Measurement of the Top Quark Mass

by

Erich Ward Varnes

B.A. (Johns Hopkins University) 1992
M.A. (University of California at Berkeley) 1994

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Mark Strovink, Chair
Professor Marjorie D. Shapiro
Professor Eugene Haller

Spring 1997

The dissertation of Erich Ward Varnes is approved:

Chair Date

Date

Date

University of California at Berkeley

Spring 1997

Abstract

Measurement of the Top Quark Mass

by

Erich Ward Varnes

Doctor of Philosophy in Physics

University of California at Berkeley

Professor Mark Strovink, Chair

This dissertation describes the measurement of the top quark mass m_t using events recorded during a 125 pb^{-1} exposure of the DØ detector to $\sqrt{s} = 1.8$ TeV $\overline{p}p$ collisions. Six events consistent with the hypothesis $t\overline{t} \rightarrow bW^+, \overline{b}W^- \rightarrow b\overline{\ell}\nu, \overline{b}\ell\overline{\nu}$ form the dilepton sample. The kinematics of such events may be reconstructed for any assumed m_t, and the likelihood of each such solution evaluated. A measurement of m_t based on these relative solution likelihoods gives $m_t = 169.9 \pm 14.8$ (stat.) ± 3.8 (syst.) GeV/c^2. A 2C kinematic fit is performed on a sample of 77 events consistent with $t\overline{t} \rightarrow bW^+, \overline{b}W^- \rightarrow b\overline{\ell}\nu, \overline{b}q\overline{q}$, and this, in combination with an estimate on the likelihood that each event is top, yields $m_t = 173.3 \pm 5.6$ (stat.) ± 6.2 (syst.) GeV/c^2. A combination of these two measurements gives $m_t = 173.1 \pm 5.2$ (stat.) ± 5.7 (syst.) GeV/c^2.

Professor Mark Strovink
Dissertation Committee Chair

In honor of my parents, and in memory of my grandparents.

Contents

List of Figures viii

List of Tables xiv

1 Introduction 1

2 Theoretical Overview 3
- 2.1 The Standard Model . . . 3
- 2.2 Top Quark Production . . . 9
- 2.3 Top Quark Decay . . . 12
- 2.4 Definition of the Top Quark Mass . . . 14
- 2.5 The Role of the Top Quark Mass . . . 14
- 2.6 Beyond the Standard Model . . . 17
- 2.7 Top Quark Measurements . . . 17

3 Experimental Apparatus 19
- 3.1 TeVatron at Fermilab . . . 20
- 3.2 DØ Detector . . . 24
  - 3.2.1 Central Detector . . . 26
  - 3.2.2 Basics of Drift Chamber Operation . . . 26
  - 3.2.3 Calorimeter . . . 38
  - 3.2.4 Muon System . . . 46
  - 3.2.5 Triggering . . . 50
  - 3.2.6 Main Ring Effects . . . 54
  - 3.2.7 Data Acquisition . . . 55
- 3.3 Detector Operation . . . 57

4 Event Reconstruction 58
- 4.1 The DØRECO Program . . . 58

4.1.1 Central Tracking Chamber Reconstruction 59
4.1.2 Event Vertex Determination 61
4.1.3 Calorimeter Reconstruction 62
4.2 Particle Identification 63
4.2.1 Electrons 63
4.2.2 Electromagnetic Energy Calibration 67
4.2.3 Muons 68
4.2.4 Missing E_T 72
4.2.5 Jets 74
4.2.6 Jet Energy Calibration 76

5 Top Quark Event Selection and Cross Section Measurement 84
5.1 Monte Carlo Modeling of Signal and Background 85
5.1.1 Model for $t\bar{t}$ Events 85
5.1.2 Model for Background Events 87
5.1.3 Model for the Detector Response 87
5.2 Data Sample 90
5.3 Event Cleanup 90
5.4 Dilepton Channels 91
5.4.1 $e\mu$ 92
5.4.2 Dielectron 94
5.4.3 Dimuon 97
5.5 Lepton Plus Jets Channels 98
5.5.1 Topological Selection 99
5.5.2 b-tag Selection 103
5.6 The $e\nu$ Channel 106
5.7 Top Quark Production Cross Section 107
5.8 Lepton Plus Jets Cuts Used for Mass Analysis 110

6 Mass Reconstruction 114
6.1 Dilepton Events 114
6.2 The Neutrino-Weighting Method 117
6.2.1 Modeling the Detector Resolution 121
6.2.2 Jet Combinatorics 123
6.3 Monte Carlo Tests 126
6.3.1 Parton-level Tests 127
6.3.2 Tests on GEANTed Monte Carlo 130
6.4 Lepton Plus Jets Events 130
6.4.1 Kinematic Fitting 130

6.4.2 Jet Corrections to the Parton Level 135
6.4.3 Effects of Gluon Radiation 137

7 Maximum Likelihood Fit **145**
7.1 Dilepton Events 147
7.1.1 Probability Density Estimation 149
7.1.2 Form of the Likelihood 151
7.2 Lepton Plus Jets Events 151
7.3 Combining Histograms 154

8 Data Analysis **158**
8.1 Dilepton Analysis 158
8.1.1 Ensemble Tests 160
8.1.2 Results 166
8.1.3 Systematic Uncertainties 174
8.1.4 Five Event Analysis 181
8.2 Lepton Plus Jets Analysis 187
8.2.1 Results 192
8.2.2 Systematic Uncertainties 196
8.3 Combined Top Quark Mass Result 197
8.3.1 Parameterization of the $-\ln L$ Curve 197
8.3.2 Systematic Uncertainties 199

9 Conclusions **202**
9.1 Production Model 202
9.2 Constraint on the Higgs Mass 202
9.3 The Future of Top Quark Physics 205

A Monte Carlo Generation **206**
A.1 Signal Events 206
A.2 Background Samples 207
A.2.1 Lepton Plus Jets Analysis 207
A.2.2 Dilepton Analysis 210

B Kinematics of the Dilepton Sample **212**
B.1 Object Resolutions 212
B.2 Tables of Dilepton Kinematics 214
B.2.1 $e\mu$ Events 214
B.2.2 Dielectron Events 215
B.2.3 Dimuon Event 216

B.3 Comparison of Dilepton Kinematics 216

Bibliography **218**

List of Figures

2.1 Leading-order diagrams for $t\bar{t}$ production in $\overline{p}p$ collisions. 10
2.2 Approximate $O(\alpha_s^4)$ calculations of the top quark production cross section in $\overline{p}p$ collisions with center-of-mass energy 1.8 TeV. Shown are the calculations of Laenen *et al.*[10] (solid), Berger *et al.*[11] (dashed), and Catani *et al.*[12] (dotted). 11
2.3 Family of curves showing the variation of m_W as a function of m_t for various Higgs masses [18]. The width of the curves is due to uncertainty in the measurement of other Standard Model parameters, and the horizontal lines show the result of recent measurements of the W mass and the one standard deviation limits [19]. 16

3.1 The Fermilab accelerator complex. (Not to scale – the Main Ring and TeVatron have the same radius, and are separated here for clarity). . 22
3.2 Cutaway view of the DØ detector, showing the tracking chambers, calorimetry, and muon system. 25
3.3 End view of one quarter of the Vertex Drift Chamber. 30
3.4 Layout of wires and cells in the Central Drift Chamber. Each cell contains seven signal wires (small points), fourteen field wires (large points) and two delay lines (open circles). 32
3.5 Exploded view showing the orientation of the Forward Drift Chambers. 33
3.6 One sector of one module of the Transition Radiation Detector. . . . 36
3.7 Structure of a calorimeter readout cell. 41
3.8 Side view of one quarter of the calorimeter, showing segmentation and tower definitions. The lines extending from the center of the detector are labelled by their pseudorapidities. 42
3.9 Total material in the calorimeter and muon toroid, as a function of polar angle. Except for the gap between the central and end toroids (at 40°) there are $\gtrsim 14\ \lambda$ in front of the outer drift tube layers. . . . 47
3.10 End view of proportional drift tubes used in the muon system. . . . 49

3.11 Sketch of the data acquisition system. 56

4.1 68% confidence intervals for the electromagnetic response parameters α and δ. The wide vertical band is the constraint from J/ψ decays, the narrow band from π^0 decays, and the ellipse from Z boson decays. The small ellipse is the combined constraint [49]. 69
4.2 Jet correction factor as a function of measured jet transverse energy [53]. 80
4.3 Percentage imbalance between photon and jet E_Ts for data and Monte Carlo samples [54]. The parameterizations shown are a triple-Gaussian for data and a double-Gaussian for Monte Carlo. 82
4.4 E_T balance in data and Monte Carlo direct photon events, after application of the post-CAFIX corrections. 83

5.1 Scatter plot of the A and H_T distribution of 170 GeV HERWIG $t\bar{t}$ events, VECBOS W plus four jet events, and multijet events in the data. . . 100
5.2 Scatter plot of expected signal and background events in the e + jets channel for many possible choices of the $(\mathcal{A}, H_T)$ cut point. Lines of constant expected uncertainty on the cross section are overlaid, and the optimal cut point (0.065, 180 GeV) is indicated by the arrow [72]. 101
5.3 Number of $W \to e\nu$ events as a function of jet multiplicity, along with the expected contribution from $t\bar{t}$ production [72]. 104
5.4 Measured $t\bar{t}$ production cross section as a function of top quark mass. The upper and lower curves bound the one standard deviation interval. 109
5.5 Separation between Monte Carlo 170 GeV/c^2 $t\bar{t}$ (solid) and background (dashed) events obtained from the variables H'_{T2} and $K_{T\min}$ 112
5.6 Top likelihood discriminant function, for signal (solid) and background (dashed) events. 113

6.1 Schematic representation of $t\bar{t}$ production and decay in the dilepton channels. 115
6.2 Distribution of neutrino rapidities from top quark decay as modeled by HERWIG, for a range of top quark masses (the two highest E_T neutrinos from each event are entered in the plots). It can be seen that the distributions are close to Gaussian, and that the width of the best-fit Gaussian decreases slightly with increasing top quark mass. 119
6.3 Width of the Gaussian that gives the best fit to the neutrino η distribution as a function of top quark mass. Also shown is the quadratic fit that parameterizes this dependence. 120

6.4 Fractional difference between reconstructed and generated jet transverse energies in HERWIG $t\bar{t}$ samples with top quark masses ranging from 110 to 190 GeV/c^2. Shown are the distributions for three E_T bins for jets in the central, intercryostat, and forward regions. Superimposed on each distribution is the double-Gaussian parameterization used in event smearing. 124
6.5 Average $\mathcal{W}(m_t|\{o_i\})$ for 130 and 190 GeV/c^2 unsmeared parton-level ISAJET $t\bar{t}$ events. The distributions were created by normalizing the area of the weight curves for each event to unity and then summing the weight curves for many events. One observes a sharp peak near the input top quark mass. 128
6.6 Average weight distribution for parton-level 190 GeV/c^2 Monte Carlo. Figure (a) shows the distribution when the events are smeared according to the expected detector resolution, Figure (b) introduces the twofold ambiguity in lepton-jet pairings, Figure (c) adds initial-state gluon radiation, and Figure (d) suppresses initial-state but allows final-state gluon radiation. 129
6.7 Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the ee channel. 131
6.8 Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the $e\mu$ channel. 132
6.9 Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the $\mu\mu$ channel. 133
6.10 Schematic representation of $t\bar{t}$ production and decay in the lepton plus jets channels. 134
6.11 Reconstructed vs. parton energies for central jets. 136
6.12 Comparison of jet invariant masses before (dashed) and after (solid) out-of-cone radiation corrections for: (a) two jets from hadronic W decay, (b) three jets from top decay (with the b untagged) and (c) with the b muon-tagged for 180 GeV/c^2 top events. In each case the distribution moves closer to the pole value following the corrections [79]. 138
6.13 Degradation in the resolution of the fitted mass in lepton plus jets events due to gluon radiation and detector resolution. The shaded histograms give the distribution of fitted masses when the correct jet assignment is input to the fit. The bottom right figure shows the result of taking a χ^2-weighted average of the three lowest χ^2 solutions of each event. 141

6.14 Result of applying the kinematic fit procedure on a sample of 170 GeV/c^2 HERWIG $t\bar{t}$ events in each of the lepton plus jets channels. . 143
6.15 Result of applying the kinematic fit procedure to samples of HERWIG $t\bar{t}$ events for a variety of top masses. The top quark mass favored by the lowest χ^2 solution for each event is plotted. The solid curves are for events in the mass analysis base sample, the dashed are those which pass the further selection described in Chapter 7, and the dotted are those which pass the cuts used in the cross section measurement[76]. 144

7.1 An illustration of the procedure for coalescing the $W(m_t)$ distribution into five 40-GeV/c^2 bins. Shown is the $W(m_t)$ distribution for a typical Monte Carlo $t\bar{t}$ event, normalized to unity. 148
7.2 Distribution of reconstructed masses for HERWIG 140 and 175 GeV/c^2 $t\bar{t}$ events, for those that pass (bin one) and fail (bin two) the LB cut described in the text [76]. 154
7.3 Distribution of background events, with binning as defined in Fig. 7.2. Both the W + multijet events, as modeled by VECBOS, and the QCD multijet events taken from data tend to have low top probability and relatively low reconstructed mass [76]. 156
7.4 Distribution of reconstructed mass vs. neural network output for HERWIG 140 and 175 GeV/c^2 $t\bar{t}$ events [76]. 156
7.5 Distribution of reconstructed mass vs. neural network output for background events [76]. 157

8.1 $\mathcal{W}(m_t)$ distributions for the dilepton candidates. For events with more than two jets, the dashed curves show the results of considering only the two highest E_T jets in the reconstruction. 159
8.2 Distribution of $\widehat{m_t}$ for ensemble tests at a range of top quark masses. 164
8.3 Pull distributions with Gaussian fits for a variety of m_t^{MC} 168
8.4 Result of the maximum likelihood fit to the dilepton sample. Plotted are the summed $w_1 \ldots w_5$ for the data sample (points), and the prediction of the best-fit model of background (dotted) and signal plus background (dashed). The error bars on the data points represent the RMS of the w_i in ensemble tests. The $-\ln L$ curve and its parameterization are inset at upper right. The curve is drawn as a solid line in the region considered in the fit. 170
8.5 Figure (a) shows the distribution of $\hat{\sigma}$ obtained from ensemble tests run with $m_t^{\mathrm{MC}} = 170$ GeV/c^2. The arrow marks the value returned by the fit to the data (14.8 GeV/c^2). The dashed lines bound the region of ensembles which enter (b). 171

8.6 Distribution of minimal $-\ln L$ obtained from 170 GeV/c^2 ensemble tests. The arrow indicates the result obtained from the data. 172
8.7 Result of the maximum likelihood fit to the ee plus $e\mu$ channels (top) and $e\mu$ only (bottom). 173
8.8 Difference in photon-jet transverse energy balance between Monte Carlo and data samples as a function of photon E_T. The dashed curves show the offset that would result if the data and Monte Carlo jet energy scales were offset relative to one another by $\pm(2.5\% + 0.5$ GeV) [54]. 176
8.9 Average $\mathcal{W}(m_t)$ distributions for a sample of HERWIG 170 GeV/c^2 $t\bar{t}$ to dilepton events, with zero (a), one (b) and two (d) additional minimum bias interactions overlaid. Figures (c) and (e) show the resolutions of the z vertex determination for the one and two additional interaction samples. 180
8.10 Result of the maximum likelihood fit for the five-event analysis. . . . 184
8.11 Distribution of minimal $-\ln L$ obtained from five-event 165 GeV/c^2 ensemble tests. The arrow indicates the result obtained from the data. 185
8.12 Distribution of candidate events (top plots) on LB plot, where events in the first bin on the likelihood axis are those most likely to be top. The bottom plots show the distribution for the best-fit signal plus background model of the data [76]. 188
8.13 Distribution of candidate events (top plots) on NN plot. The bottom plots show the distribution for the best-fit signal plus background model of the data [76]. 189
8.14 Results of ensemble tests for an input mass of 175 GeV/c^2 and assumed signal to background ratio of 1:2 [76]. 191
8.15 Results of the maximum likelihood fit to the data, for both the LB (top) and NN (bottom) binnings. In each plot the candidate distribution in reconstructed mass is shown, along with the best-fit model for signal (dotted), background (dashed) and their sum. The $-\ln L$ curve is inset at upper right [76]. 193
8.16 Distributions of $\hat{\sigma}$ from lepton plus jets ensemble tests. Arrows indicate the values obtained by fitting the DØ data [76]. 194
8.17 Reconstructed mass distribution for events passing the LB selection [76]. 195
8.18 Result of multiplying the likelihood curve from the dilepton analysis with that from the lepton plus jets LB (top) and NN (bottom) analyses. 198

9.1 Comparison of the measured top quark mass and cross section with the predictions of perturbative QCD. 203

9.2 Curves showing the minimal Standard Model Higgs boson mass as a function of the W boson and top quark masses. The cross represents the world average measurement of m_W and the DØ measurement of m_t. 204

B.1 Comparison of observed E_Ts (shaded histograms) to the expectation from the signal + background (solid) and background-only (dashed) models. The signal plus background histograms are normalized to six events, and the background-only histograms are normalized to the expected background level. The numbers on the plots are the Kolmogorov-Smirnov probabilities that the data sample was drawn from the model. 217

List of Tables

3.1 TeVatron parameters. 24
3.2 Vertex Drift Chamber parameters [26]. 31
3.3 Central Drift Chamber parameters [26]. 32
3.4 Forward Drift Chamber Θ module parameters [26]. 34
3.5 Forward Drift Chamber Φ module parameters [26]. 35
3.6 Some properties of uranium. 40

4.1 Cuts on $\mathcal{R}$ used in the lepton plus jets and dilepton analyses. The single-lepton analysis does not use TRD information in defining $\mathcal{R}$. . 67
4.2 Identification criteria for muons in Era I (the variables and Eras are defined in the text). As the dilepton channels have lower backgrounds and require greater efficiency, not all cuts are applied to these channels, as indicated in the third column. 73
4.3 Identification criteria for muons in Eras II and III. 73

5.1 Integrated luminosity for each of the $t\bar{t}$ decay channels. There is a 5.3% uncertainty in the luminosity measurement. 90
5.2 Kinematic cuts for the dilepton event selection. The cut used in place of $\not\!\!E_T$ to reject $Z \to \mu\mu$ events is described in the text, as is the H_T^e variable. Also, the muon η cut is run-dependent, as detailed in Chapter 4. 92
5.3 Number of background events expected in each dilepton channel for the dominant sources. The nature of the instrumental backgrounds differs somewhat between the channels, as described in the text. . . 93
5.4 Kinematic and fiducial cuts for objects in lepton + jets selection. The fiducial cuts for muon selection are as described in Chapter 4. 98

5.5 Summary of the contribution of each dilepton channel to the cross section measurement. The expected number of $t\bar{t}$ events (N_{top}) is based on the central value of the theoretical cross section [10] and assumes a mass of 170 GeV/c^2. 108
5.6 Contributions of the lepton plus jets channels to the cross section measurement. 108

6.1 RMS width of the distribution of peak $\mathcal{W}(m_t)$s as a function of the number of times each event is smeared for HERWIG 170 GeV/c^2 $t\bar{t}$ events. 122
6.2 Possible assignments of three observed jets (j_1, j_2, and j_3) as b jets or initial state radiation. There is an additional twofold ambiguity in pairing the leptons with the b jets. 126
6.3 Out-of-cone jet correction factors for light quarks and untagged b quarks for each η region. 139
6.4 Number of possible assignments of jets to parent partons in reconstructing a $t\bar{t}$ event in the lepton plus jets channel, as a function of jet multiplicity [80]. 139

8.1 Results of ensemble tests showing the effect of changing the PDE smoothing parameter h. All quantities are in GeV/c^2, and the Q criterion used to evaluate the relative sensitivity of the ensemble tests is defined in Eq. 8.1. 162
8.2 Results of ensemble tests generated at a wide variety of m_t^{MC}. 163
8.3 Results of ensemble tests showing the effect of different parameterizations of the $-\ln L$ curve. The fits are described by the number of points about the minimum $-\ln L$ considered, and the degree of polynomial used. 166
8.4 Pull widths for ensemble tests generated at a wide variety of m_t^{MC} with five, seven, or nine points near the minimum considered in the quadratic fit. 167
8.5 Results of several variations of the maximum likelihood fit to the data. The error estimates quoted are simply the $\hat{\sigma}$, with no correction for any non-unit pull width. 175
8.6 Sources of systematic error. See the text for details on how these errors were estimated. 176
8.7 Effect of varying the jet scale in ensemble tests with $m_t^{\text{MC}} = 170$ GeV/c^2. 177
8.8 Results of fitting ensembles of ISAJET events to signal samples modeled by HERWIG. 178

8.9 Effect of introducing dummy models for that portion of the background which has no Monte Carlo model. 178
8.10 Results of five-event ensemble tests generated at a wide variety of m_t^{MC}. 182
8.11 Pull widths for five-event ensemble tests generated at a wide variety of m_t^{MC} with five, seven, or nine points near the minimum considered in the quadratic fit. 183
8.12 Results of some variations of the five-event analysis. 185
8.13 Effect of varying the jet scale in ensemble tests with $m_t^{\mathrm{MC}} = 170$ GeV/c^2 for the five-event analysis. The symmetric excursion is ± 3.2 GeV/c^2. 186
8.14 Results of fitting ensembles of ISAJET events to signal samples modeled by HERWIG in the five-event case. The mean absolute difference in the median $\widehat{m}_t$ is 2.0 GeV/c^2. 186
8.15 Effect of unmodeled backgrounds in the five-event analysis. 186
8.16 Results of ensemble tests of the $\ell+$ jets maximum likelihood fit. Listed are the means and 68% widths of the $\widehat{m}_t$ distributions of 1000 ensembles fit using both the neural network and top likelihood discriminants [76]. 190
8.17 Systematic uncertainties in the measurement of the top quark mass in the lepton plus jets channels [76]. 196
8.18 Pull widths for different parameterizations of the combined $-\ln L$ curve. 199
8.19 Results of several parameterizations of the combined $-\ln L$ curves. . 200
8.20 Systematic uncertainties in the combined measurement of the top quark mass. 201

A.1 Summary of HERWIG $t\bar{t}$ samples used in the analysis. Samples were generated separately for the dilepton and lepton plus jets channels. . 208
A.2 Summary of VECBOS samples used in the analyses. Those marked with an asterisk represent a modified version of VECBOS which extended the range of the W Breit-Wigner distribution. This distinction is of crucial importance to the $e\nu$ channel, but nearly irrelevant for the other channels. 210
A.3 Summary of background Monte Carlo samples used in the dilepton mass analysis. 211

Acknowledgements

It was in May of 1994 that I first drove into the (as yet unpaved) parking lot at DØ. My first two years of graduate study were complete, and while my coursework had given me a sound foundation in modern physics theory, I had little knowledge of the workings of an experimental physics collaboration and could only guess at the tasks involved in performing an analysis. Writing a doctoral dissertation was a feat almost beyond my comprehension.

In fact, this document would not exist were it not for the efforts of the many people who took the time to listen to my questions, correct my misconceptions, and provide me with suggestions. First among these is my advisor, Mark Strovink. Not only has his guidance been of immeasurable importance to the completion of this work, he has been concerned from the outset with my overall development as a physicist, from encouraging me always towards the timely completion of my thesis to his tremendous assistance in my search for a postdoctoral position. In short, he has been an exemplary advisor in every possible way, and I would strongly recommend that any student seize the opportunity to work with him.

Of course, there are others whose guidance has also been invaluable. Among those who have been most helpful are Scott Snyder (whose exceptional work set the standard for this measurement, Tom Ferbel (in addition to everything else, for his wonderful St. Croix summer school), Tom Trippe, Serban Protopopescu, Uli Heintz, Meena Narain, John Hobbs, Jeff Bantly, Rich Partridge, Harrison Prosper, Natalie Roe, Ron Madaras, Al Clark, Lynn Stevenson, Hugh Montgomery, and Paul Grannis. I also must thank the many people whose years of dedication and effort made possible the construction and operation of the finest particle detector in the world.

In addition to being rewarding, my time in graduate school has also been enjoyable, thanks to the following people whom I feel fortunate to name as friends: "Lord" Jim Cochran (often an island of reason in storms of controversy), Gene Álvarez (a good guy despite being a Yankees fan), Jamal Tarazi, Ian Adam, Paul Bloom, the LBL gang

(Azriel Goldschmidt, Justin Bendich, Peter Grudberg, Danilo Pušeljić, and Emanuela Barberis), and all the members of the Prerna and Fermions soccer teams.

Among all the friends I have made at graduate school, Patrick Virador was the first and remains the closest. Many of my finest memories of this time will be of the hikes in Marin, Cal football games, and long talks about physics and life that we shared. Together we made the drive from splendor of the Bay Area to the featureless prairie to begin our research. We are not finishing the journey together, and my sadness at this fact is tempered by the knowledge that Patrick will meet with great success and happiness in his new field.

While a good part of my development into a professional physicist took place in graduate school, the process was set in motion in my junior year of high school by Mr. Robert Quail. I have yet to meet a better teacher or a finer gentlemen. It was he who introduced me to physics, to the power of formulating a mathematical description of the world, and to the inescapably addictive intellectual challenges associated with deriving and testing such a formulation.

I also must thank Michael Bernstien, who has been my friend from our days at Canterbury Elementary. We have grown up but never apart, for which I am very grateful.

Naturally, none of this would have been possible without the support of my parents, Dr. Arthur Varnes and Dr. Marie Varnes. Their sacrifices enabled me to attend Johns Hopkins for my undergraduate education, and their encouragement enabled me to follow my dream.

Last on this list but first in my heart is my fiancée, Kathryn Gonzalez. I will be eternally grateful that she somehow managed to overlook the fact that I showed up for our first date clad in an old T-shirt and blue jeans. The joy she has brought to my life (let alone the enhancement she has brought to my wardrobe) is more than I could ever express, and I can't wait for the beginning of our life together.

So, as I prepare to load up my car and leave DØ I would like to again thank everyone above, and to apologixe to those I have inadvertently omitted. The past

four and one half years have been an incredible time in my life, and they are the people who have made it so.

Chapter 1

Introduction

The study of high-energy physics fills a peculiar niche in the endeavor to expand the limits of human knowledge. The aim is to understand the workings of the universe at the most fundamental level. That is, one desires to identify those constituents of matter which may not be subdivided any further and to describe completely all of the interactions between them. The study proceeds by probing the structure of matter at progressively smaller scales of distance, from atoms to nuclei to protons to quarks, and correspondingly at higher scales of energy.

As this pursuit has progressed in the past century or so, the apparatus needed to reach the next frontier of energy has grown inexorably more complex and costly, and the time needed to design, construct, and execute an experiment has also grown accordingly. So, too, has the number of people that must collaborate in order for such an effort to succeed.

This dissertation describes an analysis carried out at the current frontier of particle energy, that provided in the collision of beams of protons and antiprotons counter-rotating in a four-mile ring of magnets located at the Fermi National Accelerator Laboratory in Batavia, IL. The accelerator is designed such that these collisions occur at two points, each of which is surrounded by a massive apparatus designed to measure the results of the collisions. One of these, the DØ detector, provides the

data used in this analysis.

One of the main goals of any experiment which probes previously unexplored regimes of energy is to search for particles too massive to have been produced at any previous experiment. Such a particle was in fact discovered in the course of the experiment, and its name befits its standing as the most massive fundamental particle known: the top quark. The goal of this dissertation is to measure of the mass of this particle with the highest possible accuracy.

The dissertation begins in Chapter 2 with a brief description of our current model of the fundamental nature of the universe. In particular, the place of the top quark in that model and the motivations for an accurate measurement of its mass, are highlighted. Chapter 3 gives a description of the accelerator and detector used in the experiment, and Chapter 4 proceeds to describe how the data are processed to provide information about the particles that were produced in the collision. In Chapter 5, the isolation of a small sample of $t\bar{t}$ events from the millions of collisions recorded is described, as is the measurement of the $t\bar{t}$ production cross section based on that sample. Chapter 6 then describes how the kinematics of the final-state particles in each candidate event are used to reconstruct the top quark mass, and Chapter 7 details the procedure for extracting the top quark mass most consistent with the sample as a whole. In Chapter 8 the application of the above machinery to the data sample is described and the results presented. Finally, in Chapter 9 the analysis is summarized, and the degree to which its motivations were satisfied is explored. The dissertation concludes with a brief discussion of what may be learned from more precise measurements of the top quark mass at future experiments.

Chapter 2

Theoretical Overview

The most recent particle to be discovered is the top quark, first seen in 1995 after a nearly two-decade search [1, 2].

This chapter provides an outline of the currently accepted model of nature, the place of the top quark in that model, and a discussion of top quark physics.

2.1 The Standard Model

The twentieth century opened with two of the greatest revolutions in the history of physics, each of which extended our knowledge into previously unexplored regimes and in so doing fundamentally altered the way in which we think of the universe. The theory of relativity posits that the nature of the space we live in is fundamentally different than our intuitive perception would lead us to believe, and that time must be understood as a dimension of that space. The theory of quantum mechanics defines a system in terms of a state which evolves in time according to a wave equation, rather than a collection of particles moving according to the Newtonian rules. This leads to a host of phenomena which defy common sense, and yet are necessary to explain the structure of the atom and the behavior of any system at very small scales. It is apparent that any fundamental description of nature must contain within it both the

concepts of quantum mechanics and relativity.

Such a synthesis has yet to be achieved. The gravitational force is described naturally in general relativity as a geometrical consequence of the true nature of space-time. On the other hand, phenomena occurring roughly at or below the scale of the atom can be understood in terms of quantum mechanics. The closest we have come to combining these ideas is to formulate theories which combine the fundamental tenets of quantum mechanics with the particle kinematics appropriate to the special theory of relativity.

The first successful attempt at such a description was Dirac's theory of the electron, put forth a mere two years after the advent of quantum mechanics. The theory had a "flaw", however: it could only work if there existed an "anti-electron" having the same mass but the opposite electric charge of the electron. The theory was redeemed when the anti-electron was discovered a few years later. This is the earliest example of a theory predicting the existence of an unobserved particle.

As successful as the Dirac theory was, there were still phenomena for which it was inapplicable. Notably, it could not conveniently describe any system for which the number of particles is not a constant, and while it could predict the energy levels of atomic states very accurately, it could not predict the lifetimes of excited states. The remedy to these shortcomings is *quantum field theory*, in which particles arise as the quanta of fields with an infinite number of degrees of freedom. By quantizing the field as a whole rather than individual particles, phenomena such as the creation and annihilation of particles can be described naturally. While our current theories encompass more particles and a greater range of interactions than did Dirac's, the concept of the quantum field still forms their basis.

The next difficulty encountered was one of calculation. The complexity introduced by quantizing the field renders exact calculation impossible. In order to compare the predictions of a theory with experiment one therefore needs an approximation technique. The most powerful such technique yet discovered is *perturbation theory*, which treats the interactions of the particles as small fluctuations around the vacuum

state, and can describe any physical process as an infinite series in increasing powers of the interaction strength (or *coupling constant*)[1]. Each term of this series can be calculated. Early results were promising, as the leading terms in the series produced results close to those observed. However, the subleading terms, which are expected to be small corrections, turned out to be infinite. Progress was stalled until a well-defined method of removing the infinities (called *renormalization*) was developed.

Among all possible quantum field theories, some guidance is needed to select the one which correctly describes nature. Such guidance can be obtained by considering only those theories whose basic equation (the Lagrangian) is invariant under generalized phase transformations (referred to by the nondescriptive term *gauge transformations*) even if the magnitude of the transformation is allowed to vary as a function of position. Such theories are called *local gauge theories*, and have two very attractive properties. First, if one begins from the kinetic energy term for any particle (which is determined uniquely by the demands of relativity), and defines the group of transformations under which the theory is invariant, one is forced to add terms to the Lagrangian which represent the coupling of the original particle to a *gauge boson*. Hence the local gauge theory *predicts* the form of the interactions, once the symmetry group and the properties of the particle in that group are defined. Secondly, it was shown by 't Hooft [3] that all such theories are amenable to renormalization.

The simplest local gauge theory which includes all known interactions (save gravity, which remains the province of general relativity) and particles, and is consistent with all experimental observations, is known as the *Standard Model*. This model, first formulated over twenty years ago, proved successful not only in explaining all experimental observations up to that time, but also in predicting (or accommodating) all subsequent observations in particle physics. It is a measure of the success and near-universal acceptance of this model that one can learn about it in textbooks

[1] Another approximation technique which has made significant strides recently is *lattice gauge theory*. In this approximation, space-time is modeled as a collection of discrete points, and the quantum field calculations are carried out in this space.

rather than journal articles. A few good examples are [4, 5, 6, 7].

The Standard Model is based on the gauge group SU(3) $\times$ SU(2) $\times$ U(1). The SU(3) sector describes the *color* interaction, which acts on the class of particle known as *quarks*. This interaction also leads indirectly to the binding of protons and neutrons to form atomic nuclei. A problem arises, however, in using the SU(3) sector of the Standard Model to calculate the interactions of the quarks bound inside a baryon or meson: at the energy scale which typifies these interactions, the color coupling constant α_s is large, which means that perturbation theory is inapplicable. Therefore comparisons between the predictions of this sector of the Standard Model (also known as Quantum Chromodynamics, or QCD) and experimental results are most easily carried out for processes at high energy. Another unique feature of QCD is that the SU(3) group requires that the gauge bosons (gluons) interact among themselves. It is this feature that allows the possibility of the coupling constant decreasing with increasing momentum scale (an effect known as *asymptotic freedom*). In SU(3), asymptotic freedom will occur if one has fewer than seventeen types of fermions of mass less than the momentum scale of the calculation. The Standard Model, with twelve known fermion types (six quarks and six leptons), meets this criterion.

Conversely, one may expect (though it is not yet proven) that in an asymptotically free theory the increase of the coupling constant at low momentum scales continues even into the nonperturbative regime, thus creating an arbitrarily strong binding between any particles with color charge. Indeed, no quark has been observed in a free state. Rather, it is color-neutral combinations of $q\bar{q}$ and qqq which appear as strongly interacting particles in experiment.

The SU(2) $\times$ U(1) piece of the Model describes the unification of the electromagnetic interaction, which is familiar in everyday life, with the weak interaction, which is responsible for certain types of radioactive decay of nuclei. The two interactions differ primarily in the ranges over which they act. The electromagnetic interaction is of infinite range, implying that the gauge boson carrying the interaction is massless. On the other hand, the range of the weak interaction is very small, implying that it is

mediated by massive bosons (this also explains why we barely notice this interaction in everyday life).

The unification of these two interactions was hard-won and stands as the crowning jewel of the Standard Model. The idea is that the underlying SU(2) $\times$ U(1) symmetry is spontaneously broken by the presence of an additional scalar field, known as the *Higgs field*. In a spontaneously broken theory, the Lagrangian itself has a certain symmetry which is not shared by the set of states which lie at the minimum of the Lagrangian. Since the perturbation expansion must be done about one of the minima, the results of the calculation will not appear to share the underlying symmetry.

In particular, the SU(2) $\times$ U(1) theory includes four massless gauge bosons, while the simplest version of the Higgs field also has four scalar degrees of freedom. However, in expanding about the minimum of the potential, these separate degrees of freedom become entangled, with three of the Higgs bosons reinterpreted as the the longitudinal components of three of the gauge bosons. These three bosons (the W^+, W^-, and Z) thereby acquire mass, while the fourth (the photon) is massless. The remaining Higgs scalar should exist as a physical particle in the theory.

The introduction of the Higgs scalar is the only known mechanism which allows a renormalizable theory to include massive gauge bosons, and also allows one to introduce masses for all other particles which are proportional to their couplings to the Higgs field (simply introducing mass terms by fiat would violate the gauge symmetry).

Lest this seem too much like black magic, it should be noted that the introduction of the Higgs mechanism leads to a prediction of the ratio of the W and Z masses:

$$\frac{m_W}{m_Z} = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} \equiv \cos\theta_W \tag{2.1}$$

where g_1 and g_2 are the coupling constants of the U(1) and SU(2) interactions. Independent measurements of m_W, m_Z, and $\cos\theta_W$ (which may be measured in a variety of electroweak processes) have confirmed this prediction, as have a number of other

sensitive tests.

On the other hand, the nature of the Higgs field is completely unknown. There is no guarantee that the simplest case outlined above is correct, and even in this case there is little information about the mass of the physical Higgs boson. Elucidation of the nature of this field is essential for the logical completion of the Standard Model.

Having discussed the symmetries, we introduce the particles in order to specify the model. The table below lists the particles and gives their electric charges (in units of the proton charge), masses (from [8]), and interactions in which they participate:

	Particle	Charge	Mass (GeV/c^2)	Interactions
Leptons	Electron (e)	-1	0.000511	EM, Weak
	Electron neutrino (ν_e)	0	$< 5.1 \times 10^{-9}$	Weak
	Muon (μ)	-1	0.1057	EM, Weak
	Muon neutrino (ν_μ)	0	$< 2.7 \times 10^{-4}$	Weak
	Tau (τ)	-1	1.771	EM, Weak
	Tau neutrino (ν_τ)	0	< 0.031	Weak
Quarks	Up (u)	$+2/3$	≈ 0.005	EM, Weak, Color
	Down (d)	$-1/3$	≈ 0.010	EM, Weak, Color
	Charm (c)	$+2/3$	≈ 1.3	EM, Weak, Color
	Strange (s)	$-1/3$	≈ 0.20	EM, Weak, Color
	Top (t)	$+2/3$	> 128	EM, Weak, Color
	Bottom (b)	$-1/3$	≈ 4.3	EM, Weak, Color

Note that the masses of the light quarks are somewhat ambiguous, as the quarks are always bound into mesons or baryons. Note also that there seems to be an unnecessary threefold redundancy in this list. For example, the electron, muon, and tau are all identical in their properties and interactions, differing only in their masses. As the Dirac theory required the anti-electron, the Standard Model demands that each of particles listed in the above must have its own antiparticle of the same mass and interactions, but opposite charge. Antiparticles are denoted by placing a bar over the symbol for the particle.

2.2 Top Quark Production

The experiment discussed in this dissertation involves the collision at high energy of protons and antiprotons. At the energies involved, the composite nature of these particles is evident, and one speaks of the actual collision as being between their constituents (either quarks or gluons). Hence the energy available in any given collision will only be some fraction of the $\overline{p}p$ system center-of-mass energy, since the interacting constituents (or *partons*) carry only a fraction of the proton or antiproton energy.

The most likely means of producing top quarks is through the color interaction, and since this interaction conserves quark flavor quantum numbers, the top quarks must be produced in $t\bar{t}$ pairs[2]. Both the strategies developed for selecting a sample of top quark events and those for measuring the top quark mass assume that top quarks are pair-produced.

The leading-order diagrams for top quark production are shown in Fig. 2.1. There are two major production channels, the $q\bar{q}$ annihilation and gluon fusion channels. Since it is more likely for a significant fraction of the proton and antiproton momenta to be carried by one of their constituent quarks than by gluons, the $q\bar{q}$ annihilation channel dominates the production rate, and the degree of this dominance increases as a function of the top quark mass.

As mentioned above, the top quark production cross section can be calculated using perturbative techniques, and such calculations have been carried out to next-to-leading order. In addition, there are nonperturbative techniques which can be used to estimate the size of the higher-order terms. Three examples of such calculations [10, 11, 12] are shown in Fig. 2.2.

Two things are important to note in Fig. 2.2. First there is the rapid decrease in cross section with increasing mass. This implies that measuring the mass of the top quark, along with its cross section, provides a sensitive test of the QCD calculations.

[2]In fact, the cross section for the electroweak production of a single top quark is of the same order of magnitude as that for $t\bar{t}$ production [9]. However, events with only one top quark in the final state are much more difficult to distinguish from background.

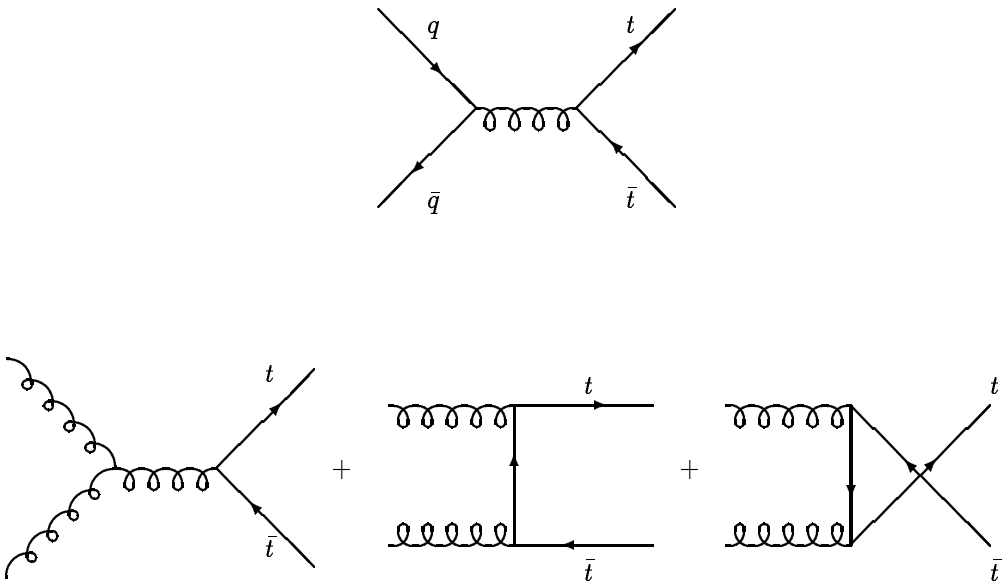

Figure 2.1: Leading-order diagrams for $t\overline{t}$ production in $\overline{p}p$ collisions.

Even with a small sample of top quark events it would be possible to detect deviations from the prediction, which could be a window to new physics. For example, a cross section significantly greater than that predicted by QCD may indicate that the top quark plays a dynamical role in the electroweak symmetry breaking [13]. On the other hand, a measured cross section lower than the prediction might imply that the top quark has additional decay channels beyond those in the Standard Model, which would indicate the presence of new particles.

Secondly, one notes that the cross section is of order several picobarns, while the total $\overline{p}p$ interaction cross section at the energy of this experiment is roughly a tenth of a barn. Hence one expects to sift through about ten *billion* events in order to find one example of top quark production.

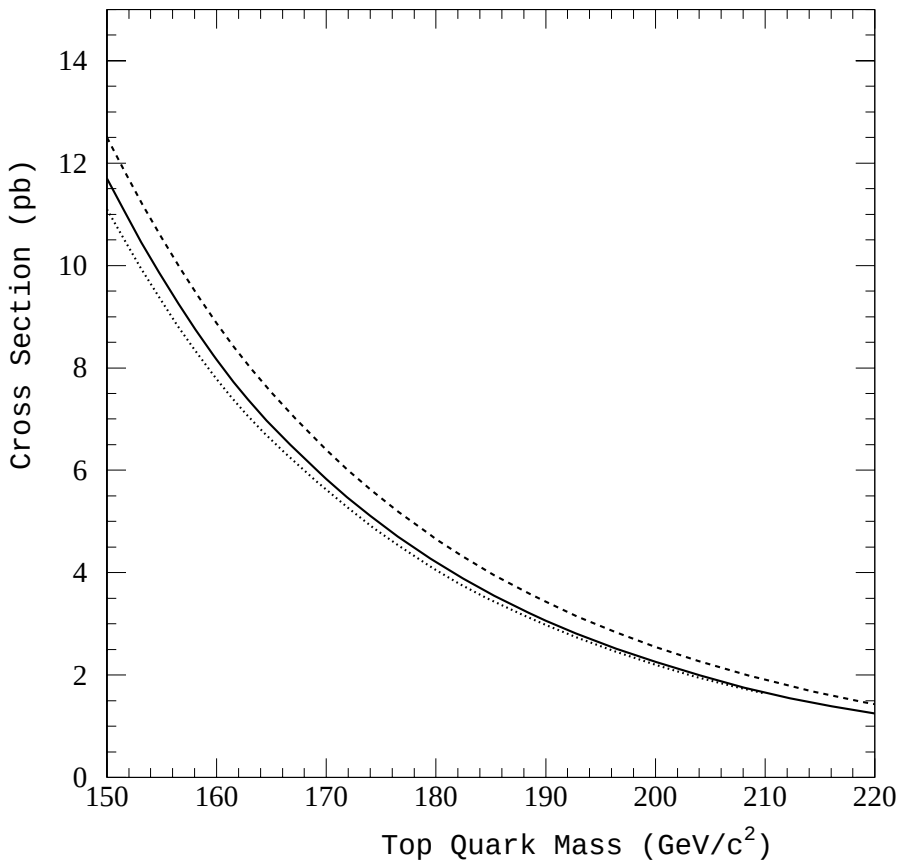

Figure 2.2: Approximate $O(\alpha_s^4)$ calculations of the top quark production cross section in $\overline{p}p$ collisions with center-of-mass energy 1.8 TeV. Shown are the calculations of Laenen *et al.*[10] (solid), Berger *et al.*[11] (dashed), and Catani *et al.*[12] (dotted).

2.3 Top Quark Decay

The only interaction which does not conserve quark type (or *flavor*) is the weak interaction, and hence this is the only route open for top quark decay in the Standard Model. Under SU(2), the fermions transform as doublets[3]:

$$\begin{pmatrix} u \\ d' \end{pmatrix} \begin{pmatrix} c \\ s' \end{pmatrix} \begin{pmatrix} t \\ b' \end{pmatrix}$$

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

The upper and lower members of any doublet can couple to a W boson, while a Z boson may only couple a particle to itself. Therefore we see that the top quark must decay to a b' through its coupling with a W. The primes after the quark names in the above grouping denote the fact that the objects listed are eigenstates of the weak interaction.

However, it is the nature of quantum mechanics that an eigenstate of one operator is in general not an eigenstate of other operators. In particular, the weak eigenstates listed above are mixtures of the mass eigenstates listed in the earlier table (it is purely convention that the mixing is defined to occur only for the lower members of the doublet). Were it not for this mixing, the b and s quarks would never decay in the Standard Model.

The degree of mixing can be parameterized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L$$

[3]This is true for the left-handed components of the fermions. The right-handed particles enter as SU(2) singlets, meaning they do not participate in the weak interactions.

None of the elements involving the top quark have been measured directly, but can be inferred by applying the necessary physical condition that the matrix be unitary. Using this, and assuming that no more than three generations of quarks exist, the direct measurements of the CKM matrix elements involving the lighter quarks imply that V_{tb} is ≈ 0.999 [8].

Combining all of the above, we see that nearly all top quarks will decay via $t \rightarrow Wb$. The b quark will fragment and hadronize, forming a *jet* of final-state particles. It the decay of the W that adds variety to the top quark final state.

The W may decay into any pair of particles forming one doublet (except of course the kinematically forbidden tb doublet). As the masses of the particles in the other doublets are far less than the W mass, the phase space available for decay into any doublet is nearly equal. Hence the rate of decay into each allowed doublet is identical, and when one takes into account that there are really three sets of quark doublets (since quarks come with three types of color charge), one finds that the branching ratio of a W into the various doublets is:

Final state	Branching Ratio
$e\nu_e$	1/9
$\mu\nu_\mu$	1/9
$\tau\nu_\tau$	1/9
ud	1/3
cs	1/3

This analysis will focus on those decay channels with electrons or muons in the final state. In particular, the *dilepton* (ee, $e\mu$, and $\mu\mu$) and *lepton plus jets* (e+jets and μ + jets) channels will be considered. The $t\bar{t}$ branching ratio for the former totals 4/81 ($\approx 5\%$), and the latter 8/27 ($\approx 30\%$).

The fact that top decays into b is exploited is some of the searches; the fact that a W is present is crucial to the mass measurement. If the top quark mass is greater than the sum of the W and b masses, as found by previous searches[4] [14], then the

[4]The lower limit of 128 GeV/c^2 applies only if the QCD prediction for the cross-section is correct, and the top quark decays according to Standard Model expectations. Since one of the motivations

final-state W will be on-shell. This provides two additional kinematic constraints which may be used in the reconstruction of the $t\bar{t}$ decay system.

2.4 Definition of the Top Quark Mass

As mentioned earlier, in order to obtain finite predictions from higher-order terms in perturbation theory, a renormalization scheme must be applied. In is often convenient in these procedures to treat the masses of quarks and leptons as *running* parameters, whose value depends on the momentum scale of the calculation. The quantity relevant to experiment, however, is the physical mass given by the pole in the quark propagator, and it is this mass which is measured in this dissertation [15]. The relation between the two definitions of the top quark mass in the commonly-used modified minimal subtraction ($\overline{MS}$) scheme is given by [16, 17]:

$$\frac{m_t^{\mathrm{Pole}}}{m_t^{\overline{MS}}\left(m_t^{\mathrm{Pole}}\right)} = 1 + \frac{4}{3}\left(\frac{\alpha_s(m_t^{\mathrm{Pole}})}{\pi}\right) + 10.91\left(\frac{\alpha_s(m_t^{\mathrm{Pole}})}{\pi}\right)^2 \tag{2.2}$$

This ratio is ≈ 1.06 if $m_t^{\mathrm{Pole}} = 170$ GeV/c^2. All references to the top quark mass in the remainder of this work refer to the scale-independent pole mass.

2.5 The Role of the Top Quark Mass

One motivation for measuring the top quark's mass, as detailed above, is to test the QCD predictions for its production. In addition, the top quark mass can provide information about the Higgs sector.

In order to see how this comes about, we begin with Eq. 2.1. The W mass is also related to the *Fermi constant* G_F, which is the effective weak coupling strength at low energy:

for measuring the top quark mass is to test these predictions, one should not assume that this limit holds. Nonetheless, it will be assumed that $m_t > m_W + m_b$, or about 85 GeV/c^2.

$$m_W = \frac{1}{2}\frac{g_1}{\left(\sqrt{2}G_F\right)^{\frac{1}{2}}}$$

In the SU(2) $\times$ U(1) model, g_1 is also related to the fine-structure constant α, yielding:

$$m_W = \frac{1}{2}\left(\frac{\pi\alpha}{\sqrt{2}G_F}\right)^{\frac{1}{2}}\frac{1}{\sin\theta_W} \equiv \frac{A}{\sin\theta_W}$$

which implies:

$$m_Z = \frac{A}{\sin\theta_W\cos\theta_W}$$

All of the above holds exactly at lowest order in the standard model. However, higher-order effects (or *radiative corrections*) cause slight modifications to the above equations. Since m_W is known less precisely than m_Z, α, G_F and $\sin^2\theta_W$, it is convenient to absorb the contributions of these higher-order effects into the equation for m_W:

$$m_W = \frac{A}{\sin\theta_W\sqrt{1-\Delta R}} \tag{2.3}$$

where ΔR represents the radiative corrections.

The radiative corrections to the W mass enter due to fermion and Higgs boson loops in the W propagator. The sizes of the corrections depend on the masses of the particles in the loops. The contributions from the fermion loops are proportional to $\frac{m_f^2}{m_W^2}$, and hence the top quark provides the dominant term. Therefore, if one measures both m_t and m_W very precisely, Eq. 2.3 provides a constraint on the allowed Higgs boson mass m_H.

Unfortunately, the Higgs-loop contribution to ΔR is proportional to $\ln\left(\frac{m_H}{m_W}\right)$, meaning that one needs to measure m_W and m_t very precisely in order to narrow down the allowed range for m_H (See Fig. 2.3.).

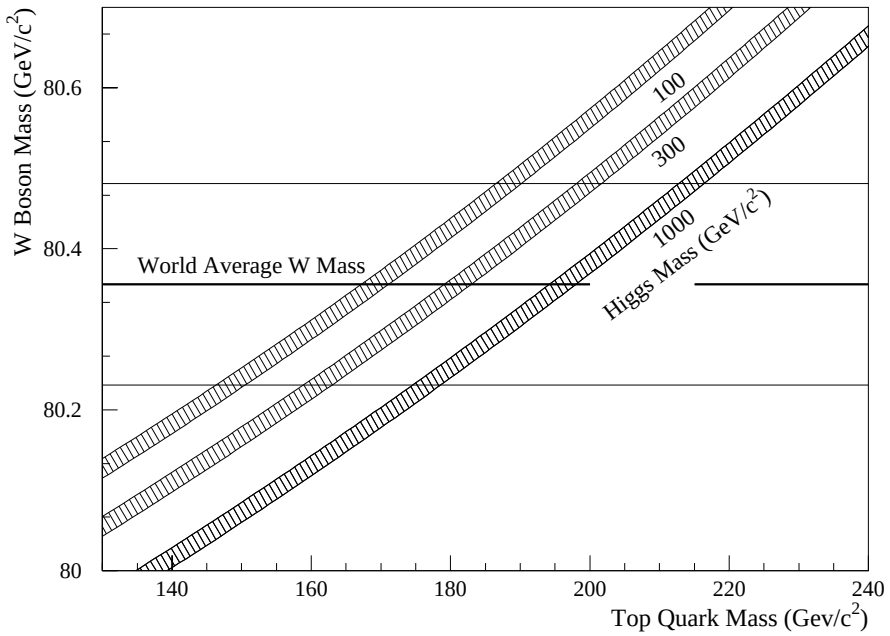

Figure 2.3: Family of curves showing the variation of m_W as a function of m_t for various Higgs masses [18]. The width of the curves is due to uncertainty in the measurement of other Standard Model parameters, and the horizontal lines show the result of recent measurements of the W mass and the one standard deviation limits [19].

2.6 Beyond the Standard Model

While there is little doubt that the Standard Model is a correct theory for the energy range over which it has been tested, there is a general consensus that the model is incomplete. This sense of incompleteness arises from the large number of parameters which must be input into the Standard Model, such as the masses of the fermions (or equivalently the magnitude of their couplings to the Higgs field), the strengths of the various interactions, and the parameters needed to define the mixing of quark states in the CKM matrix. In addition, there are some phenomena, such as CP noninvariance, which can be accommodated by the Standard Model but are not predicted by it. Perhaps most significantly, a theory with the simplest version of the Higgs field appears to be inconsistent. The couplings of this field to itself will in general produce infinities in the perturbative calculation which cannot be removed by renormalization. An additional symmetry, called *supersymmetry* has been proposed to provide a natural cancellation of these infinite terms. However, as yet there is no experimental evidence that this symmetry exists.

2.7 Top Quark Measurements

The two experiments which simultaneously discovered unambiguous evidence for top quark production, DØ and CDF, both located at Fermilab's Tevatron ring, remain responsible for all direct measurements of the top quark's mass and production cross section. While the rarity of $t\bar{t}$ production means that the sample sizes are still small (totalling around one hundred events for the two experiments), these properties are already fairly well measured.

The top quark mass presently measured by DØ using lepton plus jets events is

$$m_t = 173.3 \pm 5.6 \text{ (stat.)} \pm 6.2 \text{ (syst.)} \text{ GeV/c}^2 \tag{2.4}$$

and using dilepton events:

$$m_t = 169.9 \pm 14.8 \text{ (stat.)} \pm 3.8 \text{ (syst.) GeV/c}^2 \tag{2.5}$$

A combination of the above measurements yields:

$$m_t = 173.1 \pm 5.2 \text{ (stat.)} \pm 5.7 \text{ (syst.) GeV/c}^2 \tag{2.6}$$

The cross section is found to be (for $m_t = 170$ GeV/c^2)

$$\sigma\left(t\bar{t}\right) = 5.7 \pm 1.9 \text{ pb.} \tag{2.7}$$

The CDF collaboration [20] measures the top quark mass as

$$m_t = 176.8 \pm 4.4 \text{ (stat.)} \ \pm 4.8 \text{ (syst.) GeV/c}^2 \tag{2.8}$$

and the cross section (assuming a mass of 175 GeV/c^2) as

$$\sigma\left(t\bar{t}\right) = 7.5^{+1.9}_{-1.6} \text{ pb} \tag{2.9}$$

As both experiments are now dormant while the detectors undergo upgrades, slightly refined versions of the above numbers will represent the limit of our knowledge of the top quark for the next few years. The analyses which result in the top quark mass measurement from DØ will be described in this dissertation.

