MEASUREMENT OF TOP PAIR PRODUCTION CROSS SECTION IN ELECTRON + JETS CHANNEL IN PPBAR COLLISIONS AT $\sqrt{s} = 1.96$ TEV

by

Chunhui Han

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2003

Doctoral Committee:

Associate Professor Jianming Qian, Chairperson
Professor Martin B. Einhorn
Professor Homer A. Neal
Assistant Professor Jignesh M. Patel
Professor Leonard M. Sanders

Ere Creation, who could tell
All the changes that befell?
What strange arts can be employed
To know what passed when all was void?
— Qu Yuan, 340 B.C.- 278 B.C.

© Chunhui Han 2003
All Rights Reserved

To Yang, who chases particles in the sky.

ACKNOWLEDGEMENTS

First of all, I want to thank all the dedicated people working for the DØ experiment. The DØ Detector is so complex, that it is impossible to operate smoothly and achieve today's success without many people's relentless work to design, develop, test, and commission different parts of the detector. These largely unknown people build a solid foundation on which all the DØ physics analyses rely on.

I entered the graduate school without a firm idea about which field to choose. Prof. Homer Neal provided me the opportunity to work on the ATLAS experiment in the first summer of my graduate study. From then on, we worked on a number of software projects related to the ATLAS experiment. I wish to express my sincere gratitutde to him, for his guidance and mentoring to a novice student entering this challenging and exciting field.

I wish to send my thanks to my advisor, Prof. Jianming Qian, who supervised my research in the DØ experiment in the past four years. No matter which project I was taking on, I was always able to seek help and support from Jianming. Due to his active leadership roles in the collaboration, he has been very busy, and I am especially thankful to the numerous discussions with him about my work. His acute physics insight have been of great help in completing this analysis.

Within the DØ collaboration, I am indebted to many people in the top physics group. Many parts of this analysis rely upon the work by other people in this excellent group. I am especially thankful to the group leaders: Emanuela Barberis, Cecilia

Gerber, Marumi Kado, and Bob Kehoe. I am grateful for fruitful discussions with Ia Iashvili, Markus Klute, Nirmalya Parua, Lukas Phaf, and Reinhard Schwienhorst.

This analysis uses a new method to measure the top pair production cross section, which is adopted from the RunI method by Juan Estrada. During each stage of the analysis, Juan has provided tremendous help based on his experience. He gave numerous suggestions, helped overcome difficult hurdles, and checked the validity of the analysis. It is a great pleasure to work with him. Dr. Yi Jiang made crucial contributions to this analysis in the early stage. I am grateful for discussions on different aspects of the analysis with him.

I enjoyed working with other members of the Michigan group in DØ including Prof. Bing Zhou, Dr. Andrew Alton, Dr. Yimei Huang, and Dr. Qichun Xu. I benefited from collaborating with Dr. Sailesh Chopra and Dr. Andre Turcot on software projects.

During my stays at Fermilab, I was fortunate to meet a number of fellow Chinese students, including Jundong Huang, Chunhui Luo, Zhong-Min Wang, Xiaojian Zhang, and Junjie Zhu. It was always enlightening to have discussions in the CGS meeting.

Finally, I wish to sent my gratitude to my parents and my wife, Yang, who shared joys and burdens in this journey.

PREFACE

The DØ experiment was proposed in 1983 and was commissioned in 1992. Together with the CDF experiment, it made distinguished scientific achievements in the past decade. In 2001, the DØ experiment began a new phase of operation called RunII, after major upgrades were done during 1996 and 2001. The analysis contained in this thesis uses RunII data accumulated up to the middle of this year (2003).

I began to study experimental high-energy physics in the summer of 1999, when I worked at CERN with Prof. Homer Neal on software projects for the ATLAS experiment. The purpose of those projects was to provide database infrastructure support to the ATLAS muon chamber production at the University of Michigan.

In 2000 I gradually got involved in the DØ experiment, working closely with Prof. Jianming Qian. My first assignment was to help move RunI production database from legacy systems to the Oracle database in the SGI machine. Later I took on a number of software projects, mostly related to the Central Preshower Detector. It is through those projects that I gained deeper understanding of detector data retrieval, unpacking, online monitoring, and physics object recontruction.

My initial thesis topic is the search for single top production in RunII. However, due to limited dataset size available in RunII, I was forced to switch to top pair production cross section measurement in the e+jets channel. Given the complexity and the large size of the experiment, this analysis obviously builds upon the work of many people in the collaboration.

I feel fortunate to join this experiment and work with so many talented people throughout the world. It is a great privilege to join this scientific endeavor, in which Nature's secret is revealed gradually to humankind.

TABLE OF CONTENTS

DEDICATION . . . . . **ii**

ACKNOWLEDGEMENTS . . . . . **iii**

PREFACE . . . . . **v**

LIST OF FIGURES . . . . . **x**

LIST OF TABLES . . . . . **xiv**

CHAPTER

I. THEORETICAL OVERVIEW 1

1.1 The Standard Model 1
1.1.1 Fundamental Particles 2
1.1.2 The Electroweak Interaction 2
1.1.3 The Strong Interaction 4
1.2 The Top Quark Physics 5
1.2.1 Top Quark Production 6
1.2.2 Top Decay 8
1.2.3 Top Pair Production Cross Section 11

II. EXPERIMENTAL APARATUS 14

2.1 The Tevatron Collider 14
2.2 The DØ Detector 17
2.3 Coordinate System 17
2.4 Luminosity and Cross Section 19
2.5 Central Tracking System 20
2.5.1 Silicon Microvertex Tracker 20
2.5.2 Central Fiber Tracker 24
2.6 The Preshower Detectors 25
2.7 Calorimeter System 27
2.8 The Intercryostat Detector 33
2.9 Muon System 34

2.10 Luminosity Monitor 36
2.11 The Trigger System 38
2.11.1 The Level 1 Triggers 39
2.11.2 The Level 2 Triggers 40
2.11.3 The Level 3 Triggers 41
2.11.4 Data Acquisition 41

III. EVENT RECONSTRUCTION 43

3.1 Electron Identification 43
3.1.1 H-Matrix Chi-Squared 45
3.2 Muon Identification 45
3.3 Jet Identification 46
3.4 Neutrino Identification 48

IV. JET ENERGY CORRECTION 50

4.1 Jet Energy Scale Corrections 51
4.2 Parameterization of δE Distributions 54
4.3 Data and MC Comparison 57

V. ANALYSIS METHOD 66

5.1 Method Overview 66
5.2 Definition of Probabilities 68
5.3 Signal Fraction Extraction 72
5.3.1 Direct Likelihood Approach 72
5.3.2 Discriminator Likelihood Approach 73
5.4 $W(y, x)$ Parameterization 74
5.5 Signal and Background Simulation 75
5.5.1 Event Generation 75
5.5.2 Detector Simulation 76
5.5.3 Event Selection 77
5.6 Production Probability Calculation 77
5.7 Observation Probability Calculation 78
5.8 f_s Extraction 84
5.8.1 Direct Likelihood Analysis 84
5.8.2 Discriminator Likelihood Analysis 86
5.9 QCD Background Effects 89
5.10 Method Comparison 94

VI. DATA ANALYSIS 95

6.1 Trigger Selection 95

6.2 Data Sets 95
6.3 Offline Selection Cuts 96
6.4 Selection Efficiency 98
6.4.1 Trigger Efficiency 99
6.4.2 Acceptance And Reco. Efficiency 100
6.4.3 Electron Identification Efficiency 102
6.4.4 Jet Identification Efficiency 107
6.4.5 Kinematic Cut Efficiency 111
6.4.6 Total Efficiency 111
6.5 Direct Likelihood Analysis 113
6.5.1 Systematic Uncertainties 115
6.6 Discriminator Likelihood Analysis 118
6.7 Top Mass Effects 120

BIBLIOGRAPHY 123

LIST OF FIGURES

Figure

1.1 Dominant single top production channels at the Tevatron. 6

1.2 Leading order top pair production channels at the Tevatron. 8

1.3 DØ measurement of the top pair production cross section at Tevatron RunI in different channels. 13

2.1 A schematic drawing of the upgraded Tevatron accelerator complex. 15

2.2 A side view of the upgraded DØ Detector. 18

2.3 A side view of the upgraded DØ Central Tracking System 21

2.4 A three dimensional view of the Silicon Tracker. 23

2.5 A schematic drawing of the central region of the DØ Detector. . . . 26

2.6 A schematic drawing of the Forward Preshower Detector (FPS) mounted on one of the two Endcap Calorimeter cryostats. 28

2.7 A schematic drawing of the calorimeter system in DØ. 30

2.8 A schematic drawing of a calorimeter cell in the DØ Detector. . . . 31

2.9 The longitudinal view of one quadrant of the DØ calorimeter system. 33

2.10 A schematic drawing of one MDT layer in the forward muon system. 36

2.11 One quadrant of the forward muon scintillation counters, showing the $r - \phi$ segmentation. 37

2.12 The Level-2 Trigger system structure. 40

4.1 The reconstructed jet energy as a function of the parton energy, without Jthe et Energy Scale correction. 50

4.2 The reconstructed jet energy as a function of the parton energy, after the JES correction. 53

4.3 δE distributions in different parton energy ranges, for light quark jets. 55

4.4 δE distributions in different parton energy ranges, for b-quark jets. . 56

4.5 Test of the jet energy transfer function for light quark jets. 58

4.6 Test of the jet energy transfer function for b quark jets. 59

4.7 ΔE_T and $\Delta\phi$ distributions in Monte Carlo photon+jet events. . . . 62

4.8 Comparison of $\delta E'$ distributions between Monte Carlo (solid histograms) and data (dashed histograms). 63

4.9 Comparison between δE distributions and $\delta E'$ distributions in Monte Carlo samples, for light quark jets. 64

4.10 Comparison between δE distributions and $\delta E'$ distributions in Monte Carlo samples, for b quark jets. 65

5.1 The δE and the δR distributions for electrons in Monte Carlo events. 76

5.2 Signal production probability density ($P_s^{prd}(y)$) distributions for signal (dashed line) and background (solide line) Monte Carlo events. . 79

5.3 The uncertainty of Monte Carlo integrations in the calculation of $P_s^{obs}(x)$. 81

5.4 The uncertainty of Monte Carlo integrations in the calculation of $P_b^{obs}(x)$. 82

5.5 Comparison of observation probability distributions. 83

5.6 Observation probability ($P_s^{obs}(x)$ and $P_b^{obs}(x)$) distributions for signal and background Monte Carlo samples. 85

5.7 The Direct Likelihood Function shapes with different n_s values. . . 87

5.8 Distributions of the extracted f_s in Monte Carlo tests of the Direct Likelihood Analysis. 88

5.9 The linearity tests in the Direct Likelihood Analaysis. 89

5.10 The discriminator distributions for signal (shaded histogram) and background (open histogram) Monte Carlo events. 90

5.11 The linearity test in the Discriminator Likelihood Analysis. 90

5.12 Shapes of the Discriminator Likelihood Function for randomly selected Monte Carlo datasets with different number of signal events. 91

5.13 Distributions of the extracted f_s in Monte Carlo tests of the Discriminator Likelihood Analysis. 92

5.14 (a): The observation probability density distribution for the QCD multijet data events. (b): The discriminator distribution for the QCD multijet data events. 93

5.15 Mixture tests for QCD events extracted from data samples. 94

6.1 The jet multiplicity distribution for signal events. 98

6.2 η and p_T distributions of electrons and jets of the signal process. Parton-level information is used. 101

6.3 The invariant Z mass distributions in $Z \to e^+e^-$ events. 105

6.4 The jet identification efficiency as a function of jet E_T. 109

6.5 (a) shows the jet transverse energy distribution in Monte Carlo signal events. (b) shows the jet multiplicity distribution for Monte Carlo signal events. The jets are required to pass jet identification cuts. . 109

6.6 (a): The neutrino p_T distribution in Monte Carlo signal events. (b): The distribution of the azimuthal separation between the electron and the neutrino in Monte Carlo signal events. 112

6.7 The kinematic distributions of the W boson which decays leptonically in Monte Carlo signal events. 112

6.8 The signal and background observation probability distributions for the dataset are shown in a scatter histogram. 114

6.9 (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the final dataset in this analysis. (b) shows the linearity test result using Monte Carlo events. 114

6.10 Statistical uncertainty in the Direct Likelihood Analysis. 115

6.11 The mixture test with 8 signal events and 42 background events in the Monte Carlo test sample. This test is carried out with the Direct Likelihood Analysis. 117

6.12 (a) shows the discriminator distribution for the dataset. (b) shows the shape of the negative logarithm of the Discriminator Likelihood Function for the dataset. 119

6.13 (a) shows the Discriminator Likelihood Function for the data analysis. The shaded area covers 68% of the total integrated area covered by the function. (b) shows the f_s uncertainty distribution in Monte Carlo tests with the Discriminator Likelihood Analysis, where the number of signal events is 14, and the total number of events is 50 in the test sample. The arrow indicates the f_s uncertainty in the data analysis. 119

6.14 (a) shows the $t\bar{t}$ production cross section as a function of the top mass. This is the leading-order calculation by the Monte Carlo generator ONETOP. (b) shows the selection cut efficiency as a function of the top mass, obtained from Monte Carlo signal samples. 120

6.15 (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the dataset assuming $m_t = 170$ GeV. (b) shows the shape of the Direct Likelihood Function for the dataset. The shaded area covers 68% of the total integrated area covered by the function, so its border is equivalent to a $(-\sigma, +\sigma)$ deviation from the peak point. 121

6.16 (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the dataset assuming $m_t = 180$ GeV. (b) shows the shape of the Direct Likelihood Function for the dataset. The shaded area covers 68% of the total integrated area covered by the function, so its border is equivalent to a $(-\sigma, +\sigma)$ deviation from the peak point. 122

LIST OF TABLES

Table

1.1 Elementary particles of the Standard Model. 3

1.2 Standard Model predictions to top quark production cross sections at Tevatron RunI and RunII. 9

1.3 Top decay channels and the corresponding branching ratios. 10

4.1 Parameters for the jet energy transfer function $f(\delta E)$, obtained using Monte Carlo samples. 57

6.1 Details of the DØ top trigger EM15_2JT15. 96

6.2 Data samples used in this analysis. 97

6.3 Details of the selection cuts. 97

6.4 Numbers of signal and background events in Sample I and Sample II in the calculation of the electron identification efficiency. 106

6.5 A list of efficiency terms and the combined efficiency. 113

6.6 A list of additional systematic uncertainty sources and their values. 117

CHAPTER I

THEORETICAL OVERVIEW

Particle physics is the branch of science that aims to explain the behavior of the most basic building blocks of the universe: elementary particles. During the 20th century, physicists made great achievements in this field. This chapter first gives an overview of modern particle physics theory, then presents the basis of the top quark physics.

1.1 The Standard Model

The Standard Model [1] is the theoretical basis of modern particle physics. It came into existence in the 1970's, and numerous experimental results have confirmed its validity. At the same time, the Standard Model cannot be viewed as an ultimate theory in particle physics: although it gives correct descriptions to the kinematic behaviors of most elementary particle processes, it does not explain the origin of all the properties of elementary particles. Various theories at higher energy scales have been proposed to account for the deficiency of the Standard Model, but none has been firmly verified by experiments.

1.1.1 Fundamental Particles

In the Standard Model, the basic elements of the universe are grouped into two categories: spin-$\frac{1}{2}$ fermions which constitute matter, and spin-1 gauge bosons which act as force carriers. The fermions are further classified as quarks and leptons. So far, six quarks and six leptons have been discovered, together with the antiparticles for each type of fermion. Forces in Nature are classified into four types: the electromagnetic force, the strong force, the weak force, and the gravitational force. Each type of force is mediated by a class of gauge bosons. A list of the elementary particles and their properties is shown in Table 1.1.

The Standard Model is a gauge theory, which means that certain gauge invariance is observed in it. The Standard Model theory is symmetric under three gauge transformations: the U(1) symmetry, the SU(2) symmetry, and the SU(3) symmetry. To accomodate the three symmetry transformations, the Standard Model requires corresponding fields to interact with fermions, which leads to the existence of three of the four interactions mentioned above, except the gravitational force. The following sections discuss the Standard Model interactions individually.

1.1.2 The Electroweak Interaction

In the Standard Model, the electromagnetic interation and the weak interaction are unified into one electroweak interaction, as a result of the U(1)$\times$SU(2) symmetry.

The massless photon is the electromagnetic force carrier. The theory of Quantum Electrodynamics (QED) describes the electromagnetic interactions between charged particles. The weak interaction is mediated by the $W^{\pm}$, and the Z^0 bosons, which have relatively large masses.

To accomodate the SU(2) symmetry in the Standard Model, the following quark

Quarks	Charge(e)	Mass
up (u)	$\frac{2}{3}$	1.5 to 4.5 MeV
down (d)	$-\frac{1}{3}$	5 to 8.5 MeV
strange (s)	$-\frac{1}{3}$	80 to 155 MeV
charm (c)	$\frac{2}{3}$	1.0 to 1.4 GeV
bottom (b)	$-\frac{1}{3}$	4.0 to 4.5 GeV
top (t)	$\frac{2}{3}$	174.3± 5.1 GeV
Leptons	**Charge (e)**	**Mass**
electron (e)	-1	0.511 MeV
electron neutrino (ν_e)	0	$<$ 3 eV
muon (μ)	-1	106.7 MeV
muon neutrino (ν_μ)	0	$<$ 0.19 MeV
tau (τ)	-1	1.777 GeV
tau neutrino (ν_τ)	0	$<$ 18.2 MeV
Bosons	**Charge (e)**	**Mass**
photon (γ)	0	0
gluon (g)	0	0
$W^\pm$	± 1	80.423±0.039 GeV
Z^0	0	91.1876±0.0021 GeV

Table 1.1: Elementary particles of the Standard Model [5]. The graviton associated with the gravitational force is not included in the Standard Model theory. The u, d, and s quark masses are estimated in a mass-independent subtraction scheme. The c and b quark masses are estimated from charmonium, bottomonium, D, and B masses. The t quark mass is measured from direct observation of top events.

doublets are formed:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \qquad \begin{pmatrix} c \\ s \end{pmatrix}_L, \qquad \begin{pmatrix} t \\ b \end{pmatrix}_L;$$

and the follwing lepton doublets are formed:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L, \qquad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L, \qquad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L.$$

The subscript "L" for each doublet means that only left-handed fermions, or right-handed anti-fermions, can form doublets. In this way parity is violated in the Standard Model.

One quark doublet and the corresponding lepton doublet form a generation of fermions. Three generations have been discovered so far. Fermions in the first generation are stable particles, while fermions in the second and third generations, except neutrinos are unstable, and will decay into fermions with lighter mass.

The mass spectrum of the quarks and the weak gauge bosons had been a problem in the Standard Model framework, since an unbroken SU(2) symmetry would require all particles to be massless. This problem can be solved by introducing an additional field, called the Higgs field, to the Standard Model [2]. All massive particles interact with this field, and the mass is dependent on the coupling strength with the Higgs field. As the Higgs field is quantized, a new kind of particle, called the Higgs particle, is introduced. So far all experiments have failed to find direct evidence of the Higgs particle. As a sequel to the lengthy quest, The Tevatron RunII is charged with the major task of searching for this elusive particle.

1.1.3 The Strong Interaction

Quarks carry an intrinsic property called color charge. SU(3) symmetry requires three color charges for each quark, and they are labeled as red (r), blue (b), and green (g). Interactions due to the color charge are called color forces, or strong forces

because of the relatively large magnitude as compared to other forces. The strong force is mediated by eight types of gluons. The theory of Quantum Chromodynamics (QCD) describes the strong interaction.

An important aspect of QCD is color confinement: only color neutral objects can exist. Since quarks carry color chages, they can only exist in compound states. This property has deep implications in jet physics in high energy experiments.

1.2 The Top Quark Physics

The top quark is the last discovered quark. The search for the top quark began after the discovery of the bottom quark in 1977. Several strong reasons made physicists believe that the top quark must exist. The renormalization of the Standard Model requires that the triangle anomalies be cancelled, which is true if the electric charges of all left-handed fermions in one generation sum to zero. Applying this to the third generation requires a top quark with a charge of $+\frac{2}{3}e$.

The most compelling experimental evidence comes from the measurement of b quark weak isospin in e^+e^- collider data. Denoting $I_3^L(b)$ and $I_3^R(b)$ as the third components of the isospin for the left-handed and right-handed b quark fields, the measurements at LEP, SLC, and other low energy experiments (PEP, PETRA, and TRISTAN), show[6]:

$$I_3^L(b) = -0.490^{+0.015}_{-0.012}$$

$$I_3^R(b) = -0.028 \pm 0.056$$

This implies that the b quark must have a weak isospin partner with $I_3^L = 1/2$.

