is small (less than half of the hadronic non- W contribution calculated with the I vs \cancel{E}_t method) and has been calculated using Monte Carlo [40]. Table 8.7 lists the expected real W fraction in the data ($F_{real\ W}$) after accounting for the contribution from these processes.

The fraction of W events with a $Wc\bar{s}$ or $W\bar{c}s$ vertex (F_{Wc}) has been calcu-

lated using HERWIG Monte Carlo [39]. HERWIG is used since it includes at some level, higher order diagrams like $gg \rightarrow Wc\bar{s}$. The scaling of this fraction with jet multiplicity has also been compared to VECBOS parton level matrix element results and gives very good agreement [39]. Because many Wc events are produced from a strange quark from the proton sea, the results are sensitive to the structure functions used in the calculation. A 30 % systematic uncertainty is assigned to the Wc fraction based on studies using a variety of structure functions [44]. The final Wc sample used is generated by HERWIG and uses the MRSD0' parton distribution functions. The predicted Wc fraction (F_{Wc}) as a function of jet multiplicity is shown in Table 8.7.

The second piece of information needed is the probability of getting a soft lepton tag from the decay of the charm quark ($\epsilon_{SLT\ Wc}$) which can also be derived from the Monte Carlo. Because of low Monte Carlo statistics, the efficiency for soft lepton tagging a Wc event ($\epsilon_{SLT\ Wc}$) is calculated for all $W+\geq 1$ jet events and assumed to be constant as a function of jet multiplicity. Note that the Monte Carlo includes the trigger simulation (Section 6.2) and so should take into account any possible SLT bias due to requiring the presence of a jet over a threshold. The ($\epsilon_{SLT\ Wc}$) values are summarized in Table 8.7. Putting this all together, the Wc background in the W +jets sample is calculated as

$$N_{Wc\ tags} = N_{W\ candidates}^{[N_{jet}]} \times F_{real\ W} \times F_{Wc} \times \epsilon_{SLT\ Wc}.$$

The results are summarized in Table 8.8.

8.7 $Z \rightarrow \tau^+\tau^-$ background

Since the $Z \rightarrow \tau^+\tau^-$ decay should be straightforward to model, this background is estimated entirely from Monte Carlo. The process of concern is where both tau's decay leptonically, and one is identified as a primary lepton and the other as a (typically high P_t) soft lepton tag. Significant B_t can arise from an imbalance of the four neutrinos produced in the decay, especially when there

are additional jets in the event which can arise from initial state radiation. The ISAJET Monte Carlo is used for this calculation. CDF has previously measured the Z production cross-section $\sigma(p\bar{p} \rightarrow ZX \rightarrow e^+e^-X) = .21 \pm .01$ nb [41]. Combining this with the world average $BR(Z \rightarrow e^+e^-) = 3.366 \pm 0.008$ % [26] yields $\sigma(p\bar{p} \rightarrow ZX \rightarrow e^+e^-X) = 6.2 \pm 0.3$ nb. The ISAJET Monte Carlo program is used to generate $Z \rightarrow \tau^+\tau^-$ events and determine the number of predicted tags and the results are normalized to this cross section and the measured luminosity for the data sample of 90.3 ± 7.2 pb^{-1} .

To quantify the ability of ISAJET to model the jet multiplicity correctly, the jet multiplicity of the W candidates in data (with F_{non-W} subtracted) is compared to the multiplicity spectrum in a W sample produced by ISAJET. After normalizing to the number of observed W+0 jet events, ISAJET predicts roughly 40 % more events in each jet multiplicity bin than is observed in the data. Based on this study a 50 % systematic uncertainty is assigned to the $Z \rightarrow \tau^+\tau^-$ background estimate. The results are summarized in Table 8.8.

8.8 Diboson background

Diboson production (WW or WZ) where a W decays leptonically producing a high P_t lepton with large \cancel{E}_t and the other W or Z decays either to leptons or to heavy flavor pose a potentially serious source of background soft lepton tags in the W sample. There is no published measurement of Diboson production, and so theoretical predictions and Monte Carlo must be used for this background. The latest calculations of the WW, and WZ production cross section [43] predict 9.9 pb, and 2.8 pb respectively. Although the WW cross section is comparable to the theoretical $t\bar{t}$ cross section for $M_{top} \approx 150$ GeV, the number of expected events is small after taking into account the appropriate branching fractions. ZZ production is expected to have an even smaller cross section of ~ 1 pb and so is ignored. ISAJET and QFL are used to generate and simulate 10k WW and 28k WZ events. which are normalized to 90.1 ± 7.2 pb^{-1} , the

BACKGROUND	W+1	W+2	W+3	W+ \geq 4
(fake+W $b\bar{b}$ +W $c\bar{c}$)	202 ± 18	57 ± 5	13.2 ± 1.2	6.1 ± 0.6
Drell-Yan	-4.9 ± 3.3	1.6 ± 2.2	-0.6 ± 1.4	0.0 ± 0.6
$b\bar{b}$	16 ± 8	2 ± 1	0.4 ± 0.2	0.12 ± 0.06
Wc	8.2 ± 2.3	1.8 ± 0.5	0.3 ± 0.1	0.08 ± 0.02
$Z \rightarrow \tau^+ \tau^-$	3.3 ± 1.4	1.5 ± 1.1	0.3 ± 0.1	0.1 ± 0.05
WW, WZ	3.3 ± 1.2	1.2 ± 0.4	0.4 ± 0.2	0.23 ± 0.08
TOTAL BACKGROUND	228 ± 20	65.1 ± 5.7	14.0 ± 1.8	6.6 ± 0.8
OBSERVED TAGS	201	66	23	15

Table 8.8: Summary of the expected background and observed tags in the W sample. There is good agreement in the W+1 and W+2 jet bins and an excess in the W+ \geq 3 jet bins as expected from $t\bar{t}$ production.

integrated luminosity for the data sample. As with the $Z \rightarrow \tau^+ \tau^-$ background, a 50 % systematic uncertainty is assigned to the ability of ISAJET to model the jet multiplicity spectrum. An additional 30 % uncertainty on the theoretical cross section is added which is approximately the difference between the LO and NLO calculations. Finally there is a 30 % uncertainty due to limited Monte Carlo statistics. The final results are summarized in Table 8.8.

8.9 Summary of the backgrounds

Table 8.8 shows good agreement between the background predictions and observed tags in the W+1 and W+2 jet bins and an excess of tags over the expected background in the signal region. The results are displayed graphically in Figure 8-12.

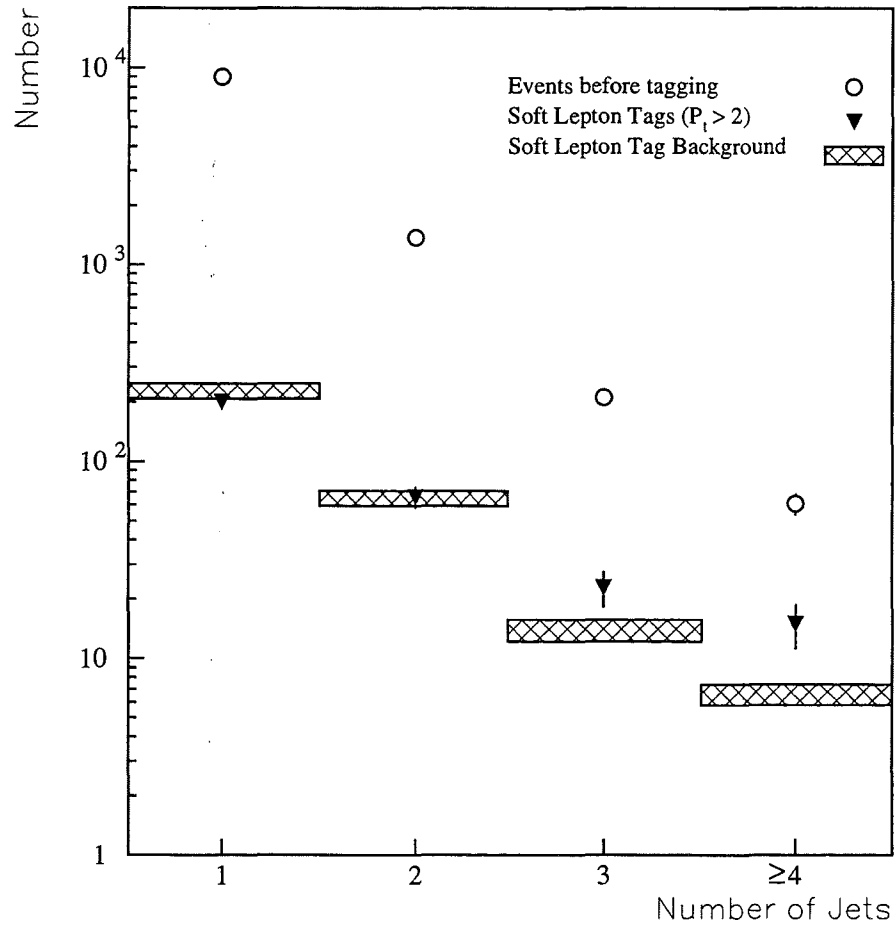


Figure 8-12: The number of events, observed tags and predicted background tags in the W +jets sample from Run 1b. Boxes show the estimated uncertainty on the background prediction. There is good agreement in the W +1 and W +2 jet bins, and an excess in the W + ≥ 3 jet bins as expected from $t\bar{t}$ production.

Chapter 9

Analysis of the Results

9.1 Significance of the Excess of Tags

The statistical significance of the excess of events in the signal region of $W+\geq 3$ jets (Table 8.8) can be quantified as the probability (\mathcal{P}) of observing the 38 tags when 20.6 ± 2.6 are expected. In general, for a mean background of N_{bkg} gaussian distributed with a width σ_{bkg} , and with N_{obs} observed, \mathcal{P} is given by

$$\mathcal{P} = 1 - \sum_{x=0}^{N_{obs}} \int_{-\infty}^{+\infty} G(N_{bkg}, \sigma_{bkg}; m) \times P(m; x) dm$$

where $G(N_{bkg}, \sigma_{bkg}; m)$ is a Gaussian distribution with mean N_{bkg} and sigma σ_{bkg} evaluated at m and $P(m; x)$ is a Poisson distribution with mean m evaluated at x . The gaussian gives the probability to have an actual background mean of m . The poisson term gives the probability of m fluctuating to give x background tags.

In practice \mathcal{P} is evaluated with a Monte Carlo procedure which generates a number of *pseudo-experiments* each with a background mean of m chosen from the gaussian. For each event the probability of observing N_{obs} tags given m is calculated as $1 - \sum_{x=0}^{N_{obs}} P(m; x)$. Iterating the procedure 10,000 times yields a value of $\mathcal{P} = 0.2 \%$.

	W+1 jet	W+2	W+3	W+ ≥ 4
RUN 1A				
Observed tags	44	16	4	2
Total background	44.2 ± 4.1	13.1 ± 1.2	2.9 ± 0.3	1.1 ± 0.1
RUN 1A+1B				
Observed tags	245	82	27	17
Total background	272 ± 24	78.2 ± 6.9	16.9 ± 2.1	7.7 ± 0.9

Table 9.1: Results of the SLT analysis for Run 1a combined with Run 1b representing $109.4 \pm 7.9 \text{ pb}^{-1}$ of data.