Chapter 3

Experimental Apparatus

The production of top quark pairs requires a large center-of-mass energy, and therefore a colliding-beam experiment is necessary. The preference of proton beams over electron beams for this purpose comes from consideration of the synchrotron radiation emitted by any accelerating charged particle. The energy dissipated by synchrotron radiation decreases as the fourth power of the mass of the accelerated particle, and hence it is far easier to accelerate proton beams to the needed energy. The drawback is that protons are themselves complex objects, comprised of quarks and gluons, which complicates the analysis of the collisions and results in only some fraction of the total proton energy being delivered to any particular collision.

One way of implementing a colliding-beam experiment is to collide beams of a particle with beams of its antiparticle. As the antiparticle shares all the characteristics of the particle but has opposite electric charge, the two beams will circulate in opposite directions in the same ring of magnets. This eliminates the need to construct a separate accelerating apparatus for each beam. Hence, the experiment described here studies proton-antiproton ($\overline{p}p$) collisions.

The detection of top quark pairs requires a detector capable of identifying and measuring the energies of electrons, muons, jets, and neutrinos. This chapter provides an introduction the both the TeVatron accelerator and DØ detector which collected

the data used for this analysis. A good reference for the former is [21], and the official reference for the latter is [22].

3.1 TeVatron at Fermilab

A total of seven acceleration devices are used to produce the colliding proton and antiproton beams, and their layout is shown in Fig. 3.1. The system is necessarily complex due to the fact that different acceleration techniques and device parameters are needed for the various energy regimes the protons pass through on the way to their final energy of 0.9 TeV.

The proton beam begins with 18 keV H^- ions, which are accelerated to 750 keV by a Cockroft-Walton electrostatic generator. While such electrostatic potential differences are the most straightforward method of particle acceleration, they are only applicable up to the relatively small potential differences which can be obtained before arcing occurs between the electrodes.

To reach higher energies, one must employ an alternating electric field, typically provided by a radio-frequency (RF) cavity. Merely introducing a beam of particles into such a cavity would be unproductive, as both the accelerating and deaccelerating phases of the field would act on the particles. However, the geometry may be arranged such that the particles are inside of a conducting shield during the deaccelerating phase, and exposed to the field only during its accelerating phase. The simplest example of such an arrangement is the *linear accelerator* in which the conductors are a series of cylinders arranged end-to-end. By carefully designing the frequency of the field and the lengths of the conductors, one can ensure that the field is always accelerating when the particles traverse the gap between cylinders. The ions from the Cockroft-Walton generator are passed to such a device (the Linac), which increases their energy to 400 MeV.

The energy reach of a linear accelerator is limited by the length of the device. In order to conserve real estate and the amount of material involved in building

an accelerator, it is advantageous to arrange the series of conductors and gaps in a circle, with a magnetic field used to bend the particles. The first example of this sort of device was the *cyclotron*, in which the conductors are separate hollow D-shaped electrodes separated by a single gap. The magnetic field permeates the entire region of the accelerator, and the particles travel in circular orbits, receiving a push at each traversal of the gap. The frequency of the orbit is constant as a function of energy (in the nonrelativistic regime), but the radius of the orbit increases as the particle accelerates.

Alternatively, one can design a device (known as a *synchrotron*) in which the particle's orbit radius remains constant. This has the advantage that the magnetic field need be applied only at the radius of the particle trajectory, allowing the construction of much larger and more energetic accelerators. Such a design is more complex than that of a cyclotron, however, since both the frequency of the accelerating potential and the magnetic field strength must be increased in concert as the particle gains energy. Also, there is a limit to the range of fields which can be achieved by a given set of electromagnets, which implies that a synchrotron has a limited dynamic range. Due to this, three synchrotrons are used at Fermilab to accelerate the 400 MeV ions from the Linac up to 0.9 TeV.

The first of these, the Booster, has a diameter of 150 m and accelerates the protons to 8 GeV (the electrons from the H^- ions are stripped before injection into this accelerator). Next the protons pass to the Main Ring, a 1000 m diameter synchrotron which shares the same tunnel as the TeVatron. While this accelerator is capable of reaching energies of 400 GeV, it needs only to accelerate the protons to 150 GeV for insertion into the TeVatron.

Besides its duties as the injector for the TeVatron, the Main Ring also provides a beam of 120 GeV protons which strike a nickel target to produce antiprotons. The target is optimized for producing antiprotons with an energy spectrum peaked at 8 GeV (the Main Ring injection energy), with an efficiency of about 10^{-5} antiprotons per incident proton. As they are produced, the antiprotons have a spread in phase

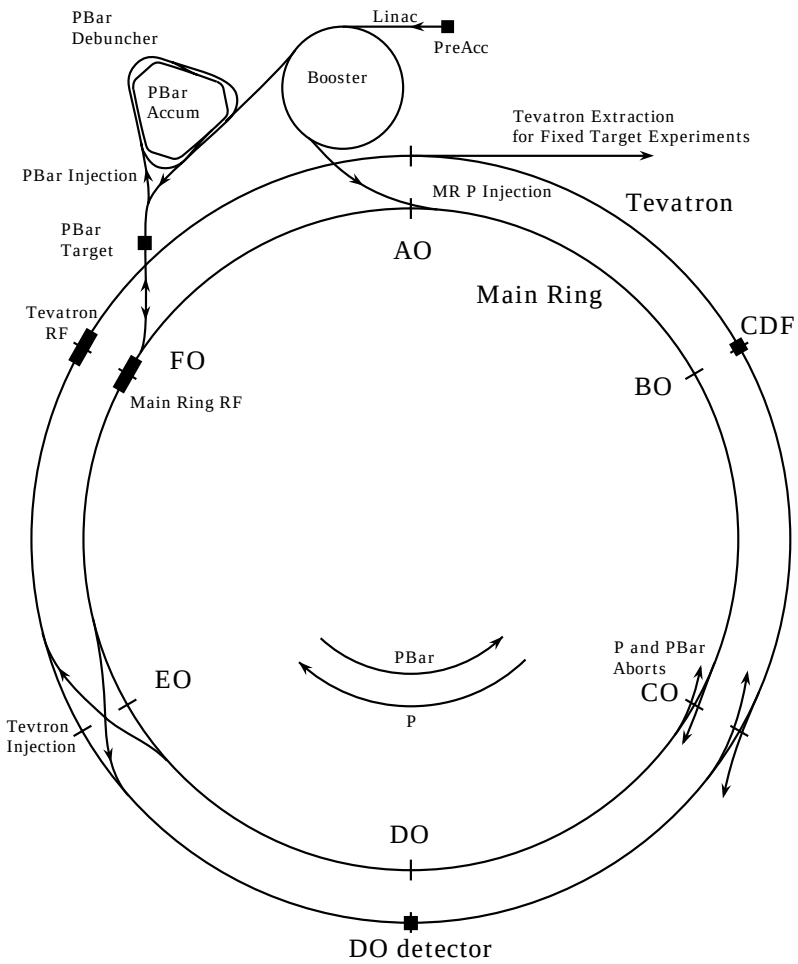

Figure 3.1: The Fermilab accelerator complex. (Not to scale – the Main Ring and TeVatron have the same radius, and are separated here for clarity).

space determined by the dynamics of the p-Ni collisions. In order for the efficiency of transfer into the Main Ring to be acceptably high, the beam must first be cooled, a process which takes place in the Debuncher and Accumulator. The Debuncher modifies the bursts of input antiprotons (which are produced in pulses corresponding to the radio frequency structure of the Main Ring) into a continuous band with a lower momentum spread. It also begins the process of *stochastic cooling*, in which deviations from the ideal orbit are measured and correction signals applied.

Antiprotons may reside in the Debuncher for only 2.4 s, after which another batch from the source comes in. The Accumulator is designed for long-term storage and continued cooling of the antiprotons. After their time in the Debuncher, the antiprotons are transferred to the Accumulator, where they are slowly merged with the stack of previously generated antiprotons and further cooled. Once the stack of antiprotons reaches a suitable size, the antiprotons are transferred to the Main Ring for injection into the TeVatron.

During typical operation, stacking of antiprotons occurs while collisions are taking place in the TeVatron, to ensure that a fresh batch of antiprotons is ready for the next injection cycle. The fact that the Main Ring is active during data-taking has non-trivial consequences for the DØ detector, which are detailed below.

The TeVatron accepts bunches from the Main Ring, and uses superconducting magnets to confine them in a circle while accelerating them to 0.9 TeV, providing 1.8 TeV of center-of-mass energy. Some of the important TeVatron parameters are given in Table 3.1. For colliding-beam operations, six bunches each of protons and antiprotons circulate in the machine. For most of the circumference, electrostatic separators keep the beams apart, but at the BØ and DØ regions, special focussing magnets reduce the transverse beam sizes to about 50 μm, providing high-luminosity collisions. (Equivalently, one may say that the magnets reduce β^*, the local wavelength of betatron oscillations about the ideal orbit at the interaction region. Typically, $\beta^* \approx 25$ cm is achieved.) The peak luminosity reached so far is $\approx 3 \times 10^{31}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The number of bunches and the size of the accelerator imply that collisions occur

Radius	1000 m
Number of dipole magnets	774
Number of quadrupole magnets	216
Number of proton bunches	6
Number of protons/bunch	$\approx 2 \times 10^{11}$
Number of antiproton bunches	6
Number of antiprotons/bunch	$\approx 7 \times 10^{10}$
Center-of-mass energy	1.8 TeV
Maximum luminosity	$\approx 3.0 \times 10^{31} \mathrm{cm}^{-2}\mathrm{s}^{-1}$

Table 3.1: TeVatron parameters.

once every 3.5 μsec, and this defines the time frame needed for the trigger system described below. Also, while the transverse size of the bunches is small at the luminous regions, the longitudinal size is not, and the distribution of collisions along the beam direction is roughly Gaussian with $\sigma \approx 30$ cm.

Once the collisions begin, the beams continue to circulate in the machine for several (10 – 20) hours, until their intensity decreases to the point where collision rates become too low, and a fresh set of bunches is injected. The time in which the bunches are kept in the TeVatron is called a *store*.

3.2 DØ Detector

As detailed in the previous chapter, the final state from $t\bar{t}$ decay may contain electrons, muons, jets, and neutrinos. The DØ detector is designed to identify and measure the energy of all of these objects. As shown in Fig. 3.2, the detector has three major subsystems: the tracking chambers, calorimeter, and muon system.

The detector design was optimized for high-resolution, hermetic calorimetry, which provides the sole measurement of the energies of electrons and jets. The inner tracking volume as a consequence is relatively small, and there is no magnetic field in the

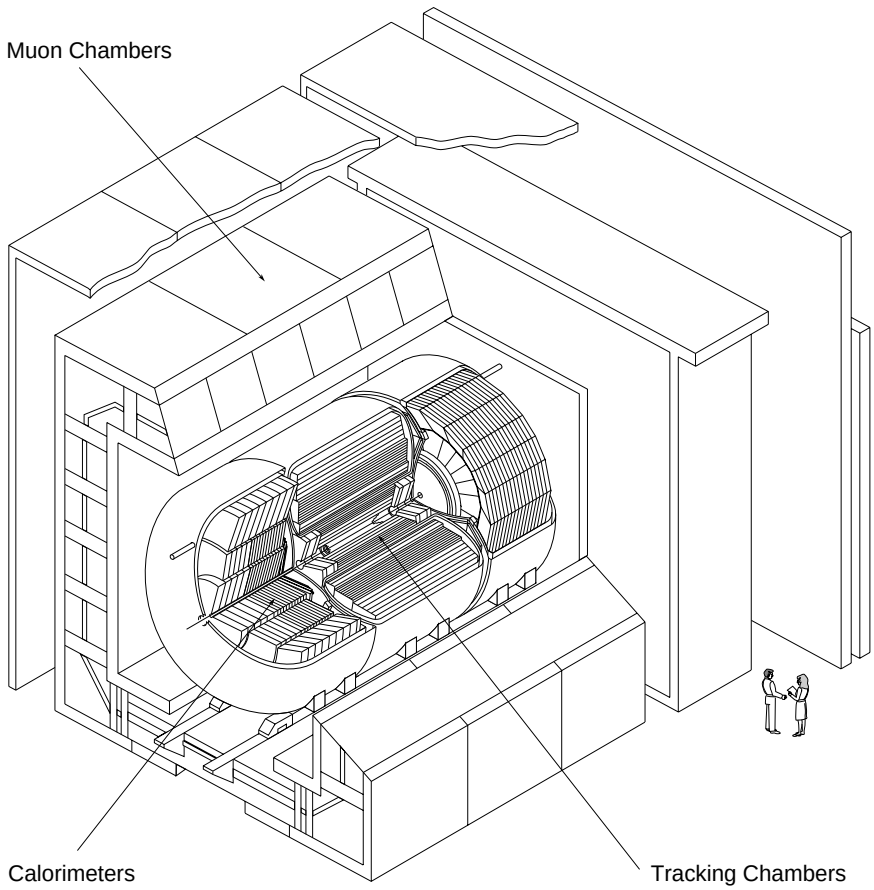

Figure 3.2: Cutaway view of the DØ detector, showing the tracking chambers, calorimetry, and muon system.

tracking region.

In discussing the positions of objects within the detector, a right-handed coordinate system is used, with the origin at the center of the detector, positive z-axis along the proton direction and y-axis upward. Due to the approximate cylindrical symmetry of the detector, it is also convenient to use the variables r (the perpendicular distance from the beamline), ϕ (the azimuthal angle with respect to the positive x-axis), and θ (the polar angle with respect to the positive z-axis). The polar direction is also often described by the *pseudorapidity* (η), defined as $\eta \equiv -\ln\tan(\theta/2)$.

3.2.1 Central Detector

The central detector systems are designed to measure the trajectories of charged particles, and also to aid in the identification of electrons. A set of drift chambers performs the former function, while the latter is handled by the drift chambers and a transition radiation detector.

3.2.2 Basics of Drift Chamber Operation

Drift chambers are designed to detect the ionization liberated by a charged particle passing through a gas-filled region. The total number of ionizations produced is given by $n_T = \frac{\Delta E}{W_i}$ where ΔE is the total energy lost by the particle and W_i is the ionization potential of the gas, and is typically of the order of 10 - 100 per cm of gas traversed. These are the *primary ionization* electrons.

If an electric field is applied across the gas region, the free electrons drift toward the anode. The energy gained by the electrons drifting in the electric field quickly comes into equilibrium with that lost due to collisions with atoms, and the drift velocity, on average, is constant[1].

[1]It is important to note that while the average drift velocity of a cluster of electrons will be constant, the drift velocities of the electrons within that cluster will vary due to the statistical nature of the collision processes. The diffusion in the spatial extent of the cluster that results gives a lower bound on the resolution attainable with a drift chamber.

While the drifting electrons will eventually reach the anode, their number is far too small to produce an observable signal. However, if the anode is a thin (typically 20 μm diameter) wire, a large electric field is created near the wire. If this field is large enough (10^4 - 10^5 V/cm) the energy gained by the electrons between atomic collisions will exceed W_i, meaning that they can liberate additional electrons, which can in turn do the same, creating an avalanche of secondary ionizations. The number of secondary ionizations is typically 10^4 - 10^6 times the number of primary ionizations (the ratio is called the *gas gain*), and is sufficient to form an observable signal.

The gas gain generally increases with increasing electric field. For moderate fields, the gain is independent of n_T, and so the signal is directly proportional to this number (the *proportional* regime). Since n_T is in turn proportional to the energy lost by the ionizing particle, such a detector allows one to measure dE/dx and hence to aid in particle identification.

As gain increases, the cloud of slow-moving positive ions surrounding the anode presents an increasing coulomb screen for additional drifting electrons. This causes a gradual loss of proportionality and culminates in the transition to the *saturated* mode of operation, in which the signal size is nearly independent of the amount of primary ionization. While this allows the largest possible signal, no measurement of dE/dx can be made.

The above description holds for any sort of ionization counter. The unique feature of a *drift chamber* is the use of a measurement of the signal time to measure the distance of the ionizing particle from the anode wire. The relation of time to distance is given by:

$$z = \int_{t_o}^{t_1} v_d(t)dt$$

If the chamber is constructed such that $v_d(t)$ is a constant, this simplifies to $z = v_d(t_1 - t_o)$. This situation will obtain if the electric field is uniform across most of the chamber (clearly this cannot be true very near the anode). To facilitate this, relatively thick field-shaping wires are placed between the anode wires.

The choice of the gas used in the drift chamber is subject to several constraints. Noble gasses are favored for two reasons: first, since ionization is the only means for dissipating energy, avalanche multiplication occurs at lower fields than it would for other gasses. Second, there is no probability that a noble gas will attach a drifting primary ionization electron and thereby attenuate the signal. Therefore, the primary gas is generally a noble gas, with argon a common choice.

However, using pure argon is not feasible. This is because the photons emitted by de-exciting atoms are energetic enough to liberate electrons from the metal cathode (positive ions striking the cathode may also do this). These electrons then create their own avalanches, and at relatively low fields (corresponding to a gain of $10^3 - 10^4$) this cycle becomes self perpetuating and the chamber is continuously "on". To allow higher gain operation, a polyatomic gas called a *quencher* is added. The many degrees of freedom available in such a molecule allow it to absorb the photons emitted by de-exciting argon atoms without liberating any electrons, and hence interrupt the above chain reaction.

The drawback to adding a complex molecule is that it can break up into simpler molecules which form polymers on the anode and cathode surfaces, degrading the chamber performance. Often a third component is added to the gas mixture to break up these polymers and prolong the useful lifetime of the chamber.

Chamber readout is generally done by differentiating the signal. The time development of the signal generated by the ions[2] is such that most of the signal develops rather quickly, with the majority of the signal in the first μsec or so of drift. If the time constant of the differentiating circuit is less than this, a sharp pulse is produced.

For further details on drift chamber principles and implementation, see [23, 24, 25].

[2] Although it is the creation of the electron avalanche that initiates the signal, the electrons are produced so close to the anode that there is little change in the energy of the system as the electrons drift the final distance to the wire. Therefore, the bulk of the signal is created by the drift of the residual ions toward the cathodes. The time development of the signal thus reflects the relatively slow-moving ions.

Vertex Drift Chamber (VTX)

The component of the DØ detector nearest the proton-antiproton interaction region is the Vertex Drift Chamber (VTX), which is composed of three cylindrical layers of drift chambers covering the region $|z| < 116$ cm and 3.7 cm $< r <$ 16.2 cm. The need to detect tracks (and resolve two tracks which lie near each other) this close to the interaction region places stringent demands on the resolution of the chamber. In order to meet these demands, a gas which has a low electron diffusion constant and slow drift velocity is needed. For the VTX, the mixture CO_2 (95%) ethane (5%) was chosen, with a small admixture of water as a cleansing agent. In order to achieve the required low diffusion constants, the gas is operated in a voltage regime such that the electron drift velocity is proportional to the drift field. Clearly, this places stringent requirements on the design of the field-shaping components of the detector. As implemented, the drift velocity was 7.3 μm/ns.

The inner layer is divided into 16 azimuthal readout cells, while the outer two layers have 32 cells (see Fig. 3.3). Each cell contains eight 25 μm diameter NiCoTin sense wires running parallel to the z-axis to provide the $r-\phi$ measurement. The field is such that the ions drift toward the sense wires in the azimuthal direction. The time of the hit allows one to determine how far the primary ionization is from the sense wire, but carries no information about which side of the wire the particle traversed (this is the *left-right ambiguity*). In order to resolve this, adjacent sense wires are staggered by $\pm 100\mu$m. Tests have demonstrated that the $r-\phi$ resolution for signal hits in the VTX varies as a function of drift distance between 30 – 60 μm, with 50 μm being a typical value.

Two grounded grid wires near each sense wire combine with the cathodes to shape the drift field in the cell. The cathode consists of aluminum traces on the inner surface of the support cylinder (coarse field-shaping) and a cage of 152 μm diameter gold-plated aluminum wires around the edges of the cell (fine field-shaping). Note that the geometry of the cell is such that the distance between the cathode surface

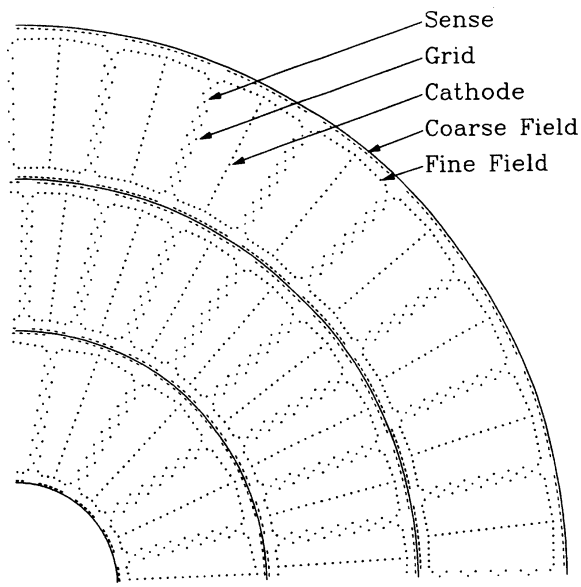

Figure 3.3: End view of one quarter of the Vertex Drift Chamber.

and anode wires increases as one moves radially outward from the beamline. In order to maintain a constant electric field in the drift region, a slightly higher voltage is applied to the cathodes at larger radii.

The fact that the NiCoTin sense wires have a resistance of 1.8 kΩ/m allows a measurement of the z position of the hit to be done by reading out both ends of the wire and comparing the signal sizes. This has the advantage of allowing a z measurement for every hit on the track, at the cost of requiring two amplifiers for each wire. A resolution of 1% of the wire length has been achieved in tests.

For more information on the VTX, see Table 3.2 and [26, 27, 28, 29].

Central Drift Chamber (CDC)

The Central Drift Chamber consists of four concentric cylinders, covering the range $|z| < 89.7$ cm, and 51.8 cm $< r <$ 71.9 cm. Each layer is divided into 32 identical azimuthal sectors, as shown in Fig. 3.4. The gas used is a mixture of argon (92.5%), methane (4%), carbon dioxide (3%) and water (0.5%). Within each sector there are

Maximum length	116.8 cm
Radial interval	3.7 - 16.2 cm
Radial wire interval	4.57 mm
Maximum drift distance	16 mm
Sense wires per cell	8
Gas used	CO_2 (95%) ethane (5%)
Drift field	2.3 kV/cm
Sense wire potential	+2.5 kV
Sense wire type	25 μm NiCoTin
Field wire type	152 μm gold-plated Al

Table 3.2: Vertex Drift Chamber parameters [26].

seven 30 μm anode wires running parallel to the beam direction. Adjacent wires are staggered by $\pm$ 200 μm in ϕ order to resolve the left-right ambiguity in the track position. There are two field wires associated with each anode wire to help produce a uniform electric field across the sector.

The drift field is about 620 V/cm, leading to an electron drift velocity of 34 μm/ns. As in the VTX, the cathode voltage must increase as a function of radius in order to keep the drift field constant. To achieve this, resistive strips are printed on the cathode surfaces, allowing the cathode voltage to vary with position.

The maximum drift distance is 7 cm, and the resolution in $r - \phi$ for a single wire varies from 150 - 200 μm as a function of the drift distance.

Measurement of the z coordinate of a track is performed by inductive delay lines, which are embedded in the cylinder walls before and after the line of sense wires. These lines are composed of a wire wrapped on a carbon-fiber/epoxy core. Signals propagate at 2.4μm/ns along the wires, so that reading out the delay lines at both ends and noting the time difference between the signals allows a measurement of z with a resolution of about 4mm.

For further details on the CDC, see Table 3.3 and [22, 26].

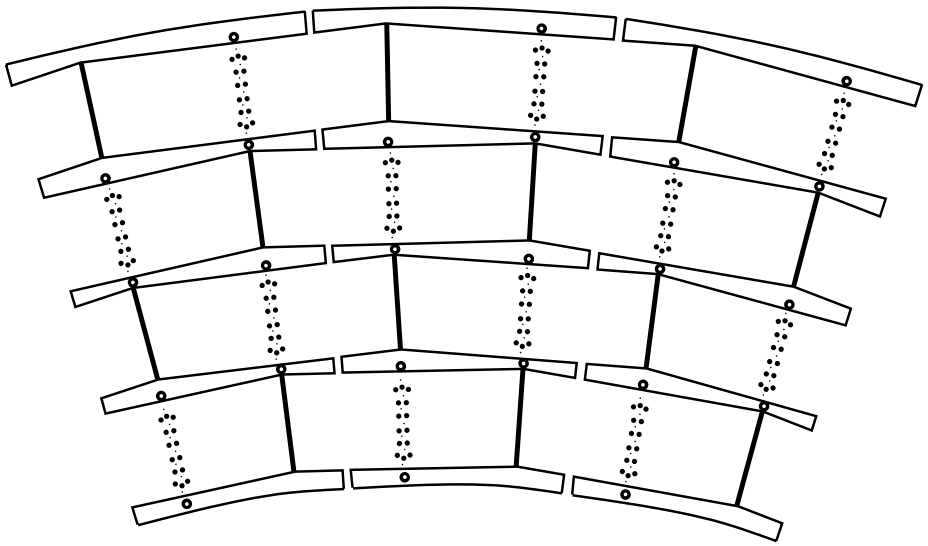

Figure 3.4: Layout of wires and cells in the Central Drift Chamber. Each cell contains seven signal wires (small points), fourteen field wires (large points) and two delay lines (open circles).

Length	179.4 cm
Radial interval	51.8 - 71.9 cm
Radial wire interval	6 mm
Maximum drift distance	7 cm
Sense wires per cell	7
Gas used	argon(92.5%) methane (4%) CO_2 (3%)
Drift field	650 V/cm
Sense wire potential	+1.5 kV
Sense wire type	30 μm gold-plated tungsten
Field wire type	127 μm

Table 3.3: Central Drift Chamber parameters [26].

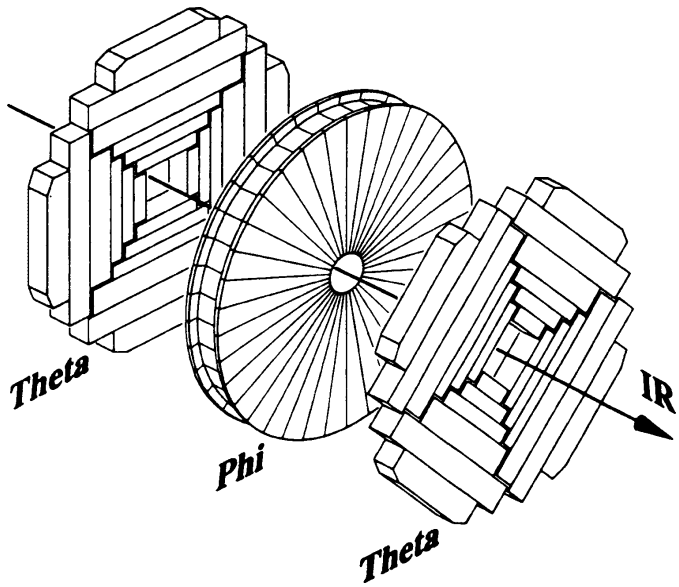

Figure 3.5: Exploded view showing the orientation of the Forward Drift Chambers.

Forward Drift Chambers (FDC)

In order to extend the tracking coverage out to $|\eta| \approx 3$, a set of Forward Drift Chambers (FDCs) are installed at each end of the cylinder defined by the VTX and CDC. The construction and operation of these chambers are similar to those of the CDC, and the gas used is identical.

Each set of FDCs is composed of three chambers: one Φ chamber between two Θ chambers. The Φ chamber has sense wires running in the radial direction to give a measurement of ϕ, while the Θ chambers have sense wires aligned (approximately) in the ϕ direction to give a measurement of θ. See Fig. 3.5 for the positioning of the FDC chambers and the signal wire directions.

The Φ module is divided into 36 azimuthal sectors, each of which contains 16 layers (in z) of 50 cm long sense wires. A single grounded field wire between each pair of sense wires, and aluminum cathode traces etched onto the cell walls, provide the field-shaping elements for the cell.

z interval	104.8 cm - 111.2 cm and 128.8 cm - 135.2 cm
Radial interval	11 - 62 cm
z wire interval	8 mm
Maximum drift distance	5.3 cm
Sense wires per cell	8
Gas used	argon(92.5%) methane (4%) CO_2 (3%)
Drift field	1.0 kV/cm
Sense wire potential	+1.5 kV
Sense wire radius	30 μm
Field wire radius	163 μm

Table 3.4: Forward Drift Chamber Θ module parameters [26].

The Θ modules are made up of four quadrants, each containing six rectangular drift cells. Each of these cells contains 8 layers of sense wires as well as one delay line (similar to those used in the CDC) to provide a measurement of the ϕ position of each hit, thus aiding the pattern recognition. The electrostatic configuration of these cells is more similar to that of the CDC, including having two field wires between each pair of sense wires. In order to further reduce ambiguities, the two Θ modules on each side of the detector are rotated in ϕ by 45° with respect to each other.

The sense wires on the inner three cells of the Θ chambers are located at one edge of the cell to eliminate left-right ambiguities. Wires in adjacent layers of all other cells in the Φ and Θ modules are staggered by $\pm$ 200 μm to resolve these ambiguities. The maximum drift distance in any cell is 5.3 cm.

Not surprisingly, the performance of the FDC is quite similar to that of the CDC. Single-hit resolutions are about 200 μm in the direction measured by drift time (the Θ chamber delay lines have a resolution of about 4 mm) and two tracks separated by 2 mm are resolved with 90% efficiency.

Further details on the Φ and Θ chambers are given in Tables 3.4 and 3.5, and in [22, 26].

z interval	113.0 cm - 127.0 cm
Radial interval	11 - 61.3 cm
z wire interval	8 mm
Maximum drift distance	5.3 cm
Sense wires per cell	16
Gas used	argon(92.5%) methane (4%) CO_2 (3%)
Drift field	1.0 kV/cm
Sense wire potential	+1.5 kV
Sense wire radius	30 μm
Field wire radius	163 μm

Table 3.5: Forward Drift Chamber Φ module parameters [26].

Transition Radiation Detector (TRD)

In order to provide additional electron identification ability, a Transition Radiation Detector (TRD) is installed between the VTX and CDC. This operates on the principle that a charged particle emits radiation when passing between media with different indices of refraction (for a discussion of transition radiation, see [30]). For a relativistic particle, this radiation is emitted in a narrow cone ($\theta \approx 1/\gamma$, where $\gamma \equiv E/(mc^2)$) around the particle's trajectory. The intensity of the radiation is proportional to γ, and reaches appreciable values only for $\gamma \approx 10^3$. As electrons and positrons are the only charged particles produced at the TeVatron which typically have such large γ factors, measuring the transition radiation allows one to discriminate between electrons and charged hadrons.

The TRD is made up of three cylindrical units, each 10.5 cm thick and containing a stack of polypropylene foils used as a radiator, and a proportional wire chamber (PWC) used for detecting the radiation (see Fig. 3.6). A stack of foils is necessary since the probability of radiation being emitted at any given transition is proportional to α, so a few hundred transitions are needed to ensure appreciable radiation. Also, judicious choice of the thickness of the foils and the gap between them allows one to

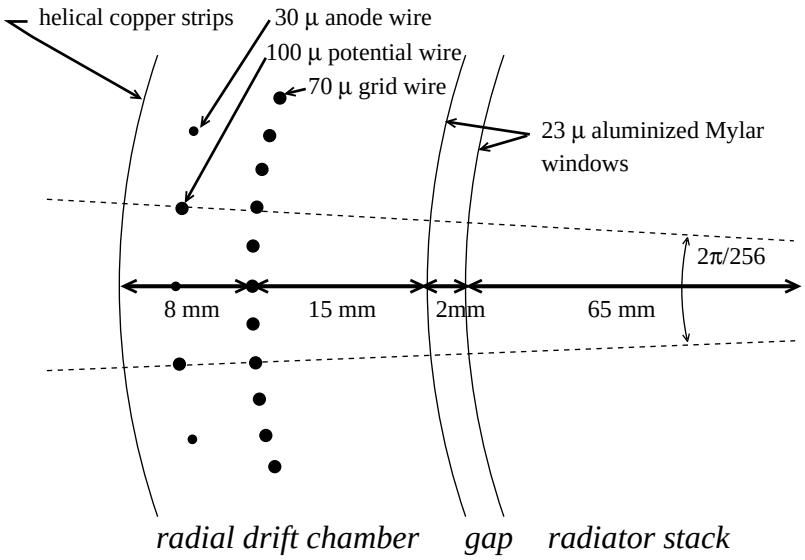

Figure 3.6: One sector of one module of the Transition Radiation Detector.

use interference effects to tailor the spectrum of the emitted radiation. Three hundred ninety-three foils, each 18 μm thick, are spaced with an average gap of 150 μm to produce a transition radiation spectrum which peaks at 8 keV.

The proportional wire chamber is divided into two sections. In the inner 15 mm region, X-rays convert to electrons (xenon is chosen as the primary gas for the PWC to ensure a high probability of conversion). After conversion, the electrons drift radially past a row of grid wires and into the detection region. Each drift cell in this region is approximately square, with dimension 8 mm $\times$ 8 mm. The drift field is radial, and the sense and field wires are parallel to the z axis.

While all charged particles will deposit energy in the PWC, electrons can be distinguished by both the magnitude and timing of the deposited charge. The magnitude will be greater both due to the presence of transition radiation and the fact that the more relativistic electrons will have somewhat larger dE/dx than charged hadrons.

The difference in timing reflects the fact that the transition radiation X-rays generally convert in the first few mm of the inner section of the PWC, so the energy from

them tends to be deposited at long drift times. For a particle producing ionization as it traverses the chamber, the energy deposited is a constant for all drift times (to the extent that the drift field is uniform).

For further information on the TRD, see [31, 32].

Central Detector Electronics

The VTX, TRD, FDC, and CDC all employ similar readout electronics. The first stage of the readout is performed by preamplifiers mounted on the detector surfaces. Besides handling the output signals during data taking, the preamplifiers are also able to inject test charges onto the sense wires in order to calibrate the entire electronics chain.

The signals then are carried about 15 m on coaxial cable to the platform beneath the detector, where they are processed by shaper circuitry which removes the long tails due to ion drift. The resultant pulse is more symmetric about its peak, and more optimal for resolving double hits [33]. After shaping the signals are carried about 45 m to the movable counting house (MCH) for digitization.

Digitization is done by flash analog-to-digital converters (FADCs) , which have an 8-bit dynamic range and operate at 106 MHz. The speed is necessary to allow two-hit separation down to small distances (and thus small signal time differences), while one desires an even larger dynamic range to allow the best possible measurement of dE/dx. In order to accomplish this, the signals enter an analog buffer before the FADC. This buffer applies an adjustable gain correction to the signal (such that small signals are amplified a factor of 8.5 less than large signals), which in effect increases the dynamic range of the digitization circuitry to 9.5 bits.

A total of 6080 channels are instrumented for the entire central detector. If one were to attempt to read out every digitization cycle for every channel for every event, data rates on the order of 325 Mbyte/s would ensue, which would clearly overwhelm the data acquisition system. In order to reduce the data set to a manageable size,

the FADCs are followed by zero suppression circuitry, which compares both the size of the signal for each digitization cycle and the difference in signal between adjacent cycles to programmable thresholds in order to define the leading and trailing edges of a signal. Only the digitization cycles lying between these edges are retained for further processing.

3.2.3 Calorimeter

The DØ detector was designed to achieve good resolution in the energies of electrons, photons and jets. Since there is no magnetic spectrometer in the central region of the detector this measurement is provided solely by calorimetry, in which the incident particles are stopped and the energy dissipated is measured.

Electrons and photons interact with material in substantially different ways than do hadrons, and thus the types of calorimeter best suited for measuring their energies are different. In DØ the two functions are carried out in separate modules. The following sections describe these types of calorimeters in general, as well as their implementation at DØ. The general descriptions are based on [34, 24, 25].

Electromagnetic Modules

The layers of the calorimeter closest to the interaction point are optimized for photon and electron measurement, and are referred to as the electromagnetic (EM) calorimeter. The operating principle of this calorimeter is that electrons (photons) with energy greater then $\approx$ 10 MeV dissipate their energy predominantly through bremsstrahlung (pair production). Hence an electron with several GeV of energy will radiate an energetic photon, which is turn will produce an electron-positron pair and so forth, creating a shower of secondary electrons and photons (the process is similar for incident photons, except that the first stage of the shower is a pair production). The distance parameter which most conveniently describes the development of this shower is the *radiation length* X_o, defined such that:

$$\left\langle \frac{dE}{dx} \right\rangle_{\text{brems}} = \frac{E}{X_o}$$

The critical energy, at which the energy loss by ionization is on average equal to that by bremsstrahlung, is given approximately by

$$\epsilon_c \approx \frac{580}{Z}(\text{MeV})$$

The mean total track length of ionizing secondary electrons in a electromagnetic shower is given by:

$$T_d \approx \left(\frac{4}{3}X_o + \frac{2}{3}s_o\right)\frac{E}{\epsilon_c}$$

where s_o is the range of electrons which have the critical energy. The proportionality between E and T_d allows measurement of the total ionization to give a measurement of the incident particle energy.

The shower reaches its maximum multiplicity at a distance of $\approx (\ln(E/\epsilon_c) - 1)\, X_o$ in the calorimeter (about 10 X_o for a 100 GeV electron in uranium), so one can see that the amount of material needed scales only as the log of the incident particle energy.

The transverse spread of the shower is determined by the typical angles of bremsstrahlung and multiple scattering, and is parameterized in terms of of the *Moliére radius* $\rho_M \approx 21X_o/\epsilon_c$. About 90% of the shower energy is contained in a cylinder of radius $2\rho_M$.

The accuracy with which a calorimeter can measure the incident energy is limited by the measurement of the total track length. Ideally, the same material would be used for initiating the shower and for measuring the ionization, so that the entire track length is visible. However, to reduce costs and allow a more compact calorimeter volume, it is often preferable to divide the calorimeter into alternating layers of dense (*absorber*) and light (*sampler*) material. Such a system is called a *sampling calorimeter*, since only that portion of the ionization which occurs in the sampling layers is

Density	18.95 g/cm^3
Radiation length (X_o)	6.00 g/cm^2
Nuclear interaction length (λ)	199 g/cm^2
Moliére radius (ρ_M)	$\approx$ 1.1 cm

Table 3.6: Some properties of uranium.

measured. The limit on resolution for such a scheme is determined by statistical fluctuations in the amount of ionization occurring in the sampling layers. Therefore the fractional uncertainty in the energy measurement will scale as one over the square root of the number of ionization tracks in the shower, or equivalently as $E^{-1/2}$.

In the DØ calorimeter, uranium is used as the absorber material and liquid argon as the sampling medium. Some of the important parameters for uranium are given in Table 3.6, and the motivation for its selection will be given in the next section. Liquid argon was chosen as the sampling medium since it allows uniform gain over the entire calorimeter, is relatively simple to calibrate, allows flexibility in the segmentation of the calorimeter into readout cells, and is not susceptible to radiation damage. However, the need to operate the calorimeter at cryogenic temperatures (and therefore to seal the modules inside of a cryostat) imposes constraints on the design of the detector. In order to facilitate construction and allow access to the tracking detectors, the DØ calorimeter modules are divided into three separate cryostats, the central calorimeter (CC) surrounded by two end calorimeters (EC), as shown in Fig. 3.2. The CC covers $|\eta| < 1.1$ and the EC extends this coverage to $|\eta| < 4$, providing the hermeticity needed for good total transverse energy resolution. The uninstrumented material between the CC and EC modules means that this region requires special attention, as detailed below.

The layers of the calorimeter closest to the interaction point are optimized for measurement of electromagnetic showers. The absorbers in these modules are thin (3 mm thick in the CC and 4 mm in the EC) plates of pure depleted uranium. In the

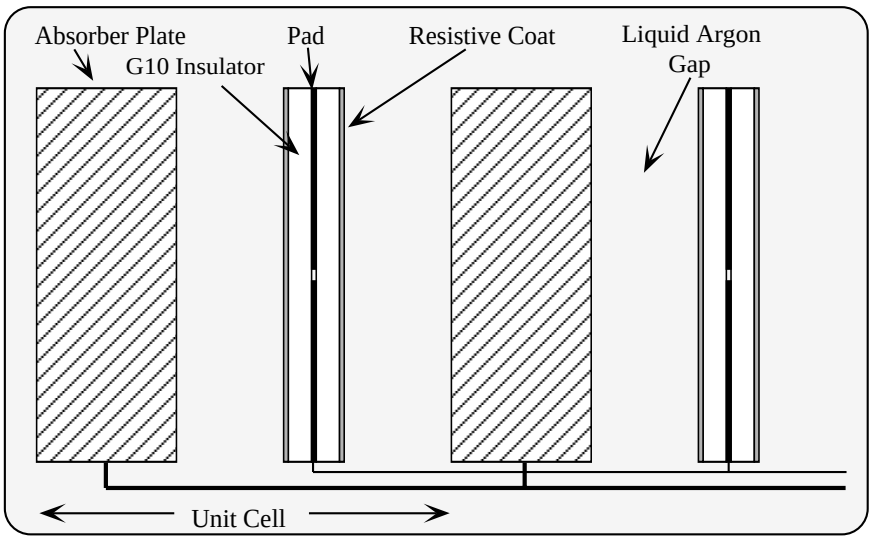

Figure 3.7: Structure of a calorimeter readout cell.

space between adjacent plates there is a signal board surrounded by two 2.3 mm liquid argon gaps to form a sampling cell (see Fig. 3.7). The signal board is a multilayer printed circuit board, the outer surfaces of which are coated with resistive epoxy and connected to positive high voltage. The absorber plates are held at ground, creating a drift field across the liquid argon gap. Signals are collected on copper readout pads in the middle later of the signal boards. The transverse segmentation of the signals is defined by the size of these pads.

Signals from several sampling cells at the same η and ϕ are ganged together in depth to form one layer for readout. The EM calorimeter is divided into four such layers. In the CC the first two are about $2X_o$ thick and measure the beginning of the shower, the third is about $7X_o$ thick and generally contains the shower maximum, and the fourth adds an additional $10X_o$ of material to complete the containment of the shower. In the EC the layers are of similar thickness except for the first, which is only $0.3X_o$ thick to compensate for the larger amount of material (about $2.3X_o$) in front of the calorimeter. The transverse segmentation is $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$ for layers 1,

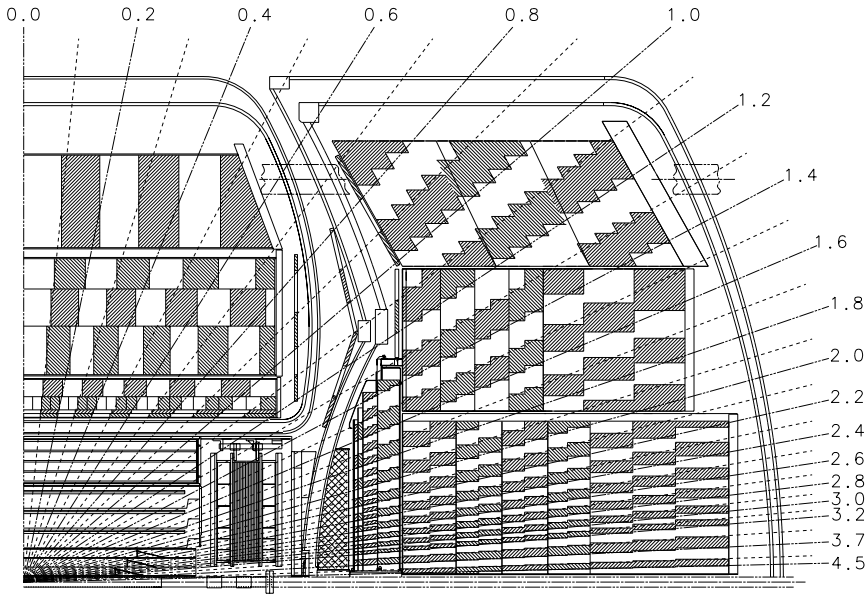

Figure 3.8: Side view of one quarter of the calorimeter, showing segmentation and tower definitions. The lines extending from the center of the detector are labelled by their pseudorapidities.

2, and 4, and is 0.05×0.05 in the third layer to allow more accurate measurement of the shower at its maximum, which is important for electron identification.

Signals from the various layers are grouped into *pseudo-projective* towers, meaning that the centers of the cells in each layer line up with the nominal interaction point, while their edges are perpendicular to the absorber plane, as shown in Fig. 3.8.

Hadronic Modules

The principles of hadronic calorimetry are similar to those for electromagnetic calorimetry. The incident particle collides inelastically with a nucleus in the absorber medium, producing a number of secondary hadrons which may also collide inelasti-

cally, thus producing a shower. However, the possible types of nuclear interactions are far more numerous and complex than the bremsstrahlung and pair production processes that cause an electromagnetic shower, thus rendering an analytic description more difficult. Nonetheless, some general features of hadronic showers can be elucidated.

The appropriate scale for nuclear processes is the nuclear interaction length λ, defined as:

$$\lambda = \frac{A}{\sigma_i N_o \rho}$$

where σ_i is the inelastic nuclear cross-section, N_o is Avagadro's number, and ρ is the density of the absorber. The average shower maximum scales as the log of the incident particle energy, occurring at $\approx (0.2 \ln E(GeV) + 0.7)\,\lambda$ (about 1.6λ for a 100 GeV hadron), and 95% of the shower is contained in a depth a little more than 2.5λ beyond this. The transverse spread of the shower is also typically significantly larger than for the electromagnetic case, with a cylinder of radius about 1λ required to contain 95% of the energy.

The limit on resolution for hadronic calorimeters comes from fluctuations in the shower composition, particularly in the fraction of π^o's produced in the first interaction. Since π^o's decay quickly to two photons, they will produce an electromagnetic shower within the hadronic shower. Most of this energy can be measured. On the other hand, nuclear interactions may produce neutrinos and muons which escape the detector, and also cause nuclear excitations and breakup, the energy from which may not be detectable. Hence typically the response of a calorimeter to hadrons is less than that for electrons of the same energy. While this difference (known as the e/π *ratio*) can be corrected for on average, a non-unity e/π ratio translates into irreducible variations in response on a shower-by-shower basis, depending on the fraction of the shower which behaves electromagnetically.

This highlights one of the main advantages of using uranium as the absorber material. Since secondary neutrons can cause fission of uranium nuclei, some of the

energy from which will be detectable, the e/π ratio can be nearly unity. Hence the limit on hadronic energy resolution from fluctuations in shower content is roughly a factor of two smaller for a uranium calorimeter than for any other feasible material.

As this lower limit tends still to be larger than the contributions of sampling statistics and incomplete shower containment, the resolution obtained is close to this limit.

The layers of calorimeter outside of the EM layers form the DØ hadronic calorimeter. In the CC, modules are of two varieties: the fine hadronic (FH) modules, which lie immediately behind the EM layers and have absorbers consisting of 6 mm thick uranium-niobium alloy plates, and the coarse hadronic (CH) modules, the absorbers for which are 46.5 mm thick copper plates. The transverse segmentation in all hadronic modules is $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$. The FH modules are divided into 3 readout layers (1.3, 1.0, and 0.9λ deep) and provide detailed information of the shower shape, while the CH modules are treated as a single 3.2λ deep layer whose primary purpose is to complete the containment of the shower.

The geometry of the EC requires a greater variety of hadronic modules, but their function is similar to that of the CC modules, with the sections nearest the interaction point containing uranium absorber plates divided into four readout layers, while the modules behind them consist of 46.5 mm thick stainless steel plates and are read out as a single layer. The total amount of material in the calorimeter ranges from 7.2λ at $\eta = 0$ to 10.3λ at the edge of the EC nearest the beamline.

Intercryostat Detector and Massless Gaps

Any particle traversing the boundary between the CC and EC encounters a substantial amount of material from the cryostat walls. In order to allow some sampling of the shower in this region, *massless gaps* and the *intercryostat detector* (ICD) have been constructed. The massless gaps have the same structure as a typical readout gap, but are bordered by copper-clad G10 rather than uranium absorber plates (the

cryostat walls thus playing the role of the absorber). Massless gaps are installed in both the CC and EC, as shown in Fig. 3.8.

ICD's are mounted on the inner surface of the EC (see Fig. 3.8) and consist of 384 scintillator tiles, each segmented into $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$ and aligned with the towers defined by the other calorimeter layers. The ICDs are the only components of the DØ calorimeter that do not use liquid argon as the sampling medium.

Calorimeter Readout and Performance

There are a total of about 5000 towers in the calorimeter, and their subdivision into layers brings the number of readout channels to $\approx$ 47000. The signals are first processed by preamplifiers and then sent to base-line subtracter (BLS) circuits on the detector platform, which sample the integrated charge just before a beam crossing and again 2.2 μs later, and define the signal as the difference between the two. The signal from the BLS is then amplified by 1 or 8 (depending on its size), allowing 15-bit dynamic range using 12-bit ADC's in the movable counting house. In order to reduce the bandwidth demands of transmitting data from every readout cell, zero-suppression is applied to remove cells without significant energy. This is done by comparing the signal to the width of the distribution of noise hits recorded between accelerator stores. If the magnitude of the signal is less than twice this width (significant signals may be positive or negative), the cell is not read out and its energy is set to zero.

Several modules of the EM and hadronic calorimeters were evaluated in beam tests prior to their installation in the calorimeter [22]. The response to both electrons and pions as a function of energy was found to be linear to within 0.5% for incident particle energies in the range 10 - 150 GeV. The resolution found may be approximated by:

$$\begin{aligned} \frac{\sigma(E)}{E} &\approx \frac{16\%}{\sqrt{E}} \text{ for electrons} \\ &\approx \frac{41\%}{\sqrt{E}} \text{ for pions} \end{aligned}$$

These approximations show the expected $1/\sqrt{E}$ dependence expected from the statistics of shower development, but do not describe the smaller contributions due to noise and calibration uncertainties (which would appear as constant terms). A more precise parameterization for electron resolution is given by [35]:

$$\frac{\sigma(E)}{E} = \frac{0.15}{\sqrt{E}} \oplus 0.03 \tag{3.1}$$

It is also important to note that the resolution for single hadrons is much better than that which can be obtained for jets made up of several hadrons of various energies, as will be discussed in Chapter 6.

Note also that the relatively good hadronic resolutions reflect the compensating nature of the calorimeter; test-beam measurements show that the e/π ratio falls from 1.11 at 10 GeV to 1.04 at 150 GeV.

Finally, the position resolution of the EM calorimeter is important for matching calorimeter energy clusters to central detector tracks in identifying electrons. This resolution is found to be 0.8 - 1.2 mm for 100 GeV electrons (better if the electron hits near a tower edge, thus sharing energy more equally) and to scale as $E^{-1/2}$, again reflecting the statistical nature of the shower.

3.2.4 Muon System

In order for a particle to pass through the material in the calorimeter, it must 1) have lifetime sufficient to travel several meters before decaying, 2) not participate in the strong interaction (and thereby cause a hadronic shower), and 3) be unlikely to lose substantial energy due to bremsstrahlung (thereby initiating an electromagnetic shower). The only charged particle known to have these properties is the muon, and therefore detectors are constructed outside of the calorimeter expressly for muon detection.

Since muons deposit little of their energy in the calorimeter, a spectrometer must be used to measure their momenta. This is formed by layers of proportional drift tubes

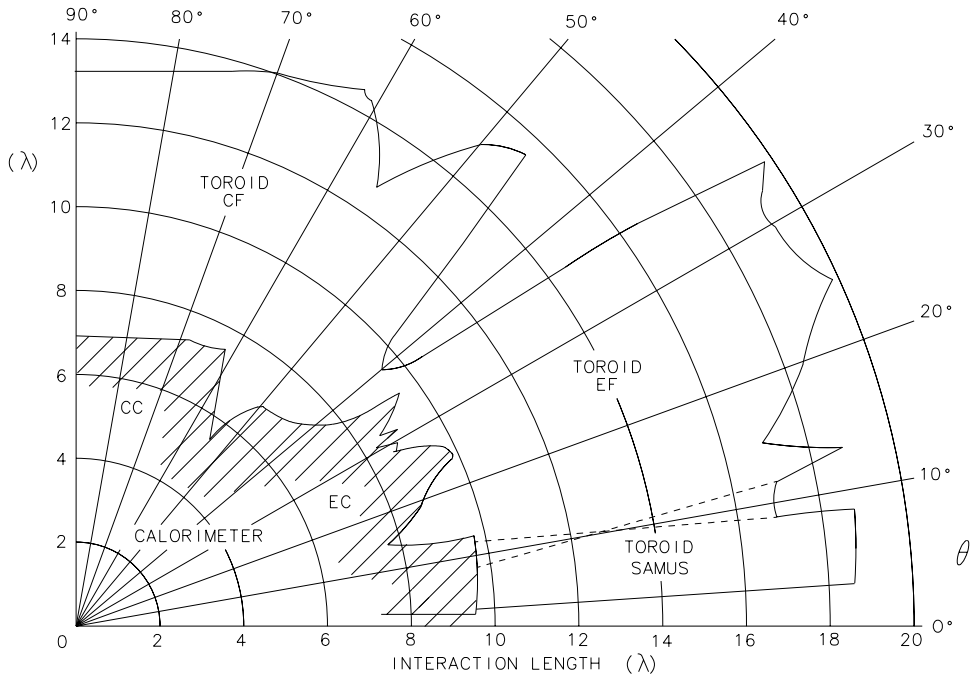

Figure 3.9: Total material in the calorimeter and muon toroid, as a function of polar angle. Except for the gap between the central and end toroids (at 40°) there are $\gtrsim 14\ \lambda$ in front of the outer drift tube layers.

(PDTs) surrounding a magnetized iron toroid. Measurement of the particle direction before and after traversing the toroid allows determination of its momentum, and the presence of the additional material outside the calorimeter makes it extremely unlikely that any particles other than muons will reach the outer layers of drift tubes (see Fig. 3.9).

The DØ Wide-Angle Muon Spectrometer (WAMUS) is formed from three planes of proportional drift tubes, the first (the A layer) mounted on the inner surface of the magnetized iron toroid, the second (B) layer on the outside surface and the final (C) layer generally 1.4 m beyond this. The A layer consists of four layers of PDTs, allowing measurement of the incident muon direction to 0.6 mrad and position to 100

μm. Additional information from the event vertex, central detector track, and muon trace in the calorimeter can improve the measurement of the initial direction. The B and C layers each have three layers of PDTs, which measure the outgoing position and direction to 0.17 mm and 0.2 mrad. The WAMUS PDTs cover the region $|\eta| < 1.7$.

The toroid itself is divided into three sections, a central piece shaped as a square annulus (the CF toroid) which covers the region $|\eta| < 1$ and two end toroids (EF) which extend to $|\eta| < 2.5$. A square hole in the center of the EF toroids accommodates the insertion of a separate toroid for use by the Small-Angle Muon Spectrometer (SAMUS) which uses another set of PDTs to extend muon coverage to $|\eta| < 3.6$. As only muons measured by the WAMUS are used in this analysis, no details of the SAMUS are included here.

The CF toroid is 1.1 m thick, with its inner surfaces 317.5 cm from the beamline. The EF toroids are slightly more than 1.5 m thick, with their inner surfaces at $|z| =$ 447 cm. Wire coils carrying a current of 2500 A induce magnetic field of about 2T in the CF and EF toroids, with the field lines running approximately in the azimuthal direction.

The PDT cells are formed from aluminum extrusions which are joined together as shown in Fig. 3.10. Each cell is 10.1 cm wide and 5.5 cm high. Cathode pads are inserted at the top and bottom of each cell, and a 50 μm diameter gold-plated tungsten anode wire is strung in the center. During operation, the aluminum walls are held at ground, with the cathodes at +2.3 kV and the anode wires at +4.56 kV, and the gas used is a mixture of Ar (90%), CF_4 (5%) and CO_2 (5%). The length of the cells varies to suit the geometry of the detector, with the longest wires being 6.1 m. All wires are aligned approximately parallel to the magnetic field direction.

