Search for Anomalous Kinematics of Top Dilepton Events in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV

by

Andrew Gennadievich Ivanov

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Paul Tipton

Department of Physics and Astronomy
The College
Arts and Sciences

University of Rochester
Rochester, New York

2004

Curriculum Vitae

The author was born in Cheboxary, Russia on March 20, 1974. He attended Moscow Institute of Physics and Technology and graduated with a Bachelor of Science degree in 1996 and a Master's Degree in 1998. He came to the University of Rochester in the Fall of 1998 and began graduate studies in Physics. He pursued his research in Physics under the direction of Professor Paul Tipton and received the Master of Arts degree from the University of Rochester in 2000.

Acknowledgments

I would like to sincerely thank my advisor professor Paul Tipton for his guidance, support and encouragement in my research work.

Abstract

We present a search for anomalous kinematics of $t\bar{t}$ dilepton events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using 193 pb^{-1} of data collected with the CDF II detector. We developed a new *a priori* technique designed to isolate the subset in a data sample revealing the largest deviation from Standard Model expectations and to quantify the significance of this departure. In the four-variable space considered, no particular subset shows a significant discrepancy and we find that the probability to obtain a data sample less consistent with the Standard Model than what is observed is 1.0−4.5%.

Table of Contents

List of Figures **ix**

List of Tables **xx**

1 Introduction **1**

2 Theoretical Perspective **7**

2.1 The Standard Model of Elementary Particle Physics 7

2.2 The Top Quark 17

2.2.1 Top Production 17

2.2.2 Top Decay 19

2.3 Beyond the Standard Model 22

3 Experimental Apparatus **25**

3.1 The Tevatron Collider 25

3.2 The Collider Detector at Fermilab 29

3.2.1 The Cherenkov Luminosity Monitor 32

3.2.2 The Silicon Tracking System 34

3.2.3 The Central Outer Tracker 40

3.2.4 Calorimetry 44

3.2.5 The Muon Detectors 50

3.2.6 Data Acquisition 52

4 Particle Identification 57

4.1 Electron Identification 57

4.1.1 Central Electrons 58

4.1.2 Plug electrons 66

4.2 Muon Identification 71

4.3 Jet Reconstruction 76

4.4 Missing Energy 80

5 Dilepton Event Selection 92

5.1 Trigger Pathes and Primary Datasets 92

5.1.1 Inclusive Electron Sample 93

5.1.2 Plug Data Sample 95

5.1.3 Inclusive Muon Sample 96

5.1.4 Good Run List 97

5.2 Event Selection Criteria 99

5.2.1 Kinematic and Topological Characteristics of the Backgrounds 101

5.2.2 Dilepton Selection and Categorization 103

5.2.3 Missing transverse energy $\not\!\!E_{\mathrm{T}}$ cut 106

5.2.4 Z veto 108

5.2.5 H_T cut 114

5.3 Top Dilepton Acceptance and Background Estimates 116

5.3.1 Top Dilepton Acceptance 116

5.3.2 Vector Boson Pair Production WW/WZ Estimate 122

5.3.3 $Z \to \tau\tau$ Background Estimate 123

5.3.4 Drell-Yan Background Estimate 127

5.3.5 Lepton Fake Rate Background Estimate 133

5.3.6 $b\bar{b}$ Background Estimate 144

5.3.7 Summary 145

6 Search Strategy 147

6.1 Assessing Deviations in Kinematic Distributions 148

6.1.1 Null Hypothesis Testing 148

6.1.2 Significance Tests Based on Cumulative Distributions . . . 152

6.1.3 Power of Test 156

6.1.4 Sensitivity of Tests to Shape Anomalies 158

6.2 Isolating Unlikely Events in a Data Sample 162

6.2.1 K-subsets 163

6.2.2 Sensitivity of Statistics to Shape Anomalies by Looking at Subsamples 166

6.2.3 Isolating a Most Unlikely Subset in a Data Sample 174

6.2.4 The Purity of a Most Unlikely Subset 176

6.2.5 Performance of the Method 180

6.2.6 Understanding the Method 184

6.2.7 Summary 186

6.3 Multi-Variate Tests 187

6.3.1 PKS (Product KS) Test 188

6.3.2 MKS (Multi-dimensional KS) Test 189

6.3.3 Effect of Correlations and Dimensionality on Multivariate Statistical Tests 191

6.3.4 Barnett and Hall SUSY Model as an Alternative Hypothesis 192

6.3.5 Testing Multi-Variate Statistics on Physics Distributions . 196

6.3.6 Topology of $t\bar{t}$ Dilepton Event. 203

6.3.7 Isolating Unlikely Events in a Multi-Variate Test 209

6.3.8 Summary 214

7 Data Analysis and Results 215

7.1 W + jets Data Set, as a Control Sample of Events. 215

7.2 Top Dilepton Sample 226

7.3 Systematic Uncertainties 232

8 Conclusions 249

Bibliography 251

List of Figures

1.1 The most peculiar kinematic distributions in Run I top dilepton sample. The missing transverse energy distribution (top left); the leading lepton P_T (top right); the angle $\phi_{\ell\ell}$ between two leptons in the transverse to colliding beams plane (bottom left); the scalar sum of transverse energies of all objects in an event H_T (bottom right). 3

2.1 The Higgs potential. The neutral component of the Higgs field acquires a vacuum-expectation value $< \phi^0 >= v/\sqrt{2}$ on the circle of minima in Higgs-field space. 15

2.2 Top-quark production via the strong interaction at Tevatron proceeds through quark-antiquark annihilation (upper diagram) and gluon fusion (lower diagrams). 19

2.3 Parton distribution functions at the scale $\mu = m_t$ relevant for top-quark production. 20

2.4 Tree level $q\bar{q} \to t\bar{t}$ production followed by the SM $t\bar{t}$ decay. 21

2.5 Scalar, fermion and vector boson loop corrections to the Higgs mass. 24

3.1 The accelerator complex. 26

3.2 Elevation view of one half of the CDF detector. 31

3.3 Schematic view of luminosity monitor inside a quadrant of the CDF detector. 33

3.4 A side view of half of the CDF silicon system on a scale in which the $\hat{z}$ coordinate is highly compressed. 35

3.5 An end view of the CDF silicon system including the SVXII cooling bulkheads and ISL support structure. 36

3.6 End view of the innermost three layers of the CDF silicon system, showing Layer 00 along with the first two layers of the SVXII region. The Layer 00 electronics (not shown) are mounted beyond the active volume for vertexing. The SVXII electronics are shown just outside and just inside of each of the layers drawn. 37

3.7 A cutaway view of one quadrant of the inner portion of the CDF detector showing the tracking region surrounded by the solenoid and endcap calorimeters. 38

3.8 A view of a portion of one COT endplate. The slots define the radii of the 8 superlayers. Within a superlayer, wire-plane slots and field sheet slots alternate. 41

3.9 Three COT cells. Each cell has 12 sense wires, 13 potential wires, 4 shaper wires and a gold-mular cathode field panel on both sides of the sense wire plane. 42

3.10 Geometry of a central calorimeter wedge and its towers. 45

3.11 One of the wedges of the central electromagnetic calorimeter. The ten $(\Delta\phi, \eta) = (15°, 0.11))$ projective towers are shown. 47

3.12 Cross section of an upper part of the plug calorimeter. 48

3.13 Tower segmentation of the plug calorimeter. 49

3.14 The $\eta - \phi$ muon coverage of the CDF detector. 51

3.15 Functional block diagram of the CDF data flow. 54

4.1 Distributions of central electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.) 62

4.2 Distributions of central electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.) 63

4.3 Distributions of plug electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.) 69

4.4 The COT exit radius ρ_{COT} of CMX muons selected as $W \to \mu\nu$ candidate events. Filled histogram is the Monte Carlo simulation with no requirement on the simulated trigger. Data and simulation agree for approximately $\rho_{COT} \geq 140$ cm. 74

4.5 Number of reconstructed vertices in the data sample. 86

4.6 $\sum E_T$ distribution. The discrepancy between simulation and the data is due to multiple interactions (top). The effect of multiple interactions is excluded by requiring no more than one vertex per an event (bottom). Here and in subsequent histograms the points with error bars are the data and the histogram is Monte Carlo simulation. 87

4.7 $U_\perp$ and $U_\parallel$ distributions. 88

4.8 The $\not\!\!E_{\mathrm{T}}$ resolution $\sigma_\perp$ dependence on $\sum E_T$. The data events with one z-vertex only (top) and no cut on number of z-vertices (bottom). 89

4.9 Mean of $U_\perp$ distributions and fractions of events outside of one, two and three σ intervals as functions of $\sum E_T$. 91

5.1 Schematic diagram of triggers, paths and datasets. 93

5.2 The missing transverse energy distribution of PYTHIA Monte Carlo $t\bar{t}$ dilepton events. The magnitude of $\not\!\!E_{\mathrm{T}}$ is required to be greater than 25 GeV. The $\not\!\!E_{\mathrm{T}}$ cut is indicated by the dash line. 107

5.3 The missing transverse energy distribution of PYTHIA Monte Carlo $Z^0 \to e^+e^-$ events (left) and $Z^0 \to \mu^+\mu^-$ events (right) reconstructed in the CDF software. The arrow indicates the $\not\!\!E_{\mathrm{T}}$ cut. . . 107

5.4 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest jet versus the $\not\!\!E_{\mathrm{T}}$ for ALPGEN Monte Carlo Z^0+2 parton events with Z^0 decays to electron pairs (left) and muons pairs (right). The dashed line indicates the dilepton 'L'-cut applied in the event selection. 108

5.5 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest lepton versus the $\not\!\!E_{\mathrm{T}}$ of PYTHIA Monte Carlo $Z^0 \to \tau^+\tau^-$ events reconstructed in the CDF software. The dashed line indicates the dilepton 'L'-cut applied in the event selection. 109

5.6 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest object (lepton or jet) versus the $\not\!\!E_{\mathrm{T}}$ of PYTHIA Monte Carlo $t\bar{t}$ dilepton events reconstructed in the CDF software. The dashed line indicates the dilepton 'L'-cut applied in the event selection. 109

5.7 The dilepton invariant mass distribution of PYTHIA Monte Carlo $Z^0 \to e^+e^-$ events (left) and $Z^0 \to \mu^+\mu^-$ events (right) reconstructed in the CDF software. 110

5.8 The dilepton invariant mass distribution of PYTHIA Monte Carlo $t\bar{t}$ events with lepton pairs of same flavor. The dashed lines indicate Z^0 mass window of 76 - 106 GeV. 110

5.9 The jet signifance distribution of ALPGEN $Z^0 + 2$ partons Monte Carlo with Z^0 decays to electron pairs(left) and muons pairs(right) falling into Z^0 mass window of 76 - 106 GeV. The arrow indicates the cut value. 112

5.10 The jet significance distribution of PYTHIA Monte Carlo $t\bar{t}$ dilepton events with at least two jets. The dashed line indicates the cut value. 112

5.11 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest jet versus the jet significance for ALPGEN Monte Carlo $Z^0 + 2$ partons events (top) and PYTHIA $t\bar{t}$ events (bottom) falling into Z^0 mass window of 76-106 GeV. The events were required to pass $\not\!\!E_{\mathrm{T}}$ requirements. The dashed lines indicate the 'Z veto' event selection cuts, which effectively reduce the Drell-Yan background with a small effect on $t\bar{t}$ acceptance. In addition, about a half of $t\bar{t}$ events have jets in the opposite hemisphere from the $\not\!\!E_{\mathrm{T}}$ and they automatically pass selection. 113

5.12 The Ht distribution of PYTHIA Monte Carlo $t\bar{t}$ events passed $\not\!\!E_{\mathrm{T}}$ and Z veto requirements. The dashed line indicates the cut value of $H_T > 200$ GeV. 115

5.13 The Ht distribution of PYTHIA WW and $Z \to \tau\tau$ events passed $\not\!\!E_{\mathrm{T}}$ and Z veto requirements. The dashed line indicates the cut value of $H_T > 200$ GeV. 115

5.14 Distribution of jet correction scale factors applied to the raw jet energies. Shifted up/down jets result in larger/smaller scale factors. 121

5.15 The missing transverse energy distribution of PYTHIA Monte Carlo in the logarithm scale of $Z^0 \to e^+e^-$ events (top) and $Z^0 \to \mu^+\mu^-$ events (bottom) compared with data. Monte Carlo distributions are normalized to data. In the distributions on the left dilepton events are accompanied with one jet and in the right distributions with at least two jets. The simulation underestimates the number of events in the high-$\not\!\!E_{\mathrm{T}}$ region. 128

5.16 Dilepton invariant mass distribution M_{ee} for central-plug dielectron events in the data. 132

5.17 The calorimeter isolation in cone 0.4 as a function of E_T for CDF EM Objects from Jet20, Jet50, Jet70 and Jet100 samples. The curves for different jet samples agree near and above jet E_T threshold. 135

5.18 E_{had}/E_{em} as a function of the calorimeter isolation in cone 0.4 cone for EM jets from Jet20, Jet50, Jet70 and Jet100 with no E_{had}/E_{em} cut (top) and with $E_{had}/E_{em} < 0.125$ (bottom). Electron fakeable jets are required to satisfy $E_{had}/E_{em} < 0.125$ 137

5.19 E/P as a function of calorimeter isolation in cone 0.4 for CDF Muon objects wth no E/P cut (top) and with $E/P < 1$ requirement (bottom). We define muon fakeable jets as tracks (Muon objects) with low calorimeter activity $E/P < 1$. 141

6.1 Two histograms having the same fit to the hypothesis of constant value for f when using the χ^2-test. 150

6.2 Two scenarios of an alternative hypothesis $H1$ having the same fit using the KS test. 155

6.3 Four scenarios for alternative hypothesis $H1$. 159

6.4 π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.8$ and requiring exactly three events from the new physics peak and six events from the $H0$ distribution in each pseudoexperiment. 168

6.5 π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.4$ and requiring exactly three events from the new physics peak and six events from the $H0$ distribution in each pseudoexperiment. 170

6.6 π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.8$ and Poisson smeared $H1$-pseudoexperiments. 172

6.7 π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.4$ and Poisson smeared $H1$-pseudoexperiments. 173

6.8 The power of various tests with isolatation a most unlikely subsample (bottom) and without (top) as a function of position of the gaussian peak for non-Poisson smeared pseudoexperiments. 177

6.9 The same plots with Poisson smeared pseudoexperiments. 178

6.10 $H0$ and $H1$ distributions, where $H1$ is the one with the longer tail. 180

6.11 Size of most unlikely subset in a sample for $H0$ and $H1$ pseudoexperiments using (Δ_+, Δ_{KS}) (top) and (Δ_-, Δ_{KS}) (bottom) tests. . 182

6.12 Power functions of the tests of (Δ_+, Δ_{KS}) (top) and (Δ_-, Δ_{KS}) (bottom) for $H0$ and $H1$ distributions given in Figure 6.10. The standard KS and AD statistics are used as benchmarks. The black line represents $\pi \equiv \alpha$. 183

6.13 Maximum deviation Δ_{KS} is equivalent to maximum deviation of y_i from expected value $< y_i >$. 185

6.14 The relation between the Δ_K of the unlikely K-subset and the order statistics y_{N-K+1}. 186

6.15 $t\bar{t}$ and Barnett/Hall SUSY distributions of $\not{E}_{\mathrm{T}}$ and leading lepton P_T. These distributions provide a good discrimination between two processes. 194

6.16 $t\bar{t}$ and Barnett/Hall SUSY distributions of some of the kinematic quantities. 198

6.17 $t\bar{t}$ and Barnett/Hall SUSY distributions of angle $\Delta\phi_{\ell\ell}$ in transverse to the beam plane between two leptons for all events passing dilepton selection (top) and for events with $\not{E}_{\mathrm{T}} + P_{T,\ell 1} > 150$ GeV (bottom). 199

6.18 $t\bar{t}$ and Barnett/Hall SUSY distributions of angle $\Delta\phi_{\ell m}$ in transverse to the beam plane between leading lepton and direction of $\not{E}_{\mathrm{T}}$. . . 200

6.19 The neutrino momentum lies on the cone with the axis along direction of $\vec{P}$ and the open angle θ, such that $\cos\theta = E/|\vec{P}|$. 204

6.20 $t\bar{t}$ and Barnett/Hall SUSY distributions of variable T. 208

6.21 Distribution of T for $t\bar{t}$ events. T can be well approximated with a gaussian with an additional bin at $T \simeq 0$ for poorly reconstructed events. 210

6.22 Distribution of T for WW events. 210

6.23 Power functions $\pi(\alpha)$ of the test for different mixtures of $t\bar{t}$ and SUSY. 213

7.1 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the leading jet versus the $\not\!\!E_{\mathrm{T}}$ of multi-jet QCD data with a non-isolated electron or muon candidate. The dashed line indicates the cut in $\Delta\phi_{jm1}$-$\not\!\!E_{\mathrm{T}}$ plane, which reduces the QCD background by 50%. 219

7.2 The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the leading jet versus the $\not\!\!E_{\mathrm{T}}$ of ALPGEN+ HERWIG Monte Carlo W + 3 partons events. The dashed line cut in $\Delta\phi_{jm1}$-$\not\!\!E_{\mathrm{T}}$ plane removes approximately 10% of W + jets events and 7% of $t\bar{t}$ events. 220

7.3 Lepton Isolation versus $\not\!\!E_{\mathrm{T}}$ distributions for events with a lepton and ≥ 3 jets. For events with $\not\!\!E_{\mathrm{T}} < 35$ GeV, $\Delta\phi_{jm1}$ cut is applied. 221

7.4 $\not\!\!E_{\mathrm{T}}$ and leading lepton P_{T,ℓ_1} distributions in $W + 3$ jets sample. . 222

7.5 $\Delta\Phi_{\ell m1}$ distribution in $W + 3$ jets sample. 223

7.6 The data and MC W + 3p events in $(\not\!\!E_{\mathrm{T}}, P_{T,\ell 1})$ plane. The PKS test isolates a most unlikely subset of three events in the tails of these distributions with $\alpha = 36.2\%$. 225

7.7 T distributions in $W \geq 4$ jets sample of events. Events are reconstructed as if they were $t\bar{t}$ dilepton events by treating the two leading jets as a second lepton and neutrino, and two other jets as b-jets (the top plot); and alternatively by treating two leading jets as b-jets and subleading jets as lepton-neutrino (the bottom plot). 227

7.8 $\not\!\!E_{\mathrm{T}}$ and $P_{T,\ell 1}$ distributions in the Run II top dilepton sample. The error bars represent Monte Carlo uncertainties. 230

7.9 $\Delta\phi_{\ell m1}$ and T distributions in the top dilepton sample. The error bars represent Monte Carlo uncertainties. 231

7.10 Top dilepton events in (P_T, T) plane with b-tagging information. . 232

7.11 Distributions of jet kinematic quantities in the Run II top dilepton sample. $E_B = \not\!\!E_{\mathrm{T}} + \sum E_{T,j}$. The first bin in the third jet E_T distribution is populated by events with 2 jets. 233

7.12 Distributions of lepton kinematic quantities in the Run II top dilepton sample. $E_S = \not\!\!E_{\mathrm{T}} + \sum E_{T,\ell}$. Notice, that the Run II dilepton sample also has a slight excess of $e\mu$ events. 234

7.13 Kinematic distribution of $F_0(x)$ and $F_{HERWIG}(x)$ templates. . . . 238

7.14 Kinematic distributions of $F_0(x)$ and $F_{noISR}(x)$ templates. The $\alpha_{noISR} = 1.2\%$. 240

7.15 Kinematic distributions of $F_0(x)$ and F_{MRST} templates. The $\alpha_{MRST} = 1.9\%$. 241

7.16 Kinematic distributions for templates with jet energies fluctuating up (left) and down (right). Leading lepton P_T distribution remains unchanged as a result of these shifts. 242

7.17 Kinematic distributions for templates with $m_t = 170$ GeV. 243

7.18 Kinematic distributions for templates with $m_t = 180$ GeV. 244

7.19 Kinematic distributions for backgrounds shifted upward. 246

7.20 Kinematic distributions for backgrounds shifted downward. 247

List of Tables

2.1 The three generations of the fundamental particles of the Standard Model. 8

2.2 Lepton quantum numbers. 12

2.3 Quark quantum numbers. 13

2.4 $t\bar{t}$ decay modes and their associated branching ratios. 22

4.1 Muon identification efficiencies in the data and the simulation and scale factors to be applied to correct acceptance measurement from Monte Carlo samples. 75

5.1 High P_T inclusive electron sample stripping cuts. 95

5.2 High P_T inclusive muon sample baseline stripping cuts. 98

5.3 Total integrated luminosities for different data samples. 98

5.4 Dilepton decay channels used in the analysis. The first lepton in the dilepton pair is the tight lepton and the second is either tight or loose. 105

5.5 The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive $t\bar{t}$ events and giving rise to the total dilepton acceptance $\epsilon_{dil}^{MC} = (0.813 \pm 0.014)\%$. Errors are statistical only. The corrections for lepton trigger efficiency and lepton ID scale factor reduce the acceptance by 15%. 118

5.6 The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of $WW \to \ell\bar{\ell}$ events. Errors are statistical only. This gives rise to $\epsilon_{WW}^{MC} = (0.177 \pm 0.005)\%$. The correction factor of 1.68 ± 0.15 is applied to account for underestimated value of ϵ_{2-jet} from PYTHIA. The lepton trigger efficiencies and ID scale factors further decrease the acceptance by 15%. 124

5.7 The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive WZ events, which give rise to $\epsilon_{WZ}^{MC} = (0.021 \pm 0.001)\%$. This is further corrected for lepton trigger efficiences and ID scale factors. Errors are statistical only. 125

5.8 The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive $Z \to \tau\tau$ events. The MC acceptance $\epsilon_{Z\to\tau\tau}^{MC} = (0.054 \pm 0.004)\%$. The correction factor of 1.68 ± 0.15 is applied to correct for ϵ_{2-jet} underestimate. Errors are statistical only. 126

5.9 Expected non-Drell Yan events inside of the Z mass window. The contributions are estimated with 30% uncertainty. 131

5.10 The fraction of events $f^j_{\not{E}_T}$ and f^j_{Zveto} in each jet multiplicity bin estimated from Monte Carlo. 131

5.11 The $R^j_{out/in}$ scale factors determined from Monte Carlo dilepton invariant mass distributions. 131

5.12 H_T cut efficiencies inside and outside of the Z mass window determined from ALPGEN Monte Carlo. Errors are statistical only. . . 132

5.13 Number of observed and predicted fakes for electron categories. We use Jet50 to predict Jet20, Jet70 and Jet100. 139

5.14 Number of observed and predicted fakes for muon categories. We use Jet50 to predict Jet20, Jet70 and Jet100. 142

5.15 Fake rate estimates using different jet samples. 143

5.16 Lepton fake rate background estimate. 144

5.17 Lepton fake rate prediction for same-sign events is in agreement with observations. 144

5.18 The breakdown of physics processes into different jet bins. 146

5.19 The breakdown of physics processes into different flavor composition.146

6.1 The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 0.0$. 159

6.2 The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 1.0$. 161

6.3 The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 2.0$. 161

6.4 The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 3.0$. 162

6.5 The KS probabilities of compatibility of $t\bar{t}$ distribution with SUSY for various kinematic variables. 197

7.1 Identification requirements for the second lepton used in the Z boson veto. 218

7.2 Results of the KS test applied to kinematic distributions. 223

7.3 Samples used for modelling the shapes of various Standard Model processes contributing to the $t\bar{t}$ dilepton signature. 228

7.4 Details of kinematics of the top dilepton candidate events.Energies and momenta are in units of GeV. Jets having b-tags are shown in boxes. 235

7.5 Details of kinematics of the top dilepton candidate events (continued). Energies and momenta are in units of GeV. Jets having b-tags are shown in boxes. 236

7.6 Values of α_X obtained upon inclusion of systematic effects. The last row shows the maximum range resulting from various combinations of the individual systematics. 248

Chapter 1

Introduction

The top quark was discovered in a variety of decay channels at Fermilab's $p\bar{p}$ collider Tevatron in 1995 by both the CDF and D0 colaborations [1]. It was a major triumph of the Standard Model, as it provided the direct experimental confirmation of the three-generation structure. The strikingly large mass of the top quark, the heaviest fundamental particle known, has caused a lot of excitement. It suggests the plausibility that the top quark may play a role in the breaking of electroweak symmetry and hence in the origin of fermion masses.

Once the existence of the top quark has been established, the extensive experimental program started to characterize the nature of top quark production and top decay properties in all possible decay channels. In the Tevatron collider top quarks are produced mainly in pairs via quark-antiquark annihilation. Within the framework of the Standard Model a top quark decays almost exclusively to a W boson and a b quark, where the W boson decays into either a lepton ℓ and its neutrino ν_ℓ or a $q\bar{q}$ pair. Experimentally $t\bar{t}$ decay channels are classified according to subsequent decay modes of the W bosons.

The topic of interest of this thesis is the dilepton decay chain $t\bar{t} \to \bar{b}W^-bW^+ \to b\ell^-\nu_\ell\bar{b}\ell^+\bar{\nu}_\ell$, which denotes the case when both W bosons decay leptonically to an

e or μ. Such events are characterized by two highly energetic leptons, two jets from the fragmentation of the b quarks and a large missing energy $\not\!\!E_{\mathrm{T}}$ from the unobserved neutrinos.

During the Tevatron Run I, which covers the period of operation from August 1992 to July 1995, there was 109 pb^{-1} of integrated luminosity accumulated at the Collider Detector at Fermilab (CDF). The $t\bar{t}$ production cross section measurement in the dilepton channel showed a slight excess over the Standard Model predictions – CDF observed nine top dilepton candidates while expecting 4.5 from $t\bar{t}$ production, assuming top mass $m_t = 175$ GeV and using the theoretical top cross-section $\sigma_{t\bar{t}} = 5.5$ pb, and 2.4 ± 0.5 events from non-top Standard Model backgrounds [2].

Perhaps more interestingly, however, several of the observed events had kinematics, which called into question their compatibility with Standard Model top quark decay. These events possessed large $\not\!\!E_{\mathrm{T}}$ and lepton E_T's quite unlike those expected from $t\bar{t}$ production (see Figure 1.1). The D0 collaboration had also observed dilepton events, one of which held similar characteristics [3].

These peculiarities led authors of Reference [4] to hypothesize that the recalcitrant events may originate from a supersymmetry source as the events populating the tails of E_T distributions are in better agreement with those, which result from the cascade decays of heavy squarks. They have also utilized the fact that the distribution of the angle $\phi_{\ell\ell}$ between two leptons in the transverse to beam plane revealed a slight excess over the Standard Model expectations in the low ϕ region (see Figure 1.1) as anticipated from decays of heavy supersymmetric particles.

Authors of Reference [5] attempted to reconstruct the dilepton events under the top hypothesis taking into acccount all kinematic quantities in an event and solving energy-momentum conservation equations, constraint with the known

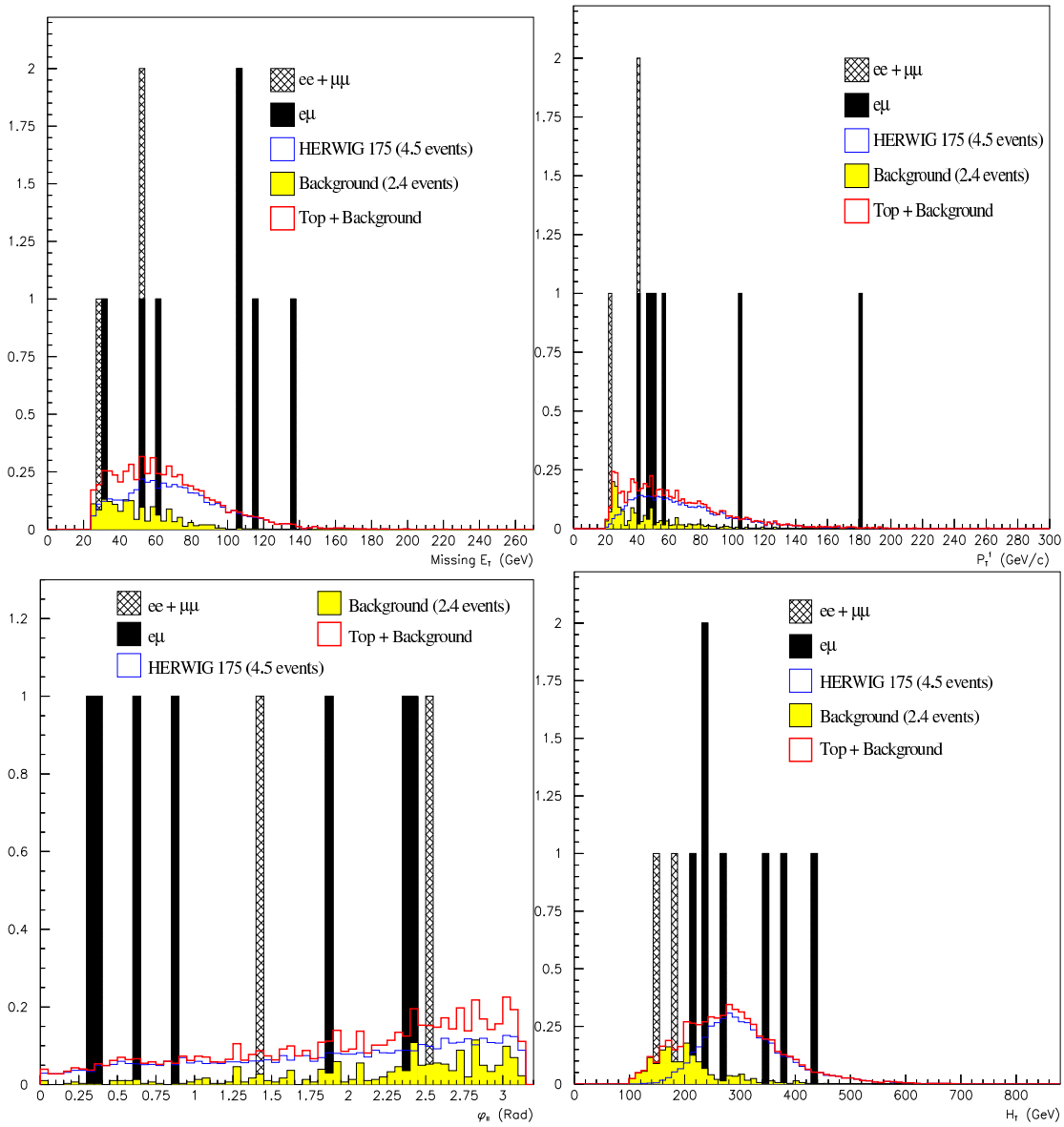

Figure 1.1: The most peculiar kinematic distributions in Run I top dilepton sample. The missing transverse energy distribution (top left); the leading lepton P_T (top right); the angle $\phi_{\ell\ell}$ between two leptons in the transverse to colliding beams plane (bottom left); the scalar sum of transverse energies of all objects in an event H_T (bottom right).

masses of primary and intermediate particles (m_t, m_W). They concluded that two of the dilepton events fit poorly to $t\bar{t}$, although it was possible that they were a statistical fluctuation ($\sim 1.5 - 2.0$ standard deviations effect).

In contrast to the event by event basis study, kinematics of the entire ensemble of top dilepton candidates was analyzed in Reference [6]. A goodness-of-fit test was performed between several kinematic distributions of observed events and those expected from combination of top and Standard Model background. Assuming that some of the data events may possibly be coming from non-standard decays, the authors introduced the concept of *subsets* of the event sample with a purpose to potentially isolate any new physics.

Although, the Run I events were found to be consistent to greater than the $\sim 6\%$ level for the most dramatic $\not\!\!E_{\mathrm{T}}$ distribution, speculations about possibility of deviations from Standard Model continued pointing to a relative excess of $e\mu$ events over same-flavor dilepton events (ee, $\mu\mu$). Indeed, only one ee and one $\mu\mu$ were observed in contrast to seven of $e\mu$ events while expecting 15% ee, 25% $\mu\mu$ and 60% $e\mu$ [7]. If one accounts for low H_T of the same-flavor dilepton events, suggesting that they are more likely to be originating from background sources (see Figure 1.1), then the flavor assymetry becomes more significant. The latter conjecture was obviously made *a posteriori* and hence does not deserve as much credibility, but nevertheless makes the Run I top dilepton data sample even more intriguing.

This thesis was born based on these curiosities in the Run I dilepton sample. It represents an attempt to build a robust systematic approach for handling and adequately quantifying deviations in the shapes of kinematic distributions, especially those exhibiting in the tails similar to the properties of the aforementioned Run I events.

The new devised technique described in this thesis is an original work. It represents a data-driven and model-independent search for new physics based on the comparison of kinematic features of observed events with Standard Model expectations. The method is designed to seek and isolate the most unlikely subset of events in a data sample and assign an *a priori* probability that quantifies its departure from the Standard Model.

As the dilepton event statistics accumulated during the Run I turned out to be not sufficient to infer new physics evidence, the interest naturally arose to whether the same tendency would be repeated with a larger amount of data expected in the next Tevatron period of data taking, Run II. And if so, would it let us come closer to potential discovery of the physics beyond the Standard Model.

The Run II started in summer of 2001. Both the accelerator complex and the CDF detector have undergone major upgrades. The Tevatron has reached a center-of-mass energy of 1.96 TeV and has achieved a much higher luminosity. The CDF detector has obtained a new tracking system, a new plug calorimeter and an upgraded data acquisition system to handle higher event rates.

In this thesis we present the results of the search for anomalous kinematics in top dilepton events observed at the upgraded Collider Detector at Fermilab using 193 pb^{-1} of Tevatron Run II $p\bar{p}$ collision data, a sample nearly twice as large as that used in Run I.

The thesis is organized as follows. In Chapter 2 we review the present understanding of the matter structure, which constitutes the Standard Model of particle physics. We discuss the top quark properties and its role in the modern particle physics and speculate what physics may lie beyond the Standard Model.

Chapter 3 describes the structure and functionality of the Tevatron accelerator complex and the CDF detector. In Chapter 4 we discuss the reconstruction and

identification of final decay products by the CDF detector components. Chapter 5 is dedicated to the Run II top dilepton selection used in this analysis.

The discussion of the new technique does not start until Chapter 6, although the technique was developed *a priori* before looking at the Run II data sample. We start the Chapter with a review of several goodness-of-fit statistical tests, discuss their sensitivity to various shape anomalies and then introduce concepts of *subsets* and demonstrate with a set of examples the benifits of this approach. We further describe the multi-variate generalization of the method, whose purpose is to extract all available information about events. We conclude the Chapter with the test on performance of the method on the physics distributions and selection of kinematic variables.

In Chapter 7 we discuss results of this test applied to the Run II dilepton sample. and then go through estimates of systematic uncertainties. Chapter 8 concludes the thesis.

Chapter 2

Theoretical Perspective

2.1 The Standard Model of Elementary Particle Physics

The Standard Model is an effective field theory accurately describing all particle physics phenomena down to scales of 10^{-18} m. Within the framework of this theory, the matter is composed from elementary particles that are subject to interactions via exchange of intermediate interaction particles, called *gauge particles.*

The fundamental constituents of the matter are classified into leptons and quarks, which are further organized into three families, called *generations.* The macroscopic matter is made of the first generation particles. The first generation consists of two leptons: electron e^- with electric charge $Q = -1$ and its corresponding neutrino ν_e with $Q = 0$, and two quarks: the up-quark u with fractional charge $Q = \frac{2}{3}$ and the down-quark d with $Q = -\frac{1}{3}$, and their anti-particles – positron e^+, $\bar{\nu}_e$, $\bar{u}$, $\bar{d}$, generally denoted by the bar above the particle name. The anti-particles have the same masses but opposite quantum numbers. For instance, an electrical charge: $\bar{Q} = -Q$.

	Leptons		Quarks	
charge Q	-1	0	$+\frac{2}{3}$	$-\frac{1}{3}$
1st Generation	e	ν_e	u	d
2nd Generation	μ	ν_μ	s	c
3rd Generation	τ	ν_τ	b	t

Table 2.1: The three generations of the fundamental particles of the Standard Model.

The quarks have an additional quantum number, the color. The color can be of three types – red R, green G and blue B, with the anti-colors $\bar{R}$, $\bar{G}$ and $\bar{B}$. Free particles with color are not observed in nature and therefore the quarks must be confined into the experimentally observed matter particles, named *hadrons*. These colorless composite particles are classified into baryons and mesons. The baryons are made of three quarks, qqq. For instance, the proton $p \sim uud$ and the neutron $n \sim ddu$. The mesons are made of one quark and one antiquark. For instance, the pions $\pi^+ \sim u\bar{d}$ and $\pi^- \sim d\bar{u}$.

The second and third generation particles have identical properties except they are heavier. The second generation particles are the muon μ^-, the muon neutrino ν_μ, the strange quark s and the charm quark c. The third generation consists of the tau lepton τ, its neutrino ν_τ, the bottom and top quarks b and t (see Table 2.1).

The gauge particles include the massless and chargeless photon γ exchanged in the electromagnetic interactions, the massive weak bosons – $W^\pm$ with charge $Q = \pm 1$ respectively and electrically neutral Z^0 carrying the weak interactions, and the eight gluons g_α mediating the strong interactions among quarks. Each gluon is massless, electrically neutral and carries two color charges, one color and one anti-color, therefore there are eight independent combinations. The consequence

of the gluons being colorful is that they interact not only with the quarks but also with themselves. The weak bosons, $W^{\pm}$ and Z^0, are also self-interacting. The gravitational interactions are ignored in the Standard Model.

The foundation of the Standard Model is a symmetry, which plays a crucial role in particle physics. The symmetry S of the physical system implies an invariance under the transformations given by S. The set of these transformations forms a symmetry group G. In addition to space-time symmetries, such as translations, rotations and etc., the Standard Model incorporates internal symmetries, i.e. those, which act on the internal quantum numbers and are hidden in the present nature.

The internal symmetries in general depend on the space-time coordinates. Such symmetries are called local or *gauge symmetries*. In order for the theory to be invariant under local transformations the gauge vector boson fields need to be introduced. These fields interact in a gauge invariant manner and the number of the gauge fields required to be equal to the number of generators of the symmetry group.

The Standard Model is based on the gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$. This gauge group includes the symmetry group of the strong interactions, $SU(3)_C$, and the symmetry group of the electroweak interactions, $SU(2)_L \times U(1)_Y$. In general, the $SU(N)$ group has N^2-1 generators, hence one obtains eight gluons, three weak bosons and a photon.

The group symmetry of the electromagnetic interactions, $U(1)_{em}$, appears in the SM as a subgroup of $SU(2)_L \times U(1)_Y$. In this respect the weak and electromagnetic interactions are unified.

The fundamental particles of matter are fermions with spin $s = 1/2$. In the absence of gauge fields their dynamics is described by Dirac equation with the

corresponding Lagrangian:

$$\mathcal{L}_{Dirac} = \bar{\Psi}(x)(i\ \partial\!\!\!/ - m)\Psi(x); \qquad \partial\!\!\!/ \equiv \partial_\mu \gamma^\mu \tag{2.1}$$

The gauge theory based on $U(1)_{em}$ is called Quantum Electrodynamics (QED), which is the most successful gauge theory particle physics tested up to an extremely high level of precision. The invariance under the $U(1)$ gauge transformations $\Psi \rightarrow \Psi e^{iQ\theta(x)}$ introduces one gauge vector boson field, the photon field $A_\mu(x)$, and the $U(1)$ gauge invariant Lagrangian becomes

$$\mathcal{L}_{QED} = \bar{\Psi}(x)(i\ D\!\!\!/ - m)\Psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x), \tag{2.2}$$

where the covariant derivative, D_μ,

$$D_\mu \Psi \equiv (\partial_\mu - ieQA_\mu)\Psi \tag{2.3}$$

contains the interaction term, and the propagation of the photon field is given in terms of the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.4}$$

The gauge theory for strong interactions based on $SU(3)_C$ is called Quantum Chromodynamics (QCD). The $SU(3)_C$ is the non-Abelian Lie group generated by the color transformations. C refers to colors and 3 refers to the three possible color states of the quarks.

The QCD invariant Lagrangian is built similarly to as in the QED case. The gauge symmetry $SU(3)$ is promoted by the QCD covariant derivative:

$$D_\mu q \equiv \left(\partial_\mu - ig_s(\frac{\lambda_\alpha}{2})A^\alpha_\mu\right) q, \tag{2.5}$$

where

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \tag{2.6}$$

are the quark fields, g_s is the strong coupling constant, $\frac{\lambda_\alpha}{2}$ are $SU(3)$ generators given by 3×3 traceless hermitian matrices, and A^α_μ are gluon fields, $\alpha = 1, .., 8$.

The QCD Lagrangian is then

$$\mathcal{L}_{QCD} = \sum_q \bar{q}(x)(i \not{D} - m_q)q(x) - \frac{1}{4}F^\alpha_{\mu\nu}(x)F^{\mu\nu}_\alpha(x), \tag{2.7}$$

with the gluon field strength tensor

$$F_{\mu\nu}(x) = \partial_\mu A^\alpha_\nu(x) - \partial_\nu A^\alpha_\mu(x) + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma} \tag{2.8}$$

containing a bilinear term, which provides the gluon self-interaction.

The gauge theory based on the composed group $SU(2)_L \times U(1)_Y$ is called the Electroweak Theory, where L is referred to *handedness* and Y is the *hypercharge.* The left-handed and right-handed fields are defined by means of the chirality operator γ_5. For instance, for the electron

$$e^-_L = \frac{1}{2}(1 - \gamma_5)e^-; e^-_R = \frac{1}{2}(1 + \gamma_5)e^-. \tag{2.9}$$

Four solutions of Dirac equation led to prediction of existence of anti-particles. These solutions in the Weyl representation are combined differently as two particles consisting of the following states – the left-handed electron with its partner

Lepton	T	T_3	Q	Y
ν_L	$\frac{1}{2}$	$\frac{1}{2}$	0	-1
e_L	$\frac{1}{2}$	- $\frac{1}{2}$	-1	-1
e_R	0	0	-1	-2

Table 2.2: Lepton quantum numbers.

the right-handed positron (e_L^-, e_R^+) and the right-handed electron with its partner the left-handed positron (e_R^-, e_L^+). This representation proved to have a deeper meaning than the one of Dirac, since the weak interaction couples only to the left-handed projections of the fermions. That fact is reflected in index L.

The left-handed fermions transform as doublets under $SU(2)_L$:

$$f_L \rightarrow e^{i\vec{T}\vec{\theta}} f_L; f_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, ... \tag{2.10}$$

whereas the right-handed fermions transform as singlets:

$$f_L \rightarrow f_R; f_R = e_R, u_R, d_R, ... \tag{2.11}$$

Here $\vec{T}$ is the *weak isospin* and $T_i = \sigma_i/2, i = 1, 2, 3$ are three generators of the $SU(2)_L$. The fourth generator is the $U(1)_Y$ generator, $\frac{Y}{2}$, which commutes with $T_i, i = 1, 2, 3$. The fermion quantum numbers satisfy the relation

$$Q = T_3 + \frac{Y}{2} \tag{2.12}$$

and are given in Tables 2.2 and 2.3.

The associated gauge bosons are $W_\mu^i, i = 1, 2, 3$ – the weak bosons of $SU(2)_L$ and B_μ is the hypercharge boson of $U(1)_Y$.

Quark	T	T_3	Q	Y
u_L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{3}$
d_L	$\frac{1}{2}$	- $\frac{1}{2}$	- $\frac{1}{3}$	$\frac{1}{3}$
u_R	0	0	$\frac{2}{3}$	$\frac{4}{3}$
d_R	0	0	- $\frac{1}{3}$	$-\frac{2}{3}$

Table 2.3: Quark quantum numbers.

The covariant derivative in the electroweak theory is

$$D_\mu f \equiv \left(\partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - ig'\frac{Y}{2}\cdot B_\mu\right) f, \tag{2.13}$$

where g and g' are coupling constants corresponding to $SU(2)_L$ and $U(1)_Y$. Similarly to QED and QCD the electroweak Lagrangian includes terms for the gauge fields

$$-\frac{1}{4}W^i_{\mu\nu}W^{\mu\nu}_i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu},$$

where the field strength tensors are

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu(x) - \partial_\nu W^i_\mu(x) + ge^{ijk}W_{\mu j}W_{\nu k} \tag{2.14}$$

and

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{2.15}$$

Thus the gauge Lagrangian contains the self-interaction terms among the three $W^i_\mu, i = 1, 2, 3$ gauge bosons, as it corresponds to a non-abelian $SU(2)_L$ group.

The physical gauge bosons $W^\pm_\mu, Z_\mu$ and A_μ are obtained from the electroweak interaction eigenstates by the following expressions

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \tag{2.16}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \tag{2.17}$$

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W. \tag{2.18}$$

Here θ_W is the *weak mixing angle*, which defines the rotation in the neutral sector. This is a parameter of the Standard Model, which is measured experimentally

$$\sin^2\theta_W = 0.2255 \pm 0.0021 \tag{2.19}$$

The problem, however, is that unlike in the QED and QCD cases the intermediate vector bosons $W^\pm$ and Z are massive, and mass terms as $m_W^2 W_\mu W^\mu$ and $\frac{1}{2}m_Z^2 Z_\mu Z^\mu$ violate the $SU(2)_L \times U(1)_Y$ gauge invariance. The gauge boson masses must be generated in a gauge invariant way.

To resolve this problem, the phenomenon of *spontaneous symmetry breaking* (SBS) is introduced, which is one of the key ingredients of the Standard Model. Spontaneous symmetry breaking implies that the Lagrangian describing the dynamics of the physical system is invariant under the symmetry transformations U, but the vacuum is not: $U|0> \neq |0>$. Or alternatively, there should exist field operators with non-vanishing vacuum expectation values: $<0|\Phi|0> \neq 0$.

The simplest Lagrangian of the SBS of the electroweak theory can be written as

$$\mathcal{L}_{SBS} = (D_\mu\Phi)^+(D^\mu\Phi) - V(\Phi), \tag{2.20}$$

where

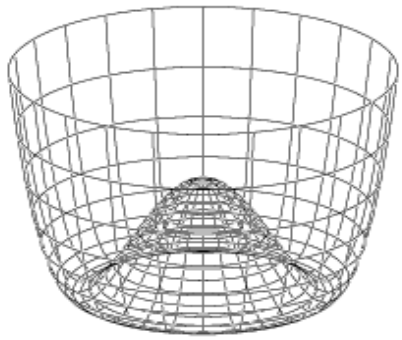

Figure 2.1: The Higgs potential. The neutral component of the Higgs field acquires a vacuum-expectation value $< \phi^0 >= v/\sqrt{2}$ on the circle of minima in Higgs-field space.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi_0 \end{pmatrix} \tag{2.21}$$

is a fundamental complex doublet with hypercharge $Y(\Phi) = 1$ and

$$V(\Phi) = -\mu^2 \Phi^+ \Phi + \lambda(\Phi^+ \Phi); \quad \lambda > 0 \tag{2.22}$$

is the simplest renormalizable potential. For $(-\mu^2) < 0$, the minima of the potential $V(\Phi)$ are achieved on a circle of radius $v \equiv \sqrt{\mu^2/\lambda}$ (see Figure 2.1):

$$| < 0|\Phi|0 > | = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{2.23}$$

Consequently, in the lowest energetic state the system spontaneously choses one vacuum expectation value, which no longer reflects the symmetry of the potential. The physical spectrum is then built by performing 'small oscillations' around the vacuum. By parameterizing $\Phi(x)$ as

$$\Phi(x) = \exp\left(i\frac{\vec{\xi}(x)\vec{\sigma}}{v}\right)\begin{pmatrix} 0 \\ (v+H(x))/\sqrt{2} \end{pmatrix} \tag{2.24}$$

and then eliminating the unphysical fields $\vec{\xi}(x)$, one obtains the tree level mass predictions:

$$m_W = \frac{1}{2}gv \qquad m_Z = \frac{1}{2}\sqrt{g^2+g'^2}v \tag{2.25}$$

and the photon stays massless. The parameter v was determined experimentally from μ-decay and is equal to 246 GeV. Thus the masses $m_W = 80.4$ GeV and $m_Z = 91.2$ GeV were anticipated before they were measured in experiment.

This procedure is called the *Higgs mechanism.* It introduces a new scalar boson H, called the *Higgs boson.* The interactions of H with fermions introduce an additional (*Yukawa*) term into the Standard Model Lagrangian

$$\mathcal{L}_{Yukawa} = -M_U^{ij}\bar{U}_L^i U_R^j - M_D^{ij}\bar{D}_L^i D_R^j - M_E^{ij}\bar{E}_L^i E_R^j + h.c., \tag{2.26}$$

where

$$U_{R,L} = \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix}_{R,L}, \quad D_{R,L} = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{R,L}, \quad E_{R,L} = \begin{pmatrix} e' \\ \mu' \\ \tau' \end{pmatrix}_{R,L}, \tag{2.27}$$

and M_U, M_D, M_E are 3×3 complex matrices in generation space. By diagonalizing these matrices one obtains the physical mass eigenstates of quarks and leptons.

Hence the Higgs boson gives masses to the gauge bosons $W^\pm$ and Z^0 as well as to the quarks and leptons. However, unlike the $W^\pm$ and Z^0 masses predicted

in terms of the coupling constants and the weak mixing parameter, every fermion mass m_f is set by a separate Yukawa coupling λ_f, such that

$$m_f = \lambda_f \frac{v}{\sqrt{2}} \tag{2.28}$$

The Yukawa couplings are unknown parameters of the Standard Model. Therefore, in contrast to the gauge boson sector, the hierarchy in the fermion masses remains completely undetermined. The SM also does not give predictions for the Higgs self-coupling λ and the Higgs boson mass $M_H = \sqrt{2v^2\lambda}$. So far the Higgs boson remains the only undiscovered particle predicted in the Standard Model.

2.2 The Top Quark

The measured quark and lepton mass spectrum is one of the most intriguing mysteries in modern particle theory. It has a remarkable feature – only the top quark has the mass ($m_t \simeq 175$ GeV) of the order of the electroweak symmetry breaking scale with $\lambda_t \approx 1$, while masses of the other quark and leptons are much smaller. In fact, the top quark is the heaviest elementary particle yet discovered.

Top's great mass suggests that it stands apart from the other quarks and leptons and indicates that it couples strongly to agents of electroweak symmetry breaking and possibly plays a role in origin of fermion masses.

2.2.1 Top Production

To date, only the Fermilab Tevatron has energy sufficient to produce particles as massive as the top quark. At hadron colliders the top quarks are produced mainly in pairs via the strong interaction. The leading order (LO) Feynman diagrams

are shown in Figure 2.2. At leading order the top pair production occurs via quark-antiquark ($q\bar{q}$) annihilation and gluon fusion (gg).

The proton is regarded as a collection of quarks, antiquarks, and gluons (collectively called partons), each carrying some fraction x of the proton's four-momentum. The probability densities of finding a parton with a given momentum fraction in a proton are called *parton distribution functions (PDFs).* The total production cross section $p\bar{p} \to t\bar{t}$ is a convolution of the PDFs for the incoming protons and antiprotons and the cross sections for the partonic processes $q\bar{q}, gg \to t\bar{t}$:

$$\sigma(p\bar{p} \to t\bar{t}) = \sum_{a,b} \int dx_a dx_b f_a^p(x_a, \mu^2) f_b^{\bar{p}}(x_b, \mu^2) \hat{\sigma}(ab \to t\bar{t}; \hat{s}, \mu^2, m_t), \tag{2.29}$$

where the summation indices a and b run over light quarks and gluons contained in the initial proton and antiproton and carrying momentum fractions of x_a and x_b, respectively. Note that parton distribution functions in perturbative QCD also depend on the factorization scale μ corresponding to a scale of the process, such as m_t.

At Tevatron the total energy in the center-of-momentum frame of $p\bar{p}$ collisions is $\sqrt{s} = 1.96$ TeV. In order to produce a $t\bar{t}$ pair at rest, the square of the total energy of the partonic subprocess $\hat{s} = x_a x_b s$ has to be greater than $4m_t^2$. Hence the typical value of x for $t\bar{t}$ production at Tevatron is

$$x \approx \frac{2m_t}{\sqrt{s}} \approx 0.18 \tag{2.30}$$

Figure 2.3 shows the parton distribution functions in the proton for different parton species for the relevant scale of the process $\mu = m_t$. One can see, that for $x \approx 0.18$ the up distribution function is larger than that of the gluon, and the

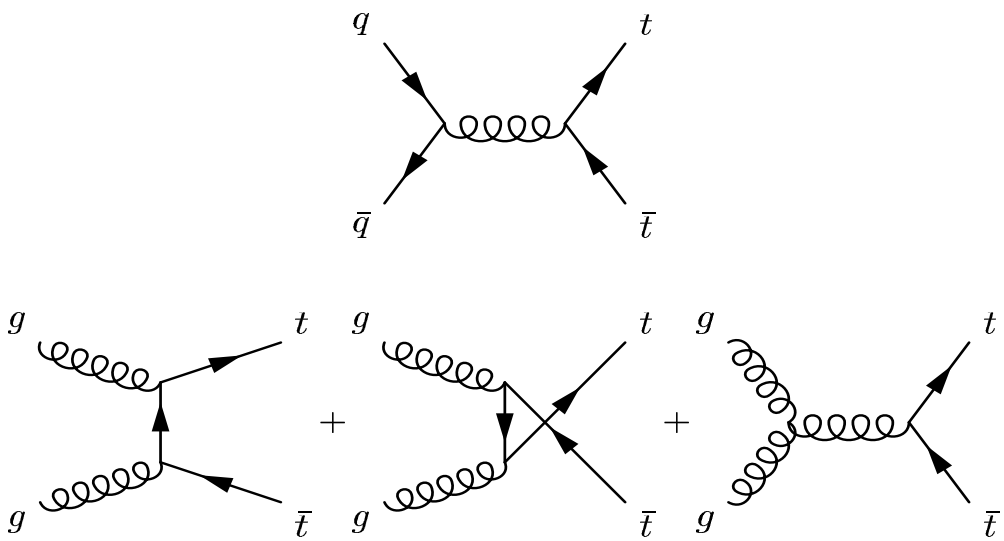

Figure 2.2: Top-quark production via the strong interaction at Tevatron proceeds through quark-antiquark annihilation (upper diagram) and gluon fusion (lower diagrams).

down distribution function is comparable to it. Therefore at Tevatron energies the relative contributions of the $q\bar{q}, gg \to t\bar{t}$ processes are about 90% and 10% respectively.

The next-to-leading order (NLO) contributions to the $t\bar{t}$ production cross section include corrections due to the soft gluon emission and evaluate $\sigma_{t\bar{t}}^{NLO} = 6.7^{+0.7}_{-0.9}$ pb [8].

2.2.2 Top Decay

Within the Standard Model, the dominant decay of a top quark is via $t \to Wb$, with a branching ratio of nearly 100%. The decays $t \to Ws$ and $t \to Wd$ are suppressed by factors of $10^{-3} - 10^{-4}$ by the Cabibbo-Kobayashi-Maskawa (CKM)

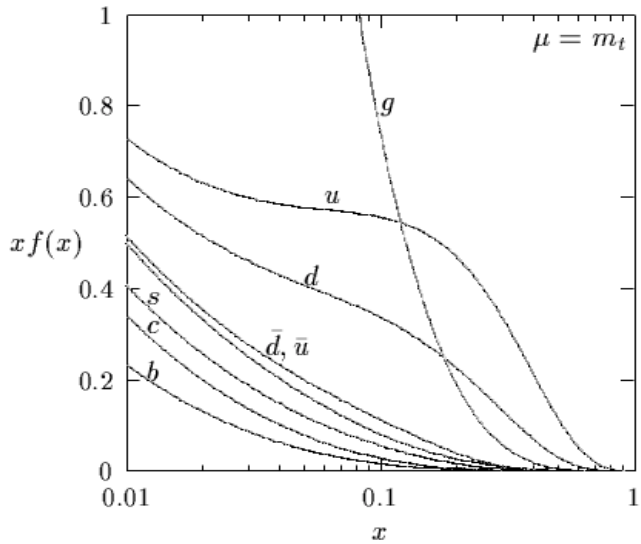

Figure 2.3: Parton distribution functions at the scale $\mu = m_t$ relevant for top-quark production.

mixings. [1] Other decays, such as neutral current decays, are loop induced and are many orders of magnitude smaller [9].

Due to the large mass of the top quark and a CKM mixing parameter V_{tb} near unity, its lifetime is extremely short ($\Gamma_t \approx 1.5$ GeV [10] corresponding to $\tau_t \sim 4 \times 10^{-25}$), so that at Tevatron energies it decays it has a chance to hadronize and can be treated as a free particle.

A $t\bar{t}$ final state contains two $t \to Wb$ decays. The $t\bar{t}$ decays are characterized by decays of W^+ and W^-. Each W boson can decay in the following ways:

[1] The CKM matrix arises in the Standard Model, as a result of the most general Yukawa interaction compatible with gauge invariance. Specifically, diagonalizing of mass matrices M_U, M_D, M_E introduced in Yukawa Lagrangian (Equation 2.26) is performed by unitary transformations of fields U, D and E (Equation 2.26) such that $U'_{L,R} = A_{U_{L,R}} U_{L,R}$, $D'_{L,R} = A_{D_{L,R}} D_{L,R}$ and $E'_{L,R} = A_{E_{L,R}} E_{L,R}$. Since in general $A_{U_L} \neq A_{D_L}$, the unitary matrix (CKM matrix) defined as $\mathbf{V} = \mathbf{A}^+_{D_L} \mathbf{A}_{U_L}$ is not unity and has four parameters (three mixing angles and one CP-violating phase).

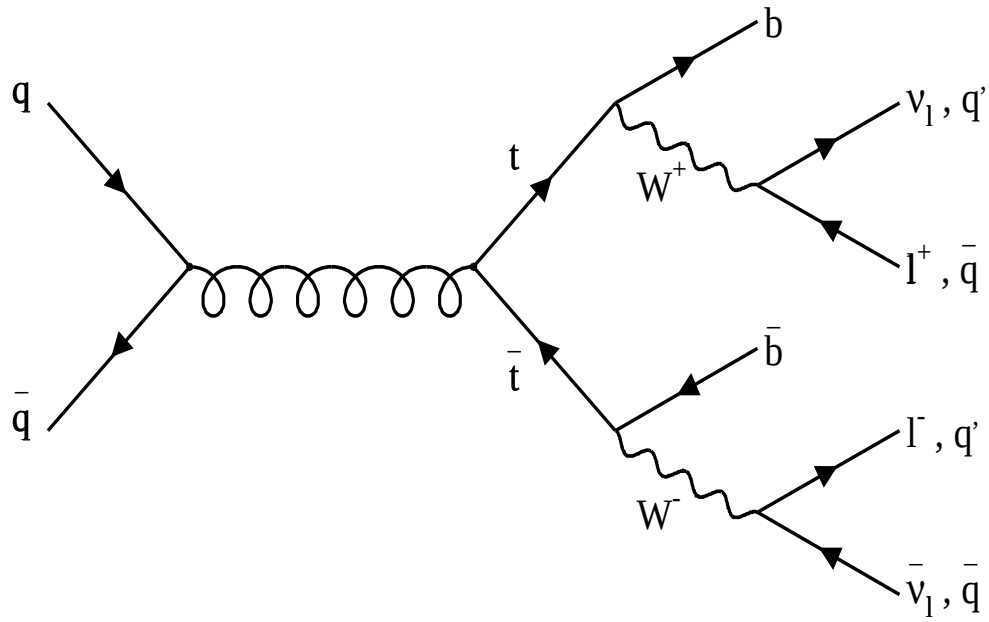

Figure 2.4: Tree level $q\bar{q} \to t\bar{t}$ production followed by the SM $t\bar{t}$ decay.

$$W^+ \to (e^+\nu_e)(\mu^+\nu_\mu)(\tau^+\nu_\tau)(u\bar{d})(c\bar{s}) \tag{2.31}$$

$$W^- \to (e^-\bar{\nu_e})(\mu^-\bar{\nu_\mu})(\tau^-\bar{\nu_\tau})(\bar{u}d)(\bar{c}s) \tag{2.32}$$

Each of the hadronic decay pairs can appear in three different color combinations $(R\bar{R}, G\bar{G}, B\bar{B})$ giving nine final state degrees of freedom for each W decay. To a good approximation, each possible decay of the W boson is equally probable.

The $t\bar{t}$ decay signatures are classified into three categories. The *dilepton* category represents the case in which both W bosons decay leptonically, the *lepton + jets* category is the case in which one W decays leptonically and the other hadronically, and the *all-hadronic* category in which both W decay hadronically. All possible $t\bar{t}$ decays are summarized in Table 2.4.

τ leptons are difficult to identify. A τ will decay into an electron or a muon about 36% of the time, and the only observable difference between the decays $W \to$

Category	Decay mode	Branching Ratio	
Dilepton	$t\bar{t} \to e\nu b e\nu \bar{b}$	1/81	4/81 (5%)
	$t\bar{t} \to \mu\nu b \mu\nu \bar{b}$	1/81	
	$t\bar{t} \to e\nu b \mu\nu \bar{b}$	2/81	
	$t\bar{t} \to e\nu b \tau\nu \bar{b}$	2/81	
	$t\bar{t} \to \mu\nu b \tau\nu \bar{b}$	2/81	
	$t\bar{t} \to \tau\nu b \tau\nu \bar{b}$	1/81	
Lepton + jets	$t\bar{t} \to q\bar{q} b e\nu$	12/81	24/81 (30%)
	$t\bar{t} \to q\bar{q} b \mu\nu$	12/81	
	$t\bar{t} \to q\bar{q} b \tau\nu$	12/81	
All-Hadronic	$t\bar{t} \to q\bar{q} b q\bar{q} \bar{b}$	36/81	36/81 (44%)

Table 2.4: $t\bar{t}$ decay modes and their associated branching ratios.

$\ell\nu$ and $W \to \tau\nu\ell\nu\nu\nu$ is the softer lepton in the latter case. The identification of hadronic τ decays suffers from large QCD backgrounds. Taking this into account it is useful to treat the signatures with W bosons decaying to $\tau\nu$ separately from other leptonic W decays.

The subject of this thesis is the top dilepton decay chain with final decay products including two leptons ee, $e\mu$ and $\mu\mu$. This channel is the purest with respect to signal over background, although has the smallest branching ratio $\sim 7\%$ (including contribution from leptonic τ decays).

2.3 Beyond the Standard Model

Although the Standard Model provides a remarkably successful description of presently known phenomena, it is very disappointing that the SM has a large

number of arbitrary parameters, such as three independent gauge couplings, six quark and three charged-lepton masses, three Cabibbo weak mixing angles and the CP-violating Kobayashi-Maskawa phase, as well as two independent masses for weak bosons.

It is commonly believed though, that a symmetry governs the world, and the gauge couplings of quarks and leptons are likely to reflect a fundamental structure with a simple description but with a complexity in dynamics. Pursuing this idea, grand unified theories (GUT) postulate that all interactions are unified into one single interaction following a single symmetry group at energies above the unification scale. While at lower energies this symmetry is broken into the $SU(3)_C \times SU(2)_L \times U(1)_Y$.

Such a more fundamental theory would provide solutions to questions that the Standard Model is not capable to answer, such as why are there so many different types of quarks and leptons, why do their weak interactions mix in the peculiar way observed, what is the origin of particle masses?

The main principle, which motivates to invoke physics beyond the Standard Model, is the search for the mechanism of the electroweak spontaneous symmetry breaking. The Standard Model Higgs boson is not theoretically satisfactory. Notice, that it is exactly because of the Higgs boson we have the multiplicity of undetermined coupling constants.

The Standard Model Higgs suffers from problems with the radiative corrections to its mass. Three diagrams contributing to the self energy of the Higgs boson are shown in Figure 2.5. Each of these corrections is quadratically divergent $\delta m_H^2 \simeq \mathcal{O}(\alpha/\pi)\Lambda^2$ [11], and consequently $\delta m_W^2 \sim \Lambda^2$, which is $\gg m_W^2$, if the cutoff Λ^2 comparable to the Planck scale M_{Pl}^2. One would need extreme fine-tuning to cancel a very large bare Higgs mass against huge loop corrections in

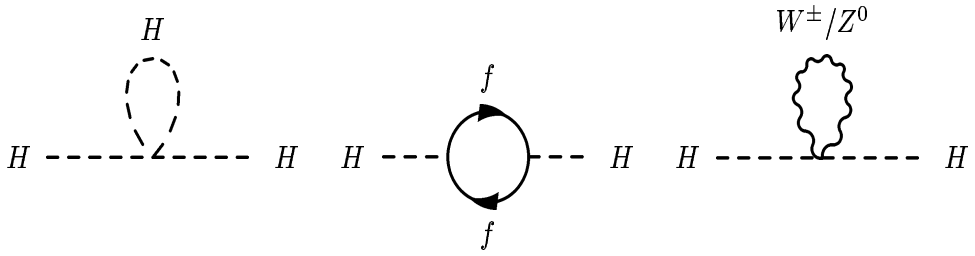

Figure 2.5: Scalar, fermion and vector boson loop corrections to the Higgs mass.

order to keep M_H at the weak scale ~ 1 TeV [12]. This strikes physicists as rather unnatural and they would prefer the corrections to be comparable to the physical value. This problem is known as a hierarchy problem.

Since creation of the Standard Model various extensions of the SM were proposed. Many of them provide what seems to be a natural solution to the hierarchy problem. Others come up with alternative to the Higgs mechanism of electroweak symmetry breaking. The most popular theories include supersymmetry (SUSY), technicolor, leptoquarks. Large resonance was recently received by the theories of extra dimensions.

It is cumbersome, however, that majority of the new theories introduce even a larger number of arbitrary parameters and by providing answers to existing questions in fact invoke many new ones. None of the current theories attempting to describe physics beyond the Standard Model has obtained yet a direct experimental evidence confirming its validity. It is quite possible that Nature will surprise us again.

Having all of this in mind, we developed an analysis not bounded by the framework of a particular new physics model. Instead this search embraces flexibility permitting Nature to give us an insight on what the new physics candidate is likely to be.

Chapter 3

Experimental Apparatus

3.1 The Tevatron Collider

The Tevatron proton-antiproton collider is currently the highest energy particle accelerator in the world. It derives its name from the "tera electron volt", or TeV (10^{12} electron volts), accentuating the amount of energy of the colliding beams in the machine. In the new data taking period started in 2001 (Run II) the protons and antiprotons reach 980 GeV. Their collisions result in a center-of-mass energy $\sqrt{s} = 1.96$ TeV.

The Tevatron is a superconducting magnet synchrotron. The accelerator is housed in a circular tunnel with a radius of 1 km. The magnets are made up of a niobium/titanium alloy that need to be kept extremely cold to remain superconducting. This allows the increased magnetic fields possible without fear of damage related to excessive resistive heating. All dipoles, quadrupoles, and correction element magnets are cryogenically cooled to about 4.6 K with liquid Helium. Before entering the Tevatron, proton-antiproton acceleration takes place over several stages. It is schematically represented in Figure 3.1.

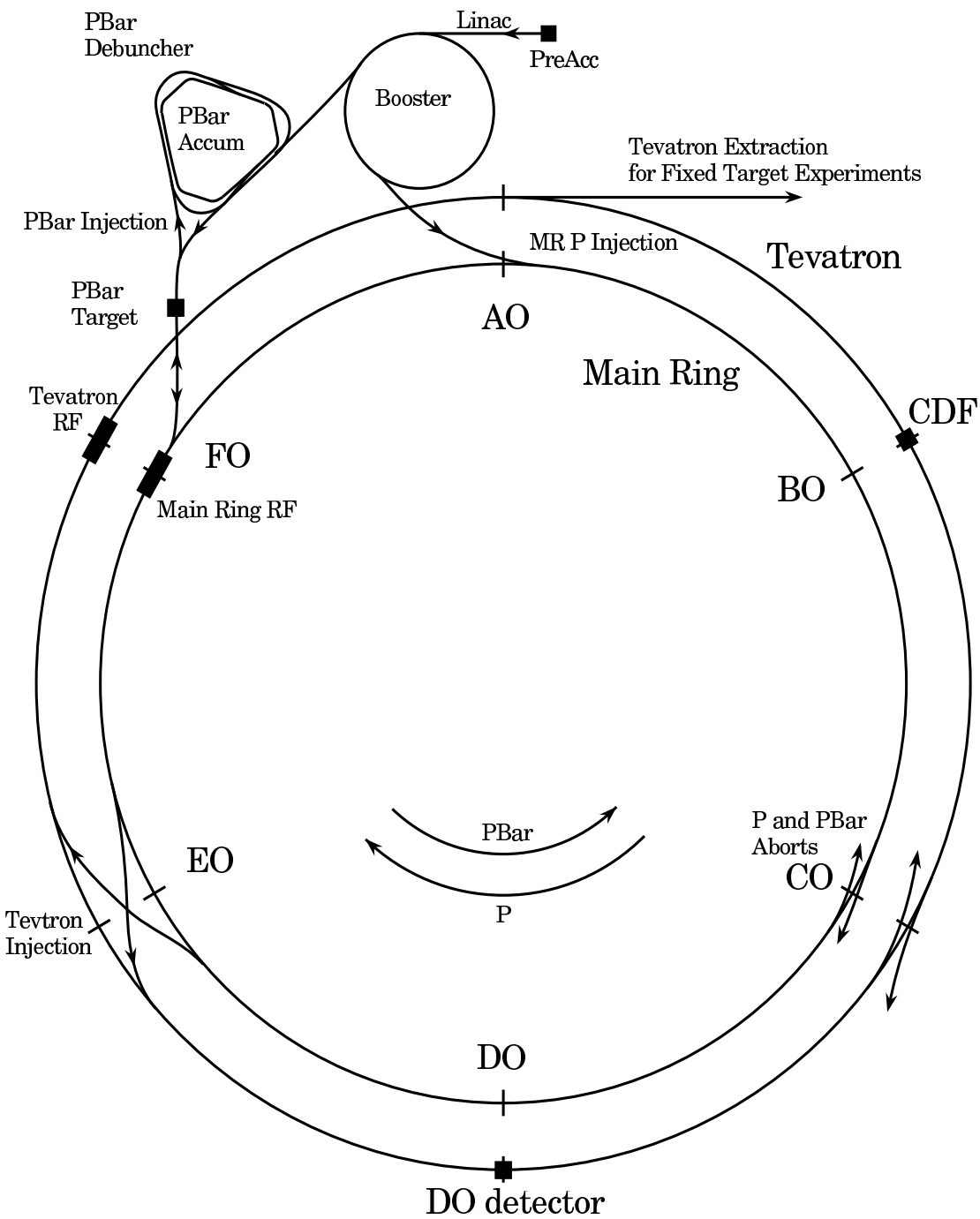

Figure 3.1: The accelerator complex.