9.2 Calculation of the $t\bar{t}$ Production Cross Section

The previous chapters have provided all the ingredients necessary to calculate the $t\bar{t}$ cross section which can be written as

$$\sigma_{t\bar{t}} = \frac{N - B}{A_{t\bar{t}} \mathcal{L}}$$

where N is the number of observed tags (38), B is the expected background (20.6 ± 2.6), and \mathcal{L} is the integrated luminosity ($\mathcal{L} = \int L dt = 90.1 \pm 7.2 \text{ pb}^{-1}$) for the dataset and $A_{t\bar{t}}$ is the acceptance. The acceptance ($A_{t\bar{t}} = 2.10 \pm 0.36\%$) is evaluated at the central value of the CDF measured top mass $M_{top} = 175 \text{ GeV}/c^2$ (See Chapter 6). The statistical uncertainty on the 38 tags is taken to be ± 6.2 tags which are the $\pm 1 \sigma$ points for a Poisson distribution with a mean of 38. The $t\bar{t}$ production cross section is determined to be

$$\sigma_{t\bar{t}} = 9.2 \pm 3.9 \text{ pb}.$$

Half of the uncertainty comes from the limited statistics on the observed number of tags (N) which contributes 1.8 pb to the final uncertainty.

9.2.1 Combined Results from Run 1a and Run 1b Data

The statistical uncertainty can be reduced by including the results of the SLT analysis performed on the Run 1a dataset of $19.3 \pm 0.7 \text{ pb}^{-1}$ which are described in detail elsewhere [12] [80]. The SLT algorithm during Run 1a did not include soft muons in the CMX but the analysis has been repeated with the CMX implemented as described in this thesis [81]. The results of the Run 1a and Run 1b analyses, representing $109.4 \pm 7.9 \text{ pb}^{-1}$ are presented in Table 9.1. The acceptance for Run 1a is identical to Run 1b with the exception of the trigger efficiency which is measured to be $\epsilon_{trig}^{1A} = 0.83 \pm 0.04\%$. The combined Run 1a and Run 1b acceptance then, is 2.09 ± 0.29 . With a total of 44 tags on a background of 24.6 ± 3.0 , the measured cross section becomes

$$\sigma_{t\bar{t}} = 8.5 \pm 3.4 \text{ pb}.$$

The probability of 24.6 ± 3.0 fluctuating up to give the observed 44 tags is 0.18 %.

9.2.2 Comparison with other Analyses

The $t\bar{t}$ production cross section has also been measured using the Run 1a and Run 1b data from an excess of events in the dilepton channel and in the lepton plus jets channel using an SVX b-tag. In the dilepton channel 9 events are observed with an expected background of 2.1 ± 0.4 events [31]. In the SVX analysis 42 tags are observed in 34 events with an expected background of 8.0 ± 1.7 events. The measured cross sections are all in good agreement and are listed in Table 9.2.

	$\sigma_{t\bar{t}}$
Dilepton analysis	$8.2^{+3.8}_{-3.3}$ pb
SVX analysis (ℓ + jets)	$6.8^{+2.3}_{-1.8}$ pb
SLT analysis (ℓ + jets)	$8.5^{+3.4}_{-3.4}$ pb

Table 9.2: The $t\bar{t}$ production cross section measured in the dilepton channel and in the ℓ + jets channel using an SVX b-tag. Within the quoted uncertainties the measurements are in good agreement.

9.3 Properties of SLT Tags in the W Sample

9.3.1 Stability of the Excess of Tags

If the observed excess of events is truly due to $t\bar{t}$ production it should be robust, and insensitive to small changes in the SLT algorithm. Restricting the analysis to soft electron or muon tags or raising the minimum P_t for SLT tags tests both the reliability of the background estimate in the W+1 and W+2 jet bins and the stability of the excess. Table 9.3 shows the tags and backgrounds broken down by soft lepton type and for SLT's with $P_t > 4$ GeV/c. In each case there is an excess of tags in the signal region, and good agreement in the W+1 and W+2 jet bins. A study of HERWIG $t\bar{t}$ events indicates that roughly 3/4 of the excess should come from SLT muon tags, and that 80 % of all tags should have $P_t > 4$ GeV/c. Within the limited statistics available the data is consistent with these predictions.

9.3.2 Double Tagged Events

A $t\bar{t}$ event contains two b-quarks and hence can contain multiple (real) SLT tags. If the observed excess of events is from $t\bar{t}$ production, some should be double tagged.

	W+1	W+2	W+3	W+ \geq 4
Soft electrons only				
Total background	53.7 ± 10	16.7 ± 3.3	3.9 ± 1.2	1.6 ± 0.8
Observed Tags	52	16	7	4
Soft muons only				
Total background	174.6 ± 17	48.3 ± 4.6	10.1 ± 1.4	5.0 ± 0.6
Observed tags	149	50	16	11
$P_t > 4$ GeV/c				
Total background	101 ± 13	32.8 ± 4.2	7.02 ± 1.4	3.5 ± 0.6
Observed tags	84	33	15	8

Table 9.3: The number of observed tags compared to the expectations from background for soft electrons, soft muons and SLT's with the minimum momentum raised to $P_t > 4$ GeV/c. In each case there is good agreement in the W+1 and W+2 jet bins and an excess in the signal region as expected from $t\bar{t}$ production.

Double Tags Expected from $t\bar{t}$

Double tags in $t\bar{t}$ events fall into three categories, *real+real* (both tags are real) *real+fake* (one tag is real, one fake) and *fake+fake* (both tags are fakes). In a HERWIG generated $t\bar{t}$ sample ($M_{top}=175$ GeV/c²) 10 % of SLT tagged top events in the $W+\geq 3$ jet bin contain a second SLT tag, so the number of *real+real* tagged events expected in the data can be estimated by multiplying the observed excess of SLT tagged events by 10%, yielding 0.76 (0.72) in the $W+3$ ($W+\geq 4$) jet bin. In the $t\bar{t}$ sample, the average fake rate per event, obtained by applying the fake rate to the simulated tracks is ~ 10 %. The number of *real+fake* tags is estimated by multiplying this by the observed excess of SLT tagged events and contributes an extra expected 0.76 (0.72) events to the $W+3$ ($W+\geq 4$) jet bin.

Double Tags from Background

Since fake tags constitute the majority of the SLT background, most double tags should arise from two fake leptons appearing in a single event. The probability for an event to produce more than one fake tag can be estimated as

$$\mathcal{P}^{fake+fake} = \left(\sum_i p_{fake}^i \right) - \left(1 - \prod_j (1 - p_{fake}^j) \right)$$

where the sum and products are over all tracks in the event and p_{fake}^i is the probability for track i to fake an SLT (as predicted by the fake rate). The first term is the total number of fakes predicted in the event, the second term is the probability of tagging the event at least once. Assuming there are no events that are triply tagged, summing $\mathcal{P}^{fake+fake}$ over all events yields the expected number of double tags. By construction, the fake+fake contribution from $t\bar{t}$ events in the sample is taken into account in this sum.

Comparison of Double Tags with Expectations

Table 9.4 compares the number of observed double tagged events with the predictions from the sources listed above. Given the low statistics of this

	W+1	W+2	W+3	W+ ≥ 4
Fake+Fake events	3.96	1.84	0.62	0.32
Real+Fake ($t\bar{t}$)	0	0	0.76	0.72
Real+Real ($t\bar{t}$)	0	0	0.76	0.72
Total background	3.96	1.84	2.14	1.76
Observed double tags	4	1	2	2

Table 9.4: Comparison, using Run 1a and Run 1b data, of the number of observed double tagged events with predictions from background and top.

comparison, the agreement is excellent.

9.3.3 Events Tagged by the SVX and the SLT

The SVX¹ and the SLT taggers provide two independent methods of identifying b quarks. The number of events tagged by both the SVX and SLT offer another handle for checking the consistency of the observed excess with the hypothesis of $t\bar{t}$ production. The W+ ≥ 3 jet sample contains 10 events tagged by both the SLT and the SVX. Figure 9-1 shows a top candidate event with one jet tagged by both the SVX and SLT. The SVX analysis has an excess of roughly 26 tagged events. With an SLT tag efficiency of 20 % and a fake rate of ~ 10 % in $t\bar{t}$ events, one would expect ≈ 8 $t\bar{t}$ events tagged by both the SLT and SVX, which agrees well with the observed 10 events. Less than one background event is expected to be tagged by both the SVX and SLT.

9.3.4 Consistency with Jet Multiplicity Expected from Top

The jet multiplicity of the excess can also be compared to expectations from simulated top events. (See Figure 3-2.) For a top mass of $M_{top} = 175 \text{ GeV}/c^2$

¹A brief description of the SVX tagger is given in Appendix A.1.

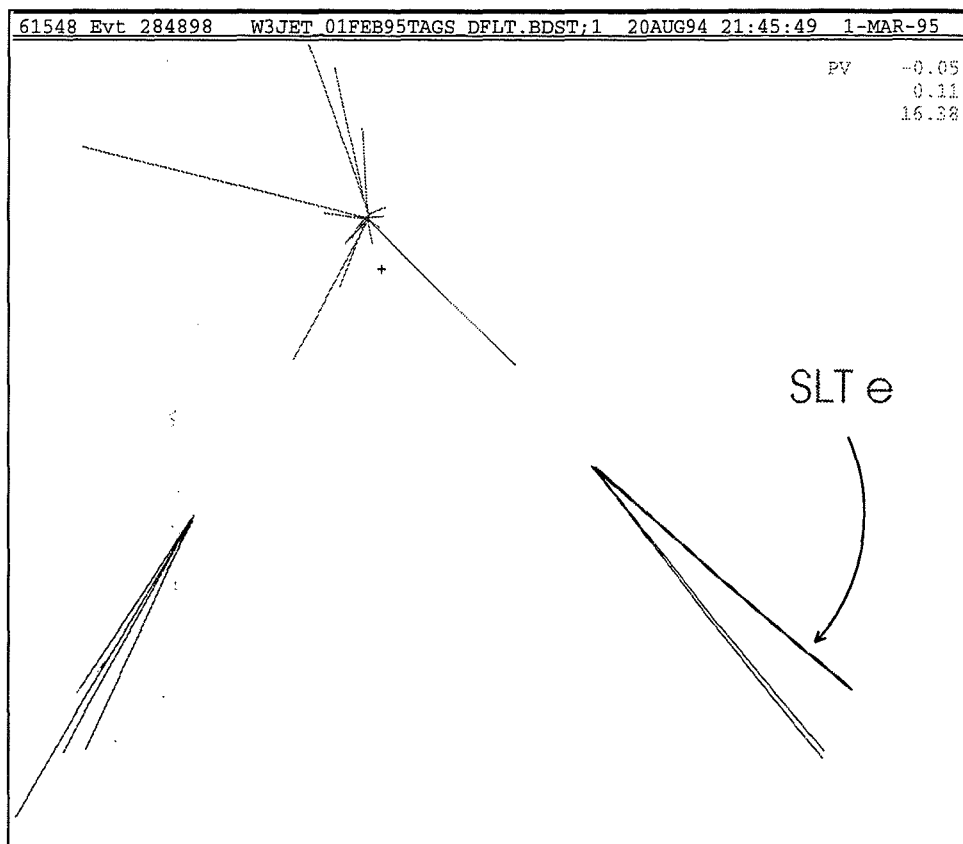


Figure 9-1: A transverse view of SVX tracks in a top candidate event with two jets tagged by the SVX, one of which contains a soft electron tag. The “+” denotes the origin ($x, y, z = 0$) and the cluster of tracks above it the primary vertex which is displaced by 0.05 cm in \hat{x} and .11 cm in \hat{y} . The other two clusters of tracks come from vertices which are displaced from the primary by ~ 6 mm and are tagged by the SVX b-tagger. The box shows an area 1.33 cm in \hat{y} by 3.86 in \hat{x} .