As in the central tracking detectors, drift time information is used to measure the track position perpendicular to the wire direction. Both timing and cathode information are used to measure the coordinate along the wire direction. In order to simplify the placement of readout electronics, anode wires from two adjacent cells in the same PDT layer are connected at one end of the chamber. Hence a muon produces

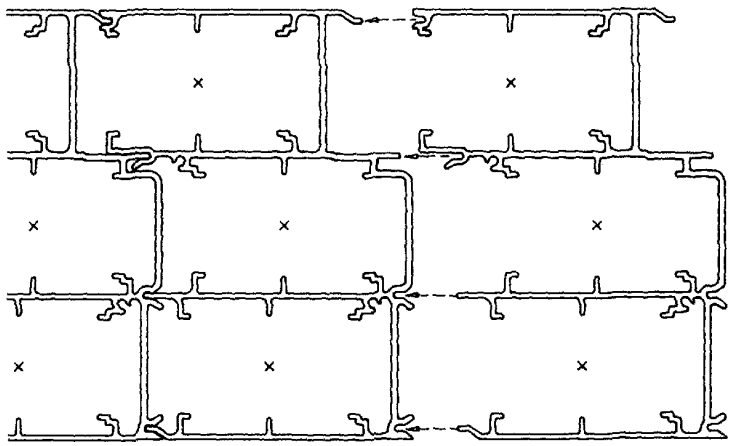

Figure 3.10: End view of proportional drift tubes used in the muon system.

a signal on two PDT wires, both the one in the cell is actually traversed and in the adjacent cell which is connected to it. Noting the difference in time between the two signals allows a rough measurement of the position of the hit along the wire direction, with a resolution of about 10-20 cm.

The cathode pads are designed to improve this resolution. The pads have two independent electrodes, which are arranged in a diamond pattern which repeats every 30 cm along the wire direction. Comparing the sizes of the signals induced on the two electrodes allows one to determine the point in the pattern at which the hit occurred, with an accuracy of about 3 mm. Since the timing information is sufficient to resolve the ambiguity which arises from the repeating nature of the pattern, the absolute position along the wire direction can be measured to 3 mm.

Much of the signal-processing electronics is mounted on the chambers themselves, with only the digitizing circuitry in the MCH. Signals from the cathodes are first sent to a pre-amplifier, and then passed to base-line subtraction circuitry similar to that used for the calorimeter. If the BLS output signal exceeds a threshold the channel is

latched for readout.

Signals from the anode are sent to both time-to-voltage circuits (for measuring the drift distance) and Δtime-to-voltage circuits (for measuring the coordinate along the wire direction). Test-beam studies have shown that the drift time is a slightly nonlinear function of distance. When this known nonlinearity is corrected for, the resolution perpendicular to the wire direction is about 0.3 mm.

The performance of the muon system in measuring a particle's momentum is determined by a number of factors. The most obvious of these is the geometrical acceptance for a muon to hit all three PDT layers. This acceptance is 60 - 70 % for most η's, but somewhat less in the CF-EF transition region (muons need to have momentum greater than ≈ 4 GeV/c to traverse the toroid and reach the outer PDT layers). For a particle which does hit all three layers, the momentum resolution is determined by the quality of information available for the original direction, the position resolution of the muon system, and multiple scattering in the toroids. The latter effect limits the resolution to 18% of the momentum. Overall, the resolution can be parameterized by [36]:

$$\sigma\left(\frac{1}{p}\right) = 0.18(p - 2\mathrm{GeV/c})/p^2 \oplus 0.003 \tag{3.2}$$

3.2.5 Triggering

As mentioned above, the proton and antiproton beams cross each other once every 3.5 μs, and at the operating luminosity of the TeVatron nearly all of these crossings produce at least one $\overline{p}p$ collision. As one cannot record data from all of these collisions, a real-time event processing system is needed to decide which events are sufficiently interesting to be preserved for offline analysis. This *trigger system* is divided into three levels.

Level Ø

The Level Ø trigger consists of two scintillator hodoscopes mounted on the front surface of each end calorimeter, covering the range $1.9 < |\eta| < 4.3$, and having greater than 99% efficiency for detecting inelastic collisions. In addition to the single bit of information on whether a collision occurred, the difference in signal time between the hodoscope signals is used as a measure of the z position of the collision. Both a fast and slow estimation of this position are made. The former is based on the analog sum of signals from a subset of the Level Ø counters, and is used to reject events with $|z| > 100$ cm, which are usually beam-gas events. This fast estimate of the z vertex is available to the Level 1 trigger for use in the calculation of transverse energies. A more accurate measurement of the z position takes into account the time and total charge from each counter, and applies known corrections and calibrations to further improve the accuracy. The result of this calculation is available to the Level 2 software filter. Of course, the z position can be determined unambiguously only for the subset of beam crossings in which only a single collision took place. By calculating the RMS deviation of the time signals from the individual counters, the Level Ø trigger can determine which beam crossings likely contained multiple collisions, and this information is also available to the higher-level triggers.

In addition to its role as primary trigger, the Level Ø system also serves as the luminosity monitor for the experiment [37, 38]. The instantaneous luminosity is given approximately by measuring the rate $R_{\mathrm{L\emptyset}}$ of Level Ø triggers:

$$\dot{\mathcal{L}}_{\mathrm{meas}} = \frac{R_{\mathrm{L\emptyset}}}{\sigma_{\mathrm{L\emptyset}}} \tag{3.3}$$

where $\sigma_{\mathrm{L\emptyset}}$ is the world average $\overline{p}p$ inelastic cross section, corrected for the LØ acceptances and efficiencies measured from Monte Carlo and data. Its value was 46.7 ± 2.5 mb in the 1992-1993 run [39], and 44.4 ± 2.3 mb in the 1994-1996 run [40]. The 5.2% uncertainty on this number is dominated by systematic differences between experimental measurements of the $\overline{p}p$ cross section.

Equation 3.3 is true if one assumes that every Level Ø firing resulted from a single $\overline{p}p$ interaction. As the instantaneous luminosity increases and multiple interactions become more common, this becomes a poor approximation and must be corrected using Poisson statistics:

$$\dot{\mathcal{L}} = \frac{-\ln\left(1 - \dot{\mathcal{L}}_{\mathrm{meas}}\tau\sigma_{\mathrm{L\emptyset}}\right)}{\tau\sigma_{\mathrm{L\emptyset}}} \tag{3.4}$$

where τ is the time between beam crossings.

The integrated luminosity is then given by numerical integration of the instantaneous luminosity measurements:

$$\mathcal{L} = \sum_{i=1}^{n} \dot{\mathcal{L}}_i f_{\mathrm{live}} \Delta t_i$$

where the live fraction f_{live} is measured using a trigger bit dedicated to this purpose.

Level 1

The next level of triggers is a hardware network which reduces the event rate to about 200 Hz. Most decisions are made within the time between beam crossing, but some events require additional confirmation from a somewhat slower network (known as the Level 1.5 trigger) which takes several beam crossings.

The Level 1 framework itself is an AND-OR network with 256 input bits provided by the calorimeter and muon system mapped into 32 output bits, or triggers. Communication with the framework is handled by the trigger control computer (TCC). Through the TCC, users can download the threshold for the firing of the AND-OR terms, and also define the pattern of AND-OR terms which causes the firing of each of the triggers. In addition, *prescale factors* can be defined for triggers whose firing rate would otherwise overwhelm the available bandwidth.

The calorimeter trigger takes its input from fast analog pickoffs from the BLS circuits. The analog sum of the energies in trigger towers of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ is computed separately for the electromagnetic and fine hadronic sections of the calorimeter. The analog input signals are digitized by an 8-bit FADC, and weighted

by the sine of the trigger tower polar angle, thus giving an approximate transverse energy (exact if the event vertex is at $z = 0$). These 8 bits, plus three bits from the Level Ø system providing z information, are used as the address for a lookup memory which returns the E_T.

Once the trigger tower E_Ts are known, the AND-OR terms are defined by comparing such quantities as the total E_T in the event, the transverse energy imbalance, and the electromagnetic and hadronic E_Ts in each trigger tower to thresholds downloaded through the TCC.

For most calorimetric information, the above processing is sufficient. However, the fact that some electromagnetic showers share their energy between two trigger towers implies that electron and photon triggers can benefit from a crude clustering algorithm, which is applied at Level 1.5 [41]. The clustering sums the electromagnetic energies from two adjacent towers, and also calculates the total energy in the 3×3 grid of towers centered on the electromagnetic trigger tower in order to allow calculation of the isolation of the electromagnetic object.

The muon trigger takes its input from the latch bits from each of the WAMUS and SAMUS cells. This information gives the bend direction coordinate with a granularity of 10 cm. By combining information from multiple layers, a centroid is defined as the center of the half-cell which was most probably hit. The OR of three chambers adjacent in the bend direction being hit is sent to a coarse centroid trigger (CCT) card, which ORs the information by another factor of 4 to create a 60 cm-wide trigger road. If the hit pattern in the A, B, and C layers is consistent with the passage of a muon, a Level 1 bit is asserted and the Level 1.5 trigger is invoked.

The Level 1.5 trigger passes information on all centroids to octant trigger cards (OTC). The OTC then compares all possible combinations of the hit centroids in the three layers to that expected from tracks above programmable p_T thresholds. This Level 1.5 processing reduces the muon trigger rate by a factor of 10-20, at the cost of about 1% deadtime. Overall, the muon and calorimeter Level 1.5 triggers reduce the event rate from about 800 Hz passing Level 1 to about 200 Hz for input into Level 2.

Level 2

Level 2 is a software filter which uses the digitized information from the event to perform a fast reconstruction, allowing the application of more sophisticated criteria to the event decision. The Level 2 system is a farm of 32 VAX Model 4000/60 and 16 VAX Model 4000/90 processors running in parallel, which reduces the event rate to about 4 Hz, which is written to tape. (Since each event is roughly 0.5 Mbyte, the detector typically writes about 2 Mbyte/s.)

During collider operations, all 48 processors are usually running the same code[3], and it is the job of a separate processor, the Supervisor, to direct each incoming event to an idle Level 2 node. The event filtering software is built around a collection of "tools", each of which has a specific task related to particle identification, or global event characteristic (such as total E_T). Which tools are invoked, and their order, are controlled by scripts, one of which is associated with each Level 1 bit. It is possible for a single Level 1 bit to cause the firing of multiple Level 2 filters; there are a total of 128 filter bits.

3.2.6 Main Ring Effects

The trigger system is also responsible for flagging events that are contaminated by Main Ring activity[42]. As described above, the Main Ring passes through the coarse hadronic calorimeter and is usually active as part of the antiproton generation system while data is being taken. Beam loss from the Main Ring can cause spurious signals in the coarse hadronic calorimeter and muon system.

Protons are injected into the Main Ring every 2.4 seconds. At the injection energy the magnetic field is of poor quality, and beam losses are large. Another large loss occurs 0.3 seconds later as the beam passes through transition. Hence a gate known as MRBS_LOSS is raised at the time of proton injections, and remains for 0.4 seconds,

[3]the exception occurs when a new version of the Level 2 code is under development. In this case, one node may be running the new code on a test basis.

until the beam has passed through transition and the muon high voltage system recovers.

During the remainder of the Main Ring acceleration cycle, losses are significant only when the passage of the proton beam through the accelerator coincides with a $\overline{p}p$ crossing the in TeVatron. Therefore, a second bit (known as MICRO_BLANK) is raised if a Main Ring beam transit occurs within ± 800 nanoseconds of a $\overline{p}p$ crossing.

Level Ø counters measure the fraction of crossings which occur during the MRBS-_LOSS or MICRO_BLANK windows, allowing analyses which veto these conditions to calculate their correct luminosity. For typical operation, this fraction was roughly 25%.

3.2.7 Data Acquisition

The data acquisition system is intertwined with the triggers. Once a Level 1 trigger is passed (and its Level 1.5 confirmation if necessary), the Supervisor is notified, and it in turn notifies another processor, the Sequencer, to begin digitizing the event. The Sequencer signals the front-end crates to begin digitization, which takes about 1 msec to complete. Data is then read out on eight unidirectional cables (corresponding to different detector systems).

The data cables, each of which can transfer 40 Mbyte/s, are connected to multi-port memory (MPM) boards on each of the Level 2 node, and the node selected by the Supervisor to process the event receives the data. The MPM are VAX memory boards with additional ports for the input cables. This direct interface between the cables and the memory enables the transmission of data to proceed at the necessary speed.

If the event passes any Level 2 filter, it is transferred to the host computer, which writes the event to a buffer disk. Once roughly 500 events have been written to a file, the file is closed and the data copied to 8 mm tape. A sketch of the data acquisition system is given in 3.11.

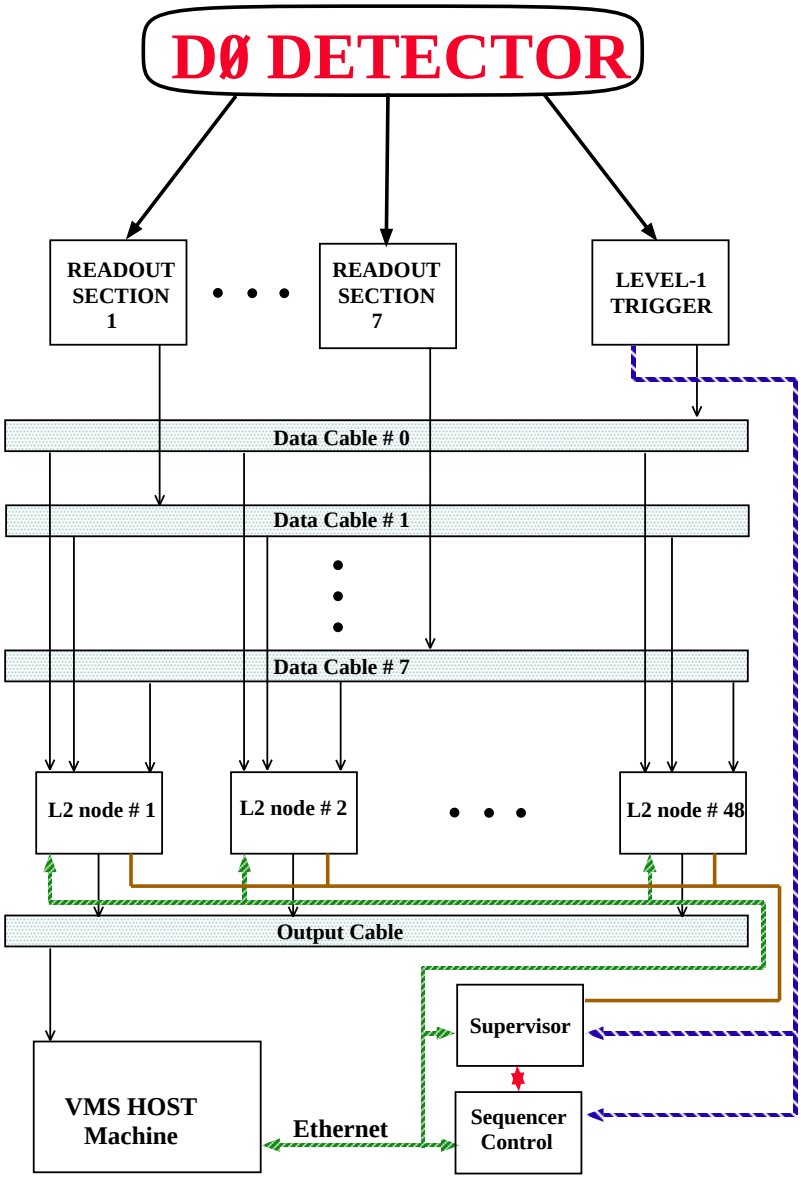

Figure 3.11: Sketch of the data acquisition system.

3.3 Detector Operation

As mentioned above, proton and antiproton beams are typically kept circulating the TeVatron for about 20 hours (a store), during which the detector is active and recording data. As the beams circulate, they gradually dissipate, resulting in lower luminosity at the collision point. This change in running conditions means that a set of prescale factors which is optimized for the beginning of a store will be unable to fill the available bandwidth near the end of the store. In order to maintain optimal throughput, data-taking is periodically paused to allow the downloading of a set of prescale factors optimized for the current luminosity.

The time in which a given prescale set is in place and the detector is running continuously is referred to as a *run*. Each run lasts roughly four hours, and events which pass Level 2 are numbered sequentially within each run, meaning that an event is labelled uniquely by its run and event number.

Chapter 4

Event Reconstruction

When a physics event is recorded by the DØ detector, the information on tape consists of ADC counts from all of the hit channels in the calorimeter and central tracking chambers, as well as analog and digital signals from the muon system. It is the task of the reconstruction program, DØRECO, to process this information into a form suitable for physics analysis, by converting the raw signals into information about the various final-state objects in the event. In this chapter, the DØRECO program is described, along with the techniques used for identifying electrons, muons, jets and missing E_T.

4.1 The DØRECO Program

DØRECO performs three major tasks. The first of these is *hit finding*, in which the signals from each sense wire of the tracking chambers are converted into the spatial location of hits, and the signals from each cell in the calorimeter are converted into energy deposits. Secondly, the tracking chamber hits are joined to form tracks while the calorimeter cells are grouped into clusters of energy. Finally, the tracking and calorimetric information is combined to reconstruct jets, and to identify electron and muon candidates (the criteria applied by DØRECO in selecting these candidates are

quite loose, and substantial rejection of spurious electrons and muons is gained by further offline processing, as described below).

4.1.1 Central Tracking Chamber Reconstruction

The first step in reconstructing the information from the central tracking chambers is identifying sense wire and delay line hits. The algorithm used depends on the difference in signal size between adjacent time bins from the FADCs. The leading edge of a pulse is found when three consecutive bins have differences above a threshold, or when two bins have differences above the threshold, the sum of which exceeds another threshold. Similarly, the trailing edge is identified by three consecutive bins whose differences fall below a threshold. The use of differences rather than absolute magnitude of the signals in defining a pulse reduces the sensitivity to fluctuations in the pedestals, and the fact that several bins are used also imposes constraints on the shape of the pulse, eliminating single-bin spikes which may arise due to bad FADC bits, and slowly rising signals due to discharge.

Once the leading and trailing edges of a pulse are found, the pulse size is calculated by summing the signals in the intermediate bins, and the width is taken as the difference between the leading and trailing edges. Time-dependent variations in the electronics gains and pedestals must be corrected for in this measurement. To do so, the response of the drift chamber channels to a calibration pulse is measured in the time between accelerator stores. The gains and pedestals are written to a database, which DØRECO accesses in order to determine the values appropriate for a given set of data.

The time of the pulse is given by

$$T = \frac{\sum_{i=1}^{N} w^{i-1} D(i) i}{\sum_{i=1}^{N} w^{i-1} D(i)}$$

where the sum is over all bins after the leading edge up to the trailing edge, w is a weight (0.5 for the VTX, 1.2 for the CDC, 0.6 or 1.0 for FDC signal lines (depending

on the pulse height) and 1.3 for FDC delay lines) and $D(i)$ is the difference between the ith bin and its predecessor. This gives the time in bins, which is easily converted into seconds since the speed of the digitizer is well-known. The weights are used to increase sensitivity to the signal from electrons arriving in the peak of the pulse. Note that this timing measurement can only give the distance of the hit from the wire, not which side the hit came from. Therefore two copies of the hit are stored, one at its actual position and the mirror image on the opposite side of the wire.

This algorithm is used for finding both sense wire and delay line hits. For the delay line, there is the additional requirement that the sum of the times for the signals read out on each end of the line be equal to the delay along the entire line, within a tolerance that allows for the resolution of the time measurements.

Once the individual hits are found, *segments* are defined which connect all of the hits within a given layer of the chamber. All of the hits are sorted in ϕ, and each possible pair of innermost and outermost wire hit combinations are considered. For each combination, a road is defined connecting the two hits, and the set of hits in the intermediate wires within this road which gives the best fit to a straight line is added to the segment. At this point the left-right ambiguity is broken by the staggering of the sense wires, which means that the true hits should have a better fit than their mirror-image hits. Once assigned to a segment, the hits are removed from the list to avoid having the same hit assigned to two segments. Up to two sense wires are allowed to be missing hits when a segment is defined.

The last step is to connect the segments in each layer to form tracks. This process begins in the outer layer of the chamber. Each segment in the outer layer is compared to the segments in the third layer which lie within a given ϕ distance, and the third-layer segment which matches most closely is added to the track. This process continues until the track extends through all four layers (one layer is allowed to have a missing segment). After this $r\phi$ fitting is done, the delay line information is added to fit the z coordinates of the track. Typical resolutions for the track direction are 2.5 mrad in ϕ and 28 mrad in θ [14]. Track-finding efficiencies are measured using

$Z \to ee$ events, and found to be $79.4 \pm 0.8\%$ in the CDC and $73.4 \pm 1.1\%$ in the FDC [43]. For more details on central tracking reconstruction, see [33, 44].

4.1.2 Event Vertex Determination

As mentioned above, the z position of a collision varies widely on an event-by-event basis, with a roughly Gaussian distribution of width 30 cm. Since it is essential to accurately measure the θ direction of final state objects (in order to assign energy vector components based on the total calorimeter energy), the z vertex of each event must be reconstructed with the highest possible accuracy.

To do this, one considers the set of CDC tracks which have an impact parameter with respect to the beam in the xy plane of less than 2.5 cm. All such tracks are projected to $x = y = 0$, and their z positions at this point are histogrammed. The peak of this histogram is used to determined the event vertex, with a resolution of about 1.2 cm. In addition, additional vertices (from multiple $\overline{p}p$ collisions) can be identified from secondary peaks in the histogram if they lie more than ≈ 7 cm away from the primary vertex [45, 44].

The xy position of the interaction is tightly constrained by the small transverse size of the colliding beams (about 50 μm). For any given store, the xy position at which the beams cross is also quite stable, so the measurement of the xy vertex is done on a store-by-store, rather than event-by-event, basis. For this measurement, a collection of about 500 events taken from the first run of a store is processed. The CDC tracks from these events are matched to VTX tracks, which improves the accuracy of the xy track position. All matched tracks are then extrapolated to either $x = 0$ or $y = 0$, depending on the azimuthal angle of the track, and the position of the orthogonal coordinate is histogrammed. The peak of each histogram gives the mean x and y interaction point for the store, which is recorded on a database for use by DØRECO [45].

4.1.3 Calorimeter Reconstruction

The reconstruction of calorimeter data begins with the conversion of the ADC counts recorded in each calorimeter cell into a value for the energy deposited there. To a first approximation, the conversion constant between counts and GeV can be taken from test beam runs performed before the calorimeter was installed for data taking. In these tests, portions of the calorimeter were exposed to electron and pion beams of known energies.

However, the test beam setup was not a perfect reproduction of the conditions that existed at the time of data-taking (some differences were in the amount of upstream material and the length of cables), and thus it is necessary to perform an *in situ* calibration, which is described later in this chapter.

As in central tracking hit reconstruction, the calorimeter signals must be corrected for time-dependent changes in the gains and pedestals of the readout channels. Calibration runs taken between stores were stored in the same database as the central tracking gains and pedestals, and accessed by DØRECO.

Once the energy deposited in each cell is determined, signals from all of the cells with the same η and ϕ indices are summed into towers. In taking this sum, it is assumed that the each cell represents a massless particle. As energy and momentum are equivalent under this assumption, each cell is assigned an energy four-vector $(E, E\sin\theta\sin\phi, E\sin\theta\cos\phi, E\cos\theta)$, where E is the signal in the cell and the θ and ϕ directions are defined by the cell centroid and the primary reconstructed z vertex. The tower energy four-vector is then given by the vector sum of each cell's four-vector.

Once this four-vector is assigned, the direction variables of the tower are calculated from it:

$$\phi = \tan^{-1}\frac{E_x}{E_y}$$

$$\theta = \tan^{-1}\frac{\sqrt{E_x^2 + E_y^2}}{E_z}$$

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

These towers are used as the starting point for jet reconstruction, while similar towers which include only the energies in the electromagnetic layers and innermost hadronic layer are used for electron and photon reconstruction.

4.2 Particle Identification

4.2.1 Electrons

The first step in the reconstruction of electrons and photons is to group electromagnetic towers into clusters of energy. Beginning with the highest-E_T tower, all neighboring towers with E_T above 50 MeV are added to the cluster, and the process repeats until no towers neighboring the cluster satisfy the energy requirement. A new cluster is then begun from the highest-E_T tower not previously assigned to a cluster.

Any cluster in the calorimeter with more than 90% of its energy in the electromagnetic layers of the calorimeter (and more than 40% in a single tower) is identified by DØRECO as an electron or photon candidate. As the typical hadronic jet is broad and deposits only about 10% of its energy in these layers, this cut alone removes most hadronic clusters while still retaining more than 99% of true electrons and photons. Electron candidates are distinguished from photon candidates by the presence in the CDC or FDC of a track within a road of size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pointing from the primary vertex to the cluster [44].

There are two primary background processes that can mimic an electron: one is π^0 decay to two photons, producing an electromagnetic cluster, with a track provided by the random overlap of a low-energy charged hadron. The other is photon conversion to e^+e^- pairs early in the tracking system. With no magnetic field in the tracking region, the electron and positron continue on nearly the same trajectory and may be identified as a single track. In order to suppress these backgrounds while retaining

high efficiency for identifying true electrons, information from the calorimeter and the tracking system is combined [14].

The first quantity considered is the *isolation* of the electromagnetic cluster. This is defined by comparing the electromagnetic energy within a cone of radius $\sqrt{\delta\phi^2 + \delta\eta^2} = 0.2$ centered on the cluster ($EM(0.2)$) to the total energy contained within a concentric cone of radius 0.4 ($E(0.4)$). The isolation fraction is defined as:

$$f_{iso} = \frac{E(0.4) - EM(0.2)}{EM(0.2)}$$

Any cluster with f_{iso} greater than 0.1 is rejected. This cut retains 98% of electrons while significantly reducing the backgrounds from random track overlaps (and also from the semileptonic decay of b or c quarks).

In addition to the overall electromagnetic fraction and isolation of the cluster, a detailed comparison is made between the shape of the cluster and that expected from an electromagnetic shower. The expected shape is characterized by a covariance matrix derived from a sample of Monte Carlo electrons:

$$M_{ij} = \frac{1}{N} \sum_{n=1}^{N} \left(x_i^n - \bar{x}_i\right) \cdot \left(x_j^n - \bar{x}_j\right)$$

where N is the total number of electrons in the sample, and the x_i are the variables which define the shape. A total of 41 variables are used:

- The fraction of the total energy contained in the first, second, and fourth layers of the electromagnetic calorimeter
- The fraction of the total energy contained in each cell of a 6×6 array around the shower center in the third layer
- The logarithm of the total energy
- The z position of the primary vertex

The matrix M is calculated individually for towers at different η, and symmetry in ϕ is assumed. Further, reflection symmetry is assumed for the positive and negative η regions of the detector, so there are 37 distinct matrices.

Once M has been calculated, the degree of agreement between an individual shower and that expected from an electron is defined by:

$$\chi^2 = \sum_{i,j=1}^{41} (x_i - \bar{x}_i)\, H_{ij}\, (x_j - \bar{x}_j) \tag{4.1}$$

where H is the inverse of M. Despite the notation, the fact that the variables are non-Gaussian means that this variable is not distributed as a χ^2 for electrons.

In order to reduce sensitivity to possible differences between data and Monte Carlo electrons, the H matrix is diagonalized and an upper limit is placed on the elements of the diagonalized matrix.

The other variables used for electron identification are provided by the tracking system. To further reject random track overlaps, the consistency between the direction of the central track and the centroid of the shower (called *track-match significance*) is calculated. The shower centroid is defined as:

$$\vec{x}_{\text{cog}} = \frac{\sum_i w_i \vec{x}_i}{\sum_i w_i}$$

where the sums are over all cells in the shower, $\vec{x}_i$ is a vector from the vertex to the cell centroid, and

$$w_i = \text{Max}\,(0, w_o + \ln\,(E_i/E))$$

The logarithmic weighting reflects the logarithmic development of a shower, and the w_o are chosen empirically to optimize the position resolution. The azimuthal resolution of the center of gravity is measured to be about 2.5 mm [14].

The track match significance for clusters in the CC (EC) is given by:

$$\sigma_{\text{TRK}}(CC(EC)) = \sqrt{\left(\frac{\Delta\phi}{\delta_{\Delta\phi}}\right)^2 + \left(\frac{\Delta z(r)}{\delta_{\Delta z(r)}}\right)^2}$$

where Δx is the mismatch in variable x between the shower centroid and the track direction, and $\delta_{\Delta x}$ is the resolution of the measurement of this mismatch.

To discriminate between prompt electrons and photon conversions, the track ionization (dE/dx) is measured. The distribution of energies deposited by a single ionizing particle has a long "Landau" tail on the high end, which reflects the energy transferred to scattered electrons (or *delta rays*). To reduce the sensitivity to these fluctuations, and thereby improve the overall resolution, the third of the CDC wires with the largest signals are removed from the measurement. The mean of the remaining signals is taken as the measurement of dE/dx.

The final variable which contributes to electron identification comes from the TRD. The TRD response is characterized by the variable ϵ:

$$\epsilon\left(E\right) = \frac{\int_E^\infty \frac{\partial N}{\partial E'}\left(E'\right) dE'}{\int_0^\infty \frac{\partial N}{\partial E'}\left(E'\right) dE'}$$

where E is the total energy recorded in the TRD minus that recorded in the layer with the largest signal (again, this is done to reduce sensitivity to delta rays) and $\frac{\partial N}{\partial E'}$ is the energy spectrum from a sample of $W \rightarrow e\nu$ events [46]. Since ϵ decreases as E increases, hadrons will tend to have values near unity while the distribution from electrons is roughly uniform over the allowed range of zero to one.

In order to extract the maximum possible background rejection and efficiency from the H-matrix χ^2, track-match significance, dE/dx, TRD ϵ and electromagnetic fraction of the cluster, these variables are combined in a vector $\mathbf{x}$ to calculate an overall consistency of the cluster with an electron.

In so doing, it is assumed that the variables are uncorrelated so that the probability of the cluster arising from hypothesis H is:

$$p\left(\mathbf{x}|H\right) = p\left(\epsilon|H\right) p\left(dE/dx|H\right) p\left(\chi^2|H\right) p\left(\sigma_{\mathrm{TRK}}|H\right) p\left(f_{\mathrm{EM}}|H\right)$$

where the possible hypotheses are electron ($H = e$), hadron overlap ($H = h$), or photon conversion ($H = ee$). Next the variable $\mathcal{R}$ is defined by:

$$\mathcal{R}(f_h) = \frac{p(\mathbf{x}|b)}{p(\mathbf{x}|e)} = \frac{f_h p(\mathbf{x}|h) + (1 - f_h) p(\mathbf{x}|ee)}{p(\mathbf{x}|e)}$$

	Lepton plus jets	Dilepton
CC	0.25	0.50
EC	0.30	0.50

Table 4.1: Cuts on $\mathcal{R}$ used in the lepton plus jets and dilepton analyses. The single-lepton analysis does not use TRD information in defining $\mathcal{R}$.

where f_h is the fraction of hadronic overlaps in the background. A cut is then placed on $\mathcal{R}(f_h)$, and clusters with values below the cut are taken to be electrons. The value of the cut differs for the single-lepton and dilepton final states, due to different optimizations between efficiency and background rejection in each analysis. Table 4.1 lists the cuts used for each analysis.

The probability densities, and f_h, are determined from the data as described in [46, 47]. f_h is found to be 0.53 in the CC and 0.62 in the EC.

4.2.2 Electromagnetic Energy Calibration

The calibration of the electromagnetic calorimeter begins by correcting for the known differences between the test beam and collider data runs. Such differences exist in the readout electronics, liquid argon purity, and voltage applied across the gap, and account for about a 5% offset between the test-beam and collider energy scales [48]. In addition, module-to-module variations are measured using large data samples with electromagnetic clusters. A minimum E_T is imposed, and the number of surviving clusters is plotted for each module. As the underlying physics is ϕ-symmetric, any nonuniformity must be due to differences in response which artificially move clusters across the threshold, and these variations are corrected. The RMS deviation is 1.3%, with the maximum excursion between any two ϕ modules being 5% [48].

Once this is done it remains to fix the overall calibration. The high degree of linearity observed in the test-beam studies allows one to infer a linear relationship

between measured and true electromagnetic energies:

$$E^{\mathrm{True}} = \alpha E^{\mathrm{Meas}} + \delta \tag{4.2}$$

Reconstructing the invariant mass spectrum of any particle which decays to electrons (or photons) is sufficient to constrain the calibration. In practice, three such calibration points are used: the Z boson, J/ψ, and π^0, and the allowed ranges for α and δ are shown in Fig. 4.1 [49]. Numerically,

$$\alpha = 0.9537 \pm 0.00086$$

$$\delta = -0.16^{+0.03}_{-0.21}\mathrm{GeV}$$

This precise calibration of the electromagnetic calorimeter serves as the starting point for calibration of the hadronic energy scale.

4.2.3 Muons

In principle the reconstruction of muon tracks is similar to the reconstruction of tracks in the central detector. However, differences in the geometry and electronics of the muon system, as well as the need to measure the bend angle (and thus the momentum) of the tracks, require that somewhat different algorithms be used. The information directly recorded by the detector is a digital pad latch indicating the presence of a hit in a given drift cell, along with analog signals which record the drift time, the difference in time between the signals read out at each end of the sense wire, and the charges on the inner and outer segments of the cathode pads.

As in the CDC, the reconstruction process begins with identifying the spatial location of hits. All hits in pad-latched channels which are associated with a drift time less than the maximum possible for the cell are located (again, two space points are associated with each hit due to the left-right ambiguity).

The hits are then joined into two straight segments, one including the hits from the A-layer cells and the other from the hits in the B and C layers (hence this will be

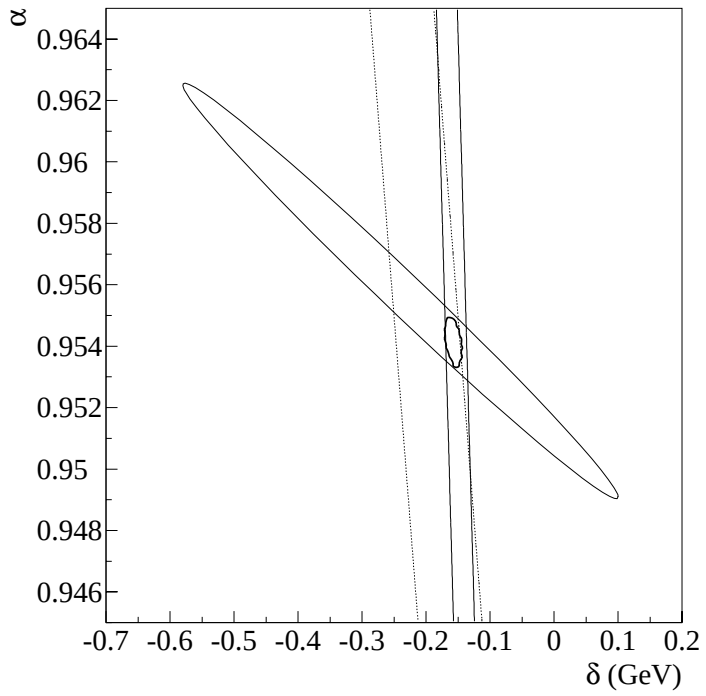

Figure 4.1: 68% confidence intervals for the electromagnetic response parameters α and δ. The wide vertical band is the constraint from J/ψ decays, the narrow band from π^0 decays, and the ellipse from Z boson decays. The small ellipse is the combined constraint [49].

denoted as the BC segment). These segments are reconstructed using a linear least-squares fit considering $r - z$ and $r - \phi$ information separately. Assuming that the set of true hits will fit more closely to a straight line than the set of mirror-image fits allows one to use the fit χ^2 to break the left-right ambiguity, and also any remaining ambiguity in the z position of the hits. BC segments are required to have four of a possible six hits, while A layer segments must have two of four, and all segments are required to point to within 5 m of the center of the detector (this cut removes randomly-oriented tracks from cosmic ray muons).

Once the segments are formed separately in the $r - z$ and $r - \phi$ views, only those segments are retained which consist of the same hits in each view. The next step is joining the segments into muon tracks. The process begins with the BC segments which are extended to the midplane of the toroid. The A-layer segment which points most closely to the point of intersection of the BC segment with the toroid midplane is then added to the track. If no A segment matches sufficiently well, the pre-toroid direction is defined as the line between the primary interaction vertex and the mid-toroid point. (Muon candidates are also formed for A segments with no matching BC segment, although since there is no momentum information available for such muons, they are discarded from this analysis).

To first order, the muon momentum is determined by the angle between the A and BC segments of the track, corrected for the energy lost by the muon in traversing the calorimeter. However, the measurement becomes more accurate if one performs a global fit making use of all possible information, including the presence of a track in the CDC. This fit makes use of a total of 16 input parameters:

- The x and y event vertex positions
- The slope and intercept of the CDC track in the $r - z$ and $r - \phi$ views
- Two angles representing the mismatch of the CDC track and calorimeter track directions

- The slope and intercept for the A and BC segments in the $r-z$ and $r-\phi$ views

The fit returns seven parameters: four for the CDC track, two representing the multiple scattering in the calorimeter, and the momentum of the track.

The primary backgrounds to muon candidates are from cosmic rays, and tracks formed from random noise hits in the muon chambers, with the contribution from hadrons punching through the calorimeter and toroid being negligible except in the gap region between the central and end toroids.

To reduce these backgrounds, several variables are used to identify good muons. These are:

- A word (IFW4) representing the quality of the track fit. Tracks with perfect fits have an IFW4 of 0, those with one failure have IFW4 of 1, and all others have IFW4 of 2.

- Muon track in the calorimeter. As the muon traverses the calorimeter, it deposits energy through ionization, and these energy traces are used in the track fit. The fraction of all possible hadronic calorimeter layers which had energy deposits large enough to be included in the fit is recorded (MTC_{fract}), along with the fraction of energy deposited in the outermost possible layer (f_{outer}) [50]. Both of these quantities are useful in rejecting muon tracks formed from random noise in the muon system.

- $\int \vec{B} \cdot d\vec{l}$. This quantity is used to reject tracks which pass through the inter-toroid crack. Not only is there a significant punchthrough background for such tracks, but their momenta are also poorly measured.

- Track impact parameter. Tracks which do not pass near the beam position are likely to be cosmic rays. The three-dimensional impact parameter is calculated, along with the impact parameters in the $r-z$ and $r-\phi$ views.

- Time offset (Δt_o): The time of the hits in the track is allowed to float in the fit, and the difference between the best-fit time and the beam-crossing time is

calculated. This helps to reject cosmic-ray muons, whose timing is independent of the beam activity.

The variables used, and the values of the cuts, vary not only for the different $t\bar{t}$ decay channels but also to reflect changes in the operating conditions of the muon chambers and in the reconstruction code. One can divide the data sample into three subsets according to the quality of muon information available:

- Era I (Runs 50000 - 65000): the reconstruction code did not perform muon tracking in the calorimeter. For these runs the MTC_{fract} and f_{outer} variables are not available, so a simpler quantity (the amount of energy in a calorimeter road around the muon track) is used in muon identification.

- Era II (Runs 65000 - 89000): some muon chambers, particularly those in the EF and Main Ring regions, were inefficient due to buildup of polymers on the anode wires.

- Era III (Runs 89000 - end): chambers cleaned to remove polymers, and efficiency restored.

Given this variety of conditions, the muon identification is necessarily complicated, as is summarized in Tables 4.2 and 4.3.

4.2.4 Missing E_T

As neutrinos do not interact in the detector, their presence is inferred from an imbalance in the transverse energy of an event as a whole, which is known as "missing E_T", and denoted by $\not\!\!E_T$. This quantity is determined by summing the transverse energy components of every calorimeter and ICD cell [14]:

$$\not\!\!E_x^{\text{cal}} = -\sum_{i=1}^{N_{\text{cells}}} E_{xi}$$

Variable	Cut	Channels
Fiducial region	CF toroid	$\ell+$ jets, $\ell\ell$
IFW4	≤ 1	$\ell+$ jets, $\ell\ell$
Cal MIP	Yes	$\ell+$ jets, $\ell\ell$
$\int \vec{B} \cdot d\vec{l}$	≥ 0.6 GeV	$\ell+$ jets
Impact Parameter	< 22 cm (3D)	$\ell+$ jets
	< 20 cm (rz)	
	< 20 cm $(r\phi)$	
Δt_o	< 100 ns	$\ell+$ jets

Table 4.2: Identification criteria for muons in Era I (the variables and Eras are defined in the text). As the dilepton channels have lower backgrounds and require greater efficiency, not all cuts are applied to these channels, as indicated in the third column.

Variable	Cut	Channels
Fiducial region	CF toroid (Era II)	$\ell+$ jets, $\ell\ell$
	CF and EF toroids (Era III)	
IFW4	≤ 1 (CF)	$\ell+$ jets, $\ell\ell$
	$= 0$ (EF)	
Cal muon track	$(MTC_{\text{fract}} = 1.0)$ OR	$\ell+$ jets, $\ell\ell$
	$(MTC_{\text{fract}} > 0.75$ AND $f_{\text{outer}} > 0.0$)	
$\int \vec{B} \cdot d\vec{l}$	≥ 0.6 GeV	$\ell+$ jets
Impact Parameter	< 20 cm (3D)	$\ell+$ jets

Table 4.3: Identification criteria for muons in Eras II and III.

$$\not{E}_y^{\text{cal}} = -\sum_{i=1}^{N_{\text{cells}}} E_{yi}$$

The magnitude of $\not{E}_T{}^{\text{cal}}$ is obtained by summing the x and y components in quadrature.

This represents the total transverse energy carried away by particles which do not interact in the calorimeter. The resolution of this measurement is greatly enhanced by near-hermeticity of the DØ detector. Based on the distribution of $\not{E}_T$ in a sample of events which were required only to pass the Level Ø trigger (this is referred to as a *minimum bias* sample), the resolution can be parameterized as [14]:

$$\sigma(\not{E}_T) = 1.08 \text{ GeV} + 0.019 \times \sum_{\text{Cells}} E_T \tag{4.3}$$

In order to isolate the portion of $\not{E}_T{}^{\text{cal}}$ which is the result of neutrinos, the contributions from any muons identified in the final state is subtracted from each component:

$$\not{E}_x = \not{E}_{Tx}{}^{\text{cal}} - \sum_{j=1}^{N_{\text{muons}}} p_x^{\mu_i}$$

$$\not{E}_y = \not{E}_{Ty}{}^{\text{cal}} - \sum_{j=1}^{N_{\text{muons}}} p_y^{\mu_i}$$

again with the magnitude of $\not{E}_T$ equal to their sum in quadrature.

4.2.5 Jets

Unlike electrons, muons, and $\not{E}_T$, there is no unambiguous way of defining a jet. In the simplest case, that of two partons produced in the $\overline{p}p$ collision, one would expect each parton to fragment and produce a number of hadrons, which travel in approximately the same direction as the original partons. Then one will observe in the calorimeter two distinct clusters of energy back-to-back in ϕ. Suppose now that one of the original partons radiates a gluon prior to the fragmentation process. This gluon will also fragment, producing another spray of hadrons. If the gluon was emitted at a small angle to the original parton direction, these particles will tend to fall in the

same calorimeter cells as the hadrons from the initial parton, and one still has two clusters in the event. On the other hand, if the gluon was emitted at a wide angle, it will produce its own distinct cluster of energy, and one would identify the event as having three jets. There is no clear line of demarcation between the two cases.

In order to proceed, one must formulate an algorithm which associates deposits in the calorimeter with jets. The jets used in this analysis are defined using a *cone algorithm*, which proceeds as follows [14, 51, 52]:

First, an E_T-ordered list of the calorimeter towers is made. For every tower with $E_T > 1$ GeV, a precluster is formed consisting of that tower and all of its neighbors which also have $E_T > 1$ GeV. The E_T-weighted η and ϕ of each precluster is stored as a starting point for jet finding.

The jet algorithm proceeds by looping over all preclusters, summing the energy from all towers within $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < R$ from the precluster center. For the jets in this analysis, R was chosen to be 0.5. In taking this sum, the energy vectors from each tower are added vectorially, and the $\eta - \phi$ centroid of the tower is calculated in the same manner as for a single tower. If the new η and ϕ position of the jet is different from that of the precluster, the cone summation is repeated with the new coordinates as the center of the cone, and the process is repeated until the jet directions are stable to within 0.001 in $\eta - \phi$ space.

The jet is then compared to the list of previously reconstructed jets to determine if its jet cone overlaps with any others. If so, the E_T's from all the shared towers are added, and compared to the E_T of the softer jet[1]. If the shared E_T is greater than half of the softer jet's E_T, the jets are merged into one object. Otherwise they are split into two jets, with each tower being assigned to the jet with the nearest center.

It is possible for two or more preclusters to give rise to identical jets after the cone algorithm. To suppress this, if two jets are within $\Delta R < 0.01$ of each other, and have $\Delta E_T < 10$ MeV, the second version of the jet is simply deleted rather than being split

[1]The definition of jet E_T used for reconstruction is different from that used in the analysis. The former uses the sum of the E_Ts of the towers comprising a jet, while the latter uses $\sqrt{E_x^2 + E_y^2}$.

or merged.

Finally, in order to suppress random noise fluctuations which can produce small energy clusters, jets are retained only if they have $E_T > 8$ GeV.

4.2.6 Jet Energy Calibration

The ambiguities inherent in the definition of jets necessarily complicate the task of their calibration. In order to perform a reconstruction of the top quark mass, one would like to identify the energy of a jet with the energy of the original parton which gave rise to the jet. In DØ the calibration to the parton level is carried out in two distinct steps, the first of which corrects for calorimeter effects so that the jet energy is on average that of the final-state particles contained within the jet cone. This procedure is applied by a post-DØRECO software package called CAFIX[53]. The average correction due to the fact that gluons can radiate from the original parton at large angles, causing some energy to fall outside the jet cone, is done after the application of CAFIX.

CAFIX

The CAFIX package corrects the jet energy reconstructed by DØRECO (hereafter referred to as $E_{\rm jet}^{\rm RECO}$) for the following effects:

- Overall hadronic response R_h. Among the factors which may cause R_h to differ from unity are the extended nature of hadronic showers, which causes some energy to be lost in intermodule cracks or other poorly instrumented regions, the nonlinearity in calorimeter response to sub-10 GeV particles which may be present in the jet, and any difference between the response measured in the test beam and that obtained during data-taking.

- Energy deposited in the jet cone by particles not produced in the hard scatter. Such particles arise for example from the fragmentation of the spectator quarks

in the collision, and their production is known as the *underlying event*.

- Noise from both electronics and the fission of uranium nuclei.
- The transverse spread of the hadronic shower, resulting in some of the energy from particles within the jet cone being deposited in calorimeter cells outside of the cone (and vice versa).

Once these correction factors have been determined, the corrected jet energy $E_{\rm jet}^{\rm corr}$ is given by:

$$E_{\rm jet}^{\rm corr} = \frac{E_{\rm jet}^{\rm RECO} - O}{(1-S)\,R_h} \tag{4.4}$$

where O is the energy-independent offset due to noise and underlying event and S is the correction factor for out-of-cone showering.

The magnitude of O is taken from a sample of minimum bias events. The energy per $\eta - \phi$ area for these events was plotted as a function of η to get the total offset term.

In order to separate the contribution from underlying event, which presumably depends on the number of collisions in a beam crossing, from that due to noise, the minimum bias sample was divided into two subsets. The first of these were events for which the Level Ø and tracking information combined to yield a high probability for the event to contain only one interaction, while the second had a high probability for multiple interactions. The relatively low instantaneous luminosity at which the sample was recorded ensures that the multiple interaction subsample consists predominantly of events with exactly two interactions.

The average difference in energy per $\eta - \phi$ area for the two subsamples was identified as the contribution to the underlying event from the second interaction in the multiple-interaction sample, and parameterized as:

$$U/\text{event} = (0.310 + 0.034|\eta_d|) \pm 0.2\text{GeV/rad}/\eta$$

where η_d is the η of the calorimeter tower containing the jet axis. In applying the correction, one needs to multiply by the expected number of interactions at the instantaneous luminosity for each event, given by

$$\langle N \rangle = 0.715 \times \dot{\mathcal{L}}(\mathrm{pb}^{-1}\mathrm{s}^{-1})$$

The remaining energy in the minimum bias sample after subtraction of the underlying event is identified as the noise contribution and parameterized as:

$$N = (0.196 + 1.44 \sin\theta_d)\mathrm{GeV/rad}/\eta$$

In contrast to the case for the electromagnetic energy scale, there is no convenient physical resonance one can use to calibrate the hadronic response. However, the hadronic scale can be related to the electromagnetic scale by considering events which contain a photon recoiling against one or more jets. Such a sample was defined by imposing the trigger requirement that the event contain a photon passing some energy threshold (ranging from 6 to 40 GeV). Additional cuts on the electromagnetic fraction and isolation of the cluster, and which veto the presence of a track (or significant CDC hits) in the road leading to the cluster improve the purity of this sample.

Events which contain real neutrinos are a negligible fraction of this sample. Therefore, we have:

$$R_\gamma \vec{E_T^\gamma} + R_h \vec{E_T^h} = -\vec{\not{E}}_T$$

where $\vec{E_T^h}$ is the net unclustered hadronic E_T. Using the fact that momentum balance implies $\vec{E_T^\gamma} = \vec{E_T^h}$, and the fact that the electromagnetic calorimeter has been calibrated so that $R_\gamma = 1$, this reduces to:

$$R_h = 1 + \frac{\vec{\not{E}}_T \cdot \widehat{n_T^\gamma}}{E_T^\gamma}$$

Since the right-hand side depends only on the well-measured photon variables (and $\not{E}_T$), this allows a direct measure of R_h.

R_h can be measured in principal as a function of any variable in the event. While it would be convenient to use the measured jet energy as the variable in which to parameterize R_h, the use of a variable with such poor resolution introduces the possibility of bias in the measurement. Hence the variable chosen is E', defined as:

$$E' \equiv E_T^{\gamma} \cosh\left(\eta_{\text{jet}}\right)$$

which depends only on well-measured quantities.

In addition, it is found that R_h varies strongly as a function of the width of the energy distribution within the jet, with narrow jets having a higher response. Hence the response correction is also determined as a function of the jet width. Adding this variable does not change the average energy scale but should improve the resolution on a jet-by-jet basis.

As most events do not contain a photon, E' is usually undefined. Therefore it is necessary to determine the response as a function of the measured jet energy. In order to do this without introducing bias, the measured jet energy as a function of E' is measured in the photon plus jet sample. Then, since R_h is known as a function of E', R_h can be found as a function of the measured jet energy.

After applying the corrections for the variation in the response as a function of jet width, the energy dependence is parameterized as:

$$R_h = a + b \ln\left(E_{\text{jet}}^{\text{MEAS}}\right)$$

where $E_{\text{jet}}^{\text{MEAS}} \equiv E_{\text{jet}}^{\text{RECO}} - O$, $a = 0.71(0.74)$ and $b = 0.025(0.031)$ for actual (Monte Carlo simulated) events.

The remaining correction factor to be determined, the out-of-cone showering S, depends explicitly on the jet definition used (cone jets with larger radius have less out-of-cone leakage). To measure the magnitude of this effect, a sample of simulated jet events is created. The locations at which the pions and photons in the jets strike the calorimeter are noted, and showers from test-beam pions and electrons of similar energies are placed at this location. A companion sample is created using the same

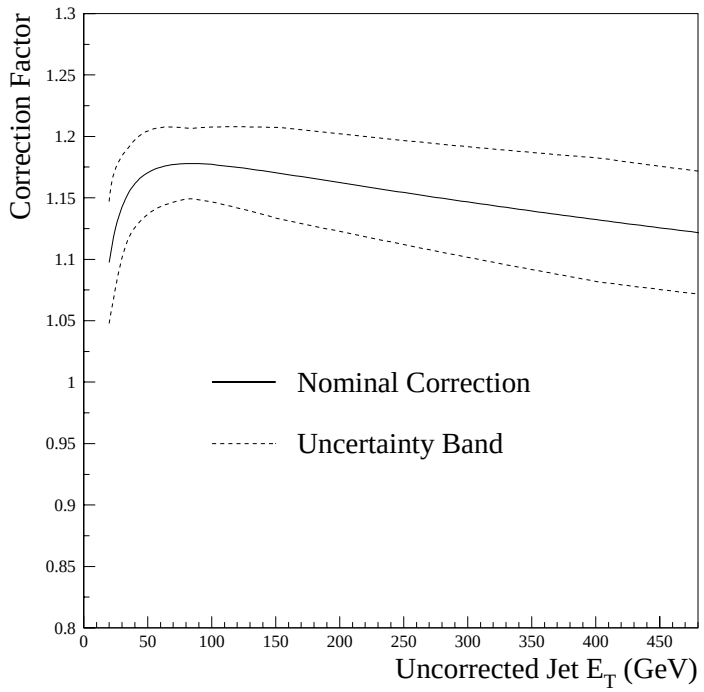

Figure 4.2: Jet correction factor as a function of measured jet transverse energy [53].

simulated jets but placing all of their energy in the first struck cell (thus eliminating the out-of-cone showering for this sample.) Comparison of the reconstructed jet energies in the two samples allows one to determine the out-of-cone correction. For 0.5 cone jets, there is found to be a 3% loss for low-energy (15 GeV) jets, which decreases to nearly 0% for 50 GeV jets.

The overall jet correction factor $E_{\rm jet}^{\rm corr}/E_{\rm jet}^{\rm RECO}$ varies as a function of jet energy and η. In the central region, it is 95% for low-energy jets, reaches a maximum of 1.18 for jets with E_T about 90 GeV, and falls to an asymptote of 1.13 for extremely high-energy jets, as shown in Fig. 4.2.

All of the corrections listed above, save for that due to out-of-cone showering, represent differences between the energies of the particles produced in a $\overline{p}p$ collision

and those energies recorded in the calorimeter. This energy must be recovered not only to obtain accurate jet energies, but also accurate $\not\!\!E_T$. In order to achieve this, the corrections applied to all 0.7 cone jets in the event are recorded, and the $\not\!\!E_T$ is changed to reflect the change in the jet energy.

Tests of the CAFIX Corrections

It is of crucial importance to the top quark mass measurement that the hadronic energy scale of the DØ detector be faithfully modeled by the Monte Carlo. In order to test that the CAFIX corrections meet this criterion, a sample of HERWIG direct photon events was generated. The subsample with the photon recoiling against a single reconstructed jet was selected, and the imbalance between jet and photon E_T along the photon direction recorded. While one should not expect the two objects to exactly balance (it is the entire hadronic recoil, not just that part reconstructed as a jet, which balances the photon), the degree of imbalance should be the same in the Monte Carlo sample and a data sample with identical selection criteria.

In fact, it is found that the degree of imbalance is not the same [54]. For jets in the central calorimeter, the discrepancy is small (about 3%) but rises to about 10% for jets in the intercryostat and forward regions. The imbalances as a function of jet detector η are shown in Fig. 4.3. An exact solution to these problems demands another iteration of the CAFIX corrections. However, due to the complexity of the corrections, many months would be required in order to derive, implement and test an improved version of CAFIX. In order to allow the timely completion of the top quark mass analysis, and also to account as much as possible for our best understanding of the jet energy scale, the decision was made to apply *ad hoc* post-CAFIX corrections to account for the discrepancy between data and Monte Carlo.