However, due to the top quark's unexpectedly large mass, it took nearly two decades to find direct evidence of its existence in particle detectors. The top quark was discovered in 1995 by the CDF and DØ collaborations at the Fermi National

Accelerator Laboratory [7] [8] [9]. Measurements on various top quark properties were carried out for different production and decay channels. In the following sections the production and decay modes of the top quark at the hadron collider are discussed.

1.2.1 Top Quark Production

At the Tevatron Collider (see Chapter II), where proton and anti-proton beams collide, the top quark can be produced in two modes: single top production mode via electroweak interactions and pair production mode via QCD interactions. The cross sections for both modes are listed in Table 1.2.

Single Top Production

Figure 1.1 shows the Feynman diagrams for dominant single top production channels at hadron colliders.

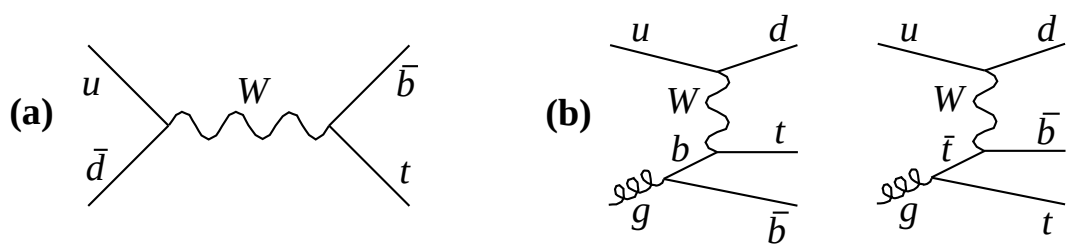


Figure 1.1: Dominant single top production channels at the Tevatron. (a) shows the s-channel process; (b) shows the t-channel processes.

Figure 1.1(a) is the s-channel single-top process, $q'\overline{q} \rightarrow t\overline{b}$ (also known as the W^* process, because a virtual W boson acts as the propagator). At Tevatron RunI with center-of-mass energy $\sqrt{s} = 1.8$ TeV, its next-to-leading-order (NLO) cross section has been calculated to be 0.73 ± 0.10 pb [11] with a top mass of 175 GeV. Figure 1.1(b) shows the t-channel process, $q'g \rightarrow tqb$ (also known as the W-gluon

fusion), and its NLO cross section is 1.70 ± 0.24 pb [12]. There is also another single top production channel of $p\overline{p} \rightarrow Wt$, but the event rate is negligible at the Tevatron.

Single top production events were not observed at Tevatron RunI. The t-channel process is the main single top production process, but its final state contains a jet in the forward direction, making it difficult to detect. The dominant background comes from $Wb\overline{b}$ events, and the rest of the background comes from $t\overline{t}$ pair production and QCD processes. Since the single top production cross section is directly proportional to the CKM [5] matrix element $|V_{tb}|^2$, it provides a unique opportunity to directly measure the matrix element value, and check the unitarity of the CKM matrix.

Top Pair Production

Figure 1.2 shows the leading order Feynman diagrams for top pair production channels, including the process $q\overline{q} \rightarrow t\overline{t}$, and the process $gg \rightarrow t\overline{t}$. There are also contributions from other channels with intermediate Z bosons and photons, but their event rate is much smaller at the Tevatron.

At Tevatron RunI, with a center-of-mass energy of 1.8 TeV and a top mass of 175 GeV, the total cross section of the top pair production processes is predicted to be around 5 pb [10], which is bigger than the single top production cross section by around a factor of 2. The top quark was discovered in the top pair production channels. Summing results from all the channels, the DØ collaboration reported a $t\overline{t}$ production cross section of 5.69±1.21(stat)±1.04(sys) pb with a top quark mass of 172.1 GeV/c^2 [15] at $\sqrt{s} = 1.8$ GeV. The CDF collaboration reported a cross section of $6.5^{+1.7}_{-1.4}$ pb with a top mass of 175 GeV/c^2 [16].

At the parton level, the cross sections of the two main top pair production subprocesses, $p\overline{p} \rightarrow t\overline{t}$ and $gg \rightarrow t\overline{t}$, are comparable. However, the relative contributions

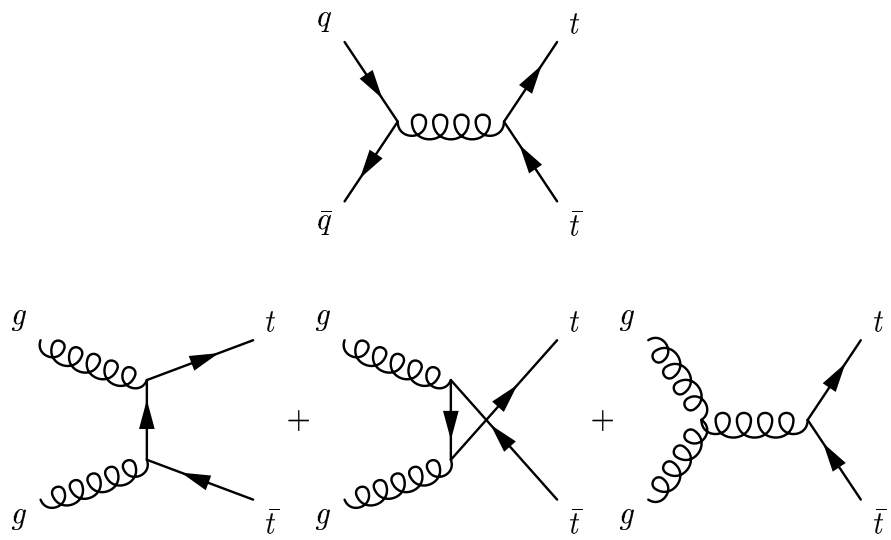


Figure 1.2: Leading order top pair production channels at the Tevatron: $q\overline{q} \to t\overline{t}$, and $gg \to t\overline{t}$.

of the two processes to the total cross section are 90% and 10%, respectively, at Tevatron RunI. This mainly arises from the difference in the parton distribution functions for quarks and gluons with $\sqrt{s} = 1.8$ TeV.

The upgraded Tevatron (Tevatron RunII) gives significant boost to the top quark production cross sections. Table 1.2 compares the cross sections for different channels at Tevatron RunI and RunII.

1.2.2 Top Decay

In the Standard Model, the top quark decays through the electroweak interaction into a W boson and a down type quark, which can be a bottom quark, a strange quark, or a down quark. The probability that the top decays into any of the three quarks (the branching ratio) is determined by the corresponding CKM matrix element. Results from previous experiments indicate that the top quark largely decays into a W and a bottom quark. In RunI, the ratio has been measured by the CDF

Analysis	top pair production	single top s-channel	single top t-channel
RunI NLO calculation	≈5 pb	0.7 pb	1.4 pb
RunII NLO calculation	≈7 pb	0.9 pb	2.4 pb
cross section improvement	40%	29%	71%

Table 1.2: Standard Model predictions to top quark production cross sections at Tevatron RunI ($\sqrt{s}$ = 1.8 TeV) and RunII ($\sqrt{s}$ = 1.96 TeV), in each production channel.

collaboration to be [14]:

$$\frac{BR(t \to Wb)}{BR(t \to Wq)} = 0.99 \pm 0.29 \equiv \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$$

The W boson subsequently decays in one of several decay modes. It can decay leptonically into a lepton and the corresponding neutrino, or it can decay hadronically into a pair of quarks. Due to the color confinement effect, we can not directly measure individual quarks in the detector. Instead, a cluster of energy, called a jet, in the direction of the quark shows up in the calorimeter.

In a top pair production event, the two top quarks ($t\bar{t}$) decay predominantly into two bottom quarks, one W^+, and one W^-. According to the decay modes of the two W's, we can categorize the top pair production events into different decay channel groups. Table 1.3 shows the branching ratio for each channel.

1. *dilepton channels*

If both W's decay leptonically, this $t\bar{t}$ event is called a dilepton event. Since the τ lepton is difficult to identify in the detector, the particular dilepton channels which have been most studied are the ee, $\mu\mu$, and $e\mu$ channels. The final state signature is two high-p_T leptons, large missing E_T, and two b-jets. The dilepton channels have

$W \rightarrow$ $\downarrow$	$W^+ \rightarrow e^+\nu_e$ (1/9)	$W^+ \rightarrow \mu^+\nu_\mu$ (1/9)	$W^+ \rightarrow \tau^+\nu_\tau$ (1/9)	$W^+ \rightarrow q\overline{q'}$ (2/3)
$W^- \rightarrow e^-\overline{\nu}_e$ (1/9)	ee (1/81)	$e\mu$ (1/81)	$e\tau$ (1/81)	e + jets (6/81)
$W^- \rightarrow \mu^-\overline{\nu}_\mu$ (1/9)	$e\mu$ (1/81)	$\mu\mu$ (1/81)	$\mu\tau$ (1/81)	μ + jets (6/81)
$W^- \rightarrow \tau^-\overline{\nu}_\tau$ (1/9)	$e\tau$ (1/81)	$\mu\tau$ (1/81)	$\tau\tau$ (1/81)	τ + jets (6/81)
$W^- \rightarrow q\overline{q'}$ (2/3)	e + jets (6/81)	μ + jets (6/81)	τ + jets (6/81)	all jets (36/81)

Table 1.3: Top decay channels and the corresponding branching ratios. Top decay channels are based on the decay modes of the two W bosons. The branching ratio of each decay channel is listed in parentheses.

small backgrounds, but have small branching fraction leading to smaller number of signal events. Moreover, since dilepton events contain two unobserved neutrinos, it is impossible to do complete reconstruction of the event kinematics.

2. *lepton+jets channels*

In this group of channels, one W boson decays leptonically, and the other decays hadronically. The final state has the signature of one high-p_T lepton, large missing E_T, and four jets, two of which are b-jets. Compared to the dilepton channels, the lepton+jets channels have a much larger cross section, as evident in Table 1.3. Since the final state contains only one neutrino, there is sufficient information to completely reconstruct the event, once a particular set of assignments of jets to the final-state partons is assumed and a particular solution of the z-momentum of the neutrino is chosen. The disadvantage of these channels is the presence of a large background from inclusive W boson productions with associated jets, plus a smaller background from QCD multijet production, where a jet is misidentified as a lepton.

3. *all-jets channel*

Events in which both W bosons decay hadronically are called all-jets events. The final state contains six jets, of which two are b-jets, no high-p_T leptons, and small missing E_T. Although this channel has the largest branching ratio, its identification is affected by the huge background coming from QCD multijet processes.

1.2.3 Top Pair Production Cross Section

One of the key quantities that have been measured at Tevatron is the top pair production cross section. Its measurement is compared with the theoretical prediction to test the validity of the Standard Model. In addition, the cross section provides a window to probe new physics beyond the Standard Model. Several hypothetical

theories beyond the Standard Model, which include multiscale technicolor [3] and electroweak isoscalar quarks [4], enhance the top pair production cross section.

The DØ collaboration measured the top pair production cross section at Tevatron RunI in all the decay channels [15]. The combined RunI results are summarized in Figure 1.3. In early 2003 the DØ collaboration released the first result of $t\bar{t}$ pair production cross section measurement at Tevatron RunII in different channels [13]. In this thesis, a new method, the Matrix Element Method, is used to measure the cross section in the e+jets channel.

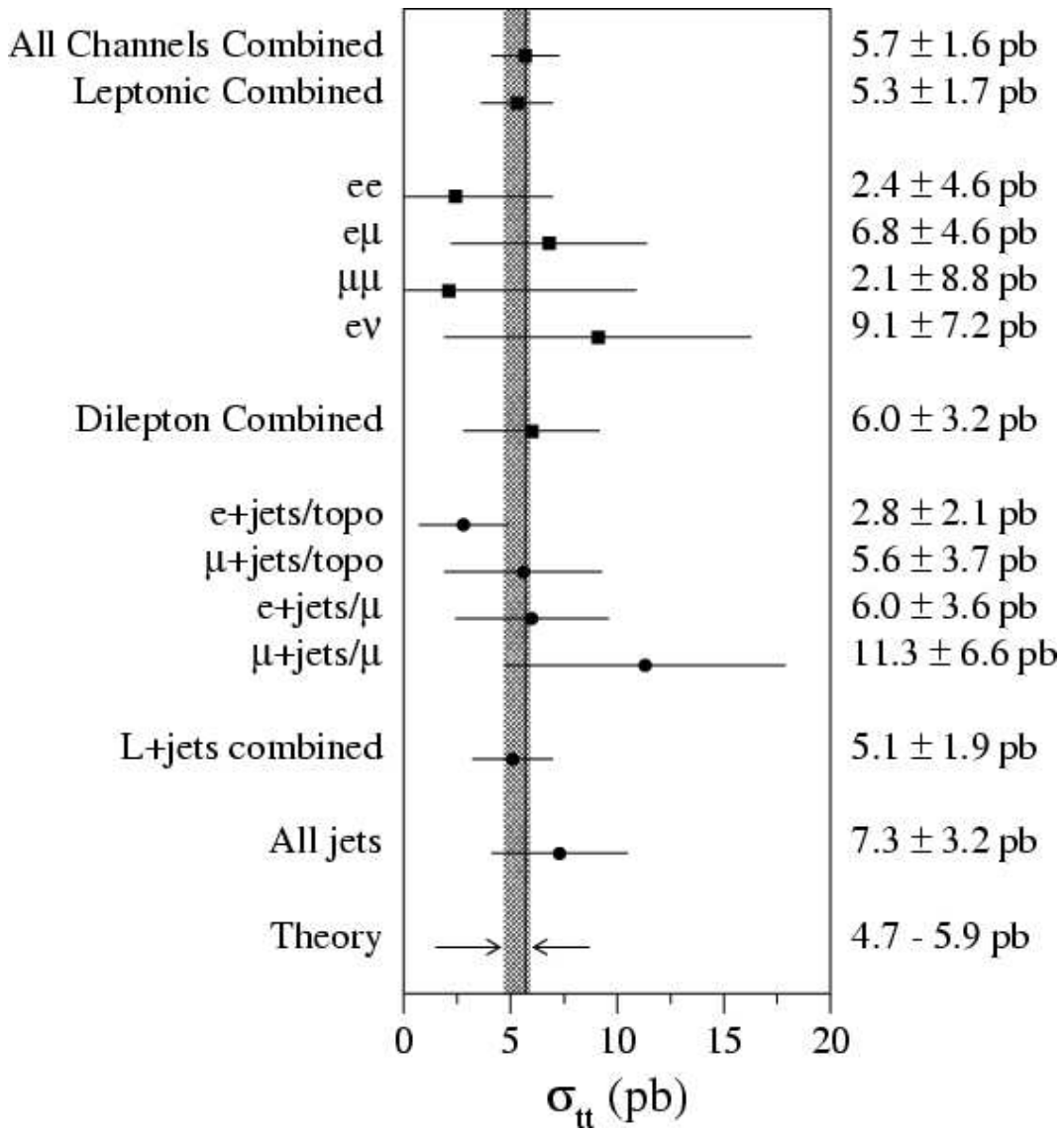


Figure 1.3: DØ measurement of the top pair production cross section at Tevatron RunI in different channels.

CHAPTER II

EXPERIMENTAL APARATUS

This chapter presents a brief description of the Fermilab accelerator complex and elements of the DØ Detector during the second operation period, Run II. Emphasis is put on the upgraded components. A complete discussion of the RunI DØ Detector can be found in Reference [19]. The reader is directed to Reference [20] for further information about the detector upgrade.

2.1 The Tevatron Collider

The Fermi National Accelerator Laboratory (Fermilab), located in Batavia, IL, is home to the Tevatron Collider. For a decade, the Tevatron has been the world's highest energy hadron collider. It uses proton and antiproton beams for hadron collisions. Two particle detectors, called CDF and DØ respectively, are built at two separate locations along the accelerator ring.

During the period of 1990 - 1996, the Tevatron operated at a center-of-mass energy of 1.8 TeV. This period is referred to as RunI. Both the CDF and DØ collaborations made distinguished achievements during RunI, including the discovery of the top quark.

The Tevatron was shut down in 1996, and has undergone a major upgrade ever

since, together with the CDF and DØ detectors. The center-of-mass energy has been boosted to 1.96 TeV after the upgrade. The Main Ring, an accelerator used in RunI, has been replaced by the Main Injector. A schematic drawing of the upgraded Tevatron accelerator complex is shown in Figure 2.1.

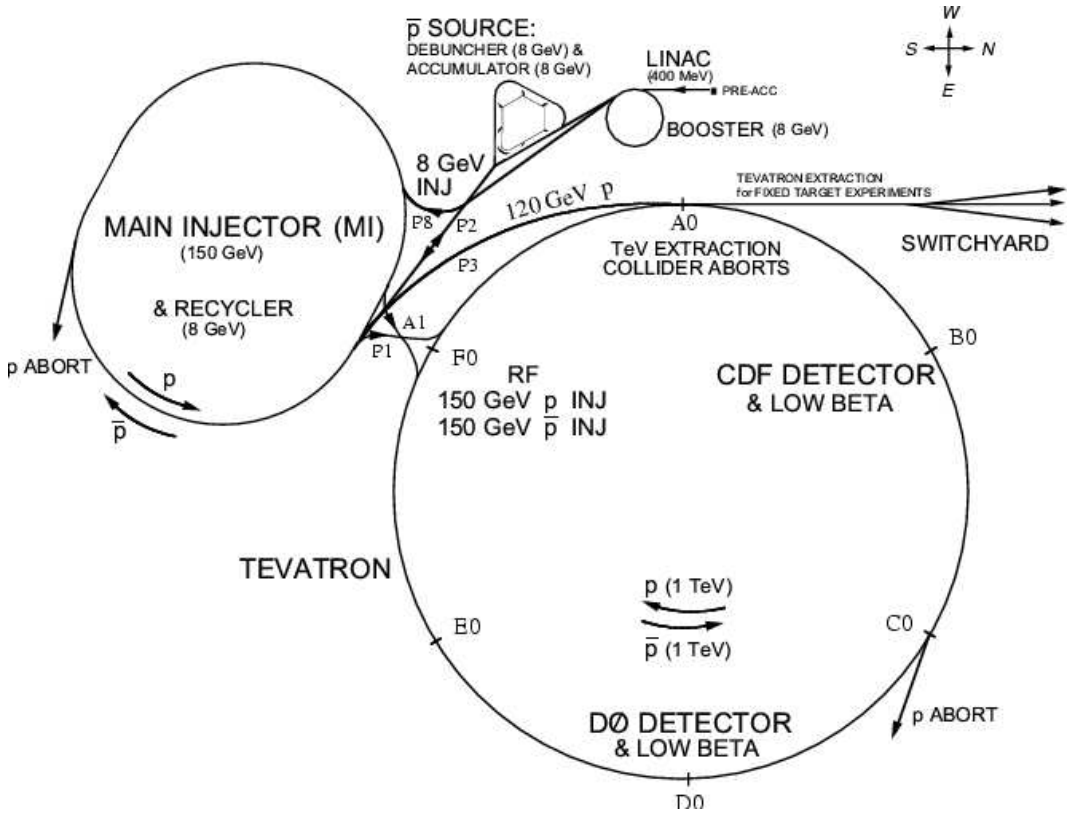


Figure 2.1: A schematic drawing of the upgraded Tevatron accelerator complex. The main injector is the most ostensible addition of the upgrade.

The proton and antiproton beams are accelerated to the maximum energy in several stages. At the earliest stage of beam preparation, negative hydrogen ions are first accelerated to 750 keV by an electrostatic Cockroft-Walton accelerator. Then they are further boosted to 400 MeV by a 150 m long linear accelerator (the linac). In the next stage, the electrons are stripped from the ions and the resulting protons

enter a 75 m radius synchrotron, the Booster, which accelerates the protons to 8 GeV.

Protons coming from the Booster are injected into the Main Injector [18]. As an important addition to the RunI Tevatron, the Main Injector is a rapid cycling synchrotron in a tunnel separate from that of the Tevatron. Its circumference is half that of the Tevatron, and it accelerates the proton beams to 120 GeV. Then the proton beams can be either extracted to the anti-proton production target or to an external physics target. For injection of either protons or anti-protons into the Tevatron, the Main Injector ramps to 150 GeV.

Anti-protons are produced by hitting a nickel target with proton beams extracted from the Main Injector. The debris from those collisions contain antiprotons. Immediately following the target, there is a lithium "lens", a cylindrical piece of lithium through which a large current (0.5 MA) can be passed. The current generates an azimuthal magnetic feld that focuses negatively-charged particles passing through its center. Following the lens, there is a bending magnet that selects negatively-charged particles with energy of 8 GeV, and transports them to the Debuncher.