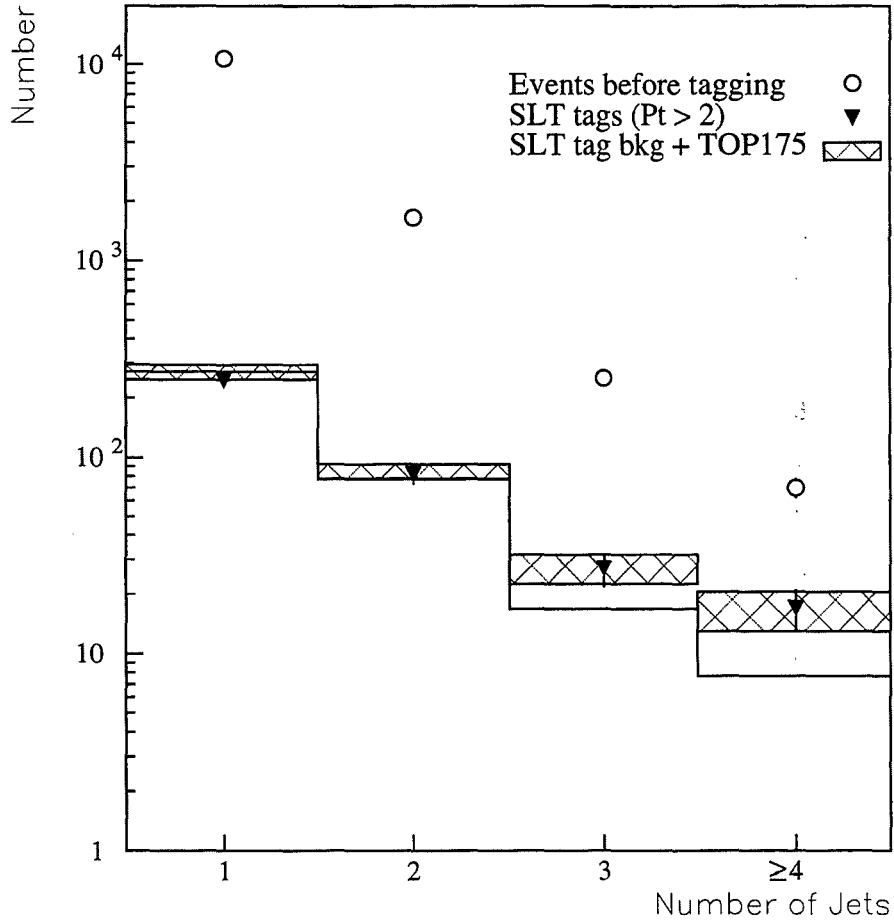


Figure 9-2: The number of events, observed tags and predicted tags from background and $t\bar{t}$ production in the W+jets sample of Run 1a and Run 1b combined. A top mass of $M_{top} = 175 \text{ GeV}/c^2$ and cross section of $\sigma_{t\bar{t}} = 8.5 \pm 3.4 \text{ pb}$ is assumed. The data fits well to a combination of $t\bar{t}$ and background.

the excess of events is expected to be evenly split between the W+3 and W+ ≥ 4 jet bins. A number of $t\bar{t}$ events equal to 1/3 of the excess are also expected in the W+2 jet bin. Figure 9-2 compares the expectations with the data, using the SLT measured cross section $\sigma_{t\bar{t}} = 8.5 \pm 3.4 \text{ pb}$. Within uncertainties, the data is well described by a combination of tags from background and $t\bar{t}$ production.

9.4 Determination of the Top Mass

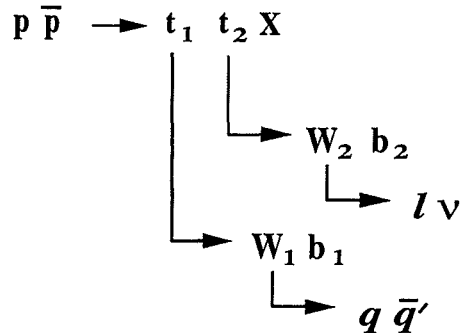
In the lepton plus jets channel, the top mass can be directly determined from the experimentally measured quantities. The details of the measurement are quite involved and beyond the scope of this thesis, but for completeness, the procedure is summarized here.

The events used to measure the top mass are a subset of the $W+\geq 3$ jet sample described in this analysis. At least four jets are required in the event so that a one-to-one match between the partons and jets can be made. To increase the acceptance the selection criteria for the fourth jet are loosened to $E_t > 8$ GeV and $|\eta| < 2.4$. Of the 325 $W+\geq 3$ jet events (Run 1a and Run 1b combined) 163 events have such a fourth jet. To reduce the number of combinatoric possibilities and improve the purity of the sample, at least one jet is required to be b-tagged by SECVTX or SLT, where the SLT is required to be within $\Delta R = 0.4$ of a jet centroid.

Top quark pairs are produced in the reaction

$$p\bar{p} \rightarrow t\bar{t} + X$$

where X represents the system recoiling against the $t\bar{t}$ pair. The events are reconstructed assuming the Standard Model decay



which is composed of five processes or *vertices*

- (1) $p\bar{p} \rightarrow t_1 + t_2 + X$
- (2) $t_1 \rightarrow b_1 + W_1$
- (3) $t_2 \rightarrow b_2 + W_2$
- (4) $W_1 \rightarrow q + \bar{q}'$
- (5) $W_2 \rightarrow \ell + \nu$

where q and \bar{q}' are light quarks. The measured energies and directions of the four jets and the primary lepton are used to infer the 4-vectors of the quarks and the lepton. The total transverse momentum of the $p\bar{p}$ system is assumed to be zero, which allows the transverse component of X to be measured from the unclustered energy and jets not used in the fit. The neutrino momentum is left as an unknown. The W mass is taken to be $M_W = 80.1 \text{ GeV}/c^2$ and the two top quark masses are assumed to be the same, leaving 18 unknown quantities. At each vertex, energy and momentum must be conserved which gives 20 equations of constraint allowing the top mass to be determined. Because the longitudinal momentum of the $p\bar{p}$ system is not measured, there are two possible solutions for the p_z of the neutrino and so two possible top masses for each fit.

Each jet is assigned to a quark, and the event is reconstructed with a program originally designed for bubble chamber experiments called SQUAW [90]. SQUAW determines a top mass corresponding to a best fit for the event and returns a chi-square (χ^2) which reflects how far the input quantities had to be adjusted within their measured uncertainties. The jets assigned to the W decay, for instance, are required to reconstruct to the W mass within the expected uncertainty from the jet energy resolution.

A fit is performed for each possible jet-quark assignment. For each assignment there are two possible solutions corresponding to the two-fold ambiguity of the \hat{z} component of the neutrino momentum. With no b-tagging information there are 12 ways to assign the jets to quarks and hence 24 solutions. If one of the four jets is b-tagged jet, there are 6 possible assignments and so 12 solutions. With two b-tagged jets there are 4 possible solutions. For each event, the solution with the lowest χ^2 is chosen. Events which have $\chi^2 > 10$

$M = 140, 160, 175, 190 \text{ GeV}$

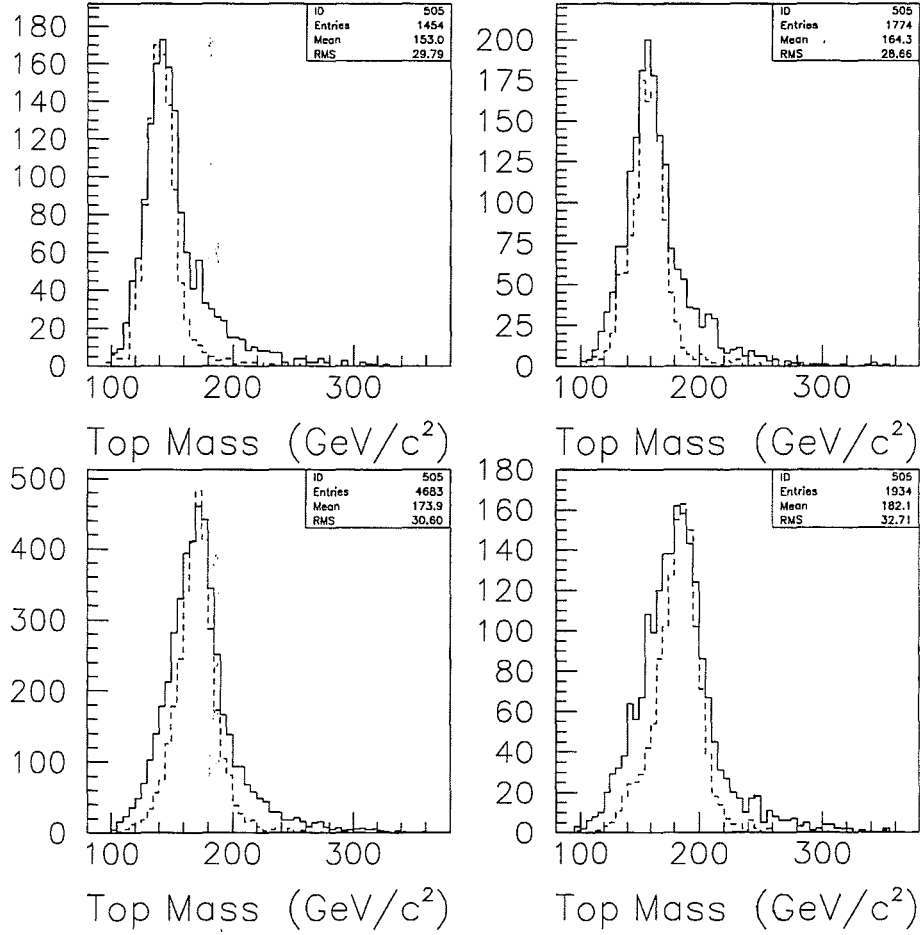


Figure 9-3: Reconstructed top mass from HERWIG generated $t\bar{t}$ samples where one jet has been b-tagged. Shown are distributions for $M_{top}=140,160,175,190 \text{ GeV}/c^2$ (clockwise). The full histogram shows the best fit obtained by the fitter. The dashed histogram shows the shape of reconstructed events with the correct jet-quark assignment.

do not fit well to the top hypothesis and are removed from consideration. This requirement is expected to be $\sim 88\%$ efficient for $t\bar{t}$ events and $\sim 73\%$ efficient for W +multijet events which pass the four jet requirement. In the data, 153 events satisfy this requirement of which 34 have an SVX or SLT b-tag.

To improve the measurement of the top mass, the jet energies used by the fitter are corrected for a variety of effects. The corrections account for reduced calorimeter response near detector boundaries, detector non-linearities, jet energy which has fallen outside of the cluster cone, multiple interactions, energy from the underlying event, and undetected energy carried by muons or neutrinos [85]. These are the standard CDF jet corrections and typically increase jet

energies by $\sim 30\%$. A second set of corrections derived from $t\bar{t}$ Monte Carlo events are also applied which make specific corrections for light quark jets as well as b-jets tagged by the SVX or SLT, and which account for the expected energy sharing between jets.