In the dilepton mass analysis, the corrections were applied to the data only, such that the post-correction data would have the same degree of imbalance as the Monte

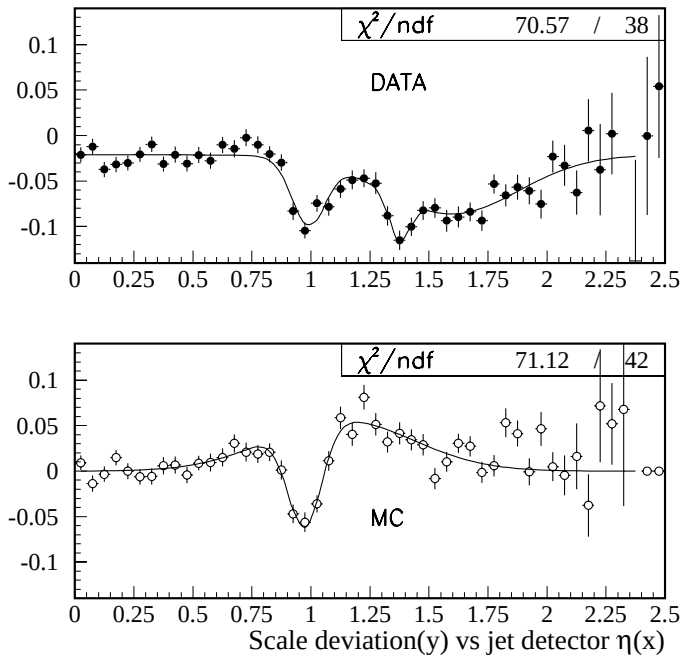

Figure 4.3: Percentage imbalance between photon and jet E_Ts for data and Monte Carlo samples [54]. The parameterizations shown are a triple-Gaussian for data and a double-Gaussian for Monte Carlo.

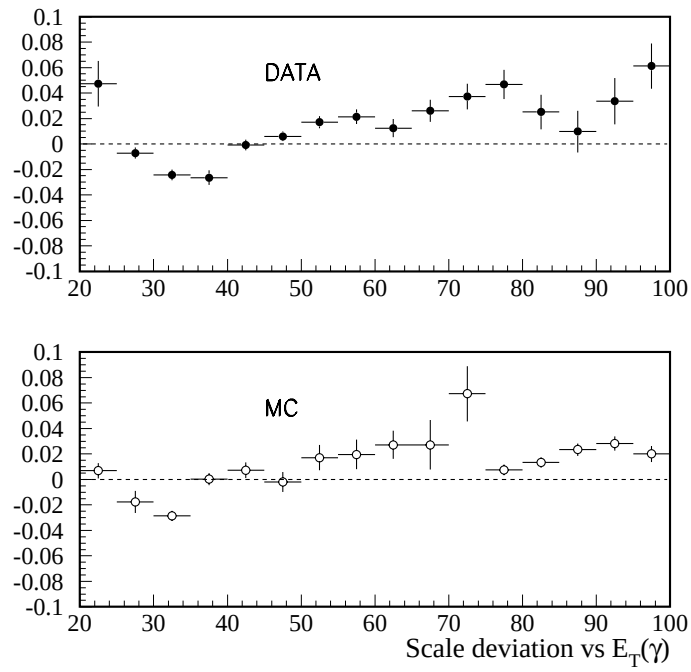

Figure 4.4: E_T balance in data and Monte Carlo direct photon events, after application of the post-CAFIX corrections.

Carlo. The data jets are corrected such that:

$$\frac{E'}{E} = \left(1 - 0.02 - 0.07 \exp\left[\frac{\eta_d - 0.99}{0.10}\right]^2 - 0.05 \exp\left[\frac{\eta_d - 1.38}{0.06}\right]^2 - 0.06 \exp\left[\frac{\eta_d - 1.59}{0.40}\right]^2\right)^{-1} \quad (4.5)$$

For the lepton plus jets analysis, both the data and Monte Carlo are corrected such that the imbalances are cancelled. This post-CAFIX correction is applied only for the mass analyses. The degree to which the Monte Carlo and data energies agree after this correction is shown in Fig. 4.4. Any residual imbalance contributes a systematic uncertainty in the top quark mass measurement, as discussed in Section 8.1.3.

Chapter 5

Top Quark Event Selection and Cross Section Measurement

Once the events have been recorded and reconstructed, the task of selecting those most consistent with $t\bar{t}$ production begins. As most $\bar{p}p$ events involve QCD interactions which do not produce leptons, those top quark decays which result in electrons and muons in the final state are easiest to distinguish. This analysis is restricted to such decays.

As discussed above, the leptonic decays of the $t\bar{t}$ events are divided into two broad categories: the lepton plus jets and dilepton channels. The former has the advantage of a large branching ratio, accounting for about 30% of all $t\bar{t}$ decays, with the disadvantage that electroweak processes or detector misidentification of final-state particles can mimic the $t\bar{t}$ signal relatively frequently. Conversely, the dilepton channels have lower backgrounds, but account for only 5% of all decays.

DØ measures the $t\bar{t}$ production cross section using events from eight channels, seven of which are also used for the top quark mass analysis. This chapter describes the selection process for events in these channels, and the manner in which the information is combined to measure the $t\bar{t}$ production cross section.

5.1 Monte Carlo Modeling of Signal and Background

In order to optimize a set of cuts to be efficient for selecting top events and for rejecting background, one needs a model for the final states expected from both signal and background events. For some sources of background, notably that arising from hadronic jets being misidentified as electrons, the model is provided by the collider data sample. On the other hand, most other sources of background and certainly the signal must be simulated using Monte Carlo programs which model both the physics of the event production and the detector response.

5.1.1 Model for $t\bar{t}$ Events

The primary model for $t\bar{t}$ production is provided by the HERWIG generator [55], version 5.7. In order to estimate the extent to which the results depend upon the uncertainty in this model, the ISAJET generator [56] is used as a cross-check.

In its modeling of QCD events, HERWIG relies upon factorization theorems which state that any process can be divided into the following independent steps:

- The elementary hard process. This is the interaction of initial-state partons giving rise to final-state partons, and is calculable in perturbative QCD
- Final state gluon emission
- Initial state gluon emission
- Formation of hadrons from final-state partons

The partons in the elementary hard process are assigned momenta based on the distribution expected from the matrix element calculation. In addition, the polarizations of the final state partons and the color connections between them are assigned, and this information is used in the further evolution of the event.

After the hard process is calculated, final-state gluon emission is added according to the rules of perturbative QCD. In particular, the energy fraction of the emitted gluon is distributed as expected from the Altarelli-Parisi splitting functions [57], and the phase space is restricted so that the angular separation between an emitted gluon and its parent is required to decrease at each successive branching. The latter effect models the interference among gluons. For heavy (b or t) quarks, the emission in an angle of size $\approx M_q/E$ is suppressed. In order to ensure that this process is restricted to the perturbative regime for which these rules are valid, a minimum threshold is placed on the transverse momentum of the emitted gluon with respect to the parent parton direction.

In addition, the initial-state partons which interact in the hard process may also emit gluons. The modeling of this emission includes all of the effects of final-state emission, plus additional corrections to account for singularities in the Altarelli-Parisi functions as the energy of the emitted gluon becomes small.

At this stage, all of the initial and final state partons which participate in the event have been modeled. However, an additional step is needed to account for the evolution of the partons into the jet of color-singlet hadrons that one observes in the detector. HERWIG models this effect by assuming that the hadronization process occurs between locally color-connected partons and is independent of the scale of the hard process. In particular, all of the final-state gluons are split into light-quark pairs, and the color lines are followed to form color-singlet clusters of partons. If a cluster is too light to decay into hadrons, it is assumed to represent the lightest hadron of its flavor. Otherwise, the cluster is fragmented into two or three hadrons selected at random from those compatible with the flavor of the cluster. Any unstable hadrons resulting from this process are assumed to decay to final state particles according to their measured branching ratios.

The remaining feature of $\overline{p}p$ collisions is the underlying event, defined as the interactions of the incident partons which do not participate in the hard scattering. The underlying event model is based on experimental observations of minimum-bias

events, which show that the multiplicity of charged clusters follows a negative binomial distribution. The hadronization of the clusters is modeled in the same manner as for final-state clusters.

5.1.2 Model for Background Events

Much of the background for $t\bar{t}$ events arises from the production of a W boson in association with multiple jets. The physics for this class of background is determined using VECBOS [58], which calculates the appropriate tree-level matrix elements exactly. The complexity of this calculation increases rapidly as a function of the parton multiplicity, requiring the calculation of only a single diagram when no final-state partons are included, but 1,286 diagrams to calculate a W + four jet event.

As complex as this calculation is, it represents only one component (the hard process) needed to fully simulate an event. A separate model must be used to add the underlying event, initial and final state emission, and to model the hadronization of the final-state partons. For most of the studies used in this analysis, the HERWIG program is used, with the partons supplied by VECBOS. As a cross-check, ISAJET was also used to model these processes.

Some additional background processes are modeled by PYTHIA [59]. A list of all Monte Carlo samples used in the analysis may be found in Appendix A.

5.1.3 Model for the Detector Response

Once the identity and momenta of all the final state particles expected from a given physics process have been simulated, it remains to simulate the detector response to these particles. This is done using the GEANT [60] program, which evolves the trajectories of the final state particles as they pass through the detector. Ionization in tracking chambers and the formation of both electromagnetic and hadronic showers are simulated in great detail, with each secondary particle also being stepped through the detector volume. Such detailed simulation is extremely CPU-intensive, and it

would not be possible to obtain the Monte Carlo statistics necessary for this analysis by using the full power of GEANT. Hence a compromise is reached by running the full GEANT simulation on a large sample of electrons, hadrons, and muons, and storing the resultant calorimeter showers in a library [61]. These showers are binned in five quantities representing the input particle:

- z vertex position (6 bins)
- η_d (37 bins, matching the calorimeter segmentation)
- Momentum (7 bins)
- ϕ region. The calorimeter is for the most part symmetric in ϕ, the exception being the cracks between modules in the central electromagnetic calorimeter, and the region through which the Main Ring passes in the hadronic calorimeter. Hence there are only two ϕ bins, representing the "good" and "bad" regions.
- Particle type. Calorimeter depositions for electrons, hadrons, and muons are stored separately.

A total of 1.2 million tracks is used to populate the library. When new Monte Carlo events are sent through the simulation, a shower from the library is selected to model the response to each particle. The total energy of the shower is scaled by the ratio of the energy of the particle to be simulated to that of the library particle which created the shower.

This still leaves some effects unaccounted for. First there is the efficiency and resolution of the muon system, both of which tend to be overestimated by GEANT. In order to account for this, a post-GEANT processing package called MUSMEAR is applied [62]. This routine smears the hit timing information simulated by GEANT so that the Monte Carlo hit position resolution matches that seen in data, and randomly discards hits to model the chamber inefficiency. In addition, the geometry file describing

the muon system is misaligned in order to reproduce the correct overall momentum resolution.

Next, one would like to model the response of the Level 1 trigger and Level 2 filter systems to the generated event. The utilities L1SIM and L2SIM provide this simulation given the input from GEANT, the former using a simulation of the trigger hardware and the latter the same code as is used in the online software filter. Unfortunately, these utilities were not implemented on the processor farm used for generating most of the samples used in this analysis. While this is not ideal, it does not significantly compromise the quality of information gained from the Monte Carlo, as the trigger efficiency for events having the $t\bar{t}$ candidate topology is quite high.

Finally, the Monte Carlo produces single events while in the actual data there is the possibility of multiple $\overline{p}p$ interactions during the same beam crossing. In order to account for this, the NOISY package allows one to overlay minimum bias Monte Carlo events on top of the Monte Carlo being processed. Since this procedure requires the events to pass through full GEANT, only a few samples are processed through NOISY. These are compared with the single-interaction Monte Carlo events in order to understand the nature and magnitude of multiple interaction effects.

Once the detector response has been simulated, the event is stored in a format identical to that for actual data, and it is processed by DØRECO to reconstruct electrons, muons, and jets. The only difference between the data and Monte Carlo events reconstructions is that determination of the z vertex is not done in the latter, since the showerlibrary procedure does not simulate particle tracks. The Monte Carlo samples are generated with a distribution of z vertices similar to that of the data sample, and the reconstruction makes use of the generated z vertex. With some exceptions which are explicitly noted, all of the Monte Carlo samples used in this analysis are passed through this procedure.

Channel	Luminosity (pb^{-1})
ee	125.3
$e\mu$	108.3
$\mu\mu$	104.5
$e\nu$	108.3
e + jets (topol.)	115.0
e + jets (μ-tag)	103.7
μ + jets (topol.)	108.3
μ + jets (μ-tag)	104.0

Table 5.1: Integrated luminosity for each of the $t\bar{t}$ decay channels. There is a 5.3% uncertainty in the luminosity measurement.

5.2 Data Sample

The data used in this analysis were collected between 1992 and 1996. The total luminosity is not identical for all top decay channels for three reasons: 1) The muon triggers were not fully efficient at the beginning of the run, leaving the channels which require only muons with slightly less luminosity, 2) the analyses of different decay channels place different cuts on the type of Main Ring activity allowed, and 3) the last period of running in early 1996 is not included by some analyses. Nonetheless, all channels have an integrated luminosity of > 100 pb^{-1}, as detailed in Table 5.1.

5.3 Event Cleanup

In order to ensure that the $t\bar{t}$ sample is not contaminated by events arising from detector pathologies, several steps are taken to remove suspicious events. A list of runs with known problems is kept, and no events from those runs are admitted to the sample. Also, in the lepton plus jets channels, if for any jet the difference between the fractions of the jet energy in the coarse hadronic and electromagnetic calorimeters is greater than 0.5 (typical of Main Ring energy deposition), the event is discarded.

In addition, events which were recorded during the MRBS or MICRO_BLANK windows are processed further to remove Main Ring energy depositions [63] and included in the cross section analysis for some of the decay channels. Due to the degradation in jet and $\not\!\!E_T$ resolution introduced by this procedure, such events are not admitted to the mass analyses.

As a general cleanup procedure during DØRECO, large hadronic deposits in isolated cells are removed, since such deposits are likely to arise from a hardware problem (a "hot cell"). The algorithm defines isolation by comparing the energy in a given cell with that in its longitudinal, but not transverse, neighbors. This can lead to the improper removal of energy from a jet, and give rise to a false $\not\!\!E_T$ signature. To protect against this, events in which a cell within a jet was removed are only retained if they would have passed the $\not\!\!E_T$ cut regardless. In addition, the events are passed though DØRECO with the hot cell removal disabled in order to calculate the proper jet energy for the top quark mass analyses.

Finally, there are two events which have, in addition to the final state objects expected from $t\bar{t}$ decay, a good photon candidate. Such events are retained in the cross section analyses, but rejected in the mass analyses as there is no kinematic hypothesis for the presence of a photon.

5.4 Dilepton Channels

While the $t\bar{t}$ branching ratio to dileptons is small, there are also relatively few background processes which produce two leptons in association with significant jet activity, allowing for the isolation of a sample of events with a reasonable signal-to-background ratio. The dilepton backgrounds come from a variety of processes, none of which has a large cross section. The importance of any given background source depends on the channel being considered.

The kinematic selection of dilepton events is summarized in Table 5.2, and is designed to isolate those events with the expected final-state signature. When selecting

	ee	$e\mu$	$\mu\mu$
Leptons	$E_T >$ 20 GeV $\|\eta\| < 2.5$	E_T(e) $>$ 15 Gev, $p_T(\mu) > 15$ GeV/c $\|\eta(e)\| < 2.5$	$p_T(\mu) > 15$ GeV/c
Jets	≥ 2 with $E_T >$ 20 GeV and $\|\eta\| < 2.5$		
$\not{E}_T$	$>$ 25 GeV	$\not{E}_T >$ 20 GeV $\not{E}_T^{\text{cal}} > 10$ GeV	N/A
H_T^e	$>$ 120 GeV	$>$ 120 GeV	$>$ 100 GeV

Table 5.2: Kinematic cuts for the dilepton event selection. The cut used in place of $\not{E}_T$ to reject $Z \to \mu\mu$ events is described in the text, as is the H_T^e variable. Also, the muon η cut is run-dependent, as detailed in Chapter 4.

muons, an isolation cut is applied in addition to the identification criteria described in Chapter 4. A muon is defined as isolated if there is no jet reconstructed within $\Delta R = 0.5$ of the muon in $\eta - \phi$ space.

In addition, the variable H_T^e, defined as

$$H_T^e = \left(\sum_{\text{jets}} E_T\right) + E_T(\text{leading electron})$$

is found to give good rejection for background processes. The sum is over all jets with $E_T >$ 15 GeV and $|\eta| < 2.5$. Muon p_T is not included in the sum due to its poor resolution, so for the dimuon channel this reduces to a sum over the jet E_Ts. This also means that the 100 GeV cut on H_T^e applied in the dimuon channel is more restrictive than the 120 GeV cut applied in the dielectron and $e\mu$ channels.

The number of background events expected from the dominant sources in each dilepton channel after the selection cuts are applied is given in Table 5.3. In the following sections, details of the analysis for each channel are given.

5.4.1 $e\mu$

The $e\mu$ channel is the most privileged dilepton channel, having twice the branching ratio of the ee and $\mu\mu$ channels while being free from much of the background from

Background Process	ee	$e\mu$	$\mu\mu$
$Z \to \ell\ell$	0.058 ± 0.013	N/A	0.56 ± 0.22
$Z \to \tau\tau \to \ell\ell$	0.078 ± 0.025	0.099 ± 0.076	0.03 ± 0.02
W pair production	0.083 ± 0.026	0.074 ± 0.019	0.007 ± 0.004
Drell-Yan	0.054 ± 0.033	0.002 ± 0.003	0.07 ± 0.03
Instrumental backgrounds	0.197 ± 0.052	0.04 ± 0.13	0.07 ± 0.03

Table 5.3: Number of background events expected in each dilepton channel for the dominant sources. The nature of the instrumental backgrounds differs somewhat between the channels, as described in the text.

Z decay. The largest remaining background is $Z \to \tau\tau \to e\mu$, the level of which is reduced by both branching ratio and kinematics from that of $Z \to \ell\ell$. Additional physics backgrounds arise from W boson pair production and other even rarer sources. For the majority of these sources, the level of background is calculated beginning with a measured cross section (for inclusive Z production, for example) and scaled down by the efficiency for Monte Carlo events of that source to pass the kinematic cuts and by the particle identification and trigger efficiencies observed in data. (For the WW background, the input cross section is a theoretical prediction rather than an experimental measurement.) By comparison of the hadronic energy in DØś $Z \to ee$ events with that in the ISAJET model of $Z \to \ell\ell$, it is found that the efficiency from the Monte Carlo must be multiplied by 1.9 ± 1.2 to properly calculate this background.

There is also the possibility of particle misidentification creating an apparent $e\mu$ event. One example of this is the production of a single W boson in association with three or more jets. If the W decays to $\mu\nu$ and one of the jets is misidentified as an electron the event may enter the sample. The amount of background from this source is obtained by multiplying the observed number of W plus three or more jet events (where one of the jets has electromagnetic energy fraction > 0.90) by the probability that such jets would pass the electron identification criteria.

Three events are observed in this channel, and the background is calculated to be

0.21 ± 0.16 events [64].

5.4.2 Dielectron

The primary sources of physics background in the ee channel arise from Z boson production with additional jets produced by gluon radiation. As these events have no neutrinos, any imbalance in the total transverse momentum must be due to the detector resolution, and hence a $\not\!\!E_T$ cut is effective at reducing this background. To further reject Z events without appreciably reducing the $t\bar{t}$ acceptance, the event is required to have $\not\!\!E_T > 40$ GeV if the dielectron invariant mass is within 12 GeV/c^2 of the Z pole.

The level of this background is calculated entirely from data. One begins with a sample of multijet events and selects those most compatible with the kinematic cuts imposed in the $t\bar{t}$ selection (*e.g.* the events must have $H_T > 70$ GeV), and counts the fraction passing the $\not\!\!E_T$ cut. This defines the fraction of the time that detector resolution will produce a false $\not\!\!E_T$ signal. Then one considers the total dielectron data sample which passes all cuts except for $\not\!\!E_T$. This sample is still predominantly composed of Z events, and so multiplying the number of events in this sample by the probability for an event to give a false $\not\!\!E_T$ signal gives the total number of $Z \to ee$ events that one expects in the signal region.

Besides the direct decay to ee, Z bosons can also create background through the decay $Z \to \tau\tau \to ee$. As there is no well-defined parent sample in the data for studying this background, one relies on a Monte Carlo simulation to calculate the rejection power of the kinematic cuts against this background. This efficiency is then multiplied by the known Z production cross section and $Z \to \tau\tau \to ee$ branching ratio to give an absolute prediction of the number of $Z \to \tau\tau$ events expected.

The last physics background considered are Drell-Yan and WW production. For the former, the cross section is based on the DØ measured value for $30 < m_{ee} < 60$ GeV/c^2, divided by the fraction of MC events that fall into this range. For the latter,

a theoretical calculation of the cross section is used. In both cases, the Monte Carlo event kinematics are used to estimate the rejection power of the selection cuts. The background from these sources is found to be quite small.

Next one must evaluate the background arising from jets being misidentified as electrons. This can happen either in $W+$ jets events where one of the jets is misidentified, or in multijet events where two of the jets are misidentified. To determine the number of events that one expects from this source, a parent sample containing two electromagnetic jets, one of which passes the electron identification criteria, is selected. Then the second electromagnetic jet is treated as an electron and all of the kinematic cuts are applied. The number of events in this sample is multiplied by the probability that a jet will pass the electron ID criteria given that the jet has large electromagnetic content to give the absolute number of expected events.

One event passes all of the selection cuts, and the background is 0.47 ± 0.09 events [65]. However, this background is estimated including the luminosity recorded during the MICRO_BLANK and MRBS_LOSS gates. As the mass analysis vetoes such events, the background estimate must be scaled to match the luminosity recorded outside of these gates. The appropriate scaling factor is 0.87 [66], so that the background to be considered in the mass analysis is 0.41 ± 0.08 events.

Event 95653/10822: A Special Case

Often when one is seeking a rare signal from a large data set, there are events which in retrospect appear to have all of the properties of the expected signal and yet fail the event selection. If the event simply falls below one of the kinematic cuts, it is clear that it must be excluded from the sample, but the situation is less certain when the event fails due to an apparent shortcoming in the selection. In the analysis of the dielectron channel, one such event (95653/10822) was found. The final state consists of two electromagnetic and two hadronic clusters, and significant $\not\!\!E_T$. One of the EM clusters was identified as a photon, as there was no reconstructed track

reconstructed pointing to it. However, if one were to project a line from the center of the cluster to the event vertex, one finds that the line traverses only the inner two layers of the CDC. If one then looks at the hit information from the CDC, one finds that hits exist in these layers pointing towards the EM cluster. Futhermore, the TRD information is very consistent with the passage of an electron in that direction. It is a clear shortcoming of the event reconstruction that one- or two-layer CDC segments are not formed into tracks when the segments pass out of the fiducial coverage of the CDC in the outer layers. Once the "photon" is reinterpreted as an electron, this event passes all of the selection criteria. In addition, one of the hadronic clusters is associated with a soft muon (this in the only dilepton candidate to have a muon-tagged jet) further lending support to the hypothesis that the event is an example of $t\bar{t}$ production. For these reasons, the event is allowed to enter into the dilepton mass analysis, although it is still excluded in the calculation of the $t\bar{t}$ cross section.

One needs an estimate of the additional background introduced by admitting this event to the sample. This is obtained by considering an extension of the dielectron analysis, in which the same kinematic criteria are applied as in the standard analysis, but only one of the two electromagnetic clusters is required to have an associated track if one of the jets is muon tagged. To estimate the background associated with this extension, one may begin with the background to the e+jets/tag channel, 0.97 events. Then, taking into account the probability for a hadronic jet to produce an isolated electromagnetic cluster ($\approx 8 \times 10^{-4}$[67, 68]), as well as the average jet multiplicity in the W+jets and multijet background models, one would expect about 0.003 e+jets/tag background events with an additional EM cluster. One must then account for the fact that the ee kinematic selection is somewhat looser than the e+jets/tag selection, as the latter applies an aplanarity cut and also restricts jets and leptons to $|\eta| < 2.0$. The background models predict that the rejection power of the ee cuts is only about one-third that of the e+jets/tag cuts, bringing the total additional background to about 0.01 events in the extended ee analysis.

Finally, one needs to account for the fact that this extension would not be con-

sidered had an event such as 95653/10822 not been observed. Based on the known efficiency for detecting a muon-tagged jet in a $t\bar{t}$ event, as well as the tracking efficiency, one would expect to observe an additional 0.15 $t\bar{t}$ events in the extended analysis. So, only in one of every six experiments would an event be observed, and the extension made. Therefore, an estimate of the additional background admitted by extending the analysis to include events such as 95653/10822 is 0.06 events, bringing the total background for the extended analysis to 0.47 ± 0.09 events.

5.4.3 Dimuon

The dimuon channel shares the $Z \rightarrow \ell\ell$ background with the dielectron channel, but the poor resolution of the muon momentum measurement makes separation of the $t\bar{t}$ signal from this background more challenging. For example, the $\not\!\!E_T$ cut used by the dielectron analysis is not applicable since both the $\not\!\!E_T$ and the dilepton invariant mass are measured so inexactly. In order to deal with the background, a kinematic fit to the $Z \rightarrow \mu\mu$ hypothesis is applied [69], and the event is required to have χ^2 probability less than 1% for this fit.

The level of background surviving this cut is determined from applying the fit to Monte Carlo $Z \rightarrow \mu\mu$ samples. In order that the Monte Carlo generation be reasonably efficient for producing events with two jets in association with the Z boson, some cut is usually placed on the Monte Carlo generation, either by requiring that the Z p_T be above a certain value, or that there be two final-state partons included in the hard process calculation. Rather than relying on a theoretical calculation of the Z cross section after these cuts, one proceeds by applying analogous cuts to the $Z \rightarrow ee$ event sample, and scaling the well-measured inclusive Z production cross section by the fraction of events which pass the cut. An identical procedure is used to estimate the contribution of $Z \rightarrow \tau\tau \rightarrow \mu\mu$ events.

Even after the Z-fit χ^2 cut, Z boson production remains the dominant background source. The level of the remaining physics backgrounds, principally $WW \rightarrow \mu\mu$ and

	e + jets	μ + jets
Lepton	$E_T >$ 20 GeV and $\|\eta\| < 2.0$	$p_t > 20$ GeV/c
Jets	Topol.: ≥ 4 with $E_T > 15$ GeV and $\|\eta\| < 2.0$ μ-tag: ≥ 3 with $E_T > 20$ GeV and $\|\eta\| < 2.0$	
$\not\!\!E_T$	> 25 GeV	Both cal. and μ-corrected $\not\!\!E_T > 20$ GeV

Table 5.4: Kinematic and fiducial cuts for objects in lepton + jets selection. The fiducial cuts for muon selection are as described in Chapter 4.

Drell-Yan, are determined solely from the Monte Carlo acceptance multiplied by the theoretical cross section.

The background due to heavy quark jets being misidentified as isolated muons is determined from the sample of events with a single muon and three or more jets. Each jet in the sample is then multiplied by its probability for appearing as an isolated muon to give the total background.

The total background in this channel is 0.73 ± 0.25 events, and one event survives the cuts. Further details on the analysis may be found in [70].

5.5 Lepton Plus Jets Channels

The process of isolating a sample of $t\bar{t} \rightarrow \ell$+jets events begins by requiring the event to contain the expected objects: a high p_T lepton, significant $\not\!\!E_T$, and high p_T jets. The kinematic and fiducial cuts placed on the final-state objects are given in Table 5.4. However, a sample chosen merely by requiring these final state objects will be dominated by events from two background sources: W plus multijet production, in which the final-state objects are identical to those expected from $t\bar{t}$ decay, and QCD multijet production, in which one of the jets is misidentified as an electron or muon and the $\not\!\!E_T$ arises from mismeasurement of the energies in the event.

Two strategies are employed to reduce these backgrounds. The *topological* selection places cuts on kinematic quantities which distinguish $t\bar{t}$ events from background, and

the *b-tagging* selection exploits the fact that $t\bar{t}$ events should have two b quarks while the background jets arise predominantly from light quarks or gluons.

5.5.1 Topological Selection

In the topological selection one needs to identify variables which cleanly distinguish signal from background. Two variables which have been found to be useful are:

- $H_T \equiv \sum_{\text{jets}} E_T$. The sum is over all jets with $E_T > 15$ GeV and $|\eta| < 2.0$.
- Aplanarity ($\mathcal{A}$) $\equiv 3/2\times$ the smallest eigenvalue of the normalized laboratory momentum tensor ($\mathcal{M}$). This tensor is defined such that

$$\mathcal{M}_{ij} = \frac{\sum_{o=\text{Objects}} p_{o,i} p_{o,j}}{\sum_{o=\text{Objects}} |\vec{p_o}|^2}$$

  where $\vec{p}_o$ is the three-momentum of an object and the objects considered in the sum are the jets and reconstructed W.

The effectiveness of both of these variables arises from their ability to distinguish jets from the decay of a heavy object (which tend to have large transverse energies and spherically symmetric directions) from those arising from gluon radiation. The separation between signal and background in these variables is shown in Fig. 5.1.

Once the proper variables have been identified, it remains to select the ideal cut points for these variables. This is done using a *random grid search* [71] in which an array of possible cut points is tested on the signal and background models. In order for the search to be efficient (*i.e.* to avoid unpopulated regions of $(\mathcal{A}, H_T)$ space) the set of cut points considered is that given by the distribution from $t\bar{t}$ Monte Carlo. For each prospective $(\mathcal{A}, H_T)$ cut the expected number of signal and background events given the data luminosity is calculated, and from this, the expected fractional uncertainty in the cross section measurement is determined, as shown in Fig. 5.2.

The cut point which minimizes this uncertainty is optimal, and is found to be:

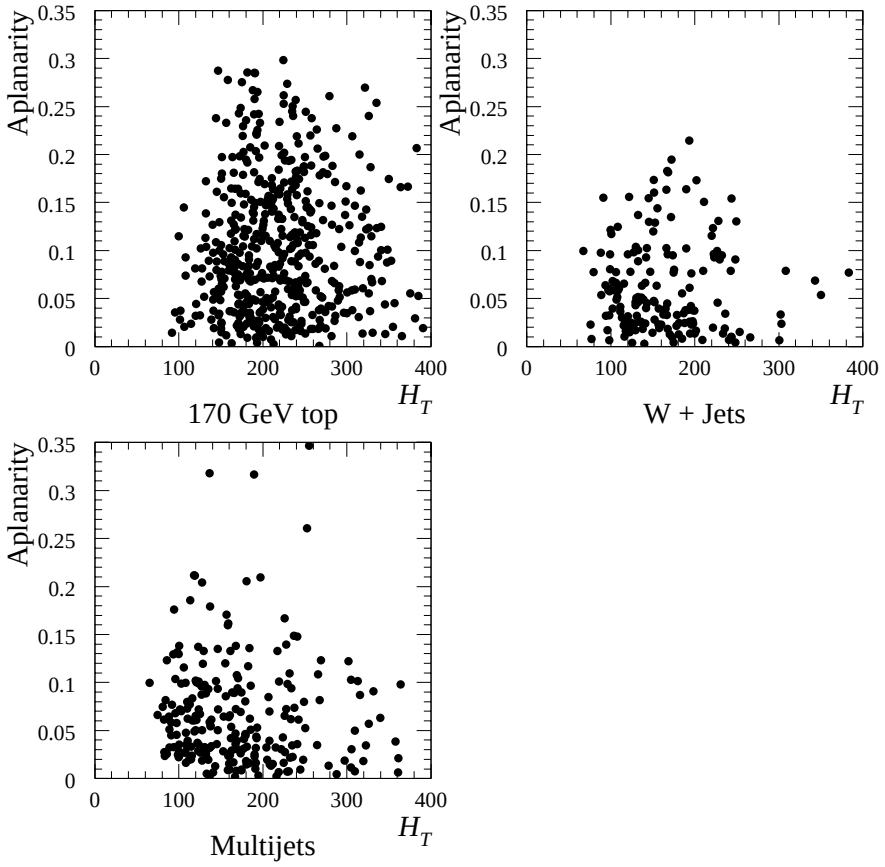

Figure 5.1: Scatter plot of the A and H_T distribution of 170 GeV HERWIG $t\bar{t}$ events, VECBOS W plus four jet events, and multijet events in the data.

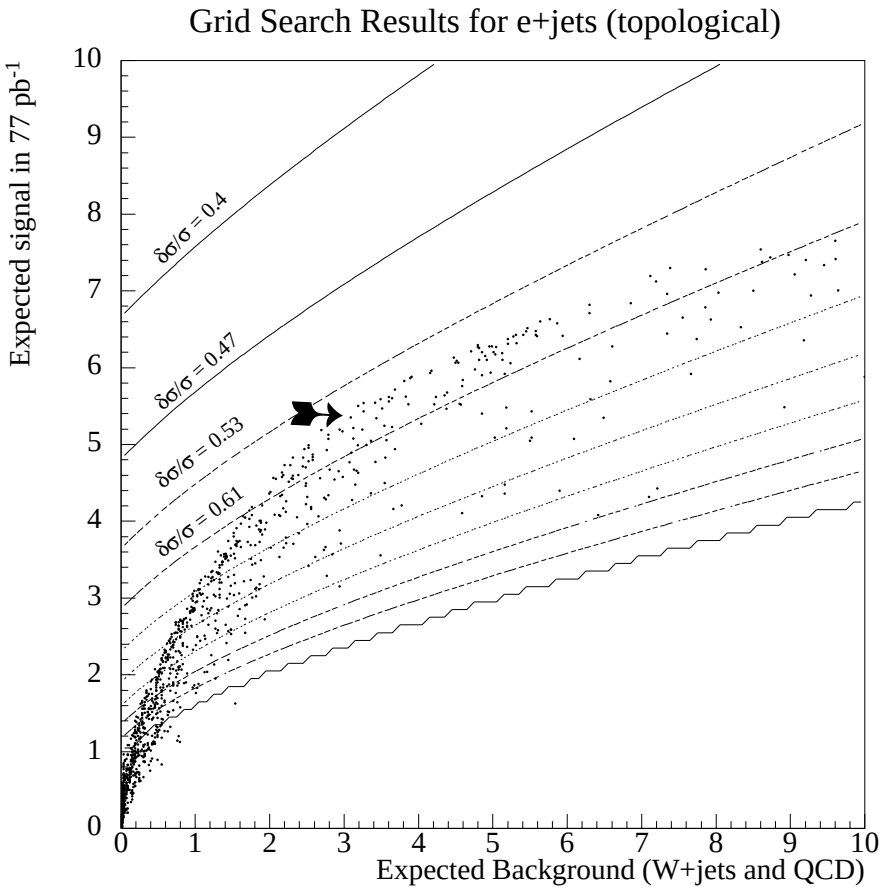

Figure 5.2: Scatter plot of expected signal and background events in the e + jets channel for many possible choices of the $(\mathcal{A}, H_T)$ cut point. Lines of constant expected uncertainty on the cross section are overlaid, and the optimal cut point (0.065, 180 GeV) is indicated by the arrow [72].

- $H_T > 180$ GeV
- $\mathcal{A} > 0.065$

In addition, the cut $E_T^W \equiv E_T(\ell) + \not{E}_T > 60$ GeV is applied in order to further reduce the multijet background.

For computing the number of expected background events, the multijet portion of the background is first calculated as a function of the inclusive jet multiplicity for a sample in which the H_T, $\mathcal{A}$, and E_T^W cuts are not applied. Different physical processes give rise to the multijet background in the e plus jets and μ plus jets channels. In the former, any jet with large electromagnetic content may appear as an electron. In the latter, the background is from the semimuonic decay of a heavy quark, in which the decay kinematics are such that insufficient hadronic energy remains to form a jet, leaving an apparently isolated muon in the event.

To reflect these differences, different techniques are used to assess the background level. For the e plus jets case, one compares the $\not{E}_T$ spectrum for samples of electrons which pass the identification criteria and those which fail them. The latter sample has a rapidly falling $\not{E}_T$ spectrum, since there are few real neutrinos and the $\not{E}_T$ arises from object mismeasurement. One then normalizes the number of events at low $\not{E}_T$ (< 10 GeV) to that in the good electron sample. The number of bad electron events which have $\not{E}_T > 25$ GeV multiplied by this normalization factor gives the multijet background in the good electron sample.

For the μ plus jets channel, one models the isolated muon plus N jet background by considering the sample of non-isolated muons plus $N+1$ jets, with again only the parent sample cuts applied. The number of events in this sample, multiplied by the probability that a muon in a jet would appear isolated, gives the multijet background for the N jet inclusive sample. The latter probability is given by the ratio at low $\not{E}_T$ of events with an isolated muon and N jets to events with a non-isolated muon and $N+1$ jets (the presumption being that W production does not contribute greatly to either sample).

Once the number of multijet events in the four-jet inclusive parent sample has been determined, it is multiplied by the efficiency of the topological cuts as measured using the samples of multijet events.

In order to estimate the magnitude of the purely physics background, that arising from the production of a W boson in association with jets, the N_{jets} scaling approximation [73, 58] is applied. The N_{jets} approximation is:

$$\frac{\sigma\,(W + N \text{ jets})}{\sigma\,(W + (N-1)\,\text{jets})} = R \tag{5.1}$$

where R is a constant and n is the inclusive jet multiplicity. Using it, one can determine the number of W plus four or more jet events in the data as follows: define N_i^{obs} as the number of events passing the lepton and $\not\!\!E_T$ cuts with i or more jets, with the expected multijet contribution subtracted. Then one can perform a fit to determine R:

$$N_i^{\text{obs}} = N_1^W \times R^{i-1} + N_i^{\text{top}};\ i = 1, \ldots, 4 \tag{5.2}$$

where N_1^W is the number of W plus one jet inclusive events (and is also an output of the fit) and N_i^{top} is the expected number of $t\bar{t}$ events with the appropriate multiplicity, determined from the theoretical cross section and Monte Carlo multiplicity distribution. Then the expected number of W plus four or more jets is given by:

$$N_4^W = N_1^W R^3 \tag{5.3}$$

The advantage of this method is that it is independent of theoretical calculations of the W production cross section, which have a large uncertainty for high jet multiplicities. The number of $W \to e\nu$ events as a function of jet multiplicity is shown in Fig. 5.3.

5.5.2 b-tag Selection

As an alternative to the application of tight topological cuts, one can also isolate a sample of $t\bar{t}$ events by requiring that at least one of the jets is likely to arise from

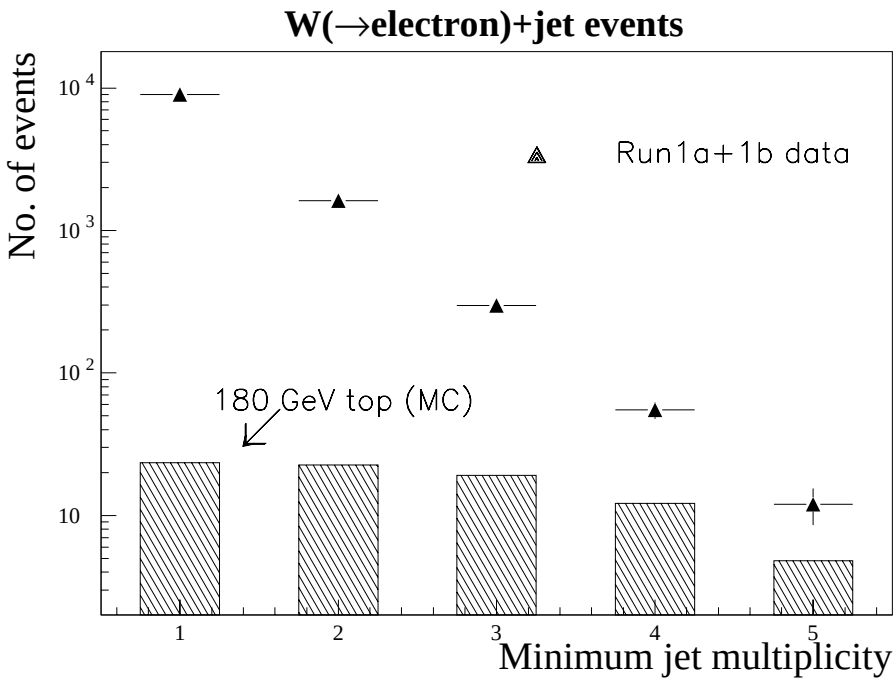

Figure 5.3: Number of $W \to e\nu$ events as a function of jet multiplicity, along with the expected contribution from $t\bar{t}$ production [72].

b decay. One factor which distinguishes b quarks from u and d quarks and gluons is their semileptonic decays (both directly and through $b \to c \to \ell$). Those that involve a muon are of of particular interest, since the DØ detector can reliably identify muons even in the midst of significant hadronic activity.

Hence the crucial requirement for the b-tagging searches is that there exist a muon within $\Delta R < 0.5$ from the axis of one of the jets in the event. The identification criteria for this muon are identical to that for the isolated muon in the topological μ plus jets search, with the exception that the $\int \vec{B} \cdot d\vec{l}$ cut is not applied. Also, the minimum p_T for the tagging muon is only 4 GeV/c.

Aside from this requirement for the tagging muon, the kinematic requirements for objects in the b-tagging channels are quite similar to that in the topological channels. The exception is the jet requirement. Since the presence of a tagging muon greatly reduces the background, one needs only to require three jets to attain a good signal-to-background ratio. The jet E_T cut, however, is raised to 20 GeV.

Additional requirements are applied to ensure that the $\not{E}_T$ is not introduced by mismeasurement of the muon momentum. The e+jets channel raises the $\not{E}_T$ cut to 35 GeV if $\Delta\phi_{\not{E}_T,\mu} < 25°$. For the μ + jets channel a contour cut in the plane of $\Delta\phi_{\not{E}_T,\mu}$ vs $\not{E}_T$ is made, where the muon considered is the highest momentum muon, be it isolated or tagging. The event passes only if:

- $\Delta\phi_{\not{E}_T,\mu} < 170°$, and
- $\dfrac{|\Delta\phi_{\not{E}_T,\mu} - 90°|}{90°} < \dfrac{\not{E}_T}{45 \text{ GeV}}$

The same Z-fit χ^2 cut used in the dimuon channel is applied to remove Z boson events.

Finally, there is still gain to be had in the muon-tagged channels from placing cuts on $\mathcal{A}$ and H_T. However, these cuts are less stringent than in the topological analyses:

- $H_T > 110$ GeV
- $\mathcal{A} > 0.04$

The sources of background in the b-tagging channels are similar to those in the topological searches ($Z \rightarrow \mu\mu$ is an additional background in the tagged μ plus jets channel), but the methods for calculating the level of background are somewhat different. It is critical to understand the probability for a given jet to have a muon within its cone. This probability is measured by looking at the number of tagged jets in a large sample collected from a multijet trigger, and is found to increase linearly with increasing jet E_T.

The calculation of the backgrounds due to multijet events with jets misidentified as electrons or isolated muons then proceeds similarly to that for the topological channels. The W + multijet backgrounds are estimated by considering the sample of events without a tagging muon which pass all of the kinematic cuts save for $\mathcal{A}$ and H_T. Then the probability for each jet in this sample producing a false tag is calculated, and the total probability is summed to give the expected number of false tags. This number is then multiplied by the efficiency of the $\mathcal{A}$ and H_T cuts when applied to this sample to give the total number of events expected in the final sample. As in the topological analysis, this provides a method of determining the W + jets background without theoretical input on the cross section for this process.

The small background in the tagged μ + jets channel from $Z \rightarrow \mu\mu$ events is calculated from Monte Carlo samples, and does rely on the calculated Z plus multijet production cross section.

5.6 The $e\nu$ Channel

The remaining channel used in the cross section measurement is the $e\nu$ channel, which has acceptance in regions of phase space rejected by the other channels. One example is a dilepton event where one W decays to an energetic electron and soft neutrino, and the other W decays to an energetic neutrino and soft charged lepton. Then is is quite likely that the soft charged lepton will fall below the kinematic cuts imposed in the dilepton selection, and the event will also have too few jets for the

single lepton selection.

The selection for this channel is based on the above scenario, and imposes tight $\not\!\!E_T$ (> 50 GeV) and $e\nu$ transverse mass (> 115 GeV/c^2) cuts to eliminate events containing only a single leptonic W as well as misidentified multijet events. In addition, the second-leading jet is required to have E_T greater than 30 GeV and the azimuthal angle between the $\not\!\!E_T$ and second-leading E_T object is required to be greater than 0.5 radians.

After these cuts are imposed, one finds that in addition to the above source of events, this channel also accepts events in which one of the W decays to a τ which decays electronically and the other W decays hadronically, as well as a small percentage of single-lepton events. In order to keep this channel orthogonal to the others, any events that pass the requirements of any of the above seven channels are explicitly vetoed.

As events in this channel contain insufficient information for the reconstruction of the event, they are not considered in the mass analysis.

Four events are observed in this channel with an expected background of 1.2 ± 0.4 [74].

5.7 Top Quark Production Cross Section

Summing the contributions from all eight of the channels discussed above, DØ observes a total of 39 $t\bar{t}$ candidate events with a background of 13.6 ± 2.2 events. This excess is a clear signal for $t\bar{t}$ production, and permits measurement of the production cross section. Input from all of the $t\bar{t}$ decay channels is combined in this measurement, using the equation:

$$\sigma_{t\bar{t}} = \frac{\sum_{i=1}^{8} N_i - B_i}{\sum_{i=1}^{8} (\epsilon\mathcal{B})_i \mathcal{L}_i} \tag{5.4}$$

where the sum is over the decay channels, N_i is the number of observed events, B_i the expected background, $(\epsilon\mathcal{B})_i$ the $t\bar{t}$ selection efficiency (including branching ratio),

	ee	$e\mu$	$\mu\mu$	$e\nu$
$\epsilon \times BR(\%)$	0.17 ± 0.02	0.35 ± 0.07	0.11 ± 0.01	0.26 ± 0.08
N_{top}	1.20 ± 0.18	2.20 ± 0.48	0.64 ± 0.09	1.66 ± 0.48
N_{bkg}	0.47 ± 0.09	0.21 ± 0.16	0.73 ± 0.25	1.16 ± 0.36
Observed	1	3	1	4

Table 5.5: Summary of the contribution of each dilepton channel to the cross section measurement. The expected number of $t\bar{t}$ events (N_{top}) is based on the central value of the theoretical cross section [10] and assumes a mass of 170 GeV/c^2.

	$e+$ jets	$e+$ jets/μ	$\mu+$ jets	$\mu+$ jets/μ
$\epsilon \times BR(\%)$	1.29 ± 0.23	0.57 ± 0.08	0.91 ± 0.27	0.37 ± 0.09
N_{top}	8.63 ± 1.57	3.59 ± 0.56	5.51 ± 1.67	2.25 ± 0.54
N_{bkg}	4.51 ± 0.91	1.05 ± 0.39	4.16 ± 1.02	1.39 ± 0.23
Observed	9	10	5	6

Table 5.6: Contributions of the lepton plus jets channels to the cross section measurement.

and $\mathcal{L}_i$ the luminosity. The individual channel results are summarized in Tables 5.5 and 5.6.

In assigning an uncertainty to this measurement, one must take care to properly account for all of the correlations between the uncertainties in each individual measurement [75]. While the channels are defined to be orthogonal, all rely on the same measurement of the luminosity, all share the uncertainty in jet energy scale, and many use the same Monte Carlo samples to model their backgrounds. The final result, assuming a top quark mass of 170 GeV/c^2, is:

$$\sigma\left(t\bar{t}\right) = 5.7 \pm 1.9 \text{ pb} \tag{5.5}$$

The measured cross section is a function of the top quark mass, as the efficiencies of the kinematic selection criteria tend to increase with increasing mass, as shown in Fig. 5.4.

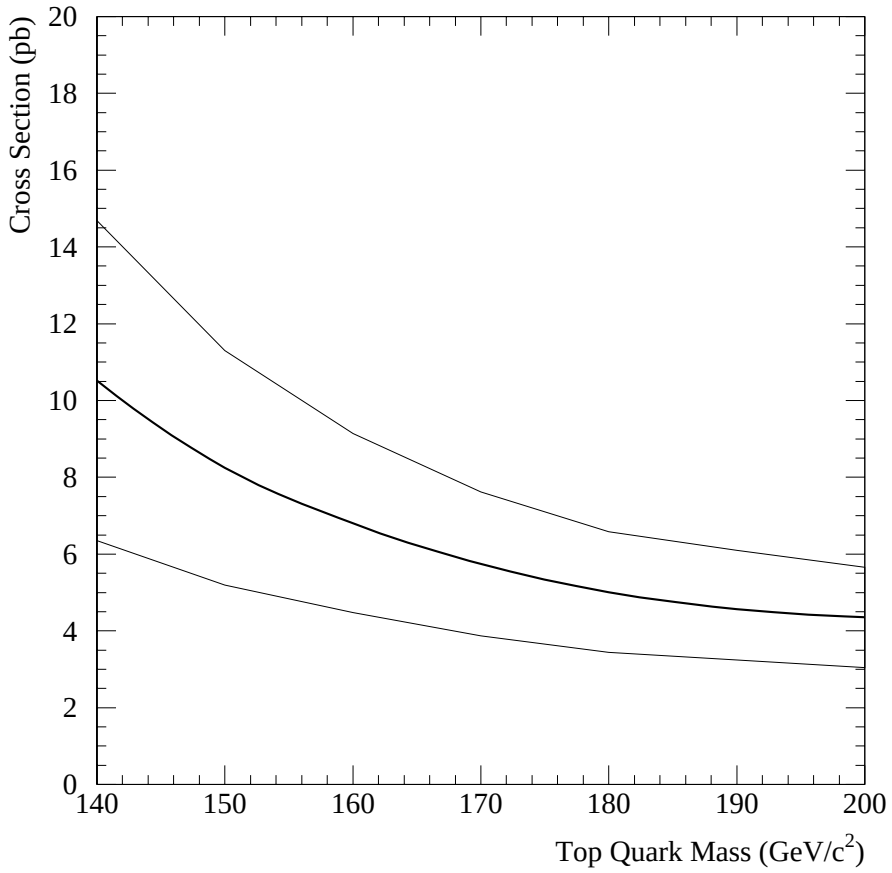

Figure 5.4: Measured $t\bar{t}$ production cross section as a function of top quark mass. The upper and lower curves bound the one standard deviation interval.

5.8 Lepton Plus Jets Cuts Used for Mass Analysis

The cuts in the lepton plus jets channels described above were selected by minimizing the expected error on the $t\bar{t}$ cross section. It is not surprising that a somewhat different selection proves optimal for measurement of the mass. In order to provide a base sample for the mass analysis, a loose selection is made which includes all of the event quality, particle identification, and trigger criteria mentioned above. Only the kinematic cuts are relaxed, the goal being to accept all events with a high-p_T lepton, four or more high-p_T jets, and significant $\not\!\!E_T$. (Also, in order to ensure that all objects are measured with the highest possible accuracy, no events with Main Ring activity are allowed in this sample).

The cuts for this sample require the objects necessary for a kinematic fit:

- One electron or muon with $E_T > 20$ GeV and $|\eta| < 2.0$
- At least four jets with $E_T > 15$ GeV and $|\eta| < 2.0$
- $\not\!\!E_T > 25$ GeV. In addition, the same contour and Z-fit χ^2 cuts are applied to events with non-isolated muons as in the cross section analysis.

Most of the events which satisfy these basic requirements are background. In order to reduce this, the following additional cuts are made:

- One of the jets is muon-tagged

or

- $E_T^W > 60$ GeV, and
- $|\eta^W| < 2.0$

The presence of a soft muon in a jet clearly favors the $t\bar{t}$ hypothesis, as detailed above. The first of the kinematic cuts for untagged events is particularly effective

at removing misidentified multijet events, while the second restricts the phase space to a region in which VECBOS provides a good model for the data. In reconstructing η^W one solves for the z-component of the neutrino momentum by imposing the W mass constraint and choosing the solution which gives the minimum longitudinal W momentum (this choice is found to be correct about 80% of the time for $t\bar{t}$ events). In the event that there is no solution for the z component (which is possible due to mismeasurement of the lepton or $\not{E}_T$), the neutrino is assumed to have the same η as the charged lepton.

While the cross section measurement applies a stringent cut on H_T to eliminate background, this is not ideal for the mass analysis, as the background events which survive this cut tend to have a reconstructed top mass in the same region as do $t\bar{t}$ events (see Chapter 6 for the definition of reconstructed mass). Clearly, one would prefer a cut that reduces the background underneath the mass peak. Therefore, one needs to identify variables that provide discrimination between signal and background independent of the reconstructed mass for the event. Four variables which are found to be useful are:

- $\not{E}_T$
- $\mathcal{A}$
- $H'_{T2} \equiv \dfrac{H_T - E_T(\text{jet } 1)}{\sum\limits_{\ell,\nu,\text{jets}} |p_z|}$
- $K'_{T\text{min}} \equiv \dfrac{\text{Min}\Delta R_{jj} \cdot E_T(\text{lesser jet})}{E_T^W}$

While the utility of the first two variables has been detailed above, the latter two require some explanation. H'_{T2} is a measure of the centrality of the event as a whole, and is effective because $t\bar{t}$ events tend to be more central than background. In the definition of $K'_{T\text{min}}$, the numerator is the smallest ΔR between any pair of jets taken from the leading four multiplied by the E_T of the softer jet in this pair. This is useful

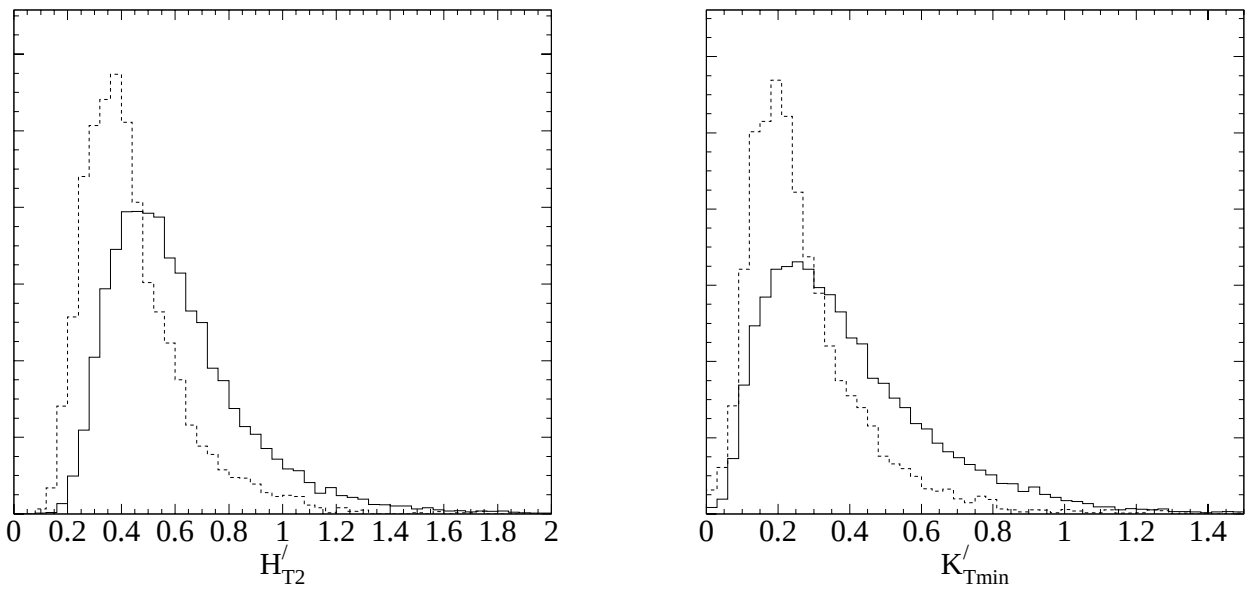

Figure 5.5: Separation between Monte Carlo 170 GeV/c^2 $t\bar{t}$ (solid) and background (dashed) events obtained from the variables H'_{T2} and $K_{T\min}$

as $t\bar{t}$ events have four independent jets, while final state gluon radiation tends to give rise to some of the jets in background events. Hence one expects the background to have more closely-spaced jet pairs. The denominators of H'_{T2} and $K'_{T\min}$ serve to eliminate most of the correlation with reconstructed top quark mass.