The hydrogen gas H_2 stored in an electrically charged dome at a potential of -750 kV is ionized to H^- and then allowed to accelerate to the grounded wall to an energy of 750 keV. Then negatively charged ions enter the Linear Accelerator or "Linac", which through the 150 meter long line of the RF cavities forms a continuous stream of particles into discrete packets (bunches), and accelerates them to an energy of 400 MeV. After that the ions pass through a thin copper foil, which strips the electrons off and the protons enter the next level of acceleration, the Booster. The Booster is the first circular accelerator, or synchrotron. It consists of a series of magnets arranged around a 75-meter radius circle and accelerates protons to 8 GeV through repeated acceleration cycles. It further loads them into the Main Injector, the next synchrotron, with seven times the circumference of the Booster.

The Main Injector accelerates protons from the Booster to either 120 GeV or 150 GeV, depending on their destination. The 150 GeV protons are coalesced into a single bunch and injected into the Tevatron. The 120 GeV protons are used for creating and storing antiprotons, called *stacking*. The protons are transported to a target area, and focused to a nickel target. By striking the target they produce a spray of various secondary particles. Among those, antiprotons are focused with a Lithium lens, and are collected by a pulsed magnet as a charge-mass spectrometer. These antiprotons are further transferred into the Debuncher.

The Debuncher is one of the two synchrotrons that make up the Antiproton Source. It has a rounded triangular-shape with a radius of 90 meters. It accepts the high momentum spread antiprotons from the target station and converts them into a beam with an energy of 8 GeV. This is accomplished by the stochastic cooling process.

The beam is then transferred to the Accumulator, which is the second syn-

chrotron of the antiproton source. It is also a triangular-shaped synchrotron of radius 75 meters, and is housed in the same tunnel as the Debuncher. It is a storage ring for the antiprotons with a number of cooling systems.

The act of extracting antiprotons from the Antiproton Source is referred to as a *shot*. *Shot setup* is the time before actually transferring the antiprotons when the various transfer lines are tuned up with protons to ensure the least amount of losses so as to preserve emittance. The proton bunches are injected one at a time with each bunch 396 ns apart. There are 3 "trains" of 12 coalesced bunches and between each train is an abort gap of 2.617 μ sec. When desired proton bunch intensity and emittance are achieved, the antiproton bunches are loaded into the Tevatron. Their path to maximum energy is similar to that for the protons. They are first loaded into the Main Injector where they reach an energy of 150 GeV and then into the Tevatron in the opposite direction to the protons. Each antiproton bunch is injected similarly to the proton bunches, 396 ns apart.

The Tevatron is the largest of the Fermilab accelerators. It ramps up protons and antiprotons from 150 GeV to 980 GeV. When they reach desirable energies, they collide at the designated bunch crossing points. At one of those points named B0, the CDF detector is located.

In Collider mode the Tevatron can store beam for hours at a time. The collision rate of interaction within the detector is given by

$$R = \sigma_{int}\mathcal{L}, \tag{3.1}$$

where $\mathcal{L}$ is the instantaneous luminosity dependent on the revolution frequency f and the area A that the beam occupies. If N_p and $N_{\bar{p}}$ are the number of particles in each bunch and n is the number of bunches in either beam, then the luminosity $\mathcal{L}$ can be expressed by

$$\mathcal{L} = \frac{fnN_pN_{\bar{p}}}{A} \tag{3.2}$$

This stable situation of 980 GeV proton and antiproton collisions is called a *store*. In Run II the Tevatron operates with 36 proton on 36 antiproton bunches and with 396 ns bunch spacing. The typical luminosity at the beginning of the store during first two years of Run II was $\sim 2 - 4 \times 10^{31} cm^{-2}s^{-1}$. The store luminosity continually decreases from its initial value as protons and antiprotons are consumed through interactions and as the bunch emittance increases. This effect in the beginning of the store is largely dominated by emittance growth due to intrabeam scattering, while after several hours of running the effect of antiproton loss due to luminosity becomes more important and the luminosity falls off exponentially. In about 20 hours the luminosity drops too low, then the store is ended and the Tevatron prepared for new beam. This sequence of the stores continues 24 hours a day except for some short periods of time allocated for maintenance.

The integrated luminosity $\int \mathcal{L}dt$ is the measure of number of collisions during a time period. This analysis corresponds to the integrated luminosity of 193 pb^{-1} during the first two years of the Run II data taking. This is almost twice as large as that collected during Run I.

3.2 The Collider Detector at Fermilab

The CDF [13] detector is a miltipurpose detector designed to observe a wide range of physics processes produced during high-energy $p\bar{p}$ collisions. It combines charged particle tracking with projective calorimetry and muon detection. The detector electronics has been significantly upgraded for Run II to accommodate

large instantaneous luminosities and to improve the performance in momentum resolution, b-tagging, electron and muon identification, etc.

The CDF has both azimuthal and forward-backward symmetry about the transverse plane passing through the interaction point, which results in the naturally arising cylindrical coordinate system. The coordinate system is right-handed with the $\hat{z}$-axis pointing in the proton momentum direction and the $\hat{y}$-axis in the upward vertical direction. It is more convenient to identify locations of the particles in $\phi - \eta$ space, where ϕ is the azimuthal angle and η is the pseudorapidity expressed in terms of the polar angle θ as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.3}$$

Since the pseudorapidity is the Lorentz invariant quantity

$$\eta = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) + \mathcal{O}(\epsilon^2) \tag{3.4}$$

where $\epsilon = \frac{m}{p_T}$, a measure of the opening angle between two particles given by ΔR:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.5}$$

is also Lorentz invariant.

The CDF detector consists of the three main functional sections going radially outwards from the beamline.

- The tracking system is used for particle charge and momentum measurements. It is immersed in a superconducting solenoid of 4.8 m length and 1.5 m radius which produces a 1.4 T magnetic field coaxial with the $p\bar{p}$ beams.

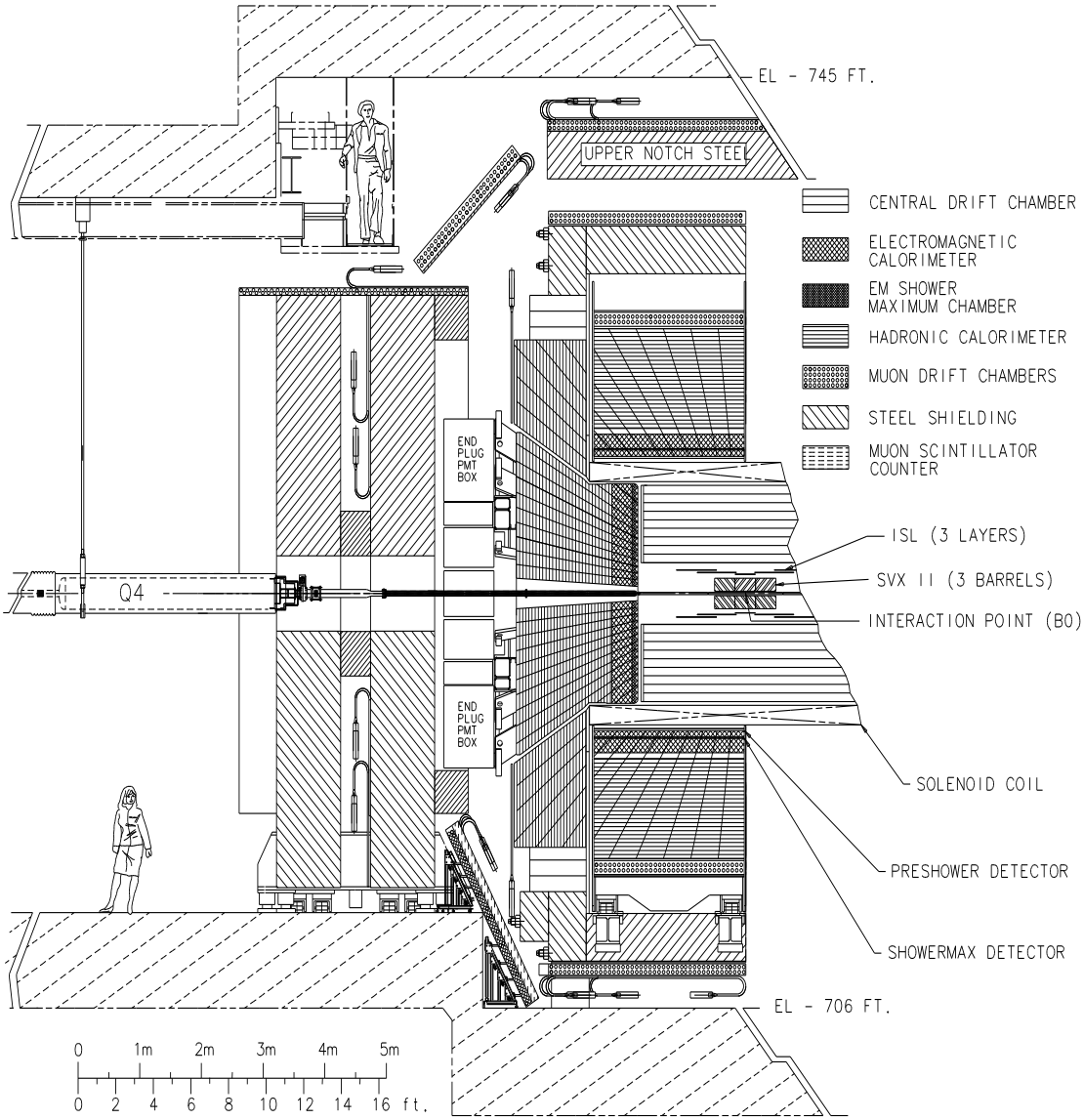

Figure 3.2: Elevation view of one half of the CDF detector.

- The solenoid is surrounded by the scintillator-based calorimeter system with separate electromagnetic and hadronic measurements, which covers the region $|\eta| \leq 3$.
- Outside of the calorimeters, layers of steel absorb the remaining hadrons leaving only muons, which are detected by the outermost muon detectors.

In the next sections we discuss the functionality of main CDF detector components.

3.2.1 The Cherenkov Luminosity Monitor

The beam luminosity (Equation 3.2) measurement is the crucial for the entire experiment. It is determined from the rate of inelastic $p\bar{p}$ interactions, called *minimum bias* events. The measurement is obtained with low pressure gaseous Cherenkov counters [14] placed in the forward and backward region at small angles $\theta \leq 3°$ relative to the beam direction, as shown schematically in Figure 3.3. The 48 thin, long, conical counters are located on each side of the detector. The counters are arranged around the beam pipe in three concentric layers with 16 counters, each oriented with their small end pointing to the center of the interaction region, and cover $3.7 < |\eta| < 4.7$ pseudorapidity range. The cones in the outer two layers (further away from the beam pipe) are about 180 cm long. The inner layer counters are shorter, about 110 cm, due to geometrical constraints. At the large aperture of the cones, furthest from the interaction region, the alluminium conical light collectors are attached. At the small end of the light collectors 2.5 cm diameter photomultiplier tubes (PMT) are placed. The completed structure is enclosed in a vessel filled with isobutane, which is used as a radiator[1].

[1]Isobutane has one of the largest indexes of refraction at atmospheric pressure for commonly available gases, $n = 1.00143$, and good transparency for photons in the ultra-violet region where

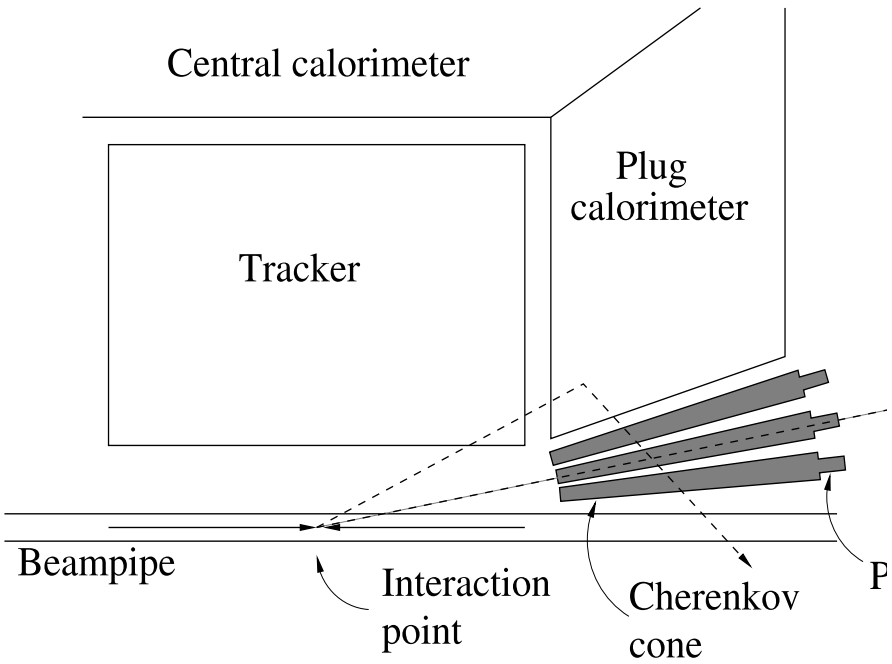

Figure 3.3: Schematic view of luminosity monitor inside a quadrant of the CDF detector.

The Cherenkov light emission angle θ_C is determined by the gas refraction index n and the particle velocity β:

$$\cos\theta_C = \frac{1}{n\beta} \tag{3.6}$$

The number of photoelectrons produced by a charged particle $N_{p.e.}$ in Cherenkov counters is proportional to the length of particle's path L inside the counter and to $\sin^2\theta_C$ [16]:

$$N_{p.e.} = N_0 \cdot L \cdot sin^2\theta_C, \tag{3.7}$$

where $N_0 \sim 200cm^{-1}$ is the counter design specific parameter.

most Cherenkov light is emitted [15]

The prompt particles from $p\bar{p}$ interactions traverse the full length of the counter and generate a large amplitude PMT signal, ~ 100 photoelectrons. While the signals from particles originating from beam-halo interaction or from secondary interactions in the detector material are significantly smaller because those particles have lower momenta and they traverse the counters at larger angles with shorter path lengths, hence their light experiences large losses due to the reflections.

The high precision of the luminosity monitor is provided due to the fact that the counters effectively measure the actual number of primary particles. If two particles pass through a single counter the resulting signal is twice that of a single particle. Therefore the Cherenkov monitor does not saturate at high operational luminosities in the Tevatron.

3.2.2 The Silicon Tracking System

The CDF tracking system consists of a series of concentric semiconductor and gaseous detectors. At large radii the charged-particle tracking in the central pseudorapidity region ($|\eta| < 1.0$) is done with a large open cell cylindrical drift chamber, The Central Outer Tracker (COT). Inside the COT a silicon micro-vertex detector surrounding the beryllium beam pipe establishes the ultimate impact parameter resolution and provides stand-alone silicon tracking in the region of $|\eta| < 2.0$

The silicon detector is comprised of eight layers of microstrip silicon sensors arranged in cylinders spanning radii from 1.35 cm to 28 cm, and lengths from 90 cm to nearly two meters. It is divided into three sub-systems: L00, SVXII and ISL with a total of six square meters of silicon and 722,000 readout channels.

A schematic view of the principal active components of the CDF silicon system is given in Figures 3.4 and 3.5. The side view shown in Figure 3.4 is a cross-section

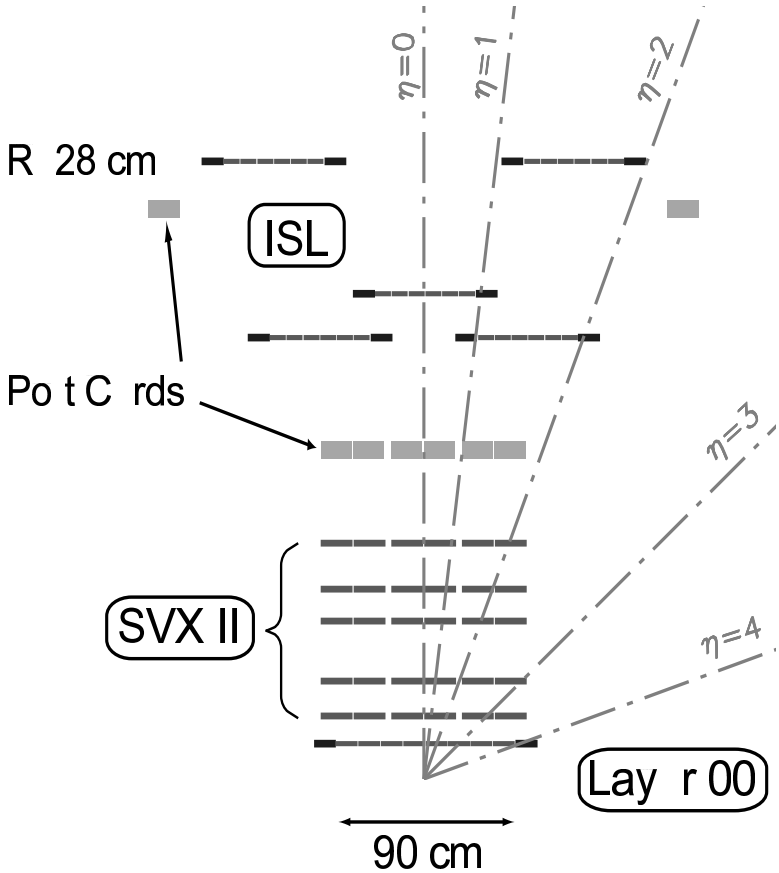

Figure 3.4: A side view of half of the CDF silicon system on a scale in which the $\hat{z}$ coordinate is highly compressed.

of one half of the silicon tracker, using a compressed z scale. Figure 3.5 shows an end view of the CDF silicon system including the SVXII bulkheads and ISL support frame.

The L00 ("Layer Zero Zero") [17] detector is the innermost single-sided layer of low-mass silicon-microstrips mounted directly over the beam pipe at a radius of $\sim$ 1.6 cm from the beamline with a total length of 80 cm. Due to its proximity to the beamline it substantially improves the impact parameter resolution. The L00 utilizes radiation tolerant axial strip sensors expected to last at least 7.4 fb^{-1}. Radiation resistance is achieved with a guard structure designed to minimize

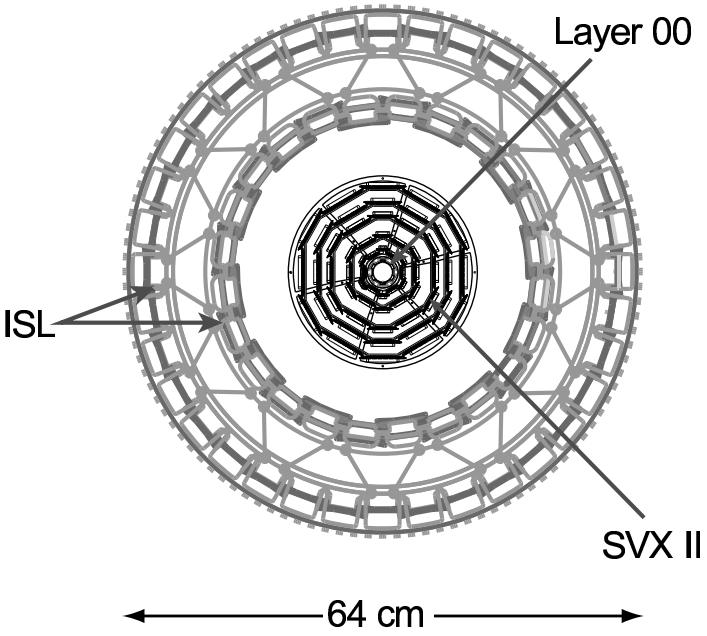

Figure 3.5: An end view of the CDF silicon system including the SVXII cooling bulkheads and ISL support structure.

leakage currents. The sensors have an implant pitch of $25\mu m$ and a readout pitch of 50 μm achieved by reading out alternate strips. The L00 is connected to electronics outside the tracking volume at $|z| > 40$ cm via fine-pitch kapton cables carrying the signals from the sensors to the readout chips.

The next five double-sided layers of the tracker positioned at radii from 2.45 to 10.6 cm comprise the Silicon Vertex Detector (SVXII) [18]. The layers of the SVXII and L00 are arranged in twelve azimuthal wedges that alternated in radius within each layer (see Figure 3.6). The SVXII is 96 cm long and is assembled in three cylindrical barrels with beryllium bulkheads at each end. A printed circuit board, called the Port Card (PC), located around the periphery of the bulkheads interfaces the hybrids and frontend chips with the rest of the readout data acquisition system by translating the readout into optical format for transmission outside the tracking volume.

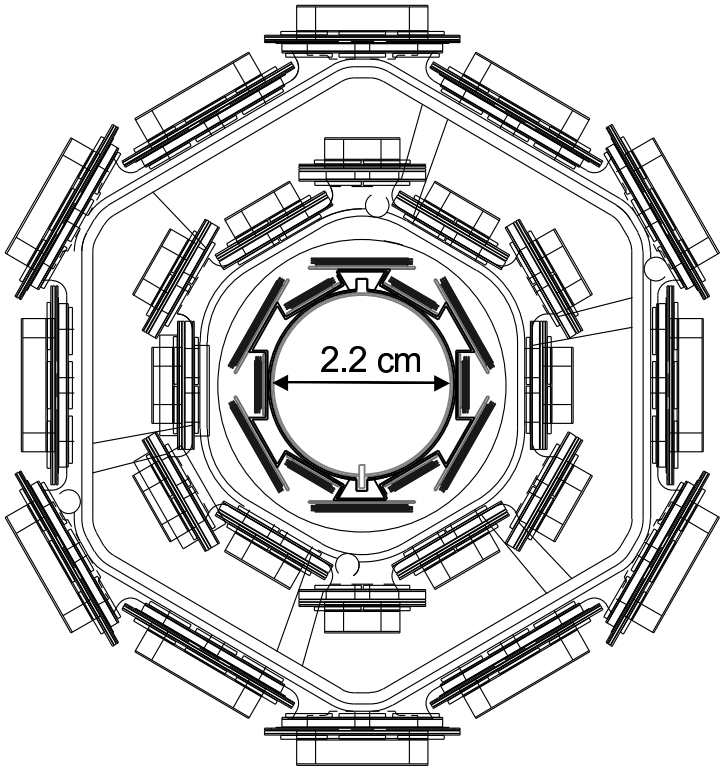

Figure 3.6: End view of the innermost three layers of the CDF silicon system, showing Layer 00 along with the first two layers of the SVXII region. The Layer 00 electronics (not shown) are mounted beyond the active volume for vertexing. The SVXII electronics are shown just outside and just inside of each of the layers drawn.

The basic structural unit of the SVXII detector is called a *ladder*, which consists of two readout units of silicon with an electrical hybrid at each end. The units are mounted directly atop the silicon surface to avoid gaps and to improve readout speed and capacitance limitations.

Both 90-degree and small-angle stereo sensors are used in the SVXII, in the pattern $(90°, 90°, -1.2°, 90°, +1.2°)$ for the n-strips from the innermost to outermost SVXII layers. SVXII sensors are AC coupled, 300 μm thick, and made from n-type high resistivity bulk silicon. They are biased using polysilicon resistors.

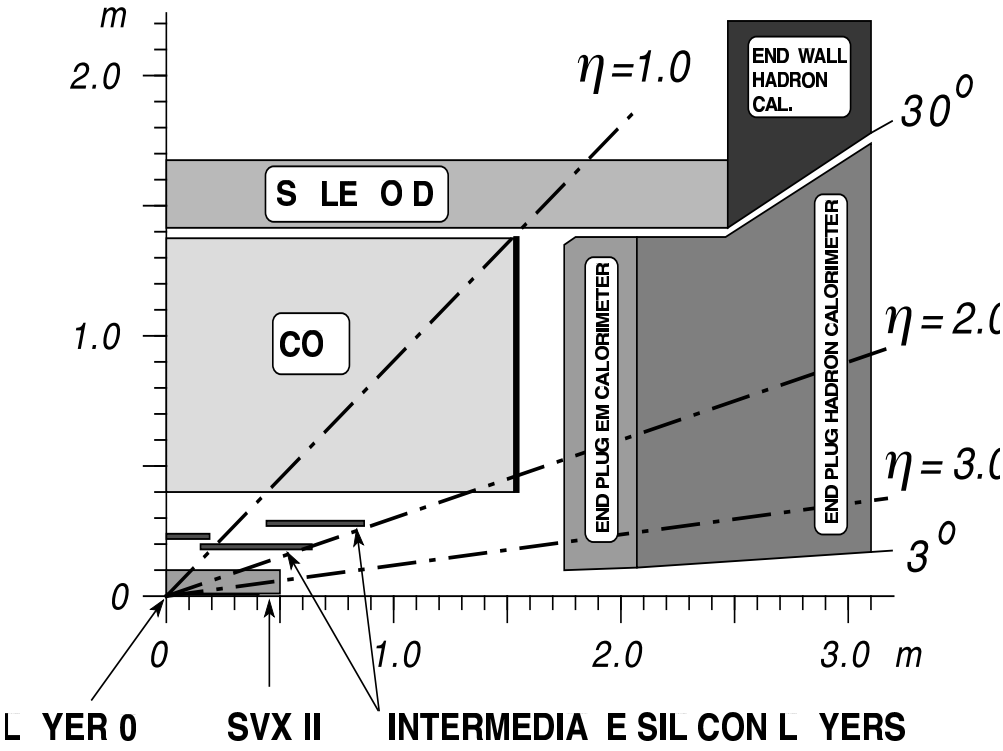

Figure 3.7: A cutaway view of one quadrant of the inner portion of the CDF detector showing the tracking region surrounded by the solenoid and endcap calorimeters.

The p-strips on the non-stereo side run in the axial direction of the detector are used to measure the azimuthal angle ϕ of the track. These strips are spaced in $r\phi$ by 60 - 65 microns, depending on the ladder, and have implant widths of 14 to 15 microns. The stereo n-strips of the SVXII are spaced by (141, 125.5, 60, 141, 65) microns, and have implant widths of 20 microns for the 90° strips and 15 microns for the small-angle stereo layers. The 90° layers have an additional layer of insulator and readout strips in the double-metal configuration; these strips carry the $\hat{z}$ signals to the readout chips.

The Intermediate Silicon Layers (ISL) detector is placed outside of the SVXII. Its space frame also supports the SVXII and all associated readout and utility

components. The ISL consists of two symmetric silicon layers in the forward and backward region ($|\eta| \geq 1.1.$) located at radii of $R \simeq 20$ cm and $R \simeq 29$ cm respectively and one in the central region ($|\eta| < 1.1.$) at $R \simeq 23$ cm. It provides one spacepoint in the central region which improves the linking between SVXII tracks and COT tracks and its fine granularity helps to resolve ambiguities in dense track environments. In the forward region, where the COT acceptance rapidly decreases, the ISL together with the SVXII constitutes a standalone 3D tracker out to $|\eta| \simeq 2.0$ (see Figure 3.7).

The ISL also utilizes AC coupled double-sided sensors with polysilicon biasing and common p-stops. Due to its large radial position radiation damage occurs more slowly and the hit occupance is lower, it is therefore possible to use longer strips and a large readout pitch to reduce the number of readout channels. A fixed strip pitch of 112 μm is used on both the axial and 1.2-degree stereo sides. Pitch adapters are used to bring the signals from the strips to the more closely spaced inputs of the readout chips.

The ISL ladders are composed of six sensors, arranged as half-ladders of three sensors each. Silicon sensors are bonded together to form a single electrical unit and glued on a carbon fiber support. At each ladder end they are read out with double-sided hybrids that extend beyond the silicon.

All components of the silicon system achieve their data readout by the SVX3D [19] chip, a radiation-hard CMOS custom integrated circuit (IC) device. Each IC has 128 parallel analog inputs and 8-bit digital output bus. Each channel contains a preamplifier, an analog delay pipeline, ADC (8 bits), and data sparsification logic. In order to maximize the physics potentiality of the CDF silicon detectors, the SVX3D has been designed to work in continuous (deadtimeless) mode, being capable of simultaneous acquisition, digitization and readout operation.

The silicon is actively cooled to remove heat due to SVX3D chip power as well as that due to leakage currents after irradiation. Water/ethylene glycol coolant mixture of - 5° C for the silicon and electronics flows within internal channels that are machined into the SVXII beryllium bulkheads at each barrel end, and in alluminium tubes attached to beryllium ledges mounted on the ISL space frame.

3.2.3 The Central Outer Tracker

The Central Outer Tracker (COT) is located outside the silicon microstrip detectors within a 1.4 T solenoidal magnetic field. It is designed to find charged tracks in the central region $|\eta| < 1.0$ with transverse momentum p_T as low as 400 MeV, and link tracks to hits in the inner silicon detectors.

The active volume of the COT spans 310 cm in the beam (axial) direction z, and between 44 cm and 132 cm in radius. It is filled with Argon-Ethane-CF_4 (50:35:15) mixture bubbled through isopropyl alcohol. This gas mixture provides a fast drift velocity $\sim 100\mu$m/ns that is essential for lowering a maximum drift time. The maximum drift time is required to be less than the 396 ns bunch spacing, and with this gas it is about 100 ns in the drift field ~ 2 kV/cm. This makes the COT immune to event pile-up, even at the highest collision rate of 1/(132 ns).

The COT is comprised of 30,240 sense wires running the length of the chamber and strung between two precision-machined alluminium end plates. The sense wires are grouped into eight *superlayers*, shown in Figure 3.8, that alternate axial-stereo with a stereo angle of $\pm 2°$. Axial superlayers provide accurate tracking information in the $r - \phi$ view. Stereo superlayers provide tracking information in the $r - z$ view, substantially less accurate, and are designated to measure η of the track. Each superlayer is divided in ϕ into *supercells* and each supercell has 12

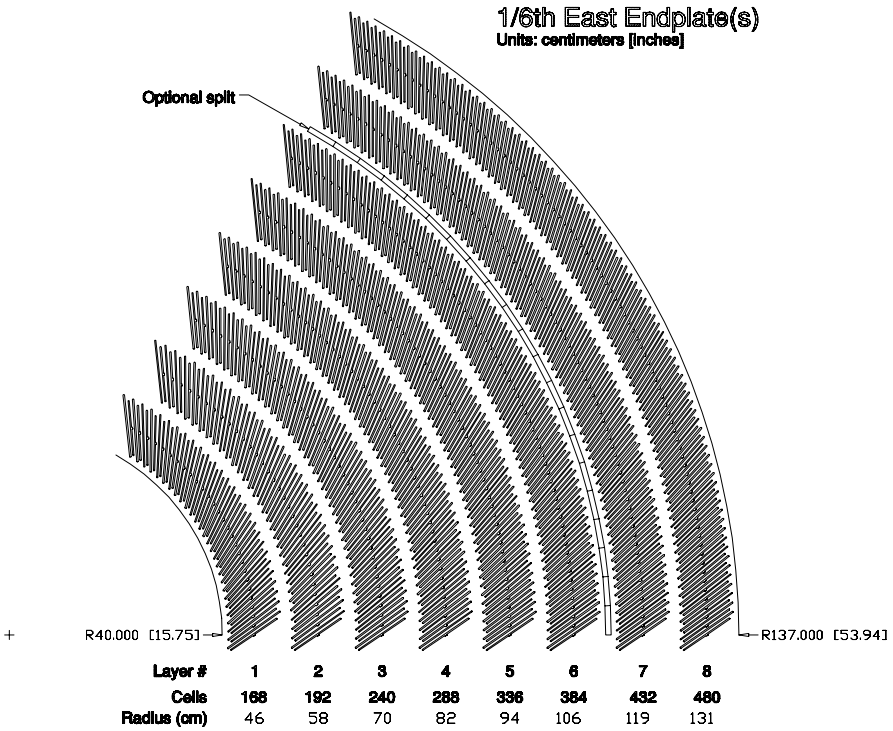

Figure 3.8: A view of a portion of one COT endplate. The slots define the radii of the 8 superlayers. Within a superlayer, wire-plane slots and field sheet slots alternate.

sense wires yielding total 96 measurement layers. A maximum drift distance is established to be approximately the same for all superlayers by scaling the number of supercells in a given superlayer with the radius.

Figure 3.9 shows the COT supercell structure, for superlayer 2. It consists of 12 sense wires and 13 potential wires that alternate. Potential wires are equispaced and sense wires are placed at the midpoint between each couple of successive potential wires. Four shaper wires serve to close the supercell electrostatically. A field sheet (cathode) is placed on both sides. Each field sheet is shared with the neighboring supercell. The supercell is tilted by 35° with respect to the radial direction to compensate for the Lorentz angle of the drifting charged particles in

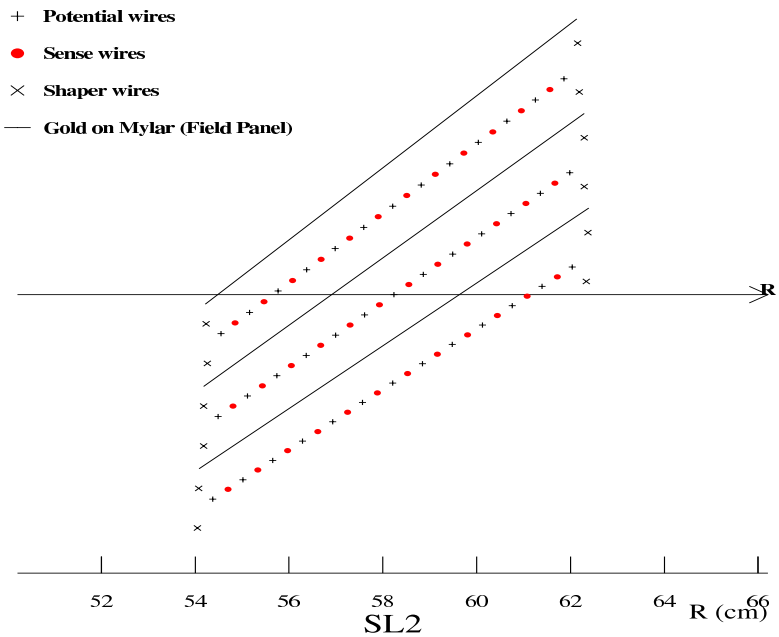

Figure 3.9: Three COT cells. Each cell has 12 sense wires, 13 potential wires, 4 shaper wires and a gold-mular cathode field panel on both sides of the sense wire plane.

the magnetic field.

The field sheet is maintained at ground potential. The potential wires are run at positive high voltage, $\sim 2kV$, and the sense wires are run at $\sim 3kV$. In order to maintain a uniform drift field actual voltages slightly differ (less than 400 V) with the variation across a cell for the sense and potential wire voltages.

Wire readout is performed via a custom-built ASD (amplifier, shaper, discriminator) radiation-hard 8-channel chip. ASD boards are placed directly on the chamber endcap. Pattern recognition logic allows multiple hits to be recorded within a single sense wire.

Particles of charge q moving in a uniform magnetic field $\vec{B}$, inside of the CDF tracker, have a helicoidal trajectory with the curvature radius

$$r = \left| \frac{p_T}{qB} \right| \tag{3.8}$$

One obtains the particle's transverse momentum p_T by reconstructing COT hits that define the particle trajectory and measuring the radius of curvature of helix r.

The momentum resolution of the COT is

$$\sigma(1/p) = 0.11\% \tag{3.9}$$

In addition, the COT provides particle identification information based on the ionization loss dE/dx measurement. To enhance particle identification capabilities, the time-of-flight detector is placed at the outer edge of the COT. It consists of an array of scintillator bars about three meters long, matching the COT active volume. Photomultiplier tubes attached to both ends of each bar provide time and pulse height measurements.

Particle identification is performed by measuring the time of arrival of a particle at the scintillator with respect to the collision time. The particle mass m can be then determined based on its momentum p, the path length L and the time of flight t.

$$m = \frac{p}{c}\sqrt{\left(\frac{ct}{L}\right)^2 - 1} \tag{3.10}$$

The time of flight t is measured with resolution $\sim$ 130 ps which provides a capability to distinguish between light $K^{\pm}$ and $\pi^{\pm}$ hadrons and search for new stable massive particles up to 500 GeV.

3.2.4 Calorimetry

Particles of interest with transverse momenta greater than 350 MeV escape the magnetic field, and are detected by the calorimeters outside of the solenoid. The calorimeters are separated into two regions. The central calorimeter provides coverage over $|\eta| < 1.1$ and the plug calorimeter covers the pseudorapidity region $1.1 < |\eta| < 3.4$, corresponding to polar angles between 37° and 3°. Each calorimeter consists of electromagnetic (EM) and hadronic (HAD) parts, both segmented into projective towers. A collision product meets first the EM part, and then traverses through the HAD part. The amounts of energy of a particle deposited in each part of the calorimeter are referred as an electromagnetic and hadronic energies respectively. The electrons and photons leave most of their energy in the EM part, while hadrons lose their energy predominantly in the HAD part of the calorimeter.

The central calorimeters are divided azimuthally into 24 wedges, each covering an azimuthal angle of 15° and extending about 250 cm along the beam axis on either side of $z = 0$. The central electromagnetic calorimeter (CEM) starts at a radius of 173 cm and is 35 cm thick, after that the hadronic calorimeter (CHA) begins. The segmentation in pseudorapidity forms calorimeter towers, which project back to the nominal interaction point. Each tower covers $\Delta\eta = 0.11$, with a CHA tower directly behind the matching tower in CEM. This allows one to measure the ratio of electromagnetic to hadronic energy for each individual tower. A schematic of a single central calorimeter wedge, showing both the CEM and the CHA, and the tower geometry is shown in Figure 3.10.

The calorimeters are made of alternating layers of absorber material with a high nuclear number Z and the active readout material. In the electromagnetic section the absorber is lead, and in the hadronic section it is iron. The polysterene

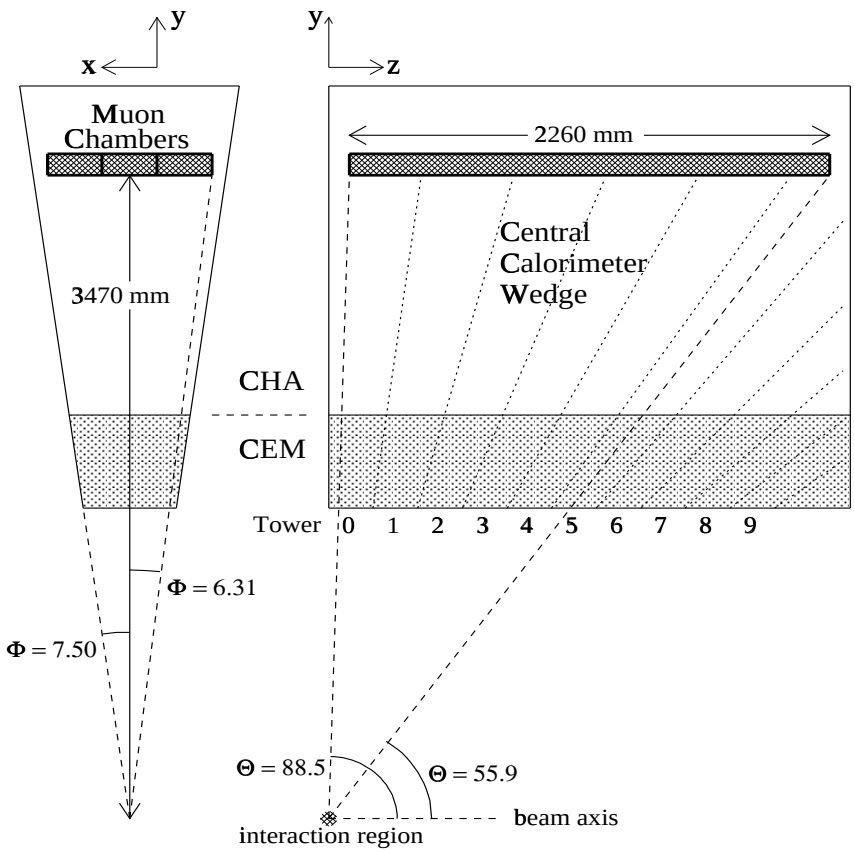

Figure 3.10: Geometry of a central calorimeter wedge and its towers.

scintillator serves as an active material. The CEM is composed of 5 mm thick layers of scintillator and 3 mm layers of lead, while for the CHA layers are thicker and arranged in 1.0 cm and 2.5 cm respectively.

As particles traverse through the absorber, they lose energy and produce cascades of secondary particles, *showers*, which then interact in the scintillators. The showers penetrate through many layers, and are sampled by the scintillators until they are completely absorbed. The scintillator's light is collected through acrylic light guides attached to photomultipliers, which are located at the rear end of each wedge, as shown in Figure 3.11. The amount of light is a measurement of the incident particle's energy.

The CEM thickness corresponds to 18 *radiation lengths* X_0, and its energy resolution is [20]:

$$\frac{\sigma_E}{E} = \frac{13.5\%}{\sqrt{E_T}} \oplus 2\%, \tag{3.11}$$

where $\oplus$ means that the constant term is added in quadrature. The first term comes from sampling fluctuations and the photostatistics of PMTs, and the second term comes from nonuniform response of the calorimeter.

The CHA is about 4.5 *attenuation lengths* λ_I thick, with the energy resolution given by [21]

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E_T}} \oplus 3\%. \tag{3.12}$$

Transverse development of the electromagnetic showers is different for electrons and hadrons. To enable a more precise measurement of the transverse profile a proportional strip and wire chamber, called the central electromagnetic shower counter (CES), is embedded in each tower of the central calorimeter at location where maximal average electromagnetic shower deposition occurs (5.9 X_0). The CES has cathode strips running in the azimuthal direction, which provide z information, and anode wires running parallel to the beamline, which provide $r-\phi$ information. In addition to the CES, the central preradiator detector (CPR) composed of proportional chambers is placed between the solenoid and the CEM. Both the CES and CPR help in distinguishing electrons from hadrons.

The plug calorimeter comprises towers with the same structure as in the central calorimeter. The electromagnetic section with shower position detection is followed by a hadronic section. The plug calorimeter is divided in 12 concentric η regions, which are further segmented in 24 for $|\eta| < 2.11$, or 12 for $|\eta| > 2.11$

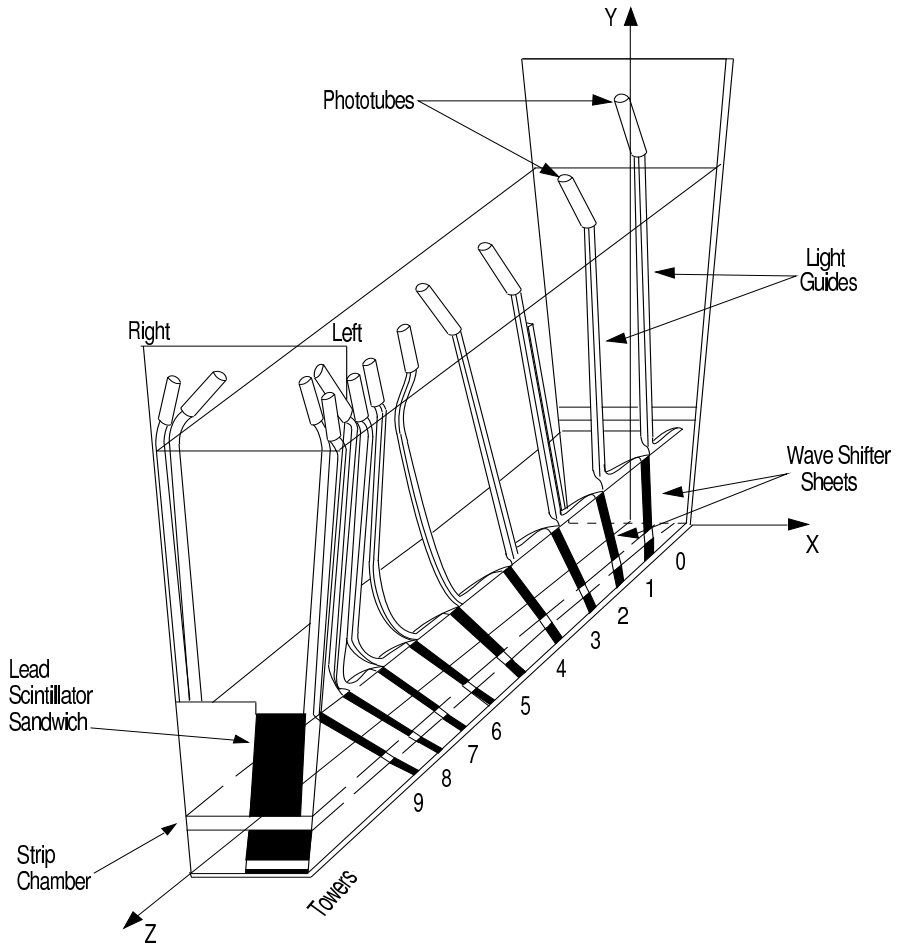

Figure 3.11: One of the wedges of the central electromagnetic calorimeter. The ten $(\Delta\phi, \eta) = (15°, 0.11))$ projective towers are shown.

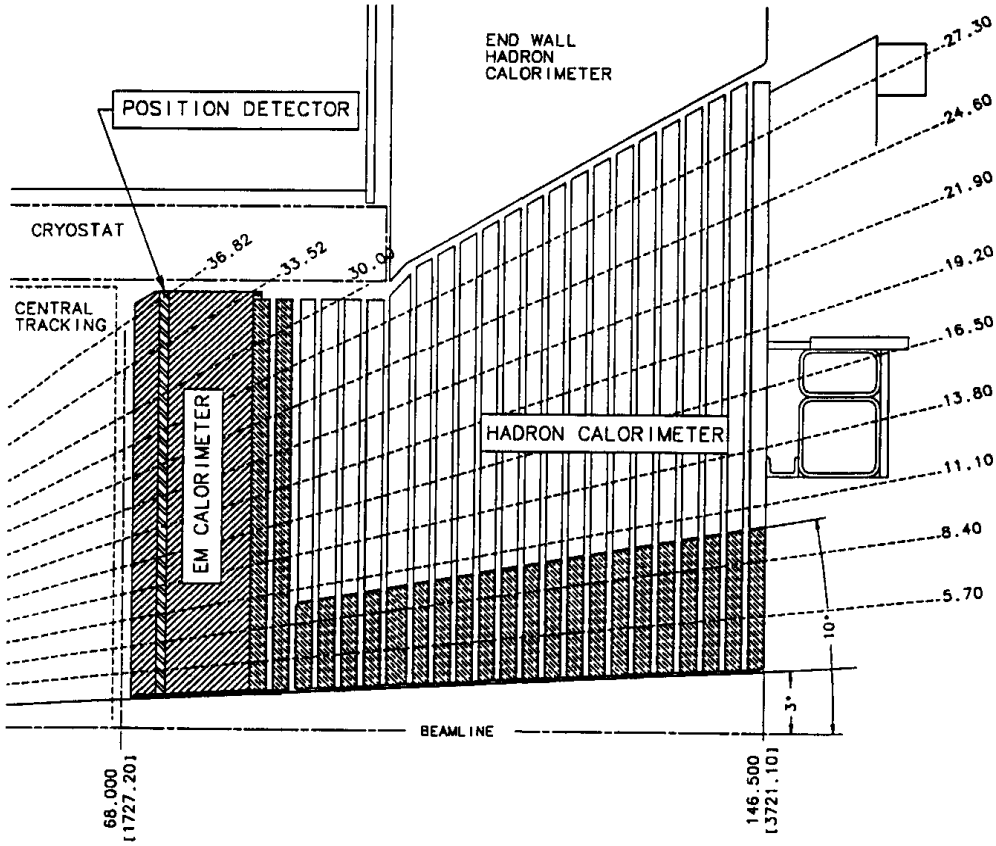

Figure 3.12: Cross section of an upper part of the plug calorimeter.

projective towers. The EM section of the plug calorimeter (PEM) consists of 23 absorber-scintillator layers. A calcium-tin-lead alloy enclosed between steel plates serves as an absorber. The total PEM thickness corresponds to 21 X_0, radiation lengths, and its energy resolution is [22]

$$\frac{\sigma_E}{E} = \frac{14.4\%}{\sqrt{E_T}} \oplus 0.7\%. \tag{3.13}$$

The first layer of the EM section is used as a preshower detector (PPR). Its structure is the same as the other EM layers except the first scintillator is thicker (10 mm instead of 4 mm) and is read out separately from the rest of the calorimeter

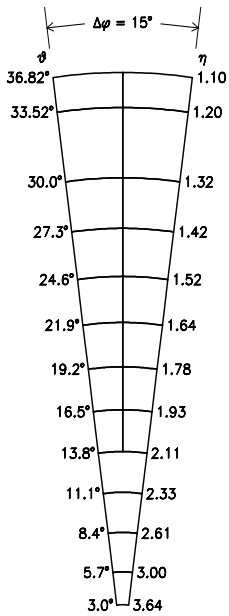

Figure 3.13: Tower segmentation of the plug calorimeter.

via multi-anode photomultiplier tubes (MAPMT). As in the central calorimeter, a shower maximum detector (PES) is also embedded in the plug EM section at a depth of $\sim 6X_0$, radiation lengths. It consists of eight 45° sectors, each covering six (or three) calorimetric towers in ϕ. Each sector is further segmented in two η regions to reduce occupancy. Within each region 5 mm wide scintillating strips are arranged in two layers in directions parallel to either edge of the sector. The two layers are denoted 'U' and 'V', and are being offset from the radial direction by $+22.5°$ and $-22.5°$ respectively. This provides a two-dimensional measurement of the shower with a position resolution of ~ 1 mm.

The PHA calorimeter is comprised of 23 unit layers composed of 5 cm iron and 6 mm scintillator. It is about 7 λ_I thick with the energy resolution

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E_T}} \oplus 5\%, \tag{3.14}$$

3.2.5 The Muon Detectors

After escaping the tracking system, muons interact minimally within the calorimeter and are detected with arrays of drift tubes in the outermost part of the detector. The CDF muon system has four separate detectors: CMU, CMP, CMX and IMU. The central muon chambers (CMU) and the central muon upgrade chambers (CMP) cover the pseudorapidity range $|\eta| < 0.6$. The central muon extension chambers (CMX) provide coverage at $0.6 < |\eta| < 1.0$. The intermediate muon detectors (IMU) are capable to identify muons in the forward region up to $|\eta| \simeq 2.0$. The ϕ coverage is not complete. The CMU alone covers 84% of the solid angle. The CMP covers 63%, and 53% is covered by both of these detectors. The CMX covers 71% of the solid angle. The complete $\phi - \eta$ coverage layout is shown in Figure 3.14.

The CMU is housed within the central calorimeter wedges, directly behind the CHA (see Figure 3.11). Each CMU wedge covers azimuthally 12.6° with a 2.4° gap between the wedges. Each wedge contains three muon towers and each muon tower consists of four radial layers of four rectangular drift cells. At the center of each drift cell a sense wire runs the length of the wedge (2260 mm). Pairs of sense wires are offset from each other by ~ 2 mm to provide unambiguous ϕ measurement by determining which sense wire was hit first. A track is measured with resolution of 250 μm in the $r-\phi$ plane. The z-position of the track is obtained by comparing the pulse heights at each end of the sense wires. The resolution in the $r - z$ plane is 1.2 mm. Tracks measured in at least 3 of the 4 layers form a track segment, called a *stub*.

Beyond the CMU there is an additional 60 cm thick steel shielding to further reduce the number of "punch-through" hadrons escaping the hadronic calorimetry. Behind the steel there are four additional layers of drift chambers, which make

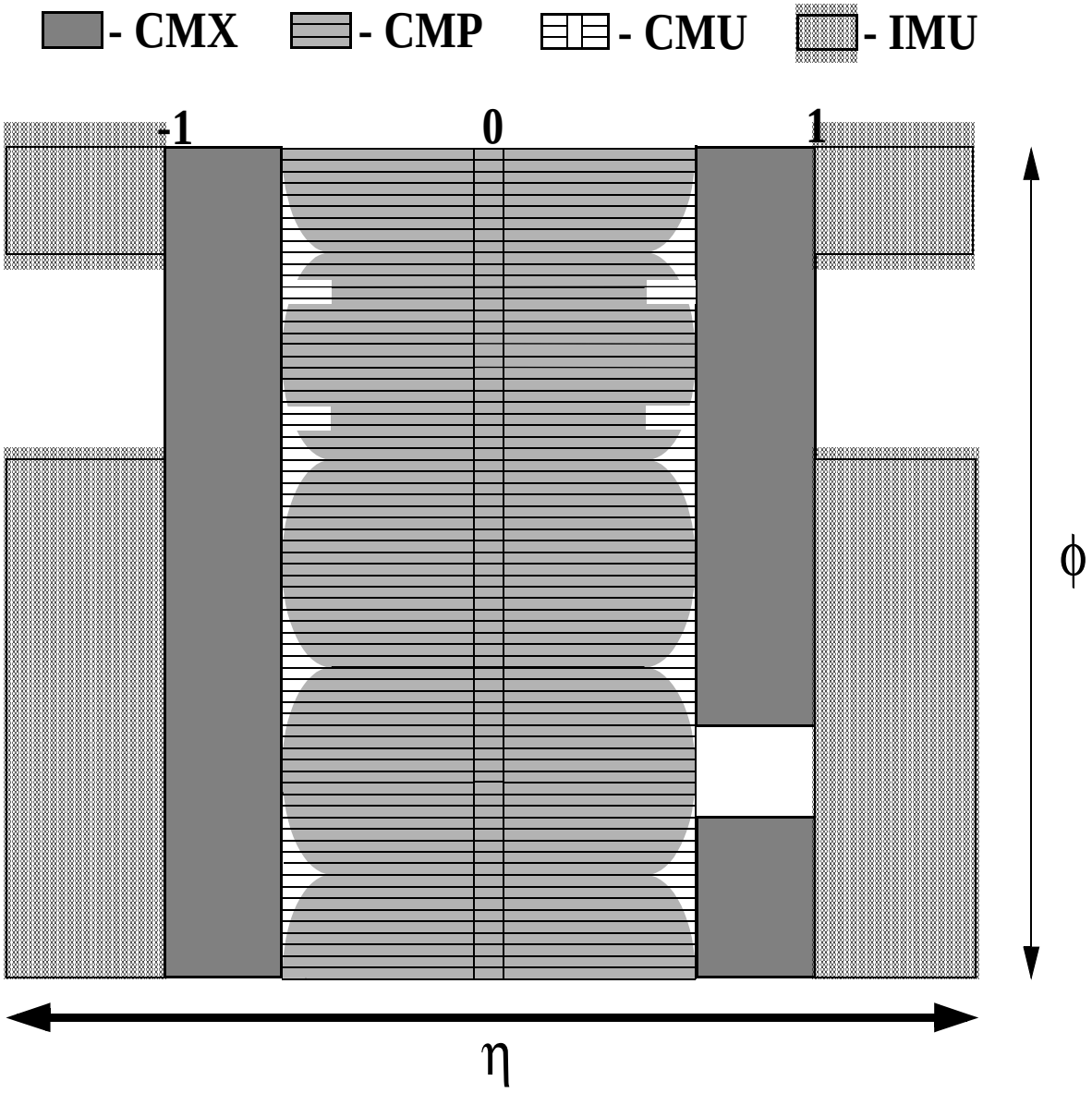

Figure 3.14: The $\eta - \phi$ muon coverage of the CDF detector.

up the central muon upgrade (CMP). The CMP operates very similarly to the CMU but contains only one anode wire per chamber, and the anode wires are not connected in couples, so that the CMP provides only the $r - \phi$ measurement of the track. Unlike the CMU chambers, which are arranged in towers, the CMP chambers are arranged in stacks positioned around the detector. The inner and outer surfaces of the CMP are lined with scintillator plates, called the CSP, used to provide timing information. The CMP partially compensate the CMU ϕ gaps but has gaps on its own in the regions $80° < \phi < 100°$ and $260° < \phi < 280°$ due to the return yoke of the solenoid.

The muon coverage is extended by additional muon chambers, which constitute the central muon extension (CMX). The CMX is comprised of four conical arches of drift tubes with layers of scintillator, called the CSX, analogous to the CSP. The CSX sandwiches the CMX drift tubes and helps in the identification of real muons.

Detection of muons in the forward region is accomplished by the IMU, which consists of four staggered layers of drift tubes and scintillator counters.

3.2.6 Data Acquisition

Due to the high rate of the beam crossings, CDF employs a trigger system to reduce the event rate down to a manageable level while retaining data from small cross section interactions and maintaining high efficiency for the broad range of the physics processes. The reduction rate is determined by the speed at which data can be stored on tape.

The trigger system is comprised of three levels and is able to function with a 132 ns bunch separation while keeping dead time as short as possible. The trigger architecture is shown in Figure 3.15. Each trigger level in fact consists of a number

of separate triggers, each with certain selection criteria. Each successive level of the trigger processes fewer events than the preceding level, only those events which pass requirements of the preceding level; but it processes them with greater sophistication and requires more time per an event. Level 1 uses custom designed hardware to find physics objects based on part of the detector information and makes a decision by simply counting the number of objects. The Level-2 trigger makes use of hardware processors to do a limited event reconstruction. The Level-3 trigger is implemented in software, and it uses a processor farm running on the full data record of each event.

In the Level-1 trigger, the information from all detectors is buffered in a 42-event deep synchronous pipeline and stored for 5.5 μsec. During this time the received data is analyzed by three parallel synchronous streams. One stream finds calorimeter based objects (electrons, photons, jets), another finds stubs in the muon chambers, while the third one, the eXtremely Fast Tracker (XFT), reconstructs tracks on the transverse plane of the COT, and an extrapolation unit (XTRP) matches these tracks to the calorimeter and muon chambers.

The calorimeter triggers are formed by applying thresholds to energy depositions in calorimeter trigger towers with a segmentation of approximately $\delta\eta \times \delta\phi = 0.2 \times 15°$. The thresholds are applied to individual towers (object triggers) as well as to sum of energies from all towers (global triggers). Electron and photon triggers are formed by applying energy thresholds to the electromagnetic (EM) energy in a tower, while jet triggers are formed using the total (EM + HAD) energy in a tower. To identify electron/muon candidates the online track processor, the XTRP, links reconstructed XFT tracks with clusters in the EM calorimeter and pairs of hits in muon drift tubes. P_T and E_T thresholds are programmable for the various detector regions, and a flexible decision module can

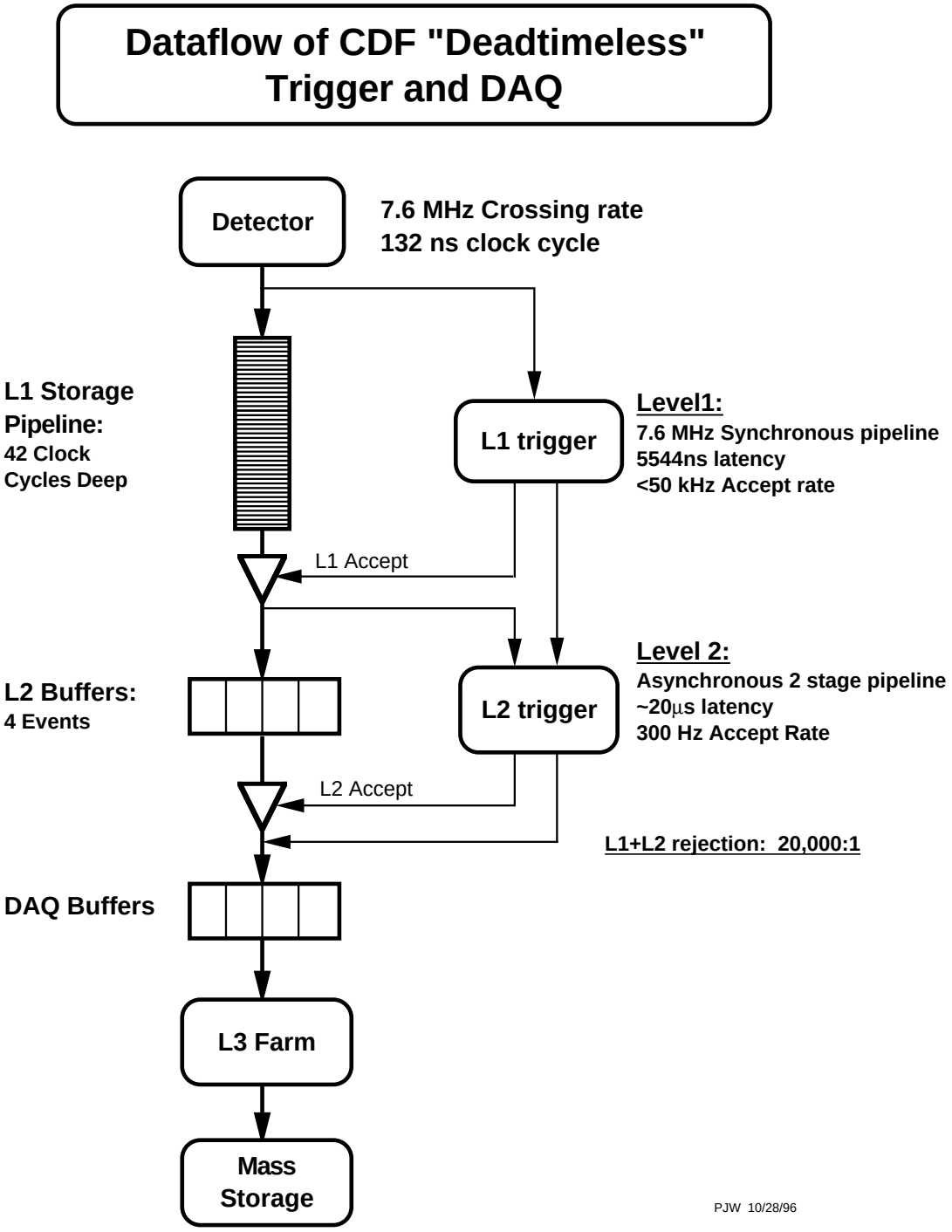

Figure 3.15: Functional block diagram of the CDF data flow.

form up to 64 different triggers by setting requirements on the number and certain features of the objects from the Level-1 trigger streams. The decision module uses simple AND and OR gates, and each trigger rate can be prescaled to the desired value.

All elements of the Level-1 trigger are synchronized to the same clock and a decision is made every 132 nsec. The rejection factor is about 150, thus decreasing the event rate from 7.6 MHz to about 50 kHz.

Events satisfying the requirements of Level-1 trigger are downloaded into one of four asynchronous event buffers and processed via programmable Level-2 hardware processors. While Level-2 analyzes the events the buffer cannot be used for additional Level-1 accepts. If all four buffers are full then the experiment starts to incur deadtime. To keep the deadtime at an acceptable level of 10% and maintain the 50 kHz Level-1 rate, the Level-2 latency is set to 20 μsec by using pipeline in two stages each taking approximately 10 μsec.

The first phase is an event building stage. Since jets usually affect more than one calorimeter tower, the cluster finder (L2CAL) combines the energies collected by single towers and forms clusters, thus providing a measurement of the total jet energy E_T as well as average ϕ and η of the jet. The shower maximum detectors (XCES) reduce the rate of fake electrons and photons by eliminating the background from single-phototube discharge and improving matching between XFT tracks and EM clusters. The Silicon Vertex Tracker (SVT) reconstructs tracks in the vertex detector and measures their impact parameter d.

On the second pipelined stage the results of the first phase are collected in the memory of the Level 2 processors, which examine the data if the criteria for any of the Level-2 triggers are satisfied. About one hundred different Level-2 triggers can be formed. The Level-2 accept rate is about 300 Hz with a rejection factor of

about 150.

After being accepted by the Level-2 trigger, event fragments are collected in the data acquisition system (DAQ) buffers and then transferred via a network switch to Event Builder CPU nodes, called *converter nodes*, where events are assembled from their fragments to complete and with appropriate data structures for analysis. The events are then passed to the Level-3 farm of parallel processor nodes, which take advantage of the full detector information and improved resolution not available to the lower trigger levels. They analyze and classify each event and then make a decision by applying filtering mechanism. If the event is accepted, it is delivered to a permanent storage by the consumer-server logger system (CSL). In addition, a sample of events is sent to the online monitoring processes verifying that the detector, trigger and data acquisition system are functioning correctly. The rate of events transferred to tape is around 75 Hz with an average event size of 250 kB corresponding to up to 20 MB/sec total output rate.

Chapter 4

Particle Identification

As outlined in Chapter 2, the $t\bar{t}$ dilepton events ($t\bar{t} \rightarrow bW^-\bar{b}W^+ \rightarrow b\ell^-\nu\bar{b}\ell^+\bar{\nu}$) have a signature that is characterized by two high-P_T leptons, electrons or muons, two high-E_T jets and two neutrinos resulting in a large transverse energy imbalance, *missing transverse energy* $\not\!\!E_{\mathrm{T}}$.

This Chapter describes in detail identification variables and criteria applied to identify final state objects in the event by the detector components. We start with the electron identification in the central and plug regions of the detector, discuss requirements for muon candidates, describe the jet reconstruction algorithm and corrections to the jet energies, and the way we determine the energy of neutrinos escaping the detector. We describe the global event selection in the next chapter.

4.1 Electron Identification

Electrons resulting from the dilepton channel of the $t\bar{t}$ decay are highly energetic. They can be identified by a high-P_T track in the drift chamber and large energy depositions in electromagnetic calorimeters. At energies of tens of GeV the dominant energy loss for electrons is bremsstrahlung. When electrons traverse the lead

absorbers in the electromagnetic calorimeter, they interact with the nuclei of the material and emit photons which produce electron-positron e^+e^- pairs. Secondary particles are also very energetic and lead to production of an electromagnetic cascade, called a *shower*. The shape and position of the electromagnetic shower is measured by shower maximum detectors, and the shower characteristics are used in electron identification. A hadronic shower is longer and much broader, as electrons deposit most of their energy in a single electromagnetic calorimeter tower, while a hadronic shower continues into the hadronic section of the calorimeter and into the adjacent calorimeter towers. In addition to certain shower properties, for electrons the momentum of the track pointing to the shower position should match the energy in the calorimeter.

4.1.1 Central Electrons

Central electron candidates traverse the central part of the detector, $|\eta| < 1.1$, leaving the track in the COT and depositing most of their energies in the CEM calorimeter. The following variables and imposed criteria are used to identify high-P_T electrons:

- $E_T = E * \sin\theta > 20$ GeV

  The transverse electromagnetic energy E_T deposited by the electron in the CEM cluster. It is given by the total electromagnetic energy of the electron cluster E multiplied by $\sin\theta$ of the COT track pointing to the seed tower of the cluster. An electron cluster is formed of a seed EM tower, where most of the energy is deposited, and a number of shoulder EM towers, which are added to the seed tower until the maximum cluster size is reached. The cluster at maximum has two towers in pseudorapidity and one tower az-

imuth. The energy E is corrected for differences in response, non-linearities and time-dependent changes.

- $P_T > 10$ GeV

  The transverse momentum of the COT track measured by its curvature in the magnetic field. Raw COT resolution is substantially improved by constraining the track to originate from the beam line. The *beam constrained* tracking introduces a P_T curvature bias in data, which is removed by correcting the signed curvature Q/P_T, where Q is a the charge of the track:

$$\frac{Q}{P_T^{corr}} = \frac{Q}{P_T} - 0.00037 - 0.0011 \times \sin(\phi + 0.28) \tag{4.1}$$

  Later in the text we refer to the P_T of the COT track when beam constrained, P_T^{corr}. This correction is applied only to data and not to sumulation of events.

- $E_{had}/E_{em} < 0.055 + 0.00045 \times E_{total}$

  The ratio of the hadronic calorimeter (CHA) energy of the cluster E_{had} to its electromagnetic energy E_{em} (CEM). The value E_{had}/E_{em} is scaled by a factor of 0.00045 multiplied by the total energy of the cluster E_{total} to compensate for inefficiency of the cut at very high energies, as higher energy electrons have more leakage into the hadronic calorimeter.

- $E/P < 2$ is applied for electrons with $E_T < 100$ GeV or $P_T < 50$ GeV.

  The ratio of the electromagnetic calorimeter energy E to the COT track momentum P. During the passage through the material inside the COT inner radius the electron might radiate a photon ("external bremsstrahlung").

The photon is collinear with the electron and it generally deposits its energy in the same calorimeter tower, thus not much affecting the value of E_T, however the momentum P_T measured in the COT after bremsstrahlung will be smaller. This causes a long tail in the E/P distribution above 1.0.

- $L_{shr} < 0.2$, "El share".

  The lateral shower profile L_{shr} is a measure of how well the lateral shower development matches that expected from the electromagnetic shower.

  This variable compares the energies of CEM calorimeter towers adjacent to the seed tower of the EM cluster with energies expected from the test beam electrons. It is defined as a sum over towers:

$$L_{shr} = 0.14 \sum_i \frac{E_i^{measured} - E_i^{expected}}{\sqrt{(0.14\sqrt{E})^2 + \sigma^2_{E_i^{expected}}}}$$

  where $E_i^{measured}$ is a measured energy in the CEM tower i; $E_i^{expected}$ is an energy deposit in the i-th tower expected from the test beam electrons and E is the energy of the EM cluster. $0.14\sqrt{E}$ represents the error on the energy measurement and $\sigma^2_{E_i^{expected}}$ is the uncertainty on the energy estimate. For a typical EM cluster L_{shr} is a two-tower sum. Any extra particles accompanying the one responsible for the main EM shower will tend to add to the energy in adjacent towers and make L_{shr} a large positive number.