The fitting procedure is tested on top Monte Carlo events generated with HERWIG, which are passed through the full simulation and analysis chain including jet corrections. The reconstructed top mass spectrum for events generated with top masses of 140, 160, 175 and 190 GeV/c² are shown in Figure 9-3. The fitter chooses the correct jet-quark assignments in for all four jets in 31% of the events. In 47% of the events, at least one of the four jets was not associated with a parton from the $t\bar{t}$ decay. In the remaining 22% of the events the jets have been assigned to the wrong quarks. Figure 9-3 shows the reconstructed mass for the subset of events which have the correct jet-quark assignment. The reconstructed masses for the 34 b-tagged events are shown in Figure 9-5.

From these events, the top mass is then estimated with a maximum likelihood technique which fits the reconstructed masses to a sum of the expected distributions from $t\bar{t}$ for a particular M_{top} and W+multijets background. The shape of the background is determined by passing VECBOS generated W+4 jet events with HERWIG fragmentation through the mass fitter and is shown in Figure 9-4. Although there are non-W sources of background in the sample these are small, and have similar shapes to the W+4 jet background. The VECBOS generated background shape has been compared to Z+jets data, and shows good agreement. The effect of varying the background shape on the final fitted mass is taken as a systematic.

The number of background events in the sample is calculated by extrapolating from the $W+\geq 3$ jet backgrounds to be $6.4^{+2.0}_{-1.5}$ using a procedure which accounts for the possibility of fake tags in top events [91]. The likelihood function, L , used for determining the mass is

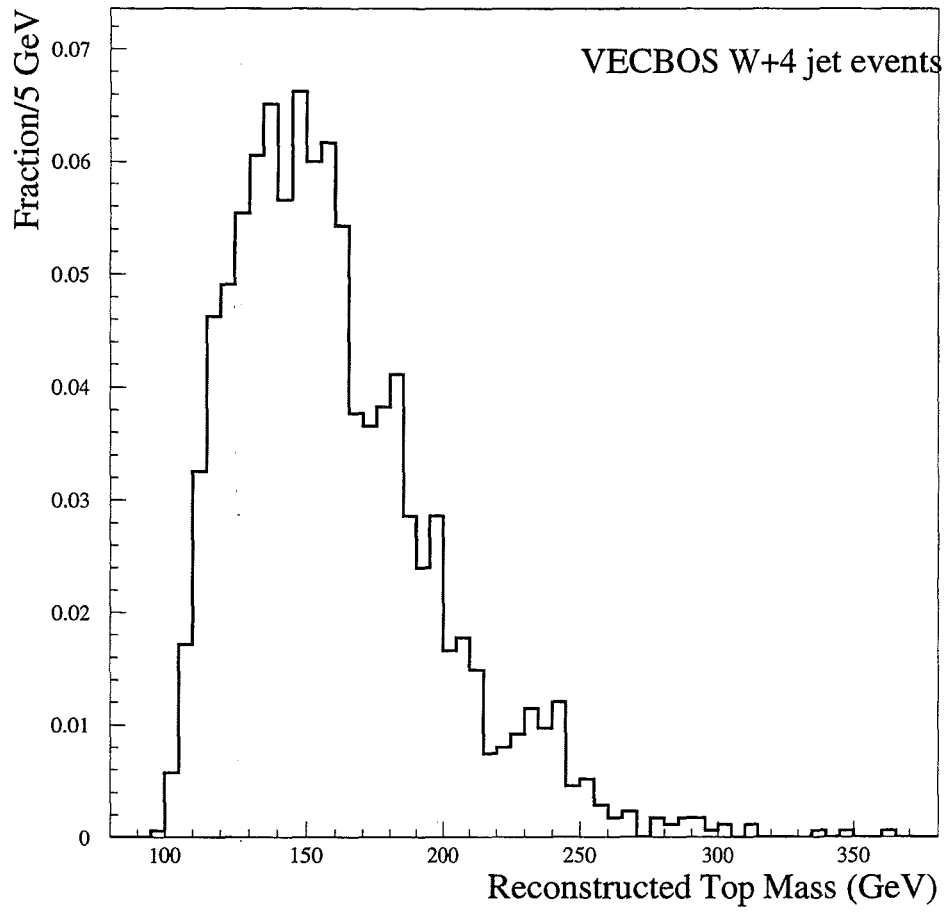


Figure 9-4: The background template used in the top mass likelihood fit. The distribution is obtained by generating W+4 jet events with VECBOS, implemented with HERWIG fragmentation, and passing them through the SQUAW program which tries to reconstruct a top mass.

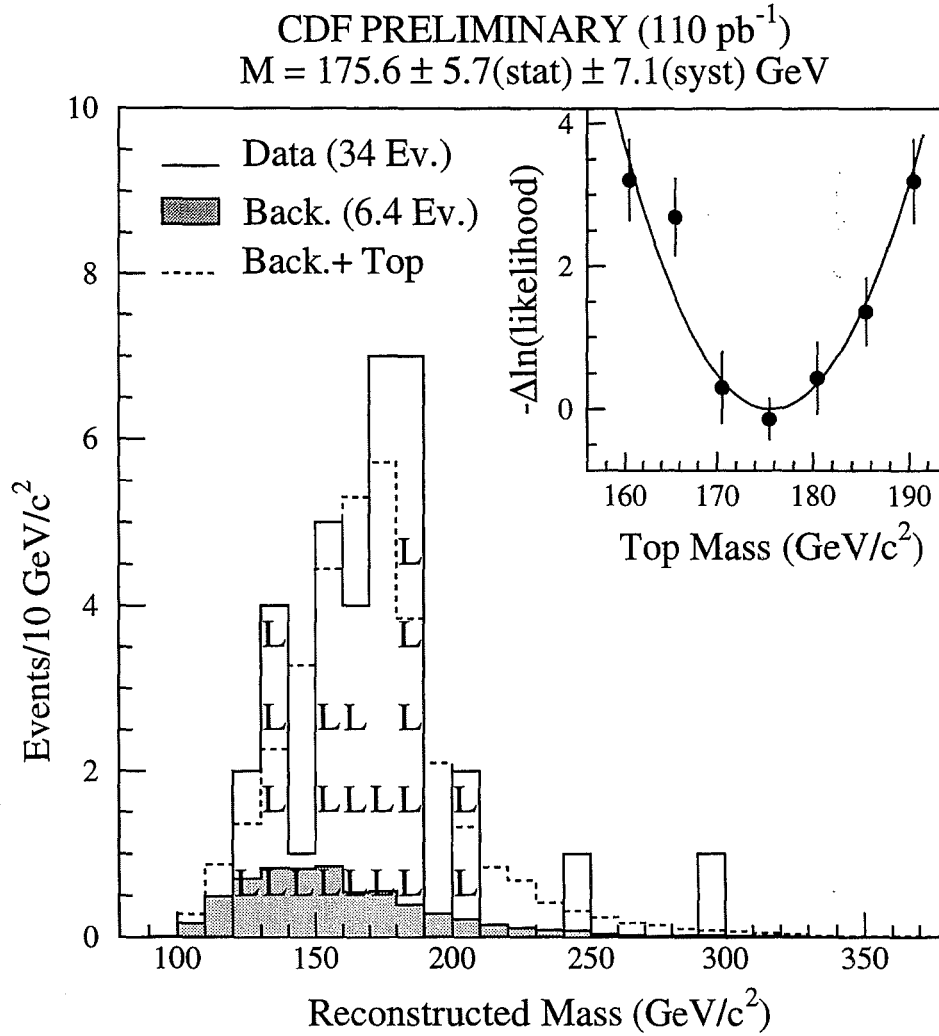


Figure 9-5: The reconstructed top mass for SVX and SLT tagged events with at least 4 jets (solid histogram). The inlay shows the change in the fit likelihood as a function of top mass. The dashed histogram shows the expected distribution from top ($M_{top} = 175 \text{ GeV}/c^2$) and background with their relative contributions given by the fit. The shaded histogram shows the background shape normalized to 6.4 events.

$$L = G(N_b, \sigma_b; n_b) \times P(n_t + n_b; N_{obs}) \times \prod_{i=1}^{N_{obs}} \frac{n_b f_b(m_i) + n_s f_s(m_i, M_t)}{(n_b + n_t)}$$

where

- n_b (n_s) are the number of background ($t\bar{t}$) events in the fit
- N_b is the expected number of background events (6.4)
- σ_b is the uncertainty on the expected background ($^{+2.0}_{-1.5}$)
- N_{obs} is the number of observed events (34)
- f_b is the normalized reconstructed mass distribution for background events
- $f_s(M_t)$ is normalized reconstructed mass distribution for top events with mass M_t .
- m_i is the reconstructed mass for event i of the data sample.
- $G(N_b, \sigma_b; n_b)$ is a Gaussian distribution with mean N_b and sigma σ_b evaluated at n_b which gives the probability of observing n_b background events when $N_b \pm \sigma_b$ are predicted.
- $P(N_{obs}; n_t + n_b)$ is a Poisson distribution with a mean of $n_t + n_b$ evaluated at N_{obs} which gives the probability of finding N_{obs} events given that $n_t + n_b$ are predicted.

For each top mass (M_t) the likelihood function is maximized which determines a value of n_b and n_s . The inlay on Figure 9-5 shows the result of the likelihood fit as a function of top mass. The error bars on each point reflect the uncertainty due to limited statistics in the Monte Carlo generated top and background templates. The points are fit to a smooth curve. The curve has a minimum at $M_{top}=175.6 \pm 5.7$ (stat) where the uncertainty on the fit is determined by varying the log likelihood by $\pm 1/2$ about the minimum.

The systematic uncertainty on the mass due to uncertainties in the jet energy scale, soft and hard gluon effects, the shape of the background template, the fitting procedure and a possible bias from the b-taggers have been evaluated and are discussed in detail in reference [14]. Table 9.5 lists the uncertainty from

Systematic Uncertainties	
Source	Value (GeV)
Jet E_t Scale	3.1
Soft Gluon Effects	1.9
Different Generators	0.9
Hard Gluon Effects	3.6
Fit configuration	2.5
b -tagging Bias	2.3
Background Spectrum	1.6
Likelihood method	2.0
Monte Carlo statistics	2.3
Total	7.1