The Debuncher is a storage ring at 8 GeV in which antiprotons are rotated in phase space from a configuration with a small time spread and large momentum spread to one with large time spread and small momentum spread (this process is called "debunching"). From the debuncher antiprotons are transferred to the accumulator ring also at 8 GeV. There they are cooled, and transferred to the 8 GeV Recycler ring before acceleration in the Main Injector, and eventually in the Tevatron, to full energy.

The Tevatron is at the final stage in the acceleration chain. It is a synchrotron with superconducting magnets, housed in a tunnel with circumference of around

4 miles. It is capable of accelerating the beam to an energy of 1000 GeV. In the collider mode, the Tevatron is filled with 36 bunches of protons and 36 bunches of antiprotons, travelling in opposite directions. The proton and anti-proton beams are then accelerated to the maximum energy and brought into collision at the B0 and D0 locations, where the CDF and DØ Detectors are located, respectively.

2.2 The DØ Detector

The DØ Detector is a general-purpose detector aimed at studying high-p_T physics at the Tevatron. Details of the RunI DØ Detector can be found in [19]. Figure 2.2 shows the schematic structure of the RunII DØ Detector. It consists of the silicon vertex detector, the scintillating fiber tracker, the preshower detectors, the calorimeter system, and the muon system. In the following sections each subdetector system will be discussed. But first let us establish the coordinate system which will be used throughout the analysis, and define some basic concepts: luminosity and cross section.

2.3 Coordinate System

The Cartesian coordinate system of the DØ Detector has the z-axis along the direction of the proton beam, and the y-axis pointing upward. The x-axis is then fixed by the right hand rule.

Based on the Cartesian coordinates, the cylindrical coordinates r and ϕ are defined as $r = \sqrt{x^2 + y^2}$ and $\tan\phi = x/y$. The polar angle θ is measured relative to the positive z direction.

The rapidity is a variable that describes the kinematic properties of particles. It

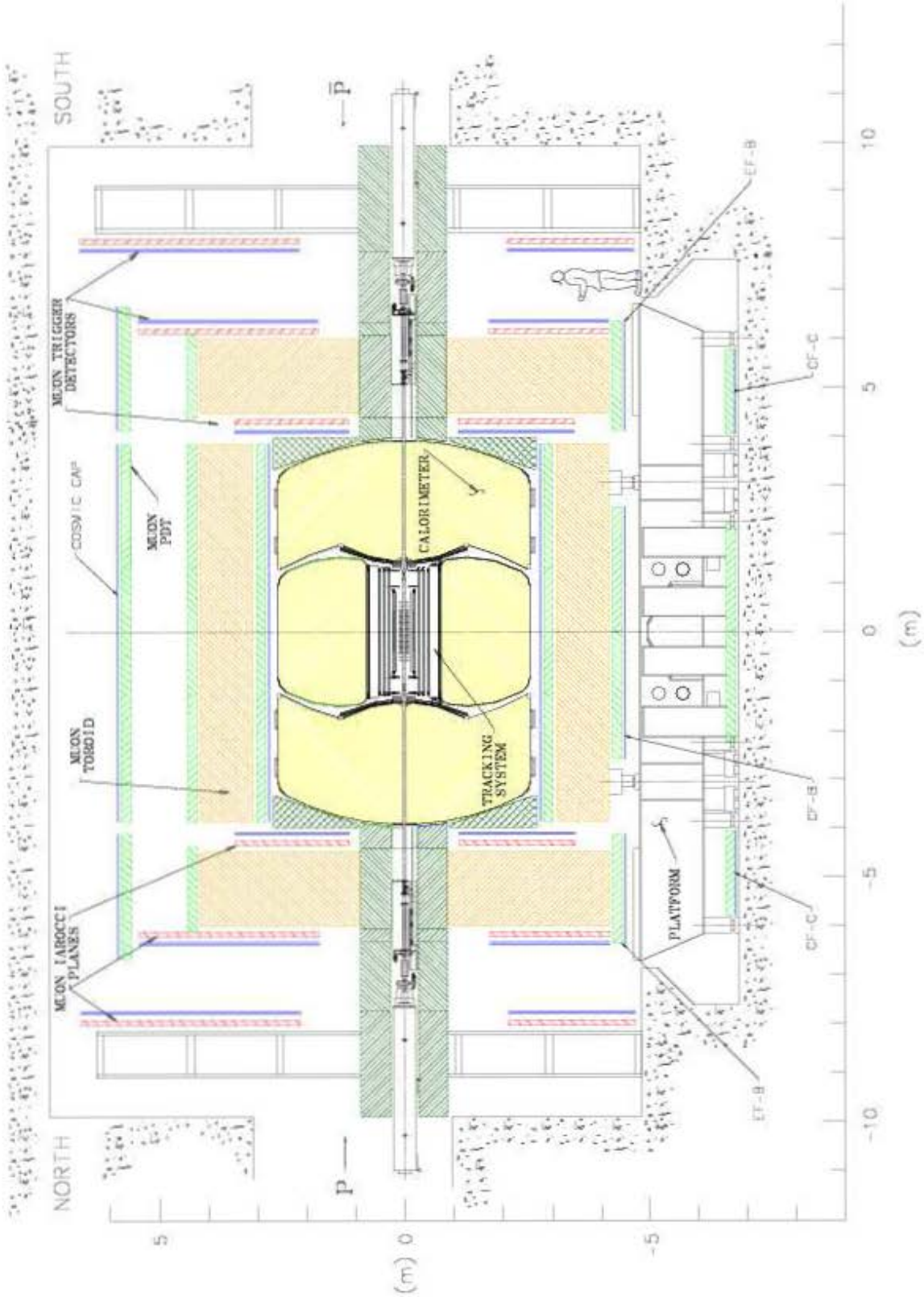


Figure 2.2: A side view of the upgraded DØ Detector.

is defined as:

$$y = \frac{1}{2} \ln \frac{E + p_L}{E - p_L}, \tag{2.1}$$

where p_L is the longitudinal momentum along the positive z direction, and E is the particle energy. Frequently, the simpler variable pseudorapidity is used. It is a function of the polar angle θ:

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

The pseudorapidity approximates the rapidity when the particle's velocity is in the relativistic regime.

2.4 Luminosity and Cross Section

In high-energy physics experiments, we are often interested in measuring the production rate of a process, and comparing it with theoretical predictions. This production rate is expressed in terms of the luminosity (L) of the machine, and the cross section (σ) of the process.

The luminosity is the product of incident particle flux (particles per unit time) with mean target density (particles per unit area), and is expressed in the unit of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Normally the luminosity varies with time, and we define the integrated luminosity ($\mathcal{L}$) by integrating the luminosity with time:

$$\mathcal{L} = \int L dt \tag{2.3}$$

The cross section is a Lorentz invariant quantity that measures the probability of interactions in a two-particle initial state. It has the dimension of area, usually with unit cm^2 or barn (1 barn $\equiv 10^{-24}\mathrm{cm}^2$). Given the integrated luminosity and the cross section for a particle interaction, the expected number of events is given

by:

$$N_{total} = \mathcal{L} \times \sigma. \tag{2.4}$$

In experiments, the number of observed events is usually smaller than N_{total} due to such factors as the geometrical acceptance of the detector, trigger efficiencies, reconstruction efficiencies, selection criteria, etc. In the data analysis, the total efficiency for a particle interaction is defined as the probability that an event from that interaction is observed, and is often denoted as ε. Then the expected number of observed events is given by:

$$N = \varepsilon \times N_{total} = \varepsilon \times \mathcal{L} \times \sigma. \tag{2.5}$$

In the equation, $\mathcal{L}$ and σ reflect intrinsic properties of the accelerator machine and the underlying physics process, respectively, while ε is determined by the kinematic properties of the physics process, the capability of the detector, and the selection criteria used in the analysis.

2.5 Central Tracking System

The new central tracking system is a key component of the DØ upgrade. A cross-sectional view of the tracking system is shown in Figure 2.3. It consists of the Silicon Microvertex Tracker (SMT), and the Central Fiber Tracker (CFT). A 2-Tesla central magnetic field is provided by a superconducting solenoid surrounding the Fiber Tracker.

2.5.1 Silicon Microvertex Tracker

Closest to the beams is the Silicon Microvertex Tracker (SMT) [21]. The basic design philosophy of the DØ Silicon Tracker is to maintain track and vertex reconstruction over the DØ full pseudorapidity acceptance. Ideally, if the Tevatron

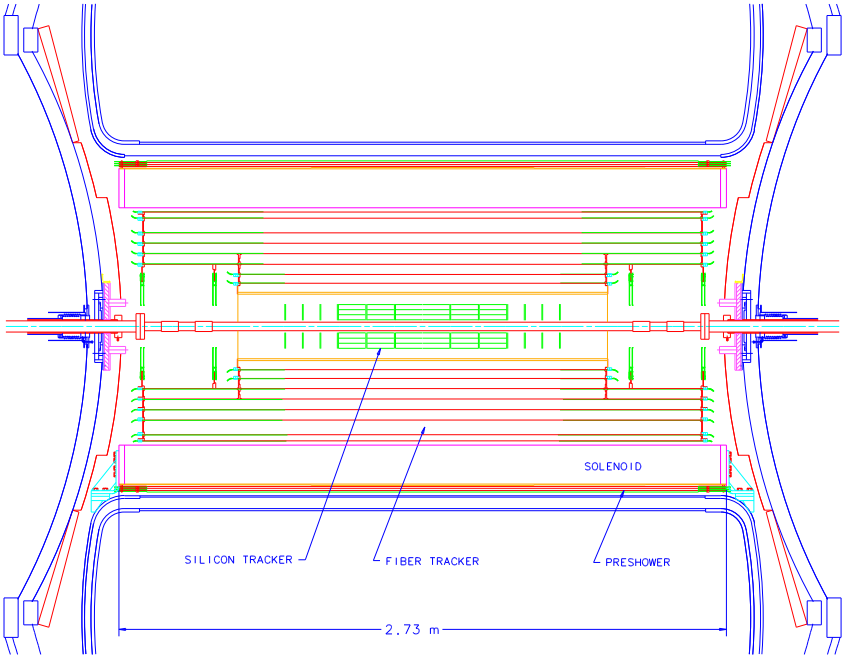


Figure 2.3: A side view of the upgraded DØ Central Tracking System

Collider's interaction point (IP) is localized to the center point of the detector, the detector would adopt a roughly spherical geometry to allow all tracks to intersect the detector planes at approximately normal incidence and provide optimal resolution. However, the upgraded Tevatron in RunII is expected to have an interaction region length of $\sigma_z \approx 25$ cm. Such situation leads to a baseline design with interspersed barrels and disks. The barrel detectors measure the $r - \phi$ coordinates and the disk detectors measure the $r - z$ coordinates as well as $r - \phi$.

The interspersed barrel-disk design is shown in Figure 2.4. The overall geometry contains:

- 6 barrel segments,
- 4 detector layers per barrel,
- 12 F (small-diameter) double-sided disks with a 30° stereo angle,
- 4 H (large-diameter) single-sided disks covering both ends with a 15° stereo angle.

Layers 1 (innermost) and 3 are divided into the central and end regions, and consist of: double-sided detectors (axial and 90° z-strips) in the central 4 barrel segments; and single-sided detectors (axial strips) in the outmost barrel segment at each end. Layers 2 and 4 are double-sided detectors (axial strips and 2° stereo strips). The total length for the barrel segments is 12 cm. Eight F disks are mounted in the 8 mm gaps between the barrel segments. Four more F disks are located at each end of the barrel. The H disks are located at $|z| \approx 100$ cm and 120 cm.

In the barrels, the basic mechanical unit is the ladder. Each ladder supports two detectors wire-bonded together, forming a 12 cm long unit. The ladders are mounted

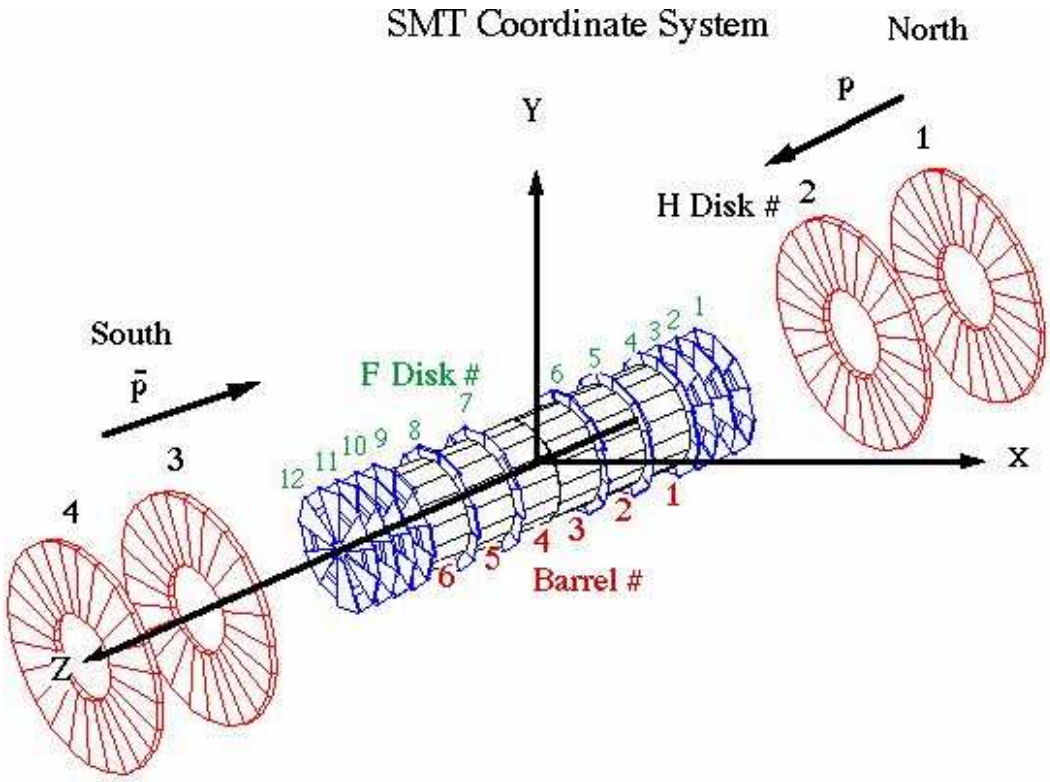


Figure 2.4: A three dimensional view of the Silicon Tracker.

on beryllium bulkheads, which serve as a support at both ends of the ladder and provide cooling at the readout end by means of an integrated coolant channel.

The individual channel count for layer 1, 2, 3, and 4 is about 46K, 83K, 92K, and 166K respectively. It is 258K for F disks, and 147K for H disks, resulting in a total channel count of around 793K. Signals from SMT channels are read out using SVX-IIe chips, which perform signal pipelining and digitization. SVX IIe chips are mounted on a kapton high density circuit (HDI).

The silicon detector currently installed in the DØ Detector is expected to survive $\approx$2 fb^{-1} of radiation. A new silicon detector will be installed for Run IIb operation. More information about the upgraded silicon detector can be found in Reference [23].

2.5.2 Central Fiber Tracker

The Cetral Fiber Tracker (CFT) [24], also called the scintillating fiber tracker, directly surrounds the silicon vertex detector and covers the central pseudorapidity region. The fiber tracker serves two main purposes. First, with the silicon vertex detector, the tracker enables reconstruction and momentum measurement for all charged particles within the range $|\eta| < 2.0$. Second, it provides fast, "Level 1" track triggering (see Section 2.11.1) within the range $|\eta| < 1.6$.

The CFT is shown in Figure 2.5. The basic element of the fiber tracker is the multi-clad scintillating fiber of 830 μm in diameter and 1.7 $\sim$ 2.6 m in axial length. A total of 76,800 scintillating fibers are mounted on eight concentric cylinders. Each cylinder supports a doublet layer of axial fibers. Counting from the inner most one, cylinders with odd-numbered radii support an additional doublet u layer that are oriented at a 2.0° to 3.0° stereo angle; cylinders with even-numbered radii support an additional doublet v layer that are oriented at a -2.0° to -3.0° stereo angle.

At either end of the CFT detector, the scintillating fibers are connected to clear fiber waveguides with identical diameters and typically 8 $\sim$ 10 meters in length. The clear fibers are grouped into 256-channel waveguide bundles and transport the optical signals to highly sensitive arsenic-doped avalanche photodiodes, known as visible light phton counters (VLPCs) [25]. Here the optical signals are converted into electrical signals and pipelined to SVX chips for digitization.

2.6 The Preshower Detectors

The preshower detectors are important additions to the DØ Detector. The system includes the Central Preshower Detector (CPS) [26] in the central region, and the Forward Preshower Detector (FPS) [28] in the endcaps.

The CPS is designed to meet several purposes: aid electron identification and triggering, and correct electromagnetic energy for the effects of the solenoid.

A drawing of the CPS is shown in Figure 2.5. The CPS is placed in the 51 mm gap between the solenoid and the central calorimeter cryostat at a radius of 72 cm, and covers the pseudorapidity region $-1.2 < \eta < 1.2$. It consists of three layers of scintillating strips with wavelength-shifting (WLS) fiber readout. Counting from the detector inside out, the three layers are referred to as the x, u, and v layers. The x-layer strips are arranged in the axial direction, while the u- and v- layer strips are arranged in stereo angles of 23° and -23°, respectively.

Each CPS layer is segmented into five supersectors, with each supersector containing 256 scintillating strips. The WLS fibers are split at $z = 0$ and are connected to clear fiber waveguides at both ends of the detector. The splitting effectively halves the occupancy of each channel and makes the detector less vulnerable to high rates. This results in a total of 1280 channels per layer per end, and a total of 7680 channels

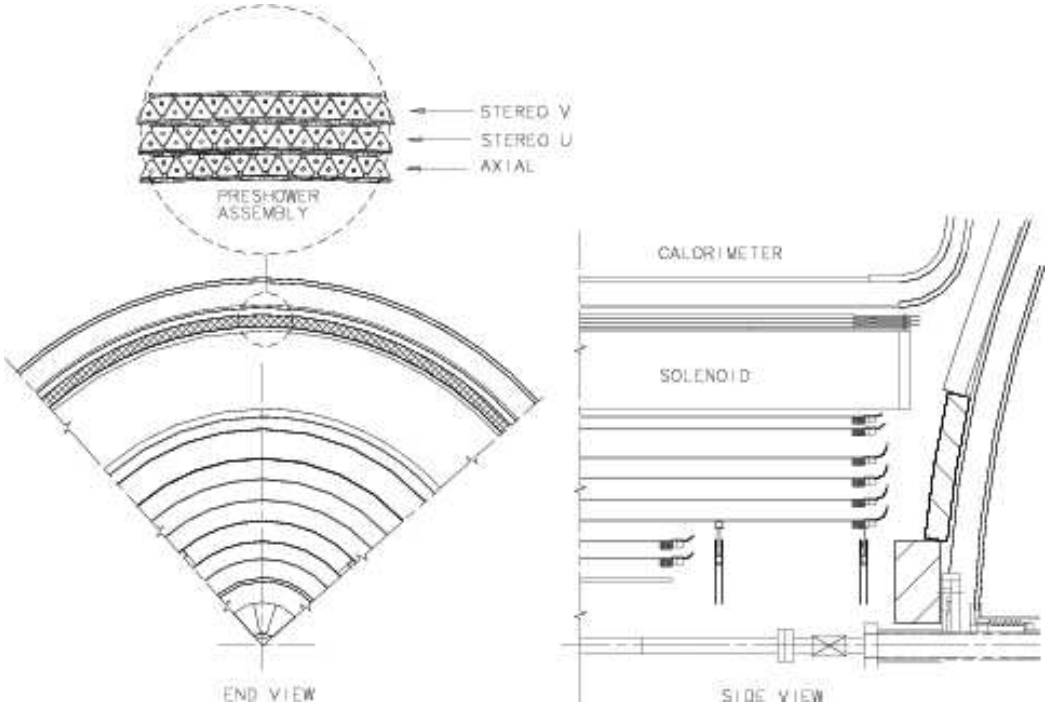


Figure 2.5: A schematic drawing of the central region of the DØ Detector. The eight-layer CFT is shown inside the solenoid. The Central Preshower Detector (CPS) is located in the thin gap between the solenoid and the calorimeter. A cross-sectional view is shown on the left, where the three layers of CPS is visible in a magnified view.

in the entire CPS. CPS strip signals are processed in a similar way as CFT fibers. Signals are relayed from clear fibers to the VLPCs on the detector platform, and digitized by SVX chips.

The Forward Preshower Detector, like the CPS, is intended to enhance electron identification in the forward region, by making precision position measurements of particle trajectories using dE/dx and showering information collected just upstream of the calorimeter.

The FPS consists of the north part and the south part, with each part mounted on the inner face of each of the End Calorimeter (EC) cryostats (see Figure 2.6), and covers the pseudorapidity range $1.4 < |\eta| < 2.5$. As the figure shows, two active scintillator planes constitute the FPS detector. A layer of lead absorber of two radiation length thick is sandwiched in between. Each scintillator plane consists of one u and one v sublayer.