- - 3.0 cm $< Q * \Delta x <$ 1.5 cm ; $|\Delta z| <$ 3 cm

  The distances Δx / $|\Delta z|$ in the $r - \phi$/ $r - z$ planes between the COT track extrapolated to the CES and the best matching CES cluster. The cut on Δx has been multiplied by the charge Q of the electron and it is asymmetric in $r - \phi$ to account for energy of the photon emitted in bremsstrahlung

radiation. This requirement on a tight match between the track and shower position removes a large number of fake electrons due to the coincidence of charged and neutral hadrons in the same tower, mainly due to π^0 and $\pi^\pm$, when the π^0 causes an electromagnetic shower in the calorimeter and the $\pi^\pm$ results in a reconstructed track that matches the electromagnetic cluster.

- $\chi^2_{strip} < 10$

  The χ^2 comparison of the CES shower profile in the $r - z$ view with the shower profile extracted from test beam electrons. The χ^2-fit is performed on the distribution of energy deposited on each of the 11 strips in the CES shower.

- $z_{vertex} < 60$ cm

  The interaction position in z, z_{vertex}, is taken from z_0, the z intersection of the track with the beam axis in the $r - z$ plane. The longitudinal spread of the event vertex about the nominal interaction point $z = 0$ is a gaussian with $\sigma = 26$ cm. The vertex position is required to be within 2σ.

- Track quality cuts

  To insure that the track associated with the electron is well reconstructed, it must pass through three axial and three stereo superlayers (SL) of the COT with at least 7 hits out of 12 in each SL.

- Fiduciality:

  This variable insures that the electron is reconstructed in a region of the detector that is well instrumented. The electron position in the CEM is determined by the CES shower position and it must satisfy the following requirements:

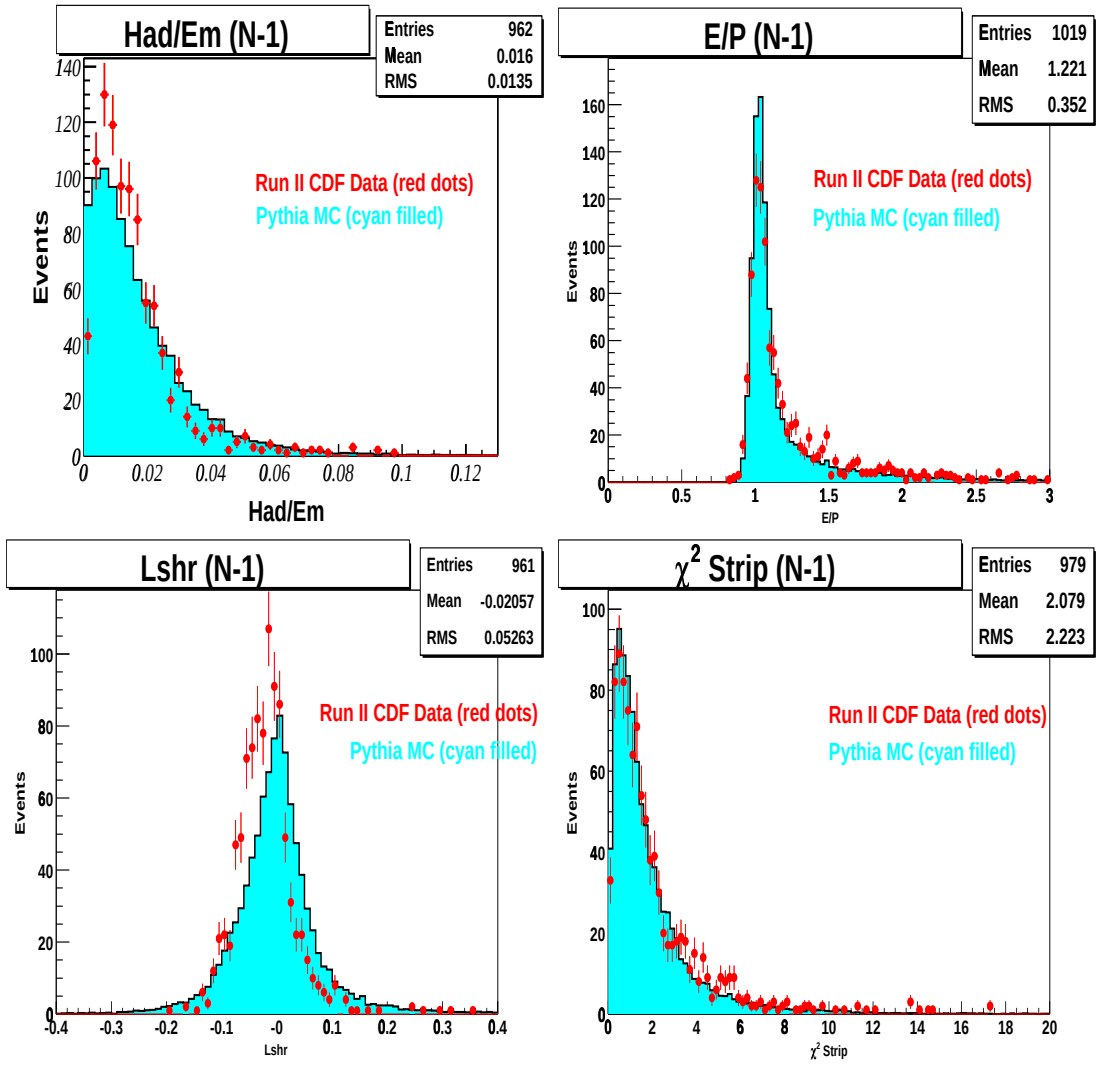

Figure 4.1: Distributions of central electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.)

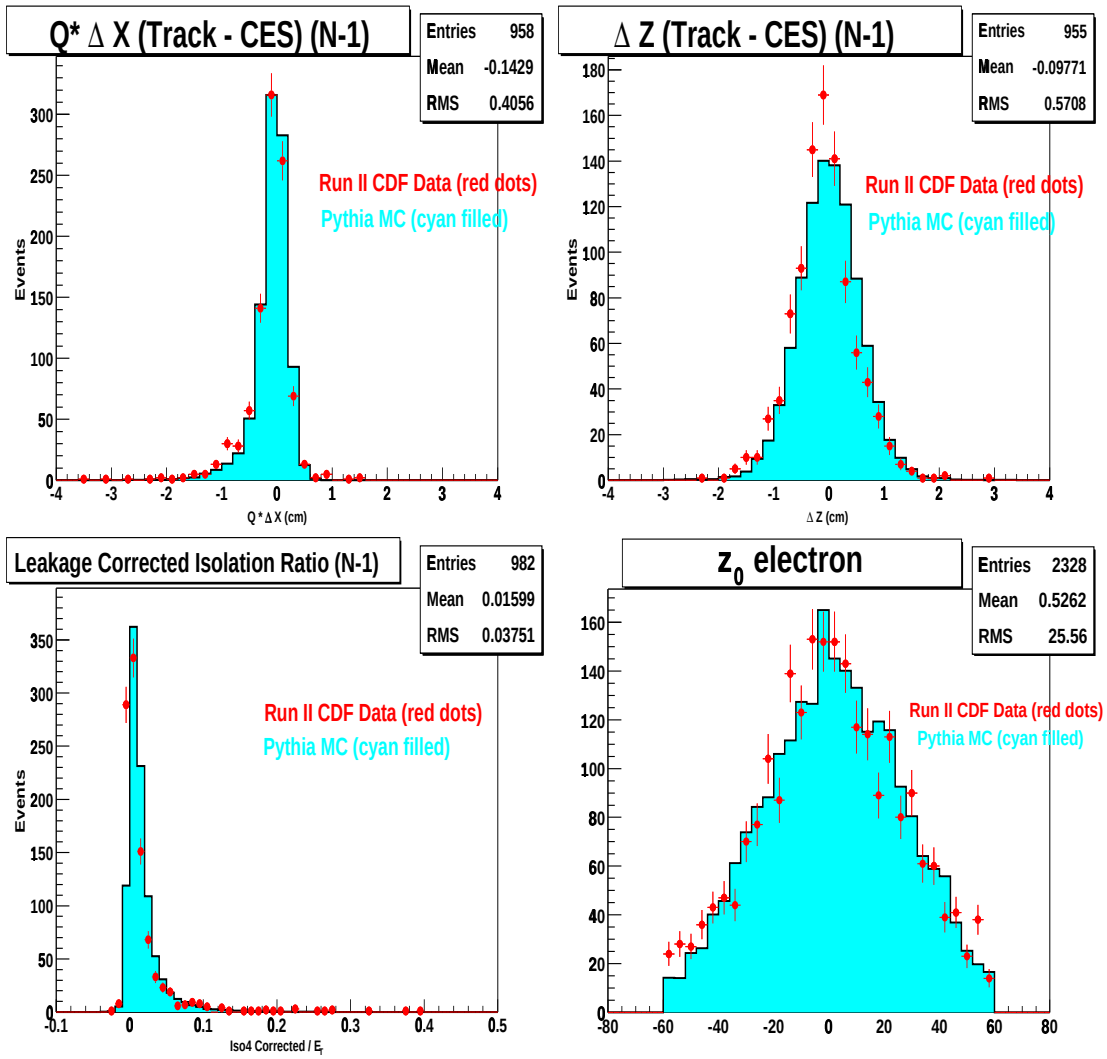

Figure 4.2: Distributions of central electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.)

- the electron must lie within 21 cm of the tower center in the $r-\phi$ view, so that the shower is fully contained in the active region; this corresponds to the cut $|x_{CES}| < 21$ cm, where the x_{CES} is the local coordinate of the calorimeter tower;

- the region $|z_{CES}| < 9$ cm, where the two halves of the central calorimeter meet is excluded, as well as the region $|z_{CES}| > 230$ cm, which corresponds to the outer half of the last CEM tower (tower 9), as it is more subjected to the leakage into the hadronic part of the calorimeter.

- the region immediately close to the point of penetration of the cryogenic connections to the solenoidal magnet, the chimney, is un-instrumented and therefore excluded. It corresponds to $0.77 < \eta < 1.0, 75° < \phi < 90°$ and $|z_{CES}| > 193$ cm.

- the region $1.05 < |\eta| < 1.10$ is excluded due to the smaller depth of the electromagnetic calorimeter.

- Not a conversion

  Photons produced either directly in the hard scattering or from π^0 decays interact with the material in the detector and convert to electron-positron pairs. These photon conversions can be identified by the presence of another track of the opposite sign near the electron candidate. If the two tracks exhibit small $r-\phi$ separation in the point of conversion $|\Delta XY| < 0.2$ cm, and the difference in their polar angle $|\Delta cot\theta| < 0.04$, the electron candidate is flagged as a conversion and the whole event is rejected.

- Isolation $= E_T^{iso}/E_T^{cluster} < 0.1$.

  where $E_T^{iso} = E_T^{0.4} - E_T^{cluster}$ is the transverse energy $E_T^{0.4}$ in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the electron cluster (hadronic + elec-

tromagnetic) excluding the electron cluster energy $E_T^{cluster}$. This cut in fact is not an electron identification requirement. It rejects the electrons that are not isolated from extra hadronic activity and could be the products of quark semi-leptonic decays, while electrons from W and Z decays are expected to be isolated. We therefore consider both categories of isolated and non-isolated electrons in the event selection (see Section 5.2.2).

The isolation is corrected for leakage energy into the neighboring ϕ wedge outside of the cone, which increases towards the edges of the ϕ wedges. The respective correction factor is determined from Monte Carlo - data comparison and parametrized in the form [24]:

$$E_{leak} = E_T^{cluster} \cdot p_0 \cdot \exp\{p_1 \cdot (|x_{CES}| - 21)\},$$

where $p_0 = 0.0511 \pm 0.0075$ and $p_1 = 0.33 \pm 0.061$.

If additional interactions occur in the same bunch crossing the energy in a cone is increased. Therefore isolation is also corrected for the number of interactions per bunch-crossing, the efficiency of the cut then becomes independent of the instantenous luminosity. This correction factor is determined similarly to jet multiple interaction energy corrections (see Section 4.3).

The corrected isolation energy is then

$$E_T^{corr} = E_T^{iso} - E_{leak} - E_{MI}.$$

and corrected isolation is given by

$$Iso_{corr} = E_T^{corr} / E_T^{cluster}.$$

Central electron ID efficiency was determined from the $Z \to e^+e^-$ data sample. $Z \to e^+e^-$ events were selected as dielectron events of opposite charge falling into

Z mass window, 75 GeV $< M_{ee} <$ 105 GeV. The number of observed same-sign dielectron events in the same Z mass range served as an estimate of QCD background contamination. At least one electron was required to be tight, i.e. passing all identification criteria. The other leg was used to test the rate of passing all electron identification cuts. The central electron ID efficiency was found to be $\epsilon_{data} = (82.5 \pm 0.2)\%$. The same calculation in Monte Carlo gives slightly higher value of $\epsilon_{MC} = (85.5 \pm 0.5)\%$. To compensate for differences between the efficiency to reconstruct an electron in the data and in the simulation the electron reconstruction efficiency scale factor $\epsilon_{data}/\epsilon_{MC} = (96.5 \pm 0.6)\%$ is applied to correct the acceptance measured in the Monte Carlo.

Distributions of electron identification variables are presented on Figures 4.1 and 4.2. For each variable all identification criteria are applied except for the one plotted.

4.1.2 Plug electrons

Electron candidates depositing their energies in the PEM calorimeter are referred to as *plug electrons*. We impose the following identification criteria to define plug electron candidates:

- $1.2 < |\eta| < 2.0$

  Although plug electron can be identified up to $|\eta| < 2.5$, this analysis considers only those with $|\eta| < 2.0$, primarily due to large charge misidentification rate at high pseudorapidities η. This cut has a small effect on the $t\bar{t}$ acceptance, as final products of $t\bar{t}$ decay events are mostly central and the acceptance falls rapidly at large η, while backgrounds considerably increase.

The low η-region $|\eta| < 1.2$ is excluded due to geometrical reasons since the PES detector does not provide usable coverage there.

The track information from the COT is unavailable in the forward region of the detector, as plug electrons do not traverse the active volume of the COT. Plug electron candidates identified based only on the presence of an EM cluster in the PEM calorimeter, i.e. with no track requirements, are called *PEM electrons*. PEM electron candidates do not provide charge sign information and are subject to a large fake rate.

To reduce the misidentification rate the track reconstruction for plug electrons is performed by utilizing the silicon layers (ISL) residing in the forward part of the detector and providing usable coverage in $1.2 < |\eta| < 1.8$ range (see Section 3.2.2). The silicon pattern recognition code extrapolates the hits in the outer layers to the inner layers of silicon and determines the location of the primary vertex. The tracking efficiency for tracks pointing to the plug region is considerably lower than for the central part of the detector. It is improved by a special algorithm, called *Phoenix algorithm* [25].

The Phoenix algorithm makes use of the information of the PES position of an EM shower. It constraints the track at two endpoints, one is fixed at the interaction vertex and the other is at the PES position of an EM shower. Adopting the corresponding PEM cluster energy as the momentum of the electron, a helix of the track can be determined. This defines two possible track trajectories, one is for negative and the other is for positive charge. The silicon pattern recognition code further attempts to reconstruct those tracks by matching hits in the layers of silicon. If one of these tracks is reconstructed, it is appended to the event record as being associated with the respective electron candidate. If both tracks are reconstructed then the

algorithm performs the χ^2-fit of a possible electron trajectory and adopts the track best matching the activity in the silicon detector. Such a track is called a *Phoenix track* and a PEM electron with a Phoenix track is referred to as a *Phoenix plug electron*, abbreviated as PHX.

The Phoenix algorithm establishes charge identification for plug electron candidates beyond the coverage of the forward silicon layers up to $|\eta| < 2.0$.

- $E_T > 20$ GeV

  Unlike the CEM clusters, the PEM clusters are limited to 2×2 - two towers in pseudorapidity by two towers in azimuth.

- $E_{had}/E_{em} < 0.05$

  Similar to the central electrons, the ratio of the hadronic calorimeter (PHA) energy of the cluster E_{had} to its electromagnetic energy E_{em} (PEM).

- $U_{5\times9} \geq 0.65$ and $V_{5\times9} \geq 0.65$

  The quantities $U_{5\times9}$ and $V_{5\times9}$ are essentially isolation variables for the shower maximum detector independently applied to both the U and V layers. The PES clustering is performed by ordering PES strips in decreasing energy with the highest-energy strips used as seeds. Then a fixed-width nine-strip cluster is formed from each seed. This is done separately for 1D U- and V-layer clusters. The quantities $U_{5\times9}$ and $V_{5\times9}$ represent the ratios of energy sum in the central 5 strips of a PES cluster to the total energy of the PES cluster (in all 9 strips).

- $\chi^2_{3\times3} < 10$

  This variable represents the goodness of χ^2-fit measure of the energy distribution in 3×3 towers around the seed tower to energy distributions from

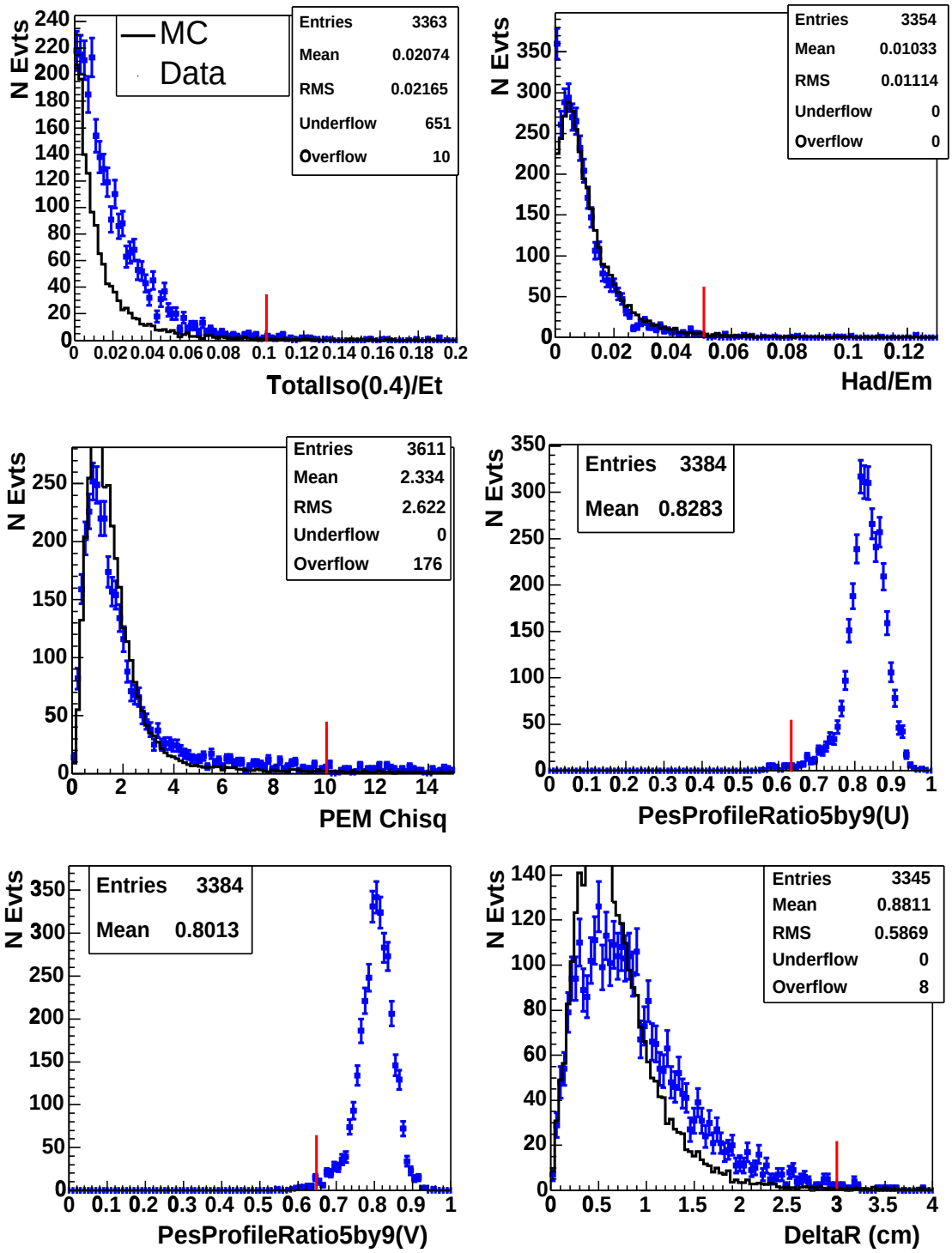

Figure 4.3: Distributions of plug electron identification variables. For each variable all identification criteria applied except for the one plotted ('N-1' distributions.)

test beam electrons.

- $|\Delta R_{PES}| < 3$ cm

  The χ^2-fit also determines the position for the center of the shower. $\Delta R_{PES} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the distance between the $\chi^2_{3\times3}$ best fit position and the intersection of the centroids in the U- and V-layer PES clusters. ΔR matching requirement is also enforced between the PES coordinates and the extrapolated Phoenix track.

- $N^{Si}_{hits} \geq 3$

  Number of Silicon hits N^{Si}_{hits}. This requirement is enforced to improve the quality of the silicon tracks at the cost of some efficiency.

- $|z_0| < 60$

  This cut is identical to the CEM electrons.

- Isolation $= E_T^{iso}/E_T^{cluster} < 0.1$

  This variable is defined identically to the CEM electrons. The corrections to E_T^{iso} are done differently, however, as the clustering algorithm in the plug region differs from the one in the central region. The leakage energy is parametrized depending on the distance of the tower from the center of the electron shower in θ and ϕ [24].

Plug electron ID efficiency was determined from the $Z \to e^+e^-$ data sample, similar to how it was done for central electrons. $Z \to e^+e^-$ events were selected as central-plug dielectron events in the Z mass range. The central electron was required to be a good electron, i.e. pass all tight identification criteria, and the other leg was required to be an EM object fiducial to the plug region. Plug electron ID efficiency was determined as the ratio of tight-tight(central-plug) candidates

over tight-loose candidates and was found to be $\epsilon_{data} = (65.1 \pm 0.8)\%$. The same calculation in Monte Carlo gives $\epsilon_{MC} = (74.9 \pm 0.2)\%$. Therefore additional PHX ID scale factor $\epsilon_{data}/\epsilon_{MC} = 0.87 \pm 0.01$ is applied to Monte Carlo simulation for acceptance estimates from Monte Carlo samples.

Distributions of plug electron identification variables are presented on Figure 4.3. For each variable all identification criteria applied except for the one plotted.

4.2 Muon Identification

Muons are minimum-ionizing particles which penetrate matter very easily. Muons resulting from the dilepton channel of the $t\bar{t}$ decay, as well as electrons, are very energetic. They are identified by the high-P_T track in COT, very little energy deposition in the calorimeters, and matching hits in the muon chambers. A muon candidate is required to have aligned hits in both $r - \phi$ and $r - z$ planes on at least 3 separate layers. These hits form a muon stub which is then matched to the COT tracks extrapolated to the muon chambers.

Muons are categorized by the detector region through which they pass. Muons reconstructed in both the CMU and CMP chambers are called CMUP muons. Due to the gaps in the muon chambers coverage there are also muons that are reconstructed only in the CMU or the CMP muon chambers. These muons are called CMU-only, CMP-only respectively. Muons with stubs in the CMX chamber are called CMX muons. The COT tracks with no muon stubs are also considered as muon candidates. Such muons are required to have minimum energy depositions in the calorimeter. These muon candidates are called CMIO's (central minimum ionizing objects).

The following criteria are applied for muon candidates:

- $P_T > 20$ GeV

  The transverse momentum of the COT track. The track is beam constrained and P_T corrections are applied identically as it is done for electrons.

- $|z_0| < 60$ cm

  As for electrons, the z-position of the muon track is associated with the event vertex and required to be within 60 cm from the geometrical center of the detector.

- $d_0 < 0.2$ cm for tracks with no silicon hits;
  $d_0 < 0.02$ cm for tracks with silicon hits

  The impact parameter d_0 is the distance between the reconstructed muon track and the beam axis in the $r - \phi$ plane. Unlike others this selection variable is based on the default muon track, not beam-constrained, and a tighter cut is applied if the track contains silicon hits. This cut forces the muon to originate from the nominal interaction region and substantially reduces the cosmic muon background. It also helps to remove muons from kons and pions that decay in flight.

- Track quality cuts

  are identical to electron track quality cuts. The track is required to have at least 7 COT hits on at least 3 axial and 3 stereo superlayers.

- $E_{em} \leq 2 + max(0, 0.0115(P - 100))$ GeV

  The energy deposited in the electromagnetic calorimeter. High P_T muons are not expected to deposit substantial amount of energy in the electromagnetic

calorimeter. The sliding cut is introduced for muons with $P > 100$ GeV to increase efficiency of the cut.

- $E_{had} \leq 6 + max(0, 0.0280(P - 100))$ GeV

  The energy deposited in the hadronic calorimeter is higher, but still quite small in comparison to strongly interacting jets.

- $E_{em} + E_{had} > 0.1$ GeV for stubless muons only.

  Stubless muons are required to have a non-zero energy deposition in the calorimeter to limit backgrounds from electrons escaping the detector through cracks in the calorimeter.

- $|\Delta x|_{CMU} \leq 3$ cm

  The distance in the $r - \phi$ plane between the extrapolated COT track and the stub segment in the relevant muon chamber. The muon candidate passing this requirement falls into CMUP or CMU-only category.

- $|\Delta x|_{CMP} \leq 5$ cm

  Muons in the CMP and CMX detectors traverse more material than in the CMU and experience greater deflections due to multiple scattering. Therefore, the track-to-stub matching cuts are looser. This muon candidate falls into CMUP or CMP-only category.

- $|\Delta x|_{CMX} \leq 6$ cm

  CMX muon category requirement.

- $\rho_{COT} > 140$ cm,

  The COT exit radius

$$\rho_{COT} = \frac{sign(\eta) \cdot z_{COT} - z_0}{cot(\theta)}$$

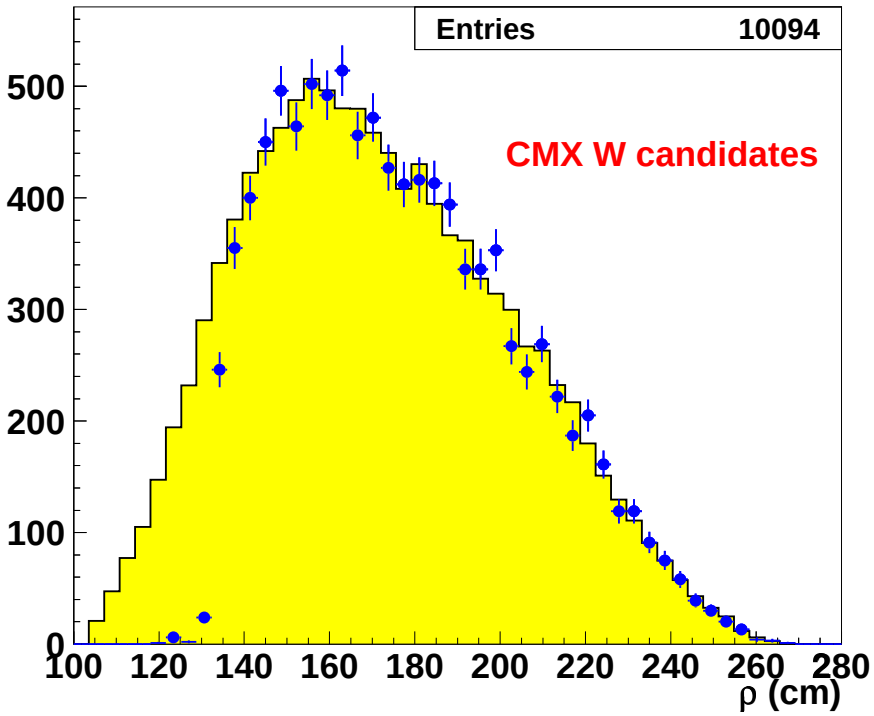

Figure 4.4: The COT exit radius ρ_{COT} of CMX muons selected as $W \rightarrow \mu\nu$ candidate events. Filled histogram is the Monte Carlo simulation with no requirement on the simulated trigger. Data and simulation agree for approximately $\rho_{COT} \geq 140$ cm.

is based on pseudorapidity η and z_0 of the track, where $z_{COT} = 155$ cm is the length of the COT and θ is the polar angle.

This cut is enforced only for CMX muons to eliminate the data bias due to the XFT trigger requirement that a track must leave hits in at least 4 COT superlayers. The data and the Monte Carlo agree for large exit radii ρ_{COT} as illustrated in Figure 4.4.

- Isolation $= E_T^{iso}/P_T < 0.1$

  where $E_T^{iso} = E_T^{cone} - E_T^{tower}$ is the difference between energy in the cone of

Muon category	$\epsilon_{data}(\%)$	$\epsilon_{MC}(\%)$	Scale Factor (%)
CMUP	86.5 ± 0.7	92.47 ± 0.17	93.5 ± 0.7
CMX	91.9 ± 0.7	90.66 ± 0.1	101.4 ± 0.7
CMU-only	90.5 ± 1.1	91.2 ± 0.3	99.3 ± 1.3
CMP-only	91.9 ± 1.1	93.5 ± 0.2	98.3 ± 1.1
CMIO	67.3 ± 1.8	67.3 ± 0.8	100.0 ± 2.0

Table 4.1: Muon identification efficiencies in the data and the simulation and scale factors to be applied to correct acceptance measurement from Monte Carlo samples.

$\Delta R = 0.4$ around the muon track E_T^{cone} and the amount of energy in the tower associated with the muon track E_T^{tower}.

- not-cosmic

  Cosmic rays can be identified in the CDF detector as dimuon events leaving a nearly straight track and therefore mimicking a very energetic $\mu^+\mu^-$ pair. Since cosmic rays do not originate from a $p\bar{p}$ collision and enter the detector at random locations, this background is reduced by a cut on the track impact parameter d_0. In addition, cosmic rays appear randomly in time and can be distinguished by a substantial time delay between the hits of the two muons in the hadronic calorimeter, measured by Time to Digital Converter (TDC), and by using timing information from the Time of Flight detector (TOF) All of this information is analyzed by the software code, named *Cosmic Ray Tagger* [26], which makes a decision on flagging an event as a cosmic.

  Events with muons identified as cosmics are rejected.

The muon ID efficiencies are measured using Z boson decays $Z \rightarrow \mu^+\mu^-$

similarly to electrons. Both legs are linked to a CMUP or CMX muon stub, where one leg is required to pass tight identification criteria and is matched to Level-1 trigger information, while the second leg is chosen to be fiducial to the tested detector region (e.g. CMUP,CMX), or non-fiducial to any (for stubless CMIO muons) and examined if it passes muon ID cuts thus being independent of a trigger requirement.

4.3 Jet Reconstruction

At high energies, as those achieved in Tevatron, jets are the dominant feature of hadron production. They result from point-like collisions of a quark or gluon from the proton with a quark or gluon from the antiproton. A jet is formed from a scattered initiating parton, which experiences fragmentation leading to the creation of a stream of energetic colorless particles emitted spatially collimated along the original parton direction.

The jets are observed as clusters of energy located in adjacent detector towers. Typically a jet contains neutral or charged pions to a lesser extent of kaons, and about 10% of light baryons such as protons and neutrons. Pions mostly deposit their energies in electromagnetic calorimeter, while kaons and baryons leave most of their energies in the hadronic section of the calorimeter. The energy of the initial parton can be approximated by summing the tower energies within a cone of specified size. This procedure is called *jet clustering*. The cone size is chosen to encompass most of the jet energy without allowing a significant contribution from other event activity. It is defined in $\eta - \phi$ space by its radius, $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and is centered at the largest calorimeter energy tower serving as a seed tower of the jet cluster. This analysis is using a cone size of $\Delta R = 0.4$.

After the jet cluster is thus formed, the E_T weighted centroid of the cluster is determined, as follows

$$\eta_{centroid} = \frac{\sum_{i=1}^{N} E_T^i \eta^i}{\sum_{i=1}^{N} E_T^i},$$

$$\phi_{centroid} = \frac{\sum_{i=1}^{N} E_T^i \phi^i}{\sum_{i=1}^{N} E_T^i},$$

where the sums are carried out over all calorimeter towers in the cluster. It defines the centroid tower and a new cone drawn around this position. This process is iterated until the cluster remains unchanged in two consecutive paths. In some cases two clusters can overlap and then they are either merged into one, if the sum of the energies in shared towers exceed 75% of the energy of the smaller cluster, or left intact.

The jet four-momentum $(E^{raw}, P_x^{raw}, P_y^{raw}, P_z^{raw})$ is then determined by the following sums over the cluster towers:

$$E^{raw} = \sum_{i=1}^{N} E_i$$

$$P_x^{raw} = \sum_{i=1}^{N} E_i \sin\theta_i \cos\phi_i$$

$$P_y^{raw} = \sum_{i=1}^{N} E_i \sin\theta_i \sin\phi_i$$

$$P_z^{raw} = \sum_{i=1}^{N} E_i \cos\phi_i$$

These quantities are referred to as raw, since they are affected by mismeasurements for a variety of reasons due to both to physics and to detector effects and

are different from the true energies of the partons which initiated jets. Therefore proper corrections need to be applied to reconstruct true momentums of partons.

The corrected value of the jet E_T is derived from the the raw value E_T^{raw} by applying the multiplicative and additive correction factors in the seven cumulative levels and is given by the following relation:

$$E_T^{corr} = (E_T^{raw} \times f_{rel} \times f_{time} \times f_{scale} - E_T^{MI}) \times f_{abs} - E_T^{UE} + E_T^{OC}, \qquad (4.2)$$

where correction factors are described and determined as follows:

1. Relative corrections

   take into account non-uniformities in the calorimeter due to differences in responses from tower to tower, and the presence of cracks between the calorimeter modules. The correction factor f_{rel} is detector-η dependent and determined from dijet studies, where one jet is central and well measured and the second one is required to balance the first jet.

2. Time dependent corrections

   are applied to the calorimeter towers due to an evidence that their responses are changing with time due to the calorimeter deterioration, e.g. due to aging of the phototubes.

3. Raw energy scale corrections

   account for non-linearities in the single-particle response of the calorimeters. The correction factor f_{scale} is determined by using photon-jet balancing since the photon energy is very precisely measured by the electromagnetic calorimeter.

4. Multiple interaction energy corrections.

   At current luminosity and with 36 bunches on average one proton-antiproton interaction per beam crossing is expected. However, this number is subject to Poisson fluctuation and the physics processes under study such as $t\bar{t}$ decays can have overlapping minimum bias events. The energy from the minimum bias events may fall into the jet clustering cone and must be subtracted. The number of additional reconstructed vertices in the event serves as an indicator of multiple interactions in the same bunch crossing. The correction factor E_T^{MI} is determined from minimum-bias event by measuring energy in a randomly chosen cone with radius $\Delta R = 0.4$ and is parametrized by the number of vertices in the event found by the SVXII.

5. Absolute energy corrections

   are introduced to get the energy estimate of the original parton generating the jet. The parton energy is usually underestimated due to nuclear absorption and particle leakage. This correction is obtained from Monte Carlo simulation by associating a parton to a respective jet. The correction factor is P_T-dependent and defined as the ratio between the parton and jet P_T:

$$f_{abs} = \left\langle \frac{P_T^{parton}}{P_T^{jet}} \right\rangle$$

6. Underlying event corrections

   take into account contributions to the jet energy not coming from the original parton. These contributions are originating from soft interactions of spectator partons, beam-beam remnants, and are referred as the *underlying event*. The contribution of the underlying event E_T^{UE} to the jet energy is estimated from minimum bias events.

7. Out-of-cone corrections

   account for energy of the jet deposited outside the clustering cone due to fragmentation effects and soft gluon radiation. The amount of energy to be added as out-of-cone energy E_T^{OC} is determined from the Monte Carlo simulation. The correction is defined as $E_T^{OC} = P_T^{parton} - P_T^{jet}$, where P_T^{parton} is the sum of the P_T of all particles originating from the parton and P_T^{jet} is the sum of the P_T of the particles falling inside the reconstruction cone. The E_T^{OC} correction is parametrized as a linear function of P_T.

Systematic uncertainties are assigned to each level of the jet corrections and the combination of these uncertainties results in the total systematic uncertainty of the jet E_T measurement. Underlying event, multiple interactions and out-of-cone corrections are ignored in quoting the central value of the jet E_T, but their effect is included into the overall systematic uncertainty of the jet energy.

In this analysis jet counting is carried out by requiring an E_T threshold of 15 GeV and $|\eta| < 2.5$. In addition, since electrons are also reconstructed as jet objects. Jets within a cone of $\Delta R < 0.4$ from any identified electron are not considered.

4.4 Missing Energy

Neutrinos interact only through weak interactions and therefore cannot be directly detected as they traverse the detector material. Production of neutrinos in an event can be spotted by the existence of the large imbalance in the calorimeter energy. The longitudinal component of the colliding partons is not known, but the transverse component is subject to conservation, and the sum of the transverse components of the neutrino momenta can be measured. This quantity is called

missing transverse energy $\vec{\not{E}}_{\mathrm{T}}$. The missing transverse energy is two-component vector $(\not{E}_{\mathrm{T}x}, \not{E}_{\mathrm{T}y})$. The raw value of $\vec{\not{E}}_{\mathrm{T}}$ is defined by the negative vector sum of the transverse energy of all calorimeter towers:

$$\vec{\not{E}}_{\mathrm{T}}^{\,raw} = -\sum_{towers} (E_i \sin\theta_i)\vec{n}_i, \tag{4.3}$$

where E_i is the energy of the i-th tower, $\vec{n}_i$ is a transverse unit vector pointing to the center of the tower and θ_i is the polar angle of the line pointing from z_0, z-coordinate of the event vertex, to the i-th tower. This sum extends to $|\eta_{detector}| < 3.6$.

The value of $\vec{\not{E}}_{\mathrm{T}}^{\,raw}$ (Equation 4.3) should be further corrected for escaping muons and jet energy mismeasurements. Muons do not deposit substantial energy in the calorimeter, but may carry out significant amount of the energy. The sum of transverse momenta of escaping muons $\sum \vec{P}_{T,\mu}$ measured in the COT has to be added to the $\vec{\not{E}}_{\mathrm{T}}^{\,raw}$ with a negative sign and the energy deposited by muons in the calorimeters $\sum \vec{E}_{T,\mu}$ has to be subtracted from that sum, as it has been already counted in the $\vec{\not{E}}_{\mathrm{T}}^{\,raw}$.

Only raw values of jet energies contribute to the $\vec{\not{E}}_{\mathrm{T}}^{\,raw}$ and these values have to be replaced in the sum (Equation 4.3) by the corrected ones (Equation 4.2). The corrected value of $\vec{\not{E}}_{\mathrm{T}}^{\,corr}$ is therefore given by the following relation:

$$\vec{\not{E}}_{\mathrm{T}}^{\,corr} = \vec{\not{E}}_{\mathrm{T}}^{\,raw} - (\sum_{muons} \vec{P}_{T,\mu} - \sum_{muons} \vec{E}_{T,\mu}) - (\sum_{jets} \vec{E}_{T,jet}^{corr} - \sum_{jets} \vec{E}_{T,jet}^{raw}) \tag{4.4}$$

Uncertainties in $\vec{\not{E}}_{\mathrm{T}}^{\,corr}$ are dominated by uncertainties in jet energies. Mismeasurements of $\not{E}_{\mathrm{T}}$ result from jets traversing through poorly instrumented regions of detectors, e.g. cracks, dead zones, and beam halo effects. They may also result

from cosmic rays, muon misidentification and mismeasurements in muon track momenta.

The resolution of the $\not\!\!E_{\mathrm{T}}$ generally depends on the response of the calorimeter to the total energy deposited in the event. It is parametrized in terms of the total scalar transverse energy $\sum E_T$, which is defined as

$$\sum E_T = \sum_{towers} E_i \sin\theta_i \tag{4.5}$$

The $\not\!\!E_{\mathrm{T}}$ resolution in the data is measured with minimum bias events, dominated by inelastic $p\bar{p}$ collisions. In minimum bias events the x and y components of $\vec{\not\!\!E}_{\mathrm{T}}$ are distributed as gaussians around zero with $\sigma_x = \sigma_y = \sigma$:

$$\frac{dN}{d\not\!\!E_{\mathrm{T}_x}} \simeq \exp\left(-\frac{\not\!\!E_{\mathrm{T}_x}^2}{2\sigma^2}\right),$$
$$\frac{dN}{d\not\!\!E_{\mathrm{T}_y}} \simeq \exp\left(-\frac{\not\!\!E_{\mathrm{T}_y}^2}{2\sigma^2}\right)$$

The $\not\!\!E_{\mathrm{T}}$ resolution $\Delta = \sqrt{<\not\!\!E_{\mathrm{T}}^2>}$ is then given by $\Delta = \sqrt{2}\sigma_{x,y}$. It is expected to scale as a square root of the total transverse energy in the event $\sum E_T$ and from minimum bias studies is determined to be $\Delta \simeq 0.64\sqrt{\sum E_T}$ [27].

Since the results of this analysis rely on the proper simulation of the events, it is essential that modelling of the $\not\!\!E_{\mathrm{T}}$ is consistent with that observed in the data.

To insure an agreement between the data and the Monte Carlo simulation we performed the $\not\!\!E_{\mathrm{T}}$ resolution measurements in both observed and simulated events of Z boson decays. $Z \to \ell^+\ell^-$ events do not possess an intrinsic $\not\!\!E_{\mathrm{T}}$, and therefore by excluding the lepton energy deposits in the calorimeter, one can measure the $\not\!\!E_{\mathrm{T}}$ resolution.

Energy deposition from electrons are more complex to model accurately due to tower-to-tower energy leakage effects. While these effects are negligible for muons,

which do not leave substantial amounts of energy in the calorimeter. Therefore more accurate measurement can be obtained from dimuon $Z \to \mu^+\mu^-$ events.

In the scattering process

$$p\bar{p} \to Z + X \to \mu^+\mu^- + X$$

the initial quarks can radiate gluons, and the Z can acquire transverse momentum. The total transverse momentum in the event is conserved, and the P_T of the boson balances the energy $\vec{U}$ that recoils against it.

In addition to the initial state radiation, the energy associated with multiple interactions and the remnants of the protons and antiprotons that are involved in the Z production will also contribute to the recoil $\vec{U}$.

The raw missing transverse energy E_T^{raw} in the event is due to the recoil energy:

$$\vec{\not{E}}_{\mathrm{T}}^{\,raw} = -\vec{U} \tag{4.6}$$

where the energy deposited by the muons is excluded from the $\vec{\not{E}}_{\mathrm{T}}^{\,raw}$.

Since $\vec{U}$ is a measure of the boson P_T, it is natural to define $\vec{U}$ in terms of the components $U_{\|}$ and $U_{\perp}$, parallel and perpendicular to the boson direction, respectively. The average value of $U_{\|}$ is the average calorimeter response balancing the boson P_T, and the average value of $U_{\perp}$ is expected to be zero, since neither side of the boson direction is preferred. $U_{\|}$ and $U_{\perp}$ can be parametrized in the form:

$$\begin{pmatrix} U_{\|} \\ U_{\perp} \end{pmatrix} = \begin{pmatrix} f(P_T^Z) \\ 0 \end{pmatrix} + \begin{pmatrix} G_1(\sigma_{\|}) \\ G_2(\sigma_{\perp}) \end{pmatrix} \tag{4.7}$$

where $G_1(\sigma_{||})$ and $G_2(\sigma_{\perp})$ are Gaussian distributed random variables of mean zero and widths $\sigma_{||}$ and $\sigma_{\perp}$, and the function $f(P_T^Z)$ is the response to the recoil energy.

The $\sigma_{\perp}$ results from the underlying event energy, which is expected to have the same dependence on $\sum E_T$ as the minimum bias data, while the $\sigma_{||}$ resolution may have contributions from low energy jets that balance the boson P_T, and in addition requires knowledge of the exact $f(P_T^Z)$ shape.

We therefore aim to determine the functional dependence of the $\sigma_{\perp}$ resolution on the total scalar transverse energy $\sum E_T$ and compare it with Monte Carlo simulation. The overall $\not{E}_{\mathrm{T}}$ resolution due to the underlying event is given by

$$\Delta = \sqrt{2}\sigma_{\perp} \tag{4.8}$$

We selected dimuon events as muon-track pairs coming from the same vertex by requiring $|z_{0,muon} - z_{0,track}| < 4.0$ cm. The tight leg was required to satisfy muon identification criteria as specified in Section 4.2, and the loose leg had the same cuts applied except for no stub requirements. To achieve high purity of the $Z \to \mu^+\mu^-$ events, we required the narrow $86 - 96$ GeV di-muon mass window.

The background from $Z \to \tau^+\tau^-$ process is significantly suppressed by the square of branching ratio $B(\tau \to \mu\nu_\mu\bar{\nu}_\tau)$. With these selection cuts the cosmic ray background estimated from the Cosmic Ray Tagger is expected to be less than 1%. A total of 1,727 data events passed these criteria with no same sign charge muons observed. That implies that QCD background is negligible.

The same selection criteria were applied to the Monte Carlo sample of the $Z \to \mu^+\mu^-$ events. The 540,000 events were generated in PYTHIA and further sumulated in the CDF software. Approximately 83,000 events survived the cuts.

The $\sum E_T$ distribution (see Figure 4.6 (top)) reveals some discrepancies between the simulation and the data. Here and in subsequent histograms the points with uncertainties are the data and the histogram is Monte Carlo simulation. This discrepncy is found to be largely accounted for by multiple interactions resulting in a larger $\sum E_T$ in the data. The number of primary vertices in an event can be found by z-vertex finding algorithm, which is based on the tracking information in the event. The z-vertex finder requires at least two COT tracks with $P_T > 300$ MeV to be associated to each vertex, i.e. be within 1 cm of silicon or within 5 cm of COT standalone vertex. The position of the vertex is obtained by weighting the z_0 of the tracks, corrected by beamline offset, by their error:

$$z = \frac{\sum_i z_i^0/\delta_i^2}{\sum_i 1/\delta_i^2},$$

where the sum is carried out over the tracks associated to the vertex, and δ_i is the error on z_i^0.

The number of reconstructed vertices follows a Poisson distribution with an average of about one vertex per event, as illustrated in Figure 4.5. About 25% of events in the data sample have two or more vertices.

Requiring no more than one vertex per an event found by the z-vertex algorithm, subtracts the effect of multiple interactions from the $\sum E_T$ in the event and gives better agreement between the data and Monte Carlo, as shown in Figure 4.6 (bottom). $U_{||}$ and $U_{\perp}$ components of the recoil energy vector are compared for data and Monte Carlo in Figure 4.7. $U_{\perp}$ distribution can be fitted with a gaussian of width $\sigma \simeq 3$ GeV around zero. The $U_{\perp}$ distribution is in fact a convolution of gaussians with different widths as functions of $\sum E_T$. To obtain $\sigma_{\perp} = \sigma_{\perp}(\sum E_T)$ dependence, we divide the sample into subsamples of events with about the same $\sum E_T$ (within 5 GeV bin width). Taking into account that due to the symmetry

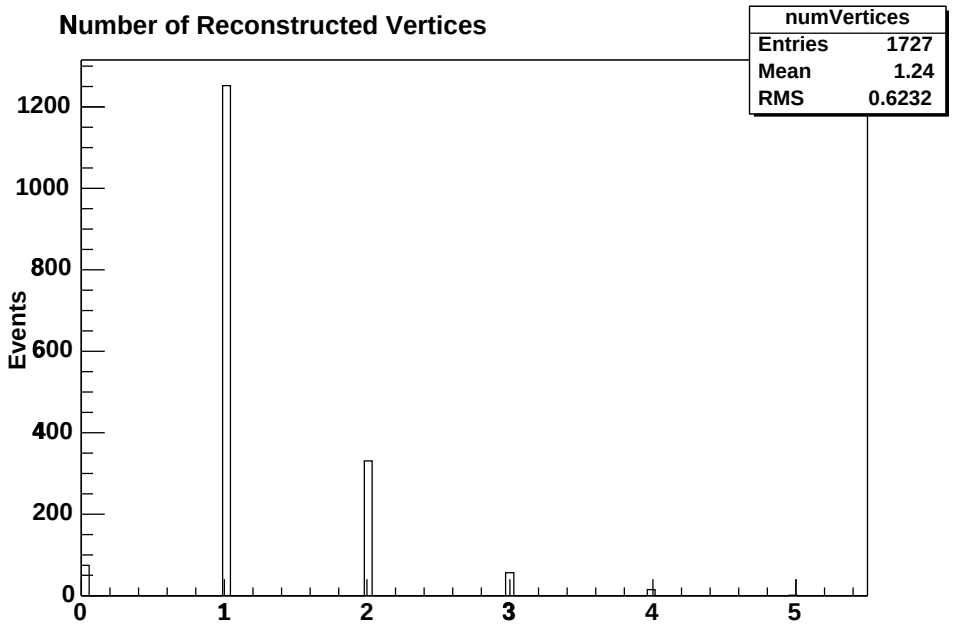

Figure 4.5: Number of reconstructed vertices in the data sample.

there is no preference in sign of $U_\perp$, i.e. $< U_\perp >= 0$, we determine RMS of $U_\perp$ distribution for each $\sum E_T$ bin. Figure 4.8 (top) illustrates the $\sigma_\perp(\sum E_T)$ dependence. The uncertainties are shown for data events. The Figure 4.8 (top) shows that the simulation slightly overestimates the $\not{E}_{\mathrm{T}}$ resolution. The effect of multiple interactions on the $U_\perp$ distribution is much smaller than it is on the $\sum E_T$, and $\sigma_\perp(\sum E_T)$ has about the same fit, if no requirement on number of vertices is applied to the data, as can be seen from Figure 4.8 (bottom).

The $\sigma_\perp(\sum E_T)$ dependence is fitted to the function $\sigma_\perp = a + b\sqrt{\sum E_T}$. The best χ^2-fit to the data points gives

$$\sigma_\perp = (-0.002 \pm 0.023) + (0.457 \pm 0.011)\sqrt{\sum E_T}$$

with $\chi^2/ndf = 0.43$. Hence the total $\not{E}_{\mathrm{T}}$ resolution is

$$\Delta = \sqrt{2}\sigma_\perp \simeq (0.646 \pm 0.016)\sqrt{\sum E_T}$$

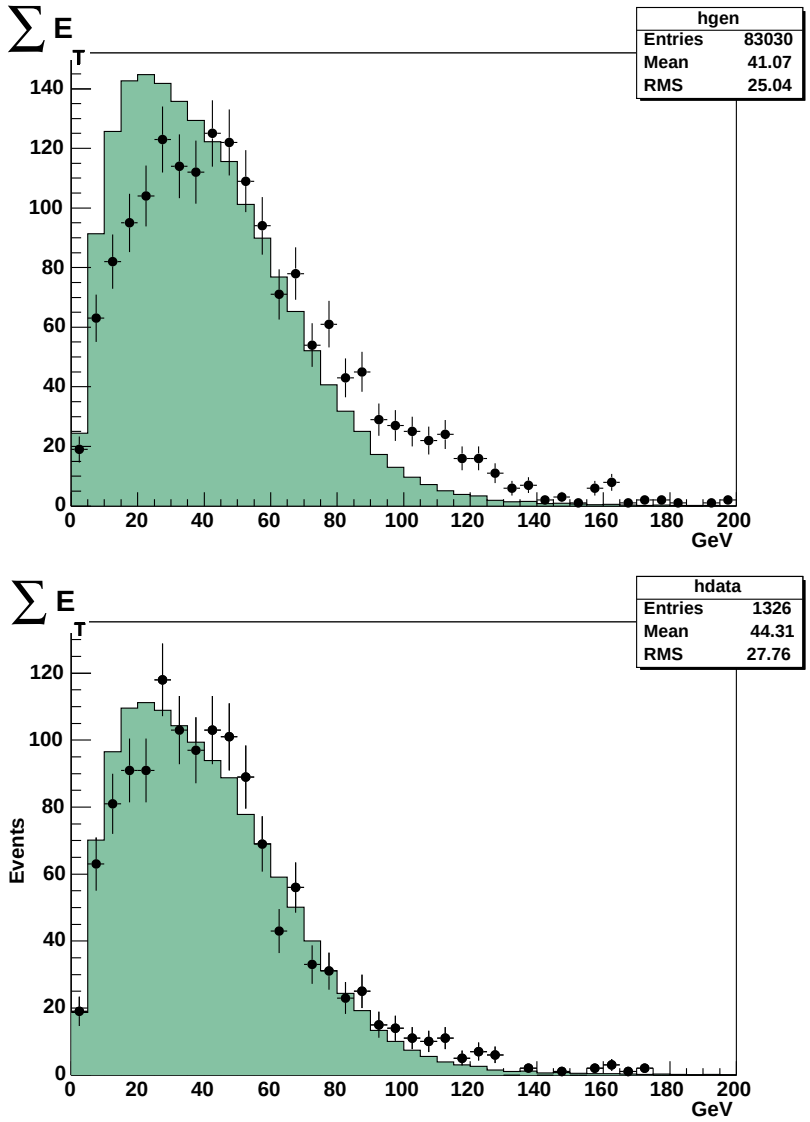

Figure 4.6: $\sum E_T$ distribution. The discrepancy between simulation and the data is due to multiple interactions (top). The effect of multiple interactions is excluded by requiring no more than one vertex per an event (bottom). Here and in subsequent histograms the points with error bars are the data and the histogram is Monte Carlo simulation.

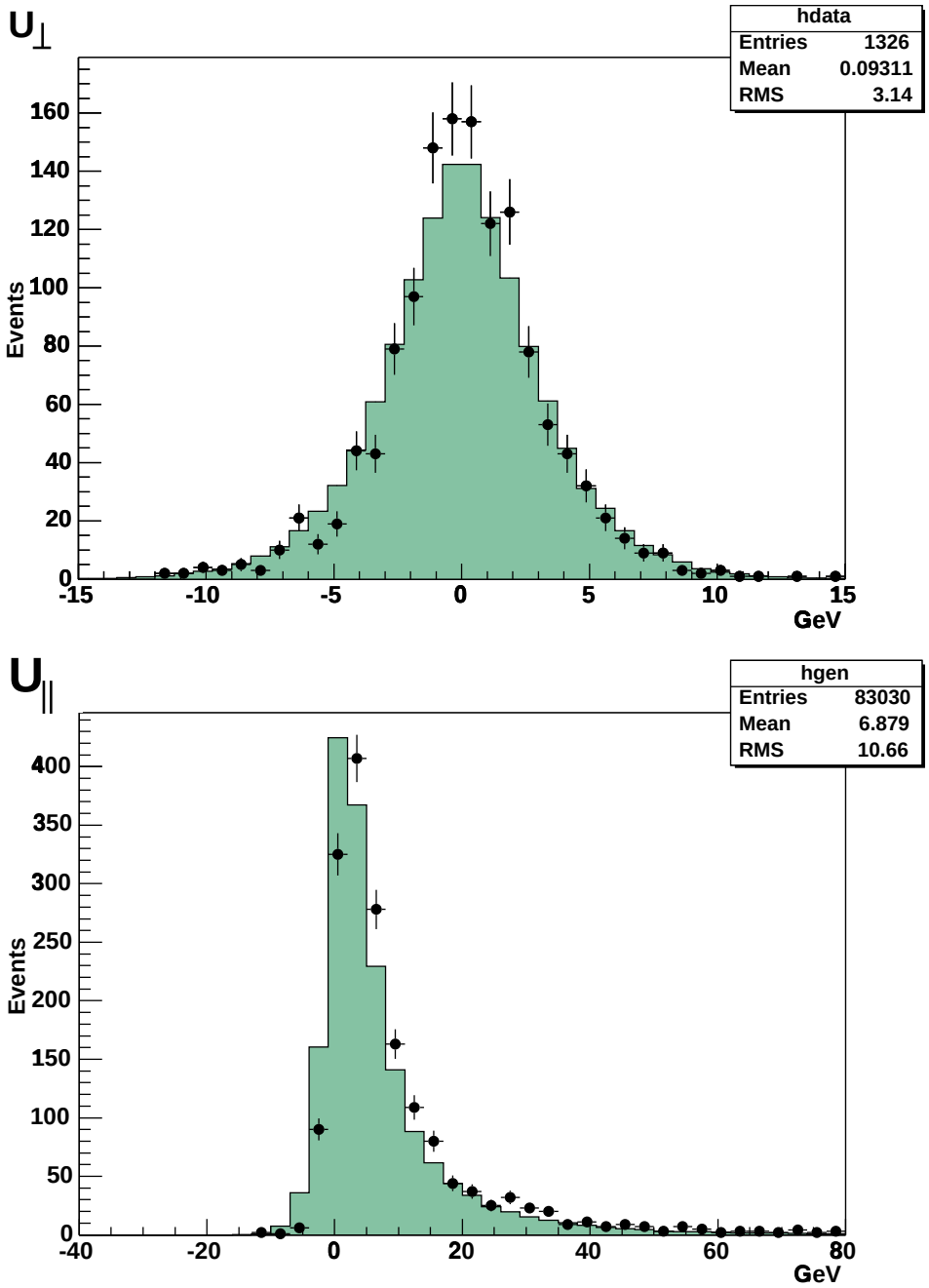

Figure 4.7: $U_\perp$ and $U_{||}$ distributions.

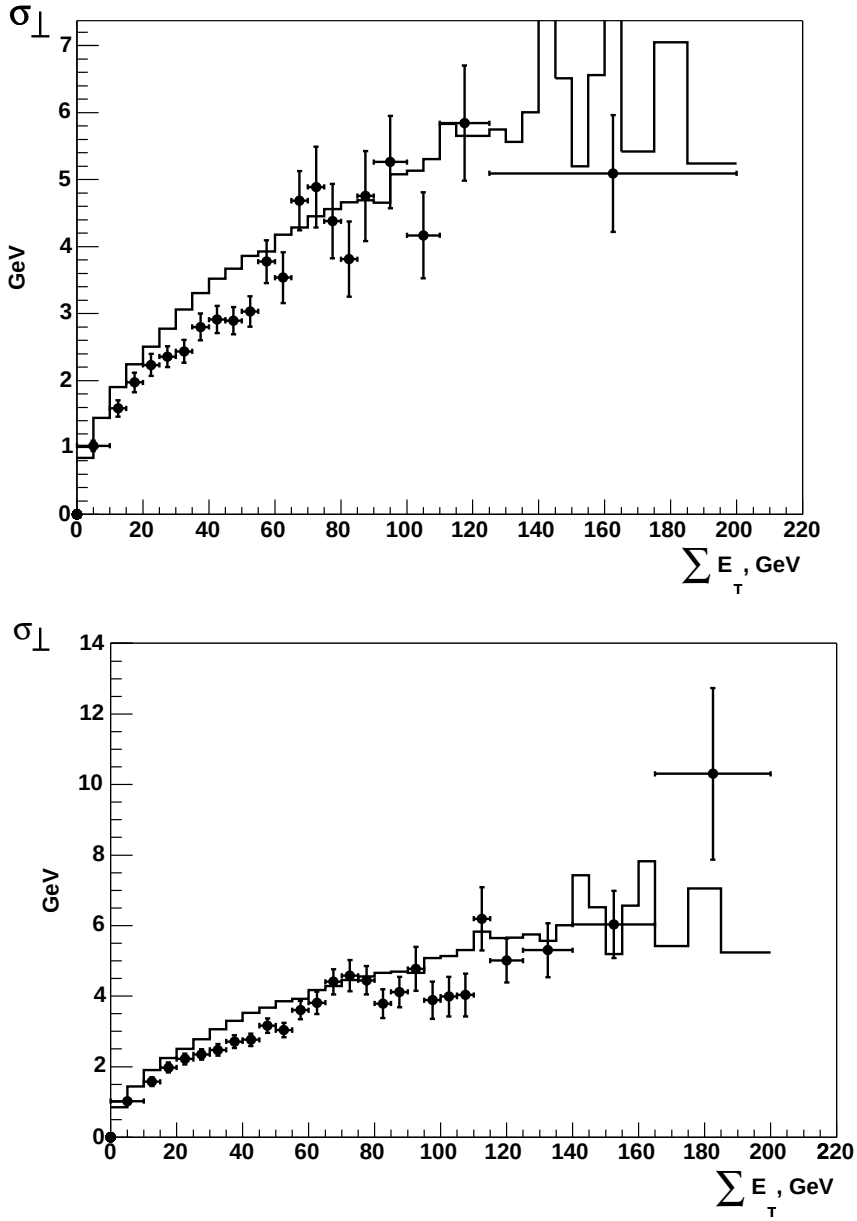

Figure 4.8: The $\not{E}_{\mathrm{T}}$ resolution $\sigma_{\perp}$ dependence on $\sum E_T$. The data events with one z-vertex only (top) and no cut on number of z-vertices (bottom).

This value agrees with minimum bias studies.

Whereas the fit to the simulation curve gives

$$\sigma_{\perp} = 0.436 + 0.438\sqrt{\sum E_T}$$

Regardless of the small discrepancy between simulated and observed $\sigma_{\perp}(\sum E_T)$ dependence, both Monte Carlo and data $U_{\perp}$ distributions for each $\sum E_T$ were found to be well described by gaussians, i.e. Monte Carlo simulation is adequately consistent with the data. This is demonstrated in Figure 4.9, where means of $U_{\perp}$ distributions are plotted as a function of $\sum E_T$, and there are shown fractions of events falling outside of one, two and three σ intervals, respectively, as functions of $\sum E_T$.

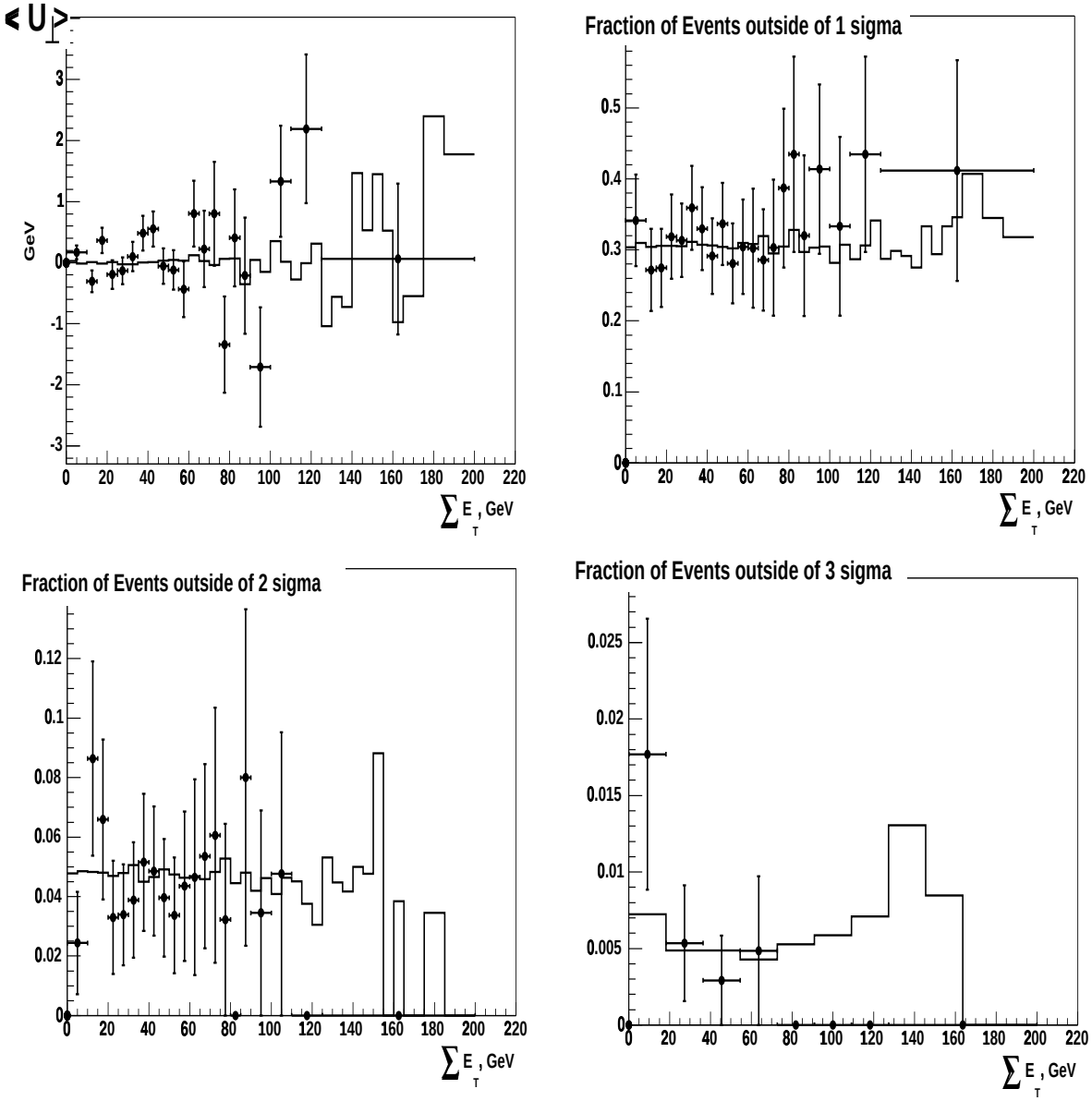

Figure 4.9: Mean of $U_\perp$ distributions and fractions of events outside of one, two and three σ intervals as functions of $\sum E_T$.

Chapter 5

Dilepton Event Selection

After describing the identification and reconstruction of the constituent objects in the final states of dilepton $t\bar{t}$ decays in the CDF detector, we discuss the global event selection criteria used to extract $t\bar{t} \to bW^-\bar{b}W^+ \to b\ell^-\nu\bar{b}\ell^+\bar{\nu}$ events in the data. The topological and kinematical event selection criteria are optimized to reduce the number of background events entering the sample while preserving the acceptance of the $t\bar{t}$ signal.

5.1 Trigger Pathes and Primary Datasets

The primary crude selection of the data sample of events is performed online at the trigger level. Events passing a certain set of rules specified at each of three trigger levels, called the *trigger path*, comprise the trigger dataset. Events satisfying the requirements of several trigger path enter the respective datasets simultaneously.

In the later stage, in the offline reconstruction of events several of the trigger datasets may be combined into one data sample with certain criteria defined on the offline level, as shown schematically in Figure 5.1. This analysis utilizes three

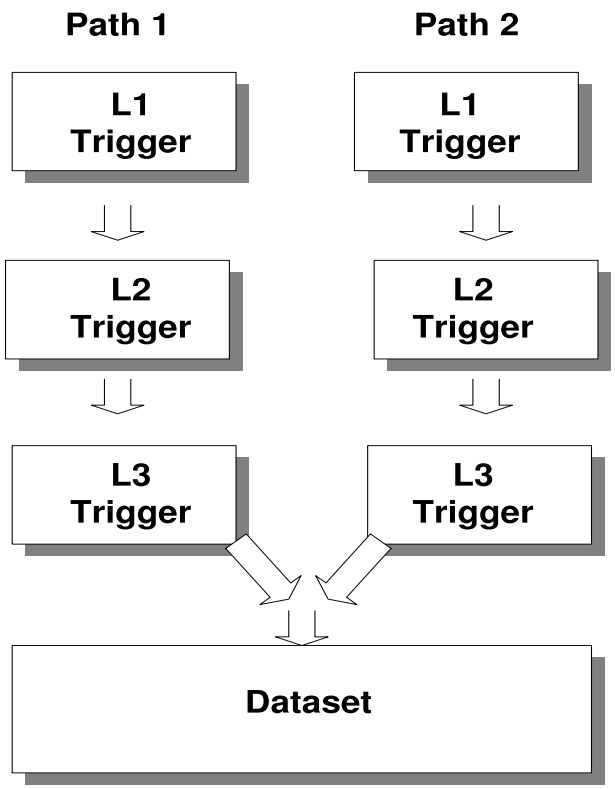

Figure 5.1: Schematic diagram of triggers, paths and datasets.

such data samples that are generally designated for studies on a broad spectrum of physics processes, and each of them contains events relevant to this study.

Two of these data samples are based on observing at least one high-P_T electron/muon in the central region of the detector ($|\eta| < 1.1$), and are called *inclusive electron/muon samples*, respectively. The third one requires at least one plug electron accompanied with large missing transverse energy and is referred to as the *plug data sample*.

5.1.1 Inclusive Electron Sample

Events from the inclusive electron data sample are selected online by requiring the following set of cuts applied on each consecutive trigger level:

- Level 1

  requires a CEM tower with EM $E_T \geq 8$ GeV and an XFT track pointing to this tower. If the tower energy EM $E_T \geq 14$ GeV, then the ratio of the hadronic to electromagnetic energy $E_{had}/E_{em} \leq 0.125$ is required. The XFT tracks must have momentum $P_T \geq 8.34$ GeV and have at least 10 hits in at least 3 layers.

- Level 2

  On the next level the online clustering is performed, the central cluster is required to have EM $E_T \geq 16$ GeV and $E_{had}/E_{em} \leq 0.125$. An XFT track of momentum $P_T \geq 8.34$ GeV must match the seed tower.

- Level 3

  requires a central EM offline cluster with $E_T \geq 18$ GeV and $E_{had}/E_{em} \leq 0.125$ with a COT track of momentum $P_T > 9$ GeV matching the EM cluster. For the data after January 2003, several additional requirements were added to the Level 3 path including $L_{shr} < 0.4$ and CES-track matching in $\hat{z}$: $|\Delta z| < 8$ cm. The transverse component of the COT track is recalculated based on the track z_0 instead of $z = 0$ and the E_{had}/E_{em} ratio is calculated for three hadronic towers instead of two.

These requirements constitute the ELECTRON_CENTRAL_18 trigger path. The combined calorimeter and track trigger efficiency for energetic electrons $E_T >$ 20 GeV was determined to be $\epsilon_{trig} = (96.6 \pm 0.1)\%$. [28].

Events with at least one trigger electron candidate are further filtered in the offline production code by applying relaxed electron quality cuts. At least one electron satisfying the requirements specified in Table 5.1 is required in each event.