Table 9.5: Sources of systematic uncertainty on the measured top mass

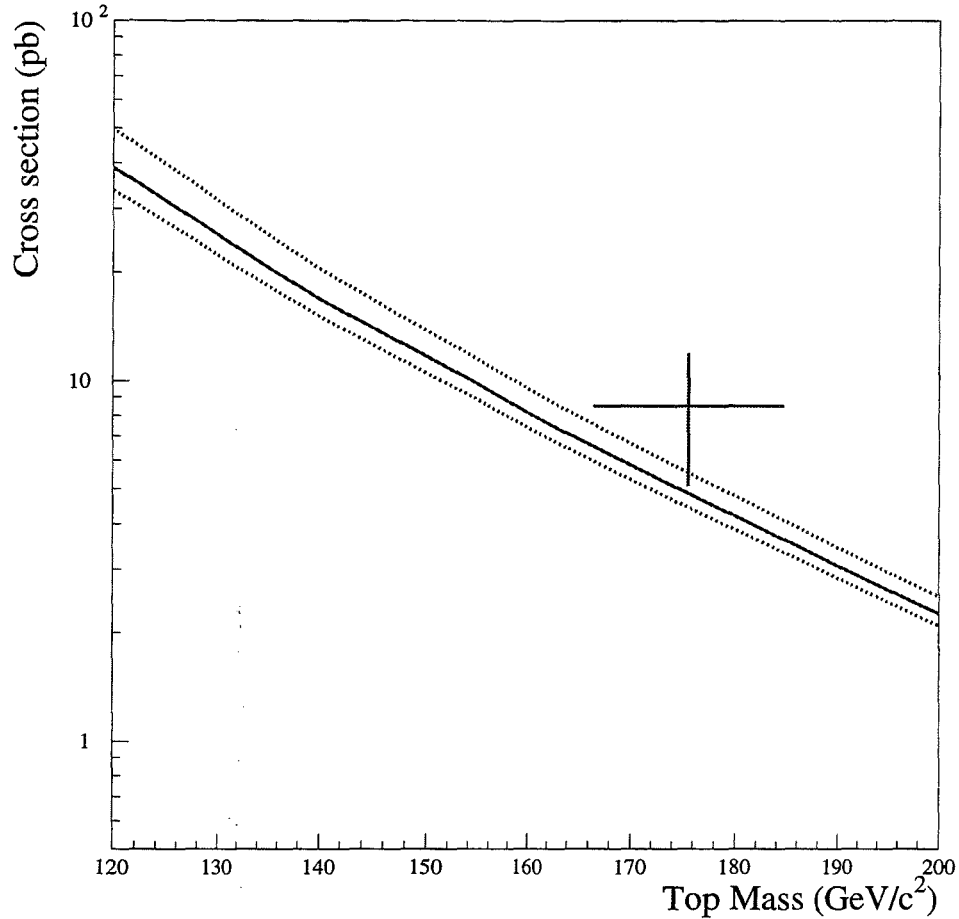


Figure 9-6: The $t\bar{t}$ production cross section measured using the SLT plotted for the CDF measured top mass using SVX and SLT tagged top candidates. The solid line shows the theoretical prediction from Laenen *et al.* [23]. The dashed lines show the variation in the result when the nonperturbative parameter μ_0 is changed.

each of these sources, which are added in quadrature to yield a total systematic uncertainty on the top mass of 7.1 GeV. The measured top mass is then

$$M_{top} = 175.6 \pm 5.7 \text{ (stat)} \pm 7.1 \text{ (syst)} \text{ GeV}/c^2.$$

The measured cross section and mass are compared to theoretical predictions in Figure 9-6 and are found to be in good agreement.

9.5 Conclusions

A search has been performed for $t\bar{t}$ events decaying in the lepton plus jets mode where the presence of a b-quark has been inferred from the observation of an

addition soft lepton in the event. The largest background to this search are W +multijet events which contain a soft lepton which is not from the semileptonic decay of a heavy flavor (b or c) quark. In $109.4 \pm 7.9 \text{ pb}^{-1}$ of data, 44 soft lepton tags are observed in 40 events with an expected background of 24.6 ± 3.0 tags. The probability that the background fluctuates up to give the observed number of tags is 0.18 %. The excess of soft lepton tagged events is observed to be stable to changes in the selection criteria, and consistent with the results from other searches. A variety of tests indicate that the background prediction and uncertainty are reliable. Assuming that the excess of events is due to top, the $t\bar{t}$ production cross section is measured to be

$$\sigma_{t\bar{t}} = 8.5 \pm 3.4 \text{ pb.}$$

These events, when combined with $t\bar{t}$ candidates identified with a displaced vertex tagger, are kinematically reconstructed and fit to give a top mass of

$$M_{top} = 175.6 \pm 5.7 \text{ (stat)} \pm 7.1 \text{ (syst)} \text{ GeV}/c^2.$$

Assuming this top mass, the measured cross section is in good agreement with theoretical calculations from the Standard Model.

Future Prospects for Top Measurements

With increased statistics CDF should be able to study in detail the characteristics of the top quark. Fermilab plans to begin the next collider run, *Run II* around or after 1999. The Main Ring will be replaced with a *Main Injector* which will allow larger anti-proton bunches, and the number of bunches will be increased from 6 to 36, giving an expected order-of-magnitude increase in luminosity to $2.5 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. The CDF detector will undergo major upgrades during this time, replacing the tracking chambers and plug and forward calorimeters. A new SVX will be installed which covers will cover the entire interaction region and enable 3D reconstruction of b vertices. The expected $t\bar{t}$ yields in 1 fb^{-1} of data are roughly 100 dilepton events, 275 lepton+jets

events tagged by the SLT and 500 tagged by the SVX. In half of the SVX tagged events both b's will be tagged which should allow the hadronic W in the event to be reconstructed and used to calibrate the calorimeter. The large number of SVX and soft lepton tagged events should enable CDF to measure the top mass to 5 GeV/c² and the $t\bar{t}$ production cross section to $\sim 10\%$. A 10% measurement of the branching ratio BR($t \rightarrow Wb$) is also expected which will constrain $|V_{tb}|$.

Improvements to the Soft Lepton Tagger

Currently the SVX-tagged lepton+jets events have a purity of $\sim 75\%$ while the purity of the SLT-tagged events is $\sim 40\%$. In the high-statistics era of RUN II the soft lepton tag algorithm should be altered to maximize the purity of tags. In particular, raising the soft lepton momentum to $P_t > 4$ GeV/c will decrease the background by a factor of two while retaining $\sim 80\%$ of the acceptance for $t\bar{t}$. The addition of a cut on the momentum of the soft lepton relative to the nearest jet axis (P_t^{rel}) should also provide a good means for rejecting background while separating direct decay leptons ($b \rightarrow \ell X$) from sequential decay leptons ($b \rightarrow c \rightarrow \ell$). This could potentially aid in the reconstruction of the top mass by using the charge of the lepton to associate the b-jet with the correct W and so reduce the combinatoric possibilities. (Figure 9-3 shows how the reconstructed mass is improved when the jet-quark assignments are correct.)

The soft electron algorithm may benefit from an improved method for reducing the background from conversions. Tracks, for instance, which do not leave hits in the VTX or SVX could be rejected. A significant reduction in the fake rate would allow the E_{had}/E_{em} cut to be removed which would double the soft electron efficiency. Detailed studies in this area have yet to be performed.

The most striking feature of the top quark is certainly its large mass which places it close to the scale of electroweak symmetry breaking. Hopefully the top quark will be as enlightening as it has been elusive.

Appendix A

Alternate Method of Estimating Fake, $Wb\bar{b}$ and $Wc\bar{c}$ Backgrounds

Since the fake rate from inclusive jets is expected to overestimate the number of tags in the W sample which come from heavy flavor, an alternate method (Method II) is presented here which attempts to calculate explicitly the fake, $Wb\bar{b}$ and $Wc\bar{c}$ backgrounds. The advantage to this approach is that it in principle does not assume that the heavy flavor content is the same in inclusive jets as in W +jets sample (without top). The disadvantage is that it depends heavily on Monte Carlo and theoretical calculations which carry large uncertainties. It serves, nonetheless, as a more conceptually satisfying check of the background estimate. The SVX top search in the W +jets channel has a much lower fake background, and so adopts this as its standard method [14][67].

A.1 The True Fake Rate

First the “fake rate” derived from the inclusive jet sample must be decomposed into two parts - tags which are truly fakes, and tags which originate from heavy flavor decay. This is done by studying the overlap of jets which are b-tagged by the SLT and the SVX. Recall that a *fake* tag is any particle identified as a lepton whose origin is not the a heavy flavor decay. This includes hadrons

which pass the lepton selection cuts (such as a pion which fakes an electron or a muon) as well as residual electrons from photon conversions or muons from pions or kaons which decay in flight. A *real* tag in what follows is an e or μ from the semileptonic decay of a b- or c-quark.

The SVX tagger

The standard CDF algorithm for finding jets with tracks which originate from a secondary vertex using the SVX is called SECVTX. The algorithm identifies two or more SVX tracks that have significantly large impact parameters and tries to fit them to a common displaced vertex. The distance in the transverse plane from the secondary vertex to the primary vertex is called L_{xy} . L_{xy} is positive if the vertex is on the same side of the primary vertex as the jet, and negative if it is on the opposite side. Jets with significantly displaced secondary vertices with positive L_{xy} are considered tags. SECVTX has a b-tag efficiency which increases from ~ 25 % for b-jets with $E_t = 15$ GeV to ~ 40 % for $E_t = 30$ GeV jets. The SECVTX fake (called *mistag*) rate is of order 1 % per jet.

The SVX and soft lepton taggers are applied to a sample of generic jets collected with the JET 50 trigger [84]. The number of SVX tags, as well as the number of soft lepton tags predicted by the SLT fake rate are summarized in Table A.1. The sample contains 453 (843) soft electron (muon) tags and 1002 SVX tags. The SVX tagged jets contain 24 (33) soft electron (muon) tags with 8.5 (16.5) predicted by the SLT fake rate. There is a clear excess of SLT's as expected in heavy flavor jets. The SECVTX algorithm also identifies 320 jets with significantly displaced *negative* values of L_{xy} which are expected to be largely mistags originating from mismeasured tracks in non-heavy flavor jets. Assuming the same number of mistags in the *positive* L_{xy} the b purity of SVX tags is ~ 70 %.

To a good approximation, the number of soft lepton tags from heavy flavor in the sample is given by

SLT and SVX tags in Jet 50 sample		
SVX tags	1002	
Soft Lepton Tags	453 e	843 μ
In SVX tags:		
SLT's	24 e	33 μ
SLT fake prediction	8.5 e	16.5 μ

Table A.1: The number of SVX and soft lepton tags found in a jet sample selected with a JET 50 trigger. The number of soft lepton tags found and predicted by the fake rate are shown for the subset of tracks in SVX tagged jets. A clear excess of soft lepton tags is observed as expected for real heavy flavor events.

$$N_{SLT}^{heavy\ flavor} = \frac{N_{SVX+SLT} - N_{SVX+fake\ SLT}}{\epsilon_{SVX+SLT}}$$

where $N_{SVX+SLT}$ is the number of jets tagged by SLT and SVX, $N_{SVX+fake\ SLT}$ is the number of soft lepton tags in SVX tagged jets predicted by applying the jet-derived fake rate, and $\epsilon_{SVX+SLT}$ is the efficiency for finding a (real) soft lepton tag in a SVX tagged heavy flavor jet. A study of the two taggers in a HERWIG generated jet sample predicts $\epsilon_{SVX+SLT} = 15 \pm 4\%$. Using the numbers from the table and subtracting the expected number of fake SLT tags, this indicates that there are $(24-8.5)/0.15 = 103 \pm 44$ real electrons and $(33-16.5)/0.15 = 110 \pm 50$ real muons in the sample. The uncertainty on these numbers include statistical uncertainties, the systematic uncertainties on the fake rate prediction, and the systematic uncertainty on $\epsilon_{SVX+SLT}$ which covers the ability of the Monte Carlo to model the SVX and SLT. Because this study was not done as a function of P_t or isolation, a further 50% systematic is assigned to the real heavy flavor fraction. Thus $23 \pm 15\%$ of the electron tags and $13 \pm 9\%$ of the muon tags in the sample are from real heavy flavor decay. Weighing the two by their relative contributions to the fake rate (see Table 8.8) and taking into account the correlations of the uncertainties, the jet derived

Jet Multiplicity	$F_{Wb\bar{b}}$	$F_{Wc\bar{c}}$	$\epsilon_{Wb\bar{b}}$	$\epsilon_{Wc\bar{c}}$
W+ 1 jet	$0.6 \pm 0.2 \%$	$1.5 \pm 0.5 \%$	$9 \pm 1 \%$	$3 \pm 1 \%$
W+ 2 jet	$1.1 \pm 0.3 \%$	$3.3 \pm 1.0 \%$	$10 \pm 2 \%$	$3 \pm 1 \%$
W+ 3 jet	$2.2 \pm 0.7 \%$	$6.0 \pm 2.0 \%$	$10 \pm 2 \%$	$5 \pm 1 \%$
W+ ≥ 4 jet	$3.6 \pm 1.7 \%$	$6.1 \pm 2.6 \%$	$10 \pm 2 \%$	$5 \pm 1 \%$

Table A.2: The fraction of W+jets events predicted from a HERWIG Monte Carlo simulation of jet events to contain a $b\bar{b}$ or $c\bar{c}$ pair. Systematic uncertainties are estimated by comparing SVX tag rates in a jet sample to HERWIG predictions. Also shown is the soft lepton tag rate for $Wb\bar{b}$ and $Wc\bar{c}$ events.

fake rate is determined to contain $16 \pm 10 \%$ real heavy flavor tags, and $84 \pm 10 \%$ fake tags. The background due to real fake soft lepton tags is then given by scaling the fake rate prediction by $0.84 \pm .10$.