The routing of FPS signals adopted the same design as for CPS. Signals from the WLS fibers are transferred to the clear fibers via connectors mounted at the end of each fiber-routing channel at the periphery of the detector. The clear fibers are routed down to the VLPCs on the platform where signals are converted into electric charge and digitized subsequently.

The design and performance of the FPS has been extensively discussed in [29].

2.7 Calorimeter System

At RunI the DØ collaboration relied heavily on the calorimeter to identify electrons, photons, and jets, due to the lack of a central magnetic field. The RunII upgrade of the calorimeter [30] mainly focused on its front-end electronics [31]. The calorimeter itself is largely preserved during the upgrade.

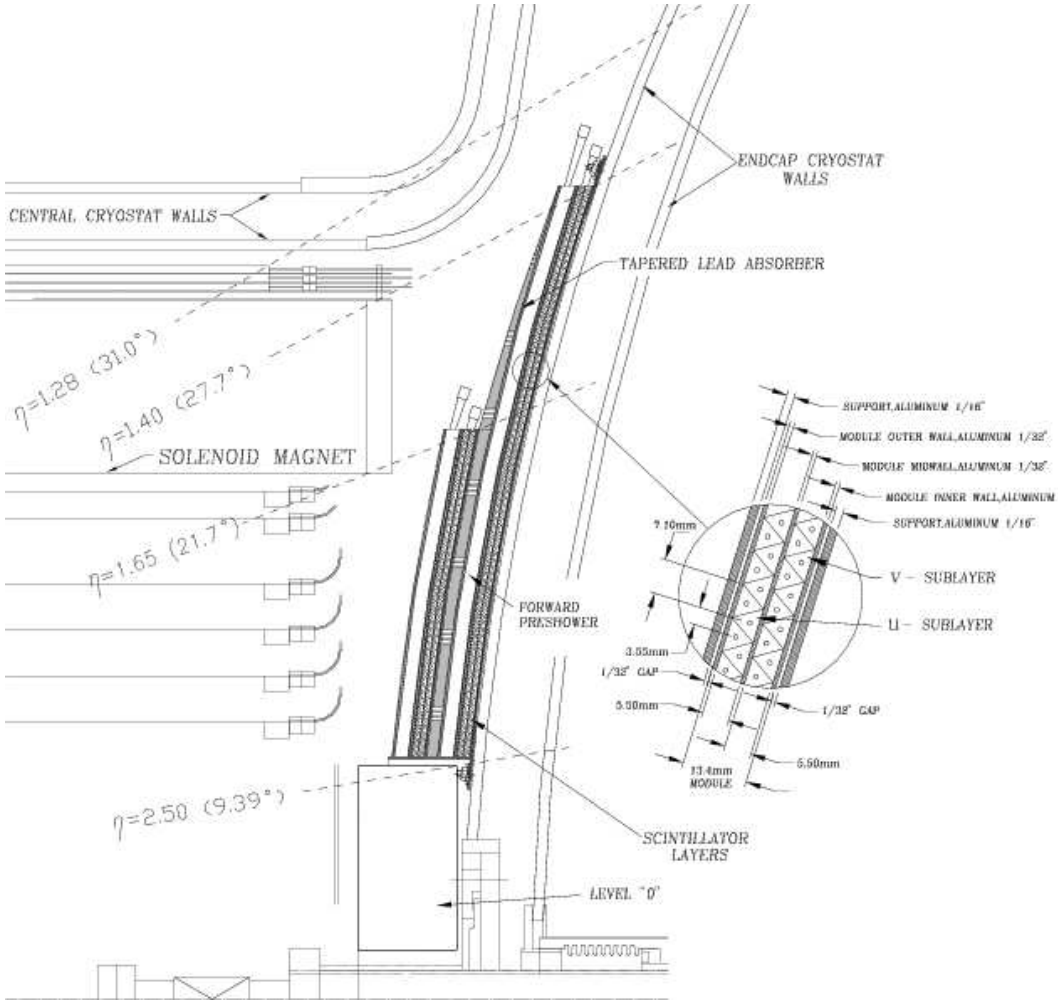


Figure 2.6: A schematic drawing of the Forward Preshower Detector (FPS) mounted on one of the two Endcap Calorimeter cryostats. The structure of the scintillation plane is seen in the magnified view.

The purpose of the calorimeter is to measure the energy and position of incident particles, and to distinguish particle types by their energy deposition patterns. The principle of the calorimeter is to have incident particles to interact with a large detector mass to produce a cascade of secondary particles, forming a cluster of low energy particles. Since most of the low-energy particles carry charges, they interact with certain parts of the calorimeter material, and the energy of the incident particle can be deduced. Within a detector, the calorimeter is the only subsystem to measure neutral particles produced in high-energy collisions.

One can classify calorimeters into two types: homogeneous and heterogeneous. In a homogeneous calorimter, the functions of passive particle absorption and active signal generation and readout are combined in a single material; while in a heterogeneous calorimeter (also called a sampling calorimeter), the functions of passive particle absorption and active signal generation and readout are realized by different materials. The DØ calorimeter is of the latter type: layers of dense, inert absorber are interleaved with layers of active medium that are sensitive to particles passing through them. When a particle enters the absorber medium, it interacts with the material and produces a cluster of low-energy particles called showers. As shower particles pass through active layers, their energy is deduced by the interaction. Since most of the incident energy is absorbed by the inert material, only a fraction of it is detected by the active medium. This fraction is known as the sampling fraction.

A cut-away view of the calorimeter system is shown in Figure 2.7. There are three cryostat systems for the calorimeter: one central calorimeter (CC), and two endcap calorimeters (EC). The CC provides coverage in the pseudorapidity range $|\eta| < 1.2$, and the EC spans the range of $1.3 < |\eta| < 4.0$. Each cryostat system is further divided into three radial layers: the electromagnetic (EM) calorimeter, the

fine hadronic (FH) calorimeter, and the coarse hadronic (CH) calorimeter.

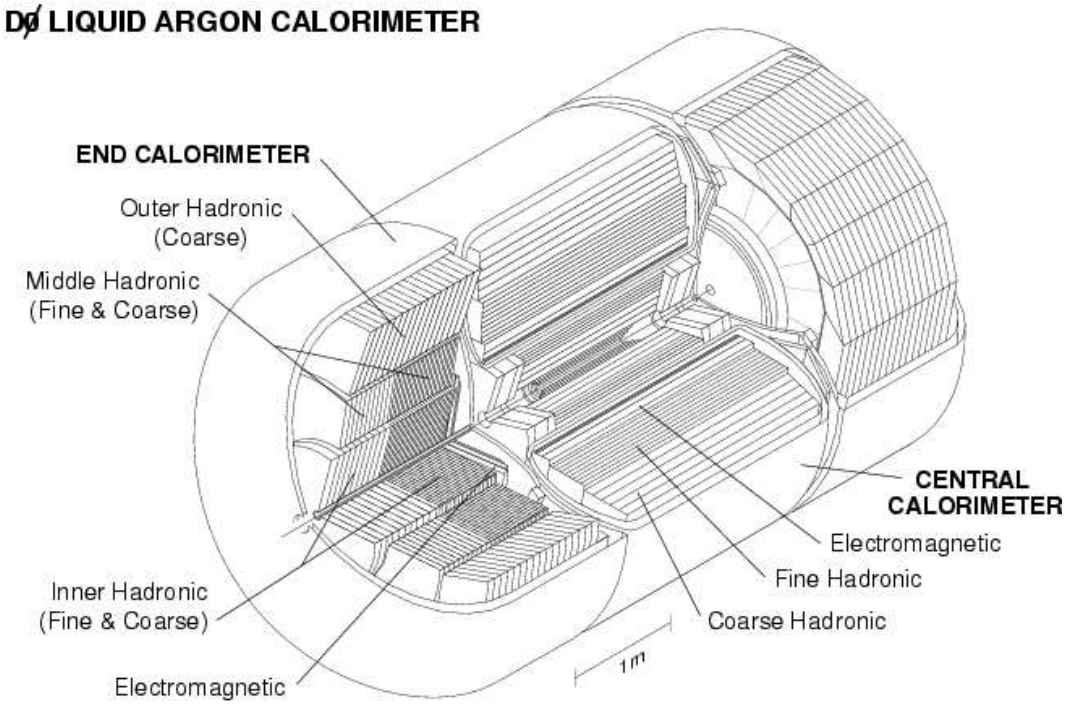


Figure 2.7: A schematic drawing of the calorimeter system in DØ. From the inside of the detector out, each of the three calorimeter cryostats consists of the following three modules: the electromagnetic calorimeter, the fine hadronic calorimeter, and the coarse hadronic calorimter.

A calorimeter cell, which is the basic unit of the calorimeter, is shown in Figure 2.8. The DØ calorimeter uses liquid argon (LAr) as the active medium, and uses plates of uranium (3 mm thick), depleted uranium (DU) mixed with 1.7% niobium (Nb) (6 mm thick), and copper/steel (46.5 mm thick) as the absorber in the EM, FH, and CH modules, respectively.

When a high-energy electron passes through a material with a high atomic number, the primary mechanism by which it loses energy is through Bremsstrahlung, in which a charged particle interacts with the Coulomb field around a nucleus and emits an energetic photon. A high-energy photon, on the other hand, will interact

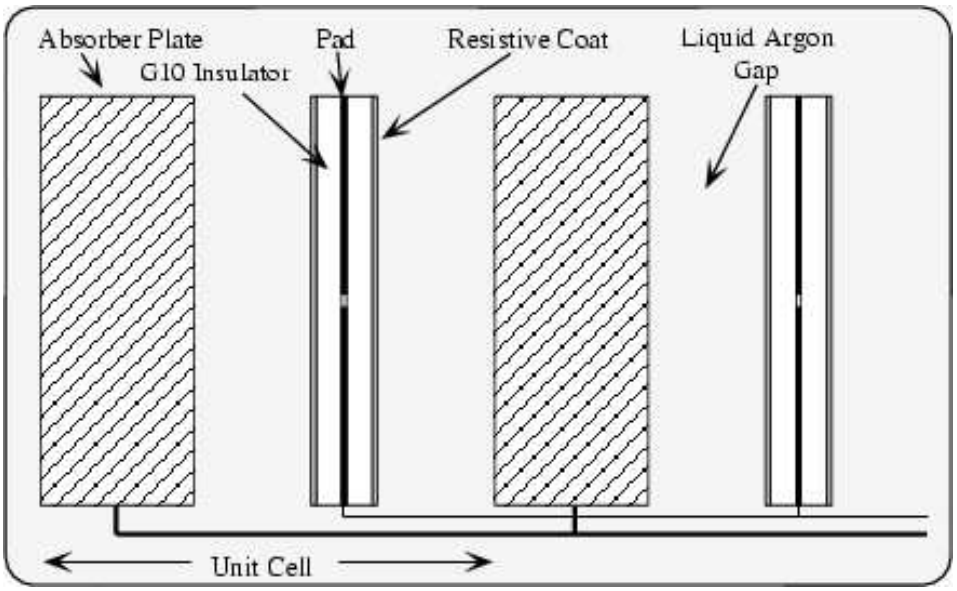


Figure 2.8: A schematic drawing of a calorimeter cell in the DØ Detector.

predominately via pair production, in which a photon converts into an electron-positron pair in the vicinity of a nucleus. The particles emitted in these interactions can themselves undergo Bremsstrahlung or pair production. Thus, an energetic electron or photon passing through a dense material will result in a shower of secondary electrons, positrons, and photons. This process is called an electromagnetic shower. The shower will continue to develop until all the secondaries have sufficiently low energies that other energy loss mechanisms (mostly ionization) become important. The rate at which an incident electromagnetic particle loses energy is a constant of the material, and is usually specified as the radiation length X_0:

$$\frac{dE}{E} = -\frac{dx}{X_0}. \tag{2.6}$$

As an example of a typical value, the radiation length for uranium is about 3.2 mm.

Hadronic particles also generate showers, but the showers are produced by different mechanisms from the electromagnetic showers and have different patterns.

Hadrons lose energy primarily through inelastic collisions with atomic nuclei. These collisions produce secondary hadrons, which can in turn undergo inelastic collisions. This process is called a hadronic shower, and it continues until all particles have either been stopped by ionization losses or absorbed by nuclear processes. As a result, hadronic showers are much more extended in space than electromagnetic showers of similar energy. The scale for the size of these showers is given by the nuclear interaction length λ for the material. For uranium, $\lambda \approx 10.5$ cm.

The EM calorimeter is designed primarily to measure the electron and photon energy. It consists of four layers of cells, and is 21 radiation length ($21X_0$) deep in total. The first two layers measure the initial longitudinal development of a shower, and is about $2X_0$ thick each in the CC (around $2.3X_0$ deep in the EC). The maximum of an EM shower is expected to occur within the third layer, which is $7X_0$ thick. The final layer contains the shower within an additional $10X_0$.

The hadronic calorimeters are 7-9 nuclear interaction length deep, and are divided into four (in CC) or five (in EC) layers. In the CC region, the FH modules make up the first three layers with depth of 0.9λ, 1.0λ, and 1.3λ. The CH make up the fourth layer and is 3.2λ deep.

The transverse segmentation of all the layers in the electromagnetic and hadronic calorimeters is $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$, except for the third layer in the electromagnetic calorimeter. Here since the maximum of an electromagnetic shower is expected, it has finer segmentation with a cell size of $\Delta\eta \times \Delta\phi \approx 0.05 \times 0.05$. Segmentation of the pseudorapidity range in the $\phi - z$ plane is shown in Figure 2.9.

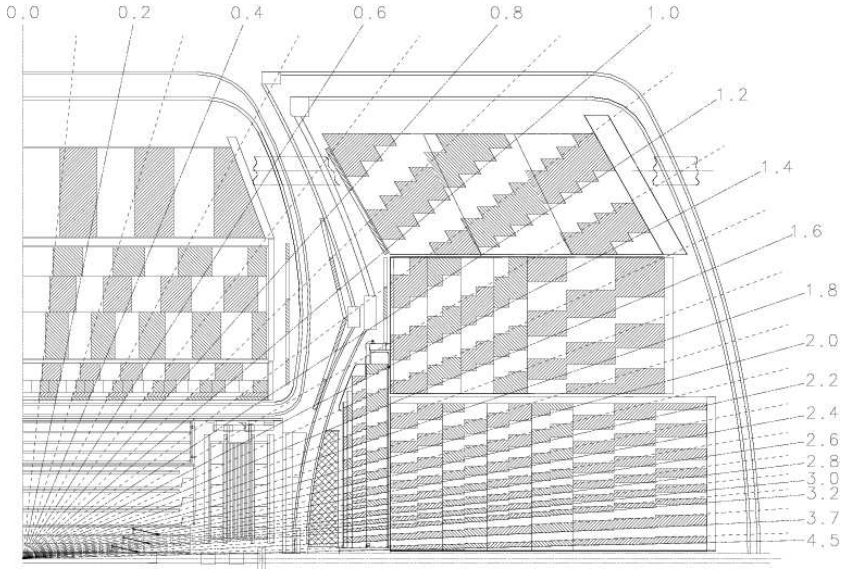


Figure 2.9: The longitudinal view of one quadrant of the DØ calorimeter system. The pseudorapidity segmentation is shown by means of the pseudorapidity lines originating from the center of the detector. The shading pattern indicates distinct calorimeter cells.

2.8 The Intercryostat Detector

The Intercryostat Detector (ICD) plays an important role in DØ calorimetry. For particles traversing the regions in the calorimeters in the range $1.1 < |\eta| < 1.4$, they pass through several layers of central and endcap calorimeter support structures that contribute to dead regions with no or minimal energy sampling. To allow measurement in this area, an Intercryostat Detector is mounted on each face of an end-cryostat wall. The detector consists of a single layer array of scintillating tiles as the active element. The light signals, picked up by wavelength shifting fibers in the tiles, are transported to the location of the photodetection readout via clear fiber ribbon cables.

This scintillating tile array not only allows for the sampling of particles exiting

the solenoid and the CC cryostat, but also enhances the measurement of hadronic jet energy. Furthermore, it allows calculating the missing transverse energy ($\not{E}_T$).

2.9 Muon System

Since Muons are minimum ionization particles which only deposit a small amount of energy in the calorimeter, the detection and measurement of muons rely on the muon system, the outmost component of the DØ detector, also the largest subdetector system. The muon system is divided into two components: the Wide Angle Muon System (WAMUS) in the central region, and the Forward Angle Muon System (FAMUS) in the forward region. During the DØ upgrade, the central muon system was supplemented with additional scintillator layers for triggering cosmic ray rejection, and low momentum muon measurements [33]. The forward muon system was completely replaced with scintillator pixels and miminum drift tube chambers [34] [35]. New shielding has been added to decrease background rates. The muon trigger has been redone to accomodate the high trigger rate and increased number of interactions per beam crossing. Detailed information about the upgraded muon system is available in [36].

The DØ muon system is shown in Figure 2.2. Both the central and forward systems consist of three detection layers, labeled A, B, and C with increasing distance from the origin of the DØ coordinate system. A 2-Tesla Toroidal iron magnet is sandwiched between the A-layer and the B-layer. The central (CF) magnet covers the pseudorapidity range $|\eta| < 1$, while the two end (EF) magnets cover the pseudorapidity range $1 < |\eta| < 2.5$. The magnets bend the tracks of muons passing through the system. By measuring the angular bending, the charge and momentum of the muon can be deduced.

The RunI central muon system consists of the CF magnet, 94 proportional drift tube (PDT) chambers in the three WAMUS layers, and 240 scintillation counters which constitute the "Cosmic Cap". The WAMUS chambers provide measurements for muons traversing the CF magnet or the outer edge of the EF magnets. The chambers are constructed from extruded aluminum tubes, with wires oriented along the toroid magnetic field direction to provide position measurement. The Cosmic Cap counters cover the top and sides of the central muon system.

In the RunII central muon system, the coverage of the Cosmic Cap was complemented by 132 additional scintillation counters underneath the toroid magnet forming a "Cosmic Bottom". In addition, a layer of "$A-\phi$" scintillation counters is placed in between the calorimeter system and WAMUS A-layer PDT, to provide a scintillation counter tag for low-p_T muons and an improved spatial match with the CFT at the hardware trigger level.

The forward muon system consists of the following major parts:

- 3 layers of Iarocci-type mini-drift tubes (MDTs) for muon track reconstruction;
- 3 layers of scintillation counters for triggering on events with muons [35];
- shielding around the beam pipe from calorimeter to the accelerator tunnel to reduce trigger rates, fake track reconstruciton, and aging of detectors.

A schematic drawing of a MDT layer is shown in Figure 2.10. The tubes are oriented along magnetic field lines. The three MDT layers are used for muon identification, to complement the central tracking detector for muon momentum measurement, and to confirm muon triggers in the forward region. The track position will be determined by drift time measurement with a coordinate accurary of $\sigma_x \approx 0.7$ mm.

There are three layers of forward muon scintillation counters (FMSC) in the forward region, labeled FMSC A-, B-, and C- layers, with $\Delta\phi = 4.5°$ and $\Delta\eta = 0.1$ segmentations shown in Figure 2.11. The pixel size is designed to trigger efficiently on muons down to $p_T = 3$ GeV. The three counter layers serves the similar function as the central $A - \phi$ counter layer does. They provide a coincident hit in all three pixel layers in conjunction with a CFT/FPS forward trigger-track match.

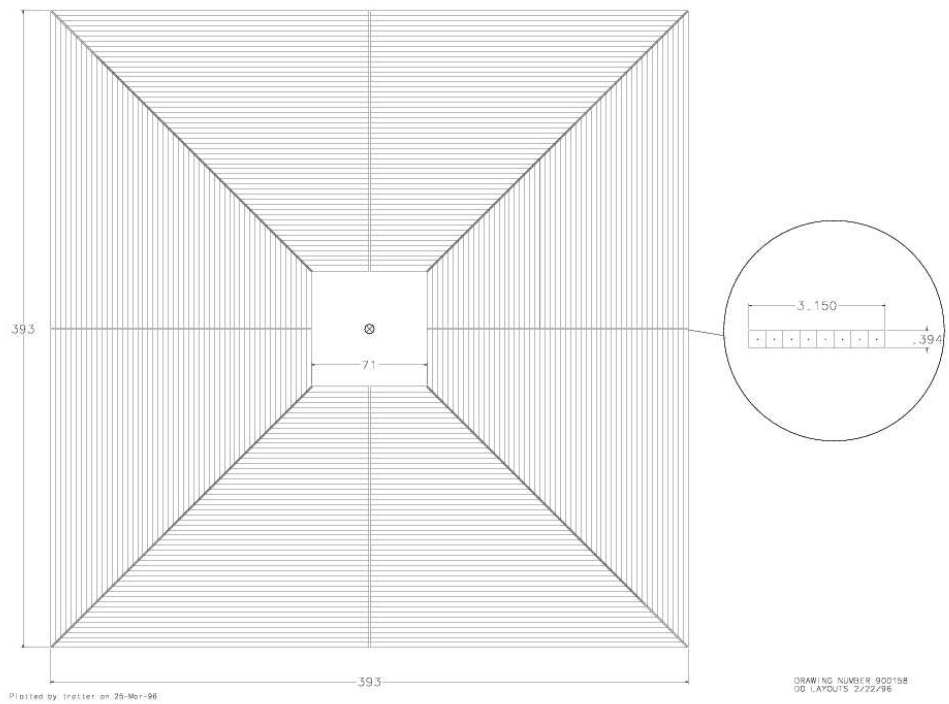


Figure 2.10: A schematic drawing of one MDT layer in the forward muon system.