Variables	Cut values for EM $E_T < 70$ GeV	Cut values for EM $E_T > 70$ GeV
E_T^{em}	> 18 GeV	> 18 GeV
E_{had}/E_{em}	< 0.125	-
E/P	< 4.0 OR $P_T > 9$ GeV	$P_T > 15$ GeV
L_{shr}	< 0.3	-
$\|\Delta x\|$	< 3.0 cm	< 3.0 cm
$\|\Delta z\|$	< 5.0 cm	< 5.0 cm

Table 5.1: High P_T inclusive electron sample stripping cuts.

5.1.2 Plug Data Sample

The plug data sample of events satisfies the requirements of the MET_PEM trigger path, where

- Level 1

  requires a PEM tower with EM $E_T \geq 8$ GeV and $E_{had}/E_{em} < 0.125$ and $\not\!\!E_{\mathrm{T}} > 15$ GeV. The missing energy is calculated summing the towers with $E_T > 1$ GeV.

- Level 2

  requires a cluster in the plug electromagnetic calorimeter with $E_T > 20$ GeV and $E_{had}/E_{em} < 0.125$

- Level 3

  requires a plug EM offline cluster with $E_T > 20$ GeV and $E_{had}/E_{em} \leq 0.125$, PEM $3 \times 3\chi^2 < 999$ and an offline $\not\!\!E_{\mathrm{T}} >$ GeV.

The additional trigger path, PLUG_ELECTRON_20, also contributes to the plug data sample. It is similar to the MET_PEM, but with no $\not\!\!E_{\mathrm{T}}$ requirement and has a prescale factor applied to Level 2.

The trigger efficiency for plug electrons was found to be $\epsilon_{trig} = (92.7 \pm 0.7)\%$ [29].

No further offline filtering is performed with the plug data sample.

5.1.3 Inclusive Muon Sample

The inclusive muon data sample is based on CMUP and CMX trigger pathes. The online CMUP muon identification follows the route, where

- Level 1

  attempts to match XFT track with both CMU and CMP hits. XFT provides tracks with $P_T \geq 4$ GeV and at least 11 hits per axial super layer.

- Level 2

  is auto-accepted for the early run period and requires a XFT track with $P_T \geq 8$ GeV for later runs.

- Level 3

  attempts to reconstruct muon COT track and links it to the CMU and CMP track segments or stubs. The P_T of the COT track must be greater than 18 GeV. The projected muon trajectory must match stubs in the plane transverse to the beamline within 10 cm for CMU and 30 cm for CMP chambers.

These requirements constitute MUON_CMUP18 trigger path. The CMUP trigger efficiency for energetic muons $P_T > 20$ GeV was determined to be $\epsilon_{trig} = (88.7 \pm 0.7)\%$ [30].

The CMX trigger path is similar:

- Level 1

  attempts to match XFT track of $P_T > 8$ GeV with with CMX hits on one of two projective wire pairs within a single CMX stack.

- Level 2

  is auto-accepted.

- Level 3

  attempts to reconstruct muon COT track of $P_T > 18$ GeV with a CMX stub. The projected muon trajectory is required to match stubs in the plane transverse to the beamline within 30 cm from the CMX stub.

The CMX trigger efficiency with a requirement on the exit radius of the track $\rho_{COT} > 140$ cm is $\epsilon_{trig} = (95.4 \pm 0.6)\%$ [30].

In order to reduce the size of these datasets, the events are filtered into a smaller dataset. At least one muon passing relaxed identification cuts as specified in Table 5.2 was required in each event.

5.1.4 Good Run List

Some runs are not acceptable for physics analyses because the beam was not stable or at least one major part of the detector or the data acquisition system did not operate properly. The functioning of every detector subsystem during

Variables	Cut values
P_T	> 18 GeV
E_{em}	$< 3 + max(0, 0.014 \times (P - 100)$ GeV
E_{had}	$< 9 + max(0, 0.042 \times (P - 100)$ GeV
$\|\Delta x\|_{CMU}$	< 5.0 cm
$\|\Delta x\|_{CMP}$	< 10.0 cm
$\|\Delta x\|_{CMX}$	< 20.0 cm

Table 5.2: High P_T inclusive muon sample baseline stripping cuts.

Data Sample	$\int \mathcal{L}dt$, pb^{-1}
CEM/CMUP	193
CEM/CMUP and CMX	175
CEM/CMUP and Si	162
CEM/CMUP and CMX and Si	150

Table 5.3: Total integrated luminosities for different data samples.

each run is monitored by the online processes. If at least one detector subsystem is observed to misfunction, such a run is marked 'bad' and is not included into this analysis. There is an exception from this rule, however, made for detector components irrelevant to this study.

For instance, the Silicon Tracking System is not crucial for this analysis and is mandatory only for Phoenix electron identification. Hence two lists of good runs are created. One requires the Silicon system properly operating during the run, named *Silicon good run list*, and the other one does not (*no Silicon good run list*). Objects are identified as Phoenix electron candidates only if the respective run is included into the Silicon good run list.

Similarly, due to operation problems with CMX chambers during early period of Run II, CMX muons do not enter the data sample for that period of time.

Due to these requirements the total integrated luminosities differ by data samples depending whether or not they include CMX muons or/and PHX electrons. The integrated luminosities range from 150 to 193 pb^{-1} depending on the detector requirements and are specified in Table 5.3.

5.2 Event Selection Criteria

The event selection criteria are preferentially imposed by the top dilepton decay chain signature, which requires two high-P_T leptons, electrons or muons, two high-E_T jets and large $\not\!\!E_{\mathrm{T}}$ due to neutrinoes in the event.

We define the top dilepton acceptance ϵ_{dil} as the fraction of all $t\bar{t}$ events that pass the event selection criteria. This implies that $t\bar{t}$ decay chains other than those, where both W bosons decay leptonically ($t\bar{t} \rightarrow bW^-\bar{b}W^+ \rightarrow b\ell^-\nu\bar{b}\ell^+\bar{\nu}, \ell = e$ or μ), also contribute to the top dilepton acceptance. These contributions come from

$t\bar{t}$ decays, in which either one W boson decays to a τ, which subsequently decays leptonically, or one W boson decays hadronically, and one of the b quarks decays to a high-P_T electron or muon directly $b \to c\ell\nu_\ell$, or through the cascade decay $b \to c \to s\ell\nu_\ell$.

The average number of observed events originating from $t\bar{t}$ decays after applying dilepton event selection cuts is then given by

$$S = \epsilon_{dil} \times \sigma_{t\bar{t}} \times \int \mathcal{L} dt, \tag{5.1}$$

where $\sigma_{t\bar{t}}$ is the $t\bar{t}$ cross section and $\int \mathcal{L} dt$ is the total intergated luminosity during a respective time period.

Various background sources of events with lepton pairs, accompanied by at least two jets, may resemble $t\bar{t}$ dilepton decays and therefore pass the event selection criteria. Such events typically involve higher order QCD processes beyond the tree level lepton pair production process. The actual number of events observed in the data would be, on average:

$$N_{obs} = S + B, \tag{5.2}$$

where B is an average contribution of the background sources. The event selection criteria are chosen to gain in the top dilepton acceptance, while concurrently reduce the relative contribution of background events B/N_{obs} thus increasing the purity of the top dilepton sample S/B.

5.2.1 Kinematic and Topological Characteristics of the Backgrounds

Before discussing the dilepton event selection we briefly overview the dominant backgrounds having the same signature as the $t\bar{t}$ dilepton events, and describe how they are kinematically and topologically different from $t\bar{t}$ decays.

One of the main backgrounds to the top dilepton sample results from pair produced electrons and muons in the Drell-Yan process $\gamma^\star/Z \to \ell\ell(\ell = e, \mu)$, where $q\bar{q}$ annihilation produces a virtual photon or a Z, which decay to lepton pairs. Higher order processes involve a QCD contribution of at least two jets in addition to the lepton pair event and thus mimic the $t\bar{t}$ dilepton decay signal.

The Drell-Yan cross section is largely dominated by the Z contribution near the Z^0 mass pole of 91 GeV. Therefore the Drell-Yan events can be efficiently eliminated by rejecting dielectron and dimuon events with the dilepton invariant mass in a Z mass window near Z^0 mass pole. In addition, due to the absense of neutrinoes in a Drell-Yan event, there is no intrinsic $\not\!\!E_{\mathrm{T}}$ and the majority of such events can be effectively removed by requiring significant $\not\!\!E_{\mathrm{T}}$.

Another source of dilepton background comes from Drell-Yan processes with decays to τ pairs, $\gamma^\star/Z \to \tau^+\tau^-$ when both τ's decay leptonically. The branching ratio for the tau leptonic decay $BR(\tau \to \ell\nu_\ell\nu_\tau), \ell = e$ or μ is 18% giving $BR(Z \to \tau^+\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau\ell^+\nu_\ell\bar{\nu}_\tau) = 13\%$. The invariant mass cut is not effective for removing $Z \to \tau^+\tau^-$ events, because the energy carried away by the neutrinoes in the τ decays diminishes the energy available to the electrons and muons, thus shifting the dilepton invariant mass lower. The four neutrinoes in the event contribute to the $\not\!\!E_{\mathrm{T}}$. However, the two τ's momenta are typically back to back if the Z does not acquire a high momentum due to recoil of a jet activity in the

event and the neutrino momenta mostly cancel leading to a small $\not\!\!E_{\rm T}$ in the event.

The next largest source of background is supplied by events where one of the leptons is faked by a hadronic particle. The muon can be misidentified due to a charged hadron punch through to the muon chambers or hadronic decay-in-flight, where a muon produced just before or in the calorimeter mimics a muon from the primary vertex. The electron can be impersonated by a highly energetic jet, for instance a jet which contains a photon or π^0 associated with a charged particle producing the track pointing to the electromagnetic cluster. The probability for a generic jet to fake a lepton is small ($\sim 10^{-4}$), therefore the fake lepton background events are mostly dominated by $W+ \geq 3$ jet events, in which the W decays leptonically, and one of the jets mimics the second lepton. The faked lepton will to the first order have no charge preference and roughly a half of the background is rejected by requiring opposite-sign dilepton events.

The next background results from vector boson pair production (WW,WZ,ZZ). The largest of these is WW with subsequent leptonic decays, $BR(W^+W^- \rightarrow \ell^+\nu_\ell\ell^-\bar{\nu}_\ell) = 5\%, \ell = e$ or μ, where decays to τ's $W \rightarrow \tau\nu_\tau$ also contribute of order 10%. The neutrinoes supply large $\not\!\!E_{\rm T}$ in the event. The production cross section is comparable with $t\bar{t}$, but only higher order diagrams produce jets in the event, thus decreasing overall background contribution from WW.

WZ events can resemble the $t\bar{t}$ dilepton events, when Z decays leptonically and W hadronically, although in this case there is no natural source for $\not\!\!E_{\rm T}$ and, in addition, leptons preferentially fall into the Z mass window. If the W decays leptonically, the branching ratio is three times smaller, and the requirement of two additional QCD jets reduces this background contribution even more.

The ZZ contribution is negligible. The cross-section $p\bar{p} \rightarrow ZZ$ and the branching ratio for one Z to decay leptonically and the other hadronically is about 9%,

while these events have no intrinsic $\not\!\!E_{\mathrm{T}}$ and dilepton invariant mass falls into the Z mass window.

The $b\bar{b}$ production can also contribute to the dilepton background primarily due to its high cross section (of order 5×10^6 times greater than the $t\bar{t}$). However, the $b\bar{b}$ kinematics are substantially different from $t\bar{t}$ decays. The neutrinoes in the event tend to be in the opposite directions and the leptons from semi-leptonic b decays will always be surrounded by hadronic activity. In addition, the P_T spectrum for leptons is rapidly falling.

These properties of the background sources motivate the dilepton event selection described below.

5.2.2 Dilepton Selection and Categorization

After preselecting events with at least one electron or muon candidate with relaxed ID requirements we create the secondary datasets of events by requiring at least one tight lepton in an event. We define the tight lepton as either the central/plug (CEM/PHX) electron passing all central/plug electron identification criteria as specified in Section 4.1 or the CMUP/CMX muon passing all muon identification criteria as discussed in Section 4.2. In the selected events the tight lepton is typically the object which actually triggers the event.

We can set looser requirements for the second lepton in order to gain in acceptance of the $t\bar{t}$ dilepton signal without significant decrease in the purity of the sample. We relax the second lepton identification requirements by dropping the isolation cut and relax track-stub match cuts for muons. This does not have a serious impact on the quality of the second lepton. As was already mentioned, the isolation is in fact not a lepton identification variable but rather a kinematic variable which separates leptons from hadronic activity in the event. Due to the

presence of some jet activity in $t\bar{t}$ dilepton events one of the two leptons is often not isolated and therefore dropping the isolation requirement allows us to gain in $t\bar{t}$ dilepton acceptance. This introduces additional backgrounds, however, from semi-leptonic b decays, as the leptons (products of b decays) will be inside of the b-jets. In addition, due to gaps in muon chambers coverage the signal acceptance can be enhanced by selecting events with second muons which have CMU/CMP-only stubs or stubless muons (CMIO), not fiducial to any of the muon detector regions. To reduce the fake rate due to lepton misidentification PHX electrons and CMIO muons are required to be always isolated. In addition, plug-plug dileptons are excluded due to negligible contributions to the signal acceptance ($\sim 1\%$), while accompanied by large fake background because of limited tracking information at high pseudorapidities (η).

Summarizing, the loose lepton category consists of non-isolated central electrons (NICEM), non-isolated CMUP/CMX muons (NICMUP/NICMX), isolated or non-isolated CMU/CMP-only (CMU/CMP/NICMU/NICMP) and stubless muons (CMIO). All possible dilepton categories are shown in Table 5.4.

Leptons in the event are required to be oppositely charged. This reduces the fake lepton background by approximately a factor of two, since there is no charge sign preference in dilepton events, when one of the leptons is misidentified. This cut almost has no effect on the dilepton acceptance. Same-sign $t\bar{t}$ dilepton events have one lepton from a W decay and the other from semileptonic decay of one of the b quarks. These events do not have a large contribution to the dilepton acceptance though, as the P_T of the lepton from a b quark decay is typically of order 8 GeV.

central-central			central-plug	
e/e	e/μ	μ/μ	e/e	e/μ
CEM-CEM	CEM-CMUP	CMUP-CMUP	CEM-PHX	PHX-CMUP
CEM-NICEM	CEM-CMX	CMUP-CMX	PHX-NICEM	PHX-CMX
	CEM-NICMUP	CMUP-NICMUP		PHX-NICMUP
	CEM-NICMX	CMUP-NICMX		PHX-NICMX
	CEM-CMU	CMUP-CMU		PHX-CMU
	CEM-NICMU	CMUP-NICMU		PHX-NICMU
	CEM-CMP	CMUP-CMP		PHX-CMP
	CEM-NICMP	CMUP-NICMP		PHX-NICMP
	CEM-CMIO	CMUP-CMIO		PHX-CMIO
	CMUP-NICEM	CMX-NICMUP		
	CMX-NICEM	CMX-NICMX		
		CMX-CMU		
		CMX-NICMU		
		CMX-CMP		
		CMX-NICMP		
		CMX-CMIO		

Table 5.4: Dilepton decay channels used in the analysis. The first lepton in the dilepton pair is the tight lepton and the second is either tight or loose.

5.2.3 Missing transverse energy $\not\!\!E_{\mathrm{T}}$ cut

The undetected neutrinoes from a $t\bar{t}$ dilepton decay result in a large missing transverse energy (see Figure 5.2). The magnitude of $\not\!\!E_{\mathrm{T}}$ is required to be greater than 25 GeV. This has the greatest reduction power for the Drell-Yan background, as can be seen in Figure 5.3. It also significantly reduces the number of dilepton events from $Z \to \tau\tau, b\bar{b}$ and fakes.

Large $\not\!\!E_{\mathrm{T}}$ in an event with no intrinsic missing transverse energy is typically due to jet energy mismeasurement. In this case the $\not\!\!E_{\mathrm{T}}$ tends to orient along a direction of mismeasured jet momentum. Such background can be reduced by adding a cut on the azimuthal angle $\Delta\phi$ between the $\not\!\!E_{\mathrm{T}}$ and the closest jet. Figure 5.4 shows the two-dimensional plots of $\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}$ angle versus $\not\!\!E_{\mathrm{T}}$ for simulated Drell-Yan events with at least two jets. It can be seen that the majority of Drell-Yan events, passing the $\not\!\!E_{\mathrm{T}} > 25$ GeV threshold, have $\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}$ below $20°$, and therefore most of them can be removed by requring $\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}} > 20°$ for $\not\!\!E_{\mathrm{T}} < 50$ GeV.

In the $Z \to \tau^+\tau^-$ events, due to the recoil of the Z boson from jet activity, the sum of the neutrino momenta can exceed the 25 GeV threshold. In those events, however, the neutrinoes from τ decay are usually boosted in the direction of the parent particle and the direction of $\not\!\!E_{\mathrm{T}}$ tends to be close to the lepton daughter of the τ, as illustrated in Figure 5.5. Hence the contribution from $Z \to \tau^+\tau^-$ events can be further reduced by applying similar $\Delta\phi_{\ell-\not\!\!E_{\mathrm{T}}} > 20°$ requirement for $\not\!\!E_{\mathrm{T}} < 50$ GeV, where $\Delta\phi_{\ell-\not\!\!E_{\mathrm{T}}}$ is an azimuthal angle between the direction of $\not\!\!E_{\mathrm{T}}$ and the nearest lepton momentum.

Therefore to reduce $Z \to \ell^+\ell^-, \ell = e, \mu$ or τ background we do not differentiate between leptons and jets and require the cut on the azimuthal angle $\Delta\phi_{obj-\not\!\!E_{\mathrm{T}}}$ between the $\not\!\!E_{\mathrm{T}}$ and the closest object (lepton or jet) momentum to be greater

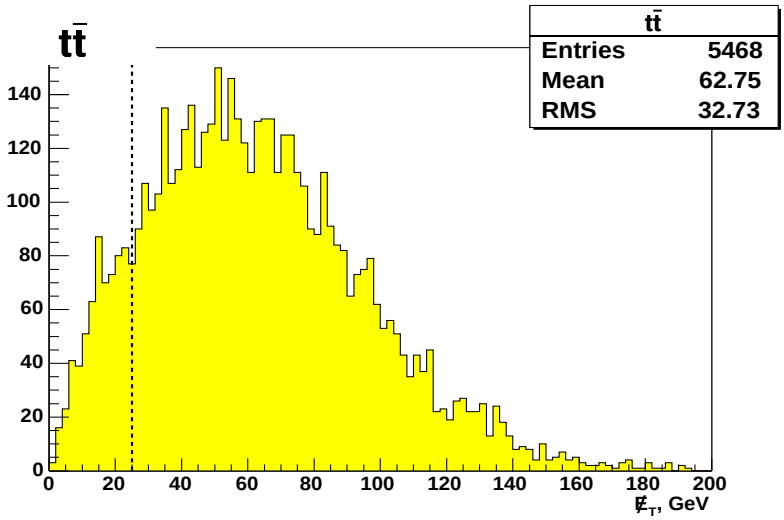

Figure 5.2: The missing transverse energy distribution of PYTHIA Monte Carlo $t\bar{t}$ dilepton events. The magnitude of $\not{E}_T$ is required to be greater than 25 GeV. The $\not{E}_T$ cut is indicated by the dash line.

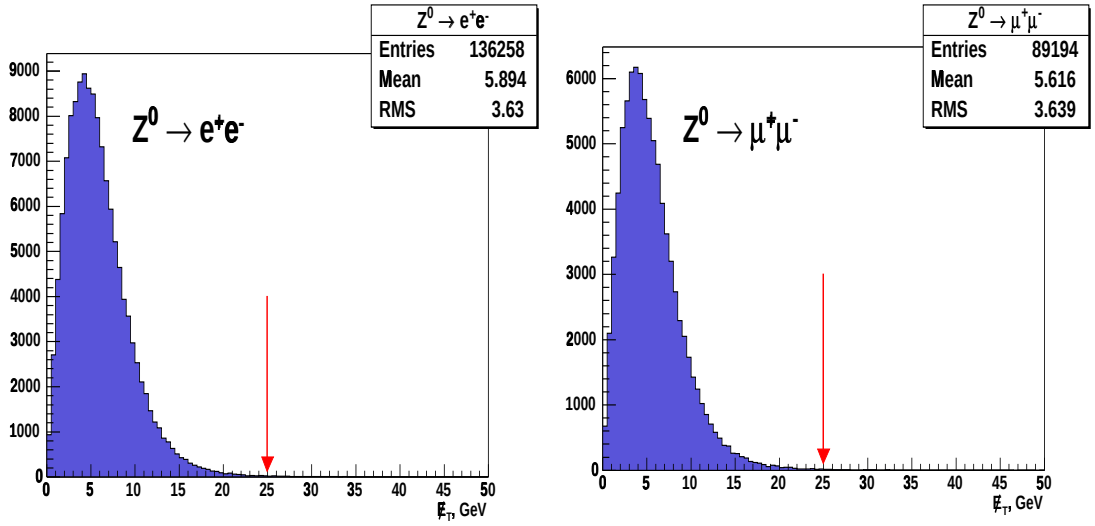

Figure 5.3: The missing transverse energy distribution of PYTHIA Monte Carlo $Z^0 \to e^+e^-$ events (left) and $Z^0 \to \mu^+\mu^-$ events (right) reconstructed in the CDF software. The arrow indicates the $\not{E}_T$ cut.

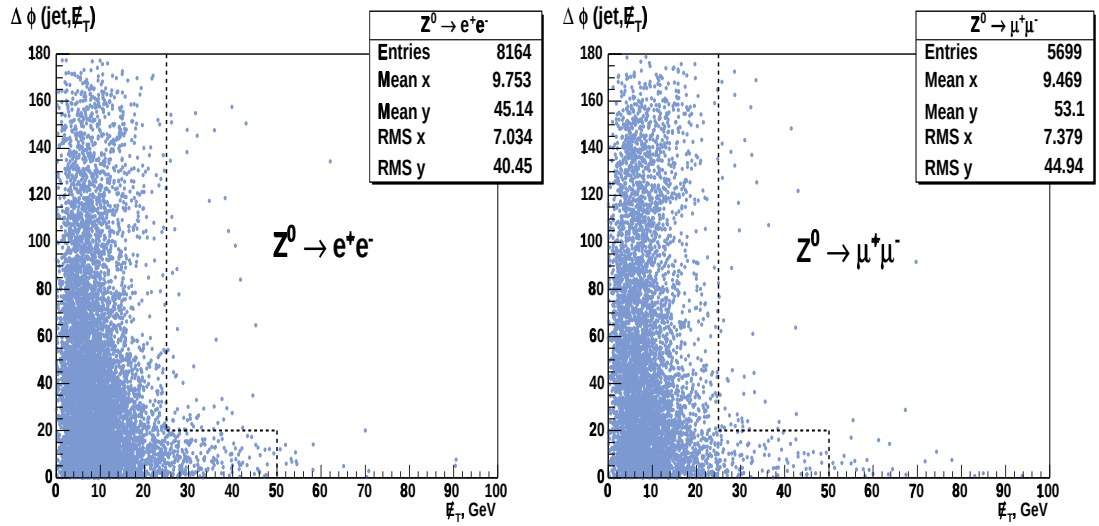

Figure 5.4: The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest jet versus the $\not\!\!E_{\mathrm{T}}$ for ALPGEN Monte Carlo $Z^0 + 2$ parton events with Z^0 decays to electron pairs (left) and muons pairs (right). The dashed line indicates the dilepton 'L'-cut applied in the event selection.

20° if $\not\!\!E_{\mathrm{T}} < 50$ GeV. The cut has an 'L'-shape in ($\Delta\phi-$ $\not\!\!E_{\mathrm{T}}$) plane and therefore it is sometimes referred as a dilepton 'L'-cut. The 'L'-cut significantly reduces Standard Model background while being 80% efficient for $t\bar{t}$ events (see Figure 5.6).

5.2.4 Z veto

The Drell-Yan background $\gamma^\star/Z \to \ell\ell(\ell = e, \mu)$ dominates the dielectron and dimuon events with about 90% of the ee and $\mu\mu$ events falling inside $76 \leq M_{\ell\ell} \leq 106$ GeV dilepton invariant mass window (see Figure 5.7). A conservative approach to reduce the Drell-Yan contribution would be to simply reject events with same flavor lepton pairs in the Z mass region. However, this cut would also result in about a 26% $t\bar{t}$ acceptance loss in the ee and $\mu\mu$ channels, as illustrated in Figure 5.8. The area on the plot between the dashed lines represents the Z invariant

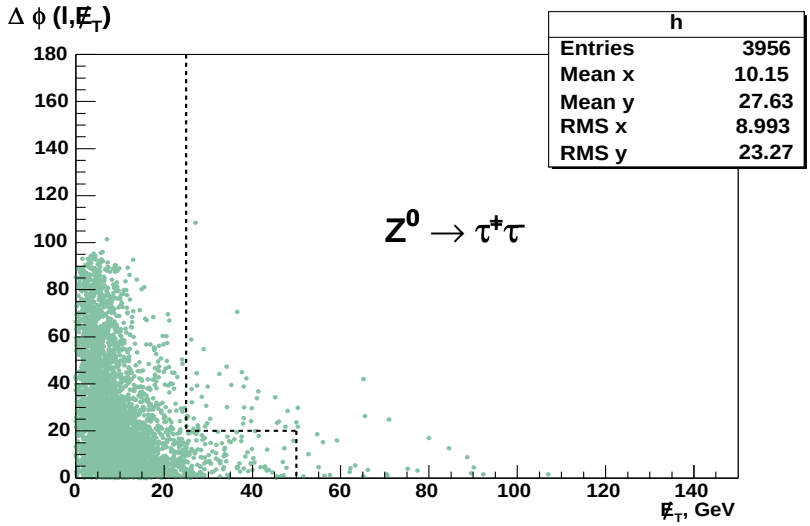

Figure 5.5: The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest lepton versus the $\not\!\!E_{\mathrm{T}}$ of PYTHIA Monte Carlo $Z^0 \to \tau^+\tau^-$ events reconstructed in the CDF software. The dashed line indicates the dilepton 'L'-cut applied in the event selection.

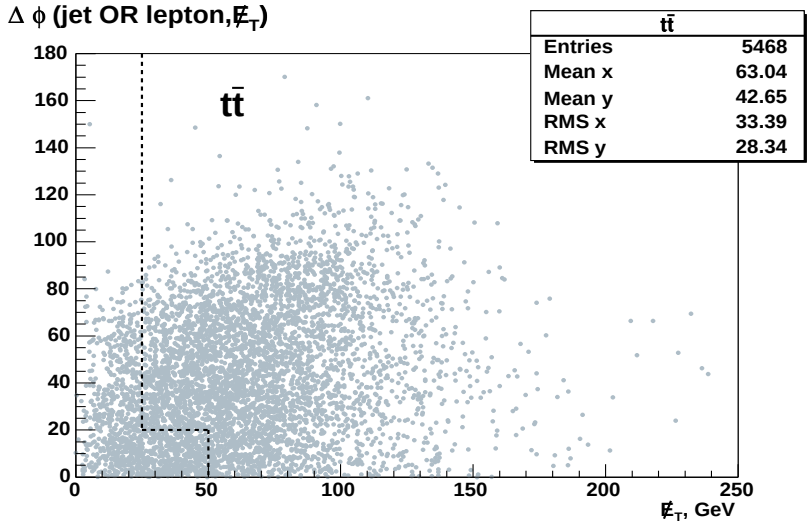

Figure 5.6: The azimuthal angle between the $\not\!\!E_{\mathrm{T}}$ and the nearest object (lepton or jet) versus the $\not\!\!E_{\mathrm{T}}$ of PYTHIA Monte Carlo $t\bar{t}$ dilepton events reconstructed in the CDF software. The dashed line indicates the dilepton 'L'-cut applied in the event selection.

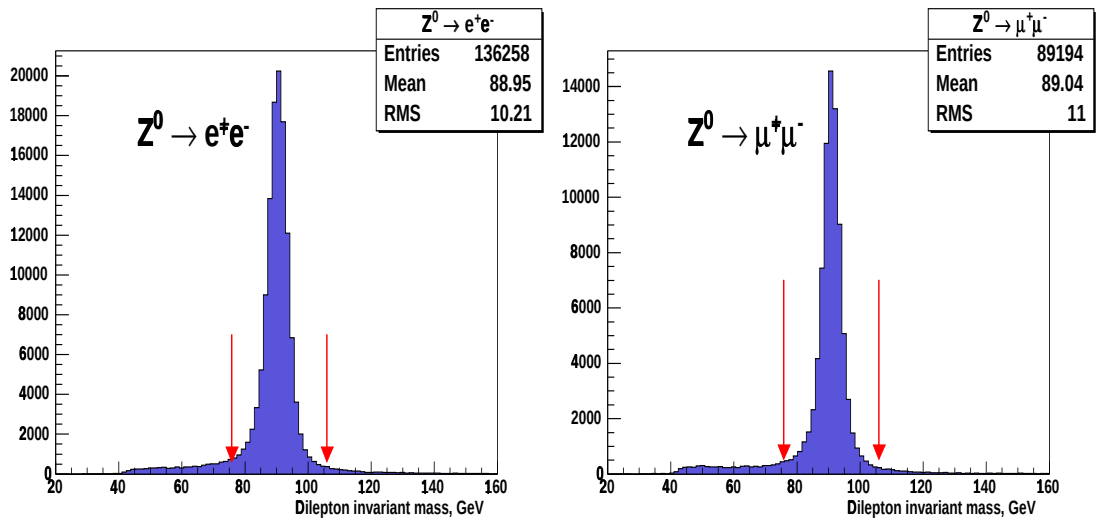

Figure 5.7: The dilepton invariant mass distribution of PYTHIA Monte Carlo $Z^0 \to e^+e^-$ events (left) and $Z^0 \to \mu^+\mu^-$ events (right) reconstructed in the CDF software.

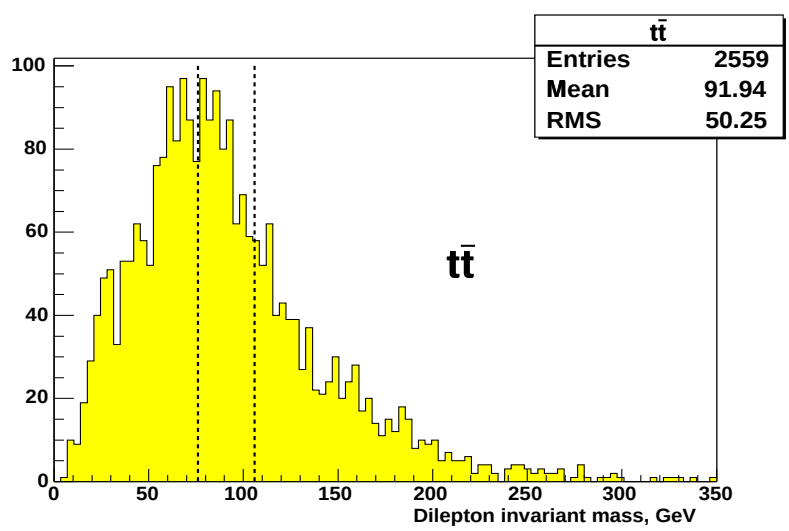

Figure 5.8: The dilepton invariant mass distribution of PYTHIA Monte Carlo $t\bar{t}$ events with lepton pairs of same flavor. The dashed lines indicate Z^0 mass window of 76 - 106 GeV.

mass window and corresponds to about a quarter of the entire distribution.

Instead, a significant fraction of the $t\bar{t}$ events in the Z mass region can be recovered by utilizing the fact that the Drell-Yan events do not possess an intrinsic $\not\!\!E_{\mathrm{T}}$. The majority of the Drell-Yan events are already left out by applying the $\not\!\!E_{\mathrm{T}}$ requirements discussed in the previous section. The events that still pass the event selection should have mismeasured magnitudes of $\not\!\!E_{\mathrm{T}} > 50$ GeV. This happens due to mismeasurements of very energetic jets $E_T > 100$ GeV, for instance, when they partially traverse through un-instrumented regions of the detector. Highly energetic jets are naturally subject to larger fluctuations in the E_T, which give rise to the $\vec{\not\!\!E}_{\mathrm{T}}$ along the mismeasured jet direction.

This effect can be quantified by making use of the variable, called *jet significance*, and defined as

$$jetsig = \frac{\not\!\!E_{\mathrm{T}}}{\sqrt{\sum \vec{E}_{T,jet} \cdot \vec{n}_{\not\!\!E_{\mathrm{T}}}}}, \tag{5.3}$$

where the sum $\sum \vec{E}_{T,jet} = \sum_{|\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}|<90^\circ} \vec{E}_{T,jet}$ is carried out over those jets, which directions are in the same hemisphere as the direction of $\vec{\not\!\!E}_{\mathrm{T}}$, i.e. $|\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}| < 90^\circ$, and $\vec{n}_{\not\!\!E_{\mathrm{T}}}$ is the unit vector collinear with $\vec{\not\!\!E}_{\mathrm{T}}$. The jet significance (Equation 5.3) takes into account possible mismeasurements in E_T of all jets in the same hemisphere as the $\vec{\not\!\!E}_{\mathrm{T}}$ and weights them depending on the magnitudes of their energies.

The Figure 5.9 shows the jet significance distributions of $Z \to \ell^+\ell^-$ events, which suggests a cut of $jetsig > 8$. This removes about 80% of the Drell-Yan events in the Z mass region while being $\sim 75\%$ efficient for $t\bar{t}$ dilepton events, as illustrated in the Figure 5.10. Events having jets in the other hemisphere are not on the plot as they have an infinite jet significance and therefore automatically pass the selection.

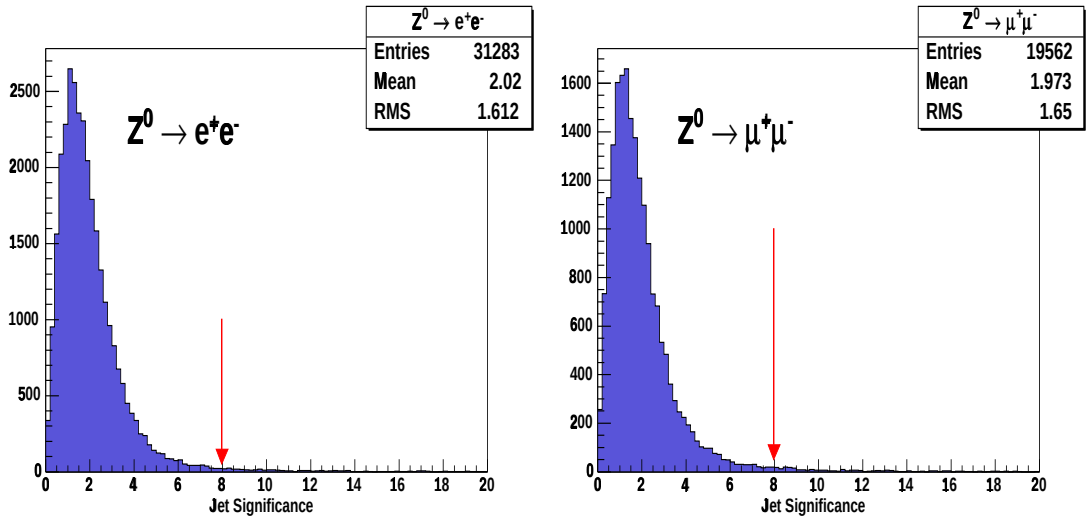

Figure 5.9: The jet signifance distribution of ALPGEN $Z^0 + 2$ partons Monte Carlo with Z^0 decays to electron pairs(left) and muons pairs(right) falling into Z^0 mass window of 76 - 106 GeV. The arrow indicates the cut value.

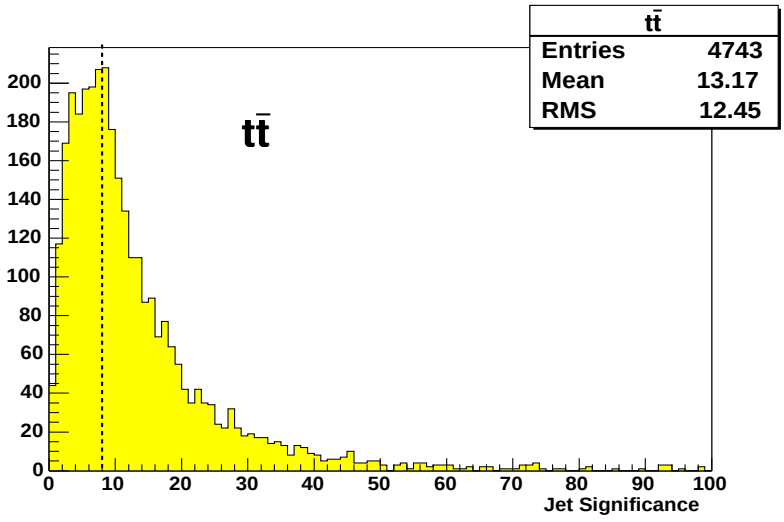

Figure 5.10: The jet significance distribution of PYTHIA Monte Carlo $t\bar{t}$ dilepton events with at least two jets. The dashed line indicates the cut value.

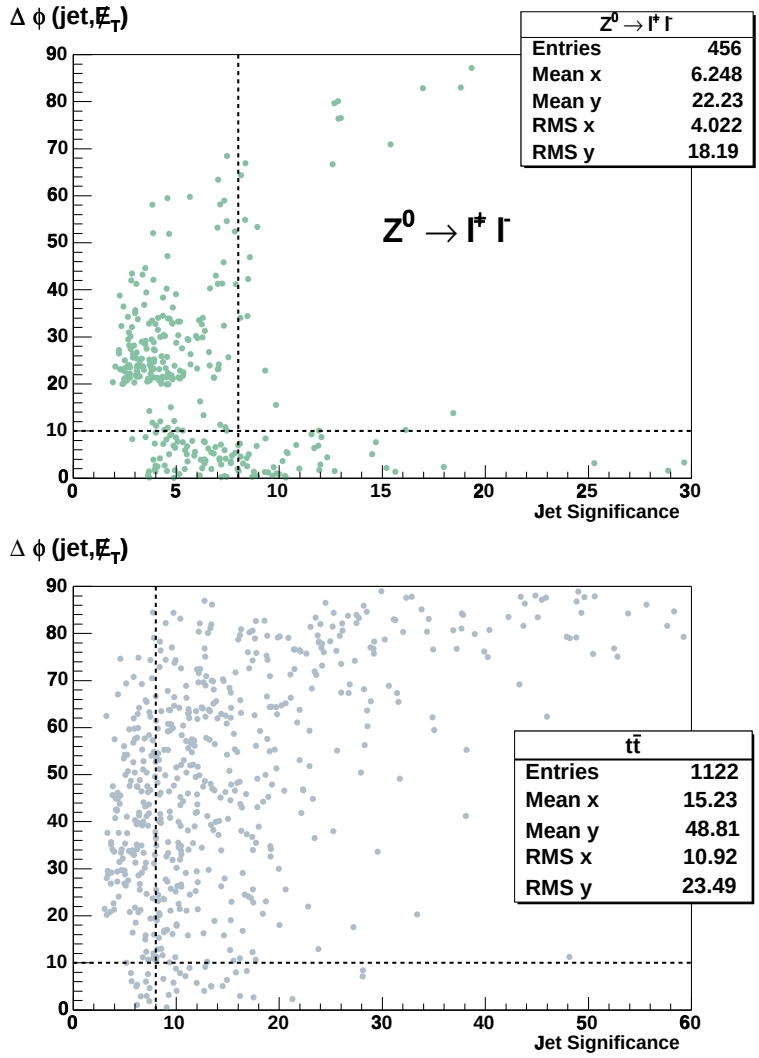

Figure 5.11: The azimuthal angle between the $\not\!\!E_T$ and the nearest jet versus the jet significance for ALPGEN Monte Carlo $Z^0 + 2$ partons events (top) and PYTHIA $t\bar{t}$ events (bottom) falling into Z^0 mass window of 76-106 GeV. The events were required to pass $\not\!\!E_T$ requirements. The dashed lines indicate the 'Z veto' event selection cuts, which effectively reduce the Drell-Yan background with a small effect on $t\bar{t}$ acceptance. In addition, about a half of $t\bar{t}$ events have jets in the opposite hemisphere from the $\not\!\!E_T$ and they automatically pass selection.

It is often the case that the $\not\!\!E_{\mathrm{T}}$ is mismeasured due to a fluctuation in E_T of only one of the jets. Then the vector $\vec{\not\!\!E}_{\mathrm{T}}$ tends to be in the direction of the mismeasured jet rather than along $\sum \vec{E}_{T,jet}$. The Figure 5.11 shows distributions of the $Z \to \ell^+\ell^-$ and $t\bar{t}$ dilepton events in the two-dimensional plane $(\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}, jetsig)$, where $\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}}$ is the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and the nearest jet. The plotted events pass the $\not\!\!E_{\mathrm{T}}$ requirements. It can be seen from this Figure that an additional requirement $\Delta\phi_{jet-\not\!\!E_{\mathrm{T}}} > 10°$ allows further reduction of the Drell-Yan contamination while retaining the $t\bar{t}$ acceptance.

5.2.5 H_T cut

The large mass of the top quark leads to higher transverse momenta of the final decay products than in most other processes. We exploit this fact by using the scalar sum of the transverse energies of all objects in the event H_T, which is defined as

$$H_T = \sum_{electrons} E_T + \sum_{muons} P_T + \sum_{jets} E_T + \not\!\!E_{\mathrm{T}} \tag{5.4}$$

and is a good variable for separation between $t\bar{t}$ and the backgrounds.

We require an event to have $H_T > 200$ GeV in order to further enhance the purity of the $t\bar{t}$ dilepton sample. This cut is 96% efficient for $t\bar{t}$ events, as can be seen in Figure 5.12, while rejects about 35% of the WW and $Z \to \tau\tau$ events, as shown in Figure 5.13.

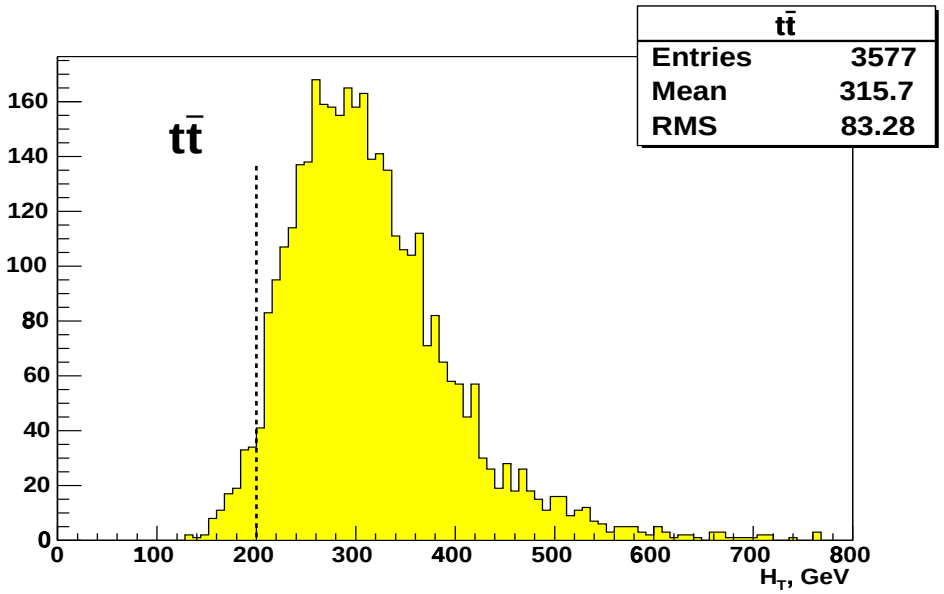

Figure 5.12: The Ht distribution of PYTHIA Monte Carlo $t\bar{t}$ events passed $\not\!\!E_{\mathrm{T}}$ and Z veto requirements. The dashed line indicates the cut value of $H_T > 200$ GeV.

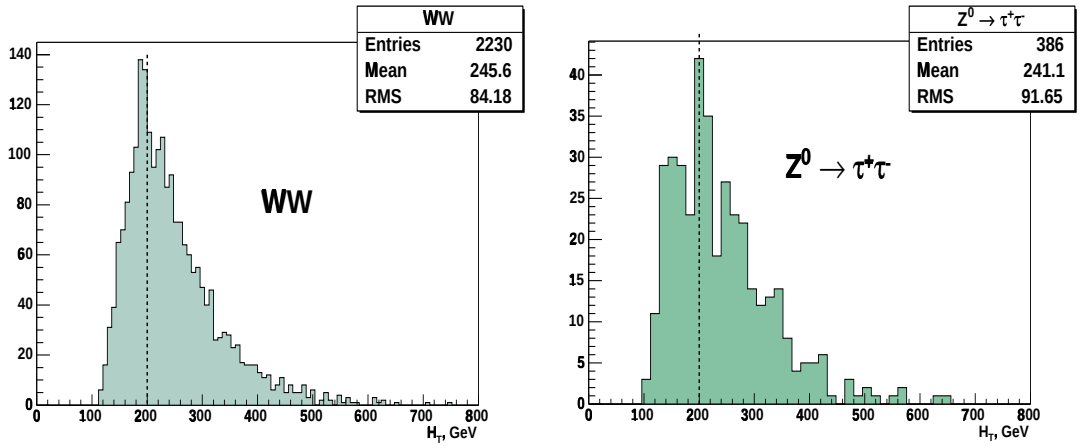

Figure 5.13: The Ht distribution of PYTHIA WW and $Z \to \tau\tau$ events passed $\not\!\!E_{\mathrm{T}}$ and Z veto requirements. The dashed line indicates the cut value of $H_T > 200$ GeV.

5.3 Top Dilepton Acceptance and Background Estimates

5.3.1 Top Dilepton Acceptance

The top dilepton acceptance ϵ_{dil} is the probability for a $t\bar{t}$ event to pass all the dilepton event selection criteria discussed in the previous section, i.e. the probability to

- produce two oppositely charged leptons passing fiducial and kinematic identification requirements described in Chapter 4 and satisfying to criteria of one of the dilepton categories as discussed in Section 5.2.2.
- produce neutrinoes such that the $\not{E}_{\mathrm{T}}$ requirements are satisfied (Section 5.2.3).
- produce at least two jets passing the jet requirements (Section 4.3).
- ee and $\mu\mu$ events must not be vetoed as Z candidates (Section 5.2.4).
- must have H_T (Equation 5.4) greater than 200 GeV.

The acceptance for $t\bar{t} \to \ell\bar{\ell}$ channel can therefore be written in terms of efficiencies to pass each requirement:

$$\epsilon_{dil} = \epsilon_{geom\cdot kin} \cdot \epsilon_{ID}^{lepton} \cdot \epsilon_{iso} \cdot \epsilon_{conv} \cdot \epsilon_{Z_{veto}} \cdot \epsilon_{\not{E}_{\mathrm{T}}} \cdot \epsilon_{2-jet} \cdot \epsilon_{H_T} \cdot \epsilon_{OS} \cdot \epsilon_{trigger}^{lepton} \cdot \epsilon_{MC/data}^{lepton}, \quad (5.5)$$

where

- $\epsilon_{geom\cdot kin}$ is the geometrical and kinematical acceptance, the ratio of dilepton events, where both leptons have P_T greater than 20 GeV and pass through

fiducial region of the detector, to the total number of $t\bar{t}$ events. This constitutes to about a half of the dilepton branching fraction as seen from Table 5.5.

- ϵ_{ID}^{lepton} is the lepton ID efficiency described in Chapter 4.
- ϵ_{iso} arises from requirement to have at least one lepton isolated.
- ϵ_{conv} accounts for electrons identified as conversions.
- ϵ_{cut} represent efficiencies of each consecutive selection cut
- Trigger efficiencies $\epsilon_{trigger}^{lepton}$ and the ID scale factors $\epsilon_{MC/data}^{lepton}$ account for differences between data and simulation. They are applied on the last stage.

These efficiencies are estimated sequentially for each dilepton category by using the PYTHIA [31] Monte Carlo sample of $\sim 400,000$ inclusive $t\bar{t}$ events, which has a leading order matrix element for the parton hard scattering, convoluted with the CTEQ5L parton disribution functions [32]. The effective luminosity of the sample, based on a theoretical cross-section $\sigma_{t\bar{t}}^{NLO} = 6.7$ pb [8], is 59.4 fb^{-1}. The event selection used for the acceptance estimates is almost the same as applied to data with a few differences outlined below:

- Curvature corrections (Equation 4.1) are not applied to Monte Carlo events.
- Monte Carlo version of jet correction factors is applied to account for discrepancies between data and simulation.
- The cosmic ray filter is not used.
- No opposite sign cut is applied to dilepton evente with PHX; the charge fake rate observed in the data is used instead.

$t\bar{t}$	
Selection cut	ϵ, %
$\epsilon_{geom \cdot kin}$	2.69 ± 0.03
ϵ_{ID}^{lepton}	49.2 ± 0.5
ϵ_{iso}	95.4 ± 0.3
ϵ_{conv}	97.4 ± 0.2
$\epsilon_{Z_{veto}}$	97.2 ± 0.2
$\epsilon_{\not{E}_T}$	81.5 ± 0.6
ϵ_{2-jet}	85.3 ± 0.6
ϵ_{H_T}	96.4 ± 0.3
ϵ_{OS}	98.1 ± 0.2

Table 5.5: The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive $t\bar{t}$ events and giving rise to the total dilepton acceptance $\epsilon_{dil}^{MC} = (0.813 \pm 0.014)\%$. Errors are statistical only. The corrections for lepton trigger efficiency and lepton ID scale factor reduce the acceptance by 15%.

z_{vertex} for Monte Carlo events is required to be within 60 cm of the nominal interaction point, in agreement with data.

The detail breakdown of $t\bar{t}$ events per each dilepton category can be found in [33]. The efficiences of each event selection cut combined for all dilepton categories are presented in Table 5.5. Their product results in the dilepton acceptance of $(0.813 \pm 0.014)\%$, which is further corrected for the trigger efficiencies $\epsilon_{trigger}^{lepton}$ and the Monte Carlo/ data scale factors $\epsilon_{MC/data}^{lepton}$ to account for differences between data and simulation. These efficiences are applied per each dilepton category and reduce the overall top dilepton acceptance by 15%.

The presented uncertainties in Table 5.5 are statistical only. The systematic uncertainties were determined by applying the event selection requirements to Monte Carlo samples with different modelling of the $t\bar{t}$ production and decay as follows:

- *Event generator.* We chose PYTHIA as the default Monte Carlo event generator. As an alternative we make use of the sample of $t\bar{t}$ events generated with HERWIG [34]. The two generators give a difference in acceptance of about 10%. Half of the difference can be explained by the value $B(W \to \ell\nu)$ assumed by two generators. HERWIG exploits the theoretical value of 0.111 and PYTHIA uses the measured value of 0.108. Therfore an effect of W leptonic branching fractions is corrected by reweighting Herwig acceptance by a factor of $(0.108)^2/(0.111)^2 = 0.947$.

  Another known difference is due to the fact that HERWIG does not model QED final state radiation. Hence the systematic error in the acceptance due to Monte Carlo modelling is performed by comparing the HERWIG reweighted sample with a PYTHIA sample generated with no QED FSR. The observed difference (5.5%) is a systematic error due to different MC generators.

- *Initial/Final state radiation.* High transverse momentum gluons can be radiated off the initial or final state partons. This gluon radiation is referred as initial/final state radiation (ISR/FSR) respectively. The radiated gluons form jets, thus adding a number of jets per an event and affecting the jet energy distributions. These effects are estimated by using two PYTHIA samples with different Λ_{QCD} values and K-factors for the transverse momentum scale of the ISR evolution. The range of variation was determined by taking extremes of a range in a study of Drell-Yan $Z \to \ell\ell$ events in

data and Monte Carlo [35]. The uncertainty from the modelling of final state gluon radiaton is estimated with the same variations applied to the FSR evolution. In addition to the hard scattering process and initial and final state radiation, remnants of the proton and anti-proton and their interactions, affect event kinematics. To estimate the uncertainty from the modeling of the underlying event, we use a PYTHIA $t\bar{t}$ Monte Carlo sample with parameters describing the charged particle multiplicity in dijet data retuned assuming less ISR [36]. The total systematic error due to ISR and FSR radiation effects is estimated to be 1.7% by taking the quadrature of the relative difference of the two samples with respect to the default one.

- *Parton Distribution Functions.* A parton distribution function (PDF) describes how the momentum fraction of the partons inside of a hadron is distributed. We make use of the CTEQ parton distribution function for the central value, as they fit CDF's inclusive jet cross section. There are several contributions due to the uncertainty in PDF. First is the choice of PDF. To estimate it, we use the sample with the MRST PDF at the same value of the strong coupling constant α_s as the default sample. The second contribution involves the difference between this MRST sample and MRST sample with a lower value of $\alpha_s = 0.112$. The third contribution was estimated by varying the CTEQ6M eigenvector to give the maximum variation between the dominance of the quark and gluon distribution functions. All of these contributions are summed up quadratically resulting in the overall 11.6% uncertainty due to PDF.

- *Jet Energy Scale.* Systematic uncertainties in the jet energy scale affect the $t\bar{t}$ acceptance since the jets may pass or not pass the $E_T > 15$ GeV threshold. This also affects the $\not\!\!E_{\mathrm{T}}$, jet significance and H_T for the event. The change

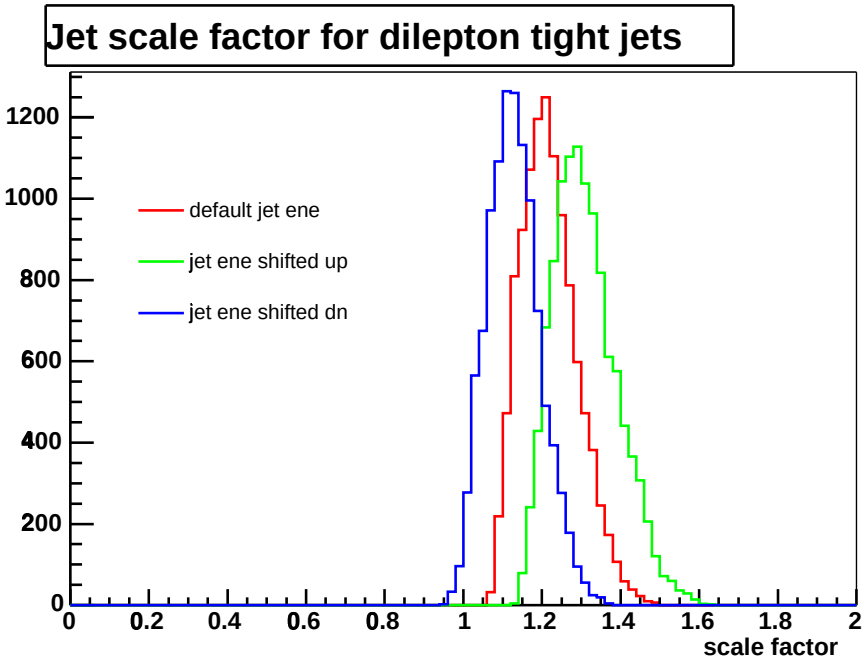

Figure 5.14: Distribution of jet correction scale factors applied to the raw jet energies. Shifted up/down jets result in larger/smaller scale factors.

in the default acceptance is evaluated by shifting the jet corrections up and down by their systematic uncertainty. The change in jet corrections scale factor due to these shifts is shown in Figure 5.14. The final systematic error due to jet corrections is taken as the half-difference of the shifted up and shifted down value divided by the default acceptance and is equal to 4.7%.

All of these systematic effects combined in quadratures result in the overall uncertainty of 14%.

The number of expected $t\bar{t}$ events in the data sample is given by Equation 5.1, which is rewritten as follows:

$$S \equiv N_{exp}^{t\bar{t}} = \sigma_{t\bar{t}}^{NLO} \times \sum_{dileptons} \left(\epsilon_{dil}^{t\bar{t}} \times \int \mathcal{L} dt \right), \tag{5.6}$$

where the sum is carried out over the dilepton categories to account for differencies in integrated luminosities, as given in Table 5.3. Using $\sigma_{t\bar{t}}^{NLO} = 6.7pb^{-1}$ [8] and the calculated uncertainties in the dilepton acceptance the expected number of $t\bar{t}$ events in the data sample is 8.2 ± 1.1 events.

5.3.2 Vector Boson Pair Production WW/WZ Estimate

The expected number of WW events is estimated by using PYTHIA Monte Carlo sample of 828,000 events with the W forced to decay leptonically $WW \to \ell\bar{\ell}$, where $\ell = e, \mu$ or τ. Using $\sigma_{WW}^{NLO} = 13.25$ pb [38] and taking into account the branching ratio for both W's to decay leptonically $BR(WW \to \ell\nu\ell\nu) \simeq 9.88\%$, the effective luminosity of this Monte Carlo sample is $\sim 630fb^{-1}$.

Table 5.6 shows efficiencies for WW events to pass each event selection requirement. Note, that higher $\epsilon_{geom \cdot kin}$ compared with the sample of inclusive $t\bar{t}$ events is because both W's are decaying only leptonically here. Also notice, that the main reduction in WW events is due to the 2-jet cut, which requires two high E_T jets, supplied by ISR and/or FSR jets in WW events. This efficiency is underestimated, as the fragmentation of jets is not modelled accurately in PYTHIA Monte Carlo. The jet multiplicity distribution from the PYTHIA sample of $Z \to e^+e^-$ events under the Z peak was tuned to match that of the data and the correction factor ϵ_{2-jet}^{corr} for events with ≥ 2 jets was found to be 1.68 ± 0.15. After this correction is applied, ϵ_{dil}^{WW} is in better agreement with the one determined using ALPGEN [37] Monte Carlo sample of WW + 2 parton events, convoluted with the HERWIG parton shower algorithm to evolve the final state partons to colorless hadrons. This Monte Carlo is believed to provide the best model for the fragmentation of events with two partons in the final state. The discrepancy between the event generators is assigned as a source of systematic uncertainty. The jet

corrections uncertainties are estimated in a similar manner as for $t\bar{t}$ events and account for 18% for WW acceptance.

The number of WW events is given by

$$N_{exp}^{WW} = \sigma_{WW}^{NLO} \times BR(WW \to \ell\nu\ell\nu) \times \sum_{dileptons} \left(\epsilon_{dil}^{WW} \times \int \mathcal{L} dt \right), \tag{5.7}$$

and is determined to be 0.61 ± 0.17.

For the WZ background estimate we make use of PYTHIA sample of 1,390,000 inclusive WZ events. Based on $\sigma_{WZ}^{NLO} = 3.96$ pb [38], the effective luminosity of the Monte Carlo sample is $\sim 350 fb^{-1}$. The list of consecutive efficiencies is given in Table 5.7. Since in WZ events which pass the event selection the jets are supplied mostly from hadronic W decays we do not exploit a correction factor for ϵ_{2-jet}. The Monte Carlo acceptance $\epsilon_{WZ}^{MC} = (0.021 \pm 0.001)\%$ is further corrected for lepton trigger efficiences and ID scale factors. Major systematic uncertainty is due to jet corrections, which is $\sim 23\%$.

The WZ estimate is given by

$$N_{exp}^{WZ} = \sigma_{WZ}^{NLO} \times \sum_{dileptons} \left(\epsilon_{dil}^{WZ} \times \int \mathcal{L} dt \right), \tag{5.8}$$

and is estimated to be 0.13 ± 0.03 events.

5.3.3 $Z \to \tau\tau$ Background Estimate

The $Z \to \tau^+\tau^-$ estimate is determined from a PYTHIA sample of 422,363 $Z/\gamma^\star \to \tau^+\tau^-$ events with τ-lepton decays handled by the TAUOLA [39] software package. The sample was filtered by requiring that the generated center-of-mass, E_{CM}, be greater than 30.0 GeV and two generator-level leptons produced with $E_T > 15$

$WW \to \ell\bar{\ell}$	
Selection cut	ϵ, %
$\epsilon_{geom \cdot kin}$	17.19 ± 0.04
ϵ_{ID}^{lepton}	58.76 ± 0.13
ϵ_{iso}	97.97 ± 0.05
ϵ_{conv}	98.07 ± 0.05
$\epsilon_{Z_{veto}}$	99.58 ± 0.02
$\epsilon_{\not{E}_T}$	69.97 ± 0.16
ϵ_{2-jet}	4.07 ± 0.08
ϵ_{H_T}	65.32 ± 1.00
ϵ_{OS}	98.25 ± 0.34

Table 5.6: The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of $WW \to \ell\bar{\ell}$ events. Errors are statistical only. This gives rise to $\epsilon_{WW}^{MC} = (0.177 \pm 0.005)\%$. The correction factor of 1.68 ± 0.15 is applied to account for underestimated value of ϵ_{2-jet} from PYTHIA. The lepton trigger efficiencies and ID scale factors further decrease the acceptance by 15%.

WZ	
Selection cut	ϵ, %
$\epsilon_{geom \cdot kin}$	1.96 ± 0.01
ϵ_{ID}^{lepton}	70.82 ± 0.28
ϵ_{iso}	97.69 ± 0.11
ϵ_{conv}	98.00 ± 0.10
$\epsilon_{Z_{veto}}$	58.76 ± 0.36
$\epsilon_{\not{E}_T}$	50.85 ± 0.48
ϵ_{2-jet}	10.09 ± 0.41
ϵ_{H_T}	70.43 ± 1.93
ϵ_{OS}	75.06 ± 2.18

Table 5.7: The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive WZ events, which give rise to $\epsilon_{WZ}^{MC} = (0.021 \pm 0.001)\%$. This is further corrected for lepton trigger efficiences and ID scale factors. Errors are statistical only.

$Z \to \tau\tau$	
Selection cut	ϵ, %
$\epsilon_{geom \cdot kin}$	24.30 ± 0.07
ϵ_{ID}^{lepton}	67.80 ± 0.15
ϵ_{iso}	99.01 ± 0.04
ϵ_{conv}	98.48 ± 0.05
ϵ_{Zveto}	98.94 ± 0.04
$\epsilon_{\not{E}_T}$	1.83 ± 0.05
ϵ_{2-jet}	30.14 ± 1.31
ϵ_{H_T}	62.53 ± 2.51
ϵ_{OS}	99.57 ± 0.43

Table 5.8: The consecutive efficiencies of each event selection requirement determined from the PYTHIA Monte Carlo sample of inclusive $Z \to \tau\tau$ events. The MC acceptance $\epsilon_{Z\to\tau\tau}^{MC} = (0.054 \pm 0.004)\%$. The correction factor of 1.68 ± 0.15 is applied to correct for ϵ_{2-jet} underestimate. Errors are statistical only.

GeV in the region $|\eta| < 2.5$. Using $\sigma^{NNLO}_{Z\to\tau\tau} = 335.3$ pb, which also includes a contribution from $\gamma^\star$, and taking filtering into account, the effective luminosity of the sample is $\sim 143 fb^{-1}$.

The list of consecutive efficiencies is presented in Table 5.8. Note, that the major reduction is due to the $\not\!\!E_{\rm T}$ requirements. The MC acceptance $\epsilon^{MC}_{Z\to\tau\tau} = (0.054 \pm 0.004)\%$. The correction factor of 1.68 ± 0.15 is applied to correct for the ϵ_{2-jet} underestimate in PYTHIA. The lepton trigger efficiencies and the ID scale factors reduce the acceptance by 15%. The systematic uncertainty in acceptance due to jet corrections is $\sim 29\%$.

The number of $Z \to \tau^+\tau^-$ events is given by

$$N^{Z\to\tau\tau}_{exp} = \sigma^{NNLO}_{Z\to\tau\tau} \times BR(both\tau \to \ell\nu_\ell\nu_\tau) \times \sum_{dileptons} \left(\epsilon^{Z\to\tau\tau}_{dil} \times \int \mathcal{L}dt\right), \qquad (5.9)$$

and found to be 0.42 ± 0.13.

5.3.4 Drell-Yan Background Estimate

The event selection requirements are very effective in reducing the Drell-Yan events $q\bar{q} \to \gamma^\star/Z \to \ell^+\ell^-, \ell = e$ or μ. However, this background is significant to the dilepton analysis due to its high production cross-section with $\sigma_Z \cdot BR(Z \to \ell\ell)$ alone over two orders of magnitude larger than the cross section for the correspondingly-flavored $t\bar{t}$ decay. The main handle in separating this background from $t\bar{t}$ is employing the fact that dielectron and dimuon events originating from Drell-Yan processes do not have an intrinsic $\not\!\!E_{\rm T}$. Therefore the estimate of the Drell-Yan contribution from Monte Carlo is dependent on the accurate modelling in the large-$\not\!\!E_{\rm T}$ tail in Drell-Yan + multi-jet events.

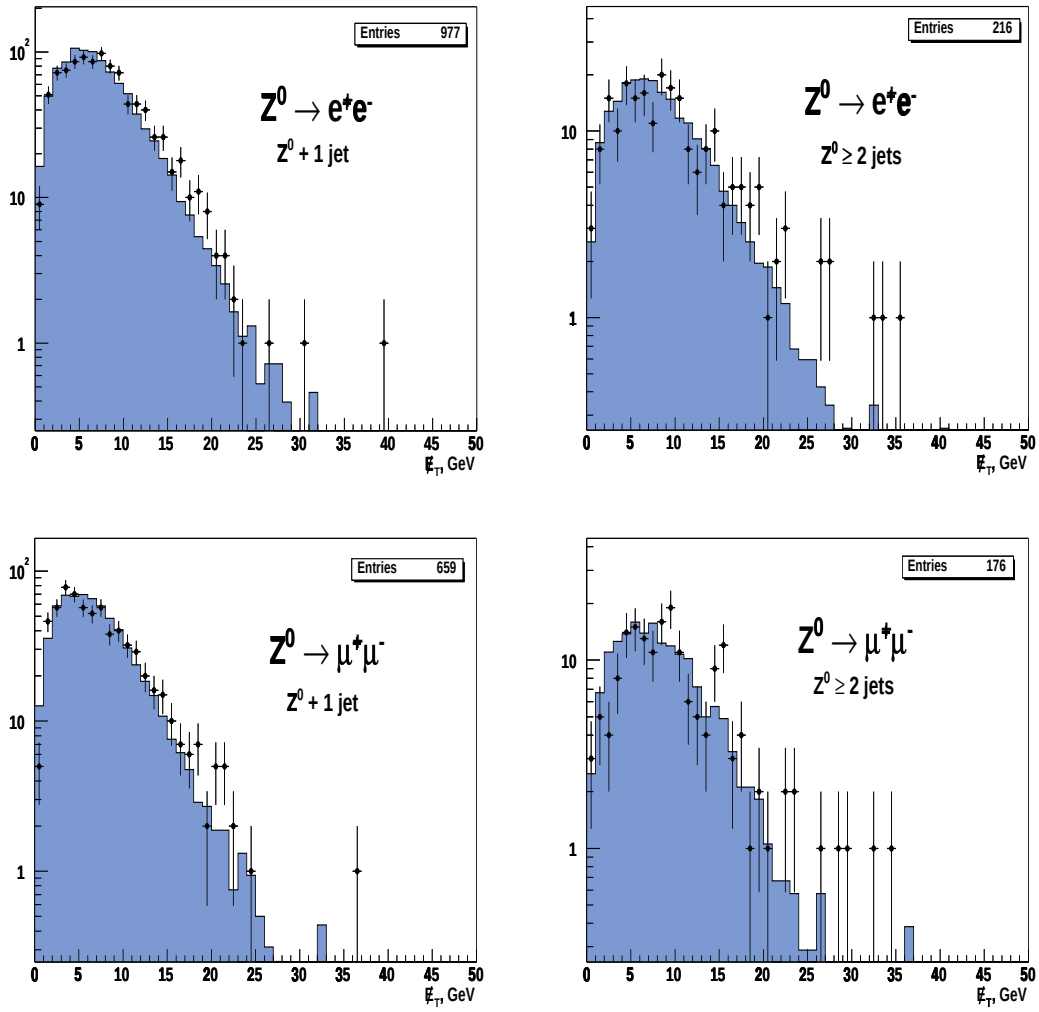

Figure 5.15: The missing transverse energy distribution of PYTHIA Monte Carlo in the logarithm scale of $Z^0 \to e^+e^-$ events (top) and $Z^0 \to \mu^+\mu^-$ events (bottom) compared with data. Monte Carlo distributions are normalized to data. In the distributions on the left dilepton events are accompanied with one jet and in the right distributions with at least two jets. The simulation underestimates the number of events in the high-$\not{E}_T$ region.

Figure 5.15 compares 'fake' $\not\!\!E_{\mathrm{T}}$ distributions for data and PYTHIA Monte Carlo, normalized to data, of $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events accompanied with one jet and at least two jets in an event. We observe that Monte Carlo underestimates the number of events in the high-$\not\!\!E_{\mathrm{T}}$ region. After subtracting contributions from WW, $Z \to \tau\tau$ and $t\bar{t}$ processes, we determine that the data excess over the Monte Carlo predictions is still significant such that scale factors ~ 2 need to be applied to the simulation in order to account for discrepancy between data and simulation. Unfortunatelly, heading this way we run into statistics limitations in determining the scale factors and hence the estimate is associated with a very large systematic uncertainty.

Instead, we determine the Drell-Yan contribution from the data. We enhance the statistics limitations by integrating over all jet multiplicity bins and then adopt the jet multiplicity distribution from Monte Carlo sculpted after both $\not\!\!E_{\mathrm{T}}$ and Z veto requirements inside of the Z mass window and after applying the $\not\!\!E_{\mathrm{T}}$ cut outside. This provides information on what fraction of events f^j falls in each jet multiplicity bin j: 0, 1 and ≥ 2.

The Drell-Yan events in the data are isolated by requiring the dilepton invariant mass to be inside of the Z mass window, and the contribution outside of the Z mass window is determined by exploiting the scale factor $R^j_{out/in}$ found in Monte Carlo.