A.2 $Wb\bar{b}$ and $Wc\bar{c}$ Backgrounds Using Monte Carlo

The fraction of events in the W+jets sample which are expected to be $Wb\bar{b}$ and $Wc\bar{c}$ has been estimated for the SVX top analysis using HERWIG and are presented in Table A.2 [39]. The uncertainties derive from a comparison of SVX tags in a HERWIG simulation of jet events and jet events in data. The second ingredient needed is the soft lepton tag rate for these events. The results of running the soft lepton tagger on the $Wb\bar{b}$ and $Wc\bar{c}$ samples are shown in Table A.2. There are limited Monte Carlo statistics in the high multiplicity bins, and so the tag rate calculated for W+2 jets is assumed to be valid for the W+3 and W+ ≥ 4 jets. The rate does not seem to change from the W+1 to W+2 jet which would tend to support this conclusion. The uncertainties in the table include the systematic uncertainty associated with the Monte Carlo implementation of the SLT algorithm and the statistical uncertainty of the

Background	W+1	W+2	W+3	W+ ≥ 4
fakes	170 \pm 25	48 \pm 7	11 \pm 1.6	5.1 \pm 0.8
$Wb\bar{b}$	4 \pm 1.2	1.3 \pm 0.4	0.4 \pm 0.1	0.2 \pm 0.1
$Wc\bar{c}$	3.6 \pm 1.1	1.1 \pm 0.3	0.5 \pm 0.2	0.2 \pm 0.1
Other sources	26 \pm 9	8.1 \pm 2.7	0.8 \pm 1.3	0.5 \pm 0.6
Total background	204 \pm 27	59 \pm 8	12.7 \pm 2.1	6.0 \pm 1.0
$\left(\begin{array}{l} \text{Total Background} \\ \text{Standard Method} \end{array} \right)$	228 \pm 20	65.1 \pm 5.7	14.0 \pm 1.8	6.6 \pm 0.8

Table A.3: Backgrounds calculated using Method II. The fake background is calculated by scaling the fake rate prediction down by 0.84 ± 0.10 . The $Wb\bar{b}$ and $Wc\bar{c}$ backgrounds are obtained from HERWIG generated Monte Carlo samples. “Other sources” includes the $Z \rightarrow \tau^+\tau^-$, $b\bar{b}$, Wc , Drell-Yan and Diboson backgrounds which are calculated as in the standard method (Table 8.8). The background calculated with the standard (fake rate) method is shown for comparison.

sample. With these inputs, the $Wb\bar{b}$ background can be calculated as

$$N_{Wb\bar{b}} = N_{events} \times F_{real\ W} \times F_{Wb\bar{b}} \times \epsilon_{Wb\bar{b}}^{SLT}$$

where the calculation is performed separately for each jet bin and where N_{events} is the number of observed events in the data, $F_{real\ W}$ is the fraction of events in the W sample expected to be real W’s (from Table 8.7), $F_{Wb\bar{b}}$ is the predicted $Wb\bar{b}$ fraction and $\epsilon_{Wb\bar{b}}^{SLT}$ is the soft lepton tag efficiency for $Wb\bar{b}$. The $Wc\bar{c}$ background is calculated in an identical manner.

The results from this method are summarized in Table A.3 and compare favorably with those of the standard method. As expected, the standard method yields slightly higher background estimates.

Appendix B

Calculation of the W Production Cross Section

While only a small fraction of the SLT backgrounds depend on Monte Carlo simulated events, this section details a quick check of the simulation by using it to measure the W production cross-section from the events in the W sample. This tests the muon trigger simulation which involves 16 different triggers, many of which are prescaled.

The acceptance for W events is calculated with a ISAJET generated sample of 120910 W's. Events which have an electron or muon over 15 GeV are simulated with QFL. The acceptance for W's is defined as the fraction of these events which have a lepton passing the selection criteria with $P_t > 20$ GeV and $E_t > 15$ GeV, and is calculated in the same manner for the $t\bar{t}$ acceptance in Chapter 6 which includes the trigger simulation and corrects for the overefficiency of the lepton id cuts when applied to Monte Carlo. In the sample, 2251 muon events and 3576 electron events are found giving a muon acceptance of 1.9 % and an electron acceptance of 3.0 %. By construction, the acceptances include the branching fractions $BR(W \rightarrow \ell\nu)$. The W cross section is calculated as

$$\sigma(p\bar{p} \rightarrow WX) = \frac{N_{observed}}{acceptance \cdot \int L}$$

	<i>Number of Events in Data</i>	<i>acceptance</i>	$\sigma(p\bar{p} \rightarrow WX)$
Electrons	57675	$3.0 \pm .2 \%$	$21 \pm 2 \text{ nb}$
Muons	38602	$1.9 \pm .1 \%$	$22 \pm 2 \text{ nb}$

Table B.1: The cross section $\sigma(p\bar{p} \rightarrow WX)$ measured using the inclusive W sample from an integrated luminosity of $90.1 \pm 7.2 \text{ pb}^{-1}$. The systematic error on the luminosity dominates the final uncertainty. The uncertainty on the acceptance from Monte Carlo is taken as 5 % to cover uncertainties in the lepton selection criteria. The results are in good agreement with previous measurements of $20.4 \pm 2 \text{ nb}$.

where $N_{observed}$ is the number of observed events and $\int L$ is the integrated luminosity for the sample ($90.1 \pm 7.2 \text{ pb}^{-1}$). In the W sample there are 38602 muon events and 57675 electron events. After correcting for the hadronic background as in Section 8.5.1 there are 37500 ± 2200 muon and 56000 ± 3300 electron W candidates giving a cross section of $21 \pm 2 \text{ nb}$ using the electrons and $22 \pm 2 \text{ nb}$ using the muons. CDF has previously published a measurement of $\sigma(p\bar{p} \rightarrow WX \rightarrow e\nu X) = 2.2 \pm 0.2 \text{ nb}$ using data from the 1988 run [41]. Dividing by the world averaged $BR(W \rightarrow e\nu) = .108 \pm .004$ [26] gives an expected W cross section of $20.4 \pm 2 \text{ nb}$ in good agreement with these findings which are summarized in Table B.1.

Appendix C

Test of SLT on SVX Tagged b's

Running the soft lepton tagger on a high purity sample of b's provides a thorough test of many components of this analysis. In this section the soft lepton tag rate is determined from an SVX selected sample of b's and compared to predictions from Monte Carlo.

The High Purity b Sample

Although SVX tags in generic jet events are expected to be roughly 70% pure heavy flavor, an even higher purity sample can be obtained by starting with events which contain an electron as from the decay $b \rightarrow e \bar{\nu}_e X$.

The low P_t electron sample is selected with a Level 2 trigger which requires a CEM cluster of $E_t > 8$ GeV matched to a CFT track of $P_t > 7.5$ GeV/c. Offline, electrons are required to have $E_t > 10$ GeV/c², $E/p < 1.5$, $|\Delta x| < 1.5$, $|\Delta z| < 3.0$, and $LSHR < 0.2$ and the event is required to have $\cancel{E}_t < 20$ GeV. The electron is not required to be isolated and is usually contained in a jet (the *electron jet*). Since there should be a second b in the event, an additional jet with $E_t > 15$ GeV which is at least $|\Delta\eta| > 2.5$ away from the electron is also required. There are 170,000 events passing these selection criteria. Figure C-1 show the electron E_t and the E/p distributions for events in the sample. Studies using the SVX indicate that $43 \pm 8\%$ of the electrons are from b's [102]. The purity can be increased to the 90% level by requiring either that

the electron or the jet be SVX tagged.

Test of Soft Electron Tags

In direct $b\bar{b}$ production (see Figure 8-1) the two b-jets are expected to be roughly back-to-back in phi. The SVX tagger, run on the sample, identifies 3272 jets with significantly displaced vertices which are at least $|\Delta R| > 2.5$ away from the electron. The distance of the displaced vertex from the primary vertex in the transverse plane (called L_{xy}) is shown in Figure C-1 and compared to the predicted shape from a HERWIG generated $b\bar{b}$ sample selected in the same manner. L_{xy} is positive if the displaced vertex is on the same side of the primary vertex as the jet, and negative if it is on the opposite side. Negatively signed L_{xy} tags are mostly fake tags from tracking errors [12]. Of the 3272 tagged jets, 199 have negative L_{xy} values. To a rough approximation the number of fakes in the $+L_{xy}$ tags may be assumed to be equal to the number of $-L_{xy}$ tags. Assigning a 30 % systematic to this assumption gives a b purity of $\sim 94 \pm 4$ % for the tagged jets. Having selected a high purity sample of b-jets, the soft electron tagger can now be tested. Considering only tracks which are within a cone of $R=0.4$ from the center of the tagged jet, 107 soft electrons are found. The electron fake rate applied to the same tracks predicts 25 tags. As discussed in Appendix A, the true fake rate is obtained by scaling this number by 0.77 ± 0.15 . The fraction of the b's with a real soft electron tag is then

$$F_{soft\ e} = \frac{N_{tags} - N_{fakes}}{N_{+L_{xy}} - N_{-L_{xy}}}$$

where N_{tags} is the number of observed soft electron tags (107), N_{fakes} is the number expected from true fakes (19.3 ± 7.0), $N_{+L_{xy}}$ ($N_{-L_{xy}}$) is the number of jets with a positively (negatively) signed displaced vertex. Including all systematic and statistical uncertainties, a value $F_{soft\ e} = 2.8 \pm 0.4$ is determined. Repeating the same procedure with a HERWIG generated $b\bar{b}$ Monte Carlo sample selected in the same manner, gives a predicted value of $F_{soft\ e} = 3.1 \pm 0.5(stat)$ in excellent agreement with the value measured from the data.