2.10 Luminosity Monitor

The DØ RunII Luminosity Monitor [37] consists of two arrays of plastic scintillation counters located on the inside face of the EC calorimeters, at a distance of ≈135 cm from the nominal interaction point. The detector covers the pseudorapidity

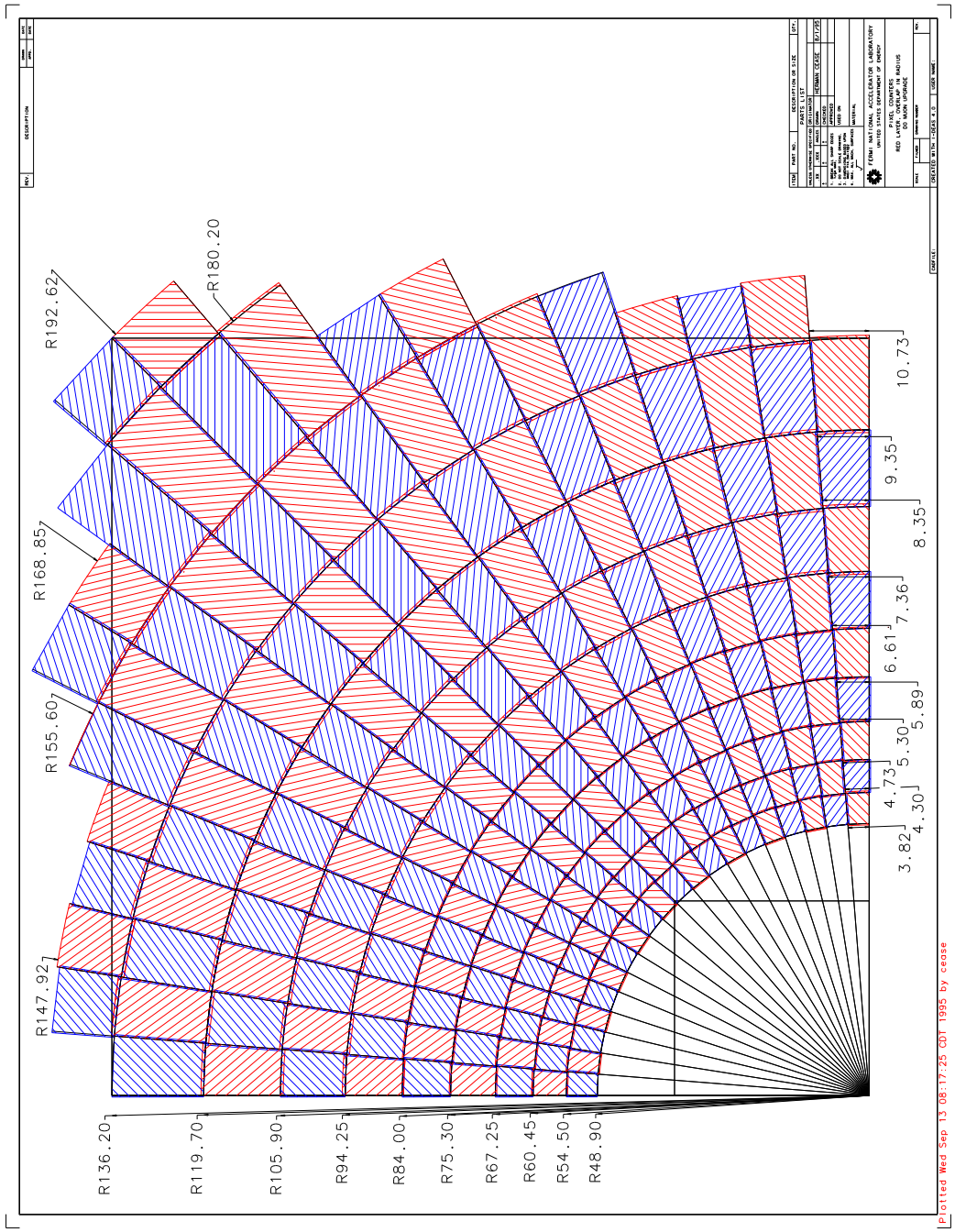


Figure 2.11: One quadrant of the forward muon scintillation counters, showing the $r - \phi$ segmentation.

region of $2.7 < |\eta| < 4.4$.

Whenever an interaction occurs, the remnants of the incident proton and antiproton give a pair of time correlated hits in the scintillators. The difference in hit times provides a fast measurement of the z vertex position with a precision of $\approx$6 cm. Scalers record the number of time correlated pairs of hits during the live time of the data acquisition system in order to calculate the exposed luminosity. The scalers are read out every minute and a unique identifier called luminosity block number (LBN) is assigned. This number flags every event recorded during the same period of time, so that the luminosity can be retrieved in offline data analysis.

2.11 The Trigger System

At a hadron collider, elastic or inelastic hadron collisions happen at every bunch crossing. It is not economical or even feasible to record every collision, due to hardware bandwidth limits and storage device constraints. So it is imperative to determine the quality and physical importance of an event right after a bunch crossing, and store the physically interesting events to permanent storage. This type of selection is called triggering.

The high luminosity ($L = 2\times 10^{32}$ cm^{-2}s^{-1}), high bunch crossing rate (396 ns or 132 ns bunch crossing time) environment in RunII demands significant improvement in the DØ trigger system over RunI. The upgrade of the trigger system includes a new trigger framework, and several new detectors (CFT, CPS, FPS, and the muon detectors). The new trigger framework consists of three levels of triggers: Level 1 (L1), Level 2 (L2), and Level 3 (L3). Eight event buffers are added between adjacent trigger stages to de-randomize the Poisson-distributed arrival times of events, and to decrease deadtime due to pile-up effects. The expected RunII trigger accept rate

limits are 5-120kHz, 1000 Hz, and 50 Hz at L1, L2, and L3, respectively.

2.11.1 The Level 1 Triggers

The upgraded L1 trigger is a hardware trigger system. It consists of the L1 Trigger Framework (L1FW) [38] and the Level 1 Trigger Subsystems, which include CFT, CPS, FPS, the calorimeter, and the muon scintillators and tracking chambers. The calorimeter, fiber tracker, and preshower detectors provide electron triggering for $|\eta| < 2.5$. The fiber tracker and the muon systems cover the region $|\eta| < 2.0$.

Each L1 trigger subsystem processes subdetector-specific information based on a number of trigger terms, and for every beam crossing produces And-Or Input Terms to the L1 Trigger Framework. For example, a calorimeter trigger will require transverse energy above a threshold in one or several calorimeter trigger elements, or require global quantities such as total transverse energy, total energy, and missing transverse energy.

Using these Input Terms and other information about the readiness of the DAQ system to begin another acquisition cycle, the L1 Trigger Framework determines for each beam crossing whether the event should be rejected, or captured for further analysis in the L2 trigger system. For every beam crossing 128 separate sets of And-Or Term combinations are evaluated. The result is used to produce 128 separate L1 Specific Trigger Decisions. The global L1 trigger decision is the logical `.OR.` of the 128 Specific Trigger Decisions. A specific trigger decision is positive for a particle beam crossing when both all its And-Or requirements and additional conditions (the final specific trigger decision conditions) are satisfied.

2.11.2 The Level 2 Triggers

The L2 trigger [39] consists of a hardware framework and a separate set of processors. The L2 trigger is aimed at reducing the 10 kHz accept event rate by roughly a factor of ten within 100 μs. It consists of two stages, the preprocessor stage and the global processor stage, as shown in Figure 2.12. The preprocessor stage prepares L1 trigger information from each subdetector system for use in the second stage. The global processor stage combines different L1 trigger objects from different subdetectors, and makes the L2 trigger decision.

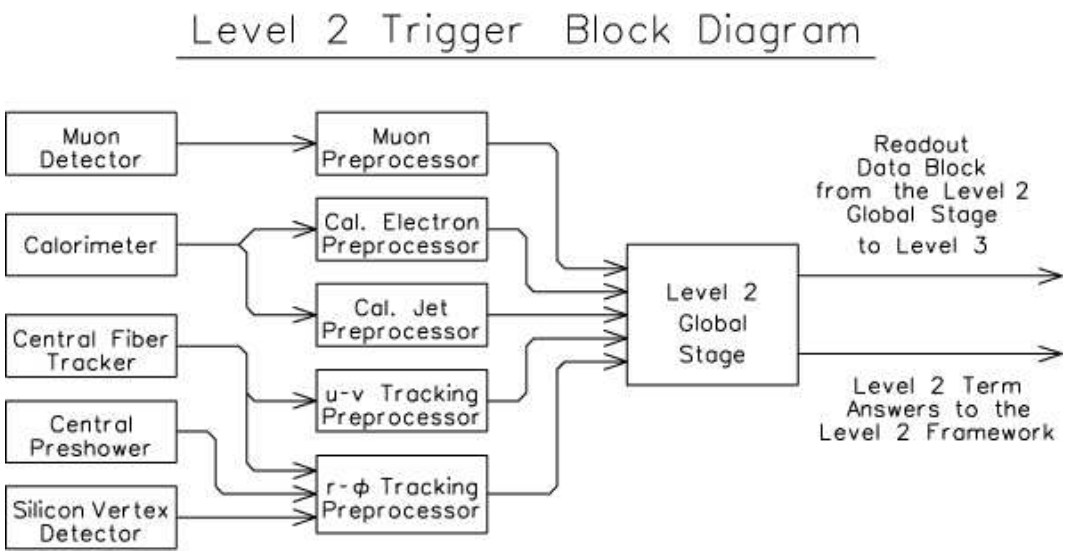


Figure 2.12: The Level-2 Trigger system structure.

The algorithm in the global processor stage is implemented by a fast Alpha processor on a VME card. It makes trigger decisions based on the objects found by the preprocessors. Such decisions are tailored separately for each fired L1 trigger bit. By either confirming or rejecting a L1 trigger bit, the L2 trigger produces a 128-bit L2 trigger bit mask. The event passes L2 if any of the 128 bits passes. The L2 decision is reported to the L2 hardware framework, and is used to guide L3 triggering.

2.11.3 The Level 3 Triggers

The DØ Level 3 system has an input event rate of 1 kHz and a 50 Hz accept rate with further possible increases in bandwidth. The Level 3 trigger uses a farm of standard, high performance commercial processors to run event filter algorithms which are written in high-level transportable code, with each filtering node running event building, event IO, and filtering software.

In the event building stage, a set of software packages (Level 3 tools) are run to unpack the raw data, identify energy clusters within each subdetector system, and then find physics objects by combining event information in different subdetectors. This event reconstruction process is done with limited data calibration due to timing constraints. The reconstructed event is analyzed by Level 3 filtering processes, which select events based on a set of criteria on physics objects. The Level 3 trigger produces a 128-bit trigger bit mask by either confirming or rejecting a Level 2 trigger bit.

2.11.4 Data Acquisition

The DØ RunII data acquisition (DAQ) system adopted an ethernet-based design using commercially available components [48]. A group of VME crates receives raw data from the detector. Each crate contains a Single Board Computer (SBC), which acts as the readout controller for event data. One of the SBC's hosts the Routing Master program which decides event data path from readout crates to Level 3 nodes. Each SBC uses VME to read event data into its main memory, and buffer the event data while waiting for routing information from the Routing Master.

The Trigger Framework crate delivers synchronization information via the Serial Command Link (SCL), which is received by the SBC hosting the Routing Master program. The same SBC also receives "buffer free" messages from the Level 3 farm.

The routing decisions by the Routing Master is transmitted to all the SBC's via ethernet. Then each SBC transmits event data via ethernet switches to the appropriate Level 3 node.

Events passing the Level 3 trigger will be processed by reconstruction programs on the FNAL processor farm system, a portion of which is dedicated to DØ and is capable of matching the data acquisition rate. Following the reconstruction stage, event data will be stored on a tightly coupled disk and robotic tape system, and made available for analysis by DØ collaborators.

CHAPTER III

EVENT RECONSTRUCTION

Raw data that pass the Level 3 triggers are written to tapes for offline analysis. It is difficult to extract meaningful physics directly out of the raw data, because they are essentially a dump of the content in memory modules in the readout crates. The raw data are processed by the DØ reconstruction program RECO, which consists of three data-processing stages. First, data in each subdetector sytem are unpacked, signals in readout channels are converted into "hits" of definite energies at specific spatial locations. Second, hits are analyzed in each subdetector system and are grouped to form primitive objects such as clusters in the calorimeter or tracks in the central tracker. Lastly, those primitive objects from different subdetectors are combined and analyzed, so that physics objects such as electrons, muons, and jets are reconstructed. This chapter gives brief explanations to the algorithms used to reconstruct different physics objects.

3.1 Electron Identification

Electrons are identified using tracking and calorimeter information. Several DØ software packages [41] [42] [40] are used in the offline electron reconstruction. First, an initial calorimeter cluster builder runs and provides a list of candidate clusters to

the EM candidate reconstruction program, EMReco. Possible algorithms that can be used to build these initial clusters are: the Nearest Neighbors algorithm, the Simple Cone algorithm, and the CellNN algorithm. The Simple Cone algorithm is used in this analysis.

During the initial cluster selection stage, EMReco calculates transverse energy (E_T), and EM fraction (f_{EM}) for each candidate. Here $f_{EM} \equiv E_{EM}/E_{tot}$, E_{EM} is the cluster energy in the EM calorimeter, and E_{tot} is the total calorimeter energy within the cone. If an initial cluster has an EM fraction greater than the EM fraction cut (by default 0.9) and $E_T > E_{Tmin}$ (by default 1.5 GeV), it is accepted to proceed to the next stage.

For each selected initial cluster, EMReco calculates its isolation (iso) as follows: find the largest p_T tower of the cluster, sum over the energy of towers within a circle centered at the largest p_T tower with a small radius and a large radius. Denote E^{tot}_{small} and E^{tot}_{large} as the energy in the small circle and the large circle, respectively, then the isolation is given as:

$$iso = \frac{(E^{tot}_{large} - E^{tot}_{small})}{E^{tot}_{large}} \tag{3.1}$$

If an initial cluster's isolation passes the cut (by default 0.2), the cluster is accepted for the next stage. Accepted initial clusters are seeds to build the final EM candidates. EMReco also looks for matching preshower clusters and matching tracks for EM candidates.

The above procedure utilizes such kinematic variables as EM fraction, E_T, and isolation to produce basic EM particle candidates. However, a photon, or a photon pair from π^0 decay, or a jet which fragments into a leading π^0 may pass the above requirements, and be misidentified as an electron. To minimize fake electron rate, additional criteria have been designed to enhance the selection of electrons. One

of the criteria is to require that the energy cluster match a central track, because photons do not leave a track in the central tracker. Another variable that is used by the top analysis, χ^2 [43], is described below.

3.1.1 H-Matrix Chi-Squared

As described in Section 2.7, the shower shape of an electron has a distinguished profile from that of a jet. A total of 41 or 8 variables are used to characterize these differences. They include the fraction of energy in layers 1, 2, and 4 of the EM calorimeter, the fraction of energy in each of the EM3 cells in a 6×6 square centered on the hottest tower, the logarithms of the cluster energy, and the vertex position. For a sample of N Monte Carlo electrons, a covariance matrix M is defined as

$$M_{ij} = \frac{1}{N} \sum_{n=1}^{N} (x_i^n - \overline{x_i})(x_j^n - \overline{x_j}) \tag{3.2}$$

where x_i^n is the ith variable of the nth electron, and $\overline{x_i}$ is the mean value of the ith variable from the N measurements [43]. The H-Matrix is defined as the inverse of the convariance matrix:

$$H_{ij} = M_{ij}^{-1} \tag{3.3}$$

For a subsequent measurement y, we define a χ^2 to quantify how consistent the shower shape is with the Monte Carlo samples:

$$\chi^2 = \sum_{i,j=1}^{41} (y_i - \overline{x_i}) H_{ij} (y_j - \overline{x_j}) \tag{3.4}$$

3.2 Muon Identification

Muons are identified in two steps using information from the muon system and the central tracker. The first step consists of finding muons identified in the muon system when a muon track is reconstructed with the following requirements:

- at least 1 wire hits in the A segment,
- at least 1 scintillator hit in the A segment,
- at least 2 wire hits in the BC segment,
- at least 1 scintillator hit in the BC segment.

A veto on cosmic muons is applied by requiring the time difference between scintillator hits in the B or C Layer and the A layer to be consistent with a muon coming from the interaction region. The muon tracks are then extended to the point of closest approach (PCA) to the beam and their parameters are compared with those of central tracks at PCA. For all central tracks within 1 radian in azimuthal and polar angle of a muon track at PCA, a global fit is performed. Only the best match is kept.

Muons are minimum ionizing particles (MIPs). It loses a fixed amount of energy per distance in the calorimeter. So a muon deposits most of its energy in the largest pieces of the calorimeter, the hadronic calorimeter. A set of energy depositions in the hadronic calorimeter, where the energies are consistent with a MIP rather than a hadronic shower or noise, contstitutes a muon signal. An algorithm is used to locate muons by searching the hadronic calorimeter for contiguous sets of cells via a histogramming technique [47]. Muon signatures in the calorimeter is only used for the purpose of measuring the identification efficiency.

3.3 Jet Identification

The QCD Theory (see Section 1.1.3) implies that color particles are confined by the strong force: quarks and gluons only exist in hadrons and cannot exist alone. This property leads to the creation of jets in high energy collisions. When an energetic

quark or gluon is created from a $p\overline{p}$ collision, and is moving away from other quarks or gluons, the potential of the strong force field between the departing partons grows linearly as the distance increases, until the energy is large enough to create a parton-antiparton pair out of the vacuum. Such processes take place as more particles are generated out of the vacuum, until the energy is too low to permit the creation of parton pairs. This hadronization process produces a group of hadrons, moving in the general direction of the original parton. These particles deposit a cluster of energy, called a jet, in the calorimeter. Since pions are the lightest hadrons, they form the majority of particles in a jet.

The jet finding algorithm consists of the following steps:

- As a first step, a list of preclusters are produced. There are several clustering algorithms available in the DØ jet reconstruction program, including the cone algorithm, simple cone algorithm, Improved Legacy Cone algorithm [45], and k_T algorithm. The Improved Legacy Cone algorithm is used in this thesis.

- A precluster that passes an energy threshold is selected from the list as a jet candidate. We initialize the jet direction along that of the precluster, and calculate the E_T-weighted (η, ϕ)-centroid using all towers within a cone of specific radius R. With the new centroid, we draw another cone of radius R and calculate the corresponding centroid again. This process is repeated until the centroid stabilizes. The resulting cluster forms a cone jet.

- The stable cone is compared to the cone jets already found in the algorithm. The E_T value from shared towers are calculated. If the shared E_T is greater than 50% of that of the jet with smaller E_T, the two jets are merged together. Otherwise, the two jets are split, and each shared tower is assigned to the

nearest jet.

- Another precluster is picked up in the remaining list, and the above steps are repeated.

This analysis uses jets reconstructed with a cone size of $R = 0.5$. We define the transverse energy E_T of a jet as:

$$E_T = \sqrt{E_x^2 + E_y^2}, \tag{3.5}$$

where E_x and E_y are calculated from all the towers within the jet cone:

$$E_x = \sum_i^{towers} E_x^i \tag{3.6}$$

$$E_y = \sum_i^{towers} E_y^i \tag{3.7}$$

The reconstructed jet energy is the sum of energies in all the calorimeter towers within the jet cone. So it clearly depends on the cone size. In reality, the reconstructed jet energy also depends on the response of the calorimeter, and an offset energy which is due to underlying events, electronic pile-up from previous $p\overline{p}$ crossings, multiple $p\overline{p}$ interactions during the same beam crossing, and noise from the natural radioactivity of uranium absorbers in the calorimeter. A Jet Energy Scale (JES) correction is applied to reconstructed jets to compensate for these effects. This procedure is discussed in detail in Chapter IV.