Figure 5.5 shows the distribution of these two variables for $t\bar{t}$ events and for the expected background. The separation between the two samples is evident. In order to make optimal use of the information obtained from all four likelihood variables, a multivariate technique is used. For each variable, the relative density of the signal and background samples as a function of the value of the variable is parameterized, to give $\mathcal{L}_i(v_i)$. The the four $\mathcal{L}_i$ are combined to give an overall likelihood:

$$\mathcal{L} = \sum_i w_i \mathcal{L}_i$$

where the w_i are chosen to cancel any residual correlation between $\mathcal{L}_i$ and the recon-

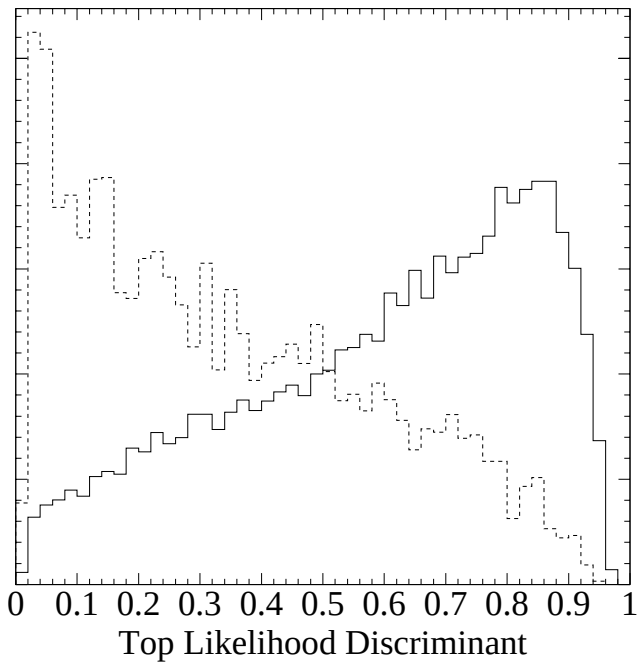

Figure 5.6: Top likelihood discriminant function, for signal (solid) and background (dashed) events.

structed top quark mass. $\mathcal{L}$ is then used to define a top likelihood discriminant:

$$\mathcal{D} = \frac{\mathcal{L}}{1 + \mathcal{L}} \tag{5.6}$$

The range of $\mathcal{D}$ is zero to one, with $t\bar{t}$ events tending to large values as shown in Fig. 5.6. The use of this variable in determining the top quark mass is detailed in Chapter 7.

As an alternative to the discriminant described above, one may also input the four mass-unbiased variables into an artificial neural network and select events based on the output of the network. Both of these approaches are used in the mass analysis to estimate the likelihood that an event in the mass sample is an example of $t\bar{t}$ production, as described in Chapter 7 and [76].

Chapter 6

Mass Reconstruction

Once a set of candidate $t\bar{t}$ events has been selected, the determination of the top quark mass based upon the characteristics of those events requires two distinct steps. The first is the determination of the top quark mass most consistent with the kinematics of each event, and the second is the comparison of the distribution of these "fitted masses" to that expected for various values of the top quark mass. This chapter describes the first step, and the next chapter the second.

6.1 Dilepton Events

In the dilepton channels, one expects the final state to consist of two charged leptons, two neutrinos, and two b jets (see Fig. 6.1) so that the final state is completely specified by knowledge of the energy four-vectors of these six particles. Since the mass of each final-state particle is known, this reduces to 18 independent quantities.

The four-vectors of the jets and charged leptons are measured directly in the detector, as is each component of the $\not{E}_T$. In addition, there are three kinematic constraints:

- The invariant mass of each lepton and neutrino pair is equal to the W mass.

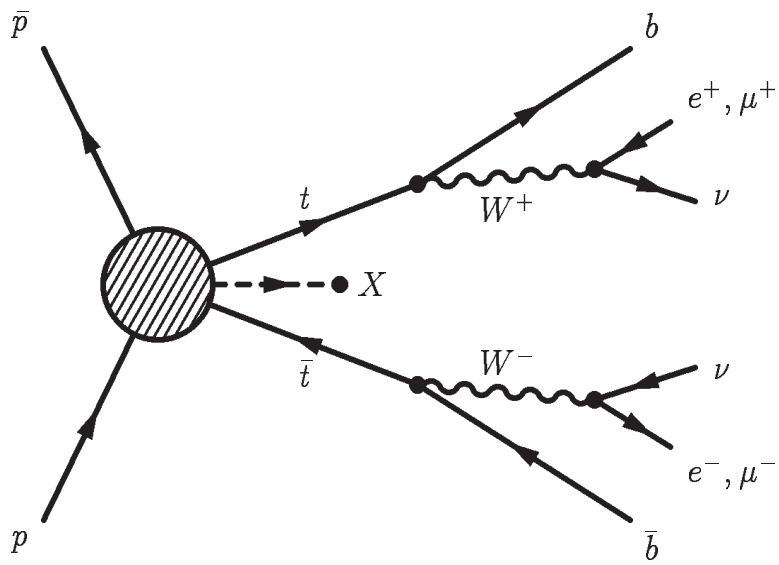

Figure 6.1: Schematic representation of $t\bar{t}$ production and decay in the dilepton channels.

- The masses of the reconstructed t and $\bar{t}$ in the event are equal.

So in total we have seventeen pieces of information, when eighteen are needed to specify the $t\bar{t}$ system. A procedure must be developed to overcome this deficit of information and extract top quark mass information from dilepton events.

To do so, note that the imposition of one additional constraint is enough to render the system soluble. One may provide this constraint by *assuming* a value for the top quark mass. Some information will then be gained from the fact that not all assumed top quark mass values will prove compatible with the observed variables. However, simply requiring that the event be soluble for a given assumed mass is insufficient, as events can in general be solved for a wide range of top masses. Therefore, one needs to calculate the likelihood of any one solution being correct. That is, one wants to measure $P(m_t|\{o_i\})$, where $\{o_i\}$ is the set of 14 variables observed in the event. This

can be achieved through Bayes' theorem [77]:

$$\mathcal{P}(m_t|\{o_i\}) = \frac{\mathcal{P}(\{o_i\}|m_t)\mathcal{P}(m_t)}{\int_0^\infty \mathcal{P}(\{o_i\}|m_t')\mathcal{P}(m_t')\mathrm{d}m_t'} \tag{6.1}$$

where $\mathcal{P}(m_t)$ represents any information one has about the top quark mass independent of the kinematics of the event. Such external information could, for example, be provided by the cross section measurement if one assumes that the calculated cross section as a function of top quark mass is correct. In this analysis, no such assumptions are made, so $\mathcal{P}(m_t)$ is taken to be a constant. The denominator on the right-hand side of Eq. 6.1 is a normalization factor which ensures that the probability that the top have some mass is unity.

The relevant feature, then, is that

$$P(m_t|\{o_i\}) \propto \mathcal{P}(\{o_i\}|m_t) \tag{6.2}$$

The right-hand side is proportional to the differential cross section into the region of phase space defined by the measured quantities and has the analytic form:

$$\mathcal{P} \propto \frac{1}{\sigma_{\mathrm{vis}}(m_t)} \int |\mathcal{M}\{v_i\}|^2 \rho_1(o_1)\ldots\rho_{14}(o_{14})\delta^4(m_i - M_i)d^{18}\{v_i\} \tag{6.3}$$

$\mathcal{M}$ is a matrix element representing $t\bar{t}$ production and decay, the ρ's are normalized detector resolution functions, and the delta function enforces the W and top quark mass constraints on each side of the decay. The $\{v_i\}$ are any set of variables in which $\mathcal{M}$ can be conveniently calculated, such as LIPS variables. Dividing by $\sigma_{\mathrm{vis}}(m_t)$, the total visible cross section for a given top quark mass, is necessary to avoid bias that may enter due to the fact that the $t\bar{t}$ production cross section is larger for smaller values of m_t.

While Eq. 6.3 is exact, its solution is quite CPU-intensive. It happens that this is not a serious drawback to its application on the few candidate events in our data sample, but renders tests of the method on large Monte Carlo samples impractical. Therefore, one needs an approximation to this probability which retains sensitivity to the top quark mass but is simpler to compute.

6.2 The Neutrino-Weighting Method

As a first simplification, let us define $\{o_i\}$ to be the set of three-momenta of each final-state particle, and assume that the detector resolution is perfect. Also, let us ignore for the moment our measurement of $\not{E}_T$. Then we can replace twelve of the ρ's in Eq. 6.3 by delta-functions for the corresponding measured variables;

$$\mathcal{P} \propto \frac{1}{\sigma_{\text{vis}}(m_t)} \int |\mathcal{M}|^2 \rho_1(o_i)\rho_2(o_2)\delta_3(v_3)\ldots\delta_{14}(v_{14})\delta^4(m_i - M_i) \left|\frac{\partial(v_3\ldots v_{14})}{\partial(o_3\ldots o_{14})}\right| d^{18}\{v_i\} \tag{6.4}$$

Further simplification is needed, and so the Jacobian term and the dependence of $\mathcal{M}$ on sixteen of the variables are ignored, allowing trivial integration over the delta functions. The motivation is to focus on the final-state phase space rather than production probability, so the $1/\sigma_{\text{vis}}(m_t)$ term is also dropped. Then one has:

$$\mathcal{W} \propto \int |\mathcal{M}|^2 \rho_1 \rho_2 d^2\{v_i\} \tag{6.5}$$

where one is no longer justified in calling the right-hand side a probability. We now need to identify the remaining variables. Clearly, they must be quantities associated with the neutrinos, but which pair of the six possible variables is a matter of choice (the remaining four neutrino variables are fixed in enforcing the mass constraints). The choice made here is to integrate over the rapidities of the neutrinos:

$$\mathcal{W} \propto \int \mathcal{P}(\eta_1|m_t)\,\mathcal{P}(\eta_2|m_t)\,\rho_{\eta 1}\rho_{\eta 2} d\eta_1 d\eta_2 \tag{6.6}$$

The prior probabilities for the neutrino ηs are the remaining contributions of the matrix element. This integral is then evaluated numerically by summing over discrete choices for the neutrino rapidities:

$$\mathcal{W}(m_t|\{o_i\}) = \sum_{\eta 1 \eta 2} \mathcal{P}(\eta_1|m_t)\,\mathcal{P}(\eta_2|m_t)\,\rho_{\eta 1}\rho_{\eta 2}\Delta\eta_1\Delta\eta_2 \tag{6.7}$$

The $\mathcal{P}(\eta|m_t)$ terms are evaluated using the neutrino η distributions from HERWIG $t\bar{t}$ samples. Since this is not a quantity that is measurable in the detector, the true

value generated is used. Examples for a range of top quark masses are given in Fig. 6.2. The shape of the distribution is reasonably approximated by a Gaussian centered at zero, with a width that depends on the top quark mass, as shown in Fig. 6.3. A quadratic is used to parameterize this dependence, giving

$$\sigma_\eta\left(m_t\right) = 5.56 \times 10^{-6} m_t^2 - 2.16 \times 10^{-3} m_t + 1.314$$

with m_t in GeV/c^2. The above is then evaluated for any top quark mass assumed in Eq. 6.7 and $\mathcal{P}\left(\eta | m_t\right)$ is taken to be a Gaussian centered at zero with width $\sigma_\eta\left(m_t\right)$.

Rather than stepping in constant intervals of neutrino η and assigning each step the appropriate probability, the step size is decreased in regions of likely neutrino η such that one may treat each step as equally probable. This means that the step sizes are chosen such that each step spans an equal area of the appropriate Gaussian. In the current implementation, $N_\eta = 10$ values (at the 0.05, 0.15,...0.95 integral points of the appropriate Gaussian) are used for each neutrino.

It remains to evaluate the resolution factor $\rho_{\eta 1}\rho_{\eta 2}$. To do so, note that for any given assumption of the values of η_1 and η_2 the system can be solved without using the measured $\not{E}_T$. (The neutrino from t decay can be solved independently of that from the $\bar{t}$. This means that the CPU time required scales only as $2N_\eta$, not N_η^2). The equation which must be solved for each side of the decay is a quadratic, so there are a total of four solutions for each set of assumptions. Then the resolution terms are evaluated by comparing each component of the $\not{E}_T$ measured in the event with the sum of the neutrino momenta required by the solution:

$$\rho_{\eta 1}\rho_{\eta 2} = \sum_{\text{Solutions}} \prod_{k=x,y} \exp\left[-\left(\frac{\left(\not{E}_k - p_k(\nu\nu)\right)^2}{2\sigma^2(\not{E}_k)}\right)\right] \tag{6.8}$$

The error on each component of $\not{E}_T$ is taken to be 4 GeV. This is intended to represent the uncertainty due to mismeasurement of the underlying event, not the uncertainty caused by mismeasurement of the jets and leptons in the event (we are for the moment still assuming perfect resolution for these objects). If either side of the decay yields no real solutions, $\rho_{\eta 1}\rho_{\eta 2}$ is set to 0.

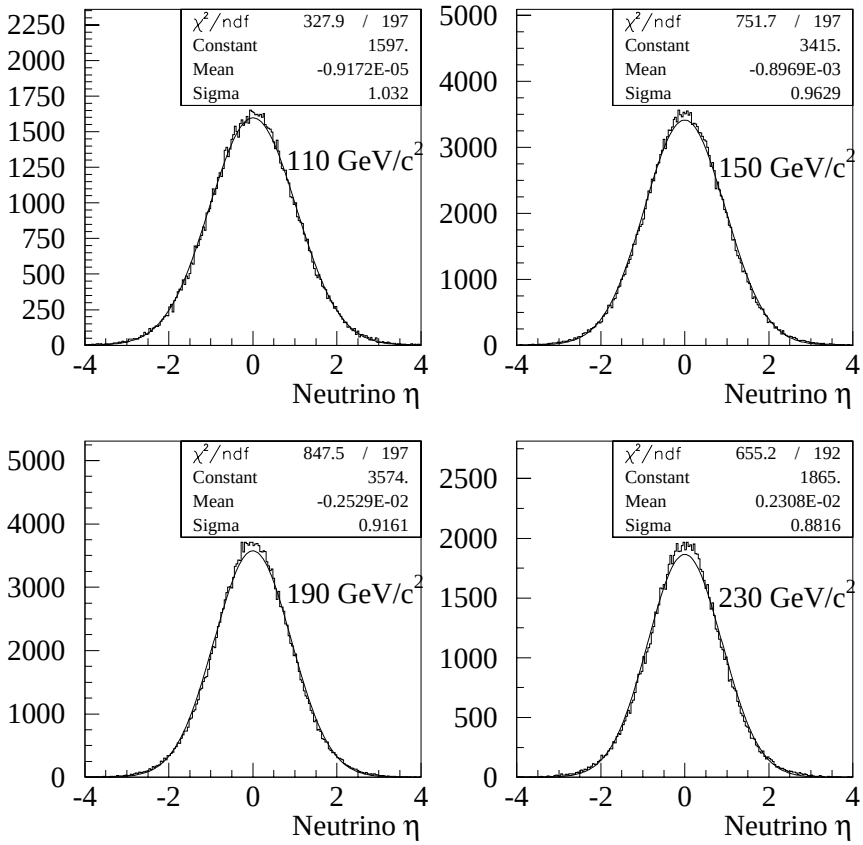

Figure 6.2: Distribution of neutrino rapidities from top quark decay as modeled by HERWIG, for a range of top quark masses (the two highest E_T neutrinos from each event are entered in the plots). It can be seen that the distributions are close to Gaussian, and that the width of the best-fit Gaussian decreases slightly with increasing top quark mass.

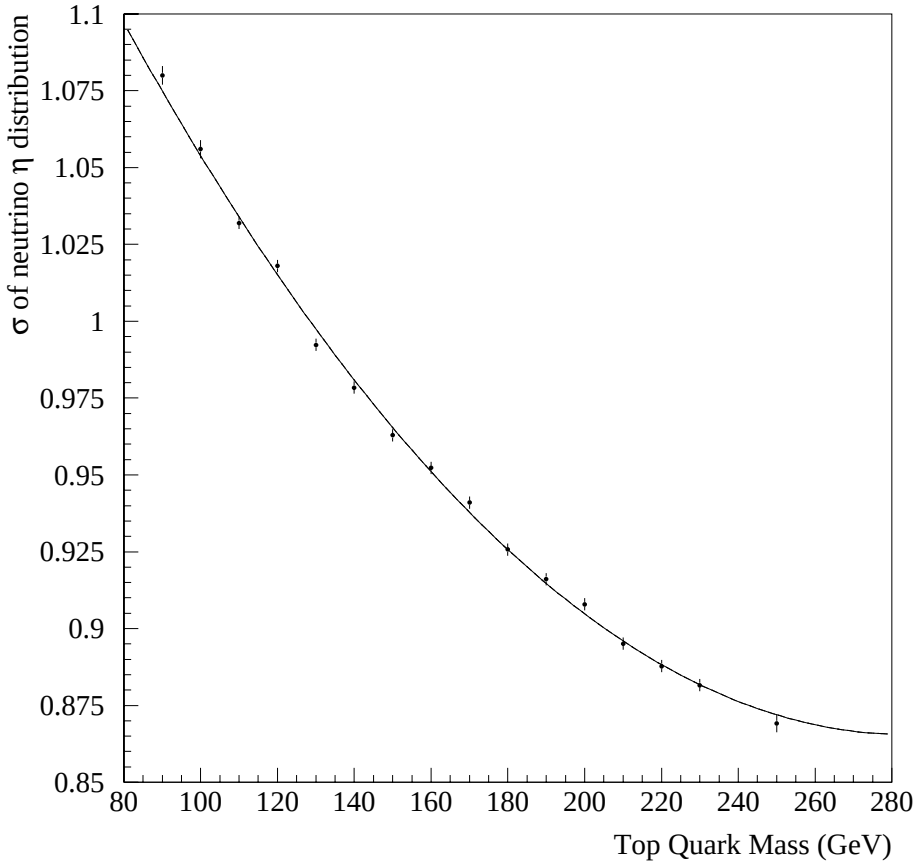

Figure 6.3: Width of the Gaussian that gives the best fit to the neutrino η distribution as a function of top quark mass. Also shown is the quadratic fit that parameterizes this dependence.

Substituting this into Eq. 6.7, one has the following definition of the weight as a function of top quark mass:

$$\mathcal{W}(m_t|\{o_i\}) = \sum_{\eta_1 \eta y_2} \sum_{\text{Solutions}} \prod_{k=x,y} \exp\left[-\left(\frac{(\not{E}_k - p_k(\nu\nu))^2}{2\sigma(\not{E}_k)^2}\right)\right] \tag{6.9}$$

The next step is to evaluate the above quantity for some range of assumed top quark masses. The W mass provides a lower bound on the allowed values (the necessary kinematic constraints do not apply for top quark masses below this). The upper bound is chosen somewhat arbitrarily based on our expectations from the calculated cross section. Unless these calculations are grossly in error, we do not expect to observe enough $t\bar{t}$ events to form a significant signal if the mass is much above 200 GeV/c^2. To be safely above this limit, the maximum value considered is 278 GeV/c^2, so that the range scanned is $82 \leq m_t \leq 278$ GeV/c^2, in 4 GeV/c^2 steps.

6.2.1 Modeling the Detector Resolution

In the above discussion we have assumed for simplicity that the momenta of the jets and leptons are measured perfectly in the detector. To approximate the integration over the ρ factors in Eq. 6.3 a Monte Carlo technique is used. The measured jet and lepton momenta are varied randomly within their resolutions, and $\mathcal{W}(m_t|\{o_i\})$ is re-evaluated. This procedure is repeated several times in order to approximate the integral over ρ.

In doing the event smearing, a compromise must be made between the desire to fully explore the solution space allowed by the measured quantities and the need to maintain reasonable throughput for Monte Carlo tests. In its current implementation the fitter smears each Monte Carlo event 100 times. The $t\bar{t}$ candidates are smeared 5000 times so that their weight as a function of assumed top quark mass is more accurately known. This is the only difference in the treatment of Monte Carlo and data events. In order to demonstrate that one is not significantly broadening the distribution of fitted masses for the Monte Carlo by only smearing 100 times, a

Number of smears	RMS (GeV/c^2)
0	31.7
10	29.5
100	28.0
1000	27.2

Table 6.1: RMS width of the distribution of peak $\mathcal{W}(m_t)$s as a function of the number of times each event is smeared for HERWIG 170 GeV/c^2 $t\bar{t}$ events.

sample of HERWIG 170 GeV/c^2 $t\bar{t}$ events was passed through the fitting procedure several times, with only the number of smears varied. As show in Table 6.1, the resolution on the top quark mass is improved by smearing, but the gains become small beyond 100 smears.

For electrons and muons, the smearing is done assuming Gaussian distributions with widths given by Equations 3.1 and 3.2 respectively. In modeling the jet resolution one needs to account for both the resolution of the hadronic calorimeter and the ambiguity in jet definition. For example, two partons with small separation in $\eta - \phi$ space may be reconstructed as a single jet, or a single parton which gives rise to a broad shower of hadrons may be reconstructed as two jets. The frequency with which these sort of confusions occur depends on the topology one is considering, and thus HERWIG $t\bar{t}$ Monte Carlo events are used to evaluate the jet resolution.

The study proceeds by comparing the reconstructed jet energy to that of the nearest cone-algorithm cluster of hadrons generated by the Monte Carlo in a sample of events ranging in top quark mass from 110 to 190 GeV/c^2. The jets are binned finely in E_T (10 GeV bins in the region where statistics are plentiful, and one overflow bin for high E_T jets) and coarsely in η (three bins representing the central ($|\eta| < 0.8$), intercryostat ($0.8 < |\eta| < 1.2$) , and endcap ($|\eta| > 1.2$) calorimeter regions). The distribution of fractional E_T difference for jets in any one of these bins typically consists of a narrow central peak due to the intrinsic calorimeter resolution and broad tails due to ambiguity in the jet definition. Therefore, a double Gaussian is fit to each

distribution, and the widths of the two Gaussians and their relative normalization are parameterized as a function of E_T. Jets are then smeared according to the double Gaussian appropriate to their E_T and η region. The fractional energy differences for some representative E_T bins for the three η regions is shown in Fig. 6.4, along with the double Gaussian approximation used in the smearing.

The Monte Carlo samples used for this study did not contain uranium noise or multiple interactions, and therefore the above resolutions are narrower than can be expected for actual data. To account for this, each jet is also smeared by a constant noise term of 5.1, 6.2, and 5.4 GeV for jets in the central, intercryostat, and endcap regions respectively. These values were determined from a study of the E_T balance in dijet events [78].

Finally, the $\not\!\!E_T$ is adjusted to account for the energy added to or removed from the other objects due to smearing. In addition, each component is smeared with $\sigma = 4$ GeV to represent uncertainty in the measurement of the underlying event.

6.2.2 Jet Combinatorics

One important problem neglected so far in the discussion of reconstructing dilepton events is that of assigning final-state jets to the appropriate partons. In the simplest case, when the only final-state objects observed are the two b jets and two charged leptons, there is a twofold ambiguity in pairing the jets with leptons. The fitter loops over both of these pairings, assigning equal weight to each.

Initial- and final-state gluon radiation (ISR and FSR) complicate the picture. One possibility for handling this is to take the leading two jets as the b jets and ignore any others in the fit.

As an alternative, one can consider all possible interpretations of the jets, and for each interpretation merge the jets assigned as bs with those assigned as their FSR and ignore those assigned as ISR, and estimate the probability of each jet interpretation being the correct one. This is done by defining a weight consisting of the following

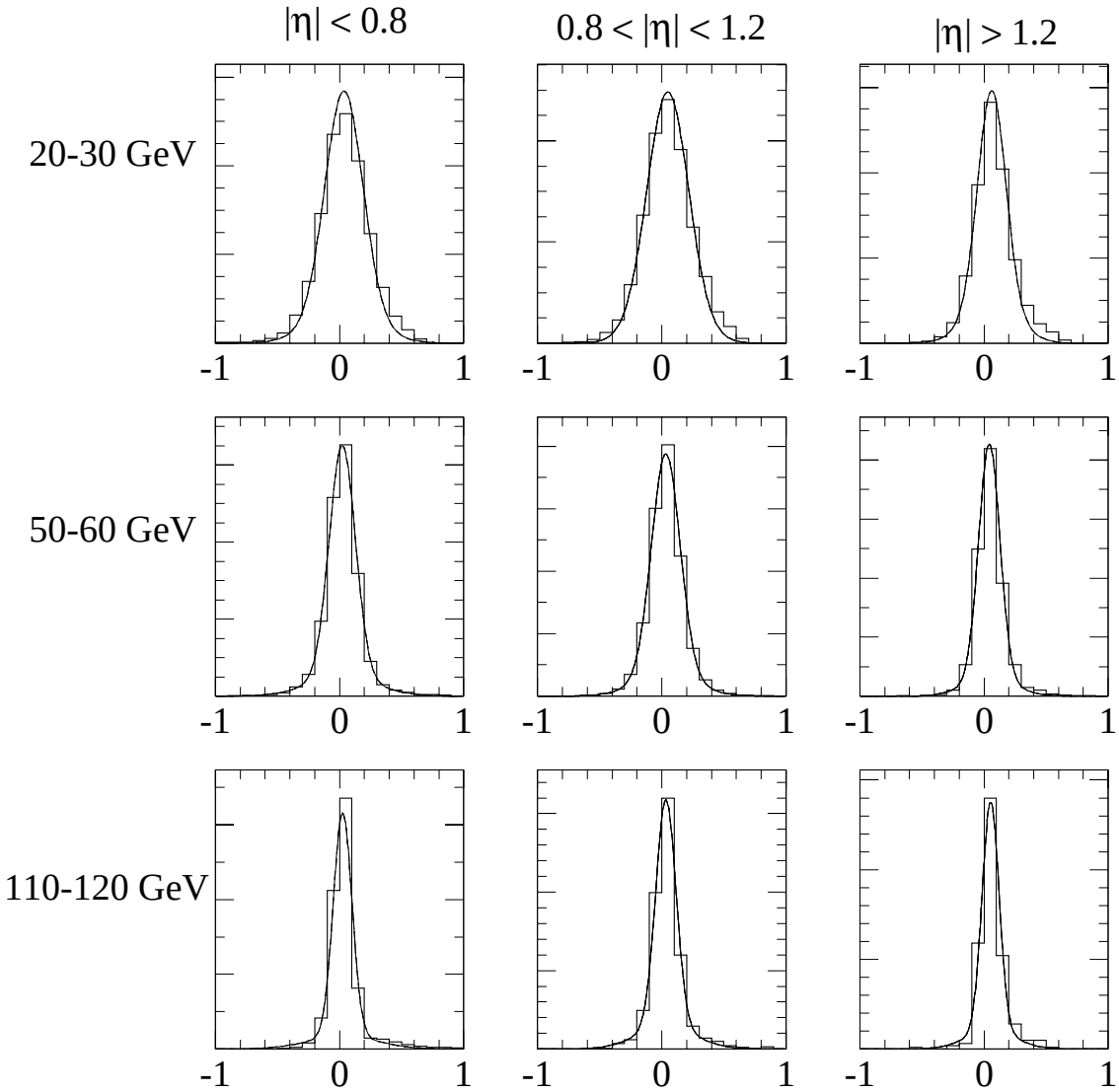

Figure 6.4: Fractional difference between reconstructed and generated jet transverse energies in HERWIG $t\bar{t}$ samples with top quark masses ranging from 110 to 190 GeV/c^2. Shown are the distributions for three E_T bins for jets in the central, intercryostat, and forward regions. Superimposed on each distribution is the double-Gaussian parameterization used in event smearing.

terms:

$$\omega_{ISR} = 0.08 \sum_{i=1}^{N_{ISR}} p_{Ti} \sin\theta_i \tag{6.10}$$

$$\omega_{FSR} = 0.1\,(m_{b1} + m_{b2}) \tag{6.11}$$

where m_{b1} and m_{b2} are the masses of the clusters of jets merged to form the b jets. The first term is designed to favor interpretations in which ISR is near the beam directions, and the second to favor the merging of jets which have low energy and/or are close to one another. The coefficients are chosen empirically based on Monte Carlo tests such that the mean reconstructed mass for two-jet and three-jet events is the same. This leads to a definition of the probability of the jth interpretation

$$P_j = \exp\left[\frac{-\left(\omega_{ISR} + \omega_{FSR}\right)}{2}\right] \tag{6.12}$$

It is found that the performance of the fitter is slightly improved by considering jets beyond the leading two, but that CPU time required increases rapidly with the number of jets one considers. As a compromise, all possible combinations of the three leading jets are considered in the fit, and any additional jets of lower E_T are ignored (there are a total of six possible assignments for three jets, as listed in Table 6.2). The weight for the jth jet interpretation is

$$\mathcal{W}(m_t, j|\{o_i\}) = \frac{P_j}{\sum_{i=1}^{N_{comb}} P_i} \sum_{\mathrm{y_1}\,\mathrm{y_2}} \sum_{\mathrm{Solutions}} \prod_{k=x,y} \exp\left[-\left(\frac{\left(\not{E}_k - p_k(\nu\nu)\right)^2}{2\sigma(\not{E}_k)^2}\right)\right] \tag{6.13}$$

The normalization term in the denominator ensures that each smear of an event contributes equally to the weight curve. Finally the weight curve for the event is obtained by summing over all jet interpretations:

$$\mathcal{W}(m_t|\{o_i\}) = \sum_{j=1}^{N_{comb}} \mathcal{W}(m_t, j|\{o_i\}) \tag{6.14}$$

b_1	b_2	ISR
j_1	j_2	j_3
j_1	j_3	j_2
j_2	j_3	j_1
$j_1 + j_2$	j_3	–
$j_2 + j_3$	j_1	–
$j_1 + j_3$	j_2	–

Table 6.2: Possible assignments of three observed jets (j_1, j_2, and j_3) as b jets or initial state radiation. There is an additional twofold ambiguity in pairing the leptons with the b jets.

For each assumed b jet, the energy is taken as that returned by CAFIX. The momentum is then adjusted such that the jet has the b quark mass (4.5 GeV/c^2). In the event that the jet is tagged by a muon, the most likely true energy of the muon and neutrino is determined based on the measured muon momentum, the Monte Carlo model of the muon and neutrino energies, and the resolution of the momentum measurement. This energy is then added back into the jet [79].

6.3 Monte Carlo Tests

The steps leading from Eq. 6.3 to Eq. 6.14 include both approximations and arbitrary choices. Therefore, while one has an *a priori* reason to expect that the probability defined in Eq. 6.3 will be sensitive to the top quark mass, tests need to be done to ensure that this sensitivity has not been compromised in the process of simplification. In addition, one needs to understand how $\mathcal{W}(m_t)$ behaves when one includes such effects as detector resolution and gluon radiation. In order to explore these issues, tests on both parton-level and reconstructed Monte Carlo samples have been performed.

6.3.1 Parton-level Tests

The first issue to be addressed is whether the approximate weight retains sensitivity to the top quark mass. To explore this, one begins by testing the method on the simplest possible case, which is parton-level Monte Carlo in which one uses the Monte Carlo information to select the correct jets and the correct lepton-jet pairing ("parton-level" means that one uses the simulation of the hard process only, and does not model gluon radiation or detector response). The event selection for these parton-level tests required only that there be two jets and two leptons with $p_T > 20$ GeV and $|\eta| < 2.5$, and was designed to restrict the sample to events broadly similar to those which enter the actual data analysis. Also, the samples used for these studies were generated by ISAJET.

One can then examine the average weight curve produced as a function of input top quark mass. This is done by normalizing the area of the weight curve for each event to unity, and then summing these normalized curves for a collection of Monte Carlo events (a sample of ten thousand events was used, about half of which passed the cuts). The results are shown in Fig. 6.5 for top quark masses of 130 and 190 GeV/c^2. On average the weight curve is sharply peaked near the input mass (there is a shift of about 1 GeV/c^2 low, reflecting the approximate nature of the weight). One also observes that the tails of the curve are asymmetric, with the high-end tail extending further than the low-end tail.

Figure 6.6 details the impact of detector resolution, jet combinatorics, and gluon radiation on the output of the fitter for 190 GeV/c^2 Monte Carlo. One sees that the distribution becomes significantly broader when one includes resolution effects and considers both lepton-jet pairings, but that the peak remains at the correct value. Inclusion of FSR lowers the peak and adds a significant low-mass tail, as expected. Turning off FSR and allowing ISR has the opposite effect. In total, detector resolution and gluon radiation tend to broaden the distribution of solution weights, and move the peak of the distribution away from the true top quark mass.

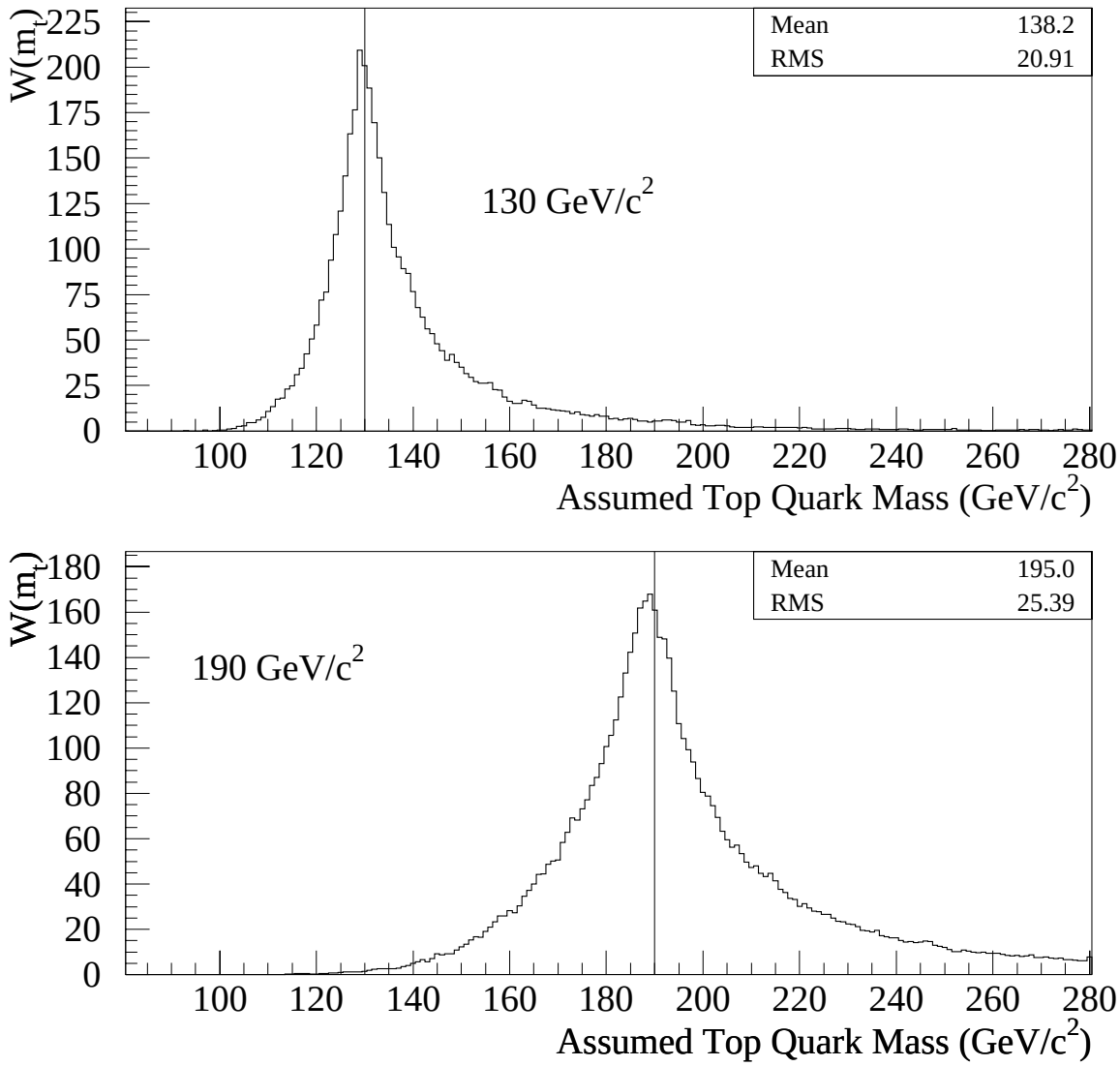

Figure 6.5: Average $\mathcal{W}(m_t|\{o_i\})$ for 130 and 190 GeV/c^2 unsmeared parton-level ISAJET $t\bar{t}$ events. The distributions were created by normalizing the area of the weight curves for each event to unity and then summing the weight curves for many events. One observes a sharp peak near the input top quark mass.

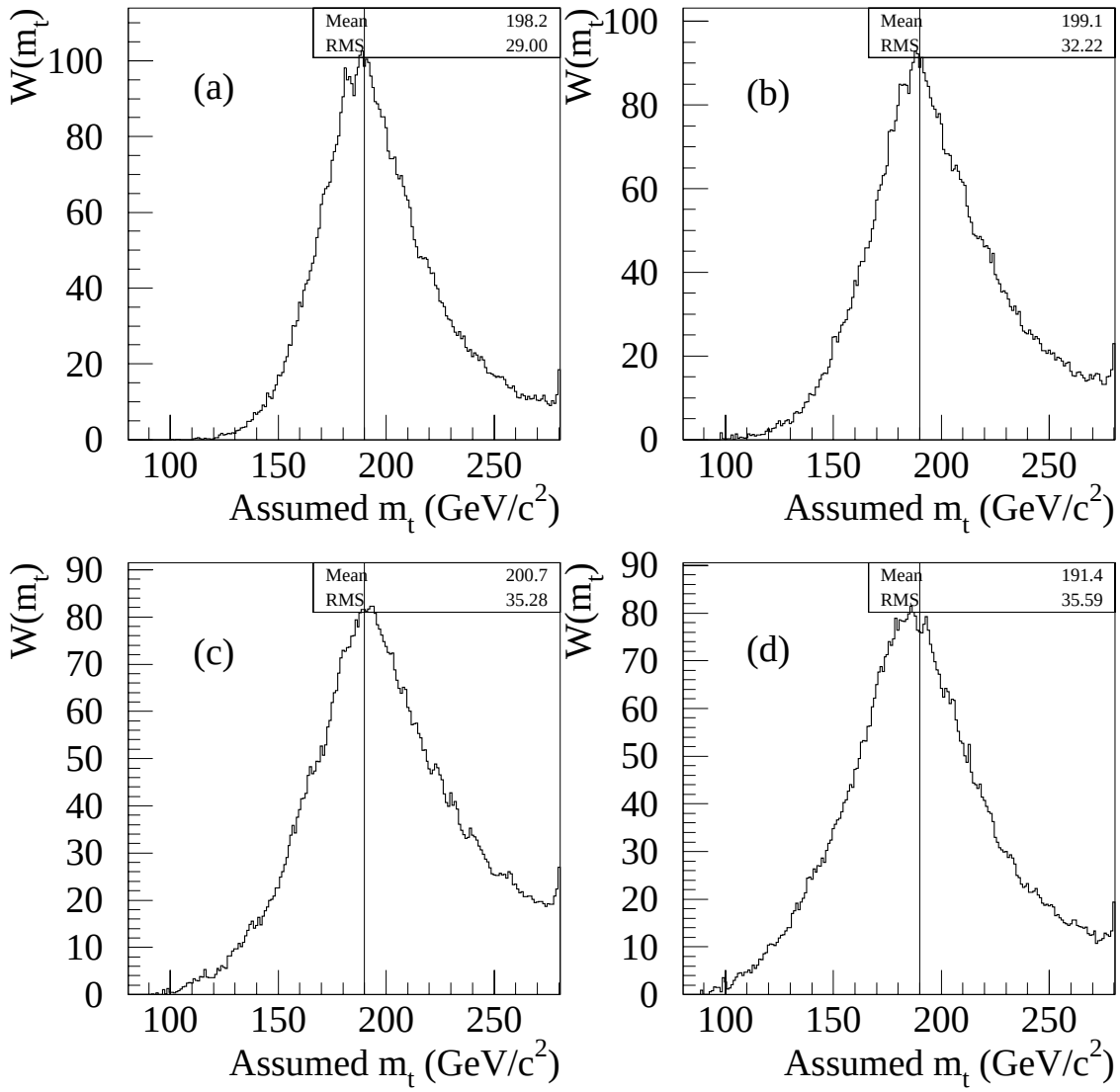

Figure 6.6: Average weight distribution for parton-level 190 GeV/c^2 Monte Carlo. Figure (a) shows the distribution when the events are smeared according to the expected detector resolution, Figure (b) introduces the two-fold ambiguity in lepton-jet pairings, Figure (c) adds initial-state gluon radiation, and Figure (d) suppresses initial-state but allows final-state gluon radiation.

6.3.2 Tests on GEANTed Monte Carlo

In order to assess the response of the fitting algorithm to the events from the DØ data sample that pass the kinematic selection described in Chapter 5, samples of HERWIG $t\bar{t}$ events are used. These samples are passed through GEANT and DØRECO, and are required to satisfy the same kinematic cuts as the candidate events (the electron identification relies solely on the $f_{iso} < 0.10$ cut, as the other variables are not modeled accurately in the Monte Carlo). For further details on the samples used, see Appendix A.

Figures 6.7 – 6.9 show the results for all three dilepton channels. Both the kinematic cuts and the additional complexity of the actual collider environment further degrade the resolution from that obtained in parton-level tests. In particular, note that for top quark masses less than about 140 GeV/c^2 the distributions are distorted significantly by the H_T^e cut (hence the distribution for 110 GeV/c^2 events looks similar to that for 140 GeV/c^2). In addition, the mean of the weight distribution moves upwards more slowly as a function of the input mass than does that for the best-case parton level tests.

Also, we see that the weight distributions become less sharp as the number of muons in the final state increases, reflecting the relatively poor measurement of their momenta. For this reason, and also due to the fact that the signal to background ratio is significantly higher for the $e\mu$ channel than for the ee or $\mu\mu$ channels, it is important to treat the three channels separately when measuring the top quark mass.

6.4 Lepton Plus Jets Events

6.4.1 Kinematic Fitting

The crucial difference between the lepton plus jets and dilepton final states is that one of the lepton-neutrino pairs in the latter is replaced by a jet-jet pair (from the hadronically decaying W), as shown in Fig. 6.10. One directly measures the

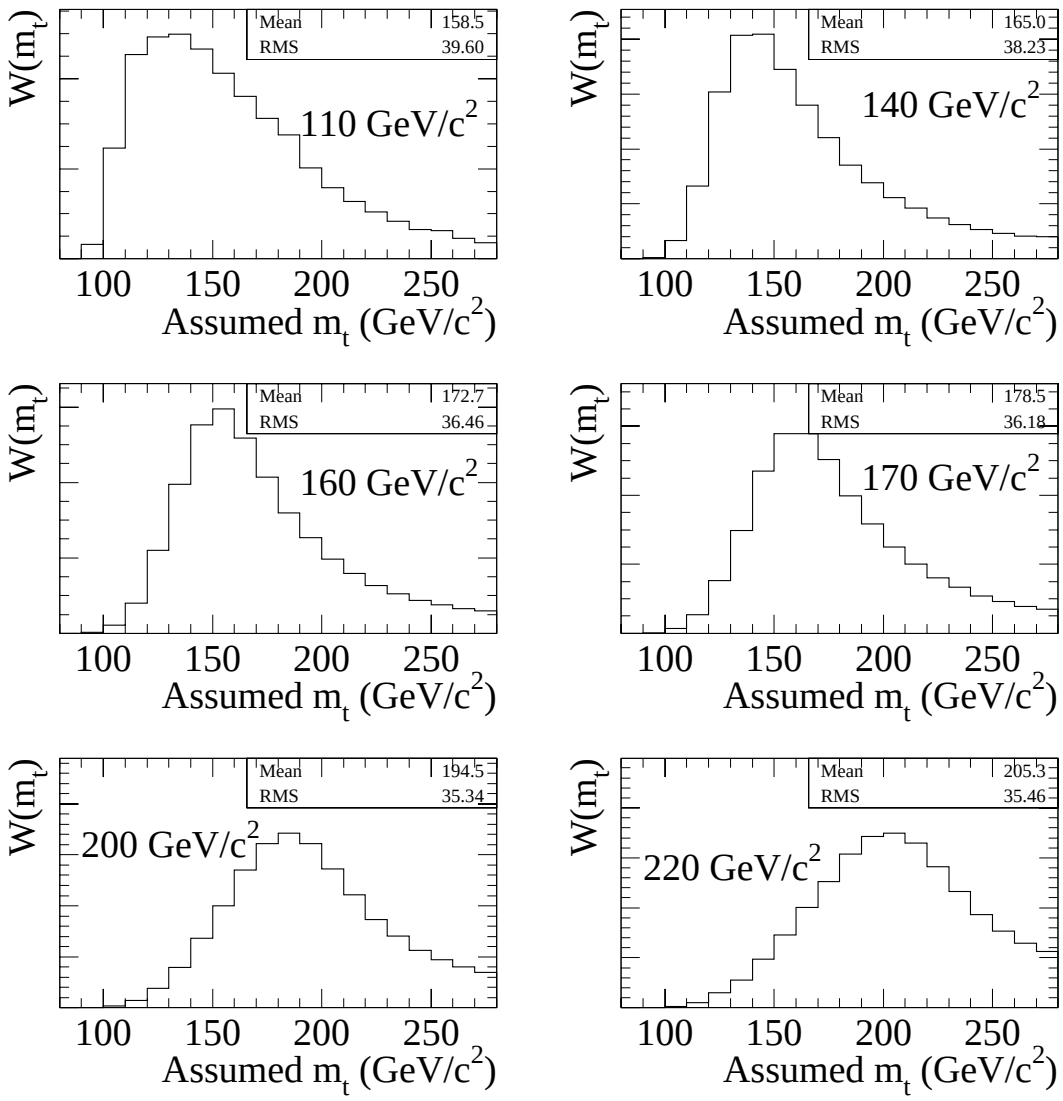

Figure 6.7: Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the ee channel.

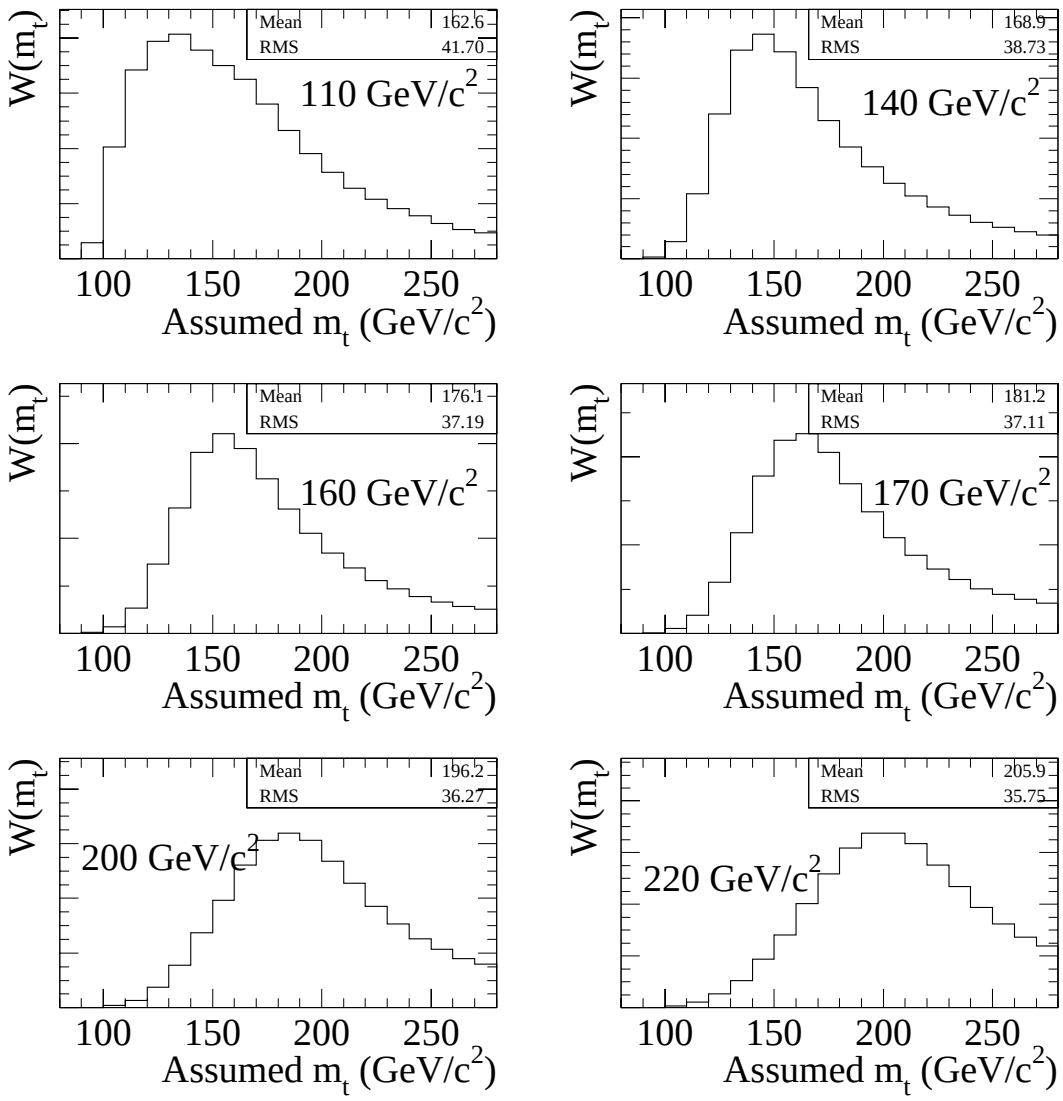

Figure 6.8: Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the $e\mu$ channel.

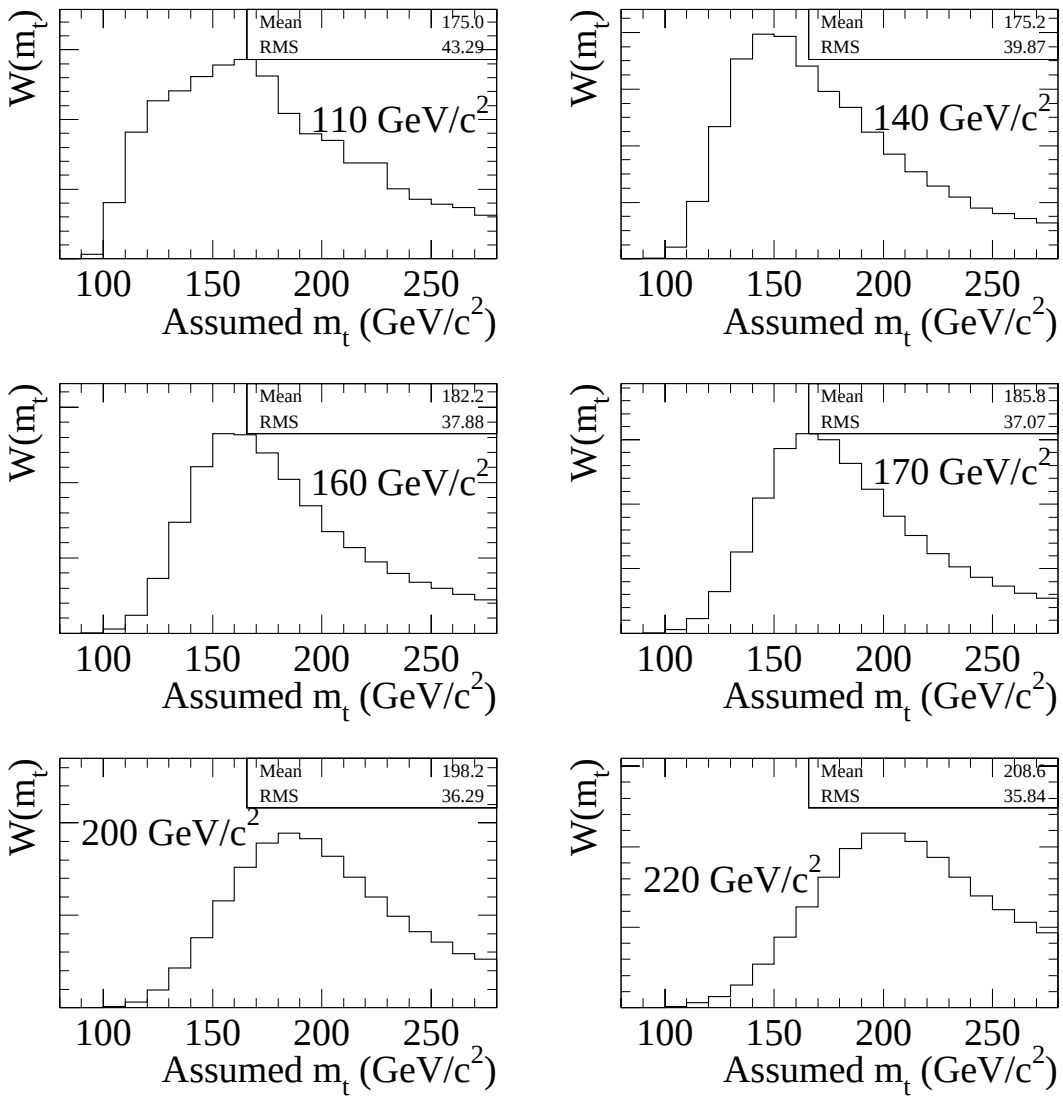

Figure 6.9: Results of applying the fitter to GEANTed HERWIG $t\bar{t}$ events in the $\mu\mu$ channel.

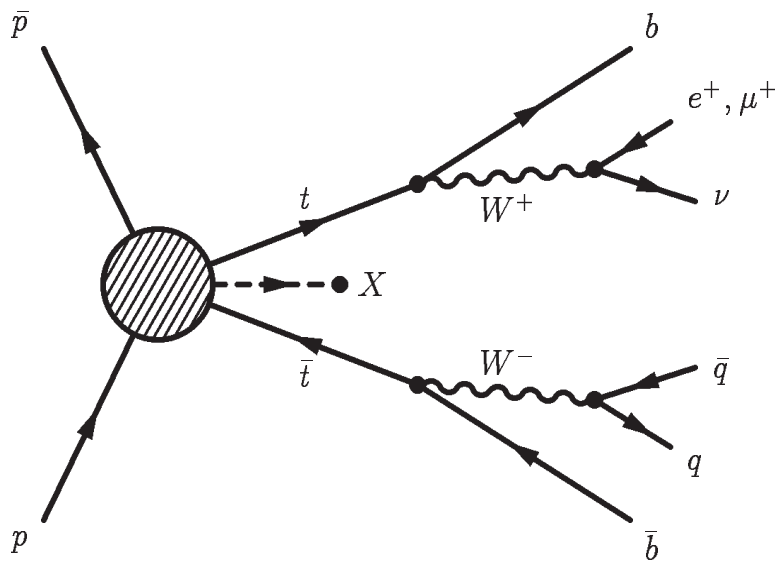

Figure 6.10: Schematic representation of $t\bar{t}$ production and decay in the lepton plus jets channels.

four-vectors of both jets in this pair.

Applying the same kinematic constraints as in the dilepton analysis thus allows the system to not only be reconstructed but also twice constrained. Given a vector $\mathbf{x}^m$ of the measured kinematic variables of the event, one can determine the vector $\mathbf{x}$ which is most consistent with the $t\bar{t}$ hypothesis and the resolutions on the quantities in $\mathbf{x}^m$ by minimizing the following χ^2, subject to the kinematic constraints:

$$\chi^2 = (\mathbf{x} - \mathbf{x}^m)^T \, \mathbf{G} \, (\mathbf{x} - \mathbf{x}^m) \tag{6.15}$$

where $\mathbf{G}$ is the inverse error matrix. In order to reduce the correlations between variables used in the fit, the vector $\mathbf{x}^m$ consists of (E_T, η, ϕ) for each observed object, and also the total energy recorded in each direction transverse to the beam that was not included in the reconstruction of a jet or electron (this carries the same information as the $\not{E}_T$, but has the advantage of being uncorrelated with the measured object energies). The method of undetermined Lagrange multipliers is used to enforce the constraints, and performing the minimization requires the solution of a system

of nonlinear equations. An iterative procedure is used in which the ideal solution is approached in incremental steps, each of which is small enough that the problem is approximately linear over the range considered. The procedure was designed and optimized explicitly for $t\bar{t}$ events, and is described fully in [80].