The total number of expected Drell-Yan events N^j_{DY} in each jet multiplicity bin j is then determined as follows:

$$N^j_{DY} = f^j_{\not\!\!E_{\mathrm{T}}} R^j_{out/in} N_{\not\!\!E_{\mathrm{T}}} + f^j_{Zveto} N_{Zveto} \tag{5.10}$$

where the superscript j refers to the jet bin, the subscript $\not\!\!E_{\mathrm{T}}$ means after $\not\!\!E_{\mathrm{T}}$ cut, the subscript 'Zveto' means after $\not\!\!E_{\mathrm{T}}$ and Z veto cuts, $R_{out/in}$ is the ratio

of events outside/inside the Z mass window, and N's refer to the number of Z events observed in the data.

We observe $N^{ee}_{\not{E}_T} = 8$ dielectron events in the Z mass window, passing the $\not{E}_T$ requirement, which reduces to $N^{ee}_{Zveto} = 7$ after applying the Z veto cut. These numbers for muons are $N^{\mu\mu}_{\not{E}_T} = 4$ and $N^{\mu\mu}_{Zveto} = 1$ respectively. These events have a non-negligible contribution from $t\bar{t}$, WW, WZ and $Z \to \tau\tau$ events that should be subtracted from the N's. The contributions in the Z mass window are estimated with $\sim 30\%$ uncertainty and presented in Table 5.9.

The estimated from Monte Carlo coefficients $f^j_{\not{E}_T}$ and f^j_{Zveto} are given in Table 5.10 and $R^j_{out_in}$ scale factors are shown in Table 5.11.

Drell-Yan events in the signal ($\geq$ 2 jets bin) are further subject to the H_T cut, the efficiency for which is taken from ALPGEN Monte Carlo $Z + 2p$ sample. Since the H_T cut strongly depends on dilepton invariant mass, the efficiencies are determined separately for outside/inside mass regions and are shown in Table 5.12. The H_T cut is particularly susceptible to the jet energy scale; a 20% effect is seen on these efficiencies when fluctuating jets up and down by 1σ, and this is taken as a systematic.

The opposite sign cut is in principle 100% efficient for Drell-Yan events. However, PHX electrons have a non-zero charge fake rate, as can be seen from Figure 5.16. The charge fake rate is quantified by estimating the fraction of central-plug (CP) electron Z event in the data with same-sign and determined to be $f = 0.130 \pm 0.07$. This factor is applied to the CP portion of the DY background.

After applying the H_T and opposite sign cut efficiencies to the $N^{\geq 2}_{DY}$ given by Equation 5.10 we estimate 0.36 ± 0.27 dielectron and 0.07 ± 0.34 dimuon events from the Drell-Yan processes.

	ee		$\mu\mu$	
Process	$\not\!\!E_T$ cut	$\not\!\!E_T$, Z veto	$\not\!\!E_T$ cut	$\not\!\!E_T$, Z veto
$Z \to \tau\tau$	0.04	0.04	0.02	0.02
WW	0.83	0.82	0.62	0.61
WZ	0.73	0.68	0.56	0.53
$t\bar{t}$	0.65	0.56	0.66	0.55

Table 5.9: Expected non-Drell Yan events inside of the Z mass window. The contributions are estimated with 30% uncertainty.

	$j = 0$	$j = 1$	$j \geq 2$
$f^{ee}_{\not\!\!E_T}$	0.39 ± 0.04	0.43 ± 0.04	0.18 ± 0.03
$f^{\mu\mu}_{\not\!\!E_T}$	0.18 ± 0.043	0.56 ± 0.05	0.26 ± 0.04
f^{ee}_{Zveto}	0.67 ± 0.05	0.26 ± 0.04	0.07 ± 0.02
$f^{\mu\mu}_{Zveto}$	0.36 ± 0.06	0.45 ± 0.07	0.19 ± 0.05

Table 5.10: The fraction of events $f^j_{\not\!\!E_T}$ and f^j_{Zveto} in each jet multiplicity bin estimated from Monte Carlo.

	$j = 0$	$j = 1$	$j \geq 2$
$R^{ee}_{out/in}$	0.41 ± 0.09	0.37 ± 0.08	0.21 ± 0.09
$R^{\mu\mu}_{out/in}$	1.33 ± 0.38	0.33 ± 0.08	0.45 ± 0.15

Table 5.11: The $R^j_{out/in}$ scale factors determined from Monte Carlo dilepton invariant mass distributions.

ϵ_{H_T}	inside	outside
ee	0.53 ± 0.13	0.89 ± 0.02
$\mu\mu$	0.82 ± 0.15	0.75 ± 0.02

Table 5.12: H_T cut efficiencies inside and outside of the Z mass window determined from ALPGEN Monte Carlo. Errors are statistical only.

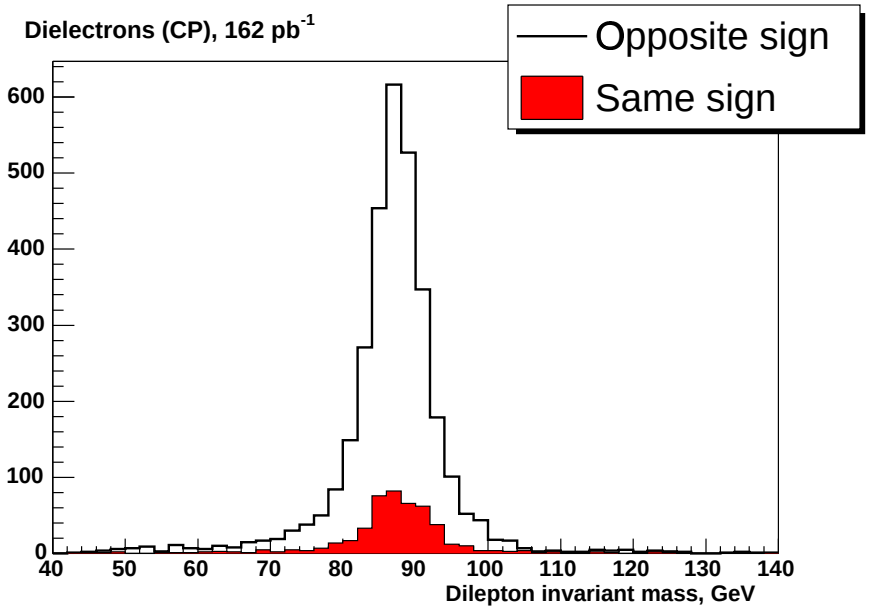

Figure 5.16: Dilepton invariant mass distribution M_{ee} for central-plug dielectron events in the data.

5.3.5 Lepton Fake Rate Background Estimate

The dilepton background contribution from events in which one lepton is faked by a hadronic particle is primarily dominated by $W+ \geq 3$ jet events, in which the W boson decays leptonically. The probability for a hadron to fake a lepton is very small $\sim 10^{-4}$. However, the large production rate of W bosons that are accompanied with at least three jets makes it one of the dominant backgrounds to the $t\bar{t}$ dilepton analysis. In these events the $\not{E}_{\mathrm{T}}$ is naturally supplied by the neutrino from the W decay, one of the detected leptons is real and the other one is faked by one of the three jets produced along with the W boson.

This background is very difficult to accurately simulate in Monte Carlo as multiple detector effects are involved. Therefore the lepton fake rate background contribution is estimated from the data. The procedure consists of two steps. First, the probability for a hadron to fake a lepton is estimated by using the QCD datasets. Secondly, the fake rates are applied to the observed $W+ \geq 3$ jets events.

QCD Datasets

To determine the lepton fake rates and to ensure ourselves that we can reliably predict the number of fake leptons in different environments, we make use of several QCD datasets, which are based on jet triggers with different jet E_T thresholds. They are referred as the *Jet20, Jet50, Jet 70* and *Jet100* samples, where the dataset *JetX* consists of events with the highest E_T jet greater than 'X' GeV.

The fake rate calculation is performed by counting the total number of *fakeable* jets [1] in all events comprising the dataset, and by counting the total number of leptons. These constitute the denominator and the numerator of the lepton

[1] Jets, which are likely to fake an electron or a muon, as defined further in the text.

fake rate, respectively. The fake rates are parametrized as a function of E_T and Isolation in cone 0.4; this parametrization leads to better consistency between different jet samples. We assume the Jet50 sample is the default one, as its jet E_T spectrum most resembles the one from $W+ \geq 3$ jets events, and apply the lepton fake rates found in the Jet50 sample to the other jet samples, thus predicting the number of leptons in those. The predictions are compared with the observed numbers and the discrepancy is taken as a systematic uncertainty.

The QCD datasets may also contain real leptons from W's and Z's, therefore for calculating electron fake rates we consider only those multijet events which have missing transverse energy less than 20 GeV, and also reject the events with two leptons falling into the Z mass window (76 GeV $< M_{ll} <$ 106 GeV). In the case of muons, the W contamination is reduced by a factor of 5 compared with electron case. This is because a W($\rightarrow$eν)+0 jet can still pass the jet trigger, while for muons only W($\rightarrow$eν)+ $\geq$1 jet will make it into the jet samples. The correcting $\not\!\!E_{\mathrm{T}}$ for a fake muon may easily result in large fake $\not\!\!E_{\mathrm{T}}$, which may resemble the W boson production. Hence the $\not\!\!E_{\mathrm{T}}$ cut is not applied for calculating muon fake rates.

Electrons and muons are treated separately also because the nature of misidentification for these leptons is very different.

Electron Fake Rates

A jet faking an electron must contain a highly energetic object, leaving most of its energy in the electromagnetic calorimeter, like a π^0, η, and another charged object with a track (π^+, π^-). It is essential that other jet fragments do not contribute significantly to the hadronic energy. Thus an 'electromagnetic' jet, i.e. a jet with low E_{had}/E_{em} ratio, has a much higher probability to fake an electron

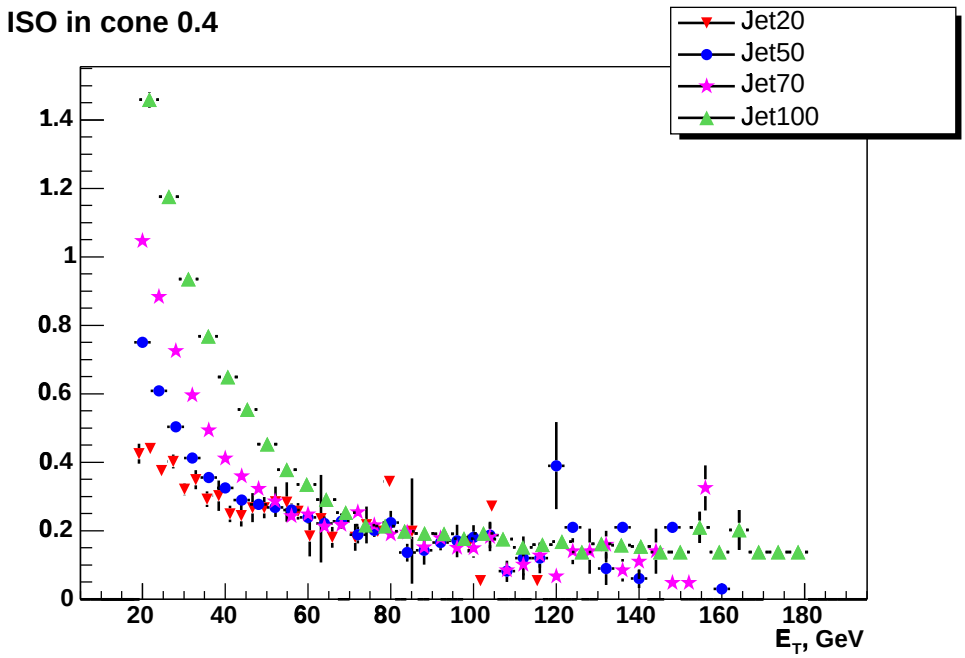

Figure 5.17: The calorimeter isolation in cone 0.4 as a function of E_T for CDF EM Objects from Jet20, Jet50, Jet70 and Jet100 samples. The curves for different jet samples agree near and above jet E_T threshold.

than a 'hadronic' jet. The trigger clustering identifies such jets as 'CDF EM Objects'. Every electron candidate on the trigger level is a CDF EM Object, therefore it is natural to define electron fakeable jets to be jets from the CDF EM Object collection. This definition permits better consistency between fake rates in different environments with distinct ratios D_{had}/D_{em} of the number of 'hadronic' (D_{had}) to the number of 'electromagnetic' jets (D_{em}). Moreover, it is more appropriate to use the CDF EM Object four-momenta instead of the one of the jet containing an electron candidate, because it is identical to what it is used in the dilepton event selection.

Thus, the electron fake rate f_e is defined as

$$f_e = \frac{N_e}{D_{em}}, \tag{5.11}$$

where N_e is the number of objects identified as electrons and D_{em} is the total number of CDF EM Objects [2].

Since a real electron can trigger the event as a jet, we remove the highest E_T jet from consideration. There is an additional trigger bias in the jet samples introduced by the jet E_T threshold requirements. Leading E_T jets in QCD samples are more likely to be formed from direct quark remnants of proton-antiproton pairs, while softer jets are more likely being produced due to gluon splitting and have higher gluon content. The gluon jets have on average higher particle multiplicity [40], which results on average in the higher hadron content and smaller isolation, hence a lower probability to fake an electron. In addition, softer jets tend to be less isolated due to activity of more energetic jets. This is illustrated in Figure 5.17, where the calorimeter isolation of a fakeable jet in cone 0.4 is plotted versus jet E_T. The curves for different jet samples coincide at energies near and above trigger E_T, while the calorimeter isolation for jets at lower energies tends to be different. To differentiate between jets at lower energies than jet E_T threshold and to account for E_T dependence, we parametrize fake rates (Equation 5.11) as a function of E_T and isolation in cone 0.4.

Due to differences between offline and trigger clustering [3] CDF EM Objects from higher trigger E_T jet samples tend to have larger hadron content (see Figure 5.18 (top)). Therefore we require the denominator to have offline $E_{had}/E_{em} <$

[2]Note, that so defined the electron fake rate is higher than a probability for a generic jet to fake an electron.

[3]Electron baseline ID cuts use two tower E_{em} and three tower E_{had}, while CDF EM Object trigger requires three tower E_{em} and two tower E_{had}, so that offline $E_{had}/E_{em} <$ trigger E_{had}/E_{em}.

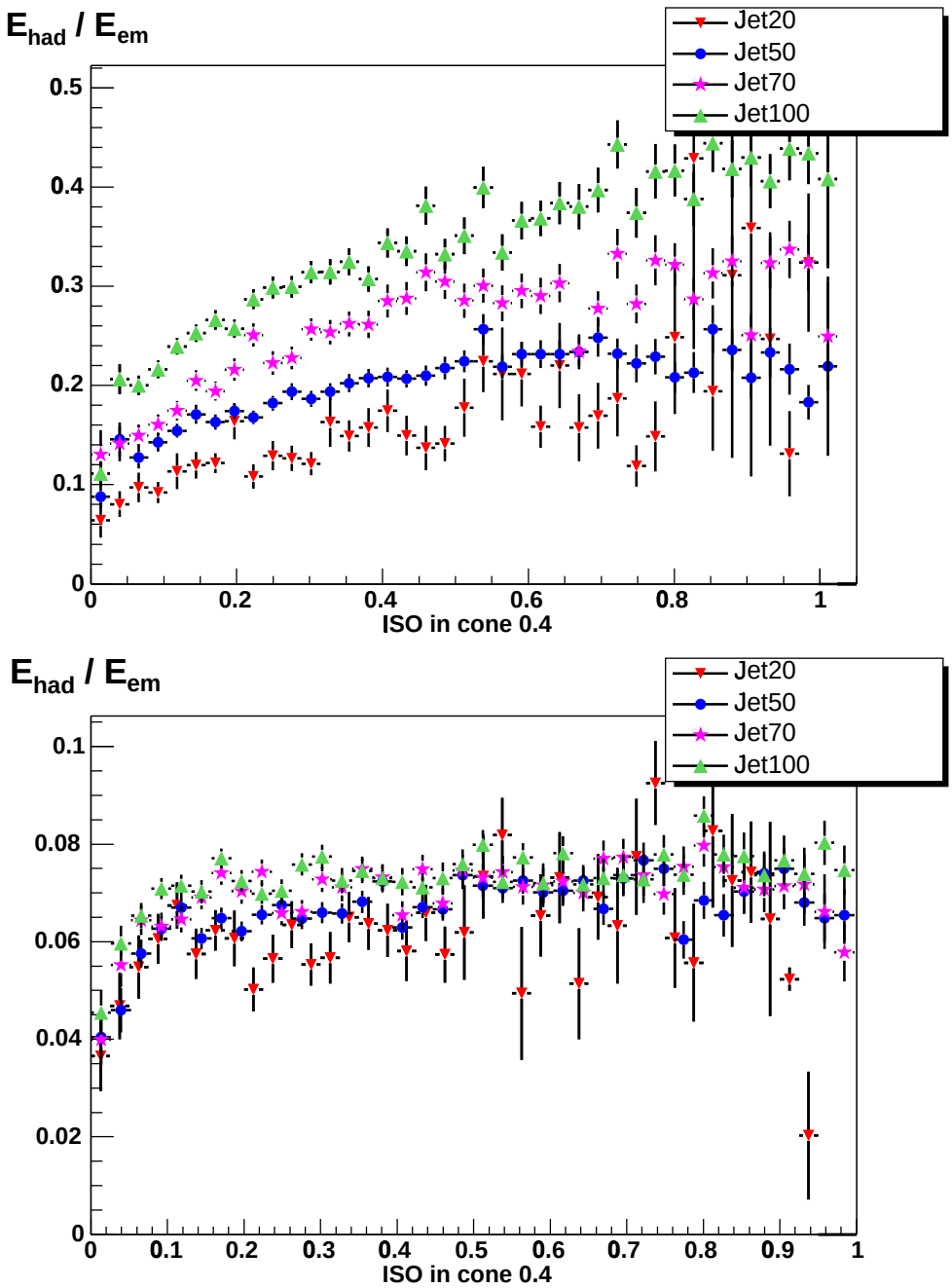

Figure 5.18: E_{had}/E_{em} as a function of the calorimeter isolation in cone 0.4 cone for EM jets from Jet20, Jet50, Jet70 and Jet100 with no E_{had}/E_{em} cut (top) and with $E_{had}/E_{em} < 0.125$ (bottom). Electron fakeable jets are required to satisfy $E_{had}/E_{em} < 0.125$

0.125, which leads to consistency between different Jet samples, as shown in Figure 5.18 (bottom). This allows parametrization in bins of isolation.

Fake rates for central and plug electrons are calculated separately. Central electron fakeable jets are also required to have a track and not to be flagged as a conversion. Plug electron fakeable jets are only required to be fiducial to the plug detector region: $1.2 < |\eta_{PES}| < 2.5$.

We test the electron fake rates by predicting the number of fakes in Jet 20, 70, 100 samples. Due to low statistics comparison per (E_T, ISO) bin is not always possible, we therefore make a cumulative comparison between the jet samples. We apply the fake rates as a function of E_T and isolation $f_e^{J50}(E_T, ISO)$ obtained from the Jet50 sample to the distributions of the fakeable jets of Jet20, 70 and 100. Precisely, the predicted number of electrons in a sample JetX is calculated as follows:

$$N_e^{pred} = \sum_{E_T, ISO} f_e^{J50}(E_T, ISO) \cdot D_{em}^{JX}(E_T, ISO), \tag{5.12}$$

where $f_e^{J50}(E_T, ISO) = N_e^{obs}/D_{em}^{J50}$ is the fake rate observed in the bin (E_T, ISO) and $D_{em}^{JX}(E_T, ISO)$ is the number of electron fakeable jets in the sample JetX in the same bin. We use the isolation bins of width 0.1, so that the fake rate parametrization reduces to E_T dependence for isolated electrons.

The result of the test are summarized in Table 5.13. We observe that the fake rate parametrization is adequate to predict fakes in the data samples with different physics composition.

Category	Data Sample	Observed	Predicted
ICEM	Jet20	15	13 ± 3
	Jet70	14	14 ± 4
	Jet100	25	15 ± 7
NICEM	Jet20	34	35 ± 4
	Jet70	63	81 ± 8
	Jet100	67	67 ± 10
PHX	Jet20	51	53 ± 8
	Jet70	75	60 ± 8
	Jet100	68	83 ± 21

Table 5.13: Number of observed and predicted fakes for electron categories. We use Jet50 to predict Jet20, Jet70 and Jet100.

Muon Fake Rates

There are many possibilities in which a jet might fake a muon: a punch-through (a hadron which reaches the muons chambers), a decay-in-flight (low P_T kaon or pion decays such that a high-P_T track is mis-reconstructed) or muons from semileptonic b or c quark decays. From the identification point of view a muon is a minimum ionizing track. Therefore muon fakeable jets are defined as tracks with a $P_T > 10$ GeV [4] and low activity in the calorimeter. Precisely, we set $E/P < 1$, where E is the calorimeter energy deposition in muon towers and P is the momentum of the track.

The muon fake rate is expected to be P_T dependent, as the higher the track momentum, the easier it is to penetrate the calorimeters and reach the muon cham-

[4] That constitutes CDF Muon collection.

bers. The fake rate is also expected to depend on the activity in the calorimeter in the region surrounding the track. Therefore we parametrize the fake rates as a function of track P_T and isolation in a cone 0.4.

The requirement $E/P < 1$ for muon fakeable jets is justified in Figure 5.19. We observe on average larger energy depositions corresponding to the same calorimeter isolation in higher E_T threshold samples (see top Figure) while after $E/P < 1$ we see better agreement between different jet samples (bottom Figure). This allows parametrization in bins of isolation.

Due to limited statistics we divide muons into three categories: TIMUO (tight isolated muons, which include ICMUP and ICMX), LMUO (loose isolated muons, which consist of ICMU, ICMP and ICMIO) and NIMUO (nonisolated muons), and determine fake rate distributions $f_\mu^{J50}(E_T, ISO)$ from Jet50 sample for each of the three muon categories. These fake rates are applied to other jet samples and the predicted numbers of fakes are determined as in the electron case by the relation:

$$N_\mu^{pred} = \sum_{E_T, ISO} f_\mu^{J50}(E_T, ISO) \cdot D_{MIP}^{JX}(E_T, ISO), \tag{5.13}$$

where subscript 'MIP' means 'minimum ionization particle', implying loose minimum ionization requirements defined for muon fakeable jets. The fake rate parametrization for isolated muons reduces to E_T dependence.

The expected numbers are compared with the observed. The results of the test are presented in Table 5.14. We see reasonable agreement between data samples with different kinematic structure.

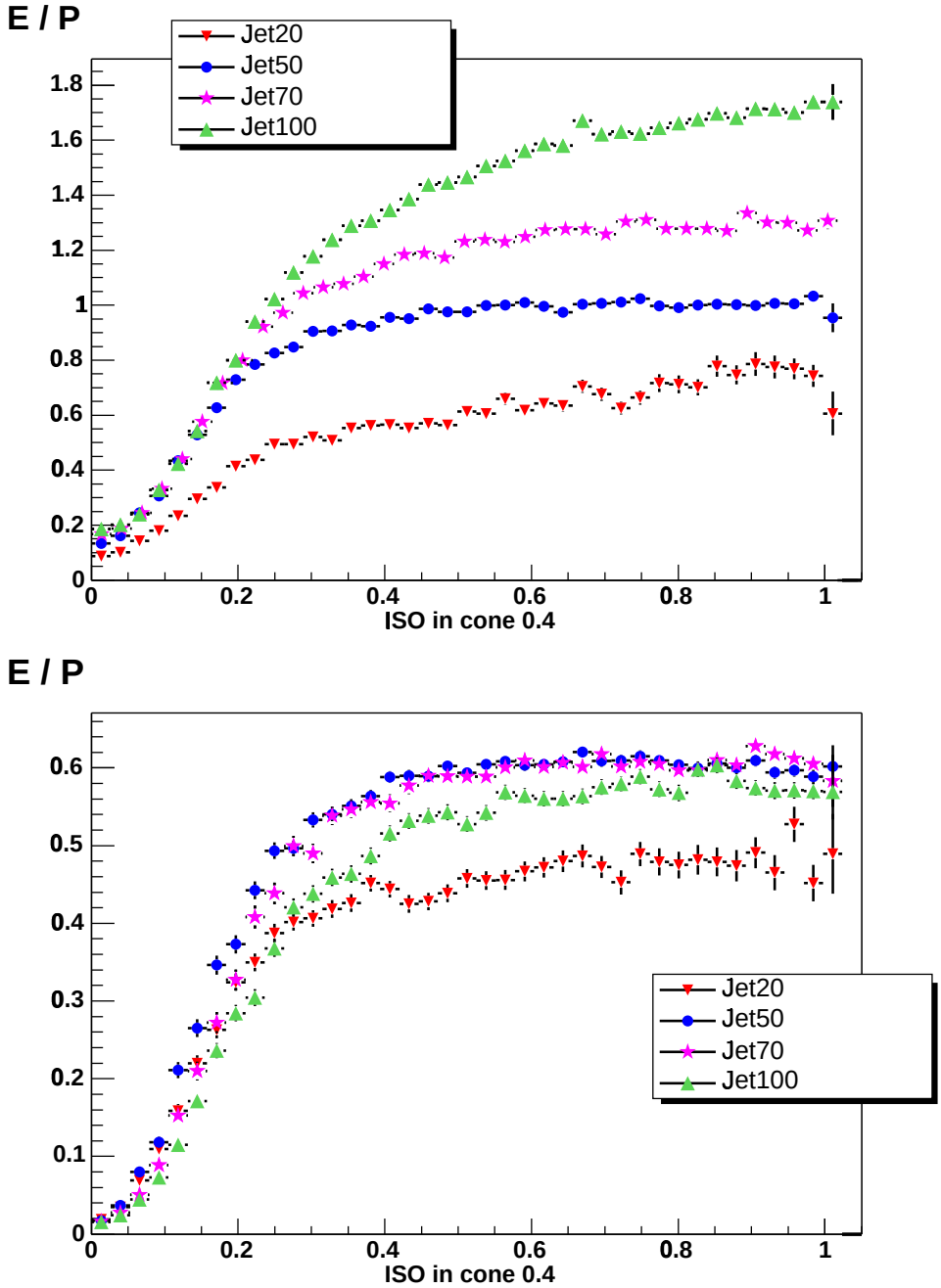

Figure 5.19: E/P as a function of calorimeter isolation in cone 0.4 for CDF Muon objects wth no E/P cut (top) and with $E/P < 1$ requirement (bottom). We define muon fakeable jets as tracks (Muon objects) with low calorimeter activity $E/P < 1$.

Category	Data Sample	Observed	Predicted
TIMUO	Jet20	22	33 ± 7
	Jet70	33	30 ± 6
	Jet100	69	76 ± 14
LIMUO	Jet20	18	42 ± 26
	Jet70	21	36 ± 12
	Jet100	50	88 ± 25
NIMUO	Jet20	72	75 ± 8
	Jet70	88	102 ± 9
	Jet100	100	148 ± 22

Table 5.14: Number of observed and predicted fakes for muon categories. We use Jet50 to predict Jet20, Jet70 and Jet100.

Applying Fake Rates to W+multijet sample

To determine the fake dilepton background we select events from the inclusive electrons, muons and plug datasets, having one and only one lepton candidate[5] and at least one fakeable jet. The fakeable jet is further treated as if it was a *real lepton*, and all respective corrections are propogated to event variables, such as $\not\!\!E_{\mathrm{T}}$, jet significance and H_T. The dilepton event selection criteria are applied to the corrected event variables. If an event happens to have two and more fakeable objects (jets or tracks), it is counted two or more times.

Due to limited numbers of events we are left with in some categories after requiring two or more jets or large H_T, we estimate the fake background in two steps:

[5]Events with two lepton candidates plus a fakeable jet would automatically fall into trilepton category and therefore are rejected.

Jet Sample	Total estimate before ≥ 2 jets, H_T and OS cut
Jet 20	7.21 ± 0.61
Jet 50	6.17 ± 1.98
Jet 70	6.49 ± 2.01
Jet 100	7.30 ± 1.38

Table 5.15: Fake rate estimates using different jet samples.

First, we require that the W+multijet events pass 'Zveto' and $\not\!\!E_{\mathrm{T}}$ cuts and estimate the fake contribution denoted by N^{MET}. The respective estimates for different jet samples are given in Table 5.15. For the central value of the total estimate we use the mean of all estimates, N_{av}^{MET}. Fake estimates are calculated per each dilepton category. The maximum difference between the average total estimate and the estimate of one of the jet samples is taken as a systematic error of measurement δN_{av}^{MET}.

Next, to get an estimate in each jet bin and after the H_T and OS cuts we multiply the obtained estimates in each category by the probability that a "fakeable dilepton event" will fall into 0,1, ≥ 2 jet bin or pass H_T, OS cuts: ϵ_X. Cut efficiencies ϵ_X and fake estimates combined from all dilepton categories are given in Table 5.16.

We test our fake rate estimate by looking at same-sign dilepton events. Before applying the dilepton selection cuts we observe $\epsilon_{SS} = (40 \pm 2)\%$ events have same-sign lepton pairs. The charge correlation between W boson and the quark produced along with W, which is usually the one responsible for faking a high-P_T lepton, is expected to bias the charge sign of lepton pairs towards opposite sign.

By applying the same-sign efficiency ϵ_{SS} to the total fake estimate we determine the number of same-sign lepton pairs expected from fake lepton background.

Cut	ϵ_X	Total estimate
0jet	0.45 ± 0.01	$5.53 \pm 1.05^{stat} \pm 0.44^{syst}$
1jet	0.35 ± 0.01	$4.35 \pm 0.83^{stat} \pm 0.35^{syst}$
$\geq$ 2 jets	0.20 ± 0.01	$2.47 \pm 0.48^{stat} \pm 0.20^{syst}$
After H_T	0.16 ± 0.01	$1.94 \pm 0.38^{stat} \pm 0.16^{syst}$
After OS	0.09 ± 0.01	$1.07 \pm 0.34^{stat} \pm 0.09^{syst}$

Table 5.16: Lepton fake rate background estimate.

Jet Bin	SS predicted	SS observed
0 jet	$2.30 \pm 0.44^{stat} \pm 0.18^{syst}$	3
1 jet	$1.79 \pm 0.35^{stat} \pm 0.14^{syst}$	2
$\geq$ 2 jets	$0.93 \pm 0.19^{stat} \pm 0.07^{syst}$	0

Table 5.17: Lepton fake rate prediction for same-sign events is in agreement with observations.

These numbers per each jet bin stand in agreement with the numbers of observed dilepton same-sign events (see Table 5.17).

5.3.6 $b\bar{b}$ Background Estimate

$b\bar{b}$ events may enter the dilepton sample as double fakes. Due to semileptonic decays the lepton fake rate from heavy flavor quarks is anticipated to be higher than from generic jets. b quark decays are identified with a SECVTX algorithm, which uses silicon tracking information to search for displaced secondary vertices. The secondary vertex is defined as the point of decay of a long-lived B hadron [6], which originated at the primary vertex.

[6] $c\tau$ for B hadrons is about 390 μm

The fake rate determined for generic jets was applied to the b-enriched sample, i.e. the sample with SECVTX tags. The heavy-flavor fraction ($c + b$ - charm + bottom) in the sample was estimated to be $\sim 12\%$ [41] We observed $N_{obs} = 188$ events while predicted $N_{pred} = 148 \pm 10$. Therefore, the lepton fake rate from heavy flavor quarks is expected to be $\beta = 3.2 \pm 0.3$ times higher than the one from generic jets.

The $b\bar{b}$ event fraction in pre-tagged $W \geq 3$ jets was determined to be $(5.6 \pm 0.8)\%$ [42]. Thus 5.6% of the obtained lepton fake rate estimate needs to be scaled up by 3.2 ± 0.3 to properly include heavy flavor quarks. That yields a fake estimate which is 12.3% higher. Since the uncertainty of the fake estimate is already $\sim 33\%$, we conclude that uncertainty due to $b\bar{b}$ pollution is well within quoted systematic error.

5.3.7 Summary

Tables 5.18 and 5.19 summarize the Standard Model expectations from various sources. The first table shows the breakdown of backgrounds and $t\bar{t}$ in different jet bins and the second table presents the breakdown of various physics processes after all event selection cuts into the flavor composition.

A search for candidates $t\bar{t} \rightarrow \ell\bar{\ell}$ events is performed by applying the event selection scheme to the three primary datasets described in Section 5.1. We find total of 13 candidates (1 ee, 9 $e\mu$ and 3 $\mu\mu$). The number of observed events stands in agreement with the Standard Model and has a little room for new physics contribution.

Source	0j	1j	$\geq 2j$	H_T, OS
WW/WZ	12.1 ± 3.5	3.5 ± 1.0	1.2 ± 0.3	0.74 ± 0.21
Drell-Yan	4.4 ± 2.0	2.2 ± 1.1	0.7 ± 0.4	0.43 ± 0.44
$Z \to \tau\tau$	0.19 ± 0.06	0.86 ± 0.26	0.69 ± 0.21	0.42 ± 0.13
Fakes	5.5 ± 2.3	4.4 ± 1.8	2.5 ± 1.0	1.07 ± 0.45
Total Background	22.2 ± 4.6	11.0 ± 2.4	5.0 ± 1.2	2.7 ± 0.7
$t\bar{t}$	0.1 ± 0.0	1.4 ± 0.2	8.7 ± 1.2	8.2 ± 1.1
Total SM	22.3 ± 4.6	12.4 ± 2.4	13.7 ± 1.8	10.9 ± 1.4
Run II data	19	11	14	13

Table 5.18: The breakdown of physics processes into different jet bins.

Source	ee	$\mu\mu$	$e\mu$	$\ell\ell$
WW/WZ	0.21 ± 0.06	0.18 ± 0.05	0.35 ± 0.10	0.74 ± 0.21
Drell-Yan	0.36 ± 0.28	0.07 ± 0.34	-	0.43 ± 0.44
$Z \to \tau\tau$	0.09 ± 0.03	0.11 ± 0.03	0.22 ± 0.07	0.42 ± 0.13
Fakes	0.26 ± 0.11	0.16 ± 0.07	0.69 ± 0.28	1.07 ± 0.45
Total Background	0.9 ± 0.3	0.5 ± 0.4	1.3 ± 0.3	2.7 ± 0.7
$t\bar{t}$	1.9 ± 0.3	1.8 ± 0.3	4.5 ± 0.6	8.2 ± 1.1
Total SM	2.8 ± 0.4	2.4 ± 0.5	5.7 ± 0.7	10.9 ± 1.4
Run II data	1	3	9	13

Table 5.19: The breakdown of physics processes into different flavor composition.

Chapter 6

Search Strategy

Motivated by the indication of a possible new physics signal among the Run I top dilepton events, we foresee that it could manifest itself more prominently in the larger statistics data sample of the Tevatron Run II. We expect that a new physics signal should result in distortions of kinematic distributions predicted according to the Standard Model.

To quantify the significance of a deviation in kinematic distributions several statistical tests can be used. In this Chapter we investigate them for sensitivity to observing various shape anomalies in kinematic distributions.

In anticipation of events with unusual kinematic properties in the new Run II data we introduce a novel statistical technique. It represents an *a priori* defined algorithm based on comparison of selected kinematic features of events. The technique is designed to isolate a subset of events revealing the largest discrepancy from the Standard Model distributions and quantify significance of their departure from the Standard Model.

We demonstrate that this method has larger sensitivity to a possible new physics signal in the data sample than the standard significance tests, and ad-

ditionally possesses a capability to isolate a subset of events with the largest concentration of new physics.

Due to a wide variety of theories going beyond the Standard Model we make only a minimal set of assumptions about the nature of non-Standard Model events, such as what observable quantities may potentially exhibit anomalous behaviour and what kind of distortions in distributions may be anticipated. Guided by possible hints from the Run I data sample we focus our search on events with large lepton P_T and large $\not{E}_{\mathrm{T}}$ resulting from the decay of an unknown massive particle.

6.1 Assessing Deviations in Kinematic Distributions

6.1.1 Null Hypothesis Testing

As experimentalists are often facing limited statistics, it is legitimate question to ask whether observed deviations from theoretical expectations are significant or can be accounted for by statistical fluctuations. Assumptions, from which the theoretical distribution is deduced, are called the *null hypothesis*. The null hypothesis is usually denoted as $H0$. The theoretical distribution itself is called the *null-hypothesis distribution*, or the $H0$-distribution. In this analysis for the null hypothesis we assume the Standard Model, and further in the text we will often interchange both of these names. Similarly, under the null-hypothesis distribution we will imply the distribution of a kinematic variable (or variables) as expected from the Standard Model.

The purpose of a significance test is to quantify how well the observed data

stands in agreement with the expected distribution. If the data is consistent with the null-hypothesis distribution then no conclusions about the validity of $H0$ can be made. Otherwise the null hypothesis is rejected.

The measure of agreement between the observed data and a given hypothesis is called a *test statistic*. The test statistic is usually chosen such that the greater its value the less likely the data is consistent with the tested distribution.

To quantify the significance of a deviation, various statistical tests can be used. The most common kinematic fitting technique is the Pearson χ^2 - test. It is designed to be applied to a binned distribution.

Suppose n is the number of bins in a histogram and $p_1, p_2, ..., p_n$ are probabilities for an event to fall into a particular bin. Then given the total number of observed events is N, the number of events in bin i has a mean Np_i and is distributed with a variance $\sigma_i^2 = Np_i$. The χ^2 test statistic is defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{(d_i - Np_i)^2}{Np_i}, \tag{6.1}$$

where d_i is the number of events observed in the bin i.

Although very useful for estimating parameters from the data, the χ^2-test possesses several undesired features as a significance test:

1. The results of the χ^2-test are dependent on an arbitrary computation choice – the bin width.

2. The χ^2-value is in general dependent on the shape of the underlying distribution. For instance, a re-parametrization of the kinematic variable would change the shape of the distribution and would affect the value of χ^2.

3. The χ^2-test ignores correlations from bin-to-bin. In the example in Figure 6.1 two histograms have the same value of χ^2 assuming the hypothesis

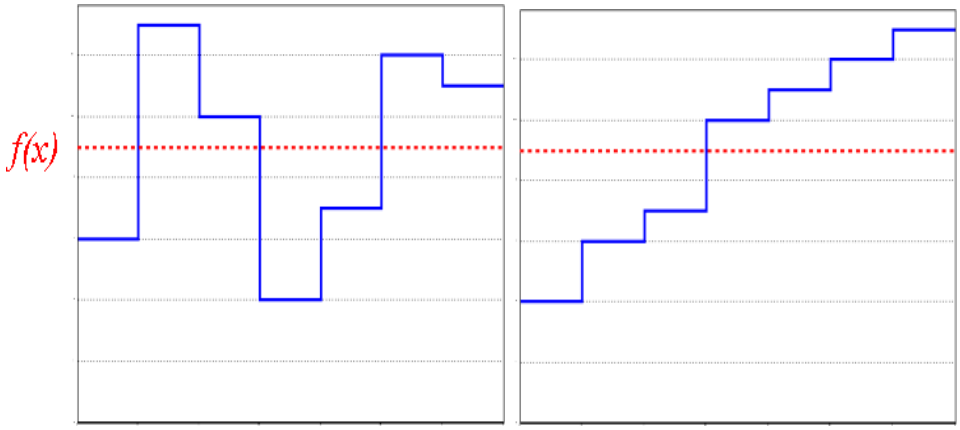

Figure 6.1: Two histograms having the same fit to the hypothesis of constant value for f when using the χ^2-test.

of constant $f(x) = c$, while the apparent asymmetry in the right histogram remains unnoticed. Because of this feature the χ^2-test is not sensitive, as we would desire it to be, to an asymmetry in the P_T spectrum caused by an excess of events in the high-P_T tail of the distribution.

4. However, the most essential problem for this analysis is that the number of observed events is relatively small to populate many bins with a sufficient number of entries. In general, the χ^2- test is not applicable for low-N statistics samples, because the variance of $\sum \frac{(d_i - Np_i)^2}{Np_i}$ grows to infinity as Np_i decreases and statistical fluctuations in bins with smaller p_i effectively hide disagreements in higher populated bins. This leads to a degradation of the test sensitivity. To resolve this issue one can reduce the number of bins in the histogram. However then the shape information is being lost.

Authors of Reference [6] alternatively overcame the last problem by developing the kinematic technique valid for distributions of any number of entries. This technique was employed for kinematic analysis of the Run I dilepton events. The

method was named the product likelihood and its test statistic was defined as the product of N probabilities:

$$PL_N = \prod_{i=1}^{N} p_i, \tag{6.2}$$

where the index i reflects the i-th event and p_i is the probability for an event i to occur. The value of p_i is determined by the expectation value in the bin and the location of the i-th event.

Unfortunately, the PL_N method has all the other flaws of the χ^2-test, such as it does not always extract all of the relevant information from the data. Indeed, a permutation of bins in a histogram does not change the product likelihood. More importantly, it does not give reliable estimates as a goodness-of-fit due to its dependence on the underlying distribution [43]. For instance, for the hypothesis of constant $f(x) = c$ every outcome is equally likely whether the data points are uniformly distributed or clumped about one point in a way inconsistent with f [1].

[1] One may argue that the product of probabilities of single events represents the probability for the entire data sample only in the limit of a large number of bins. In order to correctly weight pseudoexperiments with several events falling into the same bin, one has to assign multinomial coefficients in front of the products of probabilities. The probability for a data set is

$$PL_N = \frac{N!}{\nu_1!\nu_2!..\nu_M!} P_1 P_2 ... P_N,$$

where $M < N$ is the number of bins populated by a data sample and $\nu_k, k = 1, .., M$ is the number of events falling into the k-th bin. Although it might appear that this procedure solves the problem with dependence of the outcome on the underlying distribution, even in this form the method still ignores an order of bins in a histogram and in fact becomes more strongly dependent on the choice of binning of the null-hypothesis distribution.

6.1.2 Significance Tests Based on Cumulative Distributions

Taking into account these considerations it is generally believed that methods based on the integral probability distribution function (the Kolmogorov-Smirnov test and its varieties), which take into account the relation between data points, are more efficient for the kinds of phenomena generally of interest to high-energy physicists. These methods do not suffer from the problems above. Unlike the χ^2 test they are primarily intended for use with continuous distributions and are robust and independent of the shape of the experimental distribution and arbitrary computation choices such as a bin width.

The Kolmogorov-Smirnov statistic is defined as the maximum value of the absolute difference between two cumulative distributions. Let $f(x)$ be the null-hypothesis distribution of some kinematic quantity x, and let $F(x)$ denote its cumulative function, i.e. $F'(x) = f(x)$. Suppose we observe N events with values of x: $\{x_i\}$, $i = 1, .., N$ sorted in ascending order. The Kolmogorov-Smirnov (KS) statistic is defined as

$$\Delta_{KS} = \max_x |F(x) - S_N(x)|, \tag{6.3}$$

where $S_N(x)$ is empirical distribution function (e.d.f.) giving the fraction of data points to the left of a given value x:

$$S_N(x) = \begin{cases} 0 & \text{for } x < x_1 \\ i/N & \text{for } x_i \leq x < x_{i+1} \quad i = 1, .., N-1 \\ 1 & \text{for } x \geq x_N \end{cases} \tag{6.4}$$

The larger the value of Δ_{KS} the less likely the data matches the theory. The probability of observing $\Delta_{KS} > \Delta_0$, in the limit of large N, is given by Kolmogorov:

$$P(> \Delta_0) = 2\sum_{k=1}^{\infty}(-1)^{k-1}exp(-2k^2 N\Delta_0^2) \tag{6.5}$$

This relation is good when either of the distributions under examination has ≤ 80 entries [44]. For small sample sizes up to $N \geq 4$ more accurate approximation [45] can be obtained by replacing

$$\sqrt{N} \rightarrow \sqrt{N} + 0.12 + 0.11/\sqrt{N}$$

in Equation 6.5, or alternatively the KS tables can be consulted [46].

To properly simulate the shape of the probability distribution function $W(\Delta_{KS})$ of the KS distances we employ the Monte Carlo technique. We randomly generate N events according to the theoretical distribution $f(x)$. These events constitute a pseudoexperiment. Then this procedure is repeated a large number of times and for each pseudoexperiment the Δ_{KS} (Equation 6.3) between the H0-distribution and the e.d.f (Equation 6.4) of the pseudoexperiment is computed, which makes an entry into the $W(\Delta_{KS})$ probability distribution. The distribution $W(\Delta_{KS})$ is then used for quantifying the significance level of consistency of the data with the null-hypothesis as follows

$$P(> \Delta_0) = \int_{\Delta_0}^{\infty} W(\Delta)d\Delta. \tag{6.6}$$

Since two cumulative distributions agree at the smallest allowable value of x and the largest allowable value of x, it is the behaviour between the largest and smallest values that distinguishes distributions. The Kolmogorov-Smirnov

test quantifies a maximum excess of one distribution over the other accumulated starting from one of the tails.

One can think of any number of statistics to measure the overall difference between two cumulative distribution functions – variants of the KS test. The Smirnov-Cramer-von Mises (SCM) statistics uses the integrated square difference between $F(x)$ and $S_N(x)$:

$$\Delta_{SCM} = \int_0^1 (F(x) - S_N(x))^2 dF(x) \tag{6.7}$$

or in general

$$\Delta = \int_0^1 (F(x) - S_N(x))^2 \psi(x) dF(x), \tag{6.8}$$

where $\psi(x)$ is called a *weight function.*

We note that the sensitivity of the KS test to deviations from a null-hypothesis cumulative distribution function $F(x)$ is dependent on x. In fact, the KS test tends to be most sensitive around the median value $F(x) = 0.5$ and less sensitive at the extreme ends of the distribution. This is because the difference $|S_N(x) - F(x)|$ has a probability distribution with a variance proportional to $F(x)[1 - F(x)]$.

Consider, for instance, hypothesis of constant $f(x) = c$. Suppose we observe 100 events and superimpose expected and observed distributions in histograms of 100 bins. That is we expect on average one event per bin. Assume a situation when the maximum deviation is achieved in the median of $H0$, such that we observe 40 events in the lowest 50 bins and 60 events in highest 50 respectively. Compare it now with a scenario when the maximum deviation is achieved in the tail. One observes 90 events in the first 99 bins and the remained 10 events in the last bin. Both scenarios result in the same KS probability. For $\Delta = 0.1$ and

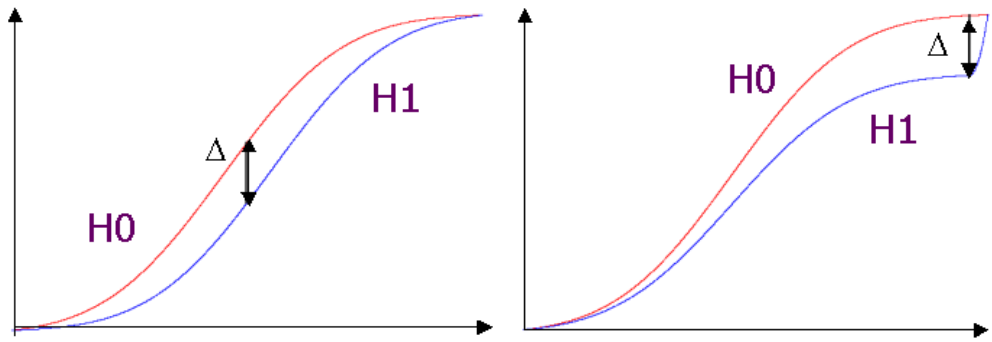

Figure 6.2: Two scenarios of an alternative hypothesis $H1$ having the same fit using the KS test.

$N = 100$ we have $P_{KS} \simeq 0.27$, i.e. quite likely, while Poisson probability with expectation $\mu = 1$ to fluctuate to 10 at least in one of 100 bins is $\simeq 1.11 \times 10^{-5}$. Thus we conclude that the KS test is not optimal for accessing deviations in the tails of kinematic distributions.

One way of increasing the sensitivity of the KS test out in the tails is to use a so-called *stabilized* or *weighted* statistic. The test statistic given by relation 6.8 with the weight function

$$\psi(x) = \frac{1}{\sqrt{F(x)[1 - F(x)]}}$$

is called the Anderson-Darling (AD) statistic [47]

$$\Delta_{AD} = \int_0^1 \frac{(F(x) - S_N(x))^2}{F(x)(1 - F(x))} dF(x) \tag{6.9}$$

It tends to be more sensitive to detection of deviations in the tails of the distribution.

Another approach is the Kuiper (K) statistic [48], which is defined as the sum of the maximum positive and maximum negative deviation between $F(x)$ and $S_N(x)$:

$$\Delta_K = \max_x(F(x) - S_N(x)) - \min_x(F(x) - S_N(x)). \tag{6.10}$$

It is equivalent to the KS statistic with an arbitrarily chosen starting integration point:

$$\Delta_K = \max_{x_1,x_2} |(F(x_1) - F(x_2)) - (S_N(x_1) - S_N(x_2))| \tag{6.11}$$

The Kuiper statistic has even more general symmetry, which guarantees equal sensitivities at all values of x, by effectively wrapping the x axis around into a circle (identifying the largest and the smallest allowable values of x). It is invariant under all shifts and parametrization on the circle.

In order to investigate which statistical test is more sensitive to the anticipated class of shape anomalies in the kinematic distributions, e.g. deviations in the tails, we need to define the quantity, called the *power of test*.

6.1.3 Power of Test

A statement about the validity of an $H0$ often involves a comparison with some *alternative hypothesis*, $H1$,$H2$,... We define pseudoexperiments randomly drawn from a null-hypothesis distribution as $H0$-*pseudoexperiments*, and pseudoexperiments randomly drawn from an alternative hypothesis $H1$-distribution as $H1$-*pseudoexperiments*.

By generating a large number of $H0$-pseudoexperiments and computing statistical deviations Δ between their e.d.f. and the $H0$-distribution, one can build the probability distribution of the statistic, $W_{H0}(\Delta)$ (p.d.s.)

Suppose our assumptions about the null hypothesis are wrong and an alternative hypothesis $H1$ is the true description of nature. Then the data will

follow the $H1$-distribution. One can similarly generate a large number of $H1$-pseudoexperiments, then compute statistical deviations Δ between their e.d.f.'s and the $H0$-distribution[2] and thus build the p.d.s $W_{H1}(\Delta)$.

The compatibility between the data and various hypotheses is determined by defining a *critical region*, for instance, $\Delta > \Delta_0$. Then the region $\Delta < \Delta_0$ is called the *acceptance region*. If Δ_{data} is observed in the critical region, the hypothesis $H0$ is rejected. The value

$$\alpha = \int_{\Delta_0}^{\infty} W_{H0}(\Delta) d\Delta, \tag{6.12}$$

is called the *significance level* of the test. Or alternatively one says that the hypothesis $H0$ is rejected with the *confidence limit*

$$C.L. = \int_{0}^{\Delta_0} W_{H0}(\Delta) d\Delta, \tag{6.13}$$

if Δ_{data} happens to be greater than Δ_0.

The choice of Δ_0 and hence the choice in acceptance or rejection of $H0$ is purely conventional. However, even if the $H0$ is valid, one can still mistakenly reject the null-hypothesis with a probability α. Therefore it is also called an *error of the first kind*. Likewise if the $H1$ is the true theory, one can mistakenly accept the $H0$ with a probability

$$\beta = \int_{0}^{\Delta_0} W_{H1}(\Delta) d\Delta. \tag{6.14}$$

β is called an *error of the second kind*. The value

[2]Note that comparison in both cases is made with the $H0$-distribution, since it is the null hypothesis, which is tested for validity.

$$\pi = 1 - \beta = \int_{\Delta_0}^{\infty} W_{H1}(\Delta) d\Delta \tag{6.15}$$

is the probability to find inconsistency of the observed data with the $H0$, given that the data satisfies the $H1$-hypothesis. This quantity is called the *power* of the test to discriminate against the alternative hypothesis $H1$. We note that the power of the test is unambiguously defined by specifying both hypotheses $H0$ and $H1$ and the test statistic t in use. Hence $\pi = \pi(H0, H1, t)$.

6.1.4 Sensitivity of Tests to Shape Anomalies

To explore the sensitivity of the discussed statistics to various shape anomalies we conducted a test using the gaussian distribution with mean $\mu_0 = 0.0$ and $\sigma_0 = 1.0$ as the null-hypothesis distribution and for simplicity assuming that a shape anomaly (new physics signal in a kinematic distribution) is localized in a certain area of the $H0$ distribution. The shape anomaly was represented by a peak − a sharp gaussian with $\sigma_1 = 0.1$. We considered four scenarios of the peak location with $\mu_1 = 0.0, 1.0, 2.0$ and 3.0, as shown in Figure 6.3. In our study the alternative hypothesis $H1$ was defined as a combination of 75% $H0$ and 25% of a 'new physics' peak.

Five statistics were investigated − Kolmogorov-Smirnov (KS) test, Smirnov-Cramer-von Mises (SCM), Anderson-Darling (AD), Kuiper tests and the PL_N-technique. Probability distributions of statistics $W_{H0}(\Delta)$ and $W_{H1}(\Delta)$ were built based on 10,000 $H0$-pseudoexperiments and the same number of $H1$-pseudoexperiments, 30 events each. We looked into three critical regions with confidence limits (given by Equation 6.13): 68%, 95% and 99% respectively.

In the first scenario, a shape anomaly is located in the center of the $H0$-distribution. The power for each test statistic is calculated for various confidence

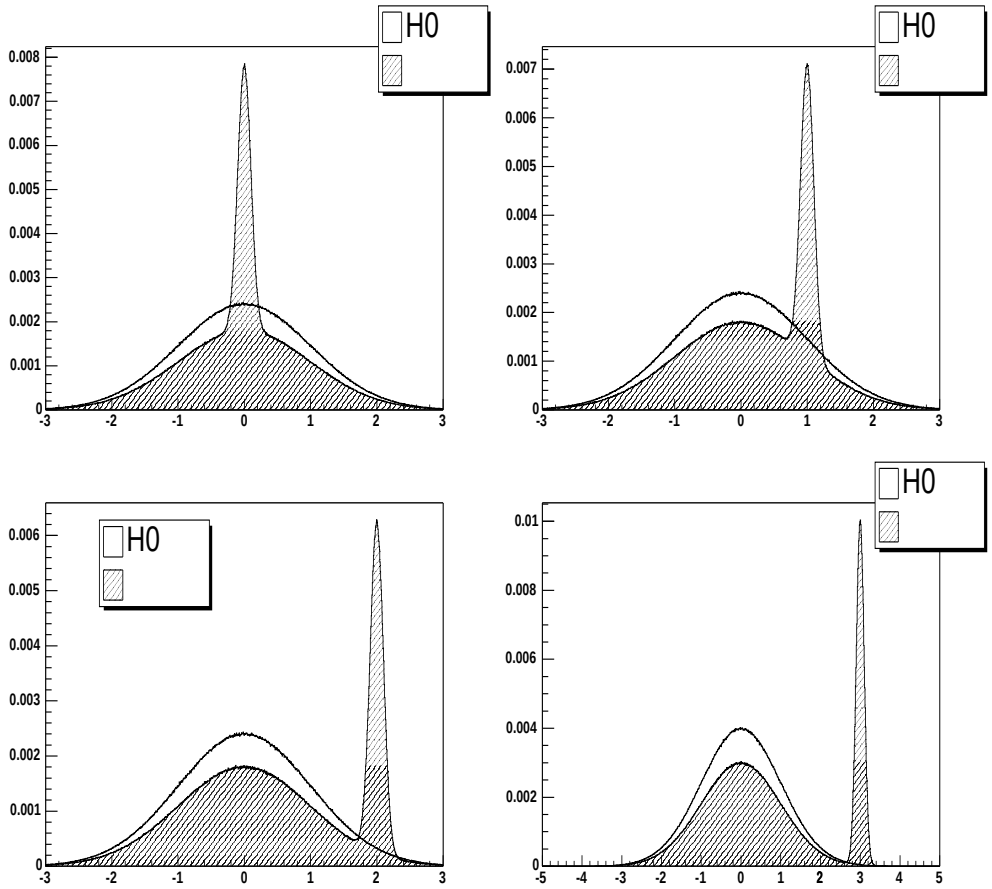

Figure 6.3: Four scenarios for alternative hypothesis $H1$.

Statistic	0.68	0.95	0.99
KS	0.698	0.286	0.102
SCM	0.639	0.107	0.008
AD	0.620	0.088	0.008
PL_N	0.077	0.004	0.000
Kuiper	0.803	0.464	0.225

Table 6.1: The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 0.0$.

limits using Equation 6.15 and results are presented in Table 6.1. One can notice that the Kuiper statistics is the most powerful in finding the peak at the center. As follows from the Table, about 22.5% of $H1$-pseudoexperiments were found to be as unlikely as the least likely 1% of $H0$-pseudoexperiments. While this number corresponds to 10.2% for the KS test. This lower sensitivity of the KS statistic to a symmetric $H1$-distribution can be explained because the KS test has its most sensitivity in the median of the distribution, while in this example both the $H0$- and $H1$- cumulative distribution functions agree at $F(x) = 0.5$.

The other noteworthy circumstance is that the PL_N test fails to find a discrepancy since $H1$-pseudoexperiments tend to have more events in the sharp peak of $H1$, which is the likely region for $H0$. An $H1$-pseudoexperiment has on average larger product likelihood than an $H0$-pseudoexperiment. This leads to fact that the $H1$-pseudoexperiments are on average more consistent with the $H0$-distribution than the $H0$-pseudoexperiments themselves. In this case one says that such a significance test does not have *merit*.

In the second scenario, μ_1 is set to 1.0 and the results of the tests are given in Table 6.2. Here the KS test is comparable with the Kuiper test, while the PL_N test remains blind to discrepancies. The results of $H1$ with the peak at $\mu_1 = 2.0$ are shown in Table 6.3. Note that the Anderson-Darling statistic starts dominating once the anomaly keeps moving toward the tail of the distribution. We note also that the Kuiper statistic retains uniform sensitivity for different scenarios and a slow rise of the Kuiper test power with μ_1 shifting to the tail of the $H0$-distribution can be explained by more significant excess of events in the area of the peak over the expected $H0$.

In the last scenario (Table 6.4) the most powerful tests are the Anderson-Darling and the PL_N. It is important to note though that the latter achieves the

Statistic	0.68	0.95	0.99
KS	0.845	0.540	0.385
SCM	0.829	0.507	0.260
AD	0.820	0.477	0.254
PL_N	0.324	0.037	0.001
Kuiper	0.858	0.522	0.274

Table 6.2: The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 1.0$.

Statistic	0.68	0.95	0.99
KS	0.955	0.753	0.608
SCM	0.932	0.717	0.463
AD	0.980	0.892	0.744
PL_N	0.988	0.842	0.587
Kuiper	0.911	0.612	0.354

Table 6.3: The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 2.0$.

highest discrimination between hypotheses by counting some of unlikely outcomes consistent with $H0$.

Based on results of this study we conclude that *an optimal search technique differs if the new physics is located in the tails of kinematic distribution or if the new physics causes a peak within the distribution.* We observe that the Kuiper statistic is more preferable when a deviation occurs in the core of the distribution and the Anderson-Darling is most sensitive when a deviation occurs in the tails.

Statistic	0.68	0.95	0.99
KS	0.976	0.822	0.660
SCM	0.944	0.730	0.511
AD	0.999	0.984	0.942
PL_N	1.000	0.996	0.983
Kuiper	0.937	0.684	0.406

Table 6.4: The power of statistical tests for various confidence limits for H1 with a peak at $\mu_1 = 3.0$.

6.2 Isolating Unlikely Events in a Data Sample

In general, the purpose of a significance test is to test consistency of a data sample with theoretical predictions and reject the null-hypothesis if the test statistic for the data is found to be in the critical region. We, however, do not aim to abolish the Standard Model, which served us so well, and successfully continues to describe observed processes at currently accessible energies. A more fundamental theory is an extension to the Standard Model and must account for SM interactions. Therefore only some of the observed events are possibly involved in interactions beyond the SM.

We expect that the subset of new physics events should reveal the largest discrepancy from the Standard Model kinematics. Indeed, suppose Q events of the total N observed events are non-SM events, exhibiting kinematics different from the Standard Model. Then disagreement of a smaller subset of anomalous events ($K < Q$) is statistically less significant, while a larger subset of events ($K > Q$) would inevitably contain SM events and hence be in better agreement with the Standard Model. Therefore a subset of size $K \simeq Q$ should reveal the

largest level of dissimilarity.

6.2.1 K-subsets

In order to isolate possible new physics events we focus on subsets of events which are maximally responsible for distortions in the kinematic distribution, and call them *unlikely subsets*. A test statistic itself can serve as a measure of their unlikeness.

We define *a most unlikely subset of K events* (*K-subset*) in a sample of N events the subset that has the largest statistical deviation from H0 among all possible subsets of K events in the sample.

This definition assumes that a search for such a subset must present a loop over all possible subsets, calculation of the statistical deviation for each subset, and thus finding the one that gives the largest deviation. However, this procedure is not always optimal, as the total number of possible subsets in a sample of N events is $N!$ – a very large number. Even for $N = 20$ it will take an enormous amount of time to process each pseudoexperiment with a modern computer.

The most straightforward way in determining K-subsets in the data sample is given by the PL_N method. In the PL_N the probability for each event is determined independently. Therefore all events can be sorted in the order of their likelihood to appear. The K-subset consists of the K events with the smallest probabilities p_i, which results in the smallest product of Πp_i.

In the case of Kolmogorov-Smirnov statistic one can easily see that the maximum deviation between the continuous null-hypothesis cumulative distribution function $F(x)$ and the empirical distribution function $S_N(x)$ is achieved at one of the data points x_i, where the step-function $S_N(x)$ increases by $1/N$. Therefore

$$\Delta_{KS} = \max_{x} |F(x) - S_N(x)| = \max_{i=1,..,N} \left(\left| F(x_i) - \frac{i}{N} \right|, \left| F(x_i) - \frac{i-1}{N} \right| \right) \quad (6.16)$$

One can notice that for a subset the largest deviation is achieved, when several sequential 'steps' are removed in the function $S_N(x)$. If we measure the deviation at x_i and q 'steps' are eliminated before x_i, then x_i becomes the $(i-q)$-th point and the KS-deviation at this point is

$$\Delta_{KS} = \max \left(\left| F(x_i) - \frac{i-q}{K} \right|, \left| F(x_i) - \frac{i-q-1}{K} \right| \right) \quad (6.17)$$

Looping over all x_i and all possible q for a given i and also taking into account that $0 \leq S_K(x) \leq 1$, the largest deviation may be presented in the following compact form:

$$\Delta_{KS,K} = \max_{\substack{1 \leq i \leq N \\ Max\{0, i-K\} \leq q \leq Min\{i, N-K+1\}}} \left| F(x_i) - \frac{i-q}{K} \right| \quad (6.18)$$

Consequently, the K-subset of the KS test consists of the K events located the farthest from the median of the $H0$ distribution at both tails of the distribution.

The same arguments applied to the Kuiper statistic using Equation 6.11 leads to

$$\Delta_{Kuiper} == \max_{\substack{i = 1,..,N \\ j = 1,..,i-1}} \left(\left| F(x_i) - \frac{i}{N} \right|, \left| F(x_i) - \frac{i-1}{N} \right|, \left| F(x_i) - F(x_j) - \frac{i-j}{N} \right|, \right.$$

$$\left. \left| F(x_i) - F(x_j) - \frac{i-j-1}{N} \right|, \left| F(x_i) - F(x_j) - \frac{i-j+1}{N} \right| \right) \quad (6.19)$$

and for the K-subset

$$\Delta_{Kuiper,K} = \max_{\substack{1 \le i \le N \\ 0 \le j \le N-1 \\ Max\{0, i-j-K\} \le q \le Min\{i-j, N-K+1\}}} \left| F(x_i) - F(x_j) - \frac{i-j-q}{K} \right| \tag{6.20}$$

As in the KS test, the largest Kuiper deviation for a subset is achieved when several sequential steps are removed. However, since the x-axis is effectively wrapped into a circle in the Kuiper test, removing a few data points with the lowest values of x and simultaneously data points with the highest values of x also leads to a larger Kuiper deviation. Therefore, the K-subset in the Kuiper test selects events resulting in the highest density over the $H0$ expectations, e.g. a possible unknown resonance peak.

To quantify the significance of the departure of the K-subsets from the null hypothesis, we ought to determine how frequently such unlikely subsets occur in $H0$-pseudoexperiments. To do this, we employ the Monte Carlo technique. We generate a large number of pseudoexperiments of N events each. In each pseudoexperiment we find K-subsets and determine Δ_K for all $K : 1 \le K \le N$. The value Δ_K becomes an entry into the probability distribution of statistic $W_K(\Delta)$. In total we build N probability distributions of statistic $W_K(\Delta_K), K = 1, .., N$. Using the same procedure one can determine deviations for K-subsets of the data sample Δ_K^{data}. The fraction of pseudoexperiments, whose K-subsets have deviations larger than Δ_K^{data}, gives a significance level P_K of departure of the K-subset of the data sample from the null hypothesis distribution:

$$P_K = \int_{\Delta_K^{data}}^{\infty} W_K(\Delta_K) d\Delta_K, \tag{6.21}$$

The values $p_K, K = 1, .., N$ determine how well each K-subset stands in agree-

ment with the Standard Model expectations.

Similarly to the discussion in Section 6.1.3, one can specify a critical region for the test with a certain significance level

$$\alpha = \int_{\Delta_{0,K}}^{\infty} W_{H0,K}(\Delta_K) d\Delta_K, \tag{6.22}$$

where $\Delta_{0,K}, K = 1, .., N$ are boundaries of the critical regions for probability distributions of statistic $W_{H0,K}(\Delta_K)$. For a given alternative hypothesis $H1$, one can determine the power of test by generating $H1$-pseudoexperiments and building probability distributions of statistic $W_{H1,K}(\Delta_K)$. In general, the power of test π_K differs for different K and is given by

$$\pi_K = \int_{\Delta_{0,K}}^{\infty} W_{H1,K}(\Delta_K) d\Delta_K. \tag{6.23}$$

If an alternative hypothesis $H1$ is the sum of the null hypothesis $H0$ (the Standard Model) and some shape anomaly (new physics signal) we expect that power of the test is maximized at $K \simeq Q$, where Q is an average number of events expected from the new physics contribution.

In the next section we test this assumption and investigate sensitivity of statistics to shape anomalies in distributions by isolating the K-subsets.

6.2.2 Sensitivity of Statistics to Shape Anomalies by Looking at Subsamples

Based on the results of Section 6.1.4 that an optimal search technique depends on the kind of shape anomaly, and the expectation that a new physics signal will exhibit itself in the tails of the P_T spectrum, we perform a test with an $H1$ containing a shape anomaly in the tail of an $H0$.

As an $H0$ distribution we choose a distribution with a long tail. We take a Landau distribution with mean $= 0.1$ and $\sigma = 0.02$ and consider a class of new physics distributions defined as the normal distributions with various means μ_{gaus} and fixed width $\sigma = 0.03$. We assume that the $H1$ distribution is a sum of 2/3 $H0$ and 1/3 of the new physics signal.

For each position of the gaussian which specifies $H1$-distribution, we generate 100,000 of $H0$- and $H1$-pseudoexperiments of $N = 9$ events, isolate K-subsets, $K = 1, ..9$, build probability distributions of statistic $W_{H0,K}(\Delta_K)$ and $W_{H1,K}(\Delta_K)$ and determine the power of the test π_K as a function of K, for a significance level $\alpha = 0.01$.

In the first test we require that each $H1$ pseudoexperiment has precisely 6 events from the Landau distribution and 3 events from a gaussian peak in the tail of $H0$. Consider an example where the normal and the Landau distribution are well separated. For this we take $\mu_{gauss} = 0.8$. The $H1$ distribution and π_K distributions for various tests are shown in Figure 6.4.

Note that in the standard method, i.e. when $K = N \equiv 9$, the PL_N is the most discriminative along with the AD statistic, as expected from the discussion in Section 6.1.4. Also note that at $K = 2$ and $K = 3$, most of the tests have near 100% probability to identify the anomaly in $H1$-pseudoexperiments, in contrast to much lower power of the tests looking at the entire sample of events. Using the whole samples the anomaly would be found only with $\simeq$ 10% probability using the KS test, $\simeq$ 15% using the SCM and $\simeq$ 50% using the AD. The low power of these tests at $K = 1$ can be explained by the much longer tail in the $H0$ distribution, such that the most unlikely event of an $H0$-pseudoexperiment is usually much farther on the tail of the $H0$-distribution than the one from an $H1$-pseudoexperiment.