3.4 Neutrino Identification

Neutrinos only weakly interact with the detector, and therefore are not detected directly. Their presence can only be inferred from the transverse energy imbalance in the measurement of an event. In this analysis, we use missing transverse energy ($\not\!\!E_T$)

as one cut in the event selection process (see Chapter VI). Its value is calculated in the following three stages.

First, the transverse energies of all the calorimeter cells with transverse energy greater than 100 MeV are added vectorially, except those cells in the coarse hadronic part. Since those coarse hadronic cells have high level of noise, they are included in the calculation only when they are clustered within good jets. The opposite vector to the resulting vector sum is the raw missing energy vector and its module is the raw missing energy ($\not{E}_{Traw}$).

In the next stage, as jet energies in an event are modifed by Jet Energy Scale corrections, so does $\not{E}_T$. This Jet Energy Scale correction accounts for the different calorimeter response for electrmagnetic objects and hadronic objects. Only the response part of the JES correction is applied to correct $\not{E}_T$. The resulting modulus is denoted the calorimeter missing transverse energy ($\not{E}_{TCAL}$).

Since the muon is a minimum ionizing particle throughout the detector, its presence in an event may cause transverse energy imbalance, and mimic the presence of a neutrino. To account for this problem, the momentum of all matched muons in an event is subtracted from the missing transverse energy vector after the expected energy deposition in the calorimeter is subtracted. The $\not{E}_T$ value used in the selection process of this analysis is obtained after this stage.

CHAPTER IV

JET ENERGY CORRECTION

As discusssed in Section 3.3, the reconstructed jet energy is different from the energy of the parton (gluon or quark) which initiates the jet. Figure 4.1 shows the average reconstructed jet energy as a function of the underlying parton energy obtained from Monte Carlo events. A large difference is observed.

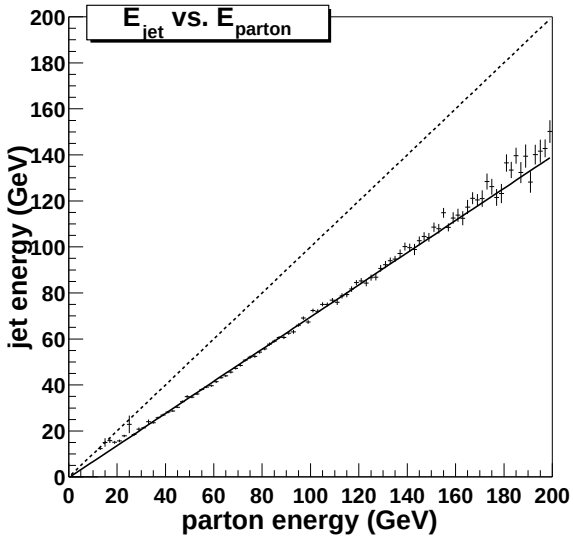


Figure 4.1: The reconstructed jet energy as a function of the parton energy, without the Jet Energy Scale correction. Each jet passes jet quality cuts, and is required to match the parton in η and ϕ to ensure that the jet originates from the parton.

In this analysis, we are interested in the energies of the underlying partons for given measured jet energies of candidate events. Two steps are taken to calibrate jet energy. The first step is the Jet Energy Scale (JES) correction, which corrects for the mean of the jet energies. The second step acccounts for the jet energy resolution of the measurement.

4.1 Jet Energy Scale Corrections

The reconstructed jet energy depends on a number of factors: the offset energy (E_{offset}), the calorimeter response (R_{jet}), and the fraction of the jet energy that is within the jet cone (F_S). The relation between the reconstructed jet energy as measured by the calorimeter (E_{jet}^{cal}) and the jet energy at the particle level ($E_{jet}^{particle}$) is:

$$E_{jet}^{particle} = \frac{E_{jet}^{cal} - E_{offset}}{R_{jet} \times F_S}, \tag{4.1}$$

where $E_{jet}^{particle}$ is defined as the energy of the particles constituting the jet before interacting with the detector material.

E_{offset} originates from the the following sources:

- uranium noise;
- multiple $p\overline{p}$ interactions during the same beam crossing;
- electronic pile-up from previous beam crossing;
- spectator parton interactions (physics underlying events).

R_{jet} is the ratio of the measured energy to the energy of an incident particle. It is a function of particle type, and particle energy. Due to the fixed cone size in the jet reconstruction algorithms, jet energy deposited outside the cone is not included in the energy measurement. The factor F_S is used to account for this effect.

The Jet Energy Scale correction tries to compensate for those effects by using Equation 4.1 so that on average the corrected jet energy matches the jet energy at the particle level. The offset energy is measured from the transverse energy density in minimum bias events as a function of detector pseudorapidity. The calorimeter response is measured using the Missing E_T Projection Fraction (MPF) method, using photon + jets events. In this method, the photon and the leading hadronic jet are required to be opposite in ϕ in order to ensure that the parton recoils in the opposite direction with respect to the scattered photon. Since energetic neutrinos are not expected in these events, and the photon energy is well measured in the electrogmagnetic calorimeter, any momentum imbalance in the transverse plane is attributed to the mismeasurement of the hadronic jet energy. The JES correction procedures with Equation 4.1 are implemented into a DØ software package called *jetcorr* [46]. Figure 4.2 shows the JES corrected jet energy versus the underlying parton energy.

As is shown in the figure, after applying the JES correction, the average reconstructed jet energy is close to the parton level energy. However, such correction process is not sufficient to return the jet energy to the parton level. While the effects of showers leaking out of the jet cone are accounted for, some of the particles radiated by the initial parton may fall outside of the cone. Such QCD radiation loss is not included in the correction process.

Denoting $\delta E = E_{corrected} - E_{parton}$ as the energy difference between the corrected jet energy and the parton energy, then the δE distribution represents the jet energy resolution. As we will see, the δE distribution depends not only on the underlying parton energy, but also on the parton type. For the purpose of studying the δE distribution, we define two quark categories: 1) light quarks, which include u, d, s,

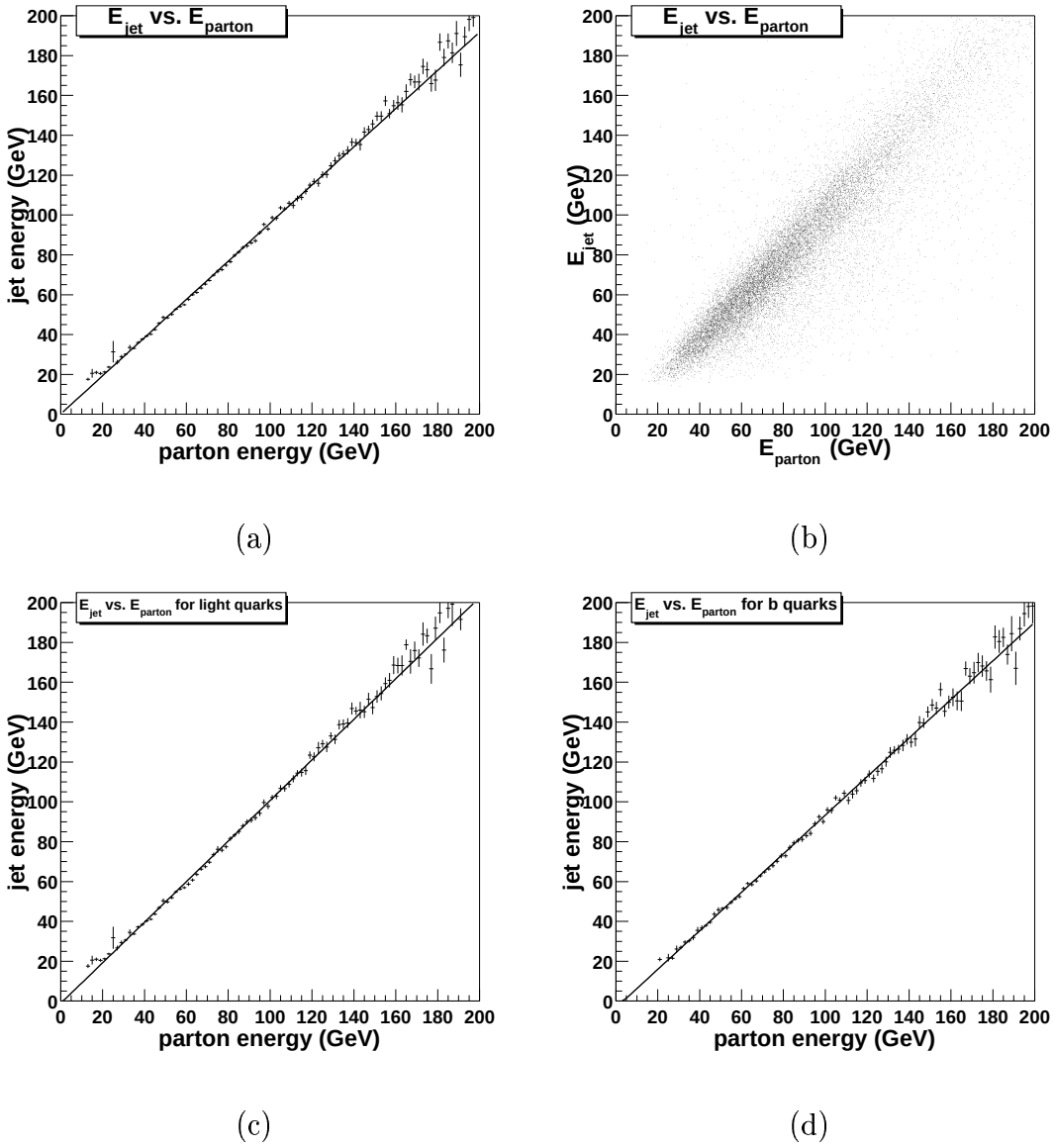


Figure 4.2: The reconstructed jet energy as a function of the underlying parton energy, after the JES correction. (a) is the profile histogram for all jets. (b) is the scatter histogram for all jets. (c) and (d) are profile histograms for light quark and b quark jets, respectively.

and c, and 2) b quarks. (The t quark is not included in consideration because it decays rapidly and leads to several jets.) Figures 4.3 and 4.4 show the δE distributions for different ranges of E_{parton}, for light quark and b quark jets, respectively. The distributions for heavy (b) quark jets are distinct from the distribution for light quarks, partly because the b quark decays semileptonically around 1/3 of the time. Several pieces of information can be obtained from these plots: the distributions of δE are not strictly Gaussian, but with asymmetric tails, and the width and degree of asymmetry of the distributions generally increase with the parton energy. These jet energy resolution features are taken into account in our analysis.

4.2 Parameterization of δE Distributions

In accordance to the δE distribution features in Figs. 4.3 and 4.4, we parameterize the distribution by the following function [50]:

$$f(\delta E) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)}[\exp\frac{-(\delta E - p_1)^2}{2p_2^2} + p_3 \exp\frac{-(\delta E - p_4)^2}{2p_5^2}]. \tag{4.2}$$

One of the Guassian functions is used to describe the peak, while the other to describe the asymmetry of the distribution. Parameters $p_1, p_2, ..., p_5$ are functions of the parton energy. For simplicity, they are assumed to have linear relations with E_{parton}:

$$p_i = a_i + b_i \times E_{parton} \tag{4.3}$$

$f(\delta E)$ is called the jet energy transfer function. When properly normalized, $f(\delta E)$ is the probability density of an energy difference δE. It describes the jet energy resolution of the detector, and includes physics effects such as soft gluon radiations.

The parameters a_i and $b_i, i = 1, 2, ..., 5$, are obtained by fitting the function to the δE distributions from Monte Carlo samples. Given N sets of (E_{jet}, E_{parton}) pairs,

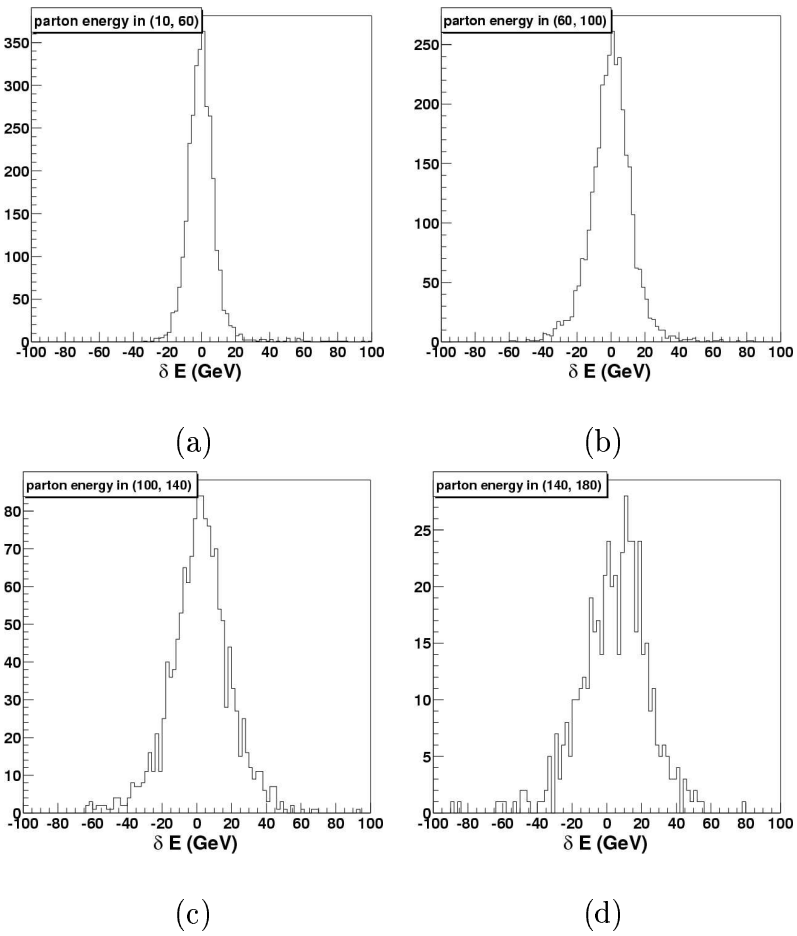


Figure 4.3: δE distributions in different parton energy ranges, for light quark jets. The energy ranges in which the parton energy lies are: (a) 10 GeV $< E_{parton} <$ 60 GeV; (b) 60 GeV $< E_{parton} <$ 100 GeV; (c): 100 GeV $< E_{parton} <$ 140 GeV; and (d) 140 GeV $< E_{parton} <$ 180 GeV.

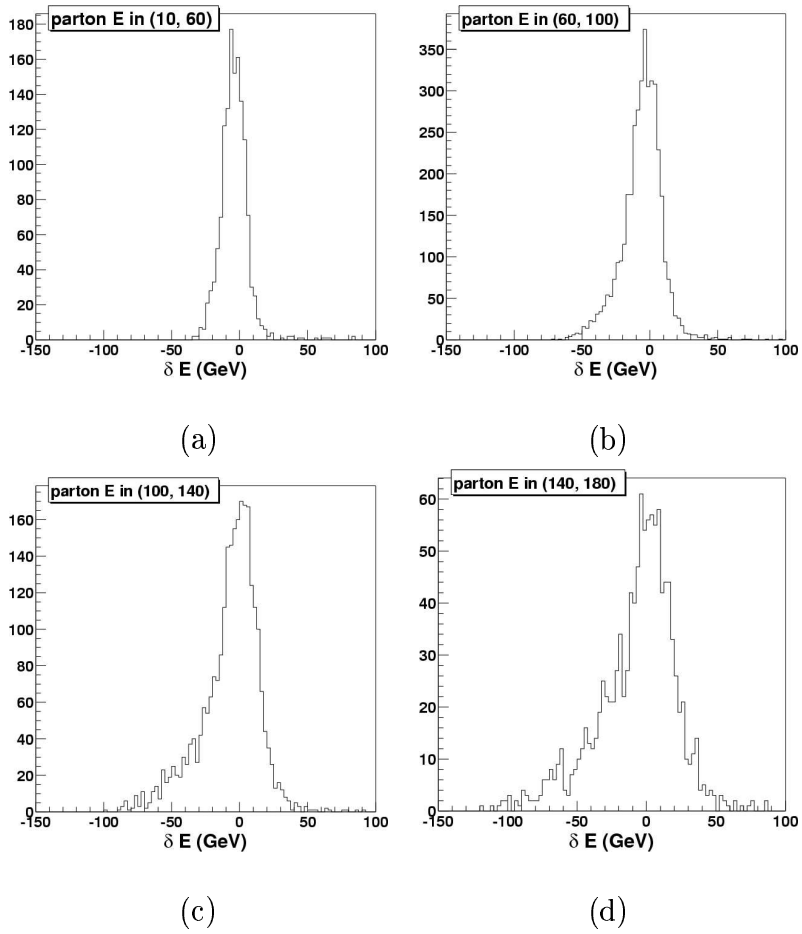


Figure 4.4: δE distributions in different parton energy ranges, for b-quark jets. The energy ranges in which the parton energy lies are: (a) 10 GeV $< E_{parton} <$ 60 GeV; (b) 60 GeV $< E_{parton} <$ 100 GeV; (c): 100 GeV $< E_{parton} <$ 140 GeV; and (d) 140 GeV $< E_{parton} <$ 180 GeV.

the probability density for having such a configuration is:

$$P = \prod_{i=1}^{N} f(\delta E_i), \tag{4.4}$$

and the likelihood function is given by:

$$-\ln L = -\ln P = -\sum_{i=1}^{N} \ln f(\delta E_i). \tag{4.5}$$

The parameters can then be extracted by using the maximum likelihood method. Due to the different δE distribution shapes of the light quark jets and the b quark

jets, we calculate two sets of parameters for the two types of jets separately. The values are listed in Table 4.1.

$p_i = a_i + b_i \times E_{parton}$ (Unit: GeV)	a_i (light quark)	b_i (light quark)	a_i (b quark)	b_i (b quark)
p_1	-1.801	3.169	-7.559	7.449×10^{-2}
p_2	1.882	1.095×10^{-1}	3.935	5.591×10^{-2}
p_3	0.0	4.042×10^{-4}	0.0	3.800×10^{-3}
p_4	1.230×10^{1}	-1.306×10^{-1}	-7.955	-2.665×10^{-2}
p_5	-1.901×10^{1}	8.558×10^{-2}	9.288	1.215×10^{-1}

Table 4.1: Parameters for the jet energy transfer function $f(\delta E)$, obtained using Monte Carlo samples. a_3 has been found to be very small, so we set it to be 0.

The jet energy transfer functions are tested using jet and underlying parton energies in Monte Carlo samples. For each parton energy, we generated a simulated jet energy based on the transfer function distribution. Figures 4.5 and 4.6 compare the δE distributions of Monte Carlo events with those calculated from the parameterization. The good agreement demonstrates that the transfer functions represent jet energy resolution in Monte Carlo samples very well.