6.4.2 Jet Corrections to the Parton Level

As mentioned in Chapter 4, the CAFIX jet calibration package is designed to return the energies of all the final-state particles contained within the jet cone. In order to impose the constraint on the mass of the hadronically decaying W, it is necessary to go one step further, so that the energy of jet represents, on average, the energy of the parton that gave rise to that jet. The difference in the two definitions is due to final-state gluon emission which may fall outside of the jet cone.

The fact that there is no unambiguous way to relate the partons in an event to the final state jets means that this out-of-cone correction can be applied only in the context of some model for the final-state emission and hadronization of the partons. The model chosen is that provided by HERWIG.

To derive the appropriate correction, the energies of reconstructed jets in a sample of HERWIG $t\bar{t}$ events were compared to the energies of the hard process partons that most closely matched in $\eta - \phi$ space [79]. To ensure that the correction factors derived were not tuned only for a particular value of the top quark mass, the sample consisted of events generated with top quark mass ranging from 160 – 210 GeV/c^2.

Since the physical size of the jet cone decreases as η increases, one must perform this comparison separately for jets in different η regions. As shown in Fig. 6.11 the relation between parton and jet energies is linear, and the slope and offset of the best-fit line defines the correction factor:

$$E_{\mathrm{corr}} = (E_{\mathrm{jet}} - \mathrm{Offset})/\mathrm{Slope} \tag{6.16}$$

where E_{jet} is the reconstructed jet energy after the CAFIX corrections.

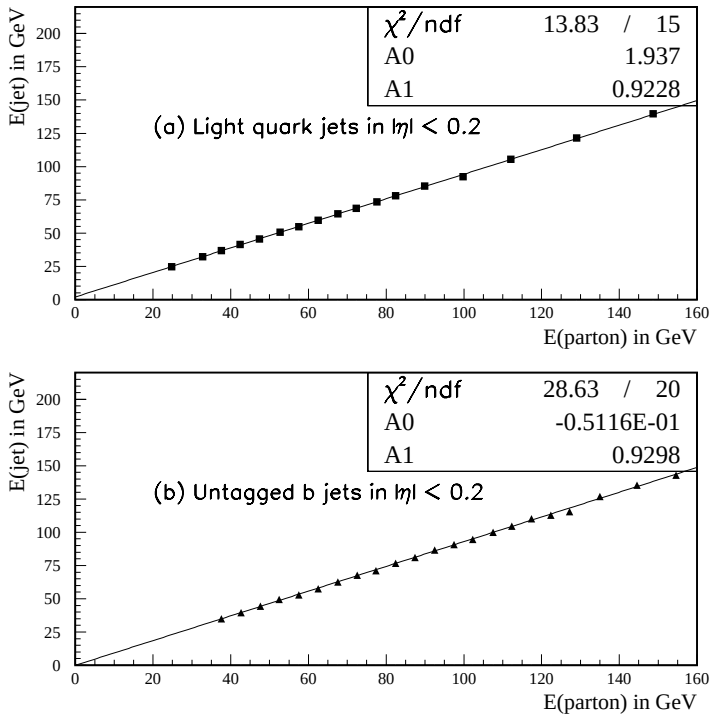

Figure 6.11: Reconstructed vs. parton energies for central jets.

In addition to the η region, the type of quark that fragmented to produce the jet must be considered in deriving the correction, as b jets may undergo semileptonic decay which produces a neutrino in the jet. Most of the semileptonic decays are undetectable (the exception being some of the decays that produce muons), and so the b jets must be corrected on average for the effects of the unmeasured neutrino, leading to a larger correction than for light quark jets[1].

In the event that a soft muon is detected within the jet cone, the most likely true muon momentum is determined (using the same algorithm as in the dilepton case). This is taken as the leptonic energy of the jet. After removing the energy contributed by ionization as the muon traversed the calorimeter, the remaining hadronic jet energy is corrected with the light quark correction factors. The leptonic energy is then added to the hadronic energy to give the reconstructed energy for the jet.

The values of the correction factors for each η region and jet type are given in Table 6.3. To demonstrate that the out-of-cone corrections are performing as designed, one can plot the invariant mass of the two jets arising from hadronic W decay and the three jets arising from top decay (using Monte Carlo information to properly identify these jets) both before and after the out-of-cone corrections. As shown in Fig. 6.12, the corrections move the mass distribution closer to the pole mass in each case.

6.4.3 Effects of Gluon Radiation

As in the dilepton case, gluon radiation and jet combinatorics complicate the ideal situation represented in Fig. 6.10. The consequences of these effects, and how they are handled, are detailed below.

While the kinematic fit assumes that the event has four jets, and that these correspond to the two b jets and two jets from W decay, this is not the case for the majority of $t\bar{t}$ events. In some cases, one or more of the desired jets is missed, and a jet arising

[1]Of course, in the absence of a soft muon tag, there is no way to tell if a particular jet is from a b or a light quark. The mass fitting procedure considers all possible identities of a jet, and assigns the correction factor consistently, as described below.

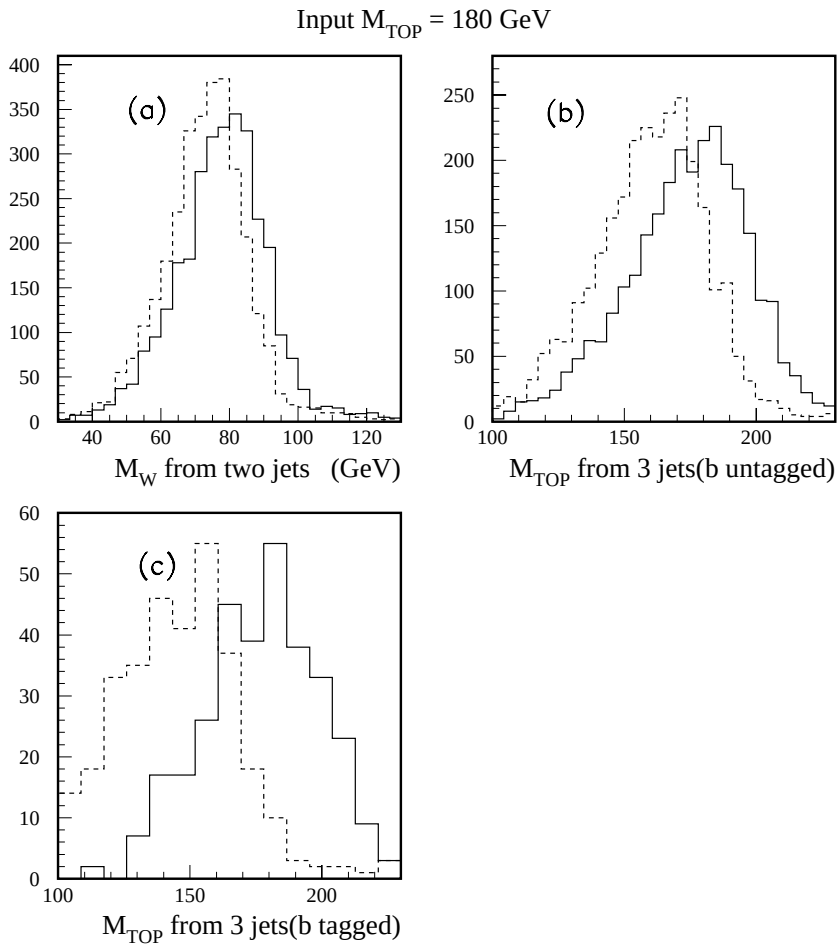

Figure 6.12: Comparison of jet invariant masses before (dashed) and after (solid) out-of-cone radiation corrections for: (a) two jets from hadronic W decay, (b) three jets from top decay (with the b untagged) and (c) with the b muon-tagged for 180 GeV/c^2 top events. In each case the distribution moves closer to the pole value following the corrections [79].

η region	Light Quark Jets		Untagged b Jets	
	Offset	Slope	Offset	Slope
$\|\eta\| < 0.2$	0.32	0.93	-0.67	0.91
$0.2 < \|\eta\| < 0.6$	0.64	0.93	-1.34	0.91
$0.6 < \|\eta\| < 0.9$	1.86	0.88	1.6×10^{-3}	0.87
$0.9 < \|\eta\| < 1.3$	1.70	0.93	-0.55	0.90
$\|\eta\| > 1.3$	4.50	0.88	2.46	0.86

Table 6.3: Out-of-cone jet correction factors for light quarks and untagged b quarks for each η region.

Number of jets	Number of permutations
4	12
5	140
6	1020
7	5992

Table 6.4: Number of possible assignments of jets to parent partons in reconstructing a $t\bar{t}$ event in the lepton plus jets channel, as a function of jet multiplicity [80].

from gluon radiation is found in its place. In other cases, all four desired jets are reconstructed along with additional gluon jets. In the latter case, more information can in principle be obtained by considering all possible origins of all the jets in the event (W, b, initial-state radiation, and final-state radiation). However, the number of possible combinations to consider grows rapidly as a function of the jet multiplicity (see Table 6.4), and the probability of picking the correct combination out of all of these becomes vanishingly small. So in practice, there is very little return for the additional complexity introduced by considering additional combinations, and the fit proceeds by considering only the four jets with highest E_T.

Even for the cases where the leading four E_T jets do correspond to the desired partons, there is still ambiguity due to the fact that one does not know the origin of each jet. Therefore the kinematic fit procedure must be repeated for all possible

assignments of jets to parent partons. There are twelve such possible assignments (this is reduced to 6 if one of the jets has a soft muon tag, in which case only combinations which assign this jet as a b are considered). As each assignment of jets to partons is considered, the out-of-cone energy correction consistent with the presumed identity of each jet is applied.

There is an additional ambiguity in that the χ^2 may have multiple local minima, and in such cases the minimum to which the fit converges depends on the initial values of $\mathbf{x}$. For the seventeen measured quantities, the starting value is provided by the measured value. The longitudinal component of the neutrino can also be assigned by requiring that the assumed t and $\bar{t}$ masses be equal at the start of the fit. However, this requires the solution of a quadratic equation, introducing a two-fold ambiguity in the starting value. Therefore, both possible starting points are attempted, and the one yielding the absolute minimum χ^2 is retained. If there is no solution with $\chi^2 < 10$, the event is discarded from the mass fit. Those events remaining after this cut form the *base sample* for lepton plus jets mass analysis.

The presence of all of these ambiguities means that one has no reason to expect that the result of the kinematic fit will represent the actual mass of the top quark. In fact, one finds that in only $\approx 40\%$ of HERWIG $t\bar{t}$ events do the leading four jets correspond to the nominal four partons, and only $\approx 17\%$ of the time does the correct jet combination provide the lowest χ^2 solution [81].

The effect on the reconstructed top quark mass is illustrated in Fig. 6.13, which shows how the distribution of fitted mass changes as the above effects are added to the simulation for HERWIG events generated with top quark mass 170 GeV/c^2. The plot at upper left shows the near delta-function at the correct mass obtained in the ideal case (perfect resolution and no gluon radiation), that at upper right shows the considerable broadening when gluon radiation is introduced, and that at lower left the additional smearing due to detector resolution.

Therefore, the fitted mass is to be understood as a kinematic variable for the event, and the relationship between this variable and the actual top quark mass must be

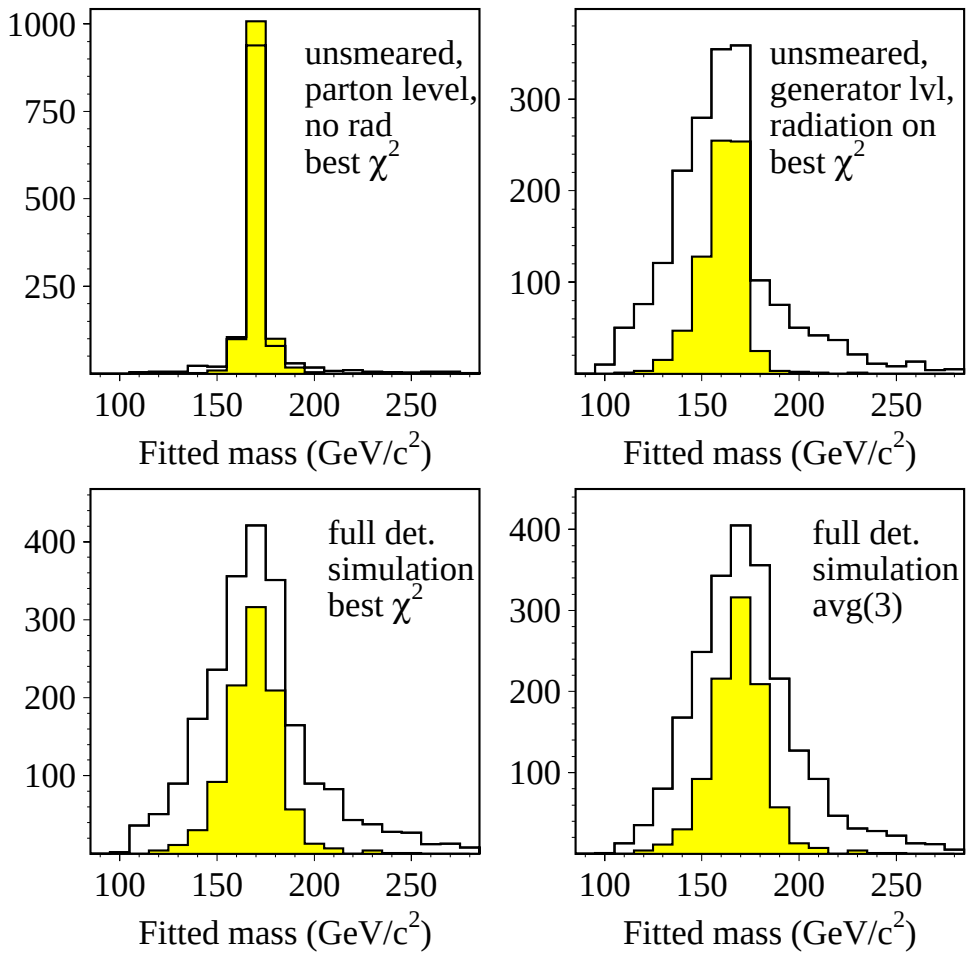

Figure 6.13: Degradation in the resolution of the fitted mass in lepton plus jets events due to gluon radiation and detector resolution. The shaded histograms give the distribution of fitted masses when the correct jet assignment is input to the fit. The bottom right figure shows the result of taking a χ^2-weighted average of the three lowest χ^2 solutions of each event.

established by performing the procedure on $t\bar{t}$ Monte Carlo events at a variety of top quark masses. Figure 6.14 shows the output of the kinematic fitting procedure on each of the four lepton plus jets channels for 170 GeV/c^2 HERWIG $t\bar{t}$ events. Events in this plot were required to satisfy the selection designed for mass analysis as described in Chapter 5. As the distributions for all four samples are very similar, the standard procedure is to combine them in the ratio expected from the $t\bar{t}$ branching fractions and DØ acceptances. The results for a range of top quark masses are displayed in Fig. 6.15. One sees that the peak of the distribution moves as a function of the true top quark mass, and also that the shape of the distribution can be altered by the cuts placed on the data sample (particularly at the lower top quark masses represented).

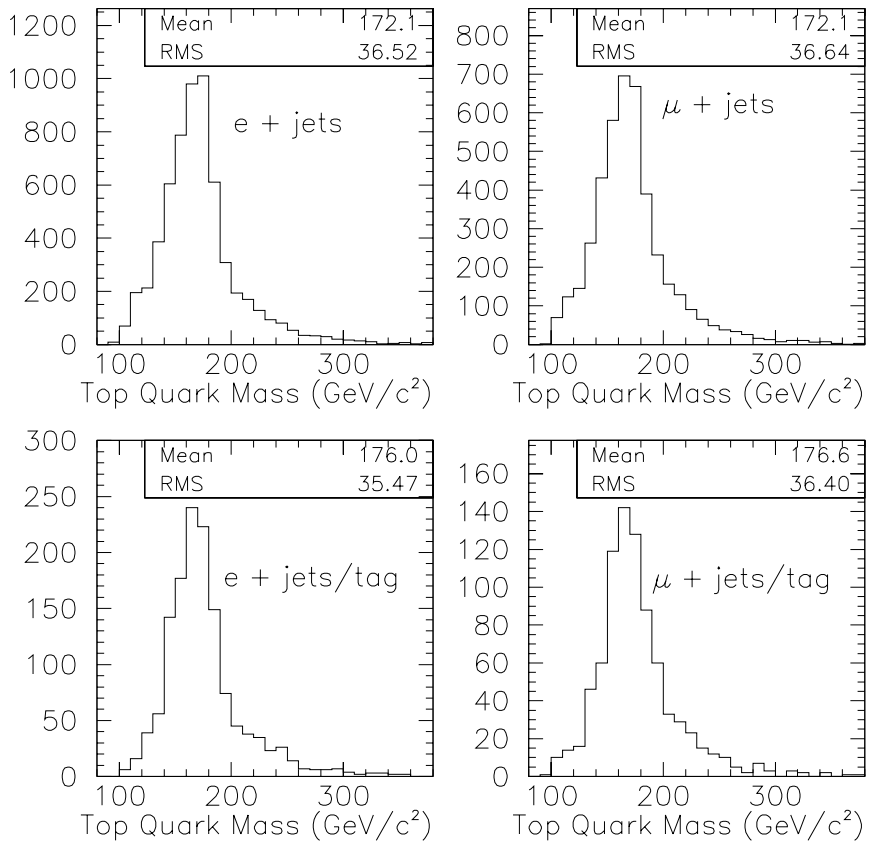

Figure 6.14: Result of applying the kinematic fit procedure on a sample of 170 GeV/c^2 HERWIG $t\bar{t}$ events in each of the lepton plus jets channels.

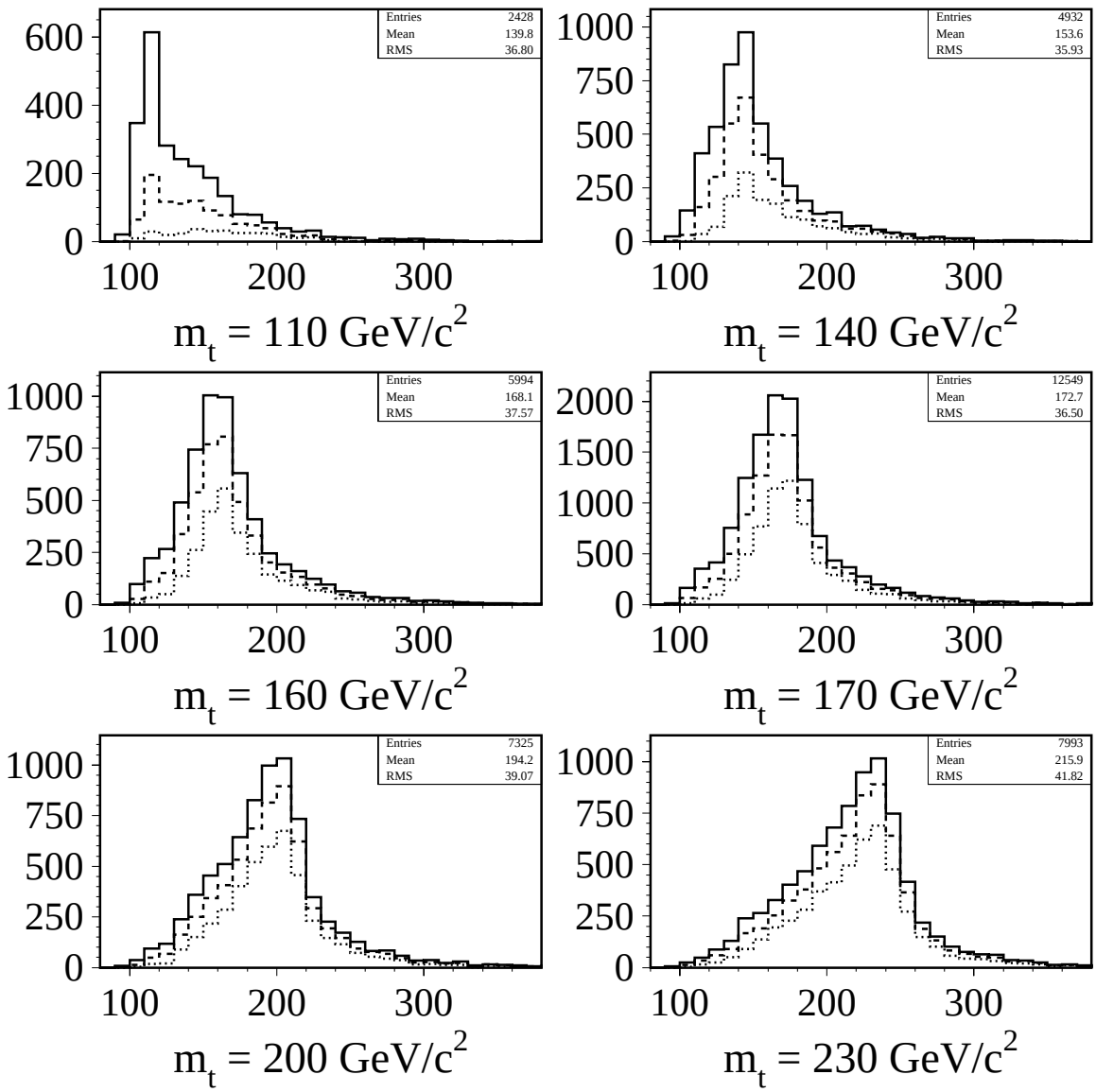

Figure 6.15: Result of applying the kinematic fit procedure to samples of HERWIG $t\bar{t}$ events for a variety of top masses. The top quark mass favored by the lowest χ^2 solution for each event is plotted. The solid curves are for events in the mass analysis base sample, the dashed are those which pass the further selection described in Chapter 7, and the dotted are those which pass the cuts used in the cross section measurement[76].

Chapter 7

Maximum Likelihood Fit

Once the mass reconstruction has been performed for each of the events in the $t\bar{t}$ candidate sample, the remaining step in the analysis is to estimate the top quark mass based on the observed distribution of reconstructed masses. A *maximum likelihood* fit is employed to achieve this goal. In general, if one wishes to estimate a set of parameters $\{\theta\}$ given a set of observations $\{v\}$ from N data events, one needs to know the probability density $f(\{v\}|\{\theta\})$. Then the maximum likelihood estimate of the parameters, $\{\hat{\theta}\}$, is given by maximizing the joint likelihood

$$L\left(\{v\}\,|\{\theta\}\right) = \prod_{i}^{N} f\left(\{v\}_i|\{\theta\}\right) \tag{7.1}$$

with respect to $\{\theta\}$. As the log of L will be maximal for the same values of $\{\theta\}$ as is L itself, it is often convenient to use the quantity

$$\ln L\left(\{v\}\,|\{\theta\}\right) = \sum_{i}^{N} \ln f\left(\{v\}_i|\{\theta\}\right) \tag{7.2}$$

In addition to providing the estimates $\{\hat{\theta}\}$, the relative value of $L\left(\{v\}\,|\{\theta\}\right)$ to the maximal $L\left(\{v\}\,|\{\hat{\theta}\}\right)$ also gives a measure of the uncertainty in the estimated $\hat{\theta}$. In fact, as the number of events tends to infinity, one can show that the set of all $\{\theta\}$ such that $\ln L\left(\{v\}\,|\{\theta\}\right)$ differs by less than half a unit from $\ln L\left(\{v\}\,|\{\hat{\theta}\}\right)$ defines

a 68% confidence region for $\{\theta\}$. However, given the low statistics available to the top quark mass measurement, it cannot be assumed that this relation holds. The issue of the proper assignment of statistical uncertainty will be revisited in detail in Chapter 8. For more information on the properties of maximum likelihood fits, see for example [82].

For the top quark mass measurement, the set $\{\theta\}$ which one would like to estimate consists of the number of signal (n_s) and background (n_b) events in the data sample, as well as m_t. The set $\{v\}$ includes information from the mass reconstruction of each event in the sample, the number of events in the sample (N), and optionally the expected number of background events ($\overline{n_b}$) as determined by the cross section analysis[1]. As discussed below, the probability density as a function of these quantities is a relatively complex object, which renders analytic maximization of L impractical. Therefore, the MINUIT [83] package is used to perform a numerical minimization of $-\ln L$.

Also, while it is natural to treat n_s and n_b as continuously varying parameters, the fact that Monte Carlo events exist only for discrete values of the top quark mass complicates a similar treatment of m_t. There are two possible procedures:

1. One may parameterize the shape of the probability density as a function of m_t. This allows evaluation of L for any m_t in the range over which Monte Carlo is generated, and therefore simultaneous minimization of $-\ln L$ with respect to n_s, n_b, and m_t.

2. One may minimize $-\ln L$ with respect to n_s and n_b for each of the top quark masses where Monte Carlo exists. In order to estimate the m_t which minimizes $-\ln L$ in this case, one must then choose a parameterization of the dependence of the resultant $-\ln L$ on m_t.

[1]One also could include the expected number of signal events in $\{v\}$, but this would require assuming a $t\bar{t}$ production cross section. Since one of the objectives of the mass measurement is to test the perturbative QCD prediction of the cross section, it is desirable to avoid making such an assumption.

While both methods require a parameterization (the exact nature of which is not uniquely defined), the fact that this occurs in the final step for the latter procedure allows one to more readily judge whether the chosen parameterization is a reasonable one. Therefore the second procedure is adopted for determining m_t, and a low-order polynomial fit to the m_t points near the minimum $-\ln L$ is used as the parameterization. The exact nature of the fit is discussed in Chapter 8.

7.1 Dilepton Events

The result of reconstructing the top quark mass for a dilepton event is the distribution $W(m_t)$, which is evaluated for 50 values of the top quark mass. In principle, one could define $\{v\}$ to include all of these measurements. However, dealing with such a large number of observations for each event would be prohibitively difficult. In order to create a more tractable set $\{v\}$, one notes that the intrinsic resolution of the dilepton mass reconstruction is much broader than the 4 GeV/c^2 interval between assumed top quark masses. Therefore one can reduce the dimensionality of the problem by coalescing adjacent weight values into broader bins without significant loss of resolution.

Of course, one must empirically determine a reasonable compromise between the desire to simplify the problem and the need to retain as much information as possible. The scale of the largest bin width one should consider is given by the RMS of the typical $W(m_t)$ distribution. As shown in Figs. 6.7 – 6.9 this RMS typically lies between 35 and 40 GeV/c^2. Therefore, the weights are coalesced into five 40 GeV/c^2 bins, as shown in Fig. 7.1. To further simplify, the total weight for each event is normalized to unity before this coalescing is performed, so that it is the fraction of the total weight which falls in each bin that is recorded. Since one now has the constraint that the coalesced weights sum to unity for each event, the values of the first four entries determine that of the fifth, this reduces the dimensionality of the problem to four. In summary, for each event we describe the result of the mass reconstruction by

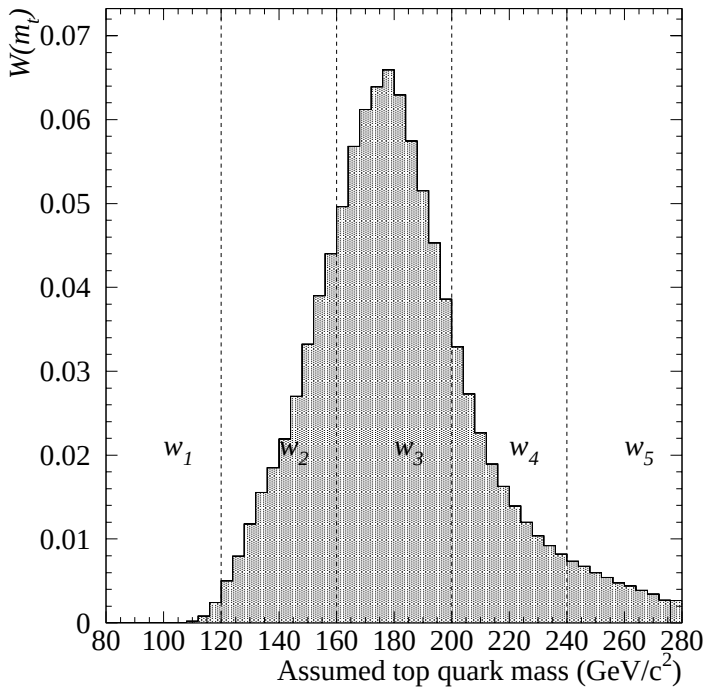

Figure 7.1: An illustration of the procedure for coalescing the $W(m_t)$ distribution into five 40-GeV/c^2 bins. Shown is the $W(m_t)$ distribution for a typical Monte Carlo $t\bar{t}$ event, normalized to unity.

the four-dimensional vector $\vec{w}$, the entries of which are the fractional weights in the bins 80 - 120 GeV/c^2, 120 - 160 GeV/c^2, 160 - 200 GeV/c^2, and 200 - 240 GeV/c^2.

We now wish to evaluate the likelihood $L\left(\vec{w}_1, \ldots, \vec{w}_N, \overline{n_b}, N | m_t, n_s, n_b\right)$. The dependence on $\overline{n_b}$ may be represented by the Gaussian constraint

$$g(n_b, \overline{n_b}, \sigma_b) \equiv \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left[\left(n_b - \overline{n_b}\right)^2 / 2\sigma_b^2\right]$$

where σ_b is the uncertainty in the background estimate. One also introduces a Poisson constraint that $n_s + n_b$ be consistent with the total size N of the data sample:

$$p(n_s + n_b, N) \equiv \frac{\left(n_s + n_b\right)^N e^{-(n_s + n_b)}}{N!}$$

Now the likelihood factorizes to

$$L\left(\vec{w}, \overline{n_b}, N | m_t, n_s, n_b\right) = g(n_b, \overline{n_b}, \sigma_b) p(n_s + n_b, N) L'(m_t, n_s, n_b | \vec{w}_1, \ldots, \vec{w}_N) \quad (7.3)$$

Evaluation of L' depends on knowledge of the signal and background probability densities $f_s(\vec{w}|m_t)$ and $f_b(\vec{w})$. The simplest manner of estimating such densities is to histogram the data. However, in a multidimensional problem such as this one, the number of bins becomes large even if one bins coarsely in each dimension. If the Monte Carlo samples are spread across too many bins, one will be commonly faced with the problem of comparing null bins in the data sample to null bins in the Monte Carlo model. For this reason, an unbinned estimation of f_s and f_b is more appropriate to the problem at hand.

7.1.1 Probability Density Estimation

The problem then is to estimate the continuous functions f_s and f_b from the discrete sample of Monte Carlo points available. The probability density estimation (PDE) technique used in this analysis is outlined below. One may find more details on the technique and its implementation at DØ in [84, 85].

The idea is to place a kernel function K at the location of each point in the Monte Carlo sample, and to estimate the probability density as

$$f(\vec{x}) = \frac{1}{N_{MC} h^d} \sum_{i=1}^{N_{MC}} K\left(\frac{\vec{x} - \vec{x}_i}{h}\right) \quad (7.4)$$

where d is the number of dimensions of $\vec{x}$, N_{MC} is the number of Monte Carlo points available, and h is the *smoothing parameter*. The latter is a free parameter and its optimization is discussed in Chapter 8.

Any function which is maximal at zero and asymptotes to zero as the absolute value of its argument becomes large would be an acceptable kernel. A Gaussian is a convenient choice. For the kernel estimation to provide a faithful model of the probability density, however, it is desirable that the kernel function have the same general

shape as the distribution being modeled. That is, one would like the covariance matrix of K to be the same as that of the input variables.

A Gaussian kernel of course has a diagonal covariance matrix, the elements of which are the variances in each dimension. Therefore a linear transformation is applied to the data:

$$\vec{w}' = \mathbf{A}\vec{w}$$

The matrix $\mathbf{A}$ is chosen so that the covariance matrix of the transformed variables from the background Monte Carlo is the identity matrix while that of the signal Monte Carlo is a diagonal matrix $\mathbf{D}$. Therefore the kernel function has the form

$$K(\vec{x}, \mathbf{M}) = \frac{1}{(2\pi)^{d/2}\sqrt{m_1 \ldots m_d}} \prod_{i=1}^{d} \exp\left[\frac{-(x_i^2)}{2m_i}\right]$$

where the matrix $\mathbf{M}$ is either the identity matrix (for background) or $\mathbf{D}$ (for signal), and the m_i are its diagonal elements.

One minor extension is needed to properly model the background with this method. As described in Chapter 5, the backgrounds in the dilepton channel arise from a variety of sources, and in general the amount of Monte Carlo events available for each source does not match its expected contribution. Hence the Monte Carlo events for each source are assigned a weight factor b such that for the jth background source

$$\frac{b_j N_j^{MC}}{\sum_{i=1}^{N_{\text{source}}} b_i N_i^{MC}} = \frac{\overline{n_{b,j}}}{\overline{n_b}}$$

where N_j^{MC} events are available to model the source and $\overline{n_{b,j}}$ is the number of events it is expected to contribute.

Then the estimate for the probability density for an event weight vector $\vec{w}$ is given by:

$$f_s(\vec{w}|m_t) = \frac{1}{N_{MC}(m_t)} \sum_{i=1}^{N_{MC}(m_t)} K\left(\frac{\vec{w} - \vec{w}_i}{h}, \mathbf{D}\right) \tag{7.5}$$

for signal[2], and

$$f_b(\vec{w}) = \frac{1}{\sum_{j=1}^{N_{\text{source}}} b_j N_j^{MC}} \sum_{j=1}^{N_{\text{source}}} b_j \sum_{i=1}^{N_{jMC}} K\left(\frac{\vec{w} - \vec{w_{ij}}}{h}, \mathbf{I}\right) \tag{7.6}$$

for background.

7.1.2 Form of the Likelihood

Now that a procedure is established for estimating f_s and f_b, one may write the form of the likelihood that is to be maximized:

$$L\left(\vec{w}_1, \ldots, \vec{w}_N, \overline{n_b}, N | m_t, n_s, n_b\right) = g(n_b, \overline{n_b}, \sigma_b) p(n_s + n_b, N) \prod_i^N \frac{n_s f_s(\vec{w}_i | m_t) + n_b f_b(\vec{w}_i)}{n_s + n_b} \tag{7.7}$$

As mentioned above, the three dilepton channels differ significantly in their expected signal-to-background ratios and somewhat in the properties of their reconstructed masses. Therefore each channel is treated separately in the likelihood fit, and the final likelihood is the product of the likelihoods from each channel.

7.2 Lepton Plus Jets Events

In the lepton plus jets channel, there are two quantities which are crucial in extracting information about the top quark mass from any particular event. These are the fitted mass for that event and the probability that the event is an example of $t\bar{t}$ production. Therefore it is natural to bin the data in two dimensions, with one axis being the fitted mass and the other some measure of the $t\bar{t}$ probability.

It remains to estimate the probability densities to be input into the likelihood fit. The fact that the problem is two-dimensional means that this estimation is simpler than in the dilepton case. When one considers that little resolution will

[2] While the evaluation of f_s for a particular top quark mass depends only on events generated at that mass, the **A** matrix used to transform the variables is computed only once, using the distribution of Monte Carlo events generated at all top quark masses

be lost by binning events coarsely along the top probability axis (one is interested mostly in the pseudo-binary information of whether an event is likely or unlikely to be top), it becomes clear that sufficient Monte Carlo statistics exist to populate a two-dimensional histogram to provide an estimate of the probability density.

While the fitted mass axis is unambiguously defined using the procedure described in Chapter 6, there are several choices available in how one bins events in the probability axis. One procedure adopted combines information from the top likelihood discriminant $\mathcal{D}$ defined in Chapter 5 with some physical intuition as to the nature of $t\bar{t}$ events, to divide the sample into those events likely and unlikely to be top. The former category contains all events with $\mathcal{D} > 0.43$ and $H_T - E_T(\text{jet } 1) > 90$ GeV, as well as every event with a muon-tagged jet, and is referred to as the "low-bias", or LB, sample, since the distribution of reconstructed masses is quite similar between events that pass and fail this selection. The distribution of signal and background events on this two-dimensional plane is shown in Figs. 7.2 and 7.3. One can observe that even for relatively low top quark masses the distributions are quite different for signal and background events, which would not be the case were one to bin the events one-dimensionally in the reconstructed mass.

Alternatively, one may use the output of the neural network to define the top likelihood. In this case, the events are grouped into ten bins of increasing neural network output, with the bin boundaries chosen such that each bin contains a roughly equal population of events in a Monte Carlo mixture with the expected signal to background ratio. The resulting two-dimensional distributions are shown in Figs. 7.4 and 7.5. Mass analyses based on both the top likelihood discriminant and neural network will be presented in Chapter 8.

Given this clear separation in shape between signal and background events, it is not necessary to impose any constraint on the amount of background in the data sample. Thus, the likelihood fit can not only determine the top quark mass, but also provide a check on the measurement of the cross section and the understanding of the background.

One then can define the likelihood contributed by a bin to be the Poisson probability that the number of events predicted in a bin could fluctuate to the number of data points observed in that bin (N_j):

$$L\left(N_j|m_t, n_s, n_b\right) = p\left(p_s a_j^s + p_b a_b^j, N_j\right) \tag{7.8}$$

where M is the number of bins, a_j^s and a_b^j are the probability densities for the signal and background models in the jth bin and p_s and p_b are normalization factors related to n_s and n_b by $p_i = n_i / \left(M + \sum_j A_j^s\right)$. The situation is complicated by the fact that the a_j themselves are not exactly known, but only approximated by the Monte Carlo distribution in which a number of events A_j fall into the jth bin. Hence one needs also to include in the likelihood the probability for the a_j to fluctuate to A_j:

$$L\left(N_j|m_t, n_s, n_b\right) = p\left(N_j, p_s a_j^s + p_b a_b^j\right) p\left(A_j^s, a_j^s\right) p\left(A_j^b, a_j^b\right) \tag{7.9}$$

One may then integrate over the unknown a_j's to eliminate them from the likelihood, and thus obtain [86]:

$$L\left(N_j|m_t, n_s, n_b\right) = \sum_{k=0}^{N_j} \frac{p_s^k}{\left(1+p_s\right)^{A_j^s+k+1}} \frac{p_b^{N_j-k}}{\left(1+p_b\right)^{A_j^b+N_j-k+1}} \begin{pmatrix} A_j^s + k \\ k \end{pmatrix} \begin{pmatrix} A_j^b + N_j - k \\ N_j - k \end{pmatrix} \tag{7.10}$$

The likelihood for a given model is then taken as the product of the likelihoods for each of the M bins:

$$L\left(D|m_t, n_s, n_b\right) = \prod_{j=1}^{M} L\left(N_j|m_t, n_s, n_b\right) \tag{7.11}$$

where D represents the distribution of the data across the two-dimensional histogram.

As a final step, the likelihood is divided by the quantity $L(D|D) \equiv \prod_{j=1}^{M} p(N_j, N_j)$. The modified likelihood then has the property that, in the limit of high statistics, $-2\ln L$ is distributed as a χ^2. This allows one to use the value of the minimal $-2\ln L$ to judge the quality of the fit of the signal and background models to the data, and therefore as a further check that the Monte Carlo models are reasonable.

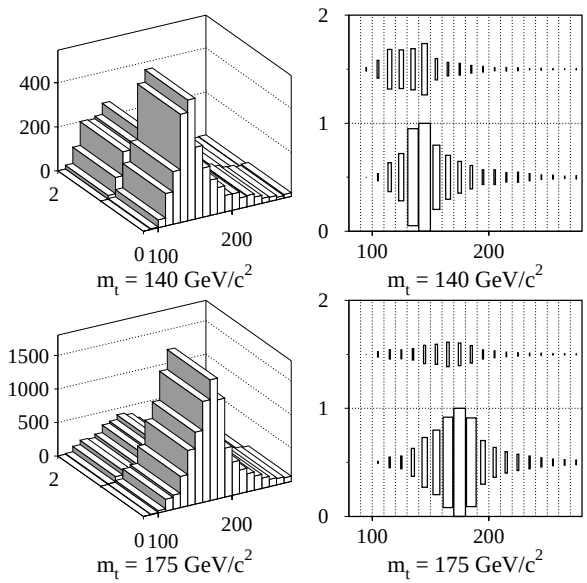

Figure 7.2: Distribution of reconstructed masses for HERWIG 140 and 175 GeV/c^2 $t\bar{t}$ events, for those that pass (bin one) and fail (bin two) the LB cut described in the text [76].

7.3 Combining Histograms

The above discussion assumes that one is comparing the data to two histograms, one representing the model for signal and the other that for background. However, it is often the case that the signal and/or background is composed of events from several different sources. (Since the fitted mass distributions and signal-to-background ratios are nearly identical for the e + jets and μ plus jets final states, the fit is not divided into channels as in the dilepton case.) While it would be possible to compare the data to a collection of histograms representing each of these sources, one would need to treat the normalization for each source as a parameter of the fit. This would result in the fit containing an unacceptable number of parameters.

In order to avoid this, the signal and background sources are combined into single histograms, with each source weighted such that the resultant histogram contains the fractional contribution expected. For example, assume that there are N_S sources of

background events, the ith of which is expected to contribute a fraction f_i of the total background, and further that that there are N_i events passing the selection cuts from the sample of events that model the source. Then the relative weight given to events from each source is determined by $w_i = f_i/N_i$.

It remains to determine the proper overall normalization S of the combined histogram [87, 81]. To do so, define

$$N_{\text{sum}} \equiv \sum_{i=1}^{N_S} w_i N_i$$

With the above definition of w_i, N_{sum} is unity. One then requires that when S is treated as a Poisson-distributed number its fractional uncertainty is equal to the uncertainty on N_{sum}, *i.e* that

$$\frac{\sqrt{S}}{S} = \frac{\sigma_{N_{\text{sum}}}}{N_{\text{sum}}} \tag{7.12}$$

or

$$S = \frac{N_{\text{sum}}^2}{\sigma_{N_{\text{sum}}}^2} = \left(\sum_{i=1}^{N_S} w_i^2 N_i\right)^{-1} \tag{7.13}$$

Additional details on the lepton plus jets likelihood fit may be found in [76].

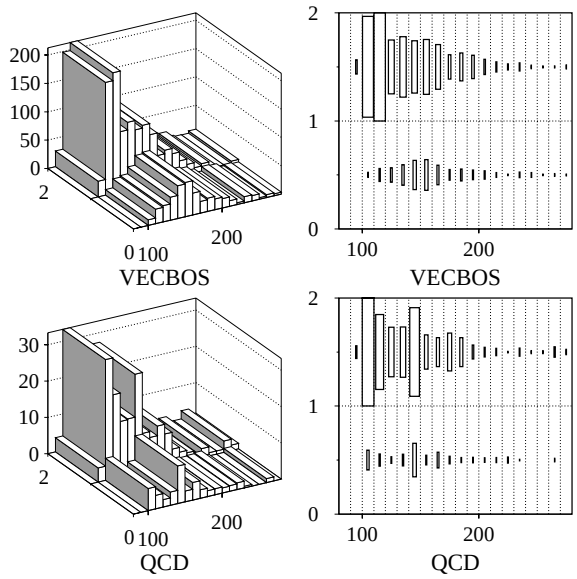

Figure 7.3: Distribution of background events, with binning as defined in Fig. 7.2. Both the W + multijet events, as modeled by VECBOS, and the QCD multijet events taken from data tend to have low top probability and relatively low reconstructed mass [76].

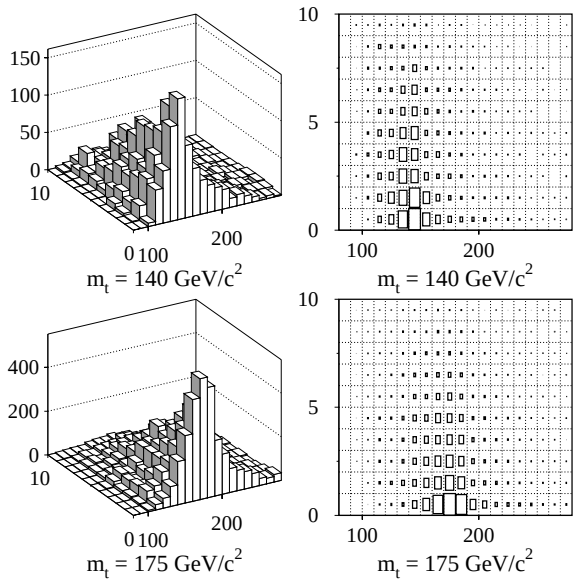

Figure 7.4: Distribution of reconstructed mass vs. neural network output for HERWIG 140 and 175 GeV/c^2 $t\bar{t}$ events [76].

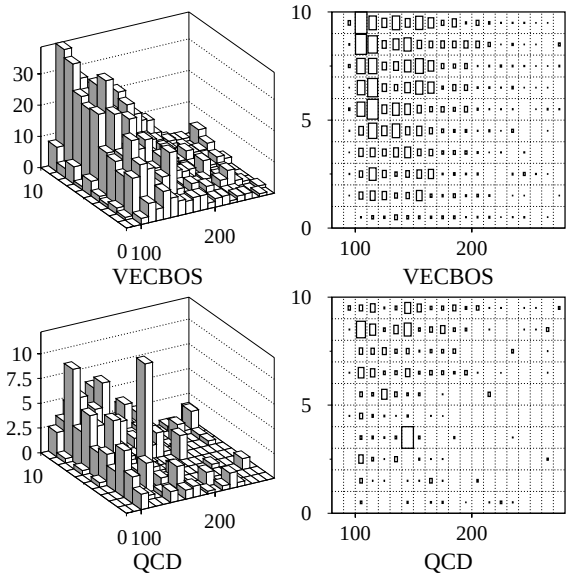

Figure 7.5: Distribution of reconstructed mass vs. neural network output for background events [76].

Chapter 8

Data Analysis

Now that the groundwork has been laid, we may proceed to the measurement of the top quark mass. This chapter describes the results of the dilepton and lepton plus jets analyses, and of their combination.

8.1 Dilepton Analysis

As detailed in Chapter 5, there are six $t\bar{t}$ candidate events in the dilepton final states. Figure 8.1 shows their $\mathcal{W}(m_t)$ distributions. Three of the events contain three jets, and in these cases the results of the fits using only the leading two jets and using all combinations of three jets are given (full details of the kinematics of these events may be found in Appendix B).

With this sample in hand, we can apply the likelihood fit methods introduced in the previous chapter in order to estimate the top quark mass. Before proceeding to this step, however, we need to verify that the maximum likelihood fit produces self-consistent results. In addition, the optimal value for the smoothing parameter h must be found, as well as the best procedure for fitting a smooth curve to the discrete $-\ln L$ points. Finally, we need to investigate the relationship between the shape of the likelihood curve and the accuracy of the top quark mass estimate (particularly, we

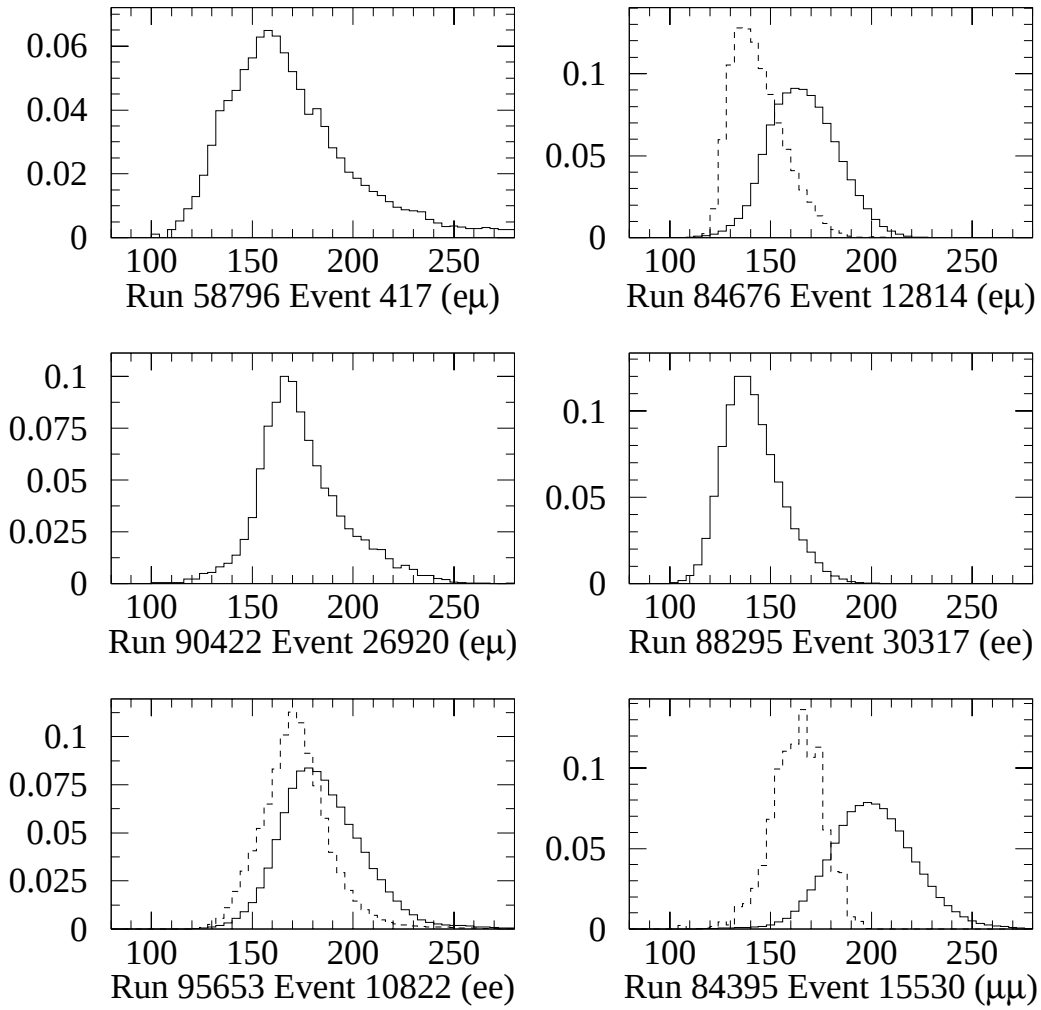

Figure 8.1: $\mathcal{W}(m_t)$ distributions for the dilepton candidates. For events with more than two jets, the dashed curves show the results of considering only the two highest E_T jets in the reconstruction.

need to test whether the interval in which $-\ln L$ changes by less than half a unit from its minimum contains the true top quark mass 68% of the time) in order to define a reasonable procedure for assigning the statistical uncertainty to the measurement. Monte Carlo *ensemble tests* provide a means to explore all of these issues.

8.1.1 Ensemble Tests

Ensemble tests are mock experiments, in which the candidate events are drawn from Monte Carlo with a known top quark mass (m_t^{MC}), but processed in exactly the same manner as the actual data. The procedure is as follows: if N_j events were observed in the jth decay channel, draw N_j events from the MC samples of the jth final state. Throw a random number as each event is drawn, and if the number is greater than $\overline{n_{b,j}}/N_j$, select an event from the signal sample. Otherwise, an event from the background sample is selected. If there are multiple sources of background, another random number is thrown in order to decide from which source to draw the event.

Once this Monte Carlo ensemble has been generated, the maximum likelihood fit procedure is performed upon it. (In order to avoid any possible bias, the Monte Carlo points that form the ensemble are not permitted to contribute to the estimation of the probability density.) This procedure is then repeated for a large number of ensembles (typically 1000). In this manner the statistical properties of the maximum likelihood estimate of the top quark mass ($\widehat{m_t}$) may be explored.

Among the desirable properties of this estimate, the foremost is that it be consistent with the input top quark mass. The necessary conditions for self-consistency include that the peak and median of the histogram of maximum likelihood masses lie at the mass of the signal MC[1]. Once a procedure has been found to be self-consistent, one needs to optimize any parameters of the procedure to maximize the accuracy of

[1]It is a subjective choice that the median, not the mean, is taken as the quantity which should reflect the input mass. The subjective goal is to quote as a result a top quark mass which has equal probability of being above or below the true value.

the estimated top quark mass.

Evaluation of Ensemble Results

In evaluating the results of ensemble tests run with a particular set of parameters, one would naturally prefer that the distribution of $\widehat{m_t}$ have narrow width and be highly responsive to the input top quark mass. Responsiveness may be evaluated by running the tests for two widely separated Monte Carlo top quark masses m_{t1}^{MC} and m_{t2}^{MC} and comparing the change in the mean $\widehat{m_t}$ to $m_{t2}^{\mathrm{MC}} - m_{t1}^{\mathrm{MC}} \equiv \Delta m_t^{\mathrm{MC}}$. The width is somewhat more ambiguous, since the distributions are generally not Gaussian. The choice made here is to calculate the narrowest range of m_t that contains 68.3% of the ensembles, and take one half of this range as the width parameter. This quantity is denoted R^{68}.

Finally, one may combine the responsiveness and width measurements into a single criterion Q, defined as:

$$Q(m_{t1}^{\mathrm{MC}}, m_{t2}^{\mathrm{MC}}) \equiv \frac{\Delta m_t^{\mathrm{MC}}}{\Delta\langle\widehat{m_t}\rangle} \cdot \frac{R^{68}(m_{t1}^{\mathrm{MC}}) + R^{68}(m_{t2}^{\mathrm{MC}})}{2} \tag{8.1}$$

where $\langle\widehat{m_t}\rangle$ is the mean of the maximum likelihood mass distribution. For some ensembles it happens that the $-\ln L$ points do not reach a minimum in the range where Monte Carlo exists, so that $\widehat{m_t}$ results from an extrapolation. As such extrapolated values are not trustworthy, all $\widehat{m_t}$ values less than 80 or greater than 260 GeV/c^2 are assigned these boundary values in calculating the mean.

The quantity Q estimates the typical resolution one would obtain in the top quark mass measurement, given that the mass lies in the range bounded by m_{t1}^{MC} and m_{t2}^{MC}. As such, those parameters which yield the smallest Q values are favored.

Optimization of the PDE h parameter

As described in Chapter 7, the probability density estimation technique used in evaluating L depends upon a free parameter h. Therefore, ensembles are processed

	$m_t^{\mathrm{MC}} = 150$ GeV/c^2			$m_t^{\mathrm{MC}} = 200$ GeV/c^2			
h	Median	Mean	R^{68}	Median	Mean	R^{68}	Q
0.1	153.3	155.2	11.9	201.1	200.8	18.6	16.7 ± 0.5
0.2	152.5	154.2	13.3	199.9	200.9	18.6	17.0 ± 0.5
0.3	151.7	153.3	13.9	198.9	200.4	19.5	17.7 ± 0.5
0.4	151.6	152.9	13.6	198.2	200.1	20.0	17.8 ± 0.5
0.5	151.7	153.0	13.5	197.7	200.1	20.7	18.1 ± 0.5
0.6	151.5	152.7	13.6	196.7	199.4	20.8	18.4 ± 0.5
0.7	151.6	152.6	13.6	195.9	198.4	21.0	18.8 ± 0.5
0.8	151.6	152.4	13.4	195.1	197.6	21.0	19.0 ± 0.5
0.9	151.5	152.2	13.5	194.0	196.6	21.1	19.5 ± 0.6

Table 8.1: Results of ensemble tests showing the effect of changing the PDE smoothing parameter h. All quantities are in GeV/c^2, and the Q criterion used to evaluate the relative sensitivity of the ensemble tests is defined in Eq. 8.1.

using a range of h values for two top quark masses (150 and 200 GeV/c^2, selected to represent a wide range about the previously measured value). For these tests, a quadratic fit to the nine $-\ln L$ points nearest the minimum is used to define $\widehat{m_t}$. Optimization of the parameterization technique is deferred to Section 8.1.1.

The results are given in Table 8.1. It is reassuring to note that the performance varies only by about 15% over the range of values studied. However, there is a preference for smaller values of h. The quality of the results from h ranging from 0.1 to 0.4 are practically indistinguishable, and $h = 0.3$ is chosen as a representative value within this range.