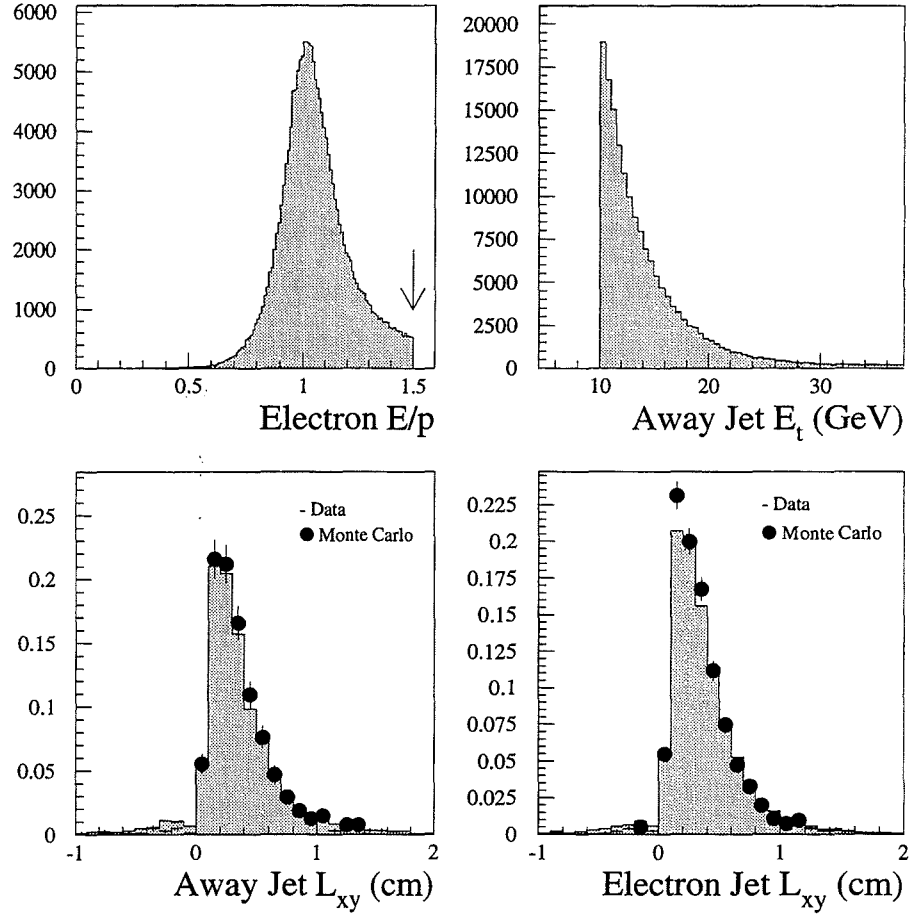


Figure C-1: Distributions from the low P_t electron sample. E/p of the electron after quality requirements have been imposed (top left), and the E_t spectrum of the jet in the event which is separated by at least $\Delta R=0.4$ from the electron. Also shown is the L_{xy} distribution for those jets which are tagged by the SVX (lower left) and for those electron jets tagged by the SVX (lower right). The expected L_{xy} distributions from HERWIG $b\bar{b}$ Monte Carlo is overlaid.

Test of Soft Muon Tagger

Unfortunately because of the bias associated with selecting a jet with energy over a threshold (see Section 8.1.3) the fake rate cannot be used to reliably estimate the number of fake muon tags in the away jets. The electron jets however, provide an alternative source of b-jets which are unbiased in this respect because they were selected using only the electron E_t . The SVX finds $9179 + L_{xy}$ tags and $447 - L_{xy}$ tags implying that the SVX tagged electron jets are $95 \pm 3\%$ pure b's. Considering only tracks within a cone of $R=0.4$ from the center of the SVX tagged electron jet, 324 soft muons are found. The fake rate predicts 91 fakes which must be scaled down by $.87 \pm .09\%$ (See Appendix A) to give the number of true fakes expected. The fraction of these SVX tagged jets with a real soft muon tag is then

$$F_{soft \mu} = \frac{N_{tags} - N_{fakes}}{N_{+L_{xy}} - N_{-L_{xy}}}$$

where N_{tags} is the number of observed soft muon tags (324), N_{fakes} is the number expected from true fakes (79 ± 12), $N_{+L_{xy}}(N_{-L_{xy}})$ is the number of electron jets with a positively (negatively) signed displaced vertex. Including all systematic and statistical uncertainties, yields a value of $F_{soft \mu}=2.8 \pm 0.3\%$. Repeating the procedure with a HERWIG generated $b\bar{b}$ sample gives a predicted value $F_{soft \mu}=2.8 \pm 0.4\%$ (stat) in excellent agreement with the measurement from the data. It should be noted that this only tests the Monte Carlo prediction for the sequential decay $b \rightarrow c \rightarrow \ell$ and not the primary decay $b \rightarrow \ell$.

Appendix D

Soft Electron Fake Rate Composition

The *fake rate* used to estimate the majority of the background for this analysis contains some leptons from real heavy flavor decay and some *fake* leptons. Fake leptons are any particles identified as leptons whose origin is not the semileptonic decay of a bottom or charm quark. The fake rate method does not require a detailed understanding of the composition of the tags in the jet sample, which is both a strength and a limitation. In order to improve the soft lepton tag algorithm in the future, or to estimate the background in a more detailed manner (See Appendix A) it is necessary to know the various components of the fake rate. A study of the jets tagged by both the SVX and SLT (Appendix A) indicate that $23 \pm 15\%$ of electron tags in generic jets come from real heavy flavor decay. The remaining portion must derive either from hadrons which fake a lepton, or from real electrons from photon conversion or Dalitz decay ($\pi^0 \rightarrow \gamma e^+ e^-$). The probability for a hadron to fake a soft electron can be determined by using a sample of known hadrons in the data. In the same jet sample used to determine the fake rate, the decays $K_s \rightarrow \pi^+ \pi^-$ and $\Lambda \rightarrow p \pi^-, \bar{p} \pi^+$ can be reconstructed by searching for opposite sign tracks that originate from a common point which is displaced from the primary interaction. First, all possible pairs of opposite sign tracks are formed.

The tracks are required to intersect in the r - ϕ plane, and the distance in z between the tracks at this point is required to be less than 5 cm. Tracks must have an impact parameter of at least 0.2 cm, and at least one of the pair must have $P_t > 2$ GeV/c. Finally the proper decay length ($c\tau$) is determined from the vector sum of the track P_t 's (P_t^{pair}) and their radius of convergence (R)

$$c\tau = \frac{R \cdot M}{P_t^{pair}}$$

where M is the mass of the K_s (0.5 GeV) or Λ (1.1 GeV), and required to be at least 0.5 cm. Figure D-1 shows the reconstructed mass peaks for each particle. For the Λ decay, the track with the highest momentum is assumed to be the proton. To improve the mass resolution, the tracks have been constrained to come from a common vertex and refit. The selection criteria are summarized in Table D.2. After subtracting the background under the mass peaks using the sidebands, there are 15664 ± 149 tracks from K_s 's and 7806 ± 140 from Λ 's which have $P_t > 2$ GeV. Table D.1 lists the number of these tracks which are identified as soft electrons, and the number predicted by the fake rate. The number of tags and the predicted fakes have been background subtracted using the sidebands. The background subtraction is particularly important here since some of it will be real electrons from photon conversions. Figure D-1 shows the radius of convergence for events in the *tails* of the K_s and Λ mass distributions which exhibit a bump around $R=30$ cm as expected for photons which convert in the outer VTX and inner wall of the CTC (compare Figure 4-4). This is not surprising since the K_s and Λ 's are selected in a manner similar to the conversion sample.

The large background under the Λ mass peak results in large systematic uncertainties in the number of tags found and predicted by the fake rate. The K_s peak is much cleaner. Here there are 20 ± 6 soft electron tags and 36 ± 9 predicted by the fake rate implying that 55 ± 22 % of the fake rate is due to pions which pass the soft electron selection criteria. If 23 ± 9 % of the fake rate comes from real heavy flavor (Section A) then this implies that $\sim 22\%$ of the

source	tracks	soft electrons	fake prediction
K_s	15664 ± 149	20 ± 6	36 ± 9
Λ	7806 ± 140	3 ± 8.5	22 ± 8

Table D.1: The number of tracks, soft electron tags found and soft electron tags predicted by the fake rate in the K_{short} and Λ mass peaks. For each number a background subtraction has been performed using the sidebands. The large uncertainties on the numbers from the Λ peak are a result of the large background present under the mass peak.

impact parameter > 0.2 cm
$c\tau > 0.5$ cm
$ \Delta z < 5$ cm

Table D.2: K_{short} and Λ selection requirements.

background is due to real electrons from photon conversions or Dalitz decay. This study was not repeated for soft muons because pions from K_s 's which decay in flight may have their momentum mismeasured and so not reconstruct to the K_s mass which would introduce a bias.

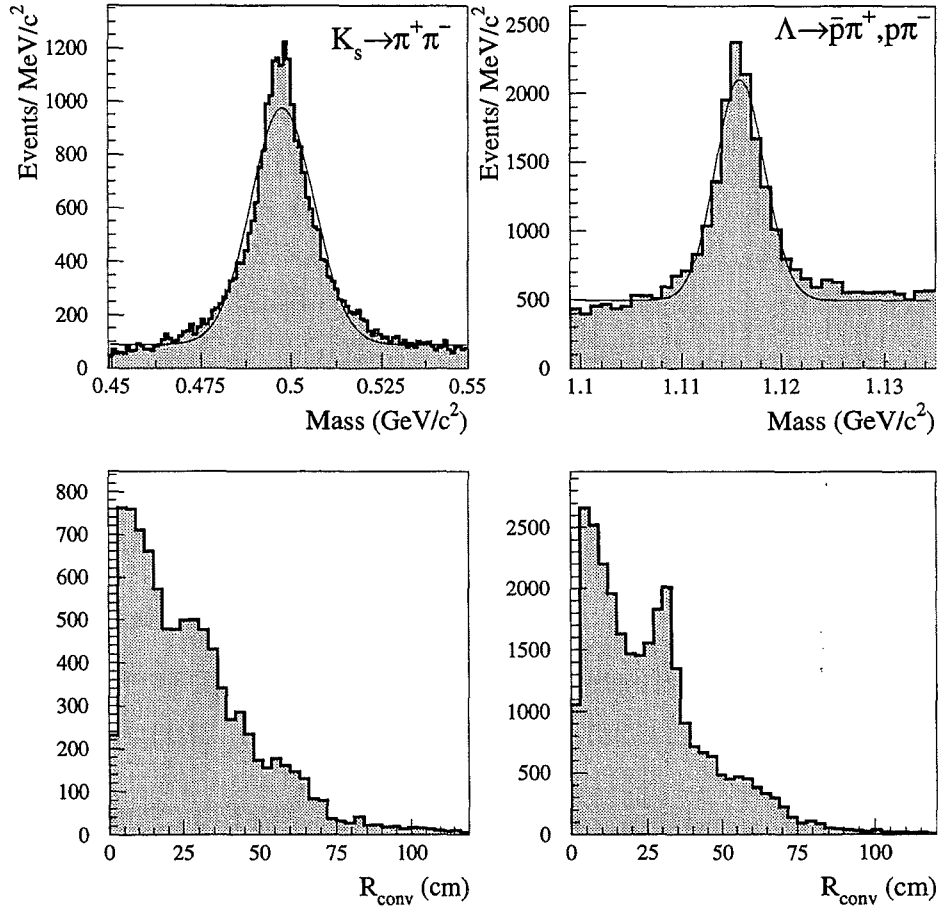


Figure D-1: **Top:** The reconstructed Kshort (K_s) and Λ mass peaks. **Bottom:** The radius of convergence for tracks in the sidebands used for background subtraction (left) and for a wider sideband region (right) showing a bump from photons which have converted in the outer VTX and inner CTC wall producing an e^+e^- pair.