4.3 Data and MC Comparison

There is no *a priori* reason to assume that the jet energy resolution in Monte Carlo events is the same as in real detector data events. Since the parton level energy is not available from data, the exact transfer functions for data cannot be obtained, and the systematic uncertainty due to inaccurate transfer functions must

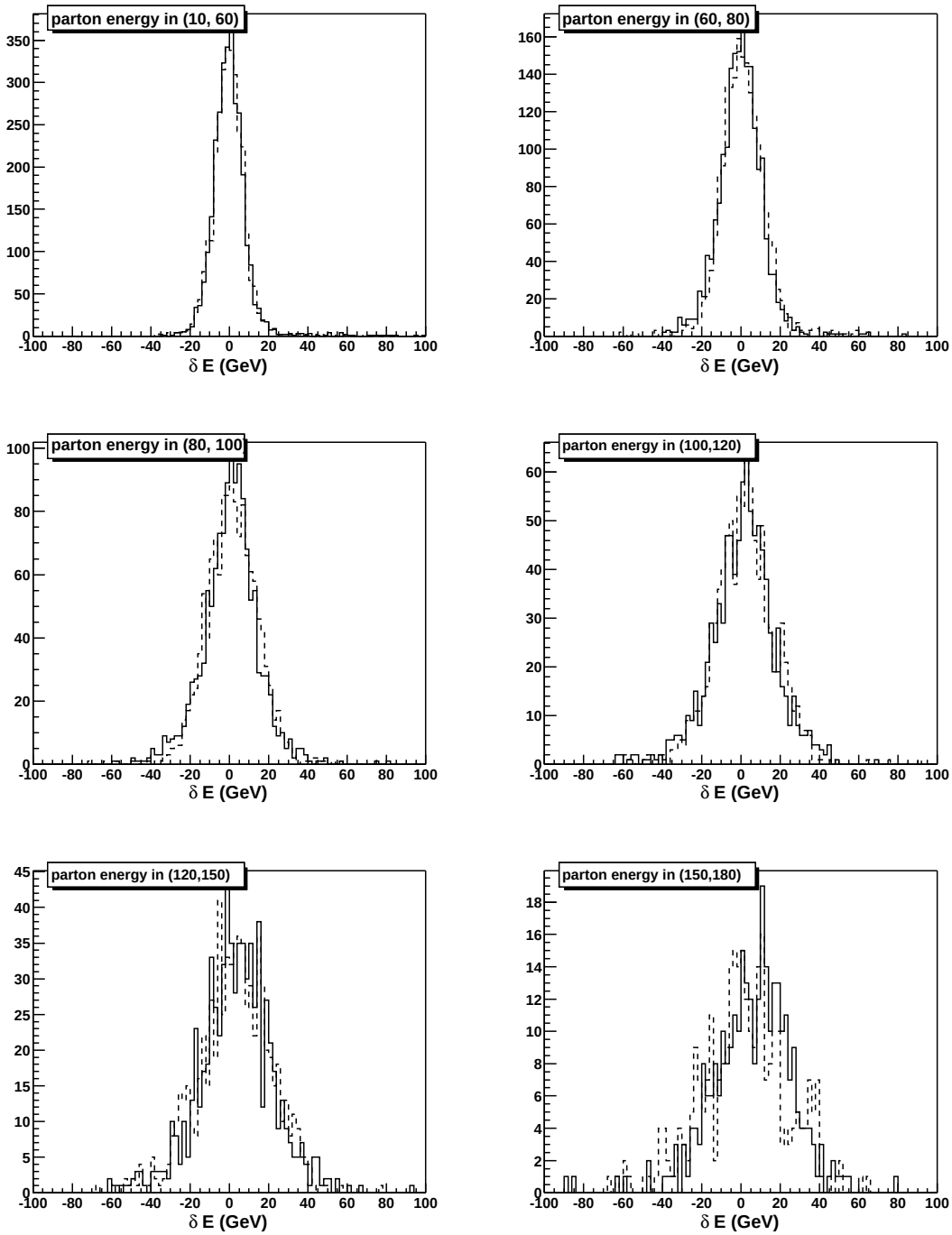


Figure 4.5: Test of the jet energy transfer function for light quark jets. In solid histograms, the jet energy is obtained from Monte Carlo samples. In dashed histograms, the jet energy is generated by the transfer function.

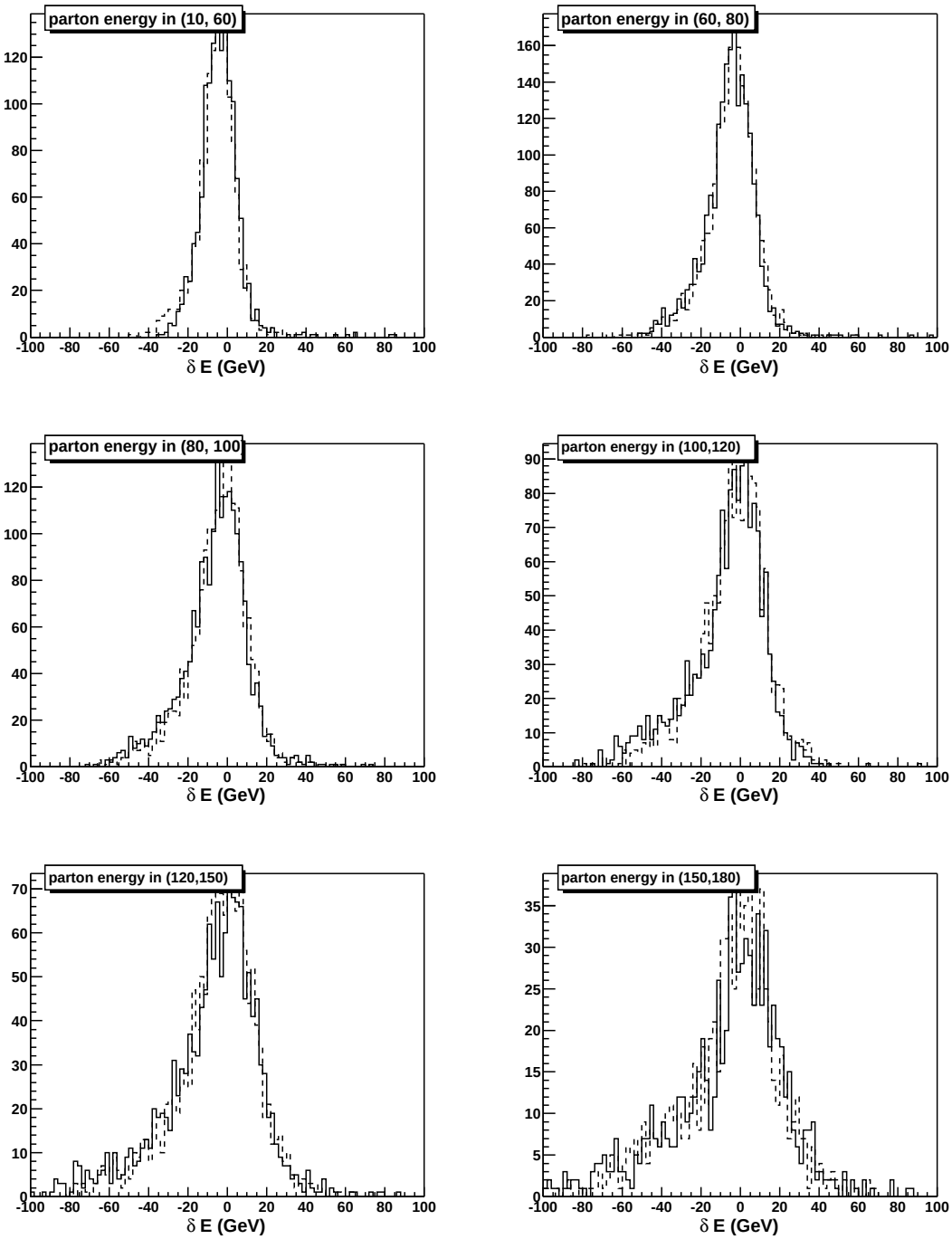


Figure 4.6: Test of the jet energy transfer function for b quark jets. In solid histograms, the jet energy is obtained from Monte Carlo samples. In dashed histograms, the jet energy is generated by the transfer function.

be estimated. We estimate the difference in jet energy resolutions between Monte Carlo and data by studying photon + 1-jet events.

The following selection criteria are applied to obtain the photon + 1-jet sample:

- Exactly one electromagnetic particle which passes quality cuts but with no associated tracks (to eliminate electrons) and $p_T > 15$ GeV. We restrict the particle to lie within the pseudorapidity region of -0.8 to 0.8, and not lie in the 16 calorimeter module boundary regions in the azimuthal direction in the central calorimeter, to ensure good energy measurement.
- Exactly one jet passing jet quality cuts, with $p_T > 15$ GeV and $|\eta| < 2.5$.
- $\not\!\!E_T$ smaller than 10 GeV.
- The azimuthal separation between the electromagnetic particle and the jet is greater than 3.0, to ensure that they are back-to-back in the ϕ direction.

Events that pass the selection criteria are mainly Compton scattering events, where the photon is scattered off a parton, and the parton recoils in the opposite direction. Assuming that the photon and the parton have the same transverse energy, and the reconstructed photon energy equals the original photon energy, the parton energy ($E_{parton}^{\gamma+jet}$) is given by:

$$E_{parton}^{\gamma+jet} = E_{jet} \cdot \frac{E_T^{\gamma}}{E_T^{jet}}, \tag{4.6}$$

where E_T^{jet} and E_T^{γ} are the reconstructed transverse energy for the jet and the photon, respectively.

However, in Compton scattering events, the transverse energy of the photon does not strictly equal the transverse energy of the parton. Figure 4.7 shows the distribution of the parton-level transverse energy difference between the photon and

the parton in Monte Carlo Compton scattering photon + 1-jet events. In addition, the detector has a finite resolution to the photon energy resolution. Therefore $E_{parton}^{\gamma+jet}$ is only an approximation to the real parton energy.

Denoting $\delta E' = E_{jet} - E_{parton}^{\gamma+jet}$, we plot the $\delta E'$ distributions from photon + 1-jet data events, as well as from Monte Carlo photon + 1-jet events in Figure 4.8. Data and Monte Carlo agree reasonably well. Figures 4.9 and 4.10 compare the $\delta E'$ and δE distributions using Monte Carlo photon + 1-jet events, for ligh quark jets and b quark jets, respectively. The difference between the $\delta E'$ and δE distributions is explained by the discrepancy betwen $E_{parton}^{\gamma+jet}$ and the true parton energy.

Based on the above discussions, we use jet energy transfer functions as obtained from Monte Carlo δE distributions in our data analysis. We extract another set of jet energy transfer functions from the $\delta E'$ distributions in photon + 1-jet data events, and use these to estimate the systematic uncertainty associated with jet energy resolutions.

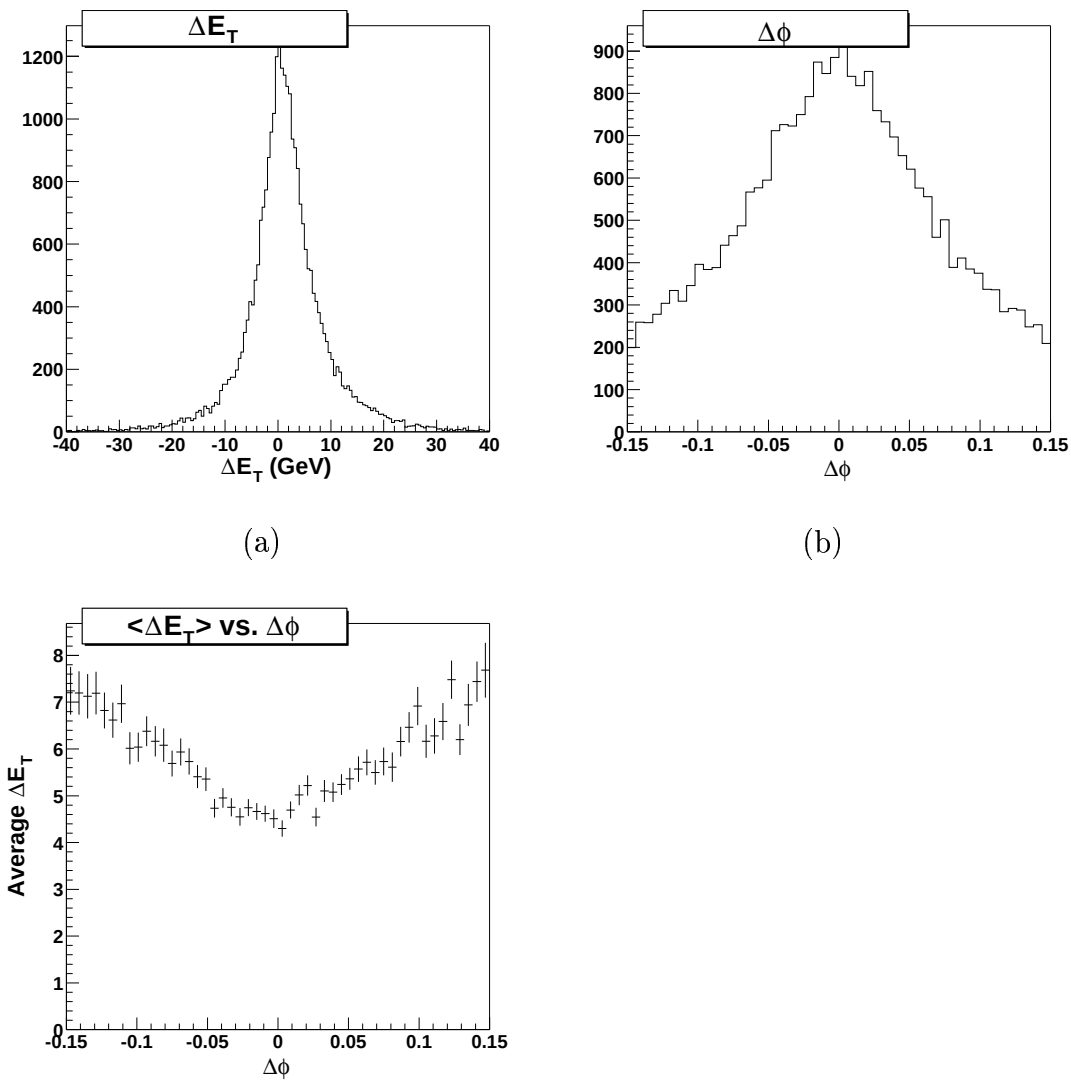


Figure 4.7: ΔE_T and $\Delta\phi$ distributions in Monte Carlo photon+jet events. (a) is the ΔE_T distribution, where $\Delta E_T = E_T^{parton} - E_T^{\gamma}$. This demonstrates that the transverse energy of the photon does not strictly equal the transverse energy of the parton. (b) is the distribution of the azimuthal separation between the photon and the parton ($\Delta\phi$). (c) shows the average ΔE_T as a function of $\Delta\phi$.

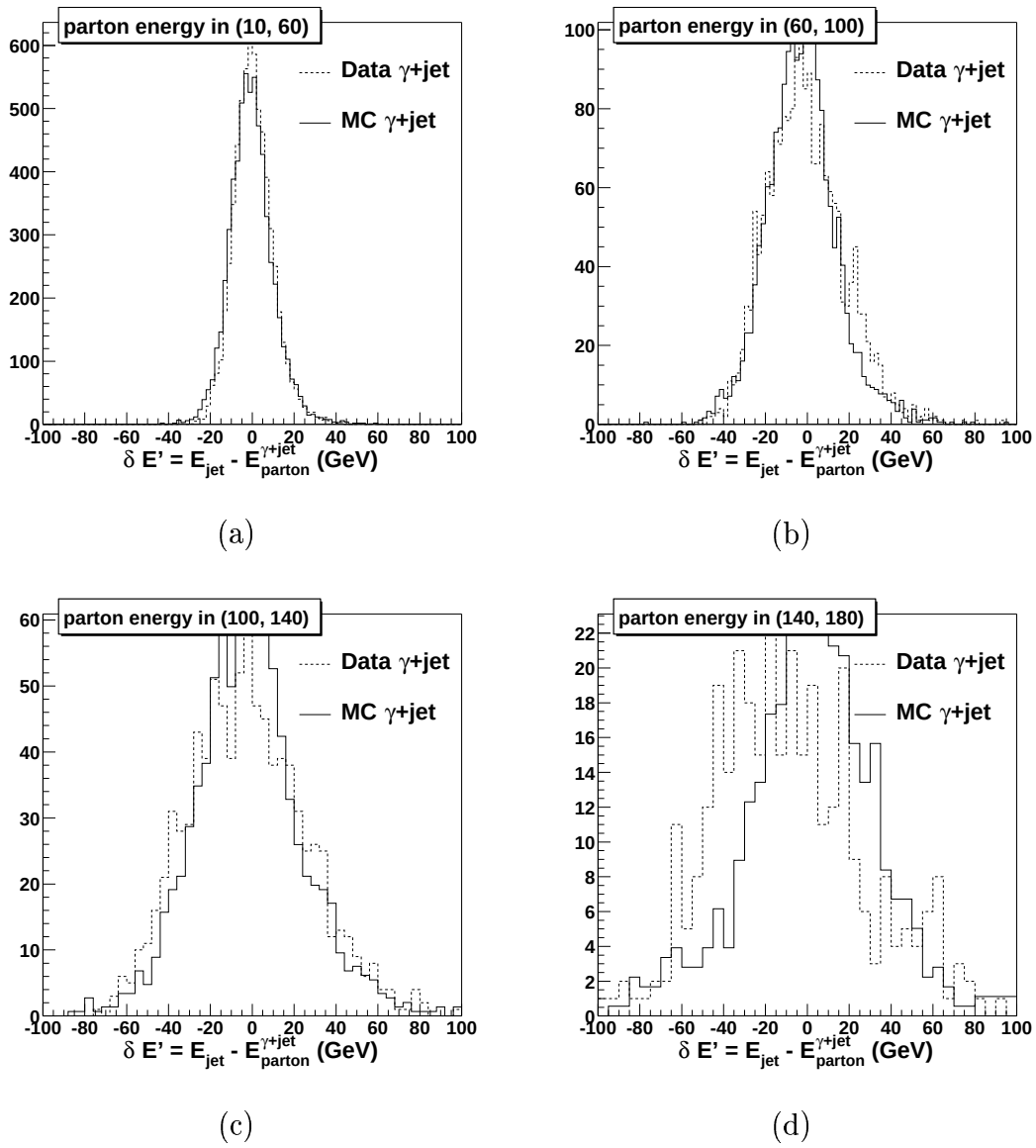


Figure 4.8: Comparison of $\delta E'$ distributions between Monte Carlo (solid histograms) and data (dashed histograms). The distributions are plotted with four different $E_{parton}^{\gamma+jet}$ ranges: (a) 10 GeV $< E_{parton} <$ 60 GeV; (b) 60 GeV $< E_{parton} <$ 100 GeV; (c) 100 GeV $< E_{parton} <$ 140 GeV; and (d) 140 GeV $< E_{parton} <$ 180 GeV.

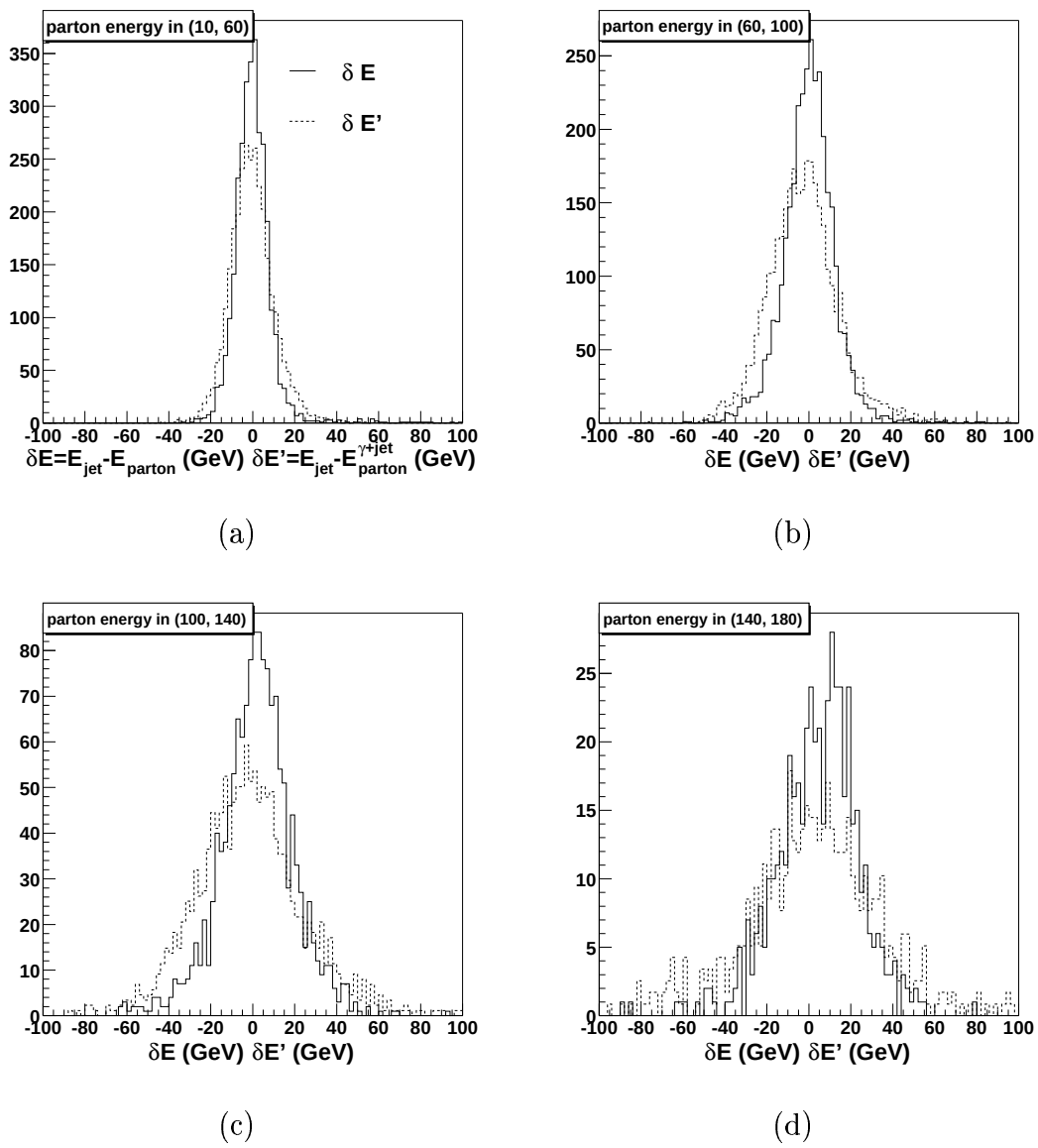


Figure 4.9: Comparison between δE distributions and $\delta E'$ distributions in Monte Carlo samples, for light quark jets. The solid histograms are the δE distributions, and the dashed histograms are the $\delta E'$ distributions.