Self-consistency of the Analysis

As shown above, ensemble tests run with input masses of 150 and 200 GeV/c^2 produce $\widehat{m_t}$ distributions centered near the input mass. To explore the relationship between $\widehat{m_t}$ and m_t^{MC} more closely, ensembles were run for 130 GeV/c^2 $\leq m_t^{\mathrm{MC}} \leq$ 210 GeV/c^2, and the results are listed in Table 8.2.

m_t^{MC}	Median	Mean	R^{68}
130	138.2	139.8 ± 0.9	18.1
135	141.7	144.1 ± 0.8	15.7
140	145.9	147.5 ± 0.7	13.9
145	149.0	151.6 ± 0.7	15.0
150	151.9	154.5 ± 0.7	13.6
155	156.8	159.2 ± 0.7	14.4
160	161.5	163.5 ± 0.6	14.4
165	167.2	167.8 ± 0.7	16.0
170	172.2	173.0 ± 0.7	16.2
175	176.6	177.8 ± 0.7	17.1
180	180.5	181.3 ± 0.7	18.1
185	186.7	186.6 ± 0.7	17.0
190	188.7	189.6 ± 0.7	17.7
195	195.4	194.5 ± 0.8	18.4
200	198.8	199.4 ± 0.7	18.9
205	204.4	203.2 ± 0.8	19.9
210	210.1	210.0 ± 0.7	20.2

Table 8.2: Results of ensemble tests generated at a wide variety of m_t^{MC}.

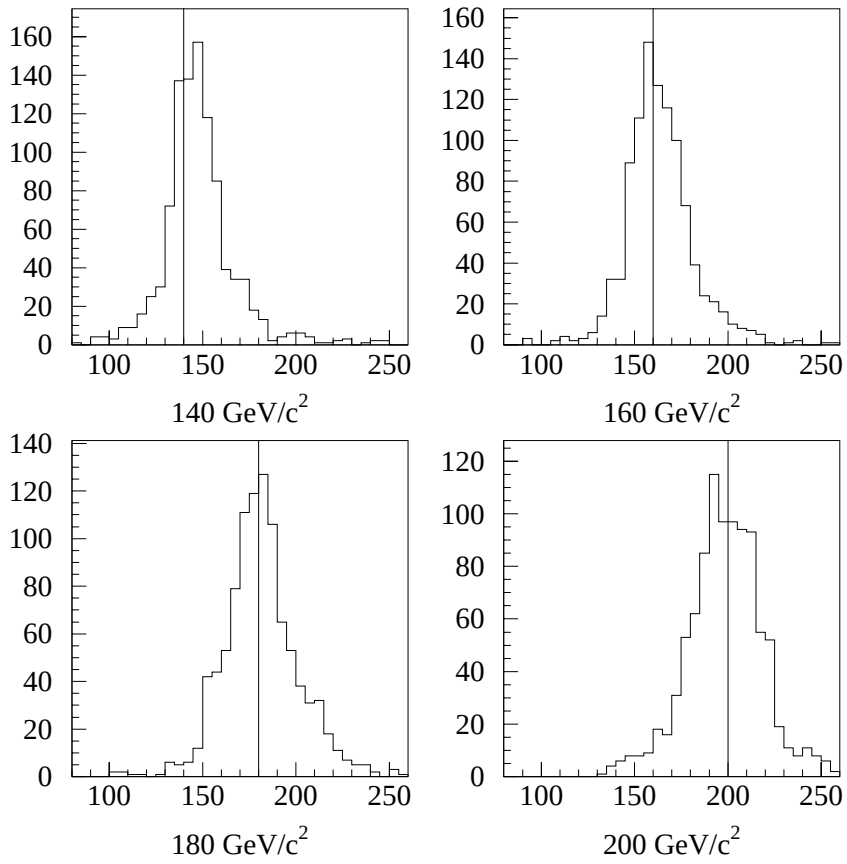

Figure 8.2: Distribution of $\widehat{m_t}$ for ensemble tests at a range of top quark masses.

There are substantial differences between $\widehat{m_t}$ and m_t^{MC} at low mass. These result from the nature of the kinematic selection, which biases the $\mathcal{W}(m_t)$ distributions at low mass. This leads to a loss of sensitivity of the likelihood fit to low m_t^{MC}.

In the high mass region (150 GeV/c^2 and above) the differences between $\widehat{m_t}$ and m_t^{MC} are less severe, with the maximum excursion in the median being 2.2 GeV/c^2. In addition, Fig. 8.2 shows that the peak of the $\widehat{m_t}$ distribution is consistent with m_t^{MC}. Therefore, if the result of the analysis exceeds 150 GeV/c^2, it is self-consistent.

Parameterization of the $-\ln L$ Curve

A quadratic fit is the simplest procedure one could use to model the shape of the $-\ln L$ curve, and also reflects the expectation that the curve should approach a parabola for a high statistics sample. However, for a six-event sample the curve cannot be parabolic over a large range of top quark masses, due to the limited significance of the excess of events over background. For example, if the event kinematics are very inconsistent with a particular top quark mass, the most likely solution will be that with n_s fluctuating to zero. At this point $-\ln L$ becomes independent of m_t. Therefore one must restrict the fit to some region near the minimum where the behavior may be approximately parabolic.

It remains to test how the result depends on the size of the region chosen, and whether fitting the shape to a more complicated function could improve the result. Therefore, fitting algorithms which differ in the number of points near the minimum considered, and in the order of the polynomial fit (quadratic or cubic) were tried, and the results are given in Table 8.3. The points that were fit were all assigned equal weight, as no attempt was made to estimate an uncertainty in $-\ln L$. One sees no gain in the quality of the result by allowing the additional complexity of a cubic fit, and little variation among the quadratic fits. Therefore a quadratic fit is favored, and the proper range of points over which to fit is discussed in the next section.

Assignment of Statistical Uncertainty

Besides returning an estimate of m_t, the maximum likelihood fit also returns an estimate of the statistical uncertainty of the measurement ($\hat{\sigma}$), defined as one half the interval over which $\ln L$ changes by less than half a unit from its maximum. The parameterization of the $-\ln L$ curve which gives the most reliable $\hat{\sigma}$ is favored. One way to investigate this is to calculate the *pull distribution* of the ensemble results. For each ensemble, the pull is defined as:

$$\text{Pull} \equiv \frac{\widehat{m_t} - m_t^{\text{MC}}}{\hat{\sigma}} \tag{8.2}$$

	$m_t^{\mathrm{MC}} = 150$ GeV/c^2			$m_t^{\mathrm{MC}} = 200$ GeV/c^2			
Type of Fit	Median	Mean	R^{68}	Median	Mean	R^{68}	Q
5-pt. Quadratic	152.2	154.1	13.4	198.1	197.8	18.6	18.2 ± 0.5
7-pt. Quadratic	151.6	154.0	13.0	198.2	198.1	19.0	18.1 ± 0.5
9-pt. Quadratic	151.9	154.5	13.6	198.8	199.4	18.9	$18,1 \pm 0.5$
9-pt. Cubic	151.6	151.8	13.3	196.0	190.0	19.6	21.5 ± 0.8
11-pt. Cubic	151.9	152.5	13.8	193.4	196.3	19.3	20.2 ± 0.6

Table 8.3: Results of ensemble tests showing the effect of different parameterizations of the $-\ln L$ curve. The fits are described by the number of points about the minimum $-\ln L$ considered, and the degree of polynomial used.

If $\hat{\sigma}$ is a good estimate of the uncertainty in the measurement, the distribution of pulls should be a unit Gaussian centered at zero. The pulls for all ensembles with $80 \leq \widehat{m}_t \leq 260$ GeV/c^2 are histogrammed, and a Gaussian is fit to the distribution. The resultant widths are tabulated for a variety of m_t^{MC} and fit procedures in Table 8.4.

In general the fits which consider only five points tend to produce curves that are too narrow. The distinction between the seven- and nine-point fits is less clear, but the nine-point fits give pull widths closer to unity over the range considered. Therefore the final results will employ the nine-point quadratic parameterization of the $-\ln L$ curve.

The pull distributions for ensemble tests at a variety of top quark masses are shown in Fig. 8.3.

8.1.2 Results

Now that the procedures for the likelihood fit have been defined, one may apply them to the data sample. Using the baseline procedure outlined above, one finds for the dilepton sample:

$$m_t = 169.9 \pm 14.8 \text{ (stat.) GeV/c}^2 \tag{8.3}$$

m_t^{MC}	5-pt. Quadratic	7-pt. Quadratic	9-pt. Quadratic
130	1.22	1.04	1.04
135	1.11	1.03	0.92
140	1.09	0.97	0.88
145	1.14	1.00	0.90
150	1.03	0.92	0.86
155	1.14	0.94	0.88
160	1.18	0.99	0.96
165	1.22	1.06	0.97
170	1.17	1.06	0.98
175	1.22	1.08	0.97
180	1.27	1.11	1.03
185	1.14	1.05	0.96
190	1.16	1.05	0.99
195	1.08	1.05	1.00
200	1.07	1.10	1.02
205	1.07	1.05	1.02
210	1.08	1.01	1.03

Table 8.4: Pull widths for ensemble tests generated at a wide variety of m_t^{MC} with five, seven, or nine points near the minimum considered in the quadratic fit.

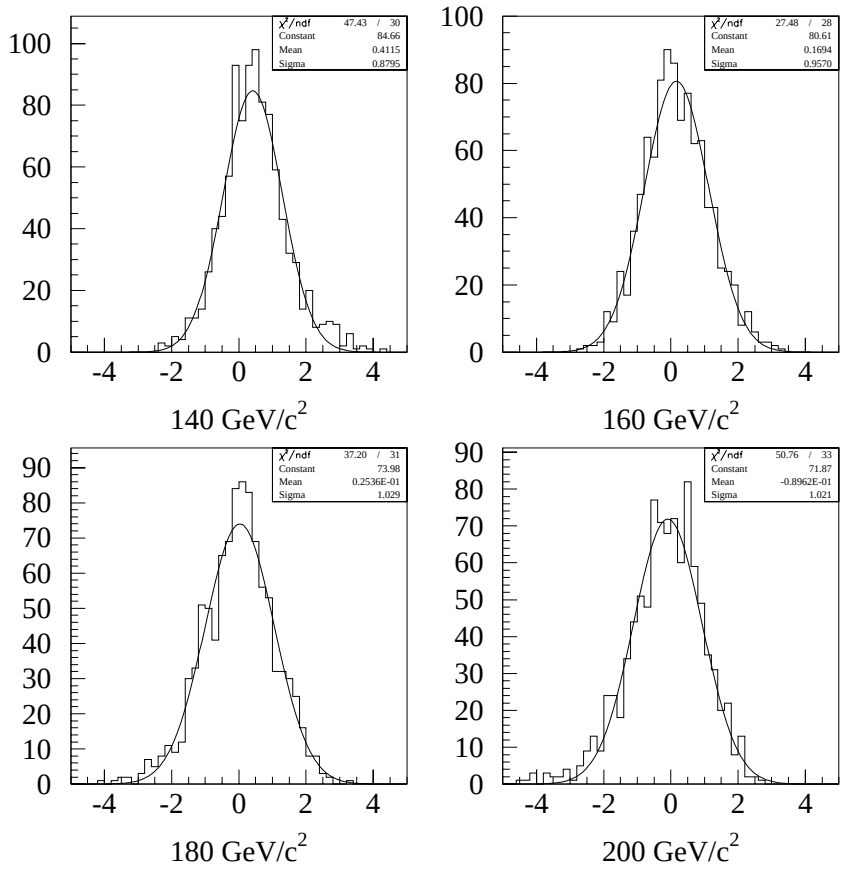

Figure 8.3: Pull distributions with Gaussian fits for a variety of m_t^{MC}

Figure 8.4 compares the $\vec{w}$'s for the data sample to the best-fit signal and background models, and shows the $-\ln L$ curve.

The statistical error quoted above is slightly smaller than the R^{68} observed in ensemble tests with $m_t^{MC} = 170$ GeV/c^2. Observing a likelihood curve of this width is very likely, as illustrated in Fig. 8.5(a). In addition, the pull distribution indicates that $\hat{\sigma}$ is a reasonable estimate. One can directly verify this by considering that subset of ensembles with $\hat{\sigma}$ consistent with the observed value. This is shown in Fig. 8.5(b), in which it is seen that the R^{68} of all such ensembles is quite consistent with $\hat{\sigma}$.

One may assess the quality of the maximum likelihood fit by comparing the minimum $-\ln L$ to that obtained in ensemble tests. As shown in Fig. 8.6, 80% of all ensembles result in a minimum $-\ln L$ greater than the 12.8 observed in the data.

In the dilepton fit there is a clear hierarchy of importance among the channels. The $e\mu$ channel, with the largest sample and smallest background, should dominate the fit, while the $\mu\mu$ channel with only one event and a sizeable background should have the least impact. Therefore it is of interest to see how the results of the fit vary as the less important channels are excluded. To begin, one may drop the $\mu\mu$ channel from consideration, yielding:

$$m_t = 173.2 \pm 14.0 \text{ (stat.) GeV/c}^2 \tag{8.4}$$

and finally consider only the $e\mu$ events, which give:

$$m_t = 170.1 \pm 14.5 \text{ (stat.) GeV/c}^2 \tag{8.5}$$

Figure 8.7 shows the maximum likelihood fit for each of these cases.

Although the ensemble tests give no indication that these results should depend strongly on the details of the procedure, one would like to see that this is in fact the case for the data sample. Table 8.5 lists the results of some variations on the analysis. These variations consist of using different parameterizations of the $-\ln L$ curve, and also performing fits in which only the two highest E_T jets are considered, rather than

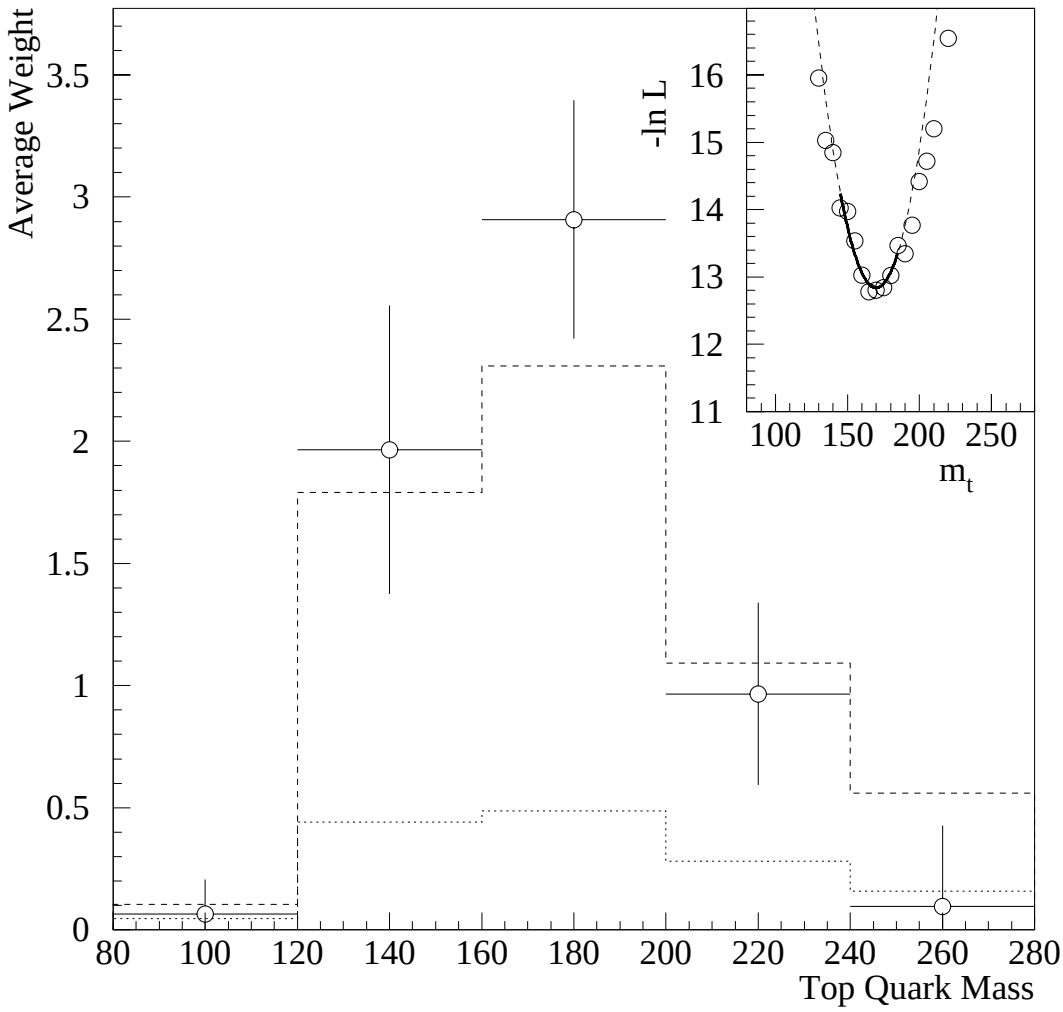

Figure 8.4: Result of the maximum likelihood fit to the dilepton sample. Plotted are the summed $w_1 \ldots w_5$ for the data sample (points), and the prediction of the best-fit model of background (dotted) and signal plus background (dashed). The error bars on the data points represent the RMS of the w_i in ensemble tests. The $-\ln L$ curve and its parameterization are inset at upper right. The curve is drawn as a solid line in the region considered in the fit.

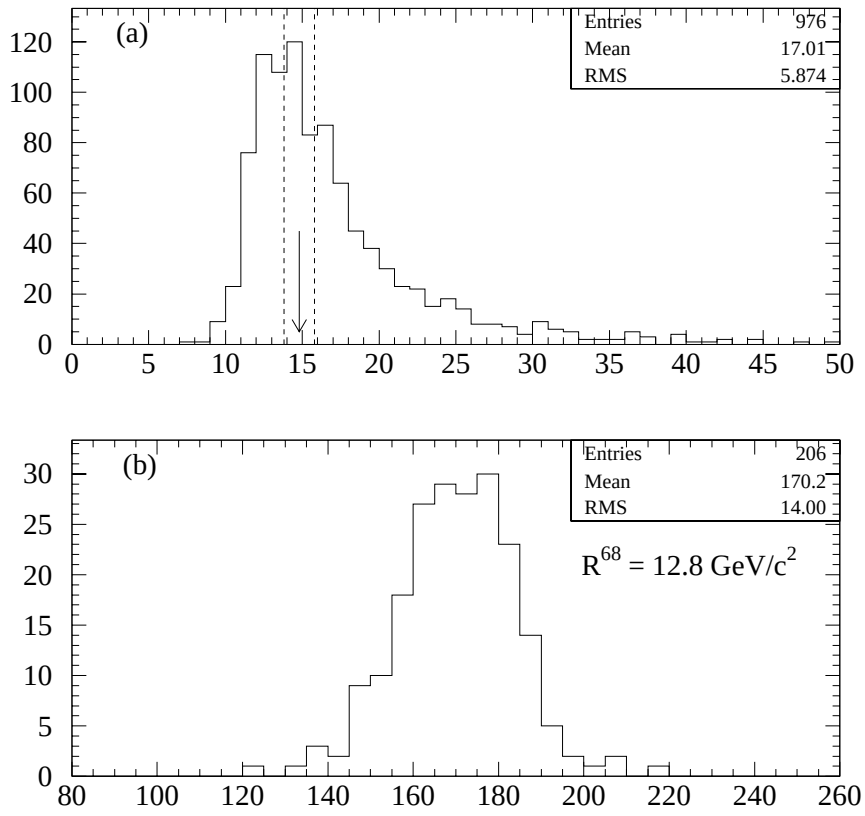

Figure 8.5: Figure (a) shows the distribution of $\hat{\sigma}$ obtained from ensemble tests run with $m_t^{\mathrm{MC}} = 170$ GeV/c^2. The arrow marks the value returned by the fit to the data (14.8 GeV/c^2). The dashed lines bound the region of ensembles which enter (b).

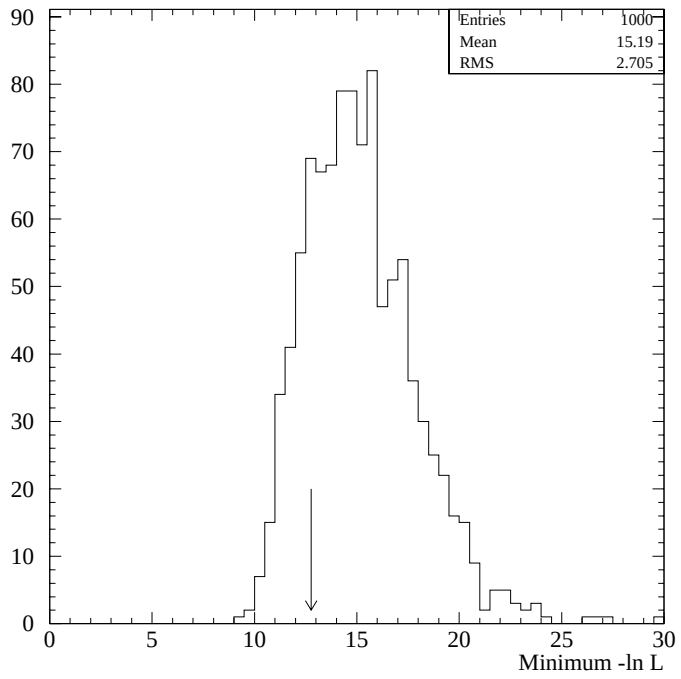

Figure 8.6: Distribution of minimal $-\ln L$ obtained from 170 GeV/c^2 ensemble tests. The arrow indicates the result obtained from the data.

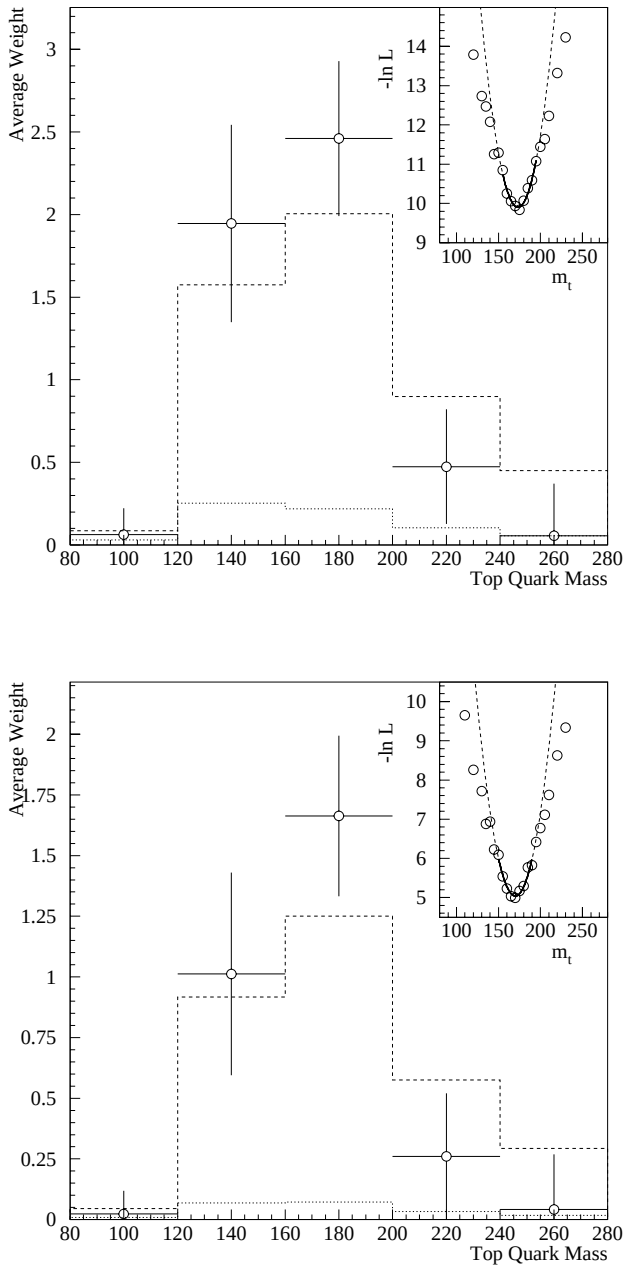

Figure 8.7: Result of the maximum likelihood fit to the ee plus $e\mu$ channels (top) and $e\mu$ only (bottom).

summing over all possible combinations of the leading three jets. No excursions larger than the statistical uncertainty of the measurement are seen in the results of any of these variations.

8.1.3 Systematic Uncertainties

The final step in the analysis is to estimate the systematic uncertainties in the above measurement. Here again the ensemble tests are valuable, as one may vary the quantity under consideration in the ensemble and directly observe the resultant change in $\widehat{m_t}$. The sources of systematic error are listed in Table 8.6, and the procedures by which they were determined are outlined below.

Jet Energy Scale

The first systematic considered is the uncertainty in the jet energy scale, which is calibrated by using photon plus jets events as described in Chapter 4. The final correction that brings the data and Monte Carlo scales to the same level is done as a function of jet $|\eta_d|$, and the degree of possible residual discrepancy is gauged by plotting the relative photon-jet balance as a function of photon E_T. Figure 8.8 shows the result of this comparison, and the curves superposed show the effect of a $\pm 2.5\%$ mismatch in jet response added linearly to a ± 0.5 GeV relative offset. As these curves contain over two-thirds of the points it seems reasonable to set this level of mismatch as the uncertainty in the jet scale.

In order to propagate this uncertainty to the measured top mass, one generates signal and background Monte Carlo samples with the jet energies moved up and down by one standard deviation. The unclustered transverse energy (the net transverse energy in the calorimeter which is not included in any jet), defined by:

$$p_i(b) \equiv -\not{E}_i + \sum_{n=\text{Objects}} p(n)_i$$

is also scaled by the same factor as the jets, and the $\not{E}_T$ is altered to reflect the scaling

Channels fit	Number of jets	Fit type	Result
$ee + e\mu + \mu\mu$	3	5-pt. Quadratic	169.2 ± 11.4
		7-pt. Quadratic	170.3 ± 12.8
		9-pt. Quadratic	$\mathbf{169.9 \pm 14.8}$
		11-pt. Cubic	$171.4^{+16.0}_{-16.3}$
	2	5-pt. Quadratic	165.3 ± 9.3
		7-pt. Quadratic	165.3 ± 12.1
		9-pt. Quadratic	163.6 ± 14.1
		11-pt. Cubic	$164.9^{+14.3}_{-13.3}$
$ee + e\mu$	3	5-pt. Quadratic	172.7 ± 12.1
		7-pt. Quadratic	172.2 ± 14.7
		9-pt. Quadratic	173.2 ± 14.0
		11-pt. Cubic	$171.8^{+13.3}_{-14.9}$
	2	5-pt. Quadratic	164.2 ± 11.0
		7-pt. Quadratic	163.9 ± 12.9
		9-pt. Quadratic	162.1 ± 15.2
		11-pt. Cubic	$163.2^{+16.1}_{-14.8}$
$e\mu$	3	5-pt. Quadratic	168.9 ± 14.3
		7-pt. Quadratic	168.8 ± 13.3
		9-pt. Quadratic	170.1 ± 14.5
		11-pt. Cubic	$169.5^{+15.4}_{-15.5}$
	2	5-pt. Quadratic	161.9 ± 18.1
		7-pt. Quadratic	159.2 ± 17.2
		9-pt. Quadratic	159.7 ± 17.1
		11-pt. Cubic	$159.1^{+16.8}_{-18.3}$

Table 8.5: Results of several variations of the maximum likelihood fit to the data. The error estimates quoted are simply the $\hat{\sigma}$, with no correction for any non-unit pull width.

Source	Error (GeV/c^2)
Jet Energy Scale	2.9
Multiple Interactions	1.0
Background Shape	1.5
Signal Generator	1.1
Likelihood Fit	1.3
Monte Carlo Statistics	0.3
Total	3.8

Table 8.6: Sources of systematic error. See the text for details on how these errors were estimated.

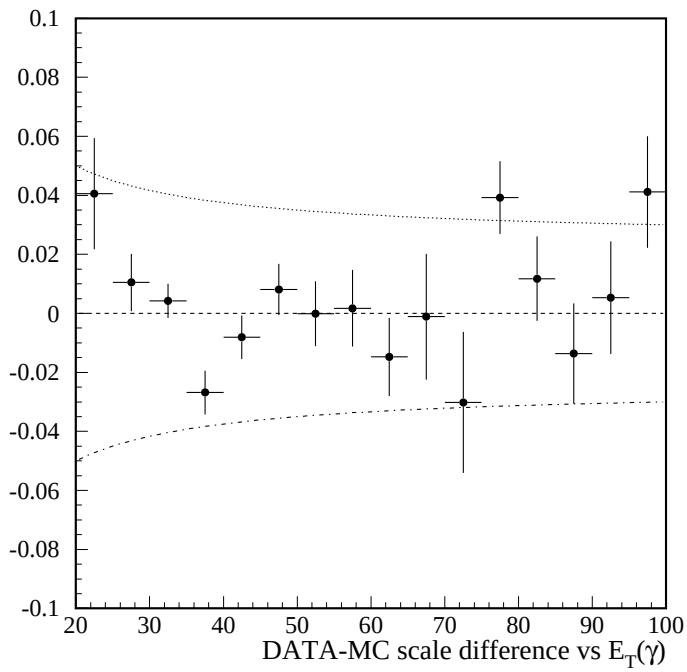

Figure 8.8: Difference in photon-jet transverse energy balance between Monte Carlo and data samples as a function of photon E_T. The dashed curves show the offset that would result if the data and Monte Carlo jet energy scales were offset relative to one another by $\pm(2.5\% + 0.5$ GeV) [54].

Monte Carlo jet scale	Median $\widehat{m_t}$
$+1\sigma$	174.0
Nominal	172.2
-1σ	168.3

Table 8.7: Effect of varying the jet scale in ensemble tests with $m_t^{\mathrm{MC}} = 170$ GeV/c^2.

of the objects. One then repeats the ensemble tests using ensembles drawn from the scaled samples, but using the unscaled samples in the probability density estimation. The results of this scaling are shown in Table 8.7 for an input mass of 170 GeV/c^2. Averaging the upward and downward excursions of the median results in a systematic uncertainty of 2.9 GeV/c^2.

Signal Monte Carlo Generator

Accurate determination of the top quark mass also depends on the signal Monte Carlo providing a faithful description of $t\bar{t}$ events. Some features, in particular gluon radiation and parton fragmentation, are only modeled approximately by HERWIG, and other reasonable approximations exist. In the absence of large samples of top events one cannot directly exclude any such approximations. To test the sensitivity of the result to changing the Monte Carlo generator, ensembles of ISAJET events were formed, and passed through the maximum likelihood fit with the baseline HERWIG used to model of the signal probability densities. The results are listed in Table 8.8. Taking the average of the absolute value of the discrepancies in the median gives a systematic uncertainty of ± 1.1 GeV/c^2.

Background Shape

The modeling of the background also depends on a Monte Carlo simulation. In addition, for some sources of background Monte Carlo statistics are very limited.

m_t^{MC}	Mean	Median	R^{68}	ΔMean	ΔMedian
140	147.8 ± 0.7	145.9	15.6	0.3 ± 1.0	0.0 ± 1.3
150	154.4 ± 0.7	152.6	15.4	-0.1 ± 1.0	0.7 ± 1.1
160	161.6 ± 0.6	160.1	15.8	-1.9 ± 0.8	-1.4 ± 1.1
170	171.6 ± 0.7	170.8	17.6	-1.4 ± 1.0	-1.4 ± 1.2
180	179.5 ± 0.7	179.1	18.2	-1.8 ± 1.0	-1.4 ± 1.2
190	188.7 ± 0.7	189.4	18.5	-0.9 ± 1.0	0.7 ± 1.1
200	198.3 ± 0.7	198.6	19.5	-1.1 ± 1.0	-0.2 ± 1.3
210	205.6 ± 0.8	206.8	20.3	-4.4 ± 1.1	-3.3 ± 1.3

Table 8.8: Results of fitting ensembles of ISAJET events to signal samples modeled by HERWIG.

Dummy background model	Median $\widehat{m_t}$
Low mass	172.7
High mass	171.2

Table 8.9: Effect of introducing dummy models for that portion of the background which has no Monte Carlo model.

To estimate how sensitive the result is to the unknown distribution of these events, dummy models were created to play their role. These models assume that the $\mathcal{W}(m_t)$ distributions are Gaussian, with a width chosen randomly between 20 and 60 GeV/c^2. In one of the models ("low mass"), the mean of the Gaussian was randomly selected between 120 and 160 GeV/c^2, and in the other ("high mass") the mean was set between 180 and 220 GeV/c^2. Ensembles were then run using the known background models plus the dummies to estimate the probability densities, with events drawn from the standard signal and background models. The results are listed in Table 8.9, and based on the observed shift in median $\widehat{m_t}$ an error of 1.5 GeV/c^2 is assigned.

Multiple Interactions

As mentioned in Chapter 5, one of the known differences between the Monte Carlo models and the actual data is that the latter includes events in which more than one $\overline{p}p$ pair interacted while the former does not. There are two ways in which multiple interactions may effect the reconstructed event. First, the additional particles deposit energy in the calorimeter, some of which falls into the jet cones. Secondly, the additional tracks may confuse the algorithm for determining the z vertex, leading to mismeasurement of the jet directions. The CAFIX package is designed to account for the former effect on average, but there is no protection against the latter.

To study these effects, a sample of 5000 HERWIG 170 GeV/c^2 $t\overline{t}$ to dilepton events was processed through NOISY. In this processing, particles from one or two simulated minimum bias events are added to those generated by the $t\overline{t}$ model, and the resultant event is passed through full GEANT simulation (full GEANT is needed to model the determination of the z vertex in such events). The samples are then processed through the same DORECO and MUSMEAR as all other Monte Carlo samples.

For this study ensemble tests are of little help, since the small sample sizes prohibit the generation of a large number of independent ensembles. One can estimate the size of the systematic effect, however, by comparing the $\mathcal{W}(m_t)$ distributions in the samples with zero, one, and two additional interactions, as shown in Fig. 8.9. Although the resolution of the z vertex degrades as interactions are added, the effect on the $\mathcal{W}(m_t)$ distribution is modest, with only a change of 0.7 GeV/c^2 observed in the mean when going from zero to two extra interactions. A change of this magnitude would be roughly equivalent to a change of 1.5 GeV/c^2 in the top quark mass. To reflect the fact that at least part of this change is accounted for by CAFIX in the actual data, an uncertainty of 1.0 GeV/c^2 is assigned to this source.

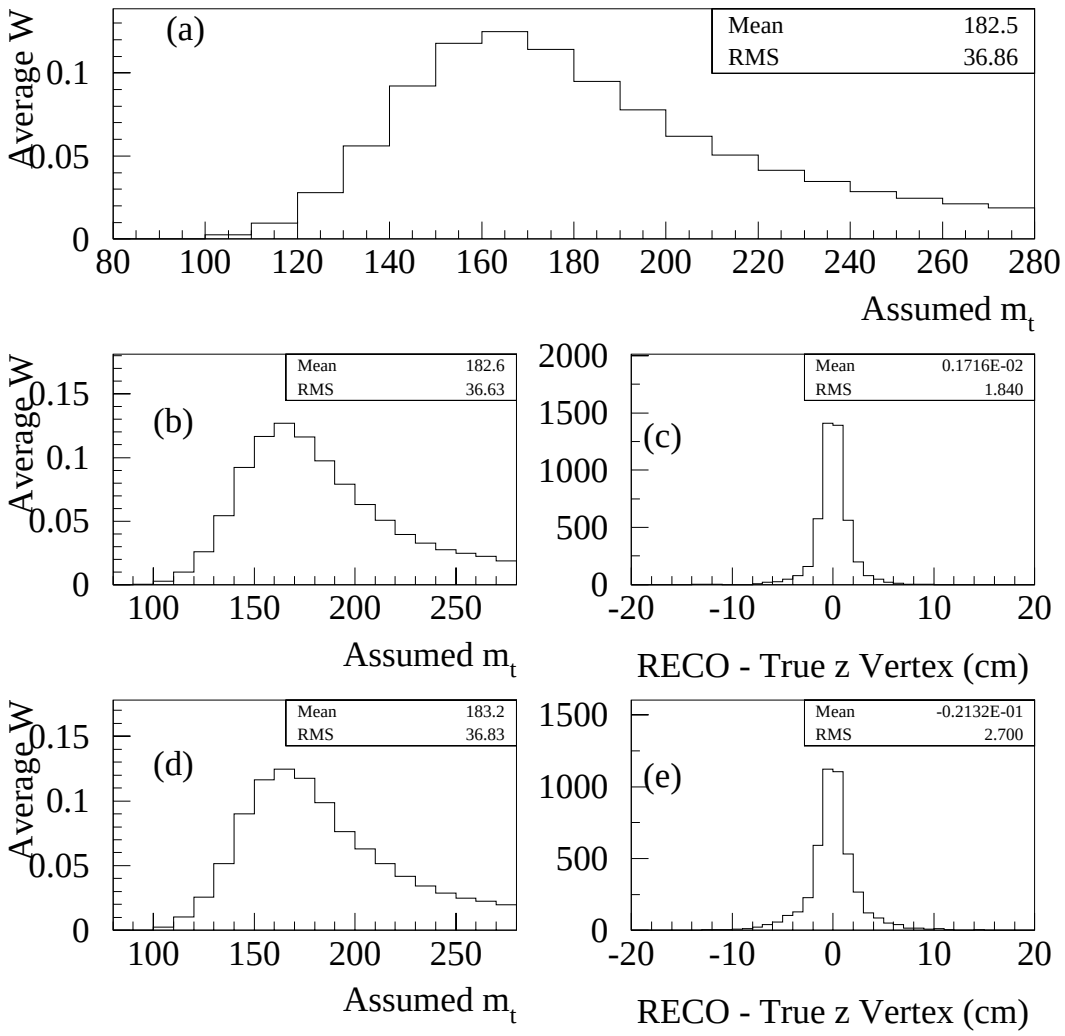

Figure 8.9: Average $\mathcal{W}(m_t)$ distributions for a sample of HERWIG 170 GeV/c^2 $t\bar{t}$ to dilepton events, with zero (a), one (b) and two (d) additional minimum bias interactions overlaid. Figures (c) and (e) show the resolutions of the z vertex determination for the one and two additional interaction samples.

Likelihood Fit and Monte Carlo Statistics

There is also a systematic uncertainty in the knowledge of the value of the top quark mass which minimizes $-\ln L$. This arises both from the limits of Monte Carlo statistics in determining the $-\ln L$ points and the use of a parabola to fit them. The former effect is estimated by splitting the signal Monte Carlo samples into five subsets and repeating the fit to the data using each subset as the signal model. The RMS variation observed in the central value is then divided by $\sqrt{5}$, yielding a systematic uncertainty of 0.3 GeV/c^2.

To estimate the uncertainty arising due to the nine-point parabolic parameterization of the $-\ln L$ distribution, $m_t^{MC} = 170$ GeV/c^2 are fit using a variety of parameterizations, and the resultant change in median $\widehat{m}_t$ observed. Five- and seven-point quadratic fits are considered, as are nine- and eleven-point cubic fits, and the maximal variation of 1.3 GeV/c^{2}is taken as the systematic.

Summing the above systematic errors in quadrature, one obtains:

$$m_t = 169.9 \pm 14.8 \text{ (stat.)} \pm 3.8 \text{ (syst.) GeV/c}^2 \tag{8.6}$$

for events in the dilepton channels.

8.1.4 Five Event Analysis

As described in Chapter 5, one of the dielectron events (95653/10822) that entered the above analysis was included by fiat even though it did not satisfy the selection criteria. While an attempt was made to calculate the additional background introduced by including events in its class, such calculation is necessarily ambiguous. Therefore, in this section the analysis is repeated with event 95653/10822 excluded.

Much of the groundwork established for the above analysis (for example, the optimization of the h parameter and the form of the parameterization of the $-\ln L$ curve) need not be repeated here. One does need to check, however, that the analysis remains self-consistent, and that $\hat{\sigma}$ remains a good estimate of the statistical uncer-

m_t^{MC}	Median	Mean	R^{68}
130	138.4	140.5 ± 0.9	21.1
135	141.2	144.8 ± 0.9	19.5
140	146.7	148.7 ± 0.8	17.3
145	150.2	153.1 ± 0.8	16.2
150	152.9	155.9 ± 0.7	15.5
155	157.3	159.6 ± 0.7	16.5
160	160.9	162.9 ± 0.7	16.8
165	168.4	169.0 ± 0.7	18.0
170	172.5	173.3 ± 0.7	18.7
175	177.4	178.7 ± 0.8	18.8
180	181.4	181.5 ± 0.7	19.3
185	185.5	185.4 ± 0.7	20.2
190	188.7	189.7 ± 0.8	19.5
195	193.8	192.7 ± 0.8	19.8
200	200.3	199.2 ± 0.8	21.0
205	204.0	202.4 ± 0.8	21.2
210	210.2	208.0 ± 0.9	22.5

Table 8.10: Results of five-event ensemble tests generated at a wide variety of m_t^{MC}.

tainty, when the sample size is reduced to five events. Therefore, ensemble tests were run with only one dielectron event in the sample, with the results listed in Tables 8.10 and 8.11. As before, one finds that the analysis is self-consistent for $m_t > 150$ GeV/c^2, and the pull widths are reasonable.

Applying the maximum likelihood fit to the data sample yields:

$$m_t = 166.4 \pm 17.6 \text{ (stat.) GeV/c}^2 \tag{8.7}$$

The distribution of $\vec{w}$s for these events, and the $-\ln L$ curve, are shown in Fig. 8.10. The minimal $-\ln L$ (10.3) is found to be less than that observed in 83% of ensemble tests (see Fig. 8.11). The $\hat{\sigma}$ is quite close to the value one would expect from the ensemble R^{68}. The results of the analysis under some variations of the procedure are

m_t^{MC}	5-pt. Quadratic	7-pt. Quadratic	9-pt. Quadratic
130	1.17	1.00	1.05
135	1.14	1.01	0.95
140	1.07	1.01	0.97
145	1.07	0.88	0.89
150	1.10	0.95	0.84
155	1.19	0.99	0.91
160	1.18	1.06	0.94
165	1.29	1.08	1.02
170	1.28	1.11	1.03
175	1.21	1.04	0.95
180	1.23	1.07	0.93
185	1.21	1.05	1.02
190	1.15	1.06	0.99
195	1.15	1.08	1.06
200	1.18	1.05	1.01
205	1.12	1.10	0.97
210	1.02	0.97	0.95

Table 8.11: Pull widths for five-event ensemble tests generated at a wide variety of m_t^{MC} with five, seven, or nine points near the minimum considered in the quadratic fit.

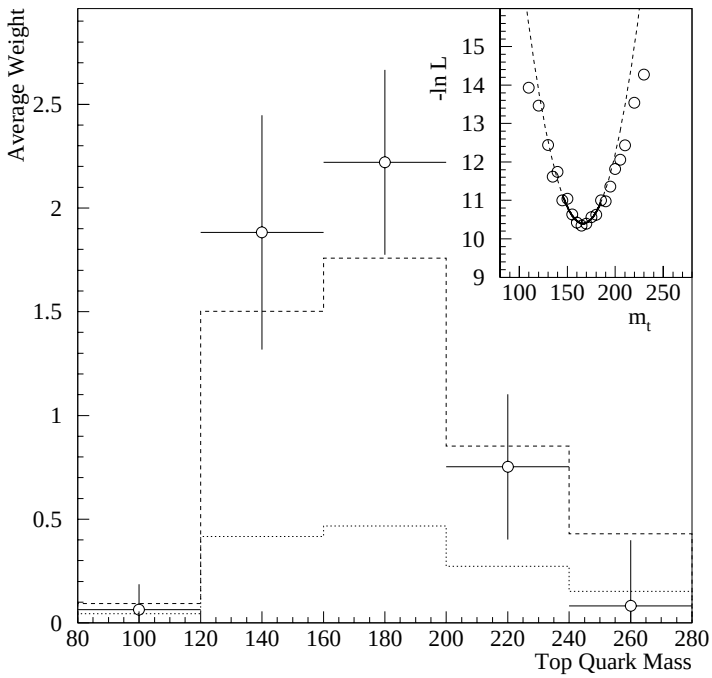

Figure 8.10: Result of the maximum likelihood fit for the five-event analysis.

shown in Table 8.12.

Systematic Uncertainties

The procedures for determining the systematic uncertainties are identical to those described for the six-event analysis. The uncertainties due to the likelihood fit procedure and multiple interactions are taken to be the same as found above. Results for the other sources are given in the following Tables: jet scale (Table 8.13); signal Monte Carlo (Table 8.14); and background shape (Table 8.15). Summing all sources in quadrature, we obtain:

$$m_t = 166.4 \pm 17.6 \text{ (stat.)} \pm 4.5 \text{ (stat.)} \text{ GeV/c}^2 \tag{8.8}$$

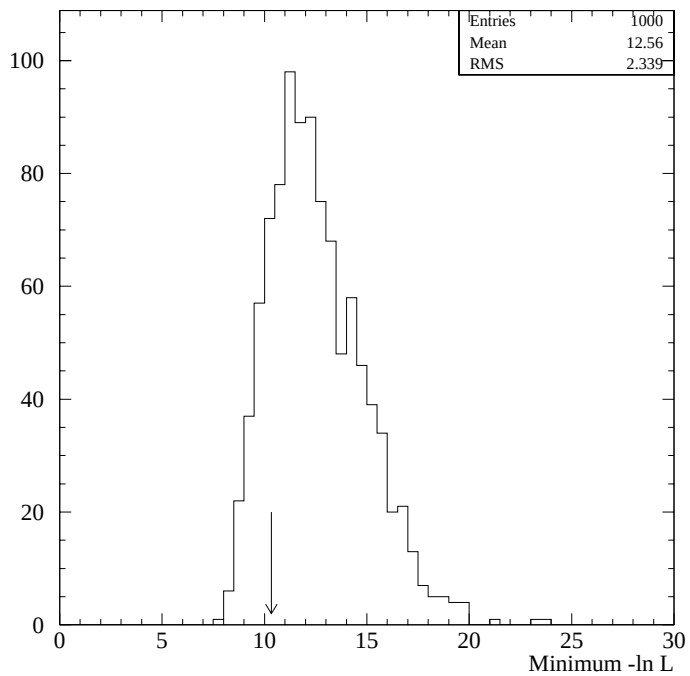

Figure 8.11: Distribution of minimal $-\ln L$ obtained from five-event 165 GeV/c^2 ensemble tests. The arrow indicates the result obtained from the data.

Number of jets	Fit type	Result
3	5-pt. Quadratic	165.7 ± 14.2
	7-pt. Quadratic	167.4 ± 15.3
	9-pt. Quadratic	$\mathbf{166.4 \pm 17.6}$
	11-pt. Cubic	$165.4^{+16.6}_{-21.0}$
2	5-pt. Quadratic	159.4 ± 17.9
	7-pt. Quadratic	155.7 ± 17.5
	9-pt. Quadratic	158.3 ± 15.9
	11-pt. Cubic	$156.6^{+13.8}_{-16.4}$

Table 8.12: Results of some variations of the five-event analysis.

Monte Carlo jet scale	Median $\widehat{m_t}$
$+1\sigma$	174.7
Nominal	172.5
-1σ	168.4

Table 8.13: Effect of varying the jet scale in ensemble tests with $m_t^{\mathrm{MC}} = 170$ GeV/c^2 for the five-event analysis. The symmetric excursion is ± 3.2 GeV/c^2.

m_t^{MC}	Mean	Median	R^{68}	ΔMean	ΔMedian
140	147.8 ± 0.9	145.9	17.9	-0.9 ± 1.2	-0.8 ± 1.3
150	157.6 ± 0.8	153.8	17.7	1.7 ± 1.1	0.9 ± 1.3
160	162.7 ± 0.8	161.5	19.0	-0.2 ± 1.1	0.6 ± 1.3
170	172.9 ± 0.7	171.1	18.9	-0.4 ± 1.0	-1.4 ± 1.3
180	178.4 ± 0.7	178.3	19.1	-3.1 ± 1.0	-3.1 ± 1.3
190	189.0 ± 0.8	187.7	20.6	-0.7 ± 1.1	-1.0 ± 1.4
200	195.8 ± 0.8	196.8	22.6	-3.7 ± 1.1	-3.5 ± 1.4
210	204.2 ± 0.9	205.8	23.0	-3.8 ± 1.3	-4.4 ± 1.7

Table 8.14: Results of fitting ensembles of ISAJET events to signal samples modeled by HERWIG in the five-event case. The mean absolute difference in the median $\widehat{m_t}$ is 2.0 GeV/c^2.

Dummy background model	Median $\widehat{m_t}$
Low mass	173.2
High mass	171.4

Table 8.15: Effect of unmodeled backgrounds in the five-event analysis.

8.2 Lepton Plus Jets Analysis

The analysis of events in the lepton plus jets channels proceeds in much the same manner as that for the dilepton channels, with the major differences being the definition of the fit likelihood, as detailed in Chapter 7. The data sample upon which the fit is performed consists of the 77 events which pass the cuts tailored for the mass analysis.

The main properties of the candidate sample are illustrated by plotting the events on the two-dimensional distributions of reconstruction mass vs. top probability, as defined in Chapter 7. Figures 8.12 and 8.13 show this distribution, both for the LB and NN methods of defining the top probability. One sees that the events which populate the low reconstructed mass region also tend to have low probability for being $t\bar{t}$ events. Further, one notes that the distribution of events is quite consistent with the expectation that the sample contains some fraction of $t\bar{t}$ events amongst the background events which dominate the sample.

In order to determine what this fraction is, and the value of the top quark mass most consistent with the observed distribution, the binned maximum likelihood fit described in Chapter 7 is performed. As in the dilepton case, the reliability of the likelihood fit results must be verified by means of Monte Carlo ensemble tests. However, as the level of background in the data sample is estimated by the likelihood fit itself without constraint from the expectations of the cross section measurement, the signal-to-background ratio to be used in the construction of the ensembles is ambiguous. For the results presented here, the ensembles are chosen to have a 1:2 expected signal-to-background ratio.

The results of these tests are listed in Table 8.16 (where again a quadratic fit to the nine $-\ln L$ points around the minimum is used to determine $\widehat{m_t}$). It is seen that the expected precision of the measurement is nearly independent of whether one uses the NN or LB discriminant to evaluate the likelihood that an event is top. Figure 8.14 shows the results of the ensemble tests for an input top quark mass of

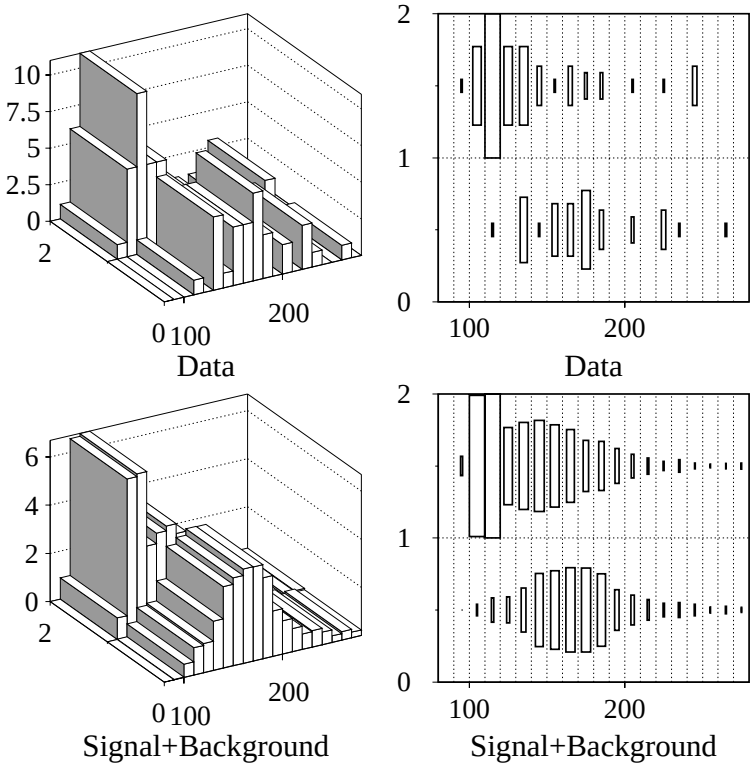

Figure 8.12: Distribution of candidate events (top plots) on LB plot, where events in the first bin on the likelihood axis are those most likely to be top. The bottom plots show the distribution for the best-fit signal plus background model of the data [76].

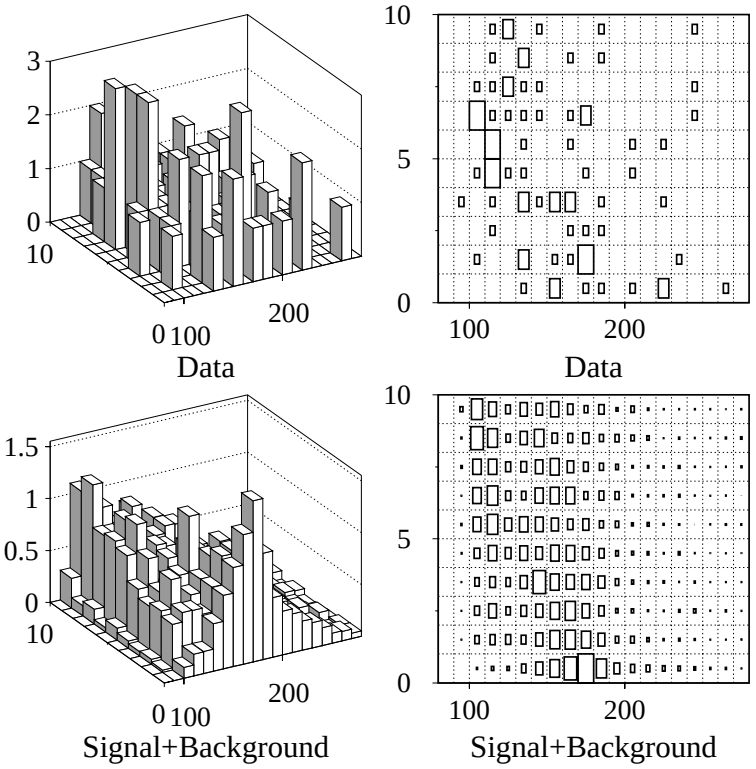

Figure 8.13: Distribution of candidate events (top plots) on NN plot. The bottom plots show the distribution for the best-fit signal plus background model of the data [76].

	NN fits		LB fits	
Input m_t	Mean $\widehat{m_t}$	R^{68}	Mean $\widehat{m_t}$	R^{68}
150	148.9	9.8	150.3	10.6
160	159.6	9.5	160.7	9.2
170	167.6	8.9	168.7	7.9
180	179.0	9.0	179.6	8.7
190	189.0	9.2	190.5	9.7

Table 8.16: Results of ensemble tests of the $\ell+$ jets maximum likelihood fit. Listed are the means and 68% widths of the $\widehat{m_t}$ distributions of 1000 ensembles fit using both the neural network and top likelihood discriminants [76].

175 GeV/c^2. For both the LB and NN methods of binning, the results have negligible bias and, just as important, reasonable pull distributions. As in the dilepton case, this indicates that the width of the $-\ln L$ curve provides a meaningful measure of the statistical uncertainty of the measurement. It is also clear that the sensitivity of the measurement is quite comparable for the two methods of binning.

The ensemble tests are also used to select an appropriate parameterization of the $-\ln L$ curve. In the case of the binned fit performed in the lepton plus jets analysis, one may analytically calculate the uncertainty in the value of $\ln L$ due to the finite Monte Carlo statistics. When fitting a curve to the $-\ln L$ points, the points are weighted according to the uncertainty due to signal Monte Carlo statistics. (The uncertainty arising from limited background Monte Carlo statistics is negligible due to the large point-to-point correlations.) As may be seen in Table 8.16 and Fig. 8.14 the nine-point quadratic fit is sensitive to the input top quark mass (the best-fit line to the distribution of mean $\widehat{m_t}$ as a function of m_t^{MC} has a slope of 0.98), and gives rise to a reasonable pull distribution. Therefore this fit method is adopted for the lepton plus jets analysis.