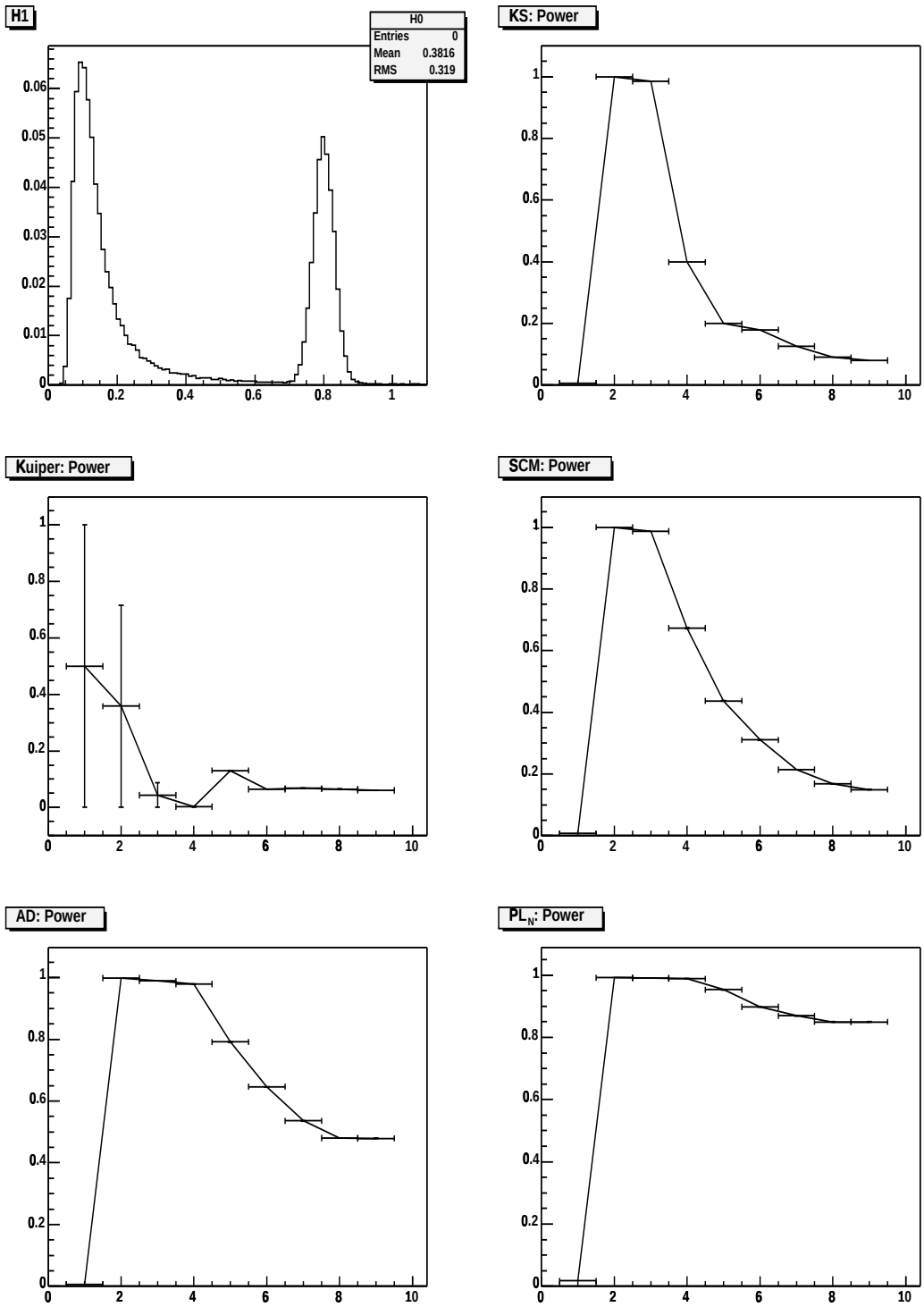

Figure 6.4: π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.8$ and requiring exactly three events from the new physics peak and six events from the $H0$ distribution in each pseudoexperiment.

The large uncertainty in the power of the Kuiper test plot is not statistical, but is associated with the fact that the Kuiper test is not very meaningful for a small number of events. For instance, for one event every outcome is equally likely, and the Kuiper statistic $\Delta_{Kuiper,K=1} \equiv 1$. Due to this feature the Kuiper statistic is not sensitive to the anomaly in question.

If a new physics peak has a larger overlap with the $H0$ distribution, then it becomes harder to isolate new physics events from the Standard Model events. In the next example $\mu_{gaus} = 0.4$ (see Figure 6.5) the KS, the SCM and the AD tests achieve maximal power at $K = 4$ due to non-negligible contribution of the $H0$-distribution in the new physics signal region. In the given example this effect is additionally enhanced by the longer tail of the $H0$-distribution, where events are counted as more unlikely than events from the gaussian peak by each of these three tests. Consequently, the discrimination is poorer between kinematic distributions of the $H0$ and the $H1$ hypotheses the more $H0$ background is contributing to the region of the anomaly. It effectively shifts the K, where the maximum sensitivity is achieved, from Q (the new physics signal contribution) to the value $Q + B$, where B is the $H0$ background contribution.

In addition, the power of the tests is much lower than in the previous example and is about $\sim 40\%$. It is still significantly higher then the one using standard methods with $K \equiv N$.

Also notice that in this example, the sensitivity of the PL_N test by looking at subsamples decreases, as the region of $x \sim 0.4$ is not considered anomalous by the PL_N, and as we discussed, the PL_N test is insensitive to an excess of events in the likely region of $H0$.

In reality, events coming from different sources are subject to statistical Poisson fluctuations and we will not have precisely 6 events coming from one channel and

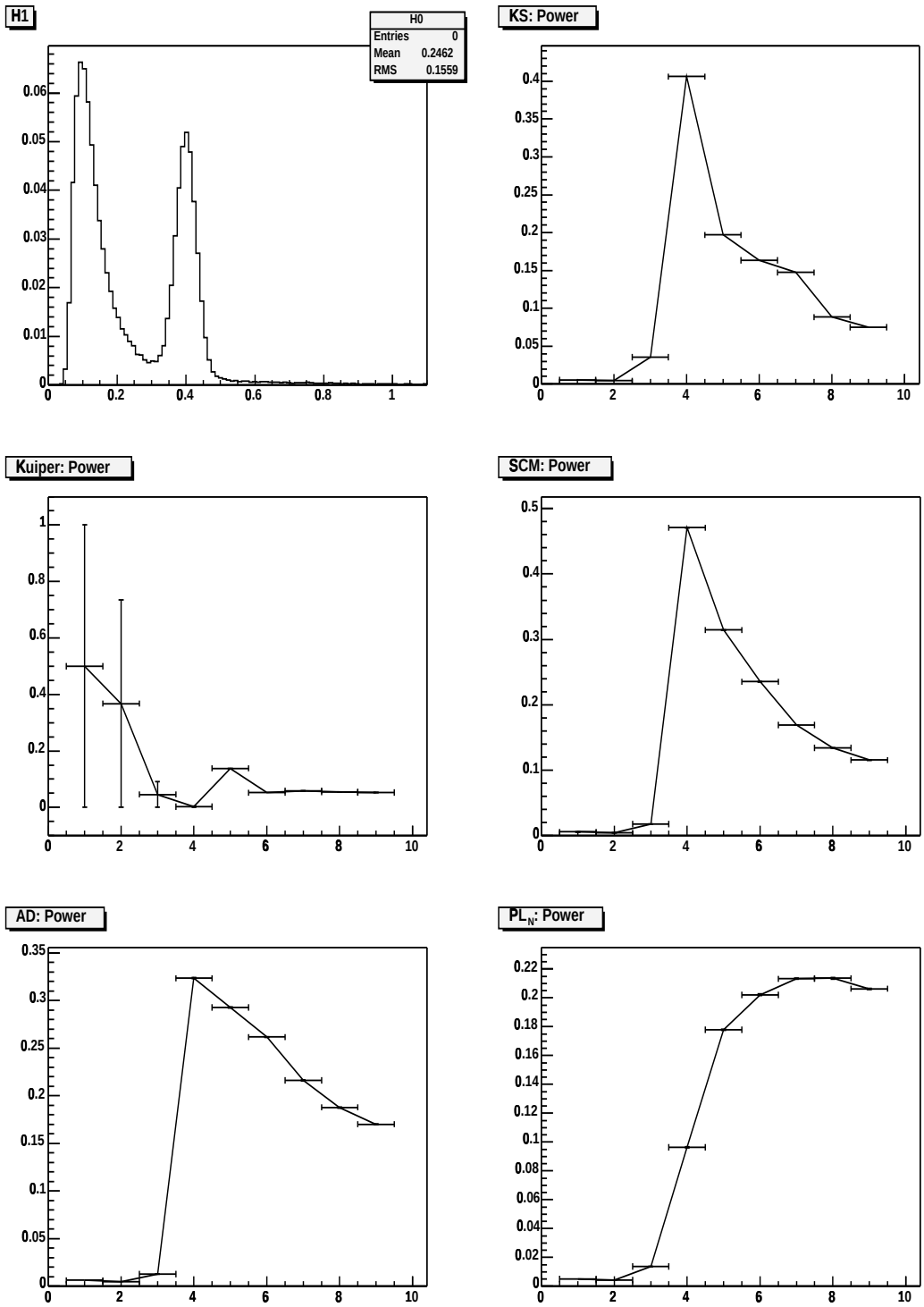

Figure 6.5: π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.4$ and requiring exactly three events from the new physics peak and six events from the $H0$ distribution in each pseudoexperiment.

3 from the other. To properly simulate a physics experiment these numbers must be Poisson smeared.

In the next example we set the gaussian peak again to $\mu_{gaus} = 0.8$ and repeat the test with Poisson smeared $H1$-pseudoexperiments. Poisson smearing is implemented by casting a random number N_1 from the Poisson distribution with the expectation $\mu_1 = 6$ and similarly casting a random number N_2 from the Poisson distribution with the expectation $\mu_2 = 3$. Only those pseudoexperiments that result in $N = N_1 + N_2 = 9$ are accepted. Then the $H1$-pseudoexperiment is constructed from random N_1 events of the $H0$ distribution and random N_2 events from the gaussian peak.

The results of the test are given in Figure 6.6. One can notice that the sensitivity of the tests to the shape anomaly is slightly lowered and the peak of maximal sensitivity is shifted to $K = 2$. This occurs because some of the $H1$-pseudoexperiments, due to Poisson smearing, can have only two events from the gaussian peak. They effectively decrease the discriminative power at $K = 3$.

If the gaussian peak is located at $\mu_{gaus} = 0.4$ and the $H1$-pseudoexperiments are Poisson smeared (see Figure 6.7) a peak of maximum sensitivity remains at $K = 4$ and, remarkably, the power of the tests increases compared with no smearing test. This effect can be explained by a larger number of unlikely events from $H0$ appearing in the tails of the distribution in the smeared $H1$-pseudoexperiments.

Summarizing, we have conducted and analyzed results of several tests on the sensitivity of various statistics to shape anomalies in the tails of kinematic distributions by looking at subsamples of events. These tests demonstrate that *the sensitivity of statistics to shape anomalies can be improved by quantifying discrepancies of unlikely subsets from the null hypothesis.*

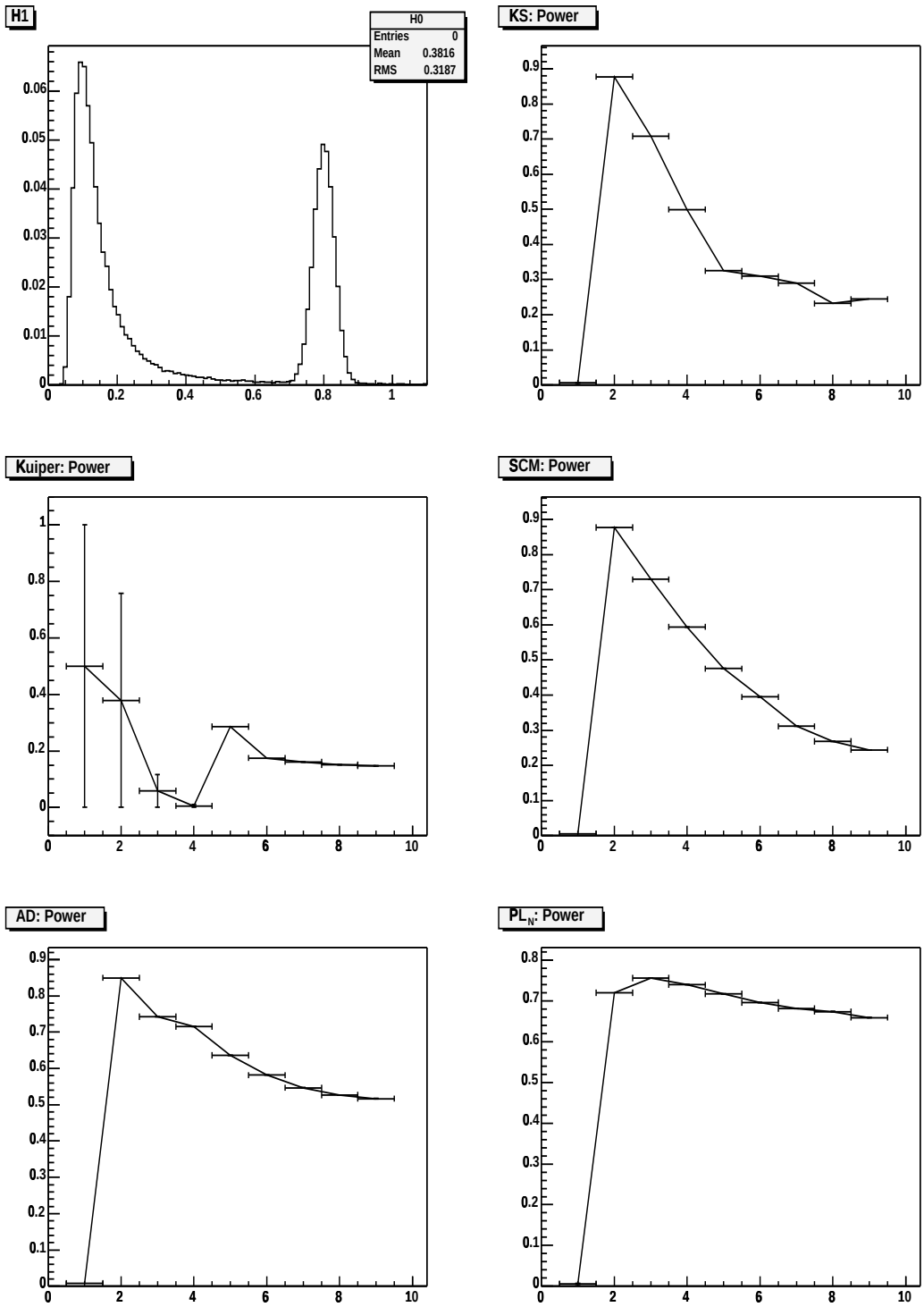

Figure 6.6: π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.8$ and Poisson smeared $H1$-pseudoexperiments.

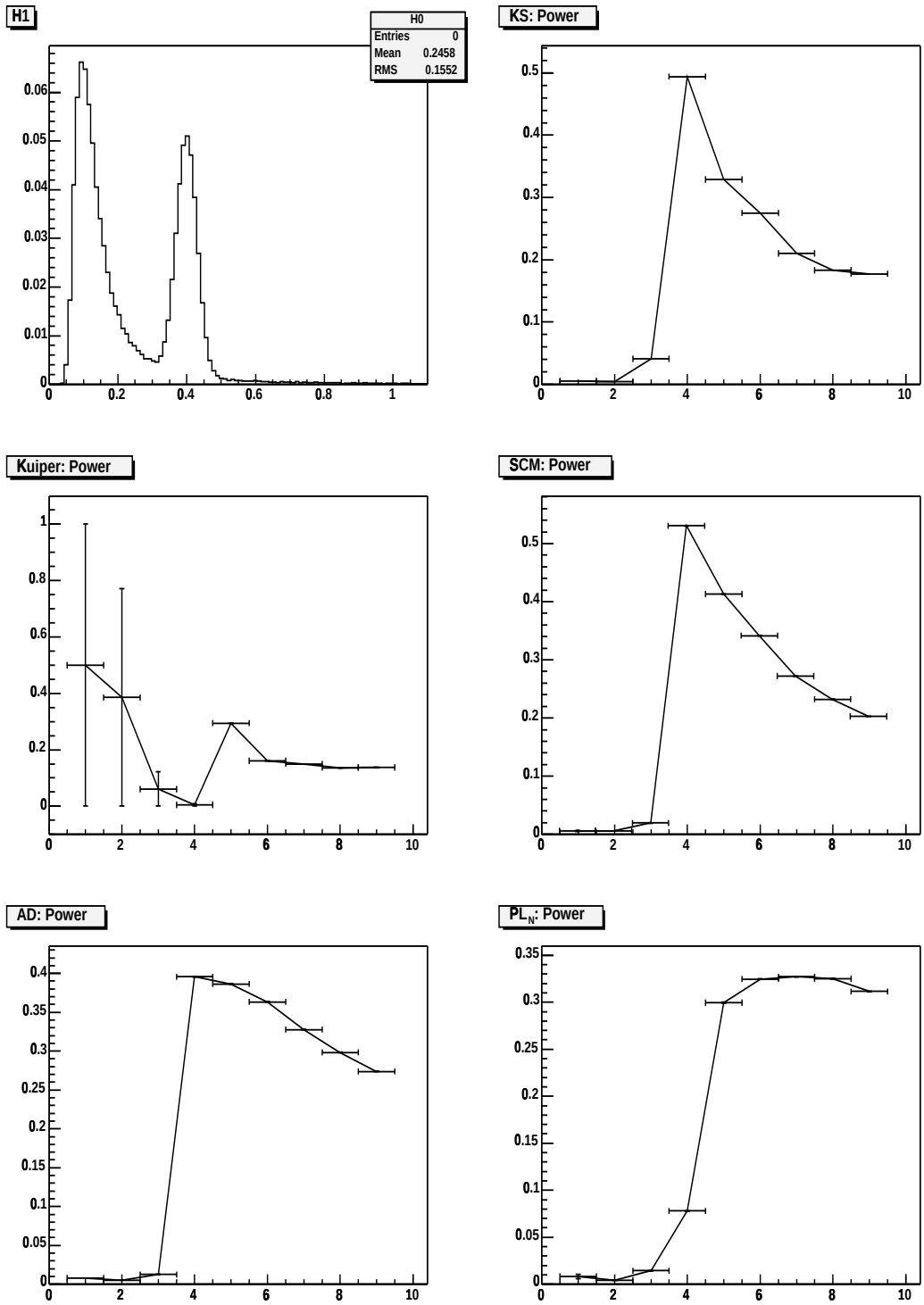

Figure 6.7: π_K distributions for various statistics in the test with the $H1$ distribution having a gaussian peak at $\mu = 0.4$ and Poisson smeared $H1$-pseudoexperiments.

6.2.3 Isolating a Most Unlikely Subset in a Data Sample

As it was illustrated in the previous Section, a subset of K isolated anomalous events exhibits on average a larger level of discrepancy than the entire sample of N data events. One, however, does not know *a priori* how many events would appear as an anomaly. The data sample is equivalent to only one pseudoexperiment and is subject to statistical fluctuations.

It can be inferred from the π_K distributions presented in Figures 6.4 – 6.7 that $W_{H0,K}(\Delta)$ and $W_{H1,K}(\Delta)$ probability distribution functions are most separated for some value of K. As a result, $H1$-pseudoexperiments tend to have the K-subset of that size in a worse agreement with $H0$ than the K-subsets of other sizes. It is natural to define a subset in the worst agreement the most unlikely.

Precisely, we define *a most unlikely subset in a data sample* the Q-subset of events revealing the largest discrepancy from $H0$ among all K-subsets, $K = 1, .., N$, so that

$$P_Q = \min_K P_K = \min_K \int_{\Delta_K}^{\infty} W_K(\Delta) d\Delta. \tag{6.24}$$

The quantity

$$\mathcal{P} = P_Q = \min_K P_K \tag{6.25}$$

then can be used to quantify the departure of the most unlikely subset from the null hypothesis. One can build the probability distribution function $\mathcal{L}(\mathcal{P})$ by generating another set of pseudoexperiments. For each pseudoexperiment one isolates the K-subsets as described in Section 6.2.1, finds P_K values using the $W_K(\Delta_K)$ probability distribution functions and determines $\mathcal{P}$. The value $\mathcal{P}$ becomes an entry into the probability distribution function $\mathcal{L}(\mathcal{P})$.

The quantity $\mathcal{P}$ serves as a test statistic. Note that unlike the majority of other test statistics, lower values of $\mathcal{P}$ correspond to larger discrepancies from the null hypothesis expectations.

Repeating the same steps for the data sample, we can determine $\mathcal{P}_{data}$ and the significance of the departure of the most unlikely subset of events from the Standard Model as follows:

$$\alpha = \int_0^{\mathcal{P}_{data}} \mathcal{L}(\mathcal{P})\ d\mathcal{P} \tag{6.26}$$

Similarly to the performed tests discussed in the previous Section one can investigate the sensitivity of the statistic $\mathcal{P}$ to shape anomalies. Remarkably, we observe that the power of $\mathcal{P}$ is equal to the maximum power of the tests $\pi = \max \pi_K$ among all $K = 1, .., N$, and therefore is significantly larger than the power of the standard tests quantifying the consistency using the entire sample.

We performed additional tests with the $H0$-distribution identical to the ones in the previous Section using the class of $H1$-distributions representing 2/3 sum of the $H0$ and 1/3 of a new physics peak - a gaussian peak with various means $\mu_{gaus} = 0.15, 0.20, ...$ and so on.

In Figure 6.8 one can see the power of various tests as a function of a gaussian position with no Poisson smearing in the $H1$-pseudoexperiments. The top plot shows the power of the standard tests. The bottom plot shows the power of the tests isolating the most unlikely subsets and quantifying their discrepancies. We observe that at gaussian positions $\mu_{gaus} > 0.6$ all statistics based on subsamples would find inconsistency with $\sim 100\%$ probability.

The same plots with realistic Poisson smeared $H1$-pseudoexperiments are given in Figure 6.9. Although due to Poisson fluctuations the resolving power of the tests is lower, the tests based on subsamples remain dramatically more powerful than

the standard ones. We observe that the tests based on the KS and the SCM statistic have better sensitivity to anomalies in question. Due to the relative simplicity of the KS test compared with the SCM, we adopt the test based on the KS statistic for searching anomalies in the tails.

6.2.4 The Purity of a Most Unlikely Subset

In addition to greater sensitivity to shape anomalies, the other advantageous property of the tests which use subsamples is their capability to identify events responsible for an anomaly in the tested distribution. In the tests looking at subsamples, we, in fact, used each statistic twice. First, we employ it to isolate subsets of unlikely events. Secondly, we use it to quantify the consistency of the subsets with the null hypothesis.

If our choice of statistic in the second case affects the sensitivity of the test to a shape anomaly, the choice of a statistic in the first case affects both the sensitivity of the test and purity of content of possible new physics candidates in unlikely subsets. Indeed, as we showed in Section 6.2.1 the KS statistic always choses an unlikely subset from events at both tails of the tested distribution. Therefore, if a new physics represents a peak in the core of the distribution, then the test based on the KS statistic would be picking up 'wrong' events. Consequently, the K-subset consistent of such events will be in agreement with the null hypothesis and the test would be insensitive to such a deviation.

From this discussion it follows that, in general, to optimize performance of the test some assumptions about properties of non-Standard Model events need to be invoked. As for our study we anticipate new physics events to preferentially populate the high-energy tail of the P_T spectrum. Therefore we ought to isolate a subset of high-P_T events.

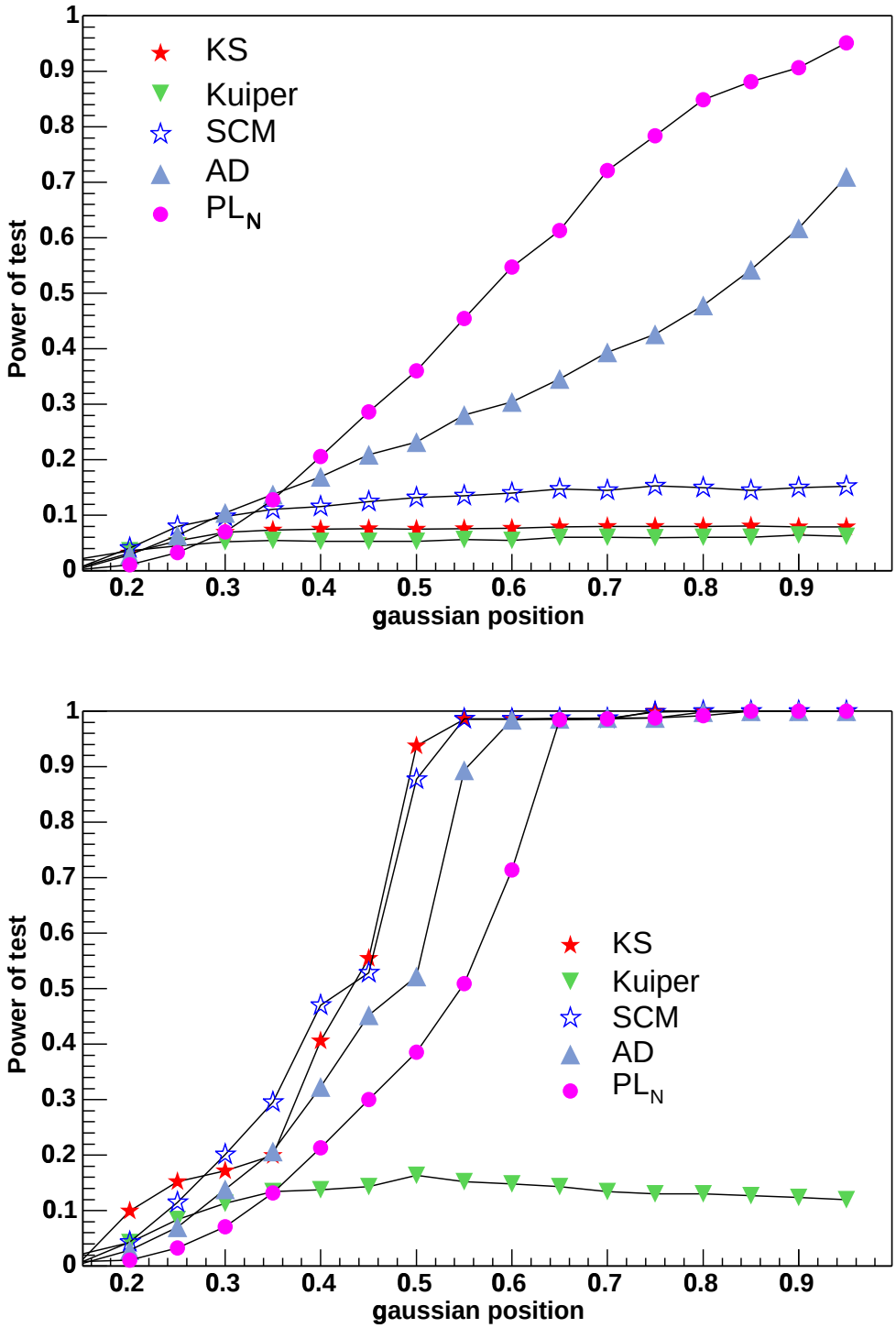

Figure 6.8: The power of various tests with isolatation a most unlikely subsample (bottom) and without (top) as a function of position of the gaussian peak for non-Poisson smeared pseudoexperiments.

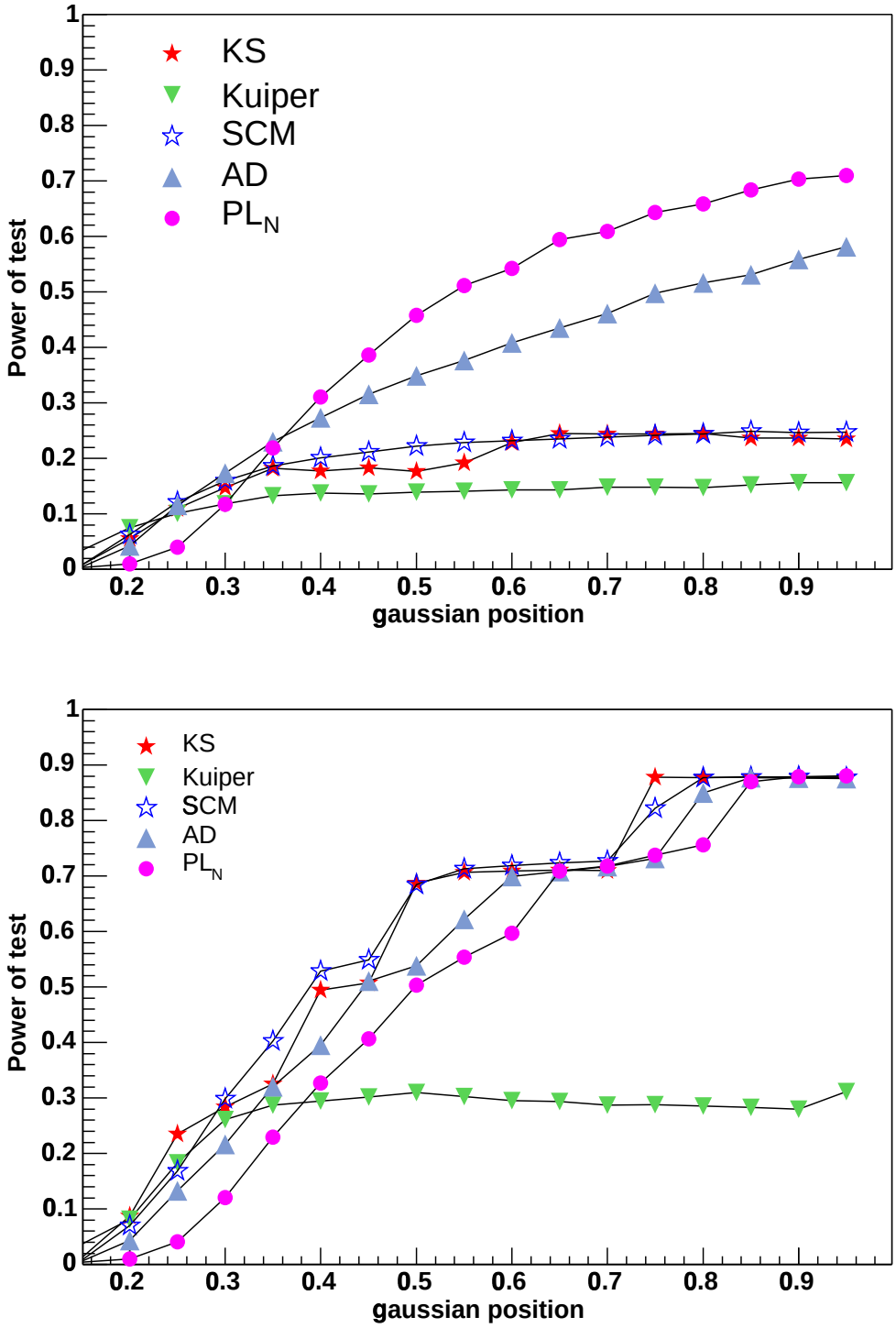

Figure 6.9: The same plots with Poisson smeared pseudoexperiments.

In the framework of the tests described above this can be done by utilizing the single-sided variants of the KS tests, Δ_+ and Δ_-, which are defined as follows [3]

$$\Delta_+ = \max_x (F(x) - S_N(x)), \tag{6.27}$$

$$\Delta_- = \max_x (S_N(x) - F(x)). \tag{6.28}$$

One can verify that each of these statistics is sensitive to anomalies in one of the tails of a kinematic distribution, and is blind to anomalies in the opposite tail. These statistics do not have merit, just as the PL_N test does not (see Section 6.1.4). Indeed, Δ_+ is large if there is an excess in the right tail, and small if there is an excess in the left tail. Consequently, if the $H1$ hypothesis favors low x values, then an $H1$-pseudoexperiment would have on average a lower Δ_+ than an $H0$-pseudoexperiment. Thus on average $H1$-pseudoexperiments would be more consistent with the null hypothesis than $H0$-pseudoexperiments.

Nevertheless, these statistics satisfy the desired properties of properly selecting the subset of possible new physics candidates from one of the tails of a kinematic distributions. Therefore for our search we adopt the statistic Δ_+ in selecting the K-subset of events, and the KS statistic to isolate a most unlikely subset, and quantify its consistency with the null hypothesis. We denote such a test (Δ_+, Δ_{KS}).

[3] The KS statistic can be expressed in terms of Δ_+ and Δ_- as $\Delta_{KS} = \max(\Delta_+, \Delta_-)$, and the Kuiper test is $\Delta_{Kuiper} = \Delta_+ + \Delta_-$.

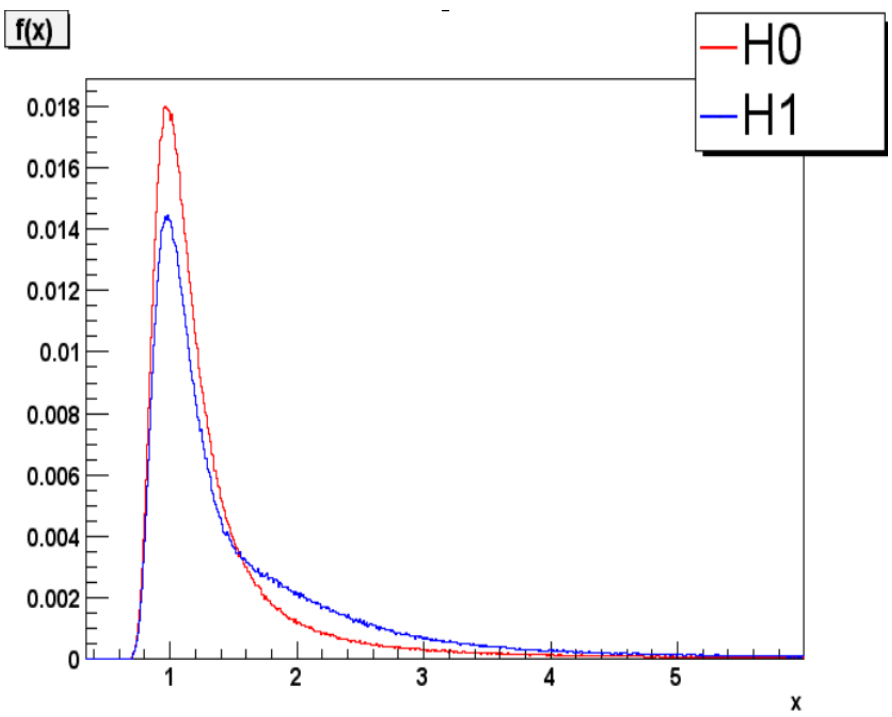

Figure 6.10: $H0$ and $H1$ distributions, where $H1$ is the one with the longer tail.

6.2.5 Performance of the Method

Alternatively to asking about the purity of events in a most unlikely subset, one can ask what would be on average a size of a most unlikely subset of events. We investigate this by performing a test with an $H0$ being a Landau distribution and adopting an $H1$ as a sum of 80% of the $H0$ and 20% of a 'new physics signal', which is also a Landau with a shifted mean (see Figure 6.10).

We generate 10,000 pseudoexperiments of 13 events from $H0$ and $H1$ distributions. In each pseudoexperiment we isolate a most unlikely subset based on (Δ_+, Δ_{KS}) statistic. The distributions of sizes of most unlikely subsets in $H0$ and $H1$ pseudoexperiments are presented in Figure 6.11 (top). One can see that the method tends to find on average 5 events in a $H1$-pseudoexperiment as the most inconsistent with the null-hypothesis, while one can naively expect that this num-

ber should be $13 \cdot 0.2 = 2.6$, i.e 2 or 3. However, as follows from the discussion in Section 6.2.2, this would be the case if a new physics contribution would be completely separated from $H0$, i.e. one uses very discriminative variable. Due to an overlap between $H0$ and a new physics signal distribution, a most unlikely subset would inevitably contain $H0$ events. The larger the overlap, the farther is the peak which is ultimately shifting to $K = 13$, as is the case for $H0$-pseudoexperiments.

It is worth mentioning that if the deviation from the null hypothesis occurs at the other tail of the distribution, then the method determines that the whole sample is unlikely. In Figure 6.11 (bottom) distributions of sizes of most unlikely subsets are given using (Δ_-, Δ_{KS}) test. In more than 50% of $H1$ pseudoexperiments, the method finds a most unlikely subset to be the entire sample. This happens because the method effectively orders events starting from one of the tails. The P_K value (Equation 6.21) decreases as more of the low-x events are included.

It is interesting to ask whether the test (Δ_+, Δ_{KS}) remains sensitive to deviations in the 'wrong' tail of the distribution. For this purpose, we determine the function $\pi(\alpha)$, called a *power function*, where π and α are given by Equations 6.15 and 6.12. In terms of a power function the test t_1 is called *more powerful* than the test t_2, if $\pi_{t_1}(\alpha) > \pi_{t_2}(\alpha), \forall\alpha$.

Figure 6.12 shows the $\pi(\alpha)$ dependence of the tests (Δ_+, Δ_{KS}) (top) and (Δ_-, Δ_{KS}) (bottom). We use as a benchmark power functions $\pi(\alpha)$ of the standard KS and AD tests. As one can see the method based on unlikely subsets has a greater sensitivity to deviations in the potentially interesting tail of the distribution than both the Kolmogorov-Smirnov and the Anderson-Darling tests. The method is also capable of detecting deviations in the other tail of the distribution, although with considerably lower sensitivity.

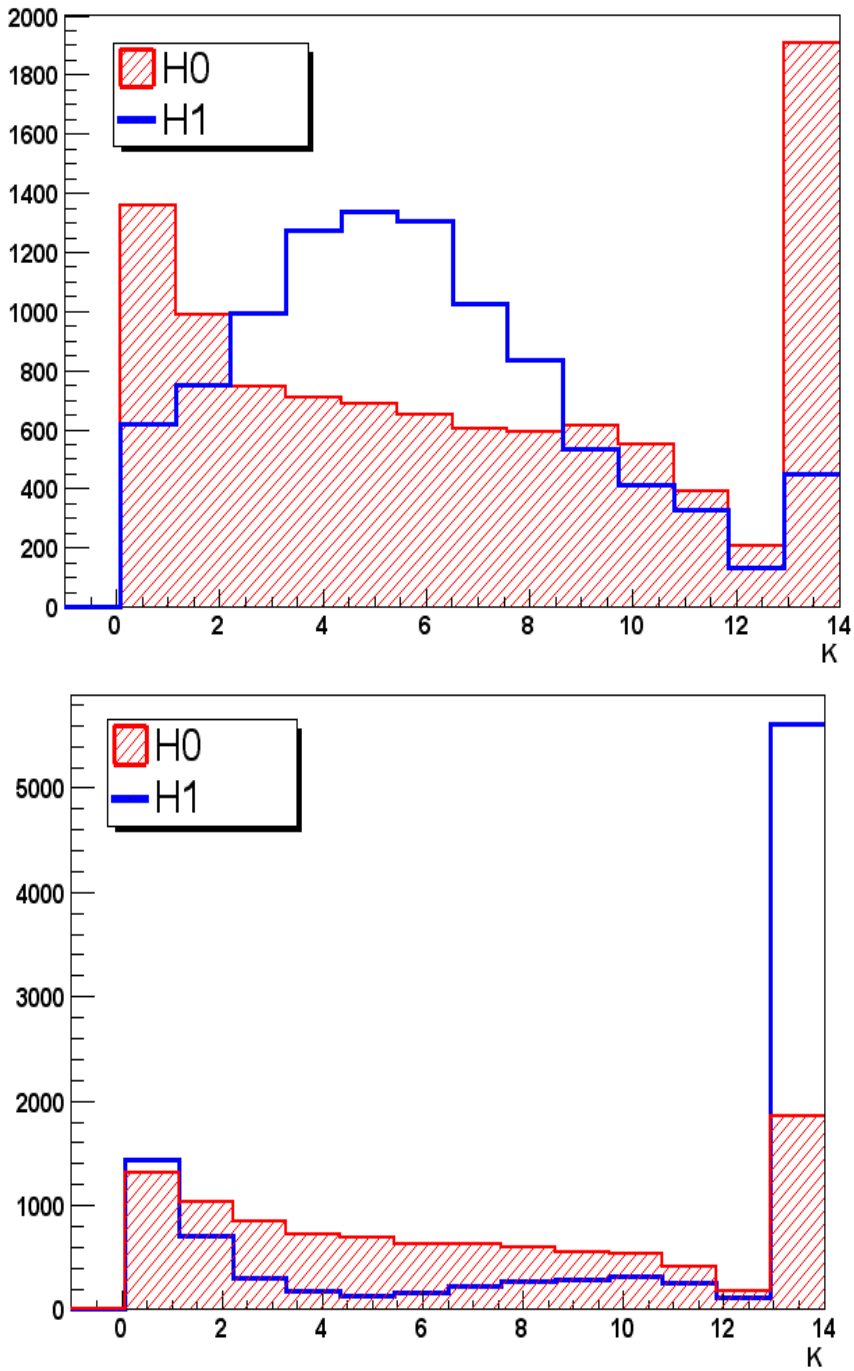

Figure 6.11: Size of most unlikely subset in a sample for $H0$ and $H1$ pseudoexperiments using (Δ_+, Δ_{KS}) (top) and (Δ_-, Δ_{KS}) (bottom) tests.

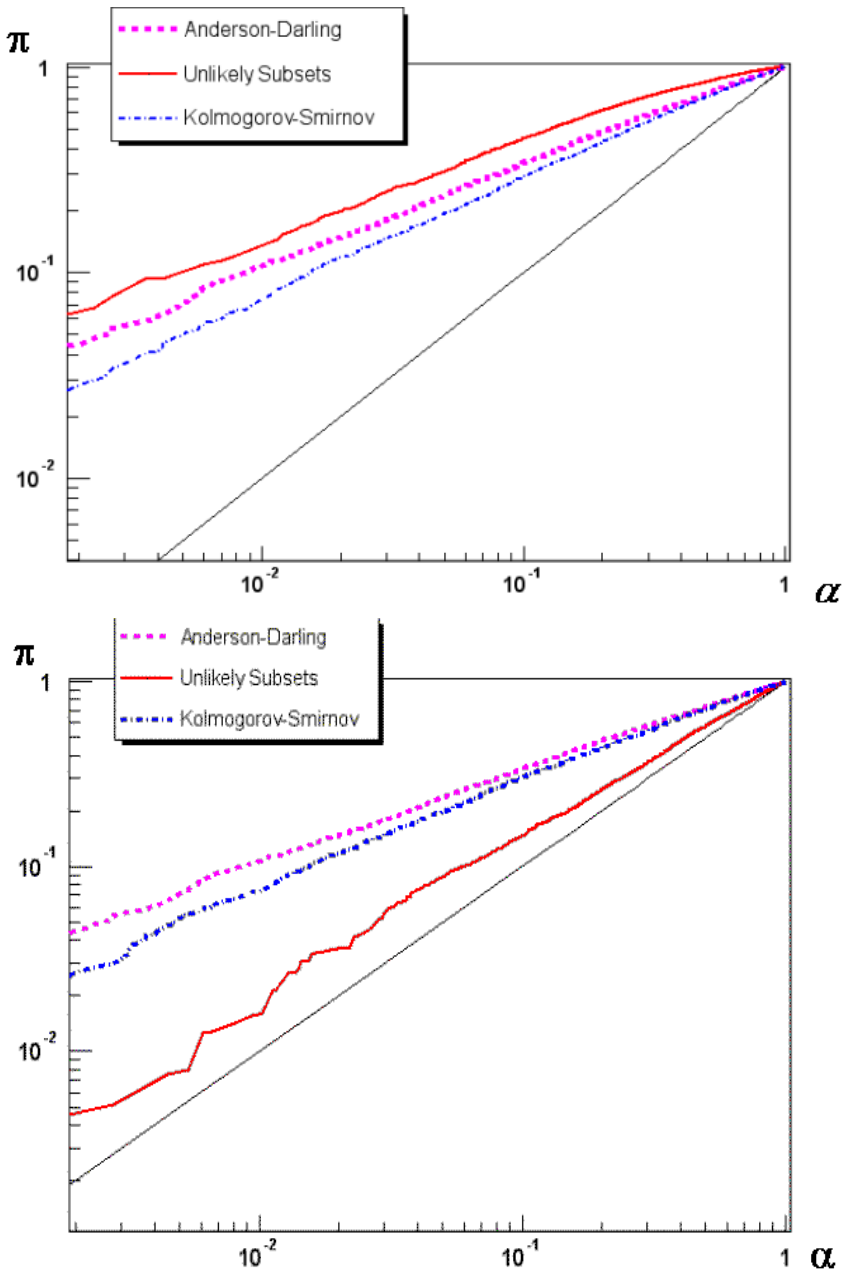

Figure 6.12: Power functions of the tests of (Δ_+, Δ_{KS}) (top) and (Δ_-, Δ_{KS}) (bottom) for $H0$ and $H1$ distributions given in Figure 6.10. The standard KS and AD statistics are used as benchmarks. The black line represents $\pi \equiv \alpha$.

6.2.6 Understanding the Method

For the statistics based on cumulative distribution functions, results of the tests are invariant to re-parametrization in the variable $x \to \psi(x)$, where $\psi(x)$ is a continuous and monotonic function. It is easier to work with a variable $y = F(x)$. In these coordinates, events of an $H0$ pseudoexperiment are distributed uniformly $\{y_i\}$, $i = 1, .., N$ in $y \in (0, 1)$ and the cumulative function is linear. The value y_i in the sample $\{y_i\}$ sorted in ascending order is called an *order statistic*.

One can see that the probability distribution function $g_i(y)$ of order statistic y_i is binomial. Indeed, if $y_i = y$, then there are $i - 1$ events such that $y_j < y$ and $N - i$ events such that $y_j > y$. Including all possible combinatorial permutations we get:

$$g_i(y) = \frac{N!}{(i-1)!(N-i)!} y^{i-1}(1-y)^{N-i} = N \begin{pmatrix} N-1 \\ i-1 \end{pmatrix} \; y^{i-1}(1-y)^{N-i}. \quad (6.29)$$

The mathematical expectation of y_i is

$$< y_i >= \frac{i}{N+1}, \quad (6.30)$$

and a variance

$$\sigma_i = \frac{1}{N+1}\sqrt{\frac{i(N+i-1)}{N+2}} \quad (6.31)$$

The non-uniform sensitivity of the KS test to deviations in the tested distribution can be understood in terms of the order statistic. The KS test effectively uses as a measure of deviation

$$\max_i (y_i - < y_i >) \quad (6.32)$$

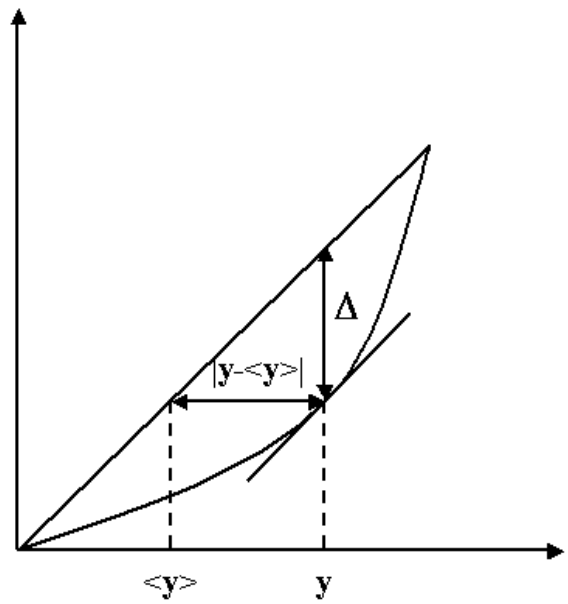

Figure 6.13: Maximum deviation Δ_{KS} is equivalent to maximum deviation of y_i from expected value $< y_i >$.

as can be seen from Figure 6.13. Indeed, if a maximum deviation Δ between two distributions is achieved at y_i, then $\Delta \simeq |y_i - < y_i > |$ [4], where y_i satisfies the pdf $g_i(y)$ (Equation 6.29). However, the width (Equation 6.31) of $g_i(y)$ is larger for y_i in the median of the $H0$ distribution than in the tail, therefore larger Δ in the core of the tested distribution is more probable. The Anderson Darling test attempts to account for this with a weight

$$\frac{1}{y(1-y)} \sim \frac{|y - < y > |}{\sigma}$$

and has more uniform sensitivity.

The order statistic probability distribution functions $g_i(y)$ have a direct relationship to $W_K(\Delta)$. As one can see on Figure 6.14 for not too large K ($K \leq N/2$)

[4] The equality is exact at $N \to \infty$

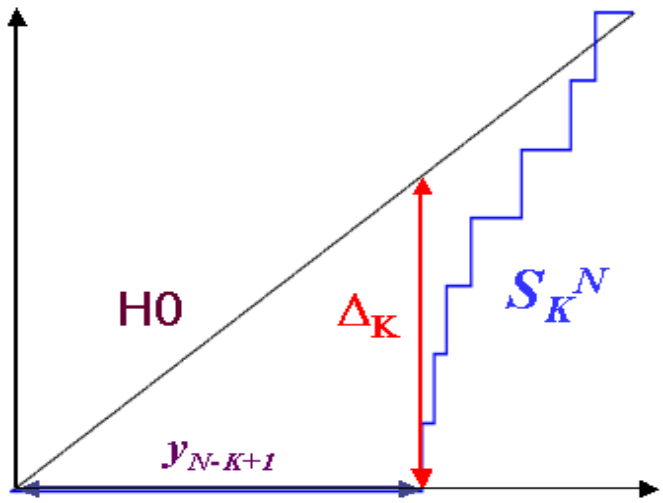

Figure 6.14: The relation between the Δ_K of the unlikely K-subset and the order statistics y_{N-K+1}.

the maximum KS deviation is achieved at $y = y_{N-K+1}$ and $\Delta_K = y_{N-K+1}$, therefore

$$W_K(\Delta) = N \begin{pmatrix} N-1 \\ N-K \end{pmatrix} \Delta^{N-K}(1-\Delta)^{K-1} \tag{6.33}$$

is binomial. For $K \to N$, $W_K(\Delta)$ approaches the Kolmogorov distribution. Since P_K (Equation 6.21) take into account the full shape of $W_K(\Delta)$ (Equation 6.33), the method properly assigns weights to events and gains an additional sensitivity at the tails.

6.2.7 Summary

We developed a single-variable kinematic technique, which is designed to isolate a most unlikely subset in a data sample and quantify its consistency with the null hypothesis. We showed that the method (Δ_+, Δ_{KS}) is very efficient at searching

for anomalies at one of the potentially interesting tails of a kinematic distribution, but at the same time it retains its sensitivity to other shape anomalies.

The method effectively selects events from one of the tails of the kinematic distribution and determines which subset of those events reveals the largest discrepancy from the $H0$. We showed that the method achieves greater performance, i.e. greater resolving power and higher purity of a new physics signal in a most unlikely subset, the better a selected kinematic variable discriminates between the $H0$ and $H1$ hypotheses. Purity of unlikely subsets and hence overall performance of this search technique can be further strengthened by incorporating additional information about events from other kinematic variables capable to provide a discrimination between both hypotheses. This is a subject of the next section.

6.3 Multi-Variate Tests

Information from several kinematic variables can be be integrated at three different levels:

1. One can construct a global kinematic variable that combines all relevant event information and serves as the best discriminator between two hypotheses. For two given hypotheses such a discriminator can be found by using the Artificial Neural Network technique (ANN) [49]. Although neural networks are good for dedicated searches, they do not provide a prescription for a generic search for shape anomalies.

2. Alternatively, one can combine the information from several variables into a single test statistic.

3. Or one can perform several one-variate statistical tests and then combine obtained probabilities into one significance measurement.

For a one-variate case we considered tests based on cumulative distribution functions due to the nice feature that results are invariant to re-parametrization of a kinematic variable $x \to \psi(x)$, which preserves the order of data points, so that for $x < y$, we have $\psi(x) < \psi(y)$. Similarly, for a multi-variate test we would like to keep this property and require that the order of events for each kinematic variable is preserved under $\vec{x} \to \psi(\vec{x})$, i.e. for any two events $\vec{x} = (x_1, x_2, ..., x_M)$ and $\vec{y} = (y_1, y_2, .., y_M)$, where a lower index j refers to a kinematic variable and M is a number of kinematic variables in use, for $x_j < y_j$ we have $\psi(x_j) < \psi(y_j)$.

This requirement will become useful for selecting unlikely subsets in a multi-variate test. It results in an invariance of a test under transformations

$$\psi(\vec{x}) = \psi_1(x_1) \cdot \psi_2(x_2) \cdot ... \cdot \psi_M(x_M). \tag{6.34}$$

We envision two possibilities for multi-variate Kolmogorov-Smirnov type tests, which we consider in the following sections.

6.3.1 PKS (Product KS) Test

If selected kinematic variables are independent then a M-variate null-hypothesis distribution can be factorized, such that

$$F(\vec{x}) = F_1(x_1) \cdot F_2(x_2) \cdot ... \cdot F_M(x_M),$$

The distributions $F_j(x_j), j = 1, .., M$ are called *marginal distributions*. In this case the problem is reduced to performing M one-variate KS-tests and obtaining

KS-probabilities $P_j, j = 1, .., M$, where the value P_j is a probability to observe $\Delta_j > \Delta_j^{data}$. One can then form a product $\Pi = P_1 \cdot P_2 \cdot ... \cdot P_M$ and determine a probability to observe Π smaller than Π_{data} using [44]

$$\alpha = \Pi_{data} \sum_{k=0}^{M-1} \frac{(-\log \Pi_{data})^k}{k!}. \tag{6.35}$$

Possible correlations between variables complicate the problem. In general,

$$F_j(x_j) = \int_0^{+\infty} dx_1 ... \int_0^{+\infty} dx_{j-1} \int_0^{+\infty} dx_{j+1} ... \int_0^{+\infty} dx_M f(x_1, x_2, ..., x_M). \tag{6.36}$$

Similarly, one can perform M one-variate tests using these cumulative distributions $F_j(x_j)$. In order to take into account the correlations, one can simulate the probability distribution function $\mathcal{M}(\Pi)$ by using pseudoexperiments and then quantify consistency of a data sample with the null-hypothesis as follows

$$\alpha = \int_0^{\Pi_{data}} \mathcal{M}(\Pi) d\Pi. \tag{6.37}$$

We call this technique *PKS (Product KS) test.* One may argue that when using marginal distributions, some information about the null-hypothesis distribution may be lost. In the next section we combine the information from all variables into a single test statistic.

6.3.2 MKS (Multi-dimensional KS) Test

In this method we make use of the full M-dimensional cumulative distribution

$$F(\vec{x}) = F(x_1, x_2, ..., x_M) = \int_0^{x_1} \int_0^{x_2} ... \int_0^{x_M} f(\xi_1, \xi_2, ..., \xi_M) d\xi_1 d\xi_2 ... d\xi_M. \tag{6.38}$$

Similarly to the one-variate test we can define the KS-statistic extended to M dimensions as follows

$$\Delta_{KS} = \max_{\vec{x}} |F(\vec{x}) - S_N(\vec{x})|, \tag{6.39}$$

where the empirical distribution function $S_N(\vec{x})$ gives the fraction of data points $y = (y_1, y_2, ..., y_M)$ such that $y_j < x_j$, for all j. If $n(\vec{x})$ is the number of data events satisfying this requirement for a given $\vec{x}$, then

$$S_N(\vec{x}) = \frac{n(\vec{x})}{N}. \tag{6.40}$$

Such a multi-variate extension to the KS test was originally proposed by Peacock [50]. He noted, however, that this definition of the cumulative distribution is ambiguous, since the directions in which we choose to order $x_1, x_2, ...$ are arbitrary. In one dimension it makes no difference, as the probability $P(X > x)$ to observe x less than the value X is related to the probability $P(X \leq x)$ to observe x greater or equal to X as follows

$$P(X > x) = 1 - P(X \leq x).$$

However, in two-dimensional case, for instance, one obtains

$$P(X > x, Y > y) + P(X \leq x, Y > y) + P(X > x, Y \leq y) + P(X \leq x, Y \leq y) = 1,$$

so that there are three independent variables. Therefore, in general, the results of the test will be dependent on the corner where the cumulation of $F(\vec{x})$ starts.

In the procedure suggested by Peacock, one finds the maximum deviations starting the cumulation from each corner, and thereafter adopts the largest of the differences.

6.3.3 Effect of Correlations and Dimensionality on Multi-variate Statistical Tests

Both of the discussed techniques are invariant to reparametrizations given by Equation 6.34 and preserve the event ordering. If kinematic variables are independent for both $H0$ and $H1$, then this is sufficient to ensure robustness of results of the test. However, if variables are correlated then a more general transformation $\psi(\vec{x})$ of coordinates will affect results of the test and consequently its performance. Depending on the transformation, the test can either provide a better discrimination between hypotheses or hide discrepancies, as it will be shown in Section 6.3.5.

For a dedicated search one can optimize variable selection to augment the resolving power. For a generic search no such handles exist, therefore we ought to make heuristic assumptions about what kinematic quantities would most likely exhibit anomalous behaviour.

One would be tempted to use as much information as possible and therefore include a large number of kinematic variables into the test. Unfortunately, non-discriminative quantities add a statistical noise and dilute the test performance.

Suppose we add to the test a meaningless variable x_{M+1}, which is uniformly distributed in $[0, 1]$ for both $H0$ and $H1$ hypotheses. In the PKS technique a flat probability distribution in x_{M+1} results in smearing of the probability distribution function $\mathcal{M}$ of $\Pi = p_1 \cdot ... \cdot p_M \cdot p_{M+1}$. This increases an overlap between $\mathcal{M}_{H0}(\Pi)$ and $\mathcal{M}_{H1}(\Pi)$ and hence dilutes the power of the test.

In the MKS technique, events falling close to the corners of the M-dimensional cube have a larger impact on the multi-variate KS-statistic than those at the center, similar to the effect of events at the tails in the one-variate KS test. The

additional degree of freedom x_{M+1} randomly re-weights events by changing the order of their importance, which results in the smearing of probability distributions functions and diminishing the sensitivity of the test.

Since selection of potentially interesting kinematic variables and possible correlations between them have a large impact on results of the test, in the next sections we investigate performance of the multivariate statistical tests when applied to real physics distributions.

6.3.4 Barnett and Hall SUSY Model as an Alternative Hypothesis

As an alternative hypothesis, we consider one by Barnett and Hall [4], which was an attempt to provide an explanation of the characteristics of unusual top dilepton events observed in CDF and D0 Run I data and discussed in Chapter 1.

In their work they selected two dilepton events with kinematic features quite unlike those expected from $t\bar{t}$ production. Specifically, the events in question were characterized with large $\not\!\!E_{\mathrm{T}}$ and large P_T of the leading lepton in an event. In addition to these peculiarities, the authors noticed that $t\bar{t}$ events with high-$\not\!\!E_{\mathrm{T}}$ and large P_T tend to have leptons going in opposite directions unlike the observed events in question. Although this could be a result of a statistical fluctuation, the authors hypothesized that these events originate from decays of supersymmetric particles and analyzed what would be the consequences for supersymmetry.

They argued that a cascade decay of heavy squarks and gluinos through charginos $\tilde{\chi}^+$ and neutralinos $\tilde{\chi}^0$ would yield events resembling the top dilepton signature, where the high lepton E_T can be achieved due to a large mass difference between initial and final superpartners. Assuming the high branching ratio of the

chargino decays to the desired leptonic mode, they suggested the decay scenario and roughly estimated the masses of the supersymmetric particles.

In the scheme they proposed the dilepton events with unusually high momenta of particles arise from decay of heavy squarks with a mass of about 300 GeV. The predominant decay of each squark is $\tilde{q_L} \to q(\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^0)$ followed by $\tilde{\chi}_1^- \to \ell'\tilde{\nu}^\dagger, \bar{\nu}\tilde{\ell}$ and $\tilde{\chi}_2^0 \to \ell'\tilde{\ell}^\dagger, \nu\tilde{\nu}^\dagger$, and finally $\tilde{\ell} \to \ell\tilde{\chi}_1^0$ and $\tilde{\nu} \to \nu\tilde{\chi}_1^0$. These events have a wide range of characteristics with large $\not{E}_{\mathrm{T}}$ and the number of isolated charged leptons N_L, varying from 0 to 4. Events with three and more leptons have higher suppression factors, while events with $N_L = 0$ and 1 have large Standard Model backgrounds.

Barnett and Hall speculate that the features of those events are better explained with the following set of parameters: $m_{\tilde{g}} \simeq 330$ GeV, $m_{\tilde{q}} \simeq 310$ GeV, $m_{\tilde{l}_L} \simeq 220$ GeV, $m_{\tilde{\nu}} \simeq 220$ GeV, $m_{\tilde{l}_R} \simeq 130$ GeV, $\mu \simeq -400$ GeV, $M_1 \simeq 50$ GeV and $M_2 \simeq 260$ GeV. In this scenario squarks masses are so high that the signal can be kinematically distinguished from the $t\bar{t}$ background.

We simulated a sample of 400,000 SUSY events in SPYTHIA Monte Carlo using parameters suggested by Barnett and Hall and then applied the top dilepton event selection criteria discussed in Chapter 5. 8,662 events survived the cuts. The Figure 6.15 shows the distribution of $\not{E}_{\mathrm{T}}$ and leading lepton P_T for both $t\bar{t}$ and SUSY with Barnett/Hall parameters. One can see that indeed these two kinematic variables provide a good discrimination between two processes.

In addition to these quantities, we explored other kinematic variables, which would help us to distinguish between the $t\bar{t}$ and this SUSY model. Due to relatively small contribution from non-top Standard Model background sources and because non-top events are on average less energetic than the $t\bar{t}$, the top decays are the major background to new physics we hope to discover.

As the top dilepton event signature is characterized by two highly energetic

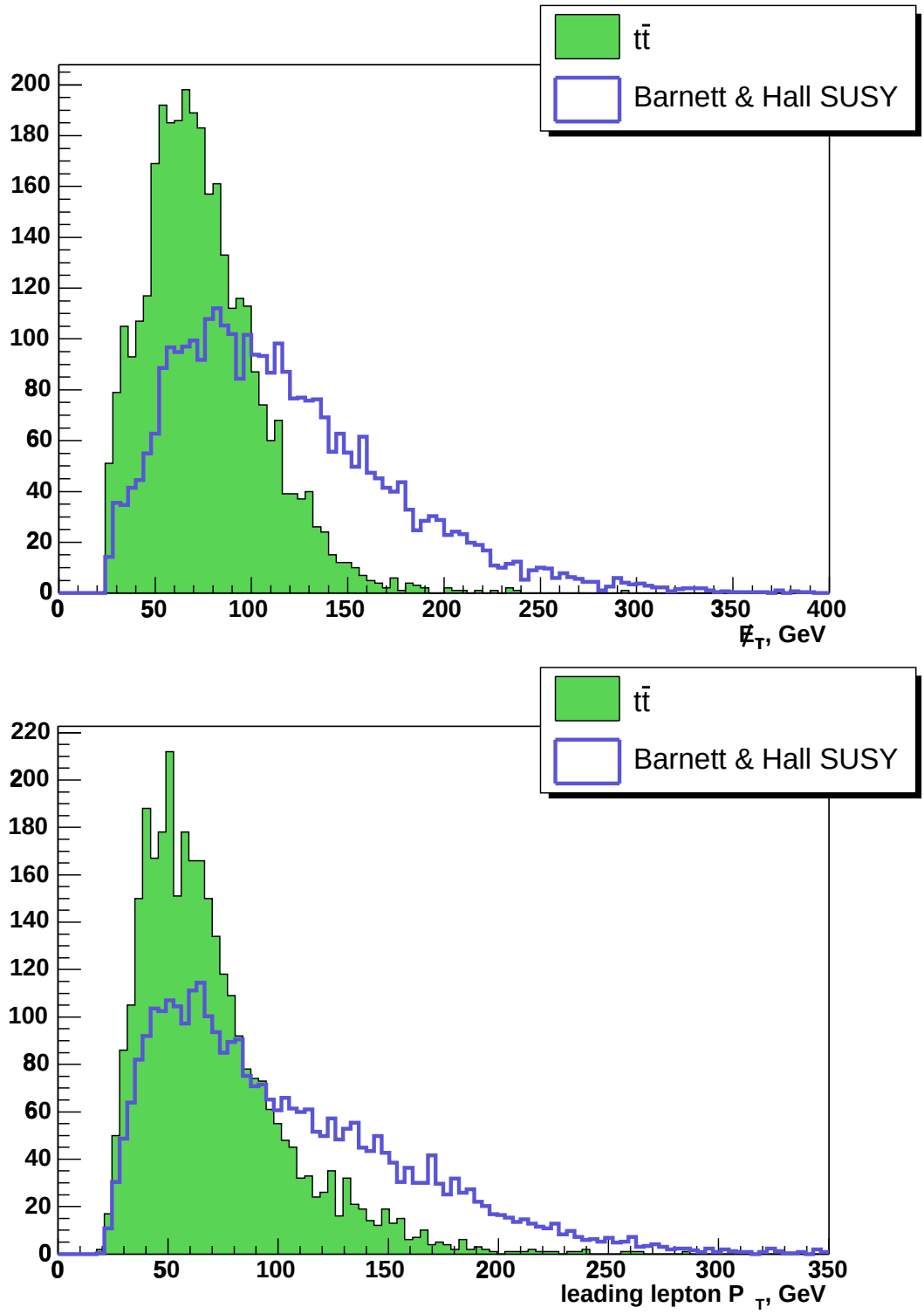

Figure 6.15: $t\bar{t}$ and Barnett/Hall SUSY distributions of $\not{E}_{\mathrm{T}}$ and leading lepton P_T. These distributions provide a good discrimination between two processes.

leptons, two jets and a large missing energy, there are four fully measured four-momentum vectors and x-y components of vector $\not{E}_{\mathrm{T}}$ in the final state. Using these quantities one can construct many kinematic variables.

Because of the boost in the z-direction absolute values of three-dimensional vectors and absolute angles between them do not have valuable information. The most interesting kinematic variables are transverse components of momenta of the particles, and the relative directions of their motion. We divide kinematic variables into three distinct families.

In the first family of variables we include transverse components of momenta and their linear combinations, since they can also provide a discrimination. Namely, we consider transverse components of missing energy $\not{E}_{\mathrm{T}}$, the momenta of the leptons $P_{T,\ell_1}, P_{T,\ell_2}$ and jets P_{T,j_1}, P_{T,j_2}, their sums $\sum E_{T,\ell} = P_{T,\ell_1} + P_{T,\ell_2}, \sum E_{T,j} = P_{T,j_1} + P_{T,j_2}$, the total transverse energy $H_T = \sum E_{T,\ell} + \sum E_{T,j} + \not{E}_{\mathrm{T}}$, and also the quantity defined by Barnett and Hall $E_S = \not{E}_{\mathrm{T}} + \sum E_{T,\ell}$ and $E_B = \not{E}_{\mathrm{T}} + \sum E_{T,j}$.

The second family are the angles between momenta in the plane transverse to the beam. These are the angles between the two leptons $\Delta\phi_{\ell\ell}$, the highest energy lepton and the leading jet $\Delta\phi_{\ell j}$, two jets $\Delta\phi_{jj}$, the highest energy lepton and missing energy $\not{E}_{\mathrm{T}} - \Delta\phi_{\ell m1}$, the same angle with the second lepton $\Delta\phi_{\ell m2}$, the leading jet and $\not{E}_{\mathrm{T}} - \Delta\phi_{jm1}$, the second jet and $\not{E}_{\mathrm{T}}$ - $\Delta\phi_{jm2}$.

The third family of variables are invariant masses. We considered $m_{\ell\ell}$, m_{jj} and $m_{\ell j}$, where the last quantity is an invariant mass of the most energetic lepton and the leading jet.

For each of those variables we conducted the KS test of the compatibility of the $t\bar{t}$ and the SUSY distributions. The respective KS probabilities are given in Table 6.5. As expected, missing energy $\not{E}_{\mathrm{T}}$ and P_{T,ℓ_1} turn out to be the most discriminative kinematic variables, along with the linear sums, which use

the information from $\not{E}_{\mathrm{T}}$ and P_{T,ℓ_1}, as, for instance, H_T and E_S. The majority of other kinematic distributions do not provide a noticeable discrimination, as is evidenced in Figure 6.16.

Barnett and Hall argued that a possible third discriminative variable for separating the $t\bar{t}$ and SUSY processes in question is the distribution of angle $\Delta\phi_{\ell\ell}$ between leptons in the plane transverse to the beam. The separation is more prominent at high energies $\not{E}_{\mathrm{T}} + P_{T,\ell_1}$, however, and it is not as significant if one considers all events passing the dilepton event selection, as it is illustrated in Figure 6.17.

Since the major differences between the event topology of $t\bar{t}$ and the SUSY process are the boosted lepton and the value of $\not{E}_{\mathrm{T}}$, conservation laws require these quantities to be back-to-back. In fact, we observe better discrimination for the distribution of the angle $\Delta\phi_{\ell m}$ in the transverse plane to the beam between the leading lepton and direction of $\not{E}_{\mathrm{T}}$ (see Figure 6.18).

6.3.5 Testing Multi-Variate Statistics on Physics Distributions

In order to find a set of the most discriminative variables we performed multivariate tests described above, using various combinations of kinematic variables. Since only some fraction of events would be associated with SUSY processes we were comparing $t\bar{t}$ distributions with 50% mixture of $t\bar{t}$ and SUSY. For this study we required a significance level of $\alpha = 0.01$ and used pseudoexperiments of $N = 30$ events. For each dimensionality M of the test, a set of the variables with the highest discriminative power π_M^{max} was found.

We observe that the power of the MKS technique degrades with increasing

kinematic variable	KS probability
$\not{E}_{\mathrm{T}}$	$< 10^{-324}$
P_{T,ℓ_1}	2.26×10^{-41}
P_{T,ℓ_2}	0.984
P_{T,j_1}	2.04×10^{-2}
P_{T,j_2}	0.293
$\sum E_{T,\ell}$	3.51×10^{-30}
$\sum E_{T,j}$	0.234
H_T	$< 10^{-324}$
E_S	$< 10^{-324}$
E_B	8.26×10^{-38}
$\Delta\phi_{\ell\ell}$	3.41×10^{-6}
$\Delta\phi_{\ell j}$	5.25×10^{-9}
$\Delta\phi_{jj}$	0.882
$\Delta\phi_{\ell m1}$	2.92×10^{-37}
$\Delta\phi_{\ell m2}$	5.52×10^{-2}
$\Delta\phi_{jm1}$	6.10×10^{-3}
$\Delta\phi_{jm2}$	0.918
$m_{\ell\ell}$	0.214
m_{jj}	5.55×10^{-03}
$m_{\ell j}$	6.31×10^{-5}

Table 6.5: The KS probabilities of compatibility of $t\bar{t}$ distribution with SUSY for various kinematic variables.