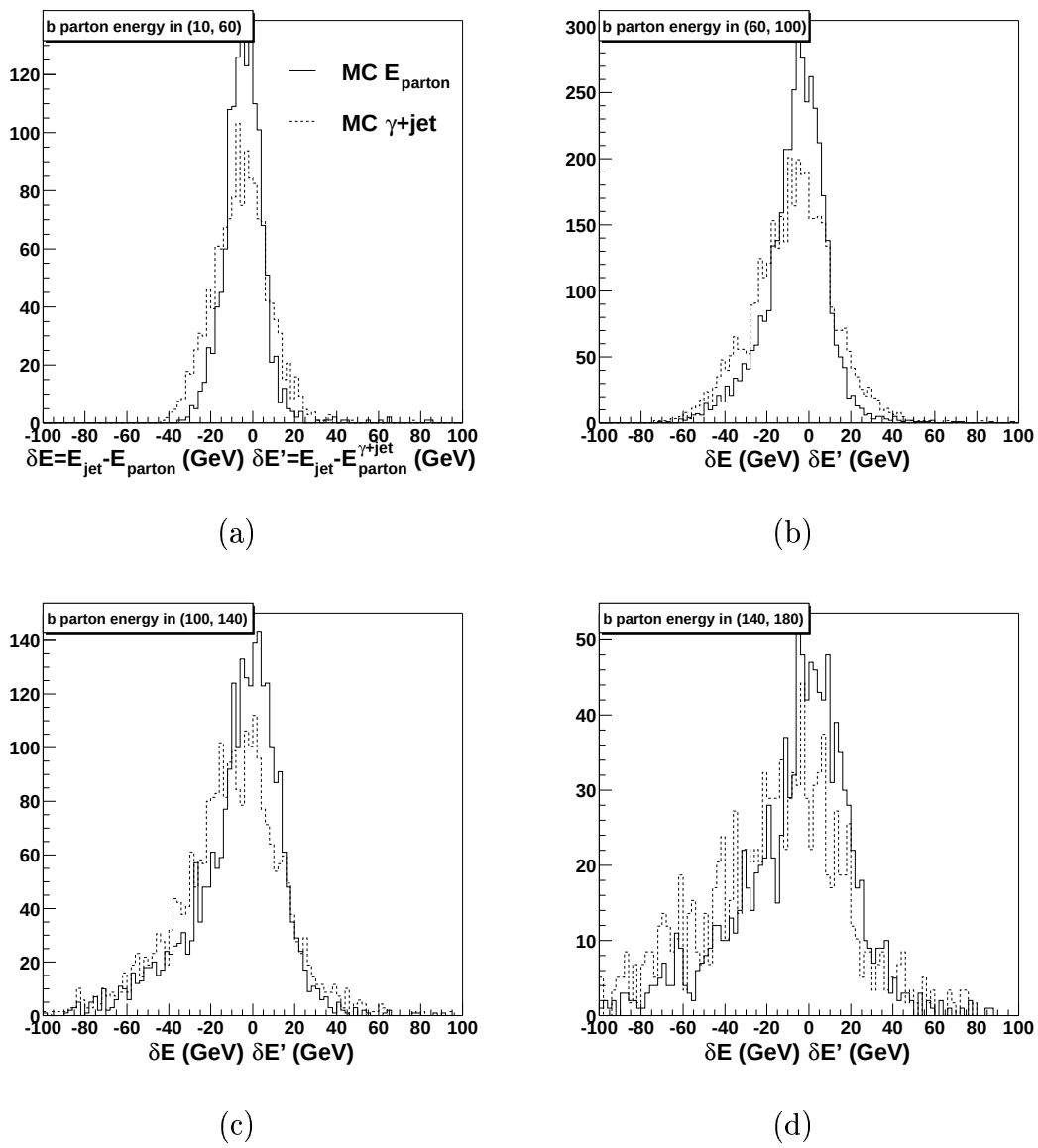


Figure 4.10: Comparison between δE distributions and $\delta E'$ distributions in Monte Carlo samples, for b quark jets. The solid histograms are the δE distributions, and the dashed histograms are the $\delta E'$ distributions.

CHAPTER V

ANALYSIS METHOD

This chapter first describes the principles of the Matrix Element Method used for top pair production cross section measurement, then it presents studies of this method using simulated events.

5.1 Method Overview

In experimental high energy physics, one prevalent problem is how to extract a small number of signal events from a data sample, given the presence of large backgrounds. In most cases, events from different processes have different kinematic properties in their final states, and background events are rejected based on these differences. In $t\bar{t} \rightarrow e$+jets decays, the event has the topology of a high p_T electron, large missing transverse momentum due to the electron neutrino, and at least four high p_T jets. The major background comes from the W+jets process, where the W decays into an electron and an electron neutrino, and some small background comes from QCD multijet processes, where a jet is misidentified as an electron. The W+jets process constitutes physics background, while QCD processes constitute instrumental (fake) background. The methods used in the DØ RunI measurement of $t\bar{t}$ cross section, as summarized in [58], first apply a set of kinematic cuts to each

final-state physics object in the event to obtain a "precut" sample. These kinematic cuts include pseudorapidity and transverse momentum cuts of the electron and jets, and a $\not\!\!E_T$ threshold. Since the "precut" sample is still rich in background events, due to the large cross section of background processes, a set of topological cuts are applied. These topological variables include the Aplanarity ($\mathcal{A}$), the scalar sum of the transverse energy of all the jets ($\mathcal{H}_T$), both of which are functions of a subset of kinematic variables in the final state. The topological cuts produce a final data sample with further reduced signal and background events. The cross section is then calculated by the following formula:

$$\sigma = \frac{S}{\varepsilon \times \mathcal{L}} = \frac{N - B}{\varepsilon \times \mathcal{L}}. \tag{5.1}$$

Here N, B, and S are the total number of observed events in the final data sample, expected background events in the sample, and number of signal events inferred from N and B, respectively. ε is the total signal efficiency, and $\mathcal{L}$ is the integrated luminosity.

The RunI method applies such optimization techniques as the neural network to find the best topological variables from a set of candidates, and the best numeric cut for each variable. The method proves to be successful in measuring the cross section. In this thesis, we study a Matrix Element Method to measure the cross section, instead of using topological cuts. The main differences between the two methods are listed below:

- Topological variables used in the previous method are functions of a subset of kinematic information of an event, while the matrix element is a function of all the kinematic variables in the final state.

- The previous method corrects jet energies on average, but ignores the energy

resolution effects. The Matrix Element Method takes it into account by integrating matrix element values over the jet energy resolution.

- The topological cuts used in RunI is about 50% efficient for signal events, so a large fraction of signal events are thrown away. The Matrix Element Method works on the "precut" sample obtained after kinematic cuts. Each event in the sample is given two probability values: signal probability and background probability. A likelihood function is constructed for all the events. The fraction of signal events in the sample can then be extracted using the maximum likelihood method.

5.2 Definition of Probabilities

Consider a generic particle interaction A involving an n-body final state. At the parton level it is represented as $p_1 + p_2 \to p_3 + p_4 + \cdots + p_{n+2}$. The differential cross section of this process is given by [5]:

$$d\sigma_A = \frac{(2\pi)^4 |\mathcal{M}_A|^2}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} d\Phi_n(p_1 + p_2; p_3, \ldots, p_{n+2}), \tag{5.2}$$

where p_1, p_2 are the four-momenta of the inital-state partons, m_1, m_2 are the masses of the two initial-state partons, $p_3, \ldots, p_{n+2}$ are the four-momenta of the final-state particles, and $\mathcal{M}_A$ is the matrix element for this process, which is a function of all the initial-state and final-state particle momenta, and determines the kinematic property of the interaction. $d\Phi_n$ is an element of the n-body phase space factor given by:

$$d\Phi_n(p_1, p_2; p_3, \ldots, p_{n+2}) = \delta^4(p_1 + p_2 - \sum_{i=3}^{n+2} p_i) \prod_{i=3}^{n+2} \frac{d^3 p_i}{(2\pi)^3 (2E_i)}. \tag{5.3}$$

To simplify notations, we denote dy as an element of the n-boday final-state phase space:

$$dy = \prod_{i=3}^{n+2} \frac{\delta(p_i^2 - m_i^2) d^4 p_i}{(2\pi)^3} = \prod_{i=3}^{n+2} \frac{d^3 p_i}{(2\pi)^3 (2E_i)}. \tag{5.4}$$

If $P_A^{prd}(y)dy$ is the probability that the interaction A leads to an event at the infinitesimal phase space element dy, then $P_A^{prd}(y)dy \propto d\sigma_A$. $P_A^{prd}(y)$ is called the production probability density function for that parton-level process.

At the Tevatron collider, initial-state partons are constituents of proton and antiprotons, and different flavors of partons can have the same type of interaction. So one has to fold $d\sigma_A$ over the intial parton energy distributions, and sum over contributions from different parton flavors to find the complete production probability. Let $f_a(q)dq$ be the probability that an initial parton of type a carries a fraction of q to $q+dq$ of the total proton (or antiproton) energy, then the production probability at the hadron collider is given by:

$$\begin{aligned} P_A^{prd}(y)dy &= \sum_{a,b} \int_{q_1} \int_{q_2} d\sigma_A(y) \times f_a(q_1) f_b(q_2) dq_1 dq_2. \\ &= \sum_{a,b} \int_{q_1} \int_{q_2} \frac{(2\pi)^4 |\mathcal{M}_A|^2}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} f_a(q_1) f_b(q_2) dq_1 dq_2 d\Phi_n. \end{aligned} \tag{5.5}$$

For the $t\bar{t}$ production process in our analysis, the above summation includes $u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$, $b\bar{b}$, and gg. The parton energy probability density function $f(q)$ is commonly called the *parton distribution function*. It has been extensively studied and its value is readily avaiable from tables.

If the momenta of all the final-state particles can be measured precisely by the detector, then $P_A^{prd}(y)$ is calculated in a straightforward way. However, we usually do not have sufficient information to pinpoint events in the experiment. As discussed in Chapter IV, the detector has a finite resolution to the jet energy measurement. Since the neutrino only weakly interacts with the detector, its momentum has to be inferred from the momentum imbalance of an event. Due to these measurement limits, the phase space position that is measured in the experiment is usually not identical to the intrinsic phase space position of that event. We thus need a different

probability term to account for this effect.

Denoting x as the phase space position of an event as measured in the experiment, the probability density for observing such an event in the detector is given by:

$$\tilde{P}_A^{obs}(x) = \int dy P_A^{prd}(y) W(y,x), \tag{5.6}$$

where the integration spans the complete final-state phase space. $W(y,x)$ is the probability density that an event with phase space position y at the production is measured to be at x by the detector. It includes all possible factors that make the measurements deviate from true values, including jet energy resolutions, angular resolutions, reconstruction errors, and misidentification of particles. We use δ-functions to represent quantities that are precisely measured in the detector, and Gaussian-like functions for quantities with finite measurement resolutions. $W(y,x)$ does not depend on the type of process that produced this event: the detector has the same energy measurement resolution to a b quark, whether it is produced from electroweak interations, or from a gluon via the QCD interaction.

It should be noted that up to this point no event is rejected. As the function $W(y,x)$ is normalized, the integrations of $P^{obs}(x)$ and $P^{prd}(y)$ over the whole final-state phase space are the same:

$$\int \tilde{P}_A^{obs}(x) dx = \int P_A^{prd}(y) dy. \tag{5.7}$$

Obviously not all events produced by the process A are included in the data analysis. For example, some events may fall out of the geometric coverage of the detector. And for those that fall within, many are rejected in the event selection process, either online by the trigger system or offline by the event selection algorithm. The probability that an event passes the selection is called the acceptance $\mathcal{A}(x)$. $\mathcal{A}(x)$ depends on such factors as detector geometry acceptance, trigger selection criteria,

reconstruction efficiency, and kinematic cuts applied to the data sample. $\mathcal{A}(x)$ is a function of the measured phase space position x, but is independent of the intrinsic process type of the event. If two events have the same final-state topology, but are produced by two different physics processes, they have the same probability of being included in the selected data sample.

Finally we arrive at the complete formula for the observation probability density function:

$$\begin{aligned} P_A^{obs}(x) &= C_A \cdot \mathcal{A}(x) \cdot \tilde{P}_A^{obs}(x) \\ &= C_A \cdot \mathcal{A}(x) \cdot \int dy P_A^{prd}(y) W(y,x) \\ &= C_A \cdot \mathcal{A}(x) \cdot \sum_{a,b} \int_{q_1} \int_{q_2} \int_y d\sigma_A(y) \cdot f_a(q_1) f_b(q_2) W(y,x) dq_1 dq_2. \end{aligned} \quad (5.8)$$

Here C_A is the normalization factor, obtained by requiring

$$\int P_A^{obs}(x) dx = 1. \quad (5.9)$$

Let $\overline{\mathcal{A}}_A$ be the average acceptance for events from process A, and σ_A be the total cross section for the process, then the integration gives:

$$C_A = \frac{1}{\int dx \mathcal{A}(x) \times \tilde{P}_A^{obs}(x)} = \frac{1}{\overline{\mathcal{A}}_A \times \sigma_A}. \quad (5.10)$$

Since $\overline{\mathcal{A}}_A$ depends on the distrubtion of $\tilde{P}_A^{obs}(x)$, it takes different values for different physics processes.

As we will show later, our analysis only depends on the relative ratios of $P_A^{obs}(x)$ for different processes. We will omit $\mathcal{A}(x)$ in the expression of $P_A^{obs}(x)$, and write Equation 5.8 as:

$$P_A^{obs}(x) = \frac{1}{\overline{\mathcal{A}}_A \times \sigma_A} \cdot \sum_{a,b} \int_{q_1} \int_{q_2} d\sigma_A(y) \times f_a(q_1) f_b(q_2) W(y,x) dq_1 dq_2. \quad (5.11)$$

5.3 Signal Fraction Extraction

In this analysis, after we apply selection cuts to the data sample, we obtain a dataset consisting of signal and background events. The signal process is $t\bar{t} \rightarrow e$+jets, and since our selection cuts require exactly four jets in each event, the main background is due to the W + 4jets process. For an event i in the dataset, we can calculate the observation probability density $P_s^{obs}(x_i)$ for the signal process, and $P_b^{obs}(x_i)$ for the background process. We denote f_s as the fraction of signal events in the dataset, then the observation probability density for event i is given by:

$$P^{obs}(x_i) = f_s \times P_s^{obs}(x_i) + (1 - f_s) \times P_b^{obs}(x_i). \tag{5.12}$$

In the following, we discuss two approaches to determine f_s.

5.3.1 Direct Likelihood Approach

$P^{obs}(x_i)$ corresponds to the overall probability that event i is included in the dataset. Then we can construct the negative logrithmic likelihood function for a given dataset of n events as:

$$-\ln L = -\sum_{i=1}^{n} \ln P^{obs}(x_i) = -\sum_{i=1}^{n} \ln[f_s \times P_s^{obs}(x_i) + (1 - f_s) \times P_b^{obs}(x_i)]. \tag{5.13}$$

We call this function the *Direct Likelihood Function.* Using the maximum likelihood method, we can extract f_s, which on average equals the true fraction of signal events *if* the distributions of signal and background events in the final-state phase space are identical to what are prescribed by the calculated probability densities $P_s^{obs}(x)$ and $P_b^{obs}(x)$. However, an examination of our procedure reveals that several factors may affect our calculations, thus making the calculated probability distributions deviate from the real distributions:

- In our calculation, we assume that the four jets in each selected event map to the four final-state partons. This is not true in all the events. Monte Carlo Studies show that only 55% of all the selected signal events have each of the four jets lie within a cone of size $\delta R = 0.3$ around a parton. This percentage drops to 25% when we restrict δR to be 0.1.

- Given four jets, there are $4! = 24$ possible ways of matching the four final-state partons. Since we don't know the flavor of the underlying parton for each jet, we compute $P_s^{obs}(x)$ and $P_b^{obs}(x)$ for all 24 permutations and the take the average value.

- The jet energy resolution puts one more smearing effect to the distribution, though it is minized by integrating over jet energy resolutions.

Due to these issues, f_s extracted by the likelihood function may be different from the true fraction of signal events. We study potential biases of the method using Monte Carlo events in the later part of this chapter.

5.3.2 Discriminator Likelihood Approach

Instead of directly using $P_s^{obs}(x)$ and $P_b^{obs}(x)$ to construct the likelihood function, we can define a discriminator variable as:

$$D(x) = \frac{P_s^{obs}(x)}{P_s^{obs}(x) + P_b^{obs}(x)}. \tag{5.14}$$

It is similar to the discriminants used in the DØ RunI top mass analysis [49]. The discriminator should be large ($\approx$1) for signal events, and small ($\approx$0) for background events. We obtain the discriminator distributions for signal and background events from Monte Carlo events. Then for an event i in our dataset, we calculate its discriminator D_i, and using the discriminator distributions as probability density

distributions, we can assign a new signal probability ($P'_s(D_i)$) and background probability ($P'_b(D_i)$) to the event. The likelihood function based on the new probabilities is defined as:

$$-\ln L' = -\sum_{i=1}^{n} \ln P'(x_i) = -\sum_{i=1}^{n} \ln[f_s \times P'_s(D_i) + (1 - f_s) \times P'_b(D_i)]. \tag{5.15}$$

This likelihood function is called the *Discriminator Likelihood Function.* The detailed steps to do analysis with $P'_s(D)$ and $P'_b(D)$ are discussed in Section 5.8 of this chapter.

5.4 $W(y, x)$ Parameterization

In $t\bar{t} \rightarrow e$+jets decays, the final-state particles include an electron, an electron neutrino, and four quarks, two of which are b quarks directly from top decays. Assuming all the final-state particles are massless, one needs $3 \times 6 = 18$ kinematic variables to fully describe an event.

However, the neutrino momentum can be deduced from the momenta of other final-state particles. Since the transverse momentum of the initial-state partons are small compared to the total interaction energy, the final-state total transverse momentum can be approximated to zero. Therefore the transverse momentum of the neutrino is given by:

$$p_x^\nu = -p_x^e - \sum_{jet} p_x^{jet} \tag{5.16}$$

$$p_y^\nu = -p_y^e - \sum_{jet} p_y^{jet} \tag{5.17}$$

where p_x^e and p_y^e are the x and y components of the electron momentum respectively, and the summation includes all the jets in the event. The z momentum of the neutrino, p_z^ν, is obtained by assuming that the invariant mass of the neutrino and

the electron equals the mass of the W boson:

$$\begin{cases} (E^e + E^\nu)^2 - (p_x^e + p_x^\nu)^2 - (p_y^e + p_y^\nu)^2 - (p_z^e + p_z^\nu)^2 = M_W^2 \\ E^e = \sqrt{(p_x^e)^2 + (p_y^e)^2 + (p_z^e)^2} \\ E^\nu = \sqrt{(p_x^\nu)^2 + (p_y^\nu)^2 + (p_z^\nu)^2} \end{cases} \tag{5.18}$$

This equation set usually leads to two solutions of p_z^ν, both of which are used in the analysis.

Since the DØ Detector has a fine-segmented Calorimeter system, it provides good position measurement to both electromagnetic particles and hadronic jets. In addition, the EM calorimeter is calibrated with $Z \to e^+e^-$ events, and measures electron energy to a high precision. In Figure 5.1 we plot the electron energy and angular resolutions using the Monte Carlo sample. The energy resolution uncertainty is less than 3 GeV, much smaller than the jet energy resolution uncertainty as shown in Figures 4.3 and 4.4. As a good approximation, we use δ-functions to represent the measurement resolutions of electron and jet angles, and electron energy. We are finally left with four kinematic variables, which are energies of the four jets. So the function $W(y, x)$ takes the form:

$$W(y, x) = \delta(\vec{p_e}(y) - \vec{p_e}(x)) \prod_{i=1}^{4} f(E_i^{parton} - E_i^{jet}) \prod_{i=1}^{4} \delta(\Omega_i^{parton} - \Omega_i^{jet}), \tag{5.19}$$

where Ω_i is the solid angle of the ith parton or jet, and $f(E_i^{parton} - E_i^{jet}) = f(\delta E)$ is the jet energy transfer function that we obtained in Chapter IV.

5.5 Signal and Background Simulation

5.5.1 Event Generation

The main Monte Carlo sample for $t\bar{t} \to e$+jets signal events is generated using PYTHIA 6.203 [55], with a top quark mass of 175 GeV. The W + 4jets background