Bibliography

- [1] S. Weinberg, Phys. Rev. Lett., **19**, 1264 (1967); (introductory textbooks from an experimental perspective) D. Griffiths, Introduction to Elementary Particles, (Harper & Row Publishers, 1987); D. Perkins, Introduction to High Energy Physics, Addison-Wesley Publishing Company, 1987); (field theory) H. Georgi, Weak Interactions, (Benjamin/Cummings Pub. Co., 1984) V. Barger and R. Phillips, Collider Physics, (Addison-Wesley Publishing Company, 1987).
- [2] C. Athanassopoulos *et al.*, nucl-ex/9605001, (1996); C. Athanassopoulos *et al.*, nucl-ex/9504022, (1995) (submitted to Phys. Rev. Lett).
- [3] F. Reines and C. L. Cowan, Jr., Physical Review **113**, 273 (1959).
- [4] S. Neddermeyer and C. Anderson, Physical Review **51**, 884 (1937).
- [5] G. Danby *et al.*, Phys. Rev. Lett. **9**, 36 (1962).
- [6] M. L. Perl *et al.*, Phys. Rev. Lett. **35**, 1489, (1975).
- [7] G. D. Rochester and C. C. Butler, Nature **160**, 855 (1947).
- [8] J. J. Aubert *et al.*, Phys. Rev. Lett. **33**, 1404, (1974); J.-E. Augustin *et al.*, Phys. Rev. Lett. **33**, 1406, (1974).
- [9] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).
- [10] D. P. Barber *et al.*, Phys. Lett. **B85**, 463 (1979); D. P. Barber *et al.*, Phys. Rev. Lett. **44**, 1722 (1980).

- [11] G. Arnison *et al.*, Phys. Lett. **122B**, 103 (1983); P. Bagnaia *et al.*, Phys. Lett. **129B**, 130 (1983); G. Arnison *et al.*, Phys. Lett. **126B**, 298 (1983).
- [12] F. Abe *et al.*, Phys. Rev. D, **50**, 2966 (1994).
- [13] F. Abe *et al.*, Phys. Rev. Lett., **73**, 226 (1994).
- [14] F. Abe *et al.*, Phys. Rev. Lett., **74**, 2626 (1995).
- [15] J. Lewis *et al.*, CDF internal document
CDF/PUB/MUON/PUBLIC/2858 (1995) (eventually in Nucl. Instrum.
Meth. Phys. Res.)
- [16] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988).
- [17] D. Amidei *et al.*, Nucl. Inst. Meth. Phys. Res., **A350**, 73 (1994); D. Amidei
et al., Nucl. Inst. Meth. Phys. Res., **A360**, 137 (1995).
- [18] D. Kestenbaum *et al.*, CDF internal document
CDF/MUON/DOC/PUBLIC (1995).
- [19] A. Daw *et al.*, CDF internal document CDF/DOC/MUON/CDFR/1614
(1991).
- [20] T. LeCompte *et al.*, CDF internal document
CDF/BOTTOM/ANAL/CDFR/2588 (1994).
- [21] T. Liss , CDF internal document CDF/TOP/DOC/CDFR/2998, (1995).
- [22] R. Casalbuoni *et al.*, hep-ph/9505212, (1995); C. Hill, FERMILAB-PUB-
94/395-T, (1994), submitted to Phys. Lett. B.
- [23] E. Laenen, J. Smith and W. L. van Neerven, Phys. Lett B **321**, 254 (1994).
- [24] D. Gerdes, CDF internal document CDF/TOP/ANAL/CDFR/2903
(1994).

- [25] R. Hughes *et al.*, CDF internal document CDF/ANAL/TOP/CDFR/2992 (1995).
- [26] Particle Data Group, Phys. Rev. D, **D50** Part I (1994).
- [27] M. L. Mangano, Nucl. Phys. **B405**, 536 (1993).
- [28] M. L. Mangano and P. Nason, Phys Lett. B **285**, 160 (1992); A. H. Mueller and P. Nason, Nucl. Phys. **B266**, 265 (1986).
- [29] G. Marchesini and B. R. Webber, Nucl. Phys. **B310**, 461 (1988); G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 265 (1992).
- [30] F. Abe *et al.*, Phys. Rev. Lett. **75**, 608 (1995).
- [31] M. Kruse, Thesis, Purdue University, (1995); J. Konigsberg proceedings electroweak interactions, XXXTH Rencontres de Moriond, France, March (1995).
- [32] S. Kopp, Thesis, The University of Chicago, (1994).
- [33] F. Ptohos, Thesis, Harvard University, (1996).
- [34] G. Unal, H. H. Williams, CDF internal document CDF/ANAL/TOP/CDFR/2315 (1993).
- [35] C. Campagnari *et al.*, CDF internal document CDF/ANAL/TOP/CDFR/2245, (1993).
- [36] M. L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).
- [37] M. L. Mangano, Nucl. Phys. **B405**, 536 (1993).
- [38] A. Caner *et al.*, CDF internal document CDF/ANAL/TOP/CDFR/2363, (1993).
- [39] G. Unal, CDF internal document CDF/ANA/TOP/CDFR/3389, (1995).

- [40] G. Unal, CDF internal document CDF/ANA/TOP/CDFR/3513, (1995).
- [41] F. Abe *et al.* Phys. Rev. D, **68**, 29, (1990).
- [42] R. Hamilton, FERMILAB-CONF-94-209-E, (1994), to be published in DPF94.
- [43] J. Ohnemus, Phys. Rev. D, **44**, 3477 (1991); J. Ohnemus, Phys. Rev. D, **44**, 1403 (1991).
- [44] M. Shaw *et al.*, CDF internal document CDF/ANAL/TOP/CDFR/2252, (1993).
- [45] B. Jacobsen, Talk presented at the XXIXth Rencontres de Moriond, CERN-PPE/94-97, (1994).
- [46] R. Akers *et al.* Z. Phys. C, **65**, 569, (1995).
- [47] P. Abreau *et al.* Z. Phys. C, **67**, 365, (1995).
- [48] P. Frazini, Physics Reports, **173**, 1, (1989).
- [49] J. Busenitz, Talk presented at the 1995 International Conference on High Energy Physics in Brussels, Belgium, (1995).
- [50] G. Kane and M. Peskin, Nucl. Phys., **B195**, 29, (1982).
- [51] C. Matteuzzi *et al.*, Physics Letters, **129**, 141, (1983).
- [52] L. Ryder, Quantum Field Theory, (Cambridge University Press, 1985).
- [53] H. Albrecht *et al.*, Phys Lett. **192B**, 245 (1987); C. Albajar *et al.*, Phys Lett. **186B**, 247 (1987).
- [54] G. Altarelli *et al.*, Nucl. Phys. , **B308**, 725, (1988).
- [55] R. K. Ellis, Physics Letters B, **259** 492, (1991).
- [56] P. Nason, S. Dawson, and R.K. Ellis, Nucl. Phys., **B303**, 607, (1988).

- [57] E.L. Berger and H. Contopanagos, Phys. Lett., **361**, 115, (1995).
- [58] S. Catani, M. Mangano, P. Nason, L. Trentadue, Cern-TH/96-21.
- [59] C. Campagnari and M. Franklin, submitted to Review of Modern Physics, (1996).
- [60] M. Glück, J.F. Owens, and E. Reya, Phys. Rev. D, **17**, 2324, (1978).
- [61] L. H. Orr, Phys Rev. D, **44**, 44, (1991).
- [62] I. Bigi *et al.*, Physics Letters B, **157**, (1986).
- [63] C. Albajar *et al.*, Z. Phys C, 505, (1988).
- [64] V. Barger and R. Phillips, Collider Physics, (Addison-Wesley Publishing Company, 1987)
- [65] M. Gallinaro, presented at Recontres du Vietnam, Saigon, October (1995); M. Hohlmann, proceedings from Lake Louise Winter Institute, (1996); M. Gallinaro, and M. Hohlmann, CDF internal document, CDF/ANAL/TOP/CDFR/3461, (1996).
- [66] P. Azzi and L. Galtieri, presented at XXXI Rencontre de Moriond (Hadronic Session), France, (1996).
- [67] D. Glenzinski, Thesis, The Johns Hopkins University, (1995).
- [68] G. Bauer *et al.*, CDF internal document, CDF/DOC/CALORIMETRY/CDFR/2591, (1994).
- [69] D. Perkins, Introduction to High Energy Physics, (Addison-Wesley Publishing Company, 1982)
- [70] J. D. Jackson, Classical Electrodynamics, (Wiley, 1962).
- [71] Ch. Berger *et al.*, Phys. Lett. **B76**, 243 (1978).

- [72] E. Rice *et al.* Phys. Rev. Lett., **48**, 906 (1982).
- [73] S. Abachi *et al.* Phys. Rev. Lett., **72** 2138 (1994).
- [74] F. Abe *et al.* Phys. Rev. D, **45**, 1448 (1992).
- [75] F. Paige and S. D. Protopopescu, BNL-38034, unpublished, (1986).
- [76] G. Marchesini and B.R. Webber, Nucl. Phys., **B238**, 1 (1984); G. Marchesini and B.R. Webber, Nucl. Phys., **B310**, 461 (1988);
- [77] H. U. Bengtsson and T. Sjöstrand, comput. Phys. Commun., **46**, 43 (1987).
- [78] Cross Section and Branching Ratios Working Group, CDF internal document, CDF/TOP/ANAL/CDFR/3403 (1995).
- [79] G. Ascoli *et al.*, Nuclear Instruments and Methods in Physics Research, **A268**, 33 (1988).
- [80] B. Farhat, Thesis, The Massachusetts Institute of Technology (1993).
- [81] C. Campagnari *et al.*, CDF internal document, CDF/TOP/ANAL/CDFR/2970 (1995).
- [82] T. Liss, CDF internal document, CDF/DOC/TOP/GROUP/2998 (1995).
- [83] F. Abe *et al.*, Phys. Rev. Lett., **75**, 608 (1995).
- [84] G. Unal *et al.*, CDF internal document, CDF/ANAL/TOP/CDFR/2281 (1993).
- [85] F. Abe *et al.*, Phys. Rev. D., **45**, 1448 (1992).
- [86] P. Avery, K. Read, and G. Trahern, Cornell Internal Note CSN-212 (1985) (unpublished).
- [87] S. Kim *et al.*, CDF internal document, CDF/BOTTOM/ANAL/CDFR/3287 (1996).

- [88] F. Abe *et al.* Phys. Rev. D., **52**, 4874 (1995).
- [89] A. Yagil, Private communication, noon, 4/23/1996.
- [90] O. Dahl, T. Day, F. Solmitz, and N. Gould, Lawrence Berkeley Lab, Group A Programming Note P-126, (1968) (unpublished).
- [91] B. Harral, CDF internal document, CDF/ANAL/TOP/CDFR/3526 (1996).
- [92] L. Galtieri *et al.*, CDF internal document, CDF/ANAL/TOP/CDFR/3026 (1995).
- [93] S. Kuhlmann, CDF internal document, CDF/ANAL/JET/CDFR/1248 (1990).
- [94] H. Keutellian, Private transaction, 2/96.
- [95] F. Abe *et al.*, Phys. Rev. Lett., **70**, 4042 (1993).
- [96] MRS Collaboration, A.D. Martin, R. G. Roberts, and W.J. Stirling, Report No. RAL-93-077 (1993).
- [97] CTEQ Collaboration, J. Botts, Phys., Lett. B **304**, 159 (1993).
- [98] F. Abe *et al.*, Phys. Rev. Lett. **74**, 850 (1995).
- [99] The Aleph Collaboration, CERN-PPPE/94-017, (submitted to Z. Phys.) (1994).
- [100] The Opal Collaboration, CERN-PPPE/93-106, (submitted to Z. Phys. C.) (1993).
- [101] J. D. Richman and P. R. Burchat, UCSB-HEP-95-08, Stanford-HEP-95-01, to be published in Reviews of Modern Physics (1996).
- [102] W. Yao *et al.*, CDF internal document CDF/ANAL/TOP/CDFR/2989 (1995).