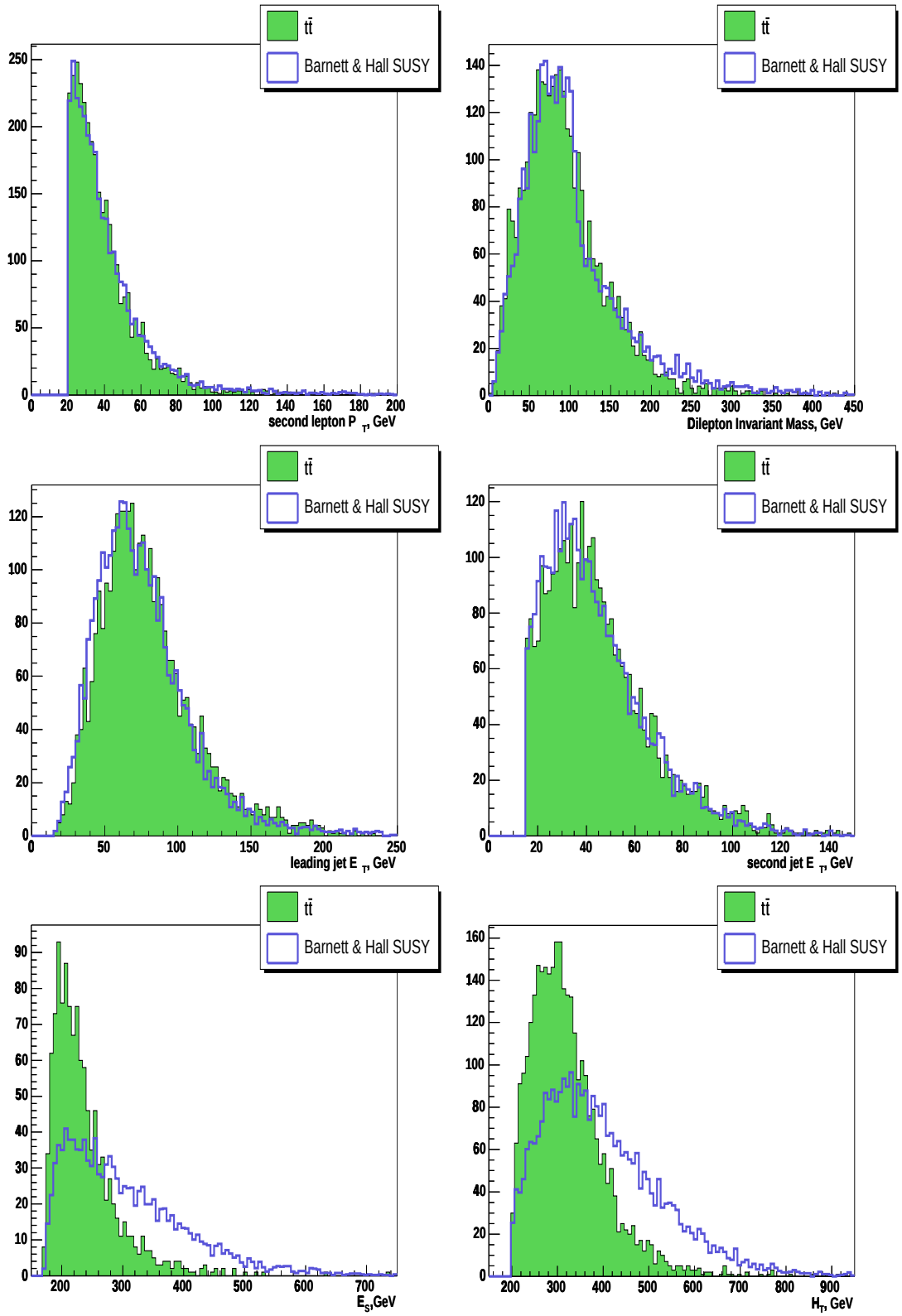

Figure 6.16: $t\bar{t}$ and Barnett/Hall SUSY distributions of some of the kinematic quantities.

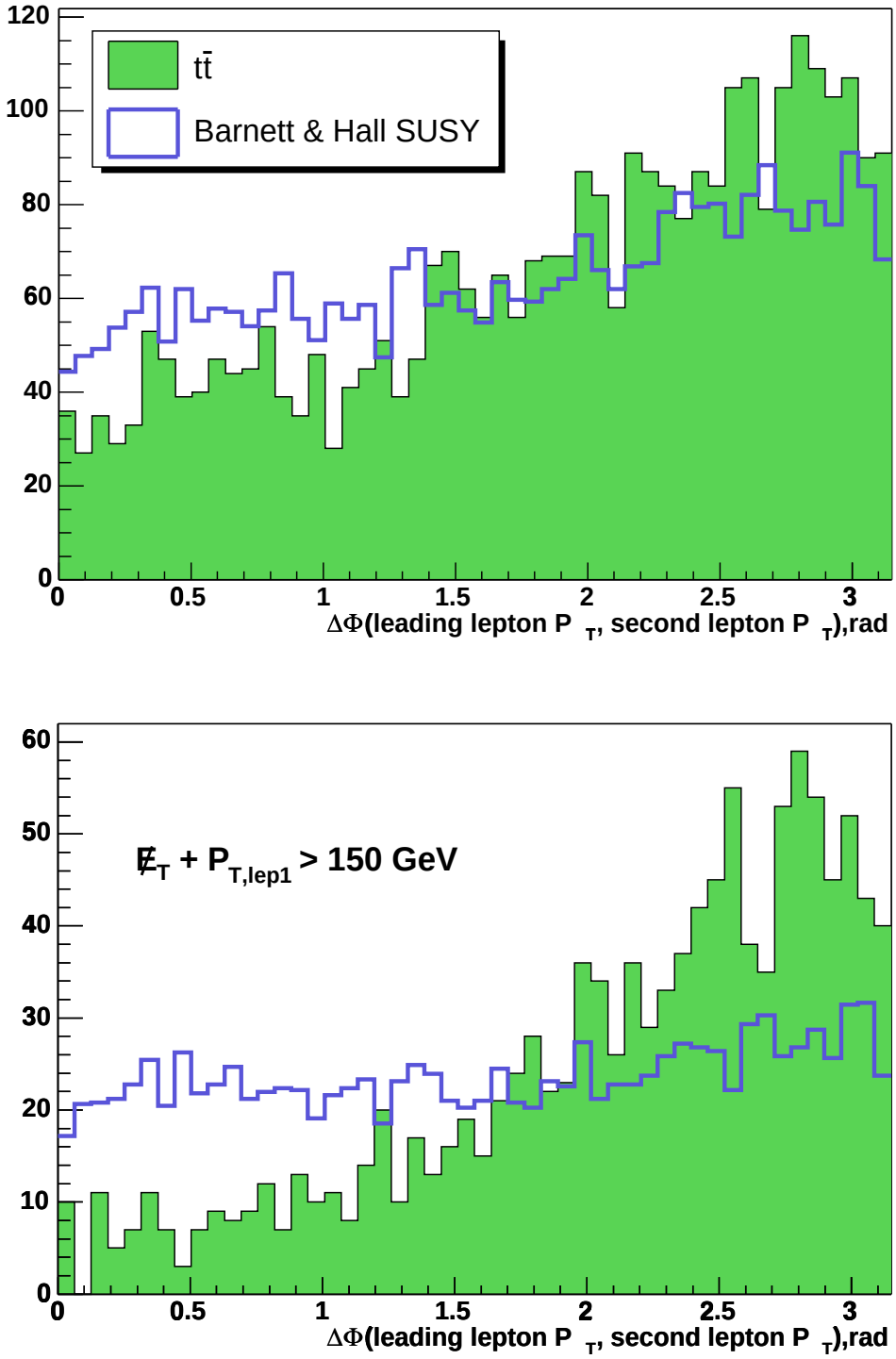

Figure 6.17: $t\bar{t}$ and Barnett/Hall SUSY distributions of angle $\Delta\phi_{\ell\ell}$ in transverse to the beam plane between two leptons for all events passing dilepton selection (top) and for events with $\not{E}_{\mathrm{T}} + P_{T,\ell 1} > 150$ GeV (bottom).

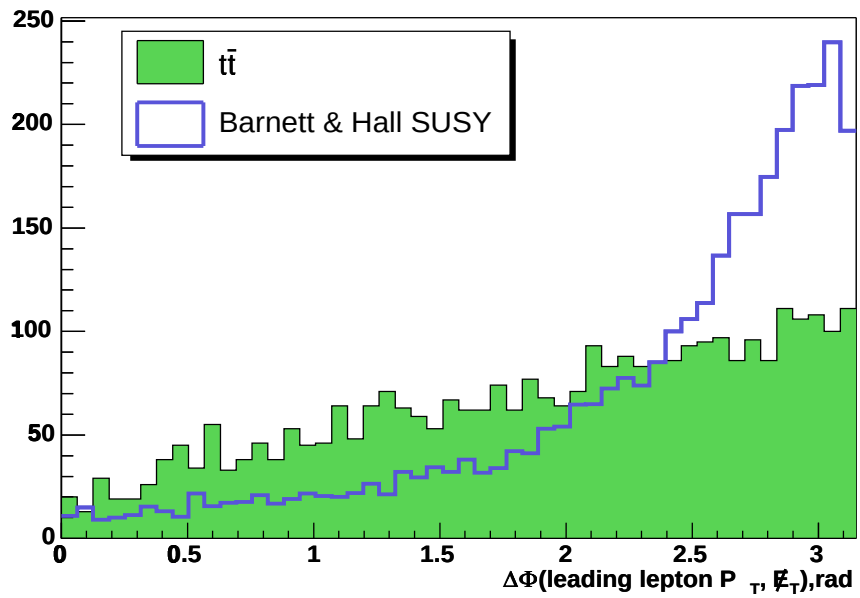

Figure 6.18: $t\bar{t}$ and Barnett/Hall SUSY distributions of angle $\Delta\phi_{\ell m}$ in transverse to the beam plane between leading lepton and direction of $\not{E}_{\mathrm{T}}$.

dimensionality. The MKS method determines that $\not{E}_{\mathrm{T}}$ is the most discriminative variable with $\pi_1^{max} = 0.33 \pm 0.01$. However, in two dimensions the power of the test remains almost the same: $\pi_2^{max} = \pi_2(E_S, \not{E}_{\mathrm{T}}) = 0.35 \pm 0.01$ and for higher dimensions it drops. We explain this effect as a bias of the MKS test to the corners of the M-dimensional cube, i.e. ascribing unjustified higher weights to events falling too close to the corners. The bias is rapidly increasing with the number of dimensions due to the fact that the number of corners is rising with the number of dimensions M as 2^M and that the corner phase space is decreasing in size. For instance, if an event falls close to the corner $(0, 0, .., 0)$, so that its location in M-dimensional parameter space is $(\epsilon, \epsilon, .., \epsilon)$, where $\epsilon \ll 1$, the effective corner phase space ϵ^M decreases with rising M.

The PKS method does not suffer from this problem. The results of this test

are the following:

$$\pi_1^{max} = \pi_1(\not{E}_{\mathrm{T}}) = 0.33 \pm 0.01$$

$$\pi_2^{max} = \pi_2(\not{E}_{\mathrm{T}}, E_S) = 0.46 \pm 0.01$$

$$\pi_3^{max} = \pi_3(\not{E}_{\mathrm{T}}, E_S, \Delta\phi_{\ell m1}) = 0.50 \pm 0.01$$

$$\pi_4^{max} = \pi_4(\not{E}_{\mathrm{T}}, E_S, \Delta\phi_{\ell m1}, P_{T,\ell_1}) = 0.51 \pm 0.01$$

$$\pi_5^{max} = \pi_5(\not{E}_{\mathrm{T}}, E_S, \Delta\phi_{lm1}, P_{T,\ell_1}, E_B) = 0.52 \pm 0.01$$

Adding more variables results in a decrease of sensitivity.

From this study it follows that the PKS technique is better suited as a generalized multi-variate KS test, since adding information from other kinematic variables provides better discrimination between two hypotheses. Therefore we adopt the PKS method for our search.

Note, that the greatest performance of the PKS was achieved when highly correlated kinematic variables were selected. In the set of five kinematic variables it used information from P_{T,ℓ_1} twice and from $\not{E}_{\mathrm{T}}$ three times. We observe that the PKS technique has greater resolving power if the kinematic variables are highly correlated in the $H1$-distribution.

For instance, the correlation coefficient $\rho_{SUSY}(\not{E}_{\mathrm{T}}, E_S) = 0.88$ while for $\not{E}_{\mathrm{T}}$ and leading lepton P_T it is smaller, $\rho_{SUSY}(\not{E}_{\mathrm{T}}, P_{T,\ell_1}) = 0.56$. As a result the power of the test using $\not{E}_{\mathrm{T}}$ and P_{T,ℓ_1} is slightly lower – $\pi(\not{E}_{\mathrm{T}}, P_{T,\ell_1}) = 0.43$ than when using $\not{E}_{\mathrm{T}}$ and E_S. Consequently, a linear combination of original kinematic variables may result in better sensitivity of the test.

We investigated the effect of correlations between variables on the power of the test by transforming coordinates to a basis in which correlations in $H0$ are minimized. Let μ_i be a mathematical expectation of variable x_i, i.e.

$$\mu_i \equiv E(x_i)$$

and

$$\sigma_i^2 \equiv E\{(x_i - \mu_i)^2\}.$$

One can form the *moment matrix* for a M-variate distribution as follows

$$\Lambda \equiv (\lambda_{i_1,i_2,..,i_M}) \equiv E\{(x_{i_1} - \mu_{i_1})^2(x_{i_2} - \mu_{i_2})^2...(x_{i_M} - \mu_{i_M})^2\}.$$

The rank of Λ is the minimum number of independent variables. Λ is a symmetric, real matrix. Hence, we can find an orthogonal transform

$$x' = Cx$$

such that the new moment matrix

$$\Lambda' = C\Lambda C^T$$

is a diagonal matrix.

For instance, in a two-dimensional case the matrix C would be the following

$$C = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

where angle α satisfies

$$\tan 2\alpha = \frac{2\lambda_{12}}{\sigma_2^2 - \sigma_2^2}$$

We made such a transformation in $(\not{E}_{\mathrm{T}}, P_{T,\ell_1})$ plane. As a result a correlation coefficient $\rho_{t\bar{t}}$ decreased from 0.12 to 0.03 and ρ_{SUSY} also decreased from 0.56 to 0.23. Although the ratio $\rho_{SUSY}/\rho_{t\bar{t}}$ got higher, the power of the test dropped from 0.43 to 0.29. This is even lower than a single-variate test using $\not{E}_{\mathrm{T}}$ only!

Note that by making these transformations of coordinates we effectively changed the event ordering. It follows from this discussion that due to correlations between kinematic variables the event ordering can be optimized for better sensitivity of the test by performing a comparison between given two hypotheses. No preferred basis based on purely $H0$ distribution was found.

Due to the absence of knowledge about the exact shape of $H1$ distribution and correlations between variables and without adhering to parameters of the discussed SUSY model, we choose basic kinematic variables for our generic search: missing transverse energy $\not{E}_{\mathrm{T}}$, the transverse component of the leading lepton P_{T,ℓ_1} and the azimuth angle between these quantities $\Delta\phi_{\ell m1}$.

6.3.6 Topology of $t\bar{t}$ Dilepton Event.

The high purity of the top dilepton sample with respect to other Standard Model backgrounds allows us to select an additional variable intrinsic to the properties of the $t\bar{t}$ dilepton decay. The initial and intermediate state particles impose constraints on the final state product properties:

$$\begin{aligned}(P_{\ell_1} + P_{\nu_1} + P_{b_1})^2 = (P_{\ell_2} + P_{\nu_2} + P_{b_2})^2 = m_t^2 \\ (P_{\ell_1} + P_{\nu_1})^2 = (P_{\ell_2} + P_{\nu_2})^2 = m_W^2,\end{aligned} \tag{6.41}$$

where index 1 refers to t, 2 to $\bar{t}$ product decays and P are four-momenta of particles respectively.

The top mass, m_t, was measured in Run I to be 174.3 $\pm$ 5.1 GeV [51]. We assume that $m_t = 175$ GeV. With this information the $t\bar{t}$ dilepton decay can be reconstructed.

The system of kinematic equations (6.41) can be written as follows:

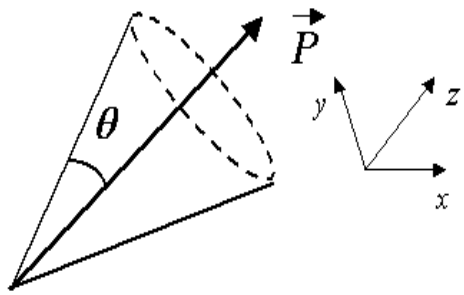

Figure 6.19: The neutrino momentum lies on the cone with the axis along direction of $\vec{P}$ and the open angle θ, such that $\cos\theta = E/|\vec{P}|$.

$$\begin{cases} (m_t^2 - m_W^2 - m_b^2)/2 - P_{\ell_i} \cdot P_{b_i} = P_{\nu_i} \cdot P_{b_i} \\ (m_W^2 - m_\ell^2)/2 = P_{\ell_i} \cdot P_{\nu_i}, i = 1, 2 \end{cases} \tag{6.42}$$

where the left side of the equations is known, based on the observed information. Therefore for each i we can create four-vector $P = (E, \vec{P})$, where E and $\vec{P}$ are energy and momentum respectively,

$$P = \left(\frac{m_t^2 - m_W^2 - m_b^2}{2} - P_\ell \cdot P_b \right) P_\ell - \frac{m_W^2}{2} P_b = (P_\nu \cdot P_b) P_\ell - (P_\ell \cdot P_\nu) P_b \tag{6.43}$$

such that $P \cdot P_\nu = 0$. Thus the four-vector P determines the direction of the neutrino momentum:

$$0 = P \cdot P_\nu = E \cdot E_\nu - \vec{P} \cdot \vec{P_\nu} = E \cdot E_\nu - \vec{P} \cdot (E_\nu \vec{n_\nu}) \tag{6.44}$$

and

$$E = \vec{n}_\nu \vec{P}, \tag{6.45}$$

where $\vec{n}_\nu$ is a unit vector collinear with $\vec{P}_\nu$.

Consequently, the neutrino momentum lies on the cone with the axis along the direction of $\vec{P}$ and the opening angle θ, such that $\cos\theta = E/|\vec{P}|$ (see Figure 6.19). If E happens to be greater than $|\vec{P}|$ then no solution for neutrino exists. Otherwise, the energy of the neutrino can be determined from the second equation of the system 6.42. Neglecting m_ℓ, as $m_\ell \ll m_W$, we have:

$$\frac{m_W^2}{4E_\ell} = E_\nu \sin^2 \frac{\theta_{\nu\ell}}{2}, \tag{6.46}$$

where E_ℓ is the lepton energy and $\theta_{\nu\ell}$ is the angle between the lepton and the neutrino momenta.

It follows that the phase space of possible neutrino momenta can be described by a single parameter $\psi, 0 < \psi < 2\pi$. Let us introduce the orthogonal system of coordinates with axis z along $\vec{P}$ and axis y perpendicular to $(\vec{P}, \vec{P}_\ell)$ plane, so that

$$\vec{n}_x = \frac{\vec{n}_P - \vec{n}_\ell(\vec{n}_P \cdot \vec{n}_\ell)}{|\vec{n}_\ell \times \vec{n}_P|} \tag{6.47}$$

$$\vec{n}_y = \frac{\vec{n}_P \times \vec{n}_\ell}{|\vec{n}_P \times \vec{n}_\ell|} \tag{6.48}$$

and

$$\vec{n}_z = \vec{n}_P. \tag{6.49}$$

Here $\vec{n}_\ell$ and $\vec{n}_P$ are unit vectors along direction of the lepton momentum and vector $\vec{P}$ respectively, i.e. $\vec{n}_\ell = \vec{P}_\ell/|\vec{P}_\ell|$ and $\vec{n}_P = \vec{P}/|\vec{P}|$.

Then we get for $P_\nu = (E_\nu, E_\nu\vec{n}_\nu)$:

$$\vec{n}_\nu = \frac{\sin\theta}{|\vec{n}_P \times \vec{n}_\ell|}\left\{ (\vec{n}_\ell - \vec{n}_P(\vec{n}_P \cdot \vec{n}_\ell))\cos\psi + (\vec{n}_P \times \vec{n}_\ell)\sin\psi \right\} + \cos\theta \cdot \vec{n}_P, \tag{6.50}$$

and

$$E_\nu = \frac{m_W^2}{2E_\ell}\frac{1}{1-\vec{n}_\ell\cdot\vec{n}_\nu}. \tag{6.51}$$

The sum of the two neutrino transverse momenta should match the observed vector of missing transverse energy $\vec{\not{E}}_{\mathrm{T}}$:

$$\vec{P}_{T,\nu_1}(\psi_1) + \vec{P}_{T,\nu_2}(\psi_2) = \vec{\not{E}}_{\mathrm{T}} \tag{6.52}$$

This sets a constraint $f(\psi_1,\psi_2) = 0$ on possible values of ψ_1 and ψ_2, which is challenging to describe analytically. However, note that due to various detector effects such as non-uniformity and poor jet energy resolution, the equality 6.52 is only approximate.

We can therefore introduce a variable which determines how well an event topology satisfies the top dilepton kinematics based on the difference between predicted and measured values of $\not{E}_{\mathrm{T}}$:

$$T = \sqrt{\int_0^{2\pi} d\psi_1 \int_0^{2\pi} d\psi_2 \quad \lambda(\psi_1,\psi_2)}, \tag{6.53}$$

where

$$\lambda(\psi_1,\psi_2) = \exp\left\{-\frac{\left(\vec{P}_{T,\nu_1}(\psi_1) + \vec{P}_{T,\nu_2}(\psi_2) - \vec{\not{E}}_{\mathrm{T}}^{\,measured}\right)^2}{2\sigma^2_{\not{E}_T}}\right\} \tag{6.54}$$

Here the integral is taken over ψ_1 and ψ_2 putting aside the constraint (Equation 6.52). $\sigma_{\not{E}_T}$ represents the uncertainty due to the mismeasurement of the unclustered energy in the event. We take the average error for each component of $\vec{\not{E}}_{\mathrm{T}}$ to be 4 GeV (see Figure 4.8). This procedure is called the neutrino weighting technique [52].

The uncertainty caused by mismeasurements of the jets and leptons in the event is taken into account by using a Monte Carlo technique. The measured jet and lepton momenta are varied randomly within their detector resolutions. Electron energies are varied with σ_E given by Equations 3.11 and 3.13. Muon momenta are varied with σ_P given by Equation 3.9. In all cases the detector resolution is assumed to be gaussian.

Jet resolution is the most difficult to model. It depends on both the resolution of the hadronic calorimeter and uncertainties in jet-parton matching. For instance, two partons with small separation in $\eta - \phi$ space may be reconstructed as a single jet, while a single parton resulting in a broad shower of hadrons may be reconstructed as two jets. We use the jet resolution determined by the jet correction systematic uncertainty (see Section 4.3). To take into account possible mismatches between the jet and the parton, we allow jet energies to fluctuate with double σ.

Due to initial and final state radiation (see Section 5.3.1) top dilepton events often have more than two jets. If this happens we select two leading jets as b-jets and ignore others in the fit. We observe that in 80% of Monte Carlo $t\bar{t}$ events passing the dilepton event selection criteria, two leading jets are associated with the b-partons.

In addition, there is an ambiguity in lepton-jet pairings. We consider both possibilities and sum contributions from each pairing when calculating $\lambda(\psi_1, \psi_2)$.

For each random drawing of lepton and jet energies the quantities $\vec{P}_{T,\nu_1}(\psi_1)$ and $\vec{P}_{T,\nu_2}(\psi_2)$ are recomputed and the integral 6.53 is re-evaluated. This procedure is repeated many times in order to approximate the value of T. We determine that smearing repeated 1,000 times for each event is sufficient to avoid statistical fluctuations in estimating the value of T by this Monte Carlo technique.

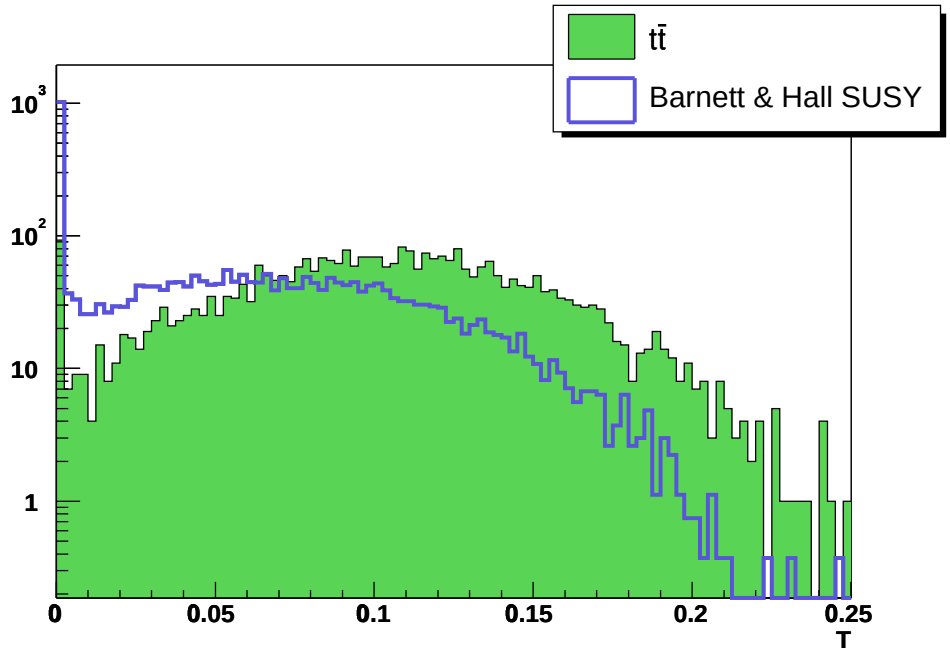

Figure 6.20: $t\bar{t}$ and Barnett/Hall SUSY distributions of variable T.

A non-$t\bar{t}$ event can also be reconstructed by this algorithm. Obviously, it is expected to give on average a smaller value of T. Figure 6.20 shows distributions of T for $t\bar{t}$ Monte Carlo events and events from the Barnett and Hall SUSY sample. When this event reconstruction algorithm is applied to Monte Carlo sample of SUSY events, no solutions are found for approximately 26% events resulting in $T \equiv 0$. Similarly, when this algorithm is applied to $t\bar{t}$ sample, no solutions are found for 3% of $t\bar{t}$ events. These events are mostly dominated by contributions from other $t\bar{t}$ decay channels resembling the dilepton signature.

Distributions of T are well approximated with gaussians, as can be seen from Figure 6.21, with an additional bin at $T = 0$ due to poorly reconstructed events. For non-$t\bar{t}$ events, this gaussian is shifted to lower values of T and has a larger fraction of non-reconstructed events. Figure 6.22 shows the T distribution for WW events.

The topological variable T distinguishes $t\bar{t}$ from Barnett/Hall SUSY events very well. The KS test between those distributions results in a value $< 10^{-324}$ making it competitive with $\not{E}_{\mathrm{T}}$ (see Table 6.5). Unlike other variables, T effectively combines the entire information about an event and quantifies how consistent each event is with the $t\bar{t}$ dilepton topology.

Adding this variable to the already selected three we gain in discrimination power of the test when comparing $t\bar{t}$ with the Barnett/Hall SUSY distribution. Performing a similar test to discussed in Section 6.3.5 with pseudoexperiments of $N = 30$ events we get $\pi = 0.60$ for the four variables in question.

Summarizing, we choose four variables for our analysis, where we anticipate new physics signal can most likely be revealed:

1) the transverse missing energy in the event $\not{E}_T$;

2) the P_T of the most energetic lepton;

3) the angle between these quantities in the transverse plane $\Delta\phi_{\ell m1}$;

4) the topological weight T (Equation 6.53) of the event.

6.3.7 Isolating Unlikely Events in a Multi-Variate Test

As discussed in the previous section we concentrate our search on events with large values of $\not{E}_{\mathrm{T}}$, P_{T,ℓ_1} and $\Delta\phi_{\ell m1}$ and small values of T. We construct K-subsets as in the one-dimensional case starting from events populating the potentially interesting tails of the kinematic distributions, i.e. we use Δ_+ statistic for $\not{E}_{\mathrm{T}}$, P_{T,ℓ_1} and $\Delta\phi_{\ell m1}$ and Δ_- for T.

Events are selected based on the combined information from each of the four kinematic variables describing an event. The information is combined by multi-

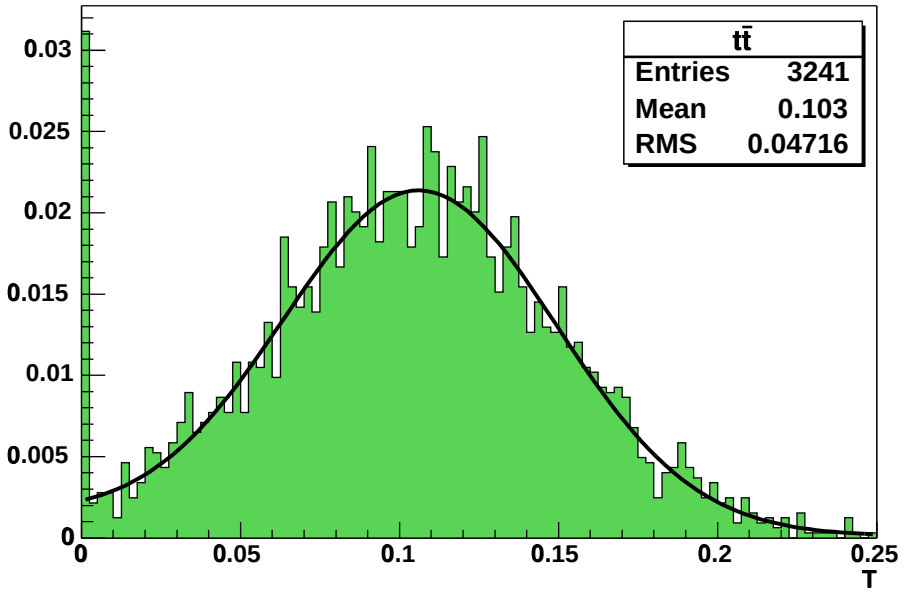

Figure 6.21: Distribution of T for $t\bar{t}$ events. T can be well approximated with a gaussian with an additional bin at $T \simeq 0$ for poorly reconstructed events.

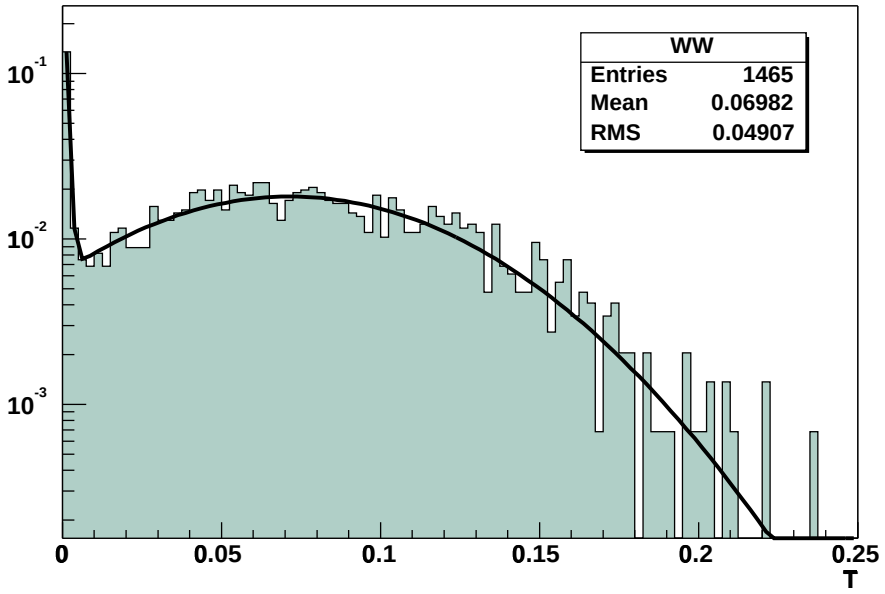

Figure 6.22: Distribution of T for WW events.

plying probabilities obtained from one-variate Δ_+ and Δ_- tests. By doing this, we effectively assign a weight to each event

$$W = w_{\not{E}_T} \cdot w_{P_{T,\ell_1}} \cdot w_{\Delta\phi_{\ell m1}} \cdot w_T, \tag{6.55}$$

where for Δ_+ - variables we integrate $H0$ distribution $f(\vec{x})$ to the right of the observed value $x_{i,event}$

$$w_i = \int_{x_{i,event}}^{\infty} f(\vec{x}) dx_i \tag{6.56}$$

and for Δ_- - variables we perform integration to the left of the observed value $x_{i,event}$

$$w_i = \int_{0}^{x_{i,event}} f(\vec{x}) dx_i \tag{6.57}$$

Thus $w_{\not{E}_T}$, $w_{P_{T,\ell_1}}$, $w_{\Delta\phi_{\ell m1}}$ and w_T represent probabilities for an event to have $\not{E}_T$, P_{T,ℓ_1} and $\Delta\phi_{\ell m1}$ larger than observed and a T smaller than observed respectively.

K-subsets are then constructed from events with the lowest weights W (Equation 6.55), such that the first subset ($K = 1$) contains the event with the lowest weight W, the second subset ($K = 2$) contains events with the two lowest weights, and so on.

After selecting K-subsets the procedure of quantifying their departure from the Standard Model is similar to what was done in the single-variate case. The KS test is performed for each of the four variables i, and $\Delta_{K,i}$ between the SM cumulative function and the cumulative function of the K-subset is computed. To assess the probability of a deviation we generate 100,000 pseudoexperiments by randomly drawing events from large Monte Carlo samples of $t\bar{t}$ and SM backgrounds. The

number of events corresponding to each SM process is a Poisson-smeared estimate based on its kinematic and geometrical acceptance. Only pseudoexperiments with a total number of $N = 13$ events, equal to the number of observed events, are accepted.

Further, in each pseudoexperiment, K-subsets are formed and the respective $\Delta_{K,i}$ for each are calculated. We thus build probability distribution functions $\mathcal{W}_{K,i}(\Delta_{K,i})$ such that the probability of observing a $\Delta_{K,i}$ greater than or equal to that observed in the data is

$$p_{K,i} = \int_{\Delta_{K,i}^{obs}}^{\infty} \mathcal{W}_{K,i}(\Delta)\, d\Delta \tag{6.58}$$

Next we calculate the product Π_K of the four $p_{K,i}$'s for each pseudoexperiment and form the probability distribution functions $\mathcal{M}_K(\Pi_K)$ such that the quantity

$$P_K = \int_0^{\Pi_K^{obs}} \mathcal{M}_K(\Pi)\, d\Pi \tag{6.59}$$

determines how well each K-subset agrees with the SM expectation based on the combined information from the four variables.

Similarly to the one-variate case, we define Q as the value of K with the smallest P_K and use the quantity $\mathcal{P} = P_Q$ as the test statistic to quantify the discrepancy of the data with the SM. Generating another set of 100,000 pseudo-experiments from SM Monte Carlo and repeating the above procedure, we determine $\mathcal{P}$ for each pseudoexperiment and build the probability distribution function $\mathcal{L}(\mathcal{P})$ such that the significance of departure of the Q-subset of events from the SM is

$$\alpha = \int_0^{\mathcal{P}_{data}} \mathcal{L}(\mathcal{P})\, d\mathcal{P} \tag{6.60}$$

α represents the probability to obtain a data sample less consistent with the SM than what is actually observed. Sufficiently low values of α would indicate the

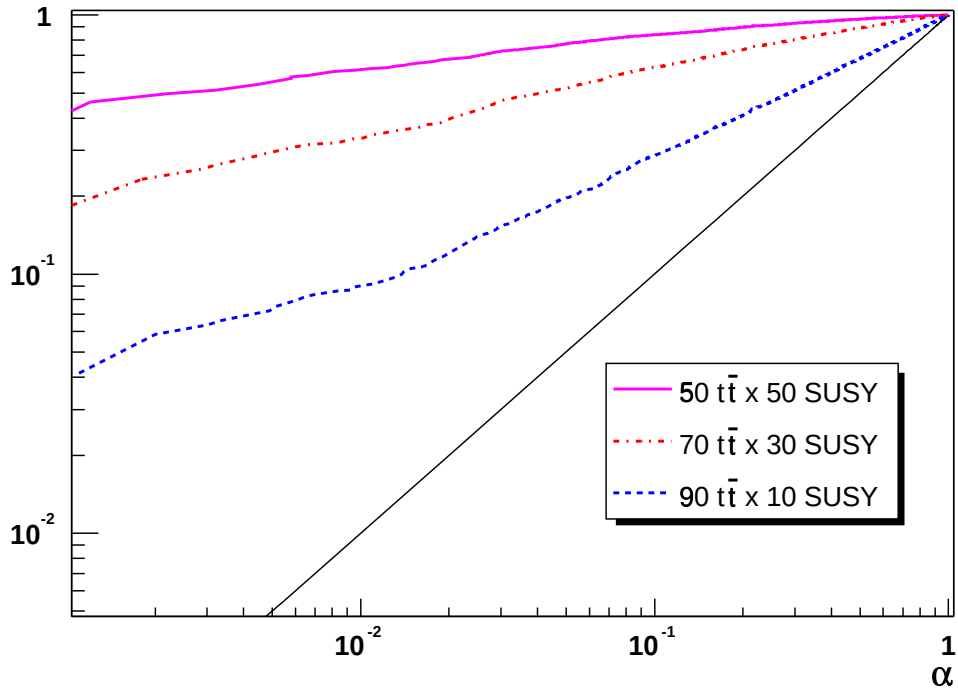

Figure 6.23: Power functions $\pi(\alpha)$ of the test for different mixtures of $t\bar{t}$ and SUSY.

presence of new physics in the data sample, and the Q events isolated would represent the subsample of the data events revealing the largest discrepancy from the Standard Model, i.e. the subset of events with the largest concentration of new physics.

We tested the performance of this method by comparing the $t\bar{t}$ distribution with the 50%:50% mixture of $t\bar{t}$ and Barnett/Hall SUSY using four selected kinematic variables. In this test we used pseudoexperiments of size $N = 13$ equal to the number of observed events in the Run II top dilepton data sample. We observe that with this technique a pseudoexperiment with a new physics contribution is found inconsistent with $\alpha < 1\%$ in 60% of the cases. In addition, the method isolates a most unlikely subset of events with on average 80% of SUSY

events, thus enhancing the new physics concentration in the unlikely subsets. In contrast, the KS-test finds $\alpha < 1\%$ only 28% of the time with no isolation of a mostly-SUSY subset.

The power functions $\pi(\alpha)$ for this technique with different mixtures of $t\bar{t}$ and SUSY are given in Figure 6.23.

6.3.8 Summary

In this Chapter we developed a novel search for new physics technique based on the comparison of kinematic features of observed events with those expected from the Standard Model. This search is designed to be sensitive to any physics that gives rise to events with kinematics different from the Standard Model processes, especially those which result in anomalies similar to the Run I top dilepton events.

The method seeks to isolate the subset of events in a data sample with the largest concentration of possible non-Standard Model events and to assign a probability that quantifies its departure from the Standard Model.

The section 6.3.7 describes the method in the form it was applied to the current Run II data. In the next Chapter we discuss the results of this search.

Chapter 7

Data Analysis and Results

For any analysis based on study of kinematics of events it is imperative that the simulation of kinematic quantities be accurate. Therefore we test the method described in the previous Chapter as well as our ability to correctly simulate kinematic variables on a control sample from data. For this purpose we use the W + $\geq$ 3 jets data set with leptonic W decays. It is a reltively high-statistics sample of events with leptons, $\not\!\!E_T$ and jets.

After verifying that the obtained results are reasonable, we open the 'blind box' and apply the search technique to the Run II top dilepton data set.

The last section of this Chapter is dedicated to an estimate of systematic uncertainties of the result.

7.1 W + jets Data Set, as a Control Sample of Events.

W + jets events are characterized by a high-momentum lepton, a substantial amount of missing energy due to the leptonic W decay, and several high-energy

jets produced in association with the W boson.

The main backgrounds contributing to the sample of W + $\geq$ 3 jets events are from:

- multi-jet QCD events,

  in which a lepton is faked by a jet, and jet energy mismeasurements cause an imbalance in the measured transverse energy leading to a large fake $\not\!\!E_{\mathrm{T}}$.

- $t\bar{t}$ production, $p\bar{p} \to t\bar{t} \to \ell\bar{\nu}_\ell q\bar{q}' b\bar{b}$,

  in which one W decays leptonically and the other W decays hadronically. The $t\bar{t}$ processes dominantly result in $\geq$ 4 jets per an event, where two jets are from the b-quarks and the other two are light quark jets from the hadronically decaying W.

Other background sources are from $W \to \tau\nu$ + 3 partons, $Z \to \ell\ell$ + 2 partons, WW/WZ + 1 parton and single top production.

For the event selection we utilize the primary datasets requiring at least one central high-momentum electron or muon. These are the same datasets which are used for extracting the $t\bar{t}$ dilepton event candidates described in Section 5.1. The total sample size is 193.5$\pm$11.4 pb^{-1} for CEM and CMUP triggers and 175.3$\pm$10.3 pb^{-1} for CMX triggers.

Event selection is performed by requiring one and only one CEM electron, CMUP muon or CMX muon with identification requirements specified in Sections 4.1.1 and 4.2. In addition to relative isolation $E_T^{iso}/E_T^{electron} < 0.1$ and $E_T^{iso}/P_T^{muon} < 0.1$, we apply an absolute isolation cut on energy inside of cone 0.4 around the lepton: $E_T^{iso} < 4$ GeV. This cut helps to reduce the QCD background due to very high-P_T fake leptons, which can still pass the relative isolation

requirement. Similar to the dilepton event selection, we remove events with an electron identified as a conversion and muon events containing a cosmic ray.

Events consistent with a leptonic Z decay are vetoed. This is done by removing events with the primary lepton and a second lepton of opposite charge forming an invariant mass within $76 < m_{\ell\ell} < 106$ GeV. The second lepton is selected with the loose identification requirements listed in Table 7.1.

Jet selection criteria are applied exactly as required in the top dilepton selection (see Section 4.3), thus after jet corrections jets are required to have $E_T > 15$ GeV and $|\eta| < 2.0$. Missing energy ($\not\!\!E_{\mathrm{T}}$) is corrected for jets and muons (see Section 4.4), and in accordance with the $t\bar{t}$ dileptons selection is required to be greater than 25 GeV.

In order to ensure that the jets and lepton are reconstructed from the same interaction we require that the event z-vertex position (see Section 4.4) is within 5 cm of the lepton COT track z_0.

As fake leptons tend to be non-isolated, we model the kinematics of QCD multi-jet background using data events that pass all of the selection requirements except lepton isolation > 0.2. Since $\not\!\!E_{\mathrm{T}}$ in these events results from jet energy mismeasurements, in order to reduce this background we require that the $\not\!\!E_{\mathrm{T}}$ is not parallel or anti-parallel to the direction of the highest E_T jet in the event. Specifically, the azimuthal angle $\Delta\phi_{jm1}$ between the leading jet in an event and $\not\!\!E_{\mathrm{T}}$ is required to be within [0.5,2.5] radians for events with $\not\!\!E_{\mathrm{T}} < 35$ GeV. This cut reduces QCD background by about 50%, while removes only 10% of W + jets events and 7% of $t\bar{t}$ events (see Figures 7.1 and 7.2).

With these selection criteria we observe 973 events in the data. To estimate the QCD background contamination in these events we remove the $\not\!\!E_{\mathrm{T}}$ and lepton isolation requirement. The lepton isolation versus $\not\!\!E_{\mathrm{T}}$ distribution of the data

Variables	Cut values
Electrons	
E_T	> 10 GeV
E_{had}/E_{em}	< 0.12
Isolation	< 0.15
Muons with a track segment in the muon chambers	
P_T	> 10 GeV
E_{em}	< 5 GeV
E_{had}	< 10 GeV
$E_{em} + E_{had}$	< 10 GeV
$\|\Delta x\|$	< 10.0 cm
$\|d_0\|$	< 0.5 cm
Isolation	< 0.15
Muons without a track segment in the muon chambers	
P_T	> 10 GeV
E_{em}	< 2 GeV
E_{had}	< 6 GeV
$E_{em} + E_{had}$	< 10 GeV
$\|d_0\|$	< 0.5 cm
Isolation	< 0.15

Table 7.1: Identification requirements for the second lepton used in the Z boson veto.

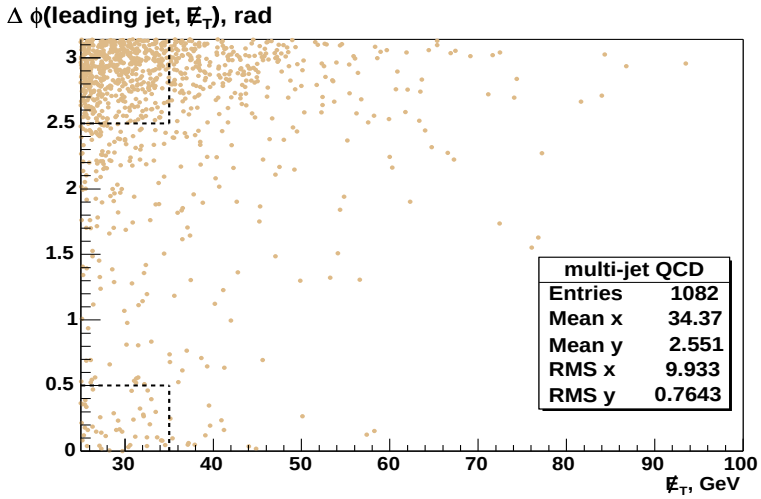

Figure 7.1: The azimuthal angle between the $\not\!\!E_T$ and the leading jet versus the $\not\!\!E_T$ of multi-jet QCD data with a non-isolated electron or muon candidate. The dashed line indicates the cut in $\Delta\phi_{jm1}$-$\not\!\!E_T$ plane, which reduces the QCD background by 50%.

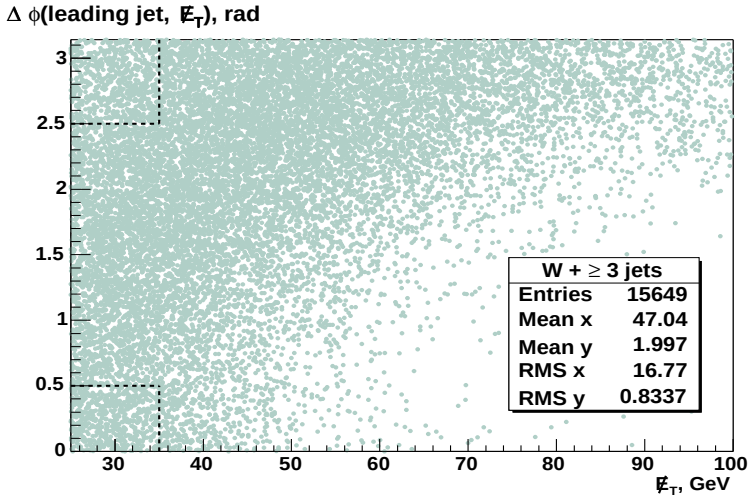

Figure 7.2: The azimuthal angle between the $\not\!\!E_T$ and the leading jet versus the $\not\!\!E_T$ of ALPGEN+ HERWIG Monte Carlo W + 3 partons events. The dashed line cut in $\Delta\phi_{jm1}$-$\not\!\!E_T$ plane removes approximately 10% of W + jets events and 7% of $t\bar{t}$ events.

events is shown in Figure 7.3. For events with $\not\!\!E_{\mathrm{T}} < 35$ GeV, we require $\Delta\phi_{jm1}$ cut between the leading jet and $\not\!\!E_{\mathrm{T}}$.

We then divide the 2D-plane into four regions:

- A: lepton isolation > 0.2 and $\not\!\!E_{\mathrm{T}} < 10$ GeV
- B: lepton isolation < 0.1 and $\not\!\!E_{\mathrm{T}} < 10$ GeV
- C: lepton isolation > 0.2 and $\not\!\!E_{\mathrm{T}} > 25$ GeV
- D: lepton isolation < 0.1 and $\not\!\!E_{\mathrm{T}} > 25$ GeV

The region D is a signal region. The QCD events are expected to have no correlations between isolation and missing transverse energy $\not\!\!E_{\mathrm{T}}$. Therefore the QCD contribution to the signal region is given by

$$N_{QCD} = N_B * N_C / N_A$$

With $N_A = 3083, N_B = 685$ and $N_C = 389$ estimated from the data, we obtain that QCD contribution is 8.8% of the data events in the signal region. The systematic uncertainty on this estimate is determined by varying borders of the four regions. We assign 75% uncertainty to the background normalization and thus obtain that the QCD accounts for $8.8\% \pm 6.6\%$ events.

The $t\bar{t}$ contribution is estimated based on its theoretical cross section 6.7 ± 0.8 pb [8] and is found to be $11.2\% \pm 1.3\%$. $t\bar{t}$ distributions are modeled using PYTHIA, assuming $m_t = 175$ GeV. The other 80% of events are mainly W + jets and are modeled with W + 3p ALPGEN Monte Carlo.

$\not\!\!E_{\mathrm{T}}$, leading lepton P_T, and the azimuthal angle $\Delta\phi_{\ell m1}$ between the directions of these quantities for data versus Monte Carlo modeled distributions are

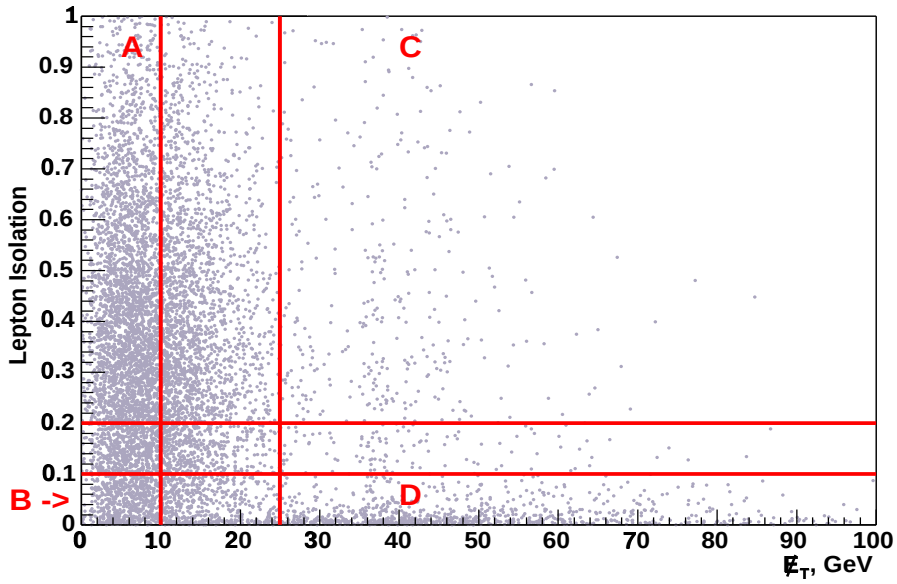

Figure 7.3: Lepton Isolation versus $\not\!\!E_{\mathrm{T}}$ distributions for events with a lepton and ≥ 3 jets. For events with $\not\!\!E_{\mathrm{T}} < 35$ GeV, $\Delta\phi_{jm1}$ cut is applied.

presented in Figures 7.4 and 7.5. To quantify consistency of the simulation with observations we make use of the KS test. To check whether these variables are sensitive to the $t\bar{t}$ signal, we perform the KS test, with $t\bar{t}$ and without $t\bar{t}$ distributions included into simulation. In the latter scenario Monte Carlo distributions consistent of 8.8% QCD and the rest is W+3p simulated events.

The results of the KS test are given in Table 7.2. Overall we observe better consistency between the distributions when the $t\bar{t}$ signal is included. However, no significant discrepancies are observed without the $t\bar{t}$. We conclude that these kinematic variables are not optimized for detecting $t\bar{t}$ events, which differ from W + jets mostly by higher jet multiplicity and harder jet E_T spectra. Therefore, by using our search technique with our chosen kinematic variables, we should not observe anomalously low values of probabilities quantifying consistency of the data with the simulation.

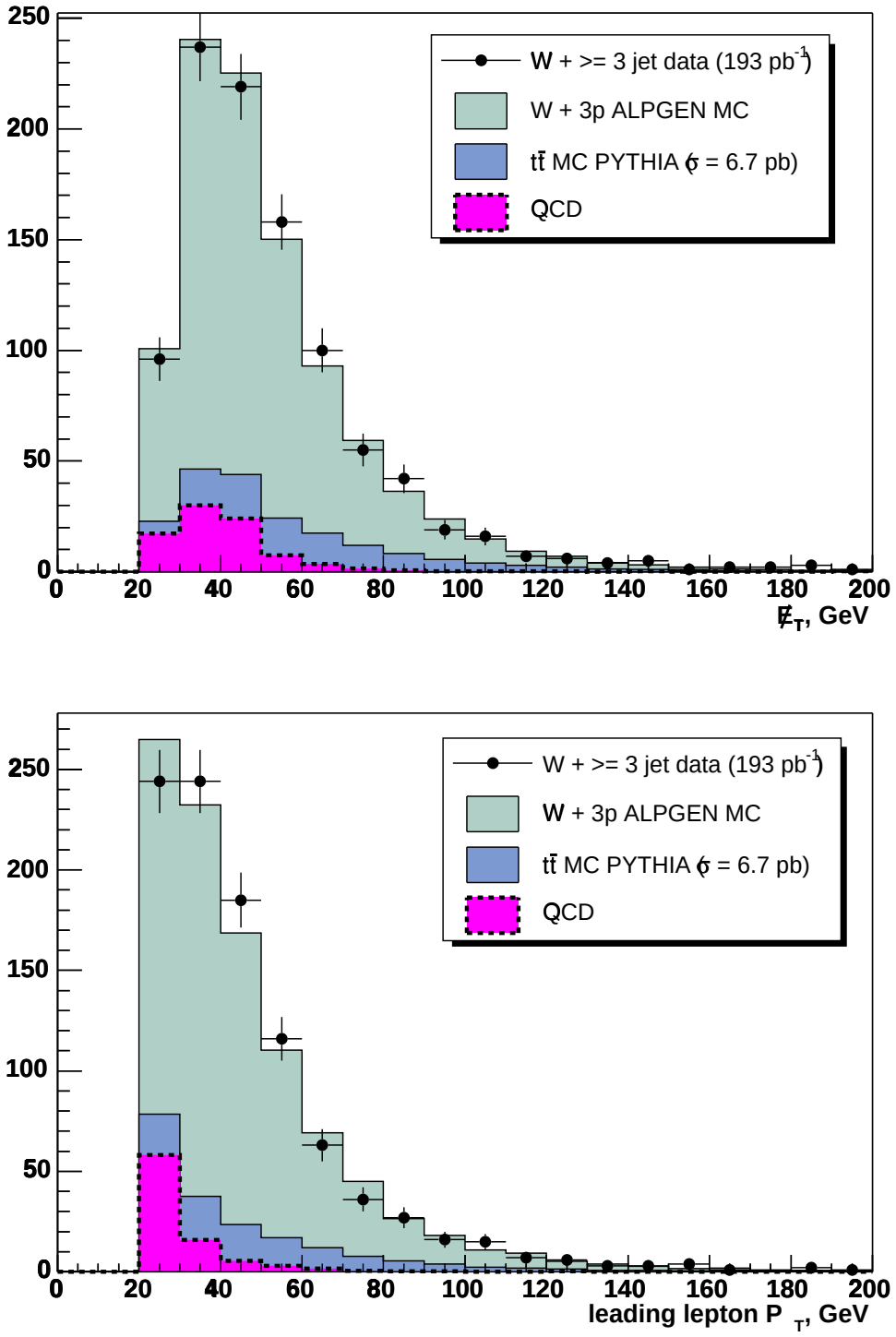

Figure 7.4: $\not\!\!E_{\rm T}$ and leading lepton P_{T,ℓ_1} distributions in W + 3 jets sample.

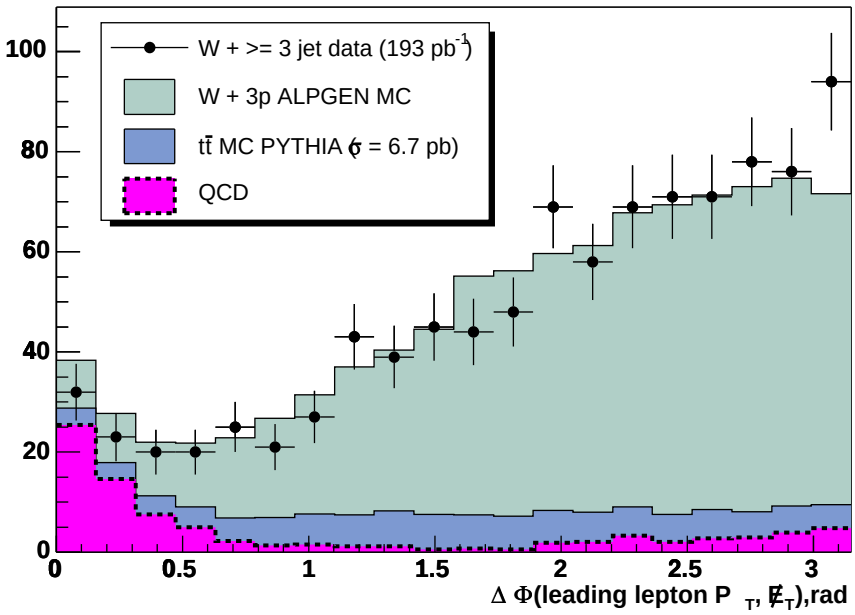

Figure 7.5: $\Delta\Phi_{\ell m1}$ distribution in W + 3 jets sample.

Variables	with $t\bar{t}$	without $t\bar{t}$
$\not{E}_{\mathrm{T}}$	75%	4%
$P_{T,\ell 1}$	18%	9%
$\Delta\phi_{\ell m1}$	6%	36%

Table 7.2: Results of the KS test applied to kinematic distributions.

We conducted several tests on the performance of our search technique in this sample. Since the variable T quantifying the topology of the $t\bar{t}$ dilepton events is not applicable to the W +jets analysis, and $\Delta\phi_{\ell m1}$ is not good in search for top because in W events the lepton and $\not{E}_{\mathrm{T}}$ tend to be back to back, we perform the PKS test on only two variables: $\not{E}_{\mathrm{T}}$ and $P_{T,\ell 1}$ of the leading lepton. We apply the technique to Monte Carlo 2D-distributions, both with and without $t\bar{t}$ signal.

We expect the following results from the PKS test applied to the W+jets sample:

1. When $t\bar{t}$ signal is included, the data should give a result indicating good consistency with the SM.

2. When $t\bar{t}$ is *not* included, the data should give a result that is less consistent with the "SM". However, we do not expect the result to be *significantly* inconsistent, i.e. to "discover top".

The pseuodexperiments in the PKS test are constructed from events with Poisson smeared event yields for the different sources.

The test of consistency of the data sample with only W + jets + QCD background finds $Q = 936$ events in a most unlikely subset, which constitutes almost the entire data sample, and sets $\alpha = 10.9\%$ probability of consistency of the data with the null hypothesis distributions. Thus the test does not find any significant deviations in the tails of the tested distributions.

The second test is performed with $t\bar{t}$ signal added to the null-hypothesis. In this case the test is effectively set to find events beyond the Standard Model. The minimum of P_K is achieved at $Q = 3$ and α was found to be 36.2%, i.e. consistent with the $H0$-distribution. That means that a less likely subset can be found in

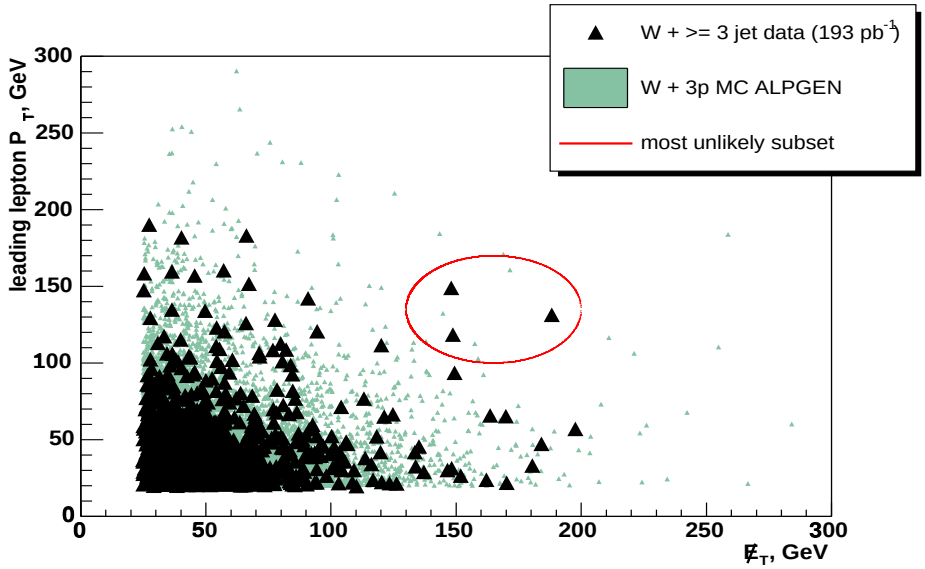

Figure 7.6: The data and MC W + 3p events in $(\not{E}_{\mathrm{T}}, P_{T,\ell 1})$ plane. The PKS test isolates a most unlikely subset of three events in the tails of these distributions with $\alpha = 36.2\%$.

at least 36.2% of $H0$-pseudoexperiments. The data and the three isolated events are shown in the 2D plane in Figure 7.6.

In the next test we add the variable $\Delta\phi_{\ell m1}$. With the $t\bar{t}$ signal included we obtain $\alpha = 35.1\%$, indicating good modelling of the data by the simulation.

The topological variable T is the most tricky, as it is defined only for events with the $t\bar{t}$ dilepton signature. At the same time, it is the most complicated variable of these four, which takes into account many quantities of the final state objects in an event. It is therefore important that the modelling of T in simulated events be accurate.

Since T requires at least six final objects in an event, we model it by using a control sample of W $\geq$ 4 jets. We estimate contributions from QCD and $t\bar{t}$ sources using the method discussed above, and find them to be 9.7% and 29.8%

respectively.

We attempt to reconstruct these events as if they were $t\bar{t}$ dilepton events by treating the leading jet as a second lepton, the subleading jet as a second neutrino, and the two other jets in the event as b-jets. We apply the reconstruction to the data and to an appropriately weighted sample of simulated events, for which we use an ALPGEN W + 4p Monte Carlo sample for simulating W + jets events. The $\not\!\!E_{\mathrm{T}}$ is recomputed in each event.

The distribution of T for the data versus Monte Carlo is given in Figure 7.7 (top) with the excellent agreement of the KS test, giving 97% probability.

In the other version of this test we treat the two leading jets as b-jets and the next two as a second lepton and a second neutrino. Such events better resemble the $t\bar{t}$ dilepton decays, as it is evidenced by larger values of T in Figure 7.7 (bottom). We have also a very good agreement in modelling of the variable T, with the KS agreement of 86%.

In conclusion, we observe that the simulation adequately models the kinematic distributions obtained from data.

7.2 Top Dilepton Sample

Finally, having established that data are adequately modeled by the simulation, we approach the exciting part of the analysis – opening the 'blind box' and applying our search technique to the sample of the Run II top dilepton events.

Top dilepton sample selection and background estimates are described in detail in Chapter 5. In the accumulated luminosity of 193 pb^{-1}, we observe total 13 events while expecting 10.9 ± 1.4, based on the theoretical $t\bar{t}$ cross section and

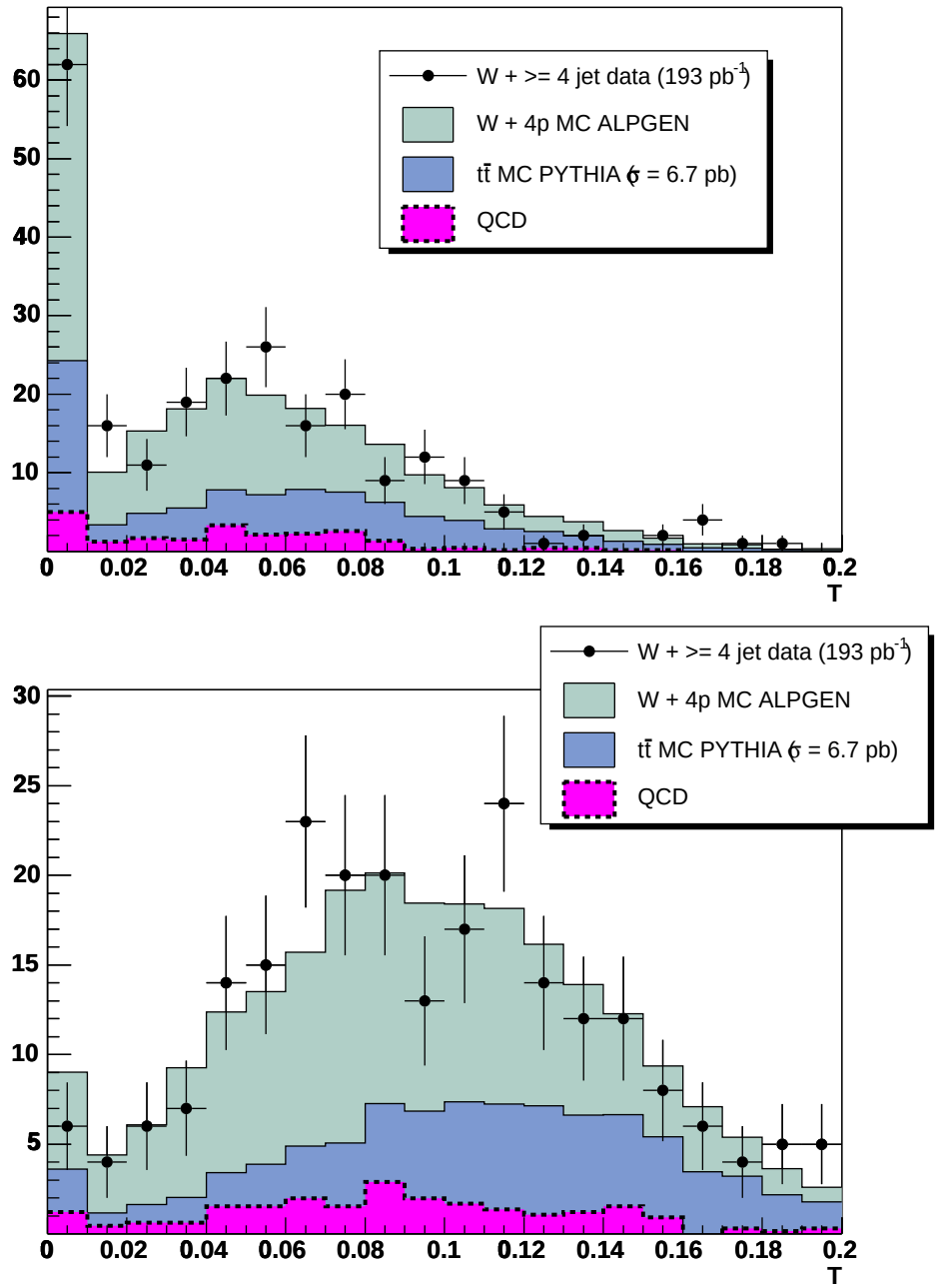

Figure 7.7: T distributions in $W \geq 4$ jets sample of events. Events are reconstructed as if they were $t\bar{t}$ dilepton events by treating the two leading jets as a second lepton and neutrino, and two other jets as b-jets (the top plot); and alternatively by treating two leading jets as b-jets and subleading jets as lepton-neutrino (the bottom plot).

Source	Shape determined from
WW	$WW \to \ell\bar{\ell}$ PYTHIA MC
WZ	inclusive PYTHIA MC
Drell-Yan	$Z/\gamma \to \ell^+\ell^-$ + 2p ALPGEN MC
$Z \to \tau^+\tau^-$	inclusive PYTHIA MC
Fakes	Data
$t\bar{t}$	inclusive PYTHIA MC

Table 7.3: Samples used for modelling the shapes of various Standard Model processes contributing to the $t\bar{t}$ dilepton signature.

taking into account the systematic uncertainties on the top dilepton acceptance and the integrated luminosity.

$t\bar{t}$ and non-top Standard Model processes are simulated with the Monte Carlo generators listed in Table 7.3. These are the same as used for estimating contributions from background sources and added up in their expected contributions (see Table 5.19).

Fake lepton background shapes are estimated from data by applying the fake lepton ratio matrix [1] to fakeable jets in the W+jets dataset. We treat a fakeable jet as a lepton and then recompute all kinematic quantities in the event. To that event it is assigned the sum of weights depending on the probability for the fakeble jets to fake various types of leptons. The fake lepton kinematic shapes are obtained from the weighted events passing the top dilepton event selection criteria.

The kinematic distributions of the four kinematic variables in use are given in Figures 7.8 and 7.9. The $t\bar{t}$ is shown summed with the non-top SM background.

[1] The fake lepton ratio is a function of lepton E_T/P_T and isolation (see Section 5.3.5)

Since the uncertainties in each bin are given by Monte Carlo simulation contents, it is more natural to picture uncertainties associated with Monte Carlo simulation.

Before applying the PKS test to these distributions one can notice that there is no excess of events with large $\not\!\!E_{\rm T}$ and leading lepton P_T distributions as observed in Run I. Simultaneously, there is no excess of events with a leading lepton in the opposite direction to the $\vec{\not\!\!E}_{\rm T}$, i.e. events with large $\Delta\phi_{\ell m1}$, and no excess of events poorly reconstructed under the top dilepton decay, i.e. events with low T. In fact, there is a slight excess of events with large T, implying that those events are likely to have the $t\bar{t}$ origin rather than accounted for by non-top background sources. And in addition, there is an excess of events in the low-P_T region.

Our search technique applied to the Run II top dilepton sample of events finds a most unlikely subset of size $Q = 13$, i.e. the entire sample of events, with a probability of 1.6% corresponding to 2.41 gaussian standard deviations. This result is entirely driven by the excess of leptons at low P_T. As the method effectively orders the subsets from high P_T to low P_T, the value P_K decreases as more of the low P_T excess is included, reaching a minimum when the entire data sample is considered. Thus, no particular subset shows a significant discrepancy and it can be concluded that new physics scenarios invoked to describe the high-P_{T,ℓ_1} and high-$\not\!\!E_{\rm T}$ events observed in Run I are not favored by the current Run II data.