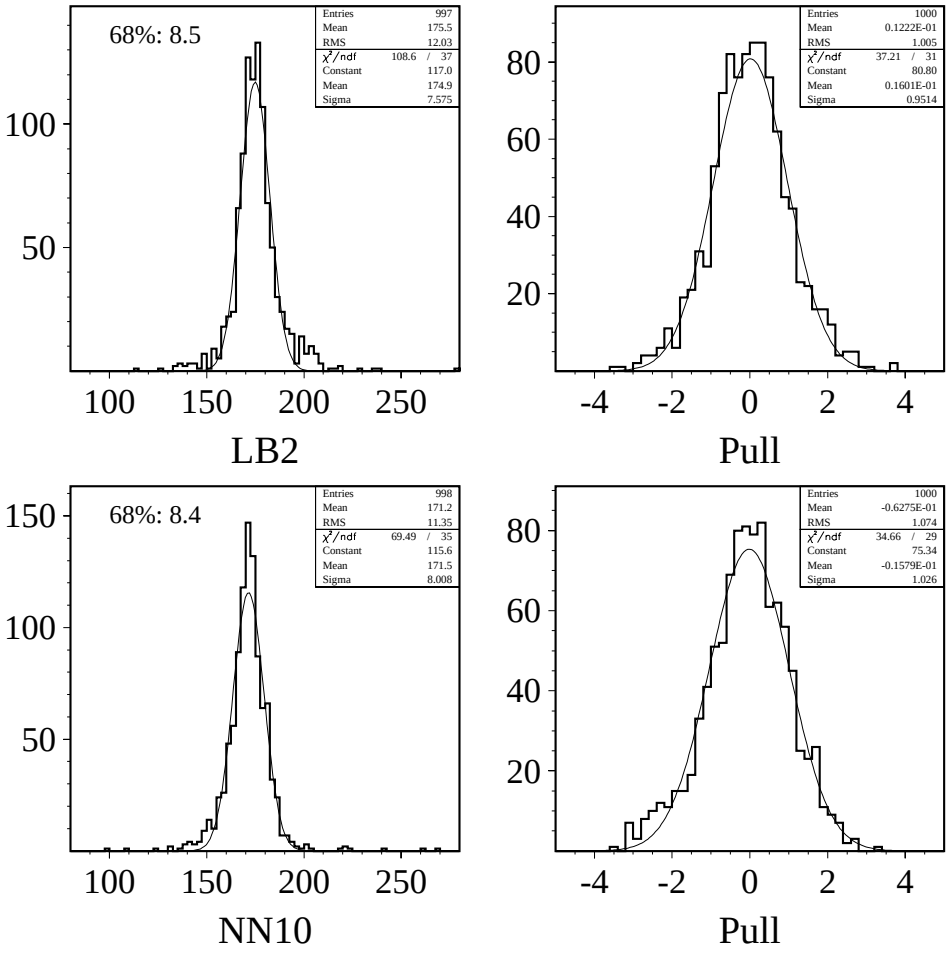

Figure 8.14: Results of ensemble tests for an input mass of 175 GeV/c^2 and assumed signal to background ratio of 1:2 [76].

8.2.1 Results

Having established that the methods for measuring the top quark mass in the lepton plus jets channel are self-consistent, we now apply them to the data sample at hand. The results for each binning method are summarized in Fig. 8.15. The figures show a one-dimensional projection of the data along the reconstructed mass axis, on which is superimposed the signal and background models with the best-fit normalization and top quark mass. Inset at upper right is the $-\ln L$ curve with its quadratic parameterization. Numerically, we find:

$$\begin{aligned} m_t &= 174.0 \pm 5.6 \text{ GeV/c}^2 \\ n_s &= 23.8^{+8.3}_{-7.8} \text{ events} \\ n_b &= 53.2^{+10.7}_{-9.3} \text{ events} \end{aligned} \tag{8.9}$$

for the LB binning and

$$\begin{aligned} m_t &= 171.3 \pm 6.0 \text{ GeV/c}^2 \\ n_s &= 28.8^{+8.4}_{-9.1} \text{ events} \\ n_b &= 48.2^{+11.4}_{-8.7} \text{ events} \end{aligned} \tag{8.10}$$

for the NN binning, where all uncertainties are statistical only. As both measurements are comparably sensitive, the final result is obtained my combining them. The correlation between the two results is determined by studying the subset of ensemble tests which have $\hat{\sigma}$ similar to that obtained from the data, and is found to be 0.88. When this is accounted for, the combined result is:

$$m_t = 173.3 \pm 5.6 \text{ GeV/c}^2 \tag{8.11}$$

There are a few checks which one would like to see the fits satisfy. Firstly, one would like to compare the number of signal events preferred by the fit to that expected given the $t\bar{t}$ cross section and detection efficiency measured by DØ. The cross section

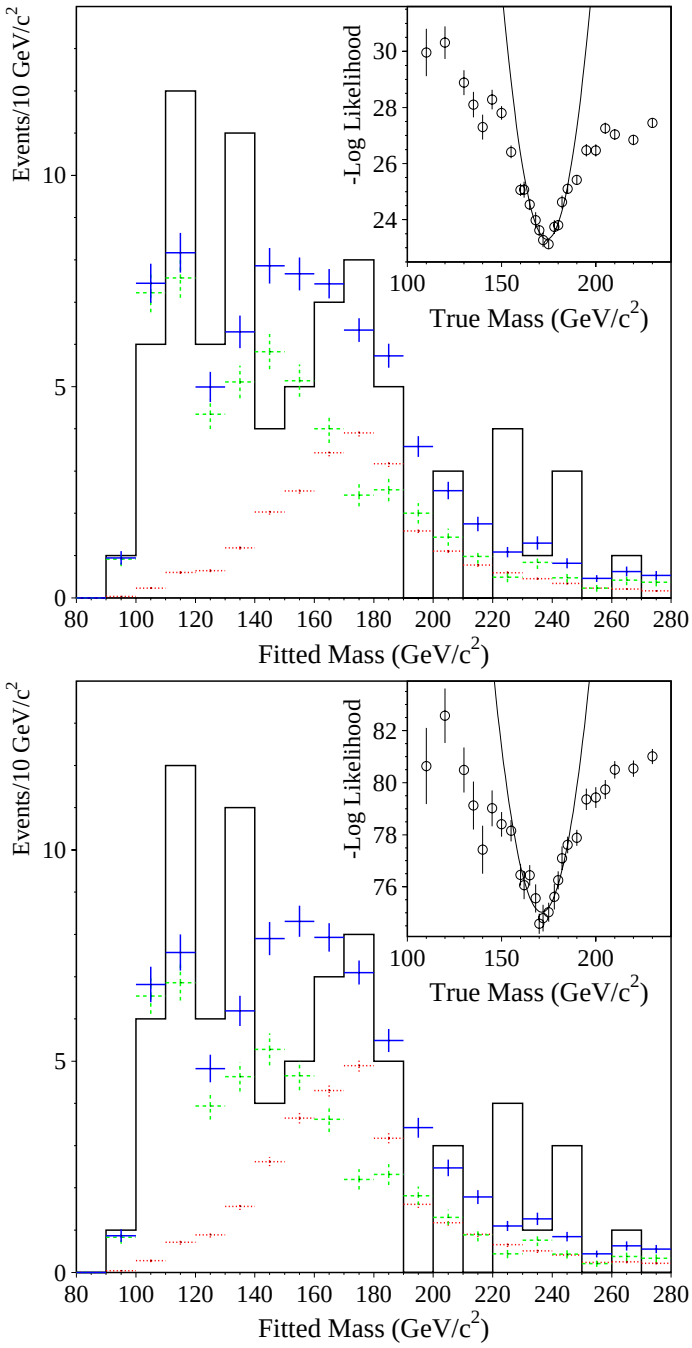

Figure 8.15: Results of the maximum likelihood fit to the data, for both the LB (top) and NN (bottom) binnings. In each plot the candidate distribution in reconstructed mass is shown, along with the best-fit model for signal (dotted), background (dashed) and their sum. The $-\ln L$ curve is inset at upper right [76].

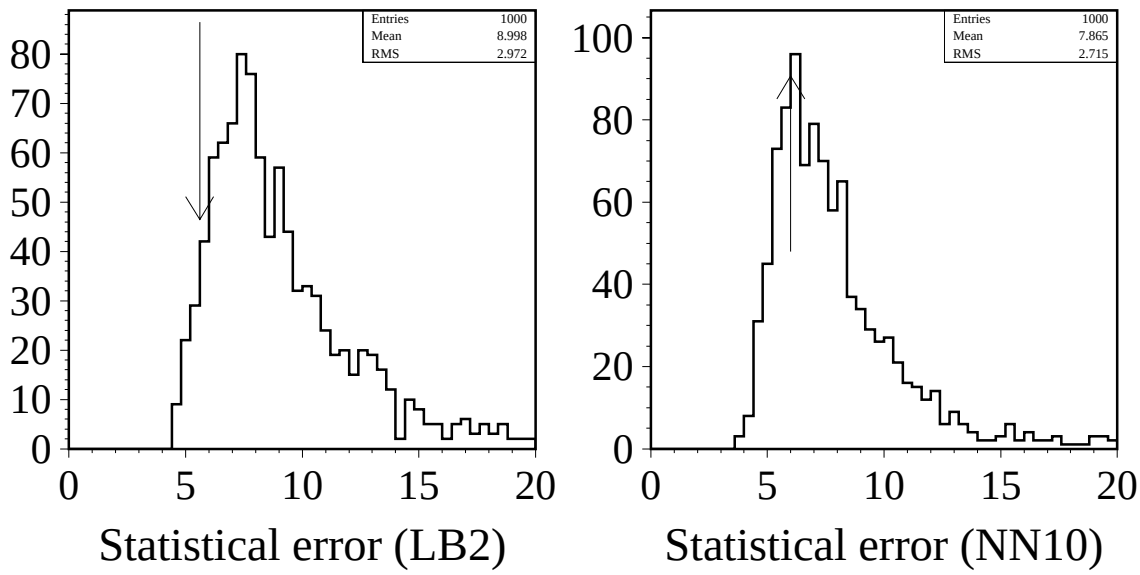

Figure 8.16: Distributions of $\hat{\sigma}$ from lepton plus jets ensemble tests. Arrows indicate the values obtained by fitting the DØ data [76].

measured at 170 GeV/c^2 is 6.0 ± 1.8 pb, which translates into 30.0 events expected in the mass sample, in agreement with the number fit in either analysis.

Secondly, one may note that the widths of the $-\ln L$ curves are narrower than expected from ensemble tests. In order to test how likely such a narrow distribution is, one may plot the distribution of $-\ln L$ widths obtained in the ensemble tests. As shown in Fig. 8.16, the width observed in our experiment is smaller than the most likely ensemble value. The probability of observing a width this small or smaller is 6.0% for the LB binning and 24% for the NN binning.

Finally, while Fig. 8.15 shows the results of the fit, it does not provide clear evidence of a peak in the mass distribution. To this end, one may look at those events that pass the LB selection. As shown in Fig. 8.17, these events do cluster near 175 GeV/c^2. (In viewing these distributions, it is important to recall that background events that satisfy the LB cuts do not peak in this region.)

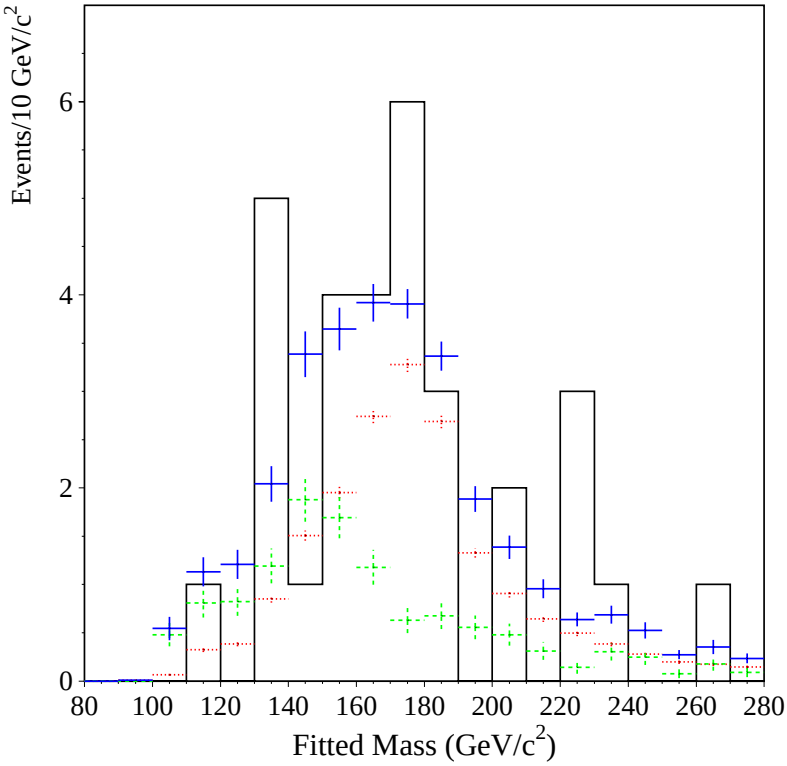

Figure 8.17: Reconstructed mass distribution for events passing the LB selection [76].

	NN	LB	Average
Jet scale	3.8	4.2	4.0
Signal MC generator	3.2	3.3	3.3
Background MC generator	2.6	2.5	2.6
Multiple Interactions	1.2	1.2	1.2
Likelihood fit	1.5	1.2	1.4
LB/NN difference	1.4	1.4	1.4
Total	6.1	6.3	6.2

Table 8.17: Systematic uncertainties in the measurement of the top quark mass in the lepton plus jets channels [76].

8.2.2 Systematic Uncertainties

Evaluation of systematic uncertainties in the lepton plus jets analysis proceeds in much the same manner as for the dilepton analysis. The parameter under study is varied in the Monte Carlo ensembles, and the resultant shift in the $\widehat{m_t}$ distribution noted. The sources of error, and the variation in $\widehat{m_t}$ that results from each, are listed in Table 8.17.

As in the dilepton case, the systematic uncertainty arising from changing the signal Monte Carlo from HERWIG to ISAJET is noted. One also has an uncertainty due to changing the parameters of the VECBOS model for W plus jets background events. The model was varied both in the dynamical scale of the generation and in the Monte Carlo used for fragmentation (again, either HERWIG or ISAJET).

Finally, one half of the difference between the NN and LB results is taken as a systematic uncertainty. Then the total systematic uncertainty is found to be 6.3 GeV/c^2 for the LB analysis and 6.1 GeV/c^2 for the NN. Averaging these, we have the result:

$$m_t = 173.3 \pm 5.6 \text{ (stat.)} \pm 6.2 \text{ (syst.) GeV/c}^2 \tag{8.12}$$

8.3 Combined Top Quark Mass Result

The final step in the analysis is to combine the dilepton and lepton plus jets mass results to yield the best possible measurement of the top quark mass. (The six-event dilepton analysis is taken.) Since the information from each channel is summarized in its $-\ln L$ curve, it is straightforward to combine the results by simply summing their $-\ln L$ points. (A slight complication arises due to the fact that for $m_t = 162$, 168, 172, 178, and 182 GeV/c^2 Monte Carlo samples were only generated for lepton plus jets events. The dilepton values taken for these points are the interpolated $-\ln L$ values from the two neighboring mass points.) The result of this combination is shown in Fig. 8.18 for both of the lepton plus jets analyses.

8.3.1 Parameterization of the $-\ln L$ Curve

Again we face the issue of parameterizing the $-\ln L$ curve in order to assign a central value and error. Since the curve for each individual channel is modeled well by a quadratic, it stands to reason that this will also be true for their combination. In addition, since the combined likelihood curve is not substantially narrower than that for the lepton plus jets events alone, one would expect nine points to be a reasonable range over which to fit. Table 8.18 shows the widths of the pull distributions obtained for fits to five, seven, and nine points for each version of the lepton plus jets analysis, which confirm this expectation.

The results are then:

$$m_t = 173.3 \pm 5.2 \text{ GeV/c}^2 \tag{8.13}$$

for the combination with the LB analysis and

$$m_t = 171.2 \pm 5.7 \text{ GeV/c}^2 \tag{8.14}$$

for combination with the NN analysis, where the errors are statistical only.

One may note in Fig. 8.18 that there is a significant scatter of the $-\ln L$ points about the fit curve. In addition, it happens that the combination of the dilepton

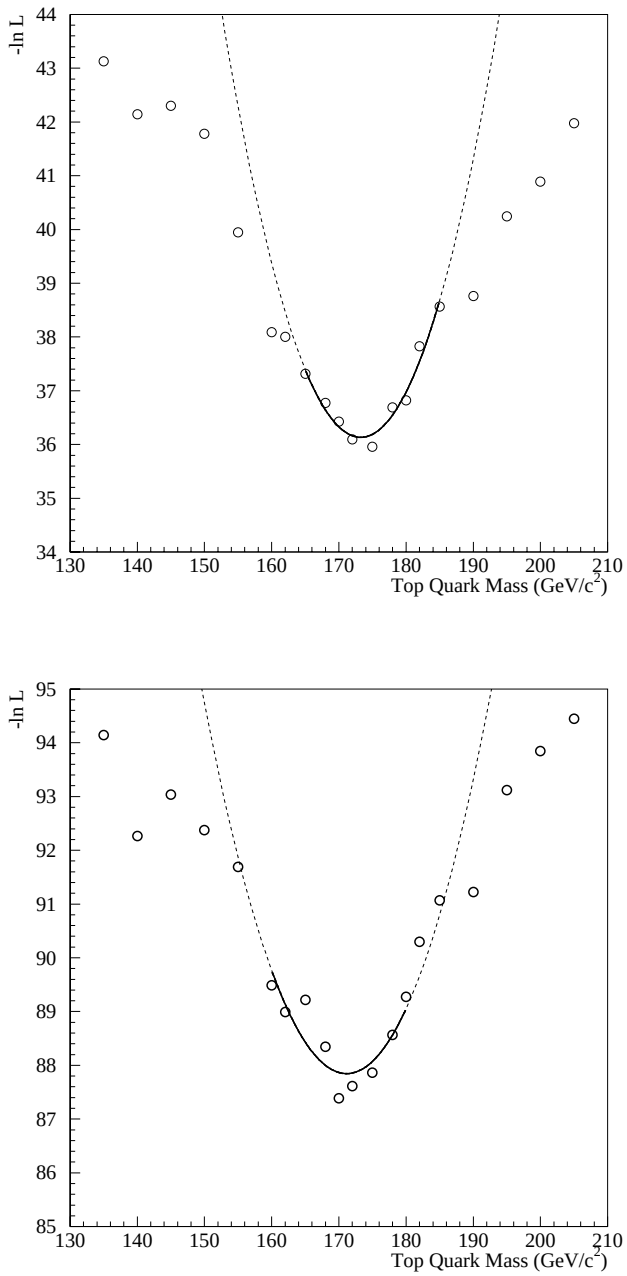

Figure 8.18: Result of multiplying the likelihood curve from the dilepton analysis with that from the lepton plus jets LB (top) and NN (bottom) analyses.

	LB			NN		
m_t^{MC}	5 points	7 points	9 points	5 points	7 points	9 points
150	1.08	0.93	0.87	0.93	0.88	0.88
155	1.13	0.96	0.90	1.03	0.97	0.91
160	1.20	1.02	0.92	1.30	1.10	0.98
165	1.25	1.05	0.98	1.52	1.18	1.02
170	1.17	0.95	0.90	1.33	1.01	0.93
175	1.19	1.01	0.97	1.49	1.10	1.03
180	1.27	1.05	1.00	1.43	1.14	1.05
185	1.22	1.06	0.93	1.26	1.15	0.96
190	1.06	1.00	1.01	0.92	0.96	0.96

Table 8.18: Pull widths for different parameterizations of the combined $-\ln L$ curve.

and NN analyses gives a broader $-\ln L$ curve than that attained for the NN analysis alone. It is therefore imperative to study the stability of the result when different parameterizations are attempted. Some of the possible variations are listed in Table 8.19.

8.3.2 Systematic Uncertainties

To evaluate the systematic uncertainties in the combined measurement, we begin with the formula for combining any two measurements x_1 and x_2 [8]:

$$x_{\mathrm{avg}} = \frac{\omega_1 x_1 + \omega_2 x_2}{\omega_1 + \omega_2}$$

where $\omega_i \equiv 1/\sigma_i^2$. The σ_i that enter into this formula are the total uncorrelated errors (statistical and systematic) in the two measurements. As the uncorrelated systematics in both the lepton plus jets and dilepton analyses are small compared to the statistical uncertainties, we approximate the σ_i by the statistical uncertainties only. Then, if the ith source of systematic uncertainty contributes δx_1 and δx_2 to

$\ell+$ jets fit	Fit type	Result
LB	5-pt. Quadratic	173.8 ± 4.7
	7-pt. Quadratic	173.7 ± 4.5
	9-pt. Quadratic	$\mathbf{173.2 \pm 5.2}$
	11-pt. Quadratic	173.3 ± 6.9
	9-pt. Cubic	$173.3^{+5.3}_{-5.2}$
NN	5-pt. Quadratic	172.0 ± 3.7
	7-pt. Quadratic	172.0 ± 5.5
	9-pt. Quadratic	$\mathbf{171.1 \pm 5.7}$
	11-pt. Quadratic	170.6 ± 5.6
	9-pt. Cubic	$171.3^{+5.7}_{-5.6}$

Table 8.19: Results of several parameterizations of the combined $-\ln L$ curves.

each individual measurement, δx_{avg} is given by:

$$\delta x_{\text{avg}} = \frac{1}{\omega_1 + \omega_2} \left[(\omega_1 \delta x_1)^2 + (\omega_2 \delta x_2)^2 + 2C_i \omega_1 \omega_2 \delta x_1 \delta x_2 \right]^{\frac{1}{2}} \tag{8.15}$$

where C_i is the correlation between the change in each measurement when this systematic effect is varied. For simplicity, it is assumed that all correlated systematics have $C_i = 1$.

For the combination of the dilepton and lepton plus jets top quark mass results, the correlated sources of systematic uncertainty arise from the jet energy scale, multiple interactions, and $t\bar{t}$ Monte Carlo model. The uncertainty due to Monte Carlo statistics is uncorrelated, as is that due to the Monte Carlo background model (due to the fact that different background generators are used in each analysis). Taking as input the systematic uncertainties listed in Tables 8.6 and 8.17, as well as the assumed correlations, one obtains the systematic uncertainties listed in Table 8.20.

With all of the systematic uncertainties evaluated (one-half of the difference between the LB and NN results is taken as an additional systematic), we may proceed to the combined result. As for the lepton plus jets analysis, a combination of the LB and NN results is performed. In this case, if we consider that subset of ensembles

	NN	LB	Average
Jet scale	3.7	4.0	3.9
Signal MC generator	2.9	3.0	3.0
Background MC generator	2.2	2.2	2.2
Multiple Interactions	1.2	1.2	1.2
Likelihood fit	1.3	1.1	1.2
LB-NN difference	1.1	1.1	1.1
Total	5.6	5.8	5.7

Table 8.20: Systematic uncertainties in the combined measurement of the top quark mass.

for which the the LB combination has $4.7 < \hat{\sigma} < 5.7$ and the NN combination has $5.2 < \hat{\sigma} < 6.2$, the correlation coefficient is found to be 0.89. Hence, the final result is:

$$m_t = 173.1 \pm 5.2 \text{ (stat.)} \pm 5.7 \text{ (syst.)} \text{ GeV/c}^2 \tag{8.16}$$

Chapter 9

Conclusions

It is interesting to note that the mass of the top is already known more precisely than that of any other quark, even though only about one hundred examples of its production have been observed. With this measurement in hand we can explore some of the questions raised in Chapter 2.

9.1 Production Model

Firstly, we can use the measured mass and production cross section to test the perturbative QCD calculation of $t\bar{t}$ production. The result is shown in Fig. 9.1, in which it is seen that the measurements are in very good agreement with the predictions.

9.2 Constraint on the Higgs Mass

Next we may superpose the measured top quark mass value on Fig. 2.3, to obtain Fig. 9.2. It is seen that the minimal Standard Model Higgs boson is quite compatible with the observed W and top masses, with the agreement being most favorable for a relatively light Higgs boson. However, improved accuracy in both the m_t and m_W measurements will be needed to make a meaningful prediction of m_H.

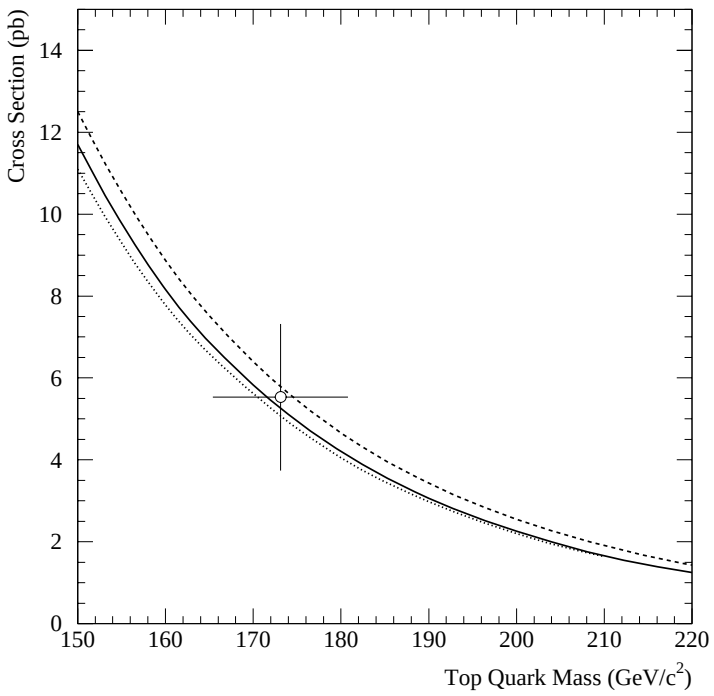

Figure 9.1: Comparison of the measured top quark mass and cross section with the predictions of perturbative QCD.

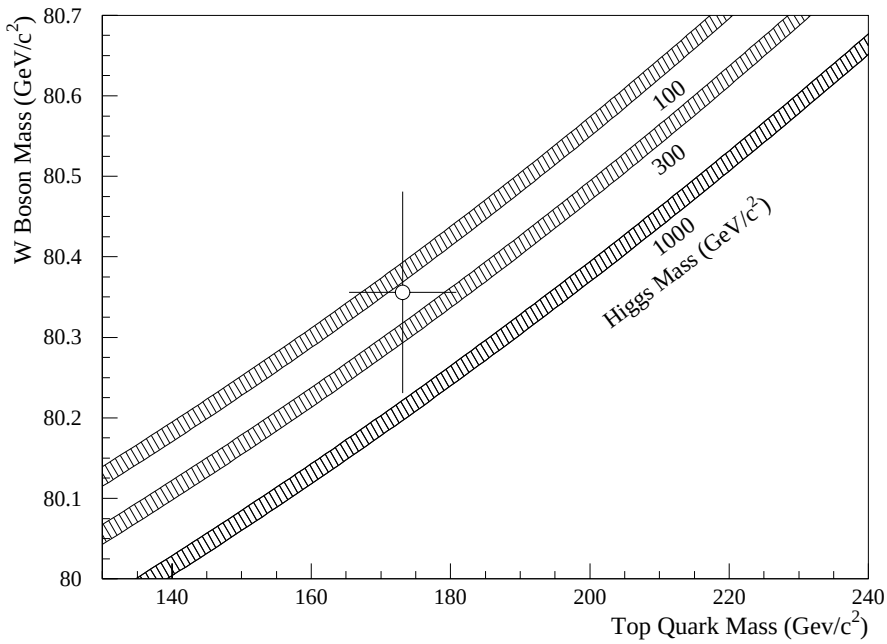

Figure 9.2: Curves showing the minimal Standard Model Higgs boson mass as a function of the W boson and top quark masses. The cross represents the world average measurement of m_W and the DØ measurement of m_t.

9.3 The Future of Top Quark Physics

The next chance for such an improvement will occur in the second run of the TeVatron collider, planned to begin in 1999. For this run, the TeVatron will be upgraded to deliver an instantaneous luminosity an order of magnitude larger than for the 1992-1996 run, and to have $\sqrt{s} = 2.0$ TeV. These improvements will result in the production of about fifteen times as many $t\bar{t}$ events as in the previous run. Furthermore, both the DØ and CDF detectors will undergo upgrades which should enhance their efficiency for observing $t\bar{t}$ pairs.

With this large sample of $t\bar{t}$ events, the measurement of the top quark mass will be limited by systematic uncertainties, particularly in the understanding of the hadronic energy scale of the detector and the model of parton fragmentation. While the additional statistics will also aid the understanding of these effects (for example allowing the possibility of directly excluding some models for $t\bar{t}$ production), the vagaries of non-perturbative QCD (in particular as they relate to the understanding of the jet energy scale) will present a major difficulty in the measurement.

It is important to note, therefore, that the measurement using dilepton events is less sensitive to these effects than that using lepton plus jets events. This implies that the methods for dilepton mass analysis set forth in this dissertation (or refinements upon them) may in the not so distant future be used to achieve the most precise measurement of the top quark mass.

Appendix A

Monte Carlo Generation

The analyses described in this dissertation depend on Monte Carlo simulation of both signal and background events, and so care must be taken to ensure that the models are as realistic as possible. This appendix gives details on the samples generated, and the choices made in generating them.

A.1 Signal Events

The primary model for $t\bar{t}$ events was HERWIG version 5.7. The user is given a choice of parton distribution function, and the CTEQ3M [88] set was chosen. In generating these events the samples were kept as unbiased as possible in order that selection efficiencies and mass spectra calculated from them be accurate. Therefore, no kinematic cuts were applied during the generation.

On the other hand, a completely unbiased sample would have over 40% of its events in the all-hadronic final state. As this channel is not used here, one would like to restrict the W decays to avoid spending the resources needed to process these events. In addition, in order that the samples of single-lepton and dilepton final state events be of similar size, it is necessary that the W decays be adjusted separately for each.

For the lepton plus jets channels, the branching ratio for the W^+ was kept at its physical value while the W^- was forced to decay to one of the three lepton families. Events with no final-state electrons or muons were vetoed as these have no chance of entering into the analysis, and one half of the dilepton events were discarded so that the ratio of lepton plus jets to dilepton events would be the same as that for an unbiased sample. This selection avoids unnecessary processing while including every possible source of events in the lepton plus jets channels.

The procedure was quite similar for the dilepton samples, with both Ws being forced to decay to leptons, and events being vetoed if they did not contain at least two electrons or muons in the final state.

All of the samples were passed through MUSMEAR to accurately model the resolution of the muon momentum measurement. In order to simulate the significant changes in the conditions of the muon chambers over the course of the run, half the files were processed with the version of MUSMEAR which modeled the chambers before they were cleaned, and half with the version appropriate to the cleaned chambers.

Tables A.1 lists the top quark mass values for which samples were generated, and the number of events in each sample.

A.2 Background Samples

A.2.1 Lepton Plus Jets Analysis

The W plus multijet background in the lepton plus jets channel was simulated using VECBOS. In order to generate events under the same conditions as the signal Monte Carlo, VECBOS was modified to interface with PDFLIB, and events were also generated with the CTEQ3M distribution functions. One must also specify the dynamical scale at which VECBOS is to carry out its calculations. Samples were generated at two of the possible choices (the average jet p_T and the W mass) to allow estimation of the systematic uncertainty arising from the fact that VECBOS calculates

Top Quark Mass	Number of Events ($\times$1000)
110	50
120	50
130	50
135	50
140	50
145	50
150	100
155	100
160	100
162	50*
165	100
168	50*
170	100
172	50*
175	100
178	50*
180	100
182	50*
185	100
190	100
195	50
200	50
205	50
210	50
220	50
230	50

*generated for ℓ+jets only.

Table A.1: Summary of HERWIG $t\bar{t}$ samples used in the analysis. Samples were generated separately for the dilepton and lepton plus jets channels.

only to leading order.

To attain reasonable efficiency in the generation, kinematic selection was applied at the generator level. All partons and leptons were required to have p_T above 10 GeV/c, and to be within $|\eta| < 3.0$.

Some complication in interpreting the output of VECBOS arises due to the fact that this generator was designed primarily to calculate the cross section of W plus jets events, not model their kinematics. Therefore, events are generated roughly uniformly across phase space, with a weight assigned to each which represents the differential cross section into its particular region of phase space.

Ideally, one would like to perform an importance sampling on this sample to select a set of events distributed naturally in phase space. However, the wide range of event weights assigned by VECBOS (typically seven orders of magnitude) renders this procedure prohibitively inefficient. Therefore, a compromise procedure is adopted in which an arbitrary cutoff weight is defined, and all events with weight greater than this value pass into the sample. An importance sampling is then done for events with weights below this cutoff.

The cutoff is chosen by fixing the percentage of events that one desires to automatically pass. For W plus four jet events, this is taken to be 40%, and for W plus three jet events it is 30% (reflecting the fact that the CPU time involved in generating the events increases rapidly as a function of the jet multiplicity). The only justification for such a procedure is that the kinematics of the resultant sample is similar to that of the original sample when events are given their assigned weight. More details on the implementation of VECBOS at DØ may be found in [89].

After the VECBOS events have been generated, one must pass them through a separate Monte Carlo which adds gluon radiation and models the fragmentation of the partons into final-state hadrons. Both ISAJET and HERWIG have been modified to accept VECBOS input rather than calculating the hard subprocess themselves.

Since HERWIG requires information about the color content of its input partons, VECBOS was also modified to assign color and flavor labels to its partons [90]. As

Number of jets	Dynamical scale	Number of events ($\times 1000$)	Fragmentation
1^*	$\langle$Jet $p_T\rangle$	1019	HERWIG
2	$\langle$Jet $p_T\rangle$	118	ISAJET
2	$\langle$Jet $p_T\rangle$	117	HERWIG
2	M_W	113	HERWIG
2	M_W	109	ISAJET
2^*	$\langle$Jet $p_T\rangle$	234	HERWIG
3	$\langle$Jet $p_T\rangle$	72	HERWIG
3	$\langle$Jet $p_T\rangle$	86	ISAJET
3	M_W	82	HERWIG
3	M_W	116	ISAJET
3^*	$\langle$Jet $p_T\rangle$	109	HERWIG
4	$\langle$Jet $p_T\rangle$	74	HERWIG
4	$\langle$Jet $p_T\rangle$	98	ISAJET
4	M_W	545	ISAJET
4^*	$\langle$Jet $p_T\rangle$	75	HERWIG

Table A.2: Summary of VECBOS samples used in the analyses. Those marked with an asterisk represent a modified version of VECBOS which extended the range of the W Breit-Wigner distribution. This distinction is of crucial importance to the $e\nu$ channel, but nearly irrelevant for the other channels.

VECBOS steps through the final-state phase space in its calculation of the cross section, it calculates the contribution of all possible diagrams. For a particular point in the phase space, then, the flavor of the partons is chosen probabilistically based on the relative contribution from each diagram. Once the flavors have been assigned, the colors are assigned randomly.

Table A.2 lists the samples of VECBOS which were generated for use in the cross section and mass analyses.

A.2.2 Dilepton Analysis

The samples used to model the various sources of background in the dilepton channels were generated by ISAJET, and PYTHIA. The generator used for each sample

Process	Generator	Kinematic Cuts
$Z \to \tau\tau$	ISAJET	$p_T(Z) > 20$ GeV
WW	ISAJET	None
$Z \to \mu\mu$	PYTHIA	$H_T > 95$ GeV

Table A.3: Summary of background Monte Carlo samples used in the dilepton mass analysis.

is listed in A.3, as are any kinematic selection applied at the generator level.

All background samples, both for the lepton plus jets and dilepton analyses, were passed through the same MUSMEAR procedure as applied to the signal events.

Appendix B

Kinematics of the Dilepton Sample

This appendix lists the kinematics of the final-state objects in each of the dilepton $t\bar{t}$ candidate events. Listed are the four-vectors as well as the commonly-used quantities E_T, η and ϕ which can be calculated from them (E_T is defined as $\sqrt{E_x^2 + E_y^2}$ for all objects). For the jet energies and $\not\!\!E_T$, two sets of values are listed. The first represents the energy after the CAFIX corrections, and the second (in parentheses) that after the post-CAFIX corrections. It is the first set that is used in the event selection, and the second that enters the mass analysis. All energies are in GeV, and angles are in radians.

The second section of the appendix then compares the distribution of the kinematic quantities to that predicted by the signal plus background models.

B.1 Object Resolutions

This section provides parameterizations of the measurement resolutions for the objects listed below. In most cases, the energy resolutions quoted are those used in the dilepton mass analysis, while the direction resolutions are taken from the lepton

plus jets mass analysis [76]. An exception is made for jets, for which all resolutions are taken from the lepton plus jets analysis.

- Electrons:

$$\begin{aligned} \frac{\sigma(E)}{E} &= \frac{0.15}{\sqrt{E}} \oplus 0.03 \\ \sigma(\phi) &= 0.0060 \\ \sigma(\eta) &= 0.0071 \end{aligned}$$

- Muons:

$$\begin{aligned} \sigma\left(\frac{1}{p}\right) &= 0.18(p - 2\mathrm{GeV/c})/p^2 \oplus 0.003 \\ \sigma(\phi) &= 0.003 \\ \sigma(\eta) &= 0.0001 \end{aligned}$$

- Jets:

$$\begin{aligned} \frac{\sigma(E)}{E} &= \frac{1.15}{\sqrt{E}} \oplus 0.04 \text{ for } |\eta_d| < 0.8 \\ &= \frac{1.26}{\sqrt{E}} \oplus 0.08 \text{ for } 0.8 < |\eta_d| < 1.4 \\ &= \frac{1.31}{\sqrt{E}} \oplus 0.05 \text{ for } |\eta_d| > 1.4 \\ \sigma(\phi) &= \sigma(\eta) = 0.04 \text{ for } |\eta_d| < 0.8 \\ \sigma(\phi) &= \sigma(\eta) = 0.05 \text{ for } |\eta_d| > 0.8 \end{aligned}$$

- Unclustered transverse energy: $\sigma(E_x) = \sigma(E_y) = 4$ GeV

B.2 Tables of Dilepton Kinematics

B.2.1 $e\mu$ Events

Run 58796 Event 417					z vertex: -1.8 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Electron	106.8	12.3	-97.8	41.1	98.6	0.41	4.84
Muon	296.6	-68.3	272.5	95.1	280.0	0.33	1.82
$\not\!\!E_T$	–	100.5	-152.7	–	182.9	–	5.29
Jet 1	33.2	-24.3	-9.5	-19.9	26.1	-0.70	3.51
	(34.7)	(-25.5)	(-9.9)	(-20.8)	(27.3)		
Jet 2	38.1	-13.3	-19.0	29.8	23.2	1.07	4.10
	(41.2)	(-14.4)	(-20.5)	(32.3)	(25.1)		

Run 84676 Event 12814					z vertex: -6.17 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Electron	81.3	-75.4	-1.1	-30.2	74.5	-0.39	3.16
Muon	30.2	-25.2	10.6	-12.8	27.4	-0.45	2.75
$\not\!\!E_T$	–	62.0	5.2	–	62.3	–	0.08
Jet 1	93.8	38.0	-83.7	-15.6	91.9	-0.17	5.14
	(95.9)	(38.9)	(-85.6)	(-16.0)	(94.0)		
Jet 2	37.8	13.9	32.3	-11.2	35.2	-0.31	1.17
	(38.8)	(14.2)	(33.1)	(-11.4)	(36.0)		
Jet 3	31.4	-1.6	28.6	11.6	28.7	0.39	1.63
	(32.2)	(-1.6)	(29.3)	(11.9)	(29.4)		

Run 90422 Event 26920					z vertex: 17.0 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Electron	148.5	-44.7	20.2	140.1	49.1	1.77	2.72
Muon	18.4	5.4	17.2	-3.3	18.1	-0.18	1.27
$\not\!\!E_T$	–	-12.5	4.5	–	13.2	–	2.79
Jet 1	50.2	38.7	-29.2	11.0	48.4	0.22	5.64
	(51.4)	(39.6)	(-29.9)	(11.3)	(49.7)		
Jet 2	39.9	18.8	-18.4	-29.6	26.4	-0.97	5.51
	(41.8)	(19.8)	(-19.4)	(-31.0)	(27.7)		

B.2.2 Dielectron Events

Run 88295 Event 30317					z vertex: 7.6 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Electron 1	53.3	2.7	50.4	17.1	50.5	0.33	1.52
Electron 2	52.7	-7.4	21.4	-47.6	22.6	-1.49	1.91
$\not{E}_T$	–	41.3	-4.0	–	41.5	–	6.19
Jet 1	58.1	-28.1	-35.5	-35.6	45.3	-0.72	4.04
	(60.4)	(-29.2)	(-36.9)	(-37.0)	(47.1)		
Jet 2	38.5	3.4	-26.0	-27.6	26.2	-0.92	4.84
	(40.2)	(3.5)	(-27.1)	(-28.9)	(27.4)		

Run 95653 Event 10822					z vertex: 31.2 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Electron 1	62.7	52.3	-4.1	-34.4	52.5	-0.62	6.20
Electron 2	38.9	-8.5	-26.6	27.0	27.9	0.86	4.40
Tagging muon	16.6	-13.3	-4.5	-8.9	14.0	-0.60	3.47
$\not{E}_T$	–	42.6	-11.3	–	44.1	–	6.02
Jet 1*	82.7	-66.7	-18.8	-44.4	70.9	-0.60	3.41
	(114.1)	(-92.4)	(-26.0)	(-61.6)	(96.0)		
Jet 2	47.5	-22.7	24.5	-32.8	33.4	-0.87	2.32
	(49.2)	(-23.5)	(25.3)	(-34.0)	(34.6)		
Jet 3	32.4	0.0	26.6	17.6	26.6	0.62	1.57
	(33.7)	(0.0)	(27.7)	(18.3)	(27.7)		

*denotes the jet tagged by the muon. In this case, the post-CAFIX corrections include the addition of the muon and neutrino energy back into the jet.

B.2.3 Dimuon Event

Run 84395 Event 15530					z vertex: 5.9 cm		
Object	E	E_x	E_y	E_z	E_T	η	ϕ
Muon 1	68.6	-63.9	12.7	-21.4	65.1	-0.32	2.94
Muon 2	34.9	-16.0	31.0	1.9	34.9	0.05	2.05
$\not\!\!E_T$	–	71.2	53.2	–	88.9	–	0.64
Jet 1	146.1	32.1	-98.2	-102.4	103.3	-0.88	5.03
	(153.5)	(33.8)	(-103.1)	(-107.6)	(108.5)		
Jet 2	35.1	-8.6	21.4	26.2	23.1	0.97	1.95
	(37.2)	(-9.1)	(22.7)	(27.7)	(24.5)		
Jet 3	47.1	-7.6	-16.8	43.0	18.4	1.58	4.29
	(52.3)	(-8.4)	(-18.6)	(47.8)	(20.5)		

B.3 Comparison of Dilepton Kinematics

In comparing the kinematics of the dilepton sample to what we would expect based on our models of the signal and background, one must decide which definition of the jet energy is to be used. For the plots that are shown here, all jets are corrected by CAFIX and the data jets have the additional post-CAFIX corrections described in Chapter 4. Muon-tagged jets in the data and Monte Carlo samples are not corrected for the muon and neutrino energies. The comparison of the transverse energies of the final-state objects is shown in Fig. B.1.

Shown on the plots is the Kolmogorov-Smirnov probability that the observed distributions are consistent with the signal plus background model. There is good agreement in general, with the exception of the second leading jet distribution. For this variable, the data are softer than the signal or background would predict. However, as observing a 3% fluctuation when looking at five variables is not uncommon, there is no reason to conclude that the second leading jet E_T distribution rules out the model.

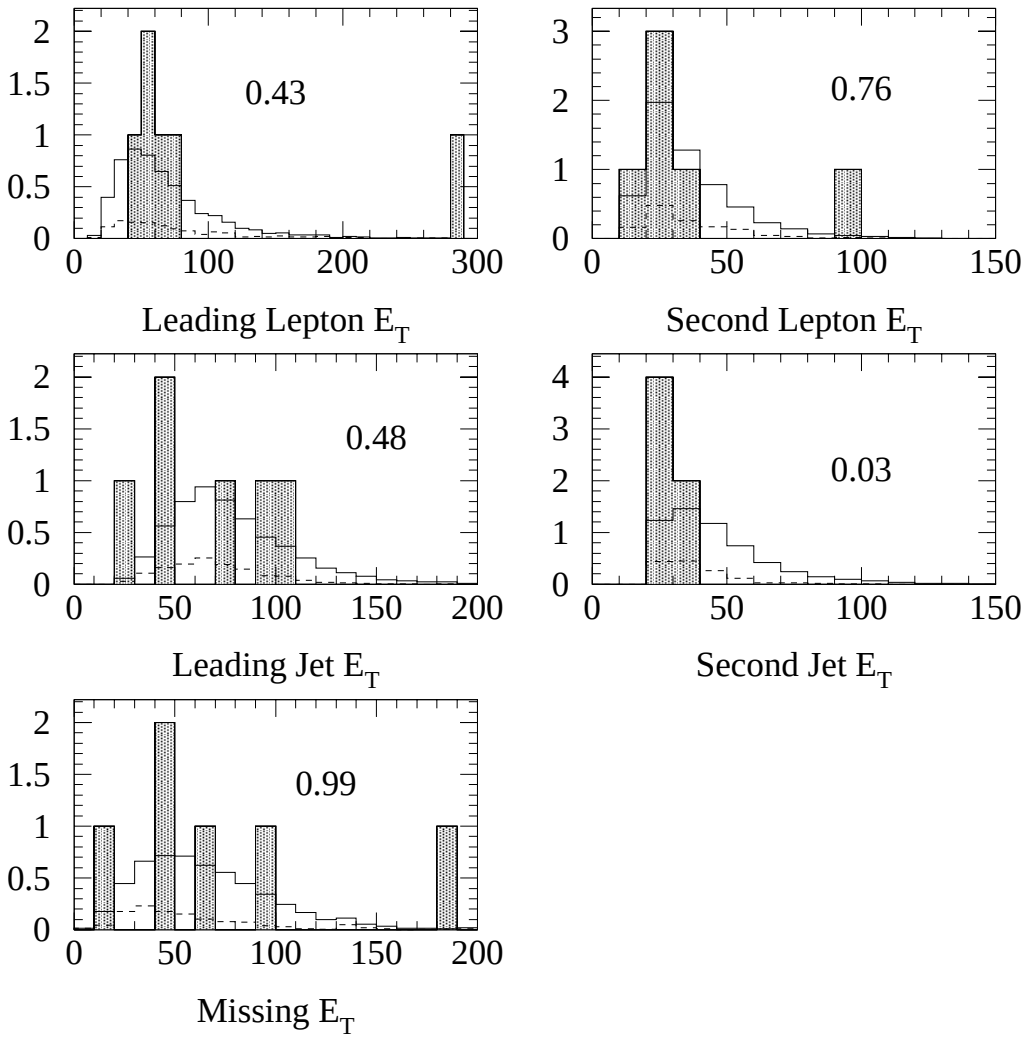

Figure B.1: Comparison of observed E_Ts (shaded histograms) to the expectation from the signal + background (solid) and background-only (dashed) models. The signal plus background histograms are normalized to six events, and the background-only histograms are normalized to the expected background level. The numbers on the plots are the Kolmogorov-Smirnov probabilities that the data sample was drawn from the model.

Bibliography

[1] S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995)

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

[3] G. 't Hooft, Nucl. Phys. **B33**, 173 (1971)

[4] Gordon L. Kane, *Modern Elementary Particle Physics: The Fundamental Particles and Forces?*, Addison-Wesley (1993)

[5] Michio Kaku, *Quantum Field Theory: A Modern Introduction*, Oxford University Press (1993)

[6] S. Pokorski, *Gauge Field Theories*, Cambridge University Press (1987)

[7] Chris Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Benjamin/Cummings (1983)

[8] Particle Data Group, Phys. Rev. **D54**, 1 (1996)

[9] A.P. Heinson, A.S. Belyaev, and E.E. Boos, UCR/95-17 (1995)

[10] E. Laenen *et al.*, Phys. Lett. **B321**, 254 (1994)

[11] E.L. Berger and H. Contopanagos, Phys. Lett. **B361**, 115 (1995)

[12] S. Catani *et al.*, Phys. Lett. **B378**, 329 (1996)

[13] C.T. Hill and S.J. Parke, Phys. Rev. **D49**, 4454 (1994)

[14] S. Abachi *et al.*, Phys. Rev. **D52**, 4877 (1995)

[15] S. Dawson, Private communication.

[16] H. Arason *et al.*, Phys. Rev. **D46**, 3945 (1992)

[17] M. Sher, Phys. Lett. **B317**, 159 (1993)

[18] D. Bardin *et al.*, Z. Phys. **C44**, 493 (1989); Comp. Phys. Comm. **59**, 303 (1990); Nucl. Phys. **B351**, 1 (1991); Phys. Lett. **B255**, 290 (1991) and CERN-TH 6443/92 (May 1992)

[19] M. Demarteau, Proc. 9th Meeting of the American Phyiscal Society Division of Particles and Fields, Minneapolis, MN (1996)

[20] M. Gallinaro, Proc. of QFTHEP, St. Petersburg, Russia (1996)

[21] J. Thompson, FERMILAB-TM-1909 (1994)

[22] S. Abachi *et al.*, (DØ). Nucl. Instr. Meth. **A338**, 185 (1994)

[23] F. Sauli, Principles of operation of multiwire proportional and drift chambers. In Thomas Ferbel, editor, *Experimental Tecnniques in High-Energy Physics*, pages 79-188. Addison-Wesley (1987)

[24] Konrad Kleinknecht, *Detectors for Particle Radiation*, Cambridge University Press (1986)

[25] Richard C. Fernow, *Introduction to Experimental Particle Physics*, Cambridge University Press (1986)

[26] A.R. Clark *et al.*, Nucl. Instr. Meth. **A279**, 243 (1989)

[27] A.R. Clark *et al.*, Nucl. Instr. Meth. **A315**, 193 (1992)

[28] A.R. Clark *et al.*, Nucl. Instr. Meth. **A261**, 420 (1987)

[29] A.R. Clark, private communication.

[30] J.D. Jackson, *Classical Electrodynamics*, 2nd. Ed., John Wiley & Sons (1975)

[31] B. Mansoulie *et al.*, Nucl. Instr. Meth. **A265**, 157 (1988)

[32] J-F. Detœuf *et al.*, Nucl. Instr. Meth. **A279**, 310 (1989)

[33] Ties Behnke, Ph. D. Thesis, State University of New York at Stony Brook, August 1989

[34] C. Fabjan, Calorimetry in High-Energy Physics. In Thomas Ferbel, editor, *Experimental Tecnniques in High-Energy Physics*, pages 79-188. Addison-Wesley (1987)

[35] S. Chopra and R. Raja, DØ internal report No. 2098, April 1994 (unpublished)

[36] C. Gerber *et al.*, DØ internal report No. 2140, October 1994 (unpublished)

[37] N. Amos *et al.*, DØ internal report No. 2031, January 1994 (unpublished)

[38] J. Bantly, Private communication.

[39] J. Bantly *et al.*, DØ internal report No. 2544, December 1995 (unpublished)

[40] J. Bantly *et al.*, FERMILAB-TM-1995, February 1997 (unpublished)

[41] J. Drinkard, Proc. 8th Meeting of the American Physical Society Division of Particles and Fields, Albuquerque, NM (1994)

[42] J. Butler, DØ internal report No. 1682, February 1993 (unpublished)

[43] J. Tarazi, Private communication.

[44] Gene Álvarez, Ph.D. Thesis, Indiana University, June 1996

[45] Myungyun Pang, Ph.D. Thesis, Iowa State University, 1994

[46] U. Heintz and M. Narain, DØ internal report No. 2355, November 1994 (unpublished)

[47] M. Narain and U. Heintz, DØ internal report No. 2386, December 1994 (unpublished)

[48] Q. Zhu, Ph.D. Thesis, New York University, April 1994

[49] S. Abachi *et al.*, Proc. 28th International Conference on High Energy Physics, Warsaw, Poland, July 1996

[50] E. Gallas, DØ internal report No. 2066 (unpublished)

[51] Andrew James Milder, Ph.D. Thesis, University of Arizona, 1993

[52] Richard Vincent Astur, Ph.D. Thesis, Michigan State University, 1992

[53] R. Kehoe, DØ internal report No. 2052, February 1994 (unpublished). R. Kehoe and R. Astur, DØ internal report No. 2908, May 1996 (unpublished)

[54] F. Hsieh, DØ internal report No. 3031, November 1996 (unpublished)

[55] G. Marchesini *et al.*, Comp. Phy. Comm. **67**, 465 (1992)

[56] F. Paige and S. Protopopescu, Proc. Summer Study on the Physics of the Superconducting Super Collider, Snowmass, Colo., (1986)

[57] G. Altarelli, Phys. Rept. **81**, 1 (1982)

[58] F.A. Berends *et al.*, Nucl. Phys. **B357**, 32 (1991)

[59] T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994)

[60] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished)

[61] W.G.D. Dharmaratna *et al.*, DØ internal report No. 1730, May 1993 (unpublished)

[62] The MUSMEAR program was written by T. Diehl and P. Quintas.

[63] R. Kehoe and M. Narain, DØ internal report No. 2480, April 1995 (unpublished)

[64] J. Cochran, DØ internal report No. 3188, February 1997 (unpublished).

[65] R. Kehoe, Ph.D. Thesis, University of Notre Dame, January 1997

[66] R. Kehoe, Private communication.

[67] S. Abachi *et al.*, Phys. Rev. Lett. **75**, 1034 (1995)

[68] P. Tamburello, Private communication.

[69] R.E. Hall *et al.*, DØ internal report No. 2591, October 1996 (unpublished)

[70] R.E. Hall and J.W. Bantly, DØ internal report No. 3088, July 1996 (unpublished)

[71] N. Amos *et al.*, DØ internal report No. 2791, January 1996 (unpublished)

[72] M. Narain, Proc. Les Rencontres de Physique de la Vallée d'Aoste, La Thuile, Italy (1996)

[73] S.D. Ellis *et al.*, Phys. Lett. **B154**, 435 (1985)

[74] K. Genser *et al.*, DØ internal report No. 3067, January 1997 (unpublished)

[75] H. Greenlee and N. Hadley, DØ internal report No. 2808, December 1995 (unpublished)

[76] P. Bhat *et al.*, DØ internal report No. 3160, January 1997 (unpublished)

[77] Bayes, Rev. T., Philos. Trans. R. Soc. London **53**, 370 (1763)

[78] M. Bhattacharjee *et al.*, DØ internal report No. 2887, May 1996 (unpublished)

[79] F. Hsieh and H. Lan, DØ internal report No. 3055, August 1996 (unpublished)

[80] Scott Stuart Snyder, Ph.D. Thesis, State University of New York at Stony Brook, May 1995

[81] S. Snyder, private communication.

[82] A.G. Frodesen, O. Skjeggstad, and H. Tøfhte, *Probability and Statistics in Particle Physics*, Universitetsforlaget (1979)

[83] F. James, CERN program library long writeup D506, CERN (1978)

[84] L. Holmström, R. Sain, and H.E. Miettinen, Comp. Phys. Comm. **88**, 195 (1995)

[85] G. Eppley, H.E. Miettinen, and P.R.G. Virador, DØ internal report No. 2526, May 1995 (unpublished)

[86] P.C. Bhat *et al.*, FERMILAB-Pub-96/367 (1996)

[87] T. Trippe, private communication.

[88] CTEQ Collaboration, MSUHEP-41024

[89] Dhiman Chakraborty, Ph.D. Thesis, State University of New York at Stony Brook, September 1994

[90] The code for assignment of color and flavor to VECBOS partons was developed by the CDF collaboration.