In order to better understand the nature of the low-P_T events, such as whether they can be attributed to non-$t\bar{t}$ SM processes, we have investigated the events in detail. Since top quark events are characterized by the presence of long-lived b-hadron decays in the final state, we identify events with jets containing a displaced secondary vertex as likely top events.

We present the bottom quark content of the sample, as well as the distribution

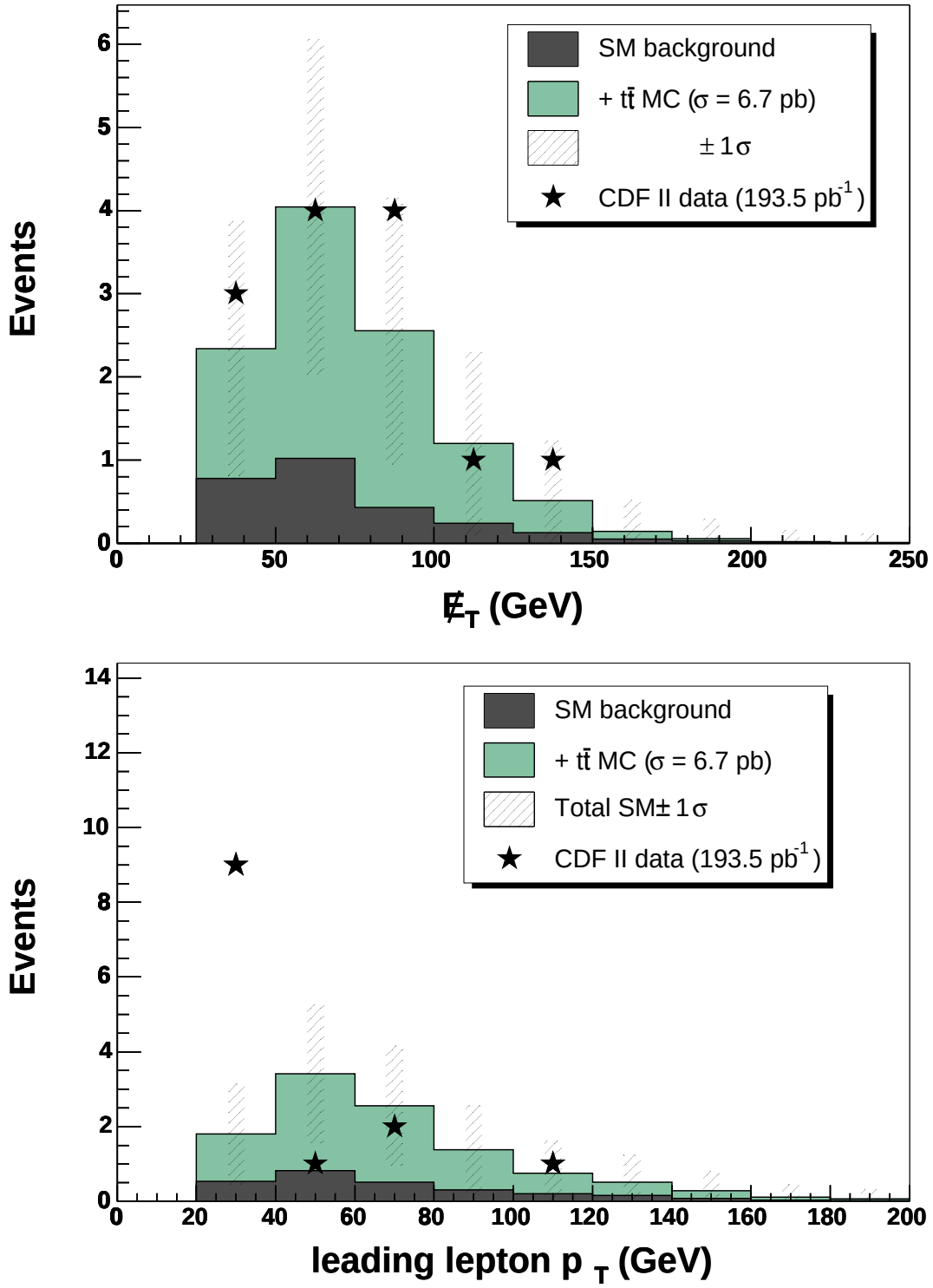

Figure 7.8: $\not\!\!E_{\mathrm{T}}$ and $P_{T,\ell 1}$ distributions in the Run II top dilepton sample. The error bars represent Monte Carlo uncertainties.

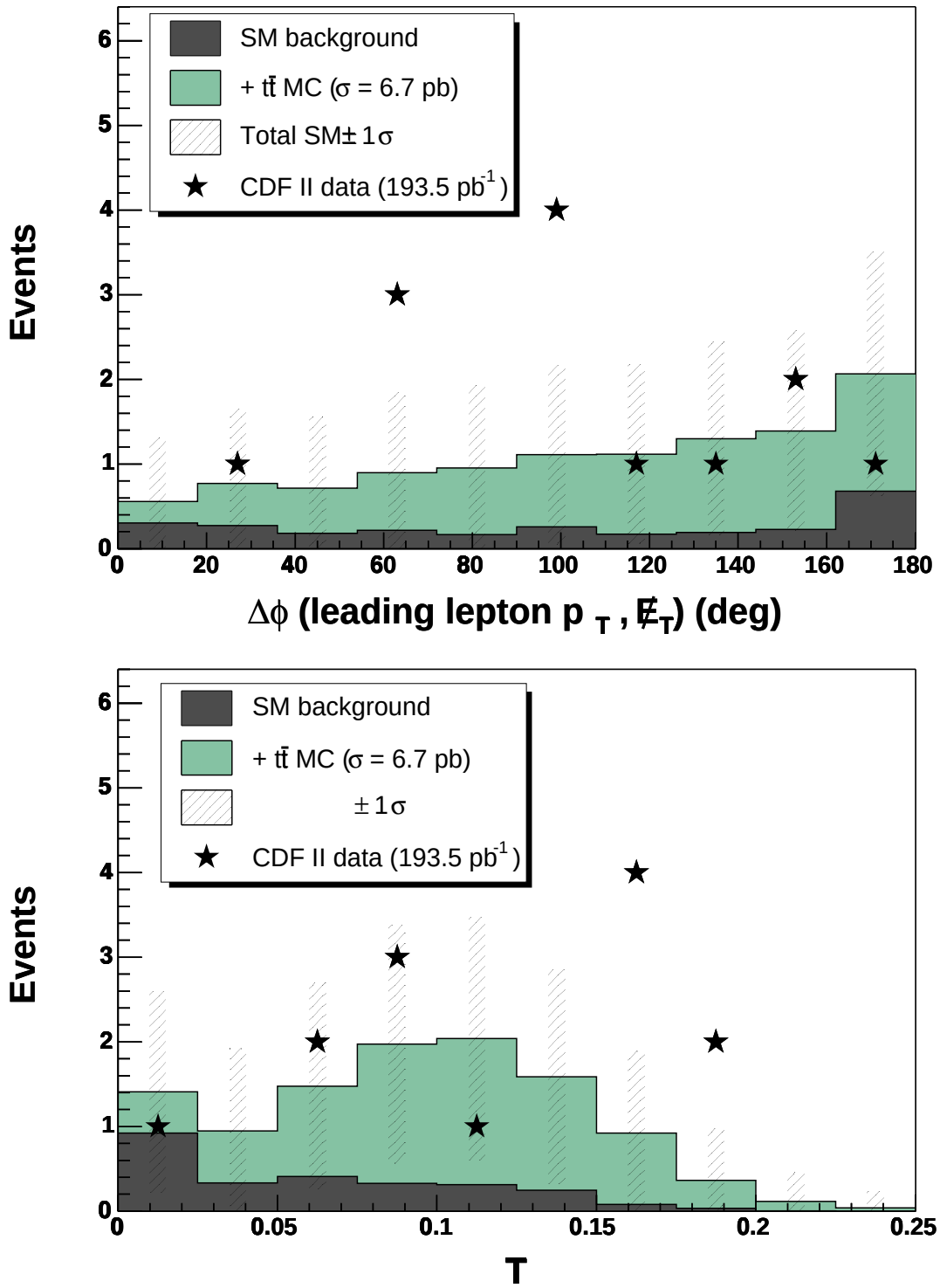

Figure 7.9: $\Delta\phi_{\ell m1}$ and T distributions in the top dilepton sample. The error bars represent Monte Carlo uncertainties.

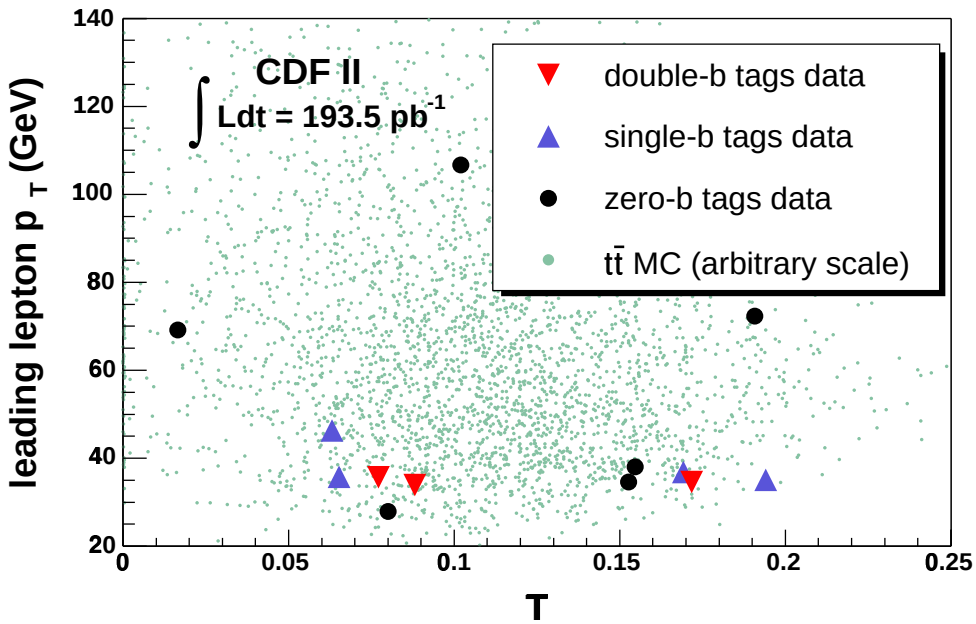

Figure 7.10: Top dilepton events in (P_T, T) plane with b-tagging information.

of events in the (P_T, T) plane, in Figure 7.10. Six of the nine low-P_T (< 40 GeV/c) events contain at least one identified b-jet, and more than half are quite consistent with the $t\bar{t}$ kinematic hypothesis as evidenced by their large values of T. In addition, the leptons do not cluster near the 20 GeV P_T threshold. We also note that the leptons are well-isolated from any additional calorimeter activity, disfavoring the possibility that they arise from semileptonic b decays. We thus conclude that the low-P_T events are more likely to be top than background.

Details of the thirteen $t\bar{t}$ dilepton candidates are listed in Tables 7.4 and 7.5. Distributions of other various kinematic quantities are shown in Figures 7.11 and 7.12.

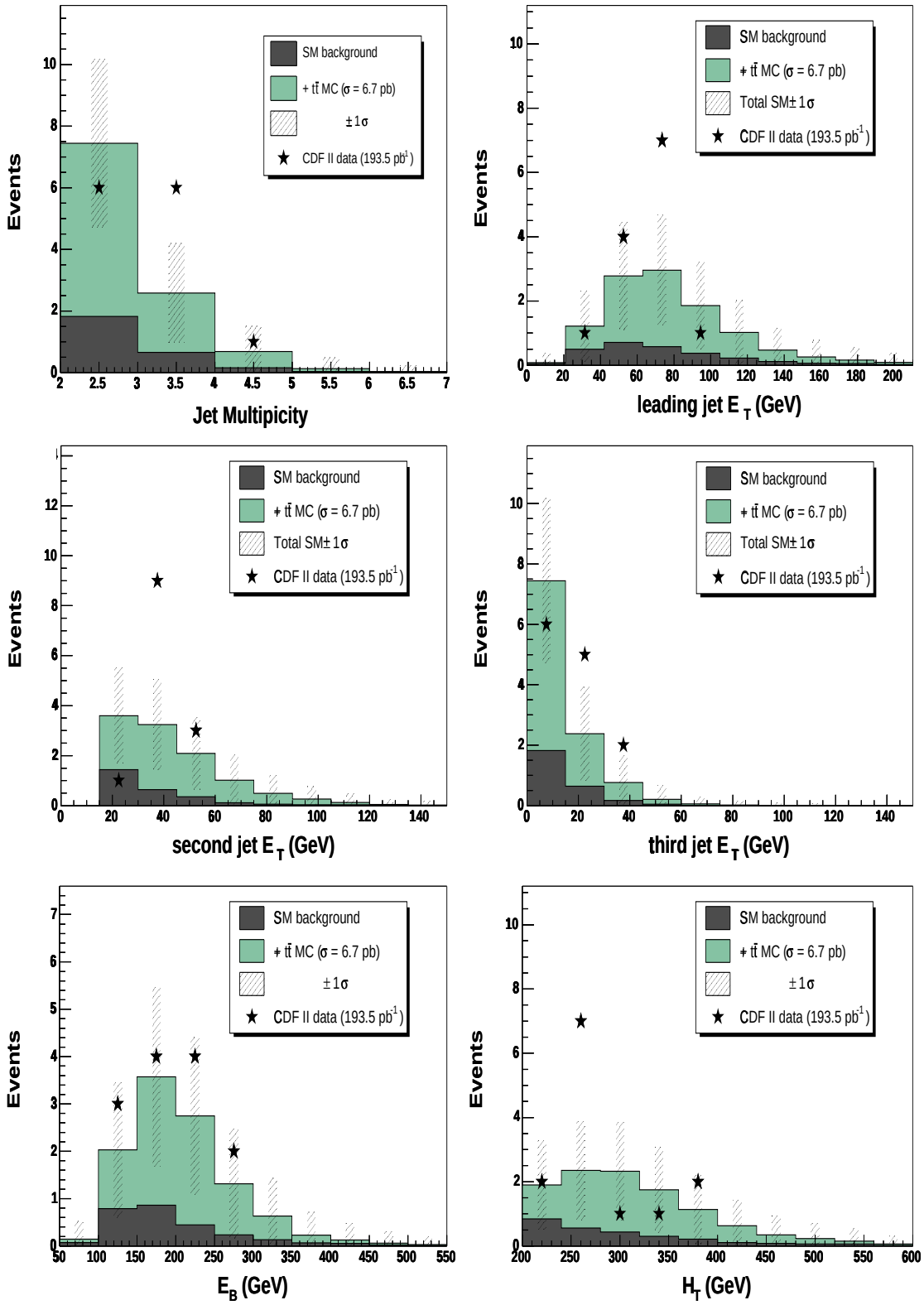

Figure 7.11: Distributions of jet kinematic quantities in the Run II top dilepton sample. $E_B = \not{E}_{\mathrm{T}} + \sum E_{T,j}$. The first bin in the third jet E_T distribution is populated by events with 2 jets.

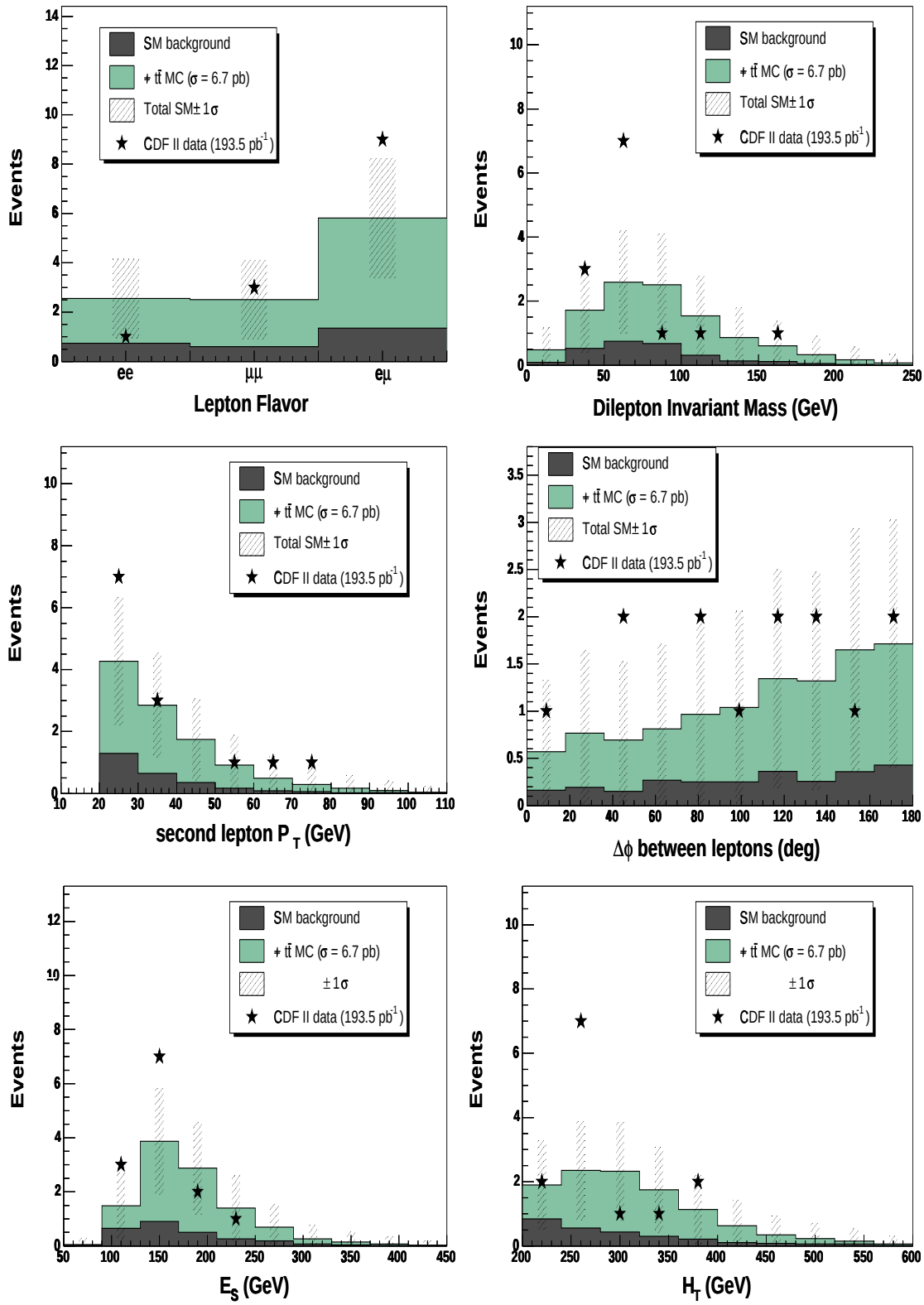

Figure 7.12: Distributions of lepton kinematic quantities in the Run II top dilepton sample. $E_S = \not{E}_{\mathrm{T}} + \sum E_{T,\ell}$. Notice, that the Run II dilepton sample also has a slight excess of $e\mu$ events.

Category Run/Event z_{vertex}	$P_{T,\ell_1}, \eta_{\ell_1}, \phi_{\ell_1}$ $(E, \vec{P})_{\ell_1}$ ℓ_1^{charge}, ISO	$P_{T,\ell_2}, \eta_{\ell_2}, \phi_{\ell_2}$ $(E, \vec{P})_{\ell_2}$ ℓ_2^{charge}, ISO	$E_{j_1}, E_{T,j_1}, \eta_{j_1}, \phi_{j_1}$ $E_{j_2}, E_{T,j_2}, \eta_{j_2}, \phi_{j_2}$ $\vdots$	H_T $\not{E}_T, \phi_{\not{E}_T}$ T
CMX-CMX 153325/599511 4.01 cm	37.5, 0.75, 218° (48.7, -29.6, -23.1, 31.1) μ^- , 0.000	25.2, 0.69, 50° (31.3, 16.1, 19.3, 18.7) μ^+, 0.026	114.3, 52.1, 1.42, 96° 68.6, 35.3, -1.28, 192° 45.2, 18.2, 1.56, 344°	216.7 48.5, 309° 0.169
CEM-CEM 153374/2276742 9.95 cm	72.6, 0.62, 305° (87.3, 41.2, -59.8, 48.5) e^+ , 0.032	56.8, -0.24, 320° (58.4, 43.7, -36.3, -13.5) e^- , 0.008	91.2, 91.2, 0.01, 83° 67.3, 56.7, -0.60, 139° 202.5, 44.2, -2.20, 127°	383.2 61.6, 249° 0.191
CMX-CMUP 165198/1827962 -26.88 cm	35.9, 0.97, 20° (54.1, 33.7, 12.4, 40.5) μ^+ , 0.006	24.2, -0.02, 73° (24.2, 7.2, 23.1, -0.5) μ^- , 0.075	68.0, 66.3, 0.23, 121° 43.0, 33.4, 0.74, 172°	264.5 104.7, 286° 0.077
CEM-CMUP 167629/180103 16.46 cm	107.2, -0.48, 145° (119.6, -87.4, 62.1, -53.0) e^- , 0.009	79.2, -0.26, 17° (82.0, 75.9, 22.8, -21.1) μ^+, 0.009	102.1, 82.9, -0.67, 323° 49.1, 36.9, 0.79, 188°	333.4 27.2, 235° 0.102
CEM-CMX 167631/2058969 25.56 cm	33.8, -0.38, 259° (36.3, -6.2, -33.2, -‘13.3) e^- , 0.002	32.5, -1.02, 125° (50.8, -18.8, 26.5, -39.0) μ^+, 0.006	67.8, 39.2, -1.14, 315° 38.7, 36.6, 0.34, 349°	201.5 59.4, 133° 0.088
NICEM-CMUP 151978/507773 -3.95 cm	36.2, 0.54, 262° (41.6, -5.1, -35.9, 20.3) e^- , 0.139	34.9, -0.32, 156° (36.7, -32.0, 14.0, -11.4) μ^+, 0.004	68.8, 48.8, 0.88, 331° 45.5, 44.7, 0.19, 286°	257.0 92.4, 108° 0.065
CMX-CEM 155114/478702 32.35 cm	34.8, 0.64, 203° (42.3, -31.9, -13.9, 23.9) μ^-, 0.027	28.1, -0.75, 23° (36.5, 25.8, 11.1, -23.2) e^+ , 0.025	88.2, 77.5, -0.52, 333° 72.5, 37.6,-1.28, 255° 30.0, 24.8,-0.64, 227°	295.8 93.0, 103° 0.194
CMX-CEM 156484/3099305 -12.70 cm	35.6, 0.97, 24° (53.8, 32.4, 14.6, 40.4) μ^+, 0.000	34.6, 0.70, 302° (43.5, 18.5, -29.3, 26.3) e^- , 0.009	67.6, 67.1, 0.13, 131° 38.5, 34.6, -0.47, 150° 22.8, 22.6, -0.13, 181°	269.0 74.6, 313° 0.153
CMUP-CMP 162820/7050764 -22.27 cm	34.8, -0.34, 322° (36.8, 27.4, -21.4, -12.1) μ^- , 0.043	25.9, -0.43, 165° (28.3, -25.0, 6.6, -11.5) μ^+ , 0.040	70.6, 61.9, 0.53, 56° 72.3, 55.5, -0.76, 110°	274.6 95.6, 275° 0.172
CMUP-PHX 163064/10576918 -1.75 cm	68.9, -0.08, 143° (69.1, -55.2, 41.2, -5.3) μ^-, 0.007	60.4, -1.33, 94° (122.0, -4.1, 60.2, -106.1) e^+ , 0.012	180.3, 71.7, 1.57, 278° 154.4, 29.6, 2.33, 282°	262.4 31.8, 360° 0.017

Table 7.4: Details of kinematics of the top dilepton candidate events.Energies and momenta are in units of GeV. Jets having b-tags are shown in boxes.

Category Run/Event z_{vertex}	$P_{T,\ell_1}, \eta_{\ell_1}, \phi_{\ell_1}$ $(E, \vec{P})_{\ell_1}$ ℓ_1^{charge}, ISO	$P_{T,\ell_2}, \eta_{\ell_2}, \phi_{\ell_2}$ $(E, \vec{P})_{\ell_2}$ ℓ_2^{charge}, ISO	$E_{j_1}, E_{T,j_1}, \eta_{j_1}, \phi_{j_1}$ $E_{j_2}, E_{T,j_2}, \eta_{j_2}, \phi_{j_2}$ $\vdots$	H_T $\not{E}_T, \phi_{\not{E}_T}$ T
CEM-CMIO 161633/963604 -2.76 cm	38.1, 0.10, 279° (38.2, 6.1, -37.6, 3.6) e^+ , 0.004	26.3, -0.60, 30° (31.1, 22.7, 13.3, -16.7) μ^-, 0.037	66.6, 65.9, -0.14, 119° 95.9, 36.3, 1.63, 81° 40.0, 18.1, 1.43, 326°	249.9 65.2, 256° 0.155
CEM-CMU 165364/592961 -20.09 cm	45.7, 0.62, 337° (54.9, 42.1, -17.6, 30.4) e^- , 0.053	21.4, 0.56, 97° (24.9, -2.6, 21.3, 12.6) μ^+, 0.006	116.2, 80.6, -0.91, 266° 103.3, 49.7, 1.36, 10° $\boxed{49.6, 40.9, -0.64, 45°}$	366.5 128.1, 167° 0.063
CMP-CEM 143257/760520 10.77 cm	27.7, -0.04, 34° (27.7, 22.8, 15., -1.2) μ^-, 0.006	21.1, 0.29, 114° (22.0, -8.5, 19.3, 6.3) e^+ , 0.009	124.0, 44.1, 1.69, 12° 39.3, 37.5, -0.31, 149° 22.6, 22.5, -0.11, 194° 19.6, 15.7, -0.69, 56°	260.3 91.7, 251° 0.080

Table 7.5: Details of kinematics of the top dilepton candidate events (continued). Energies and momenta are in units of GeV. Jets having b-tags are shown in boxes.

7.3 Systematic Uncertainties

Due to systematic uncertainties the shape of the null hypothesis distribution $F(x)$ is not precisely defined. This leads to an uncertainty in the probability distribution function $\mathcal{L}(\mathcal{P})$, which in turn affects the significance level of the measurement.

To estimate the effects of the systematic uncertainties we model deviations in the shape of $\mathcal{L}(\mathcal{P})$ due to various systematic effects causing distortions in the null hypothesis distribution.

First, we consider each source of systematic uncertainty individually. We define the $H0$-distribution used for the top dilepton acceptance estimate and the shape analysis study as a *default distribution*: $F(x) = F_0(x)$. Any template $F(x)$ generated differently from the default conditions is considered as a "systematic" shift of $F_0(x)$. We denote such a distribution as $F_X(x)$.

The template $F_X(x)$ is used for generating pseudoexperiments H_X [2], which are then compared with the default distribution $F_0(x)$, i.e. K-subsets in H_X- pseudoexperiments are formed, probability distributions $\mathcal{W}_{K,i}(\Delta_{K,i})$ and $\mathcal{M}_K(\Pi_K)$ are built and the respective P_K values are computed using the distribution $F_0(x)$. Thus a probability distribution of statistic $\mathcal{L}_X(\mathcal{P})$ is constructed.

The significance of the departure of the data sample from the Standard Model is recomputed based on the distribution $\mathcal{L}_X(\mathcal{P})$ as follows:

$$\alpha_X = \int_0^{\mathcal{P}_{data}} \mathcal{L}_X(\mathcal{P})\, d\mathcal{P}, \tag{7.1}$$

where $\mathcal{P}_{data}$ is the same value used in quoting the central value of $\alpha = \alpha_0$:

$$\alpha_0 = \int_0^{\mathcal{P}_{data}} \mathcal{L}(\mathcal{P})\, d\mathcal{P}, \tag{7.2}$$

The value α_X is defined as a 1σ shift in the measured value of α_0. It is in general larger than α_0, since H_X-pseudoexperiments are on average in worse agreement with the null hypothesis distribution $F_0(x)$ than the H_0-pseudoexperiments. However, some systematic effects leading to a relative shortage of events in the potentially interesting kinematic region, i.e. events with large $\not{E}_{\mathrm{T}}$, leading lepton P_T, $\Delta\phi_{\ell m1}$ and small T, can cause lower values of α_X, as can be seen below.

The systematic uncertainties due to different modelling of the $t\bar{t}$ production and decay are described in Section 5.3.1. In addition to these, we estimate effects due to uncertainty in the top mass value and the non-top Standard Model background contributions.

For each source of systematic uncertainty we build templates $F_X(x)$ as follows:

- **A Monte Carlo event generator.**

[2] We use 10,000 pseudoexperiments of H_0 and H_X for estimating systematic uncertainties.

We use PYTHIA as a default generator for simulation of $t\bar{t}$ processes and use a sample of HERWIG $t\bar{t}$ events as a $F_X(x)$ kinematic template. We get the value $\alpha_{HERWIG} = (1.6 \pm 0.1)\%$. The distributions of kinematic variables for the templates $F_0(x)$ and $F_{HERWIG}(x)$ are shown in Figure 7.13.

- **Radiative corrections uncertainties.**

  Incorrect modeling of radiative jets originating from initial or final state partons will certainly affect the shape of $F(x)$. We make use of the PYTHIA samples with different modelling of initial and final state radiation and the underlying event, which were used in estimating the $t\bar{t}$ dilepton acceptance uncertainty. We determine values of $\alpha_{ISR/FSR}$ to be in the range $1.2 - 1.6\%$. Comparison of distributions $F_0(x)$ and $F_X(x)$ giving the lowest $\alpha_X = 1.2\%$ are shown in Figure 7.14.

- **Uncertainties in the parton distribution structure functions.**

  We use samples generated using the MRST parton distribution functions with different values of the strong coupling constant α_s. These give a range of $\alpha_{PDF} = 1.4 - 1.9\%$. Distributions of $F_X(x)$ giving the highest $\alpha_X = 1.9\%$ are given in Figure 7.15.

- **Jet Energy Scale.**

  The templates $F_X(x)$ are obtained by shifting jet energies up and down by a 1σ of their systematic uncertainty determined from the jet corrections. All kinematic quantities are recomputed to take into account these shifts. We get $\alpha_{jets,-1\sigma} = 2.6\%$ by shifting jet energies to lower values and $\alpha_{jets,+1\sigma} = 2.1\%$ by shifting jet energies upward. The distributions for templates with shifted jet energy scale are presented in Figure 7.16.

- **Top Mass uncertainty.**

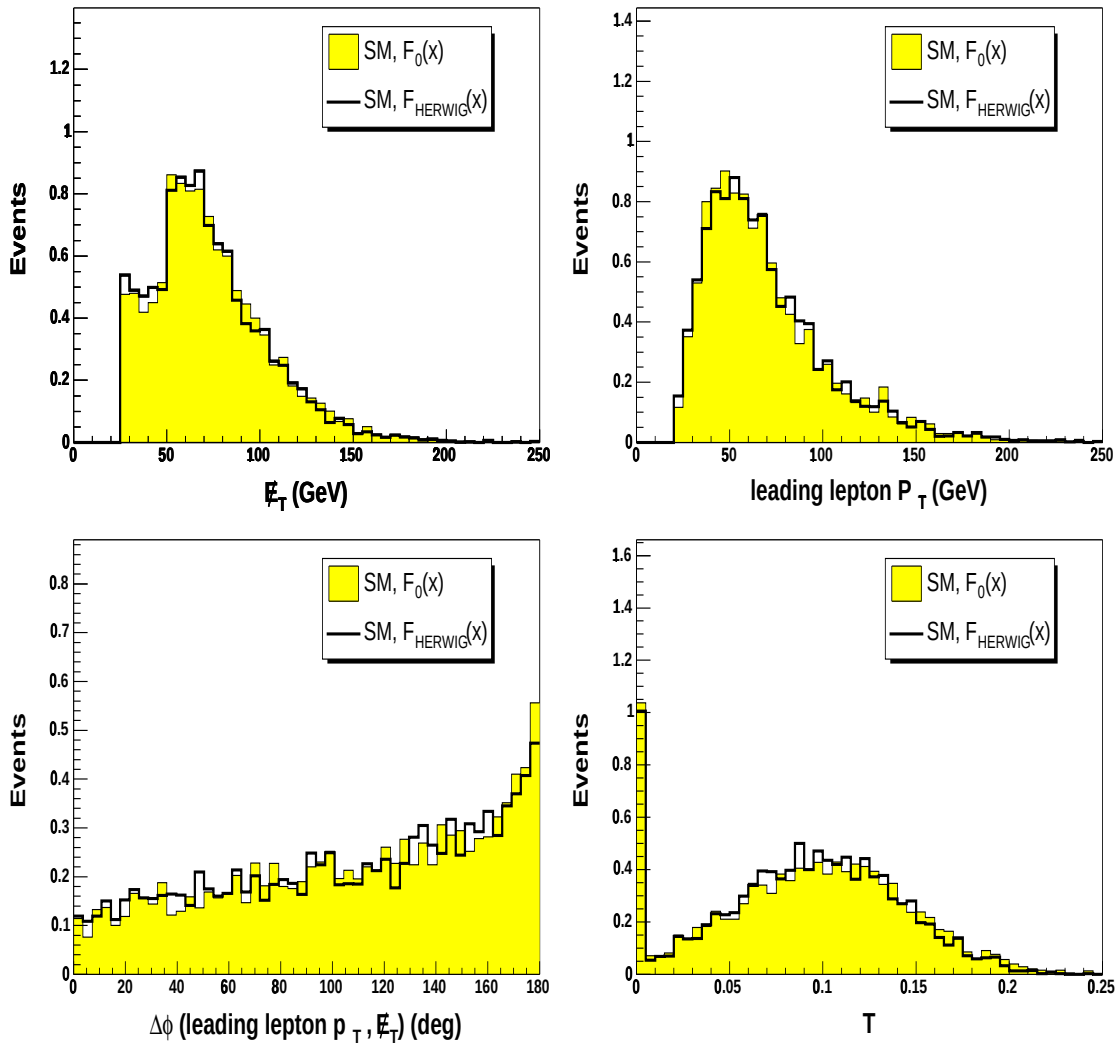

Figure 7.13: Kinematic distribution of $F_0(x)$ and $F_{HERWIG}(x)$ templates.

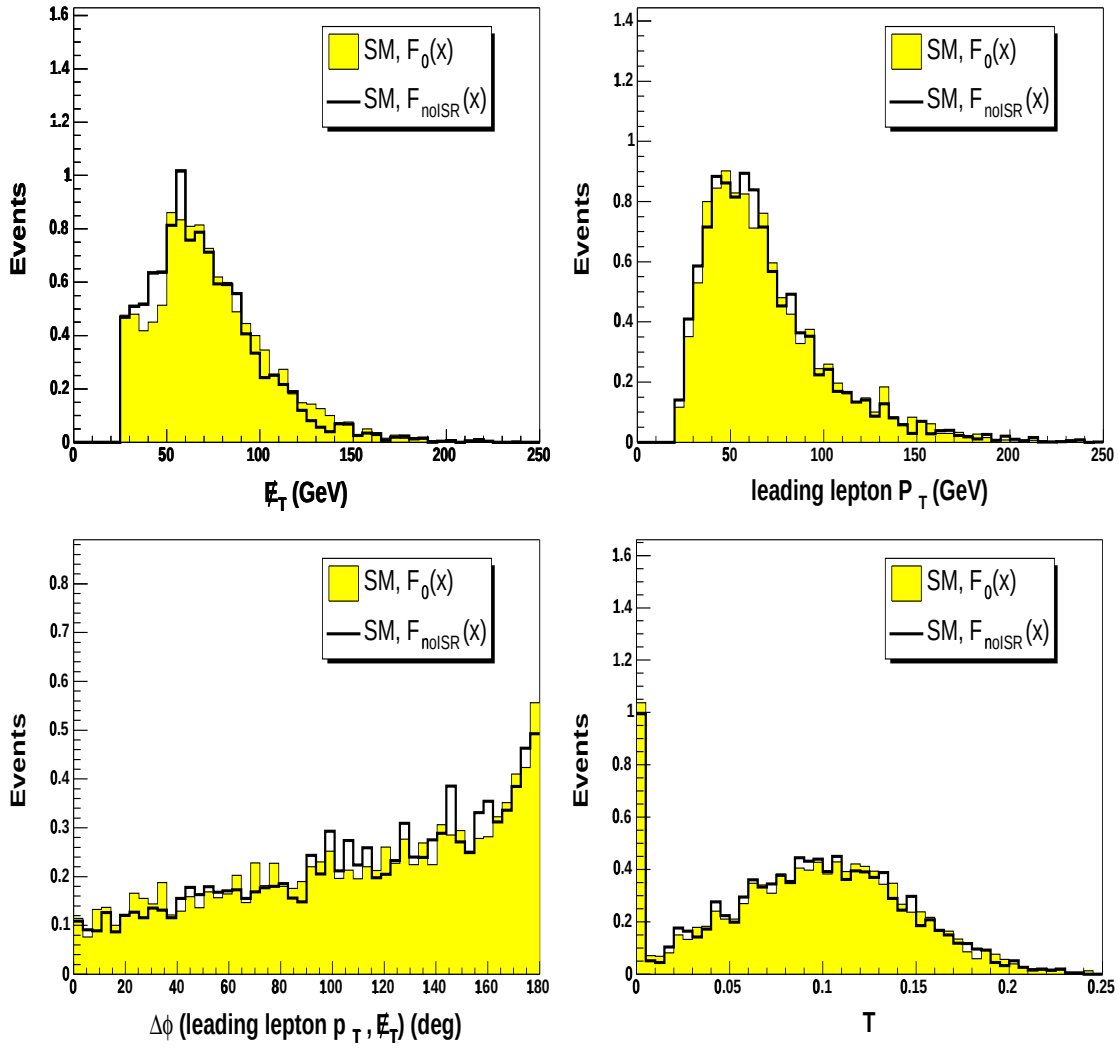

Figure 7.14: Kinematic distributions of $F_0(x)$ and $F_{noISR}(x)$ templates. The $\alpha_{noISR} = 1.2\%$.

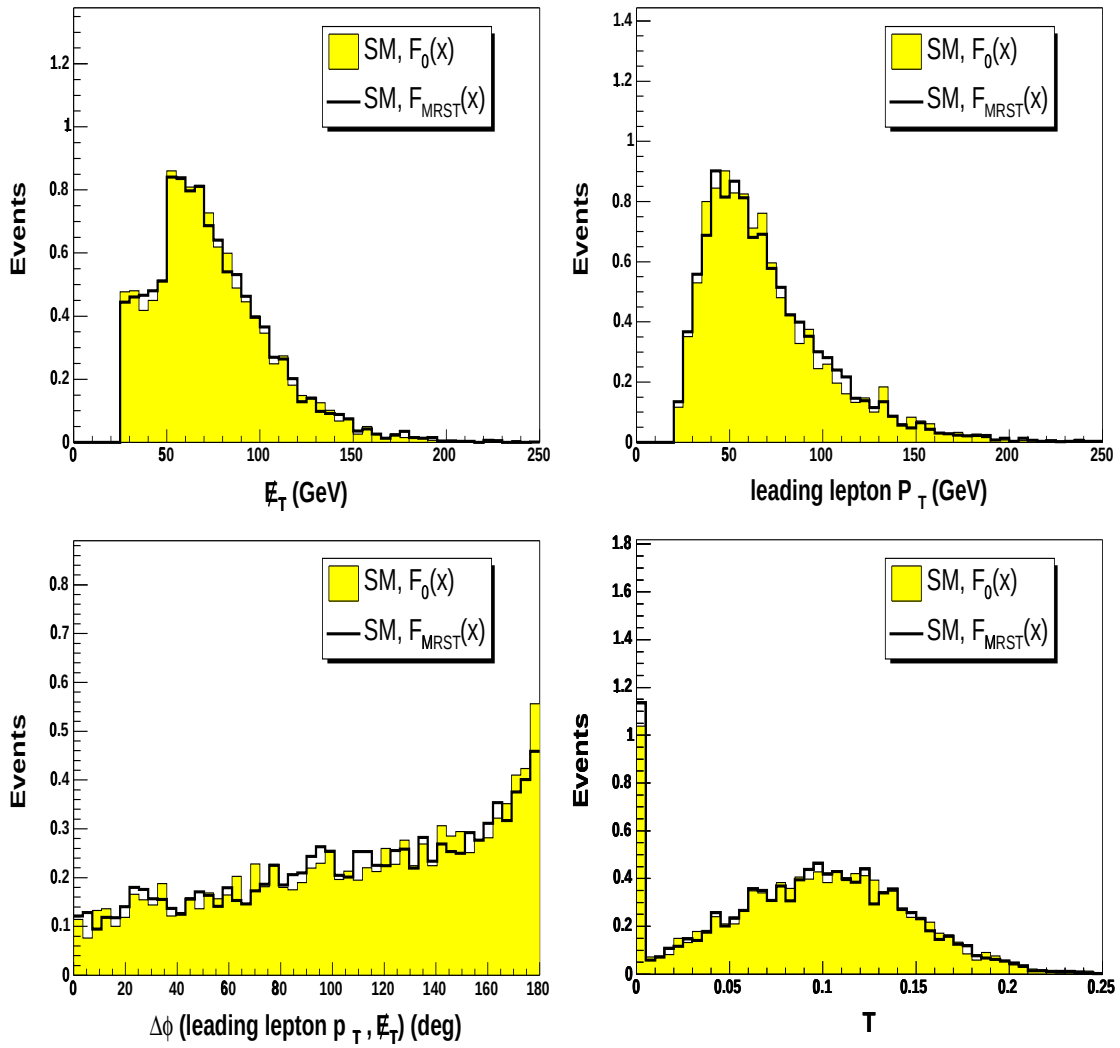

Figure 7.15: Kinematic distributions of $F_0(x)$ and F_{MRST} templates. The $\alpha_{MRST} = 1.9\%$.

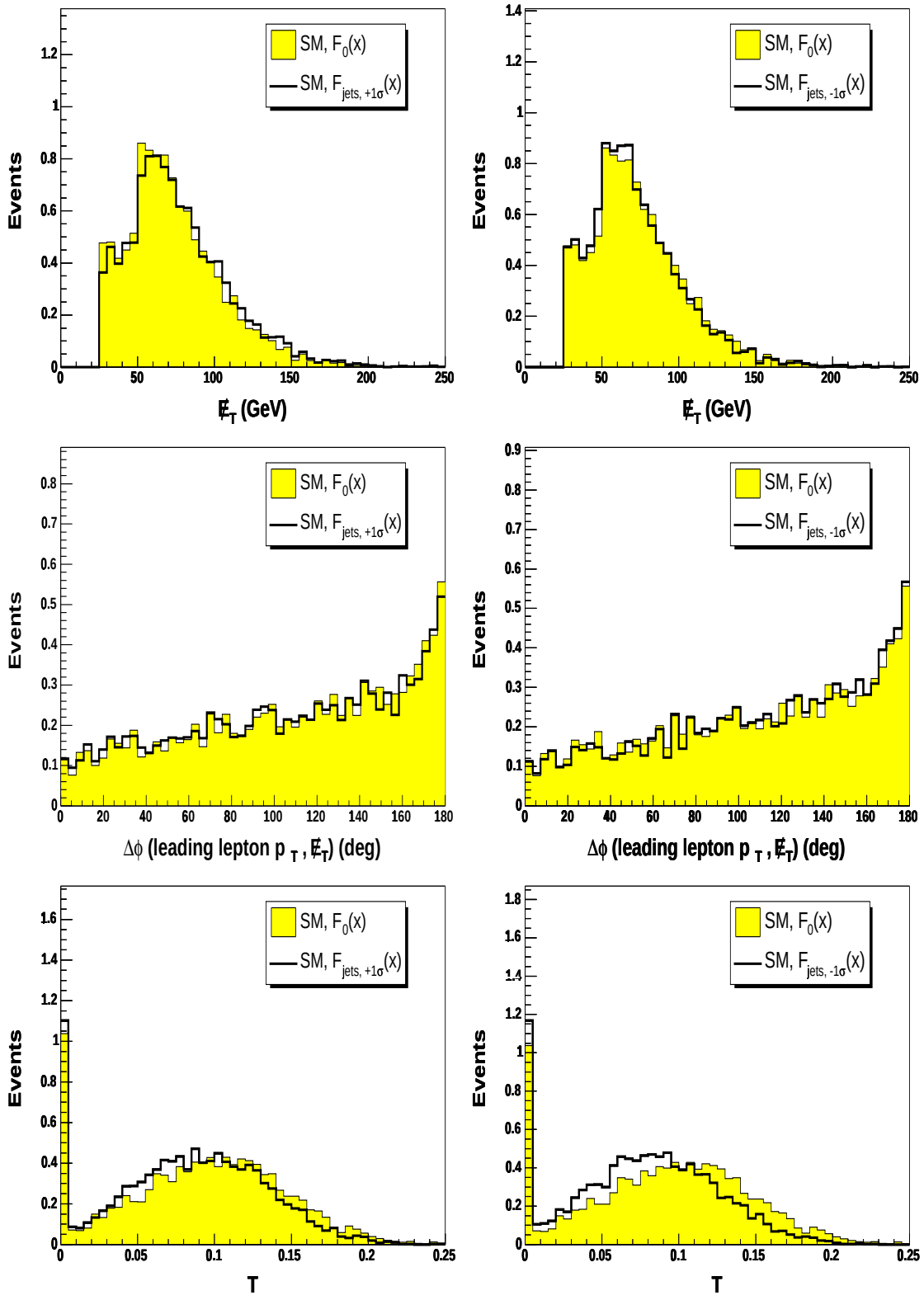

Figure 7.16: Kinematic distributions for templates with jet energies fluctuating up (left) and down (right). Leading lepton P_T distribution remains unchanged as a result of these shifts.

The up-to-date world average of the top quark mass obtained from Tevatron Run I analyzes is 175 GeV with an uncertainty of 5 GeV. To estimate this effect on our result we make use of the PYTHIA samples with $m_t = 170$ GeV and $m_t = 180$ GeV. We obtain $\alpha_{170} = 1.4\%$ and $\alpha_{180} = 2.7\%$. The distributions for templates with various top masses are given in Figures 7.17 and 7.18.

- **Uncertainties in contributions to the $t\bar{t}$ dilepton signature from other Standard Model processes.**

  As mentioned above the central value of α_0 is determined with $H0$- pseudoexperiments, where the number of events from each background contribution is Poisson smeared with mean μ given by its estimated value (see Section 5.3.1). The value α_X is obtained with pseudoexperiments where average background contributions are shifted upward/downward by 1σ. We get $\alpha_{BG,+1\sigma} = 2.7\%$ for upward fluctuated background contribution and $\alpha_{BG,-1\sigma} = 1.0\%$ for downward fluctuation. The kinematic distributions with shifted background estimates are presented in Figures 7.19 and 7.20.

  These values are obtained without varying the background shapes themselves. Due to the high purity of the $t\bar{t}$ dilepton sample, variations of the background shapes have a small effect on the overall shape of $t\bar{t}$ + non-top SM distributions. Indeed, as for $t\bar{t}$ processes we expect that the largest systematic uncertainty is due to the jet energy scale. By varying jet energies up and down for backgrounds only we obtain $\alpha_{jets,+1\sigma} = 1.7\%$ and $\alpha_{jets,-1\sigma} = 1.8\%$, which are much smaller than effects from variations of background estimates.

Lastly, we combine all of the systematic effects into one measurement. To

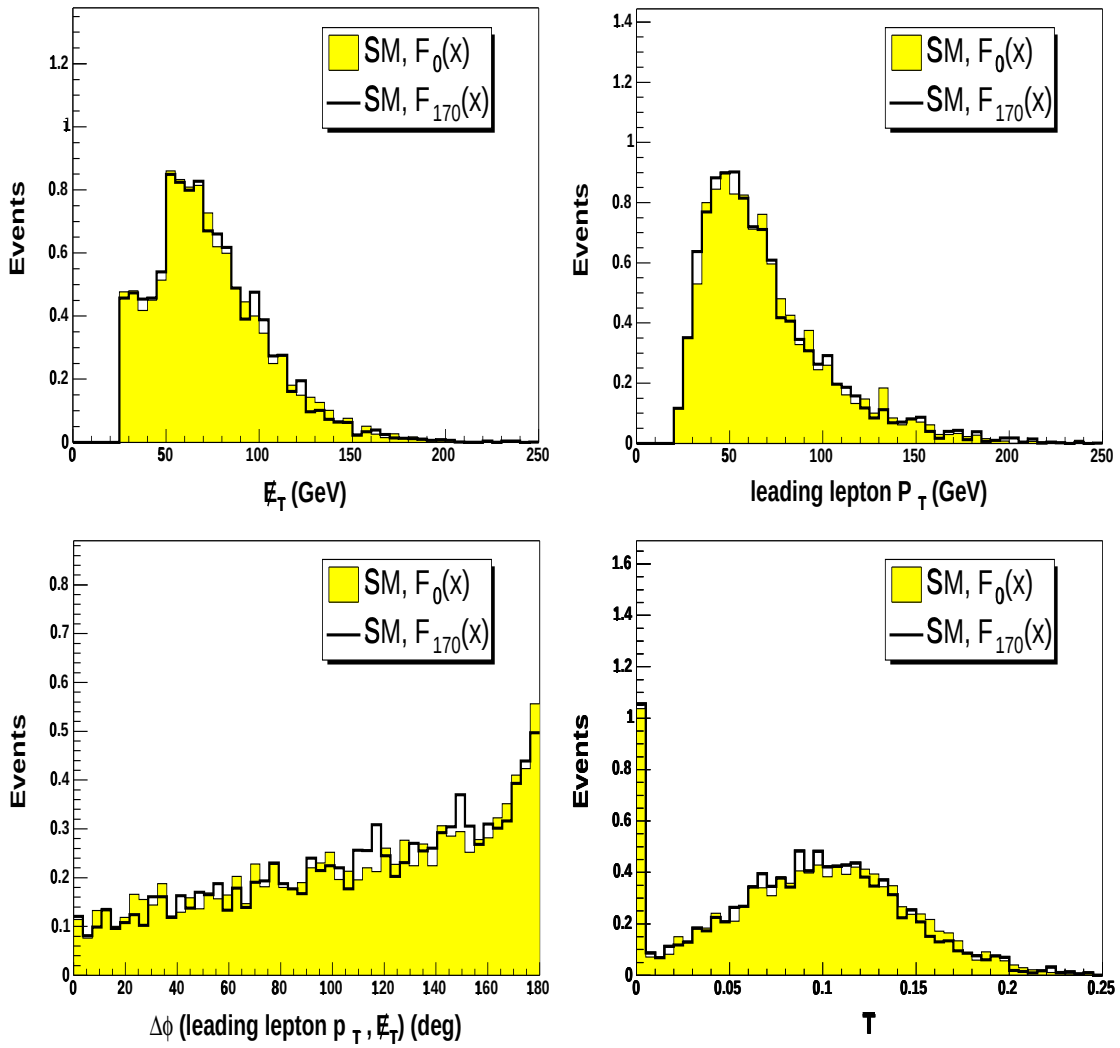

Figure 7.17: Kinematic distributions for templates with $m_t = 170$ GeV.

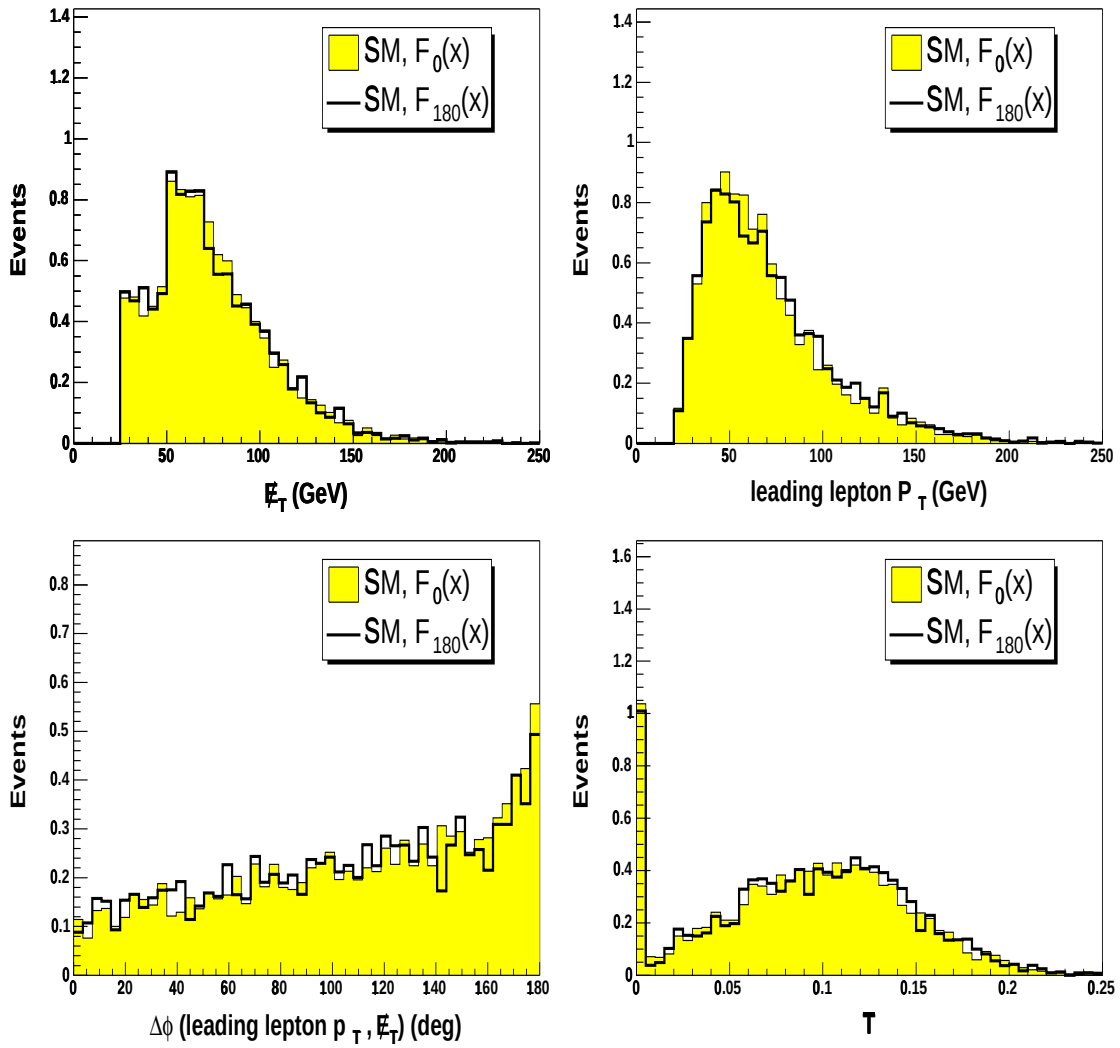

Figure 7.18: Kinematic distributions for templates with $m_t = 180$ GeV.

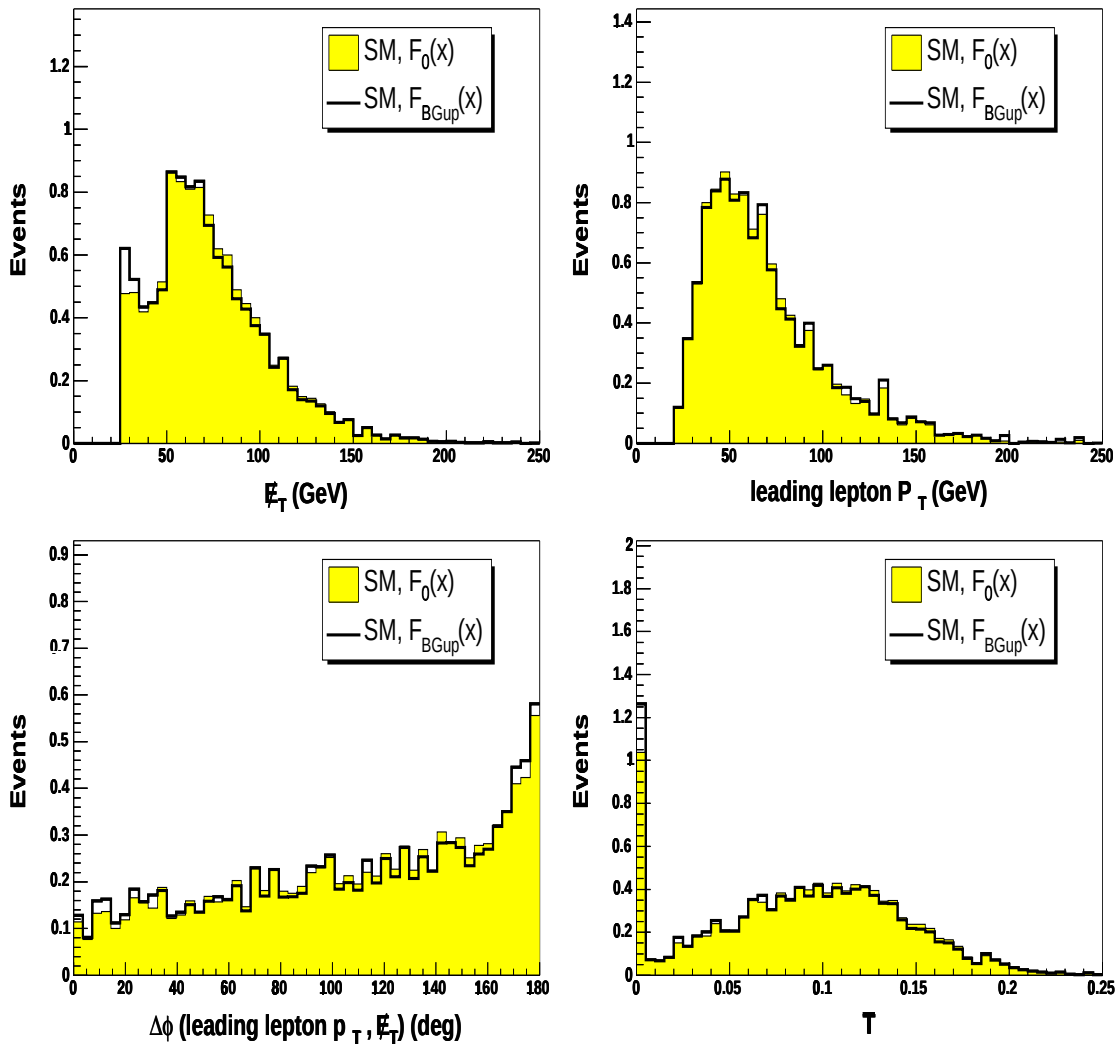

Figure 7.19: Kinematic distributions for backgrounds shifted upward.

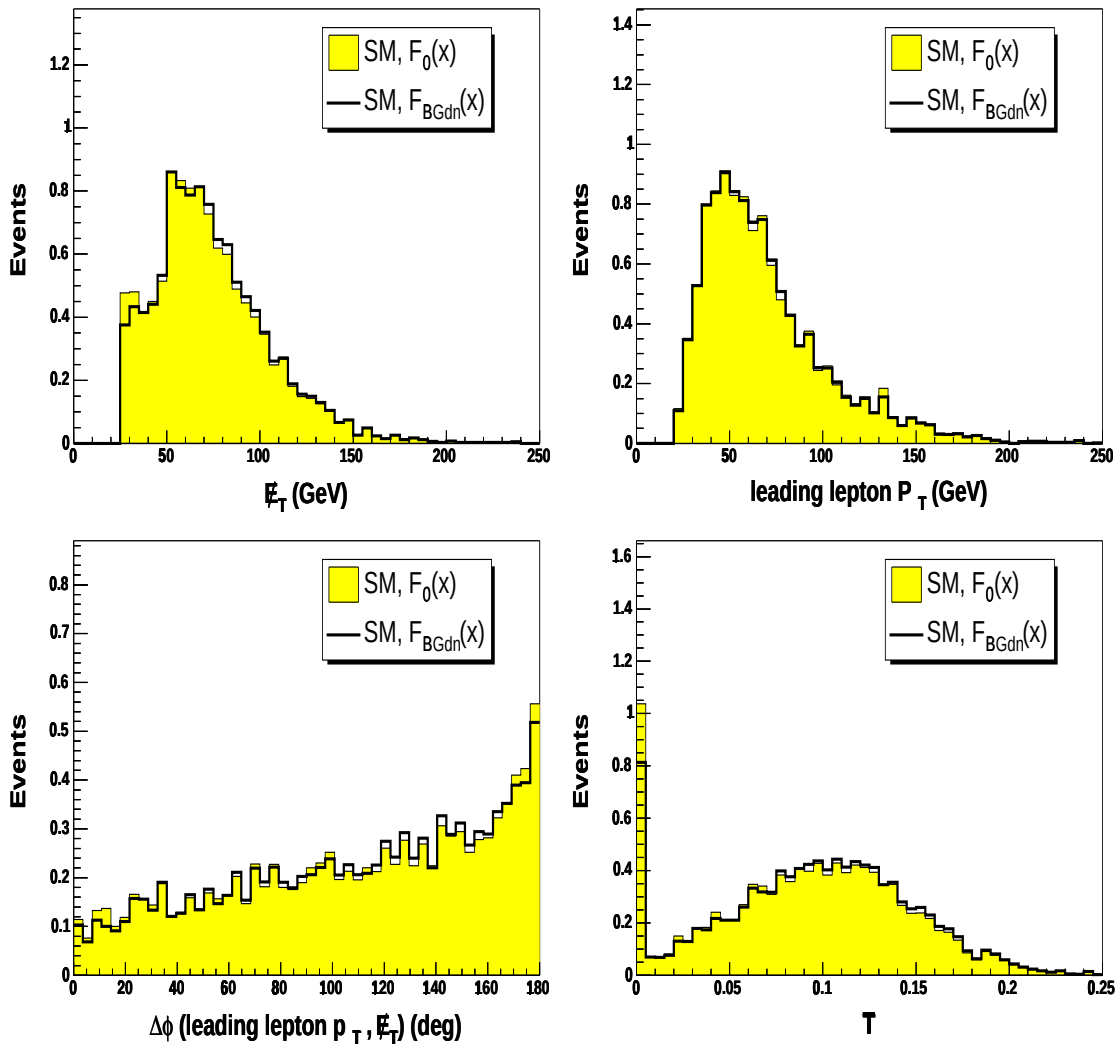

Figure 7.20: Kinematic distributions for backgrounds shifted downward.

Source of uncertainty	α' (%)
MC generator	1.6
Initial (final) state radiation	1.2 (1.6)
Parton distribution functions	1.9
M_{top} = 170 (180) GeV	1.4 (2.1)
Jet energy scale, +1 (-1) σ	2.1 (2.6)
Background estimates, +1 (-1) σ	2.7 (1.0)
combined	1.0–4.5

Table 7.6: Values of α_X obtained upon inclusion of systematic effects. The last row shows the maximum range resulting from various combinations of the individual systematics.

take into account possible correlations between them we generate a distribution $\mathcal{L}'(\mathcal{P})$ with the inclusion of all systematic sources that give a value of α_X larger than $\alpha_0 = 1.6\%$. Thus, we generate a Monte Carlo sample of $t\bar{t}$ events with the top mass $m_t = 180$ GeV and using the MRST parton distribution function with parameters resulting in the largest value of α_X. Then we generate a template $F_X(x)$ by shifting jet energies downward and background estimates upward. This combination gives the largest value of $\alpha_X = 4.5\%$.

Similarly, we repeat this procedure by including all systematic sources that result in a lower value of α_X. The smallest value of α_X is determined to be 1.0% by fluctuating backgrounds downward. All other effects result in a higher value of α_X. Uncertainties due to various systematic effects are summarized in Table 7.3.

We conclude that the probability to obtain a data sample less consistent with the Standard Model than what is observed is 1.0-4.5%. All combinations of systematic effects result in values of α lying within this range.

Chapter 8

Conclusions

In this work we have developed a novel search for new physics technique designed to isolate a subset of data events with anomalous kinematic properties and to quantify the significance of this deviation from the Standard Model expectations.

We apply this algorithm to the data sample of the $t\bar{t}$ dilepton events collected with the CDF II detector and corresponding to the integrated luminosity of 193 pb^{-1}.

We observe 13 top dilepton events in agreement with the Standard Model expectation of 10.9 ± 1.4 events. We have assessed the consistency of the $t\bar{t}$ dilepton sample in the four-variable space described. Our method is designed to be especially sensitive to data subsets that preferentially populate regions where new physics can be expected. No such subset was found in the data. We found that the probability to obtain a data sample less consistent with the Standard Model than what is observed is $1.0 - 4.5\%$.

We note that the leading lepton P_T distribution exhibits a mild excess at low P_T that is consistent with a statistical fluctuation of the Standard Model top. However, it can be concluded that new physics scenarios invoked to describe

high-P_T/high-$\not\!\!E_{\mathrm{T}}$ events observed in Run I are not favored by the current Run II data.

Bibliography

[1] F. Abe *et al.*, Phys. Rev. **D50**, 2966 (1994);
F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995), [arXiv:hep-ex/9503002];
S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995), [arXiv:hep-ex/9503003].

[2] F. Abe *et al.*, Phys. Rev. Lett. **80**, 2779 (1998), [arXiv:hep-ex/9802017].

[3] S. Abachi *et al.*, Phys. Rev. Lett. **79**, 1203 (1997), [arXiv:hep-ex/9704015].

[4] R. Barnett and L. Hall, Phys. Rev. Lett. **77**, 3506 (1996), [arXiv:hep-ph/9607342].

[5] B. Bevensee *et al.*, CDF internal note # 3159 (1995).

[6] J. Cassada *et al.*, CDF internal note # 4268 (1997).

[7] M. Kruse, PhD thesis, Purdue University (1996).

[8] M. Mangano *et al.*, arXiv:hep-ph/0303085.

[9] G. Eilam *et al.*, Phys. Rev. **D44**, 1473 (1991).

[10] I. Bigi *et al.*, Phys. Lett. **B181**, 157 (1986); L. Orr, Phys. Rev. **D 44**, 88 (1991).

[11] M. Drees. KEK-TH-501, arXiv:hep-ph/9611409.

[12] D. Dicus and V. Mathur, Phys. Rev. **D7**, 3111 (1973);
B. Lee *et al.*, Phys. Rev. **D16**, 1519 (1977).

[13] *The CDF-II Technical Design Report*, FERMILAB-PUB-96-390-E, Nov 1996.

[14] J. Elias *et al.*, Nucl. Instr. and Meth. **A 441**, 366 (2000).

[15] C. Gruppen, Particle Detectors, Cambridge University Press, 196 (1996).

[16] R. Barnett *et al.*, Phys. Rev. **D54**, 1 (1996).

[17] The CDF Collaboration, FERMILAB-Proposal-909 (1998).

[18] J. Valls, Proceedings of the 6th International Conference on Advance Technology and Particle Physics at Villa Olmo, Como (Italy), October 1998.

[19] T. Zimmerman *et al.*, Nucl. Instr. Meth. **A 409**, 369 (1998).

[20] L. Balka *et al.*, Nucl. Instr. Meth. **A 267**, 272 (1988).

[21] L. Balka *et al.*, Nucl. Instr. Meth. **A 267**, 301 (1988).

[22] Y. Seiya *et al.*, Nucl. Instr. Meth. **A 480**, 524 (2002).

[23] R. Oishi, Proceedings of the 7th International Conference on Instrumentation for Colliding Beam Physics (INSTR99), Hamamatsu, Japan, November, 1999. FERMILAB-CONF-00/036-E.

[24] A. Wyatt and B. Heinemann, CDF internal note #6167 (2002).

[25] J. Goldstein *et al.*, CDF internal note #5970 (2002).

[26] A. Taffard, CDF internal note #6255 (2003).

[27] D. Tsybychev *et al.*, CDF internal note # 6112 (2002).

[28] Y.- K. Kim *et al.*, CDF internal note # 6234 (2004).

[29] J.Y.Han *et al.*, CDF internal note # 6696 (2003).

[30] V. Martin and L. Cerrito, CDF internal note # 6825 (2004).

[31] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[32] G. Sterman *et al.*, Rev. Mod. Phys. **67**, 157 (1995).

[33] D. Goldstein *et al.*, CDF internal note # 6830 (2004).

[34] G. Corcella *et al.*, JHEP **01**, 10 (2001).

[35] Y.- K. Kim and U.- K. Yang, CDF internal note # 6804 (2004).

[36] A. Cruz and R. Field, Les Houches (2003).

[37] M.L. Mangano *et al.*, JHEP **0307**, 001 (2003).

[38] J. Campbell and R. Ellis, Phys. Rev. **D60**, 1 (1999), [arXiv:hep-ph/9905386].

[39] S. Jadach and Z. Was, Comp.Phys.Commun. **76**, 361 (1993), CERN TH-6793 preprint.

[40] D. Acosta *et al.*, FERMILAB-PUB-04/113-E. Submitted to Phys. Rev. Lett. (2004)

[41] H. Bachacou *et al.*, CDF internal note # 6904 (2004).

[42] J. Miller and D. Amidei, CDF internal note # 6865 (2004).

[43] J. Heinrich, CDF internal note # 6123 (2002).

[44] B. Roe, "Probability and Statistics in Experimental Physics",Springer-Verlag, New York, 2001.

[45] M. Stephens, Journal of the Royal Statistical Society, ser. **B**, vol. 32, 115 (1970).

[46] F. Massey, Journal of the American Statistical Association, Vol. 46 (1951).

[47] T. Anderson and D. Darling, Annals of Mathematical Statistics, vol. 23, 193 (1952).

[48] N. Kuiper, Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, ser. **A**., vol. 63, 38 (1962).

[49] L. Lönnblad *et al.*, Nucl. Phys. **B 349**, 675 (1991).

[50] J. Peacock, Monthly Notices Roy. Astron. Soc. (1983) **202**, 615

[51] L. Demortier *et al.*, FERMILAB-TM-2084 (1999)

[52] E. Varnes, Ph.D. thesis, University of California, Berkley (1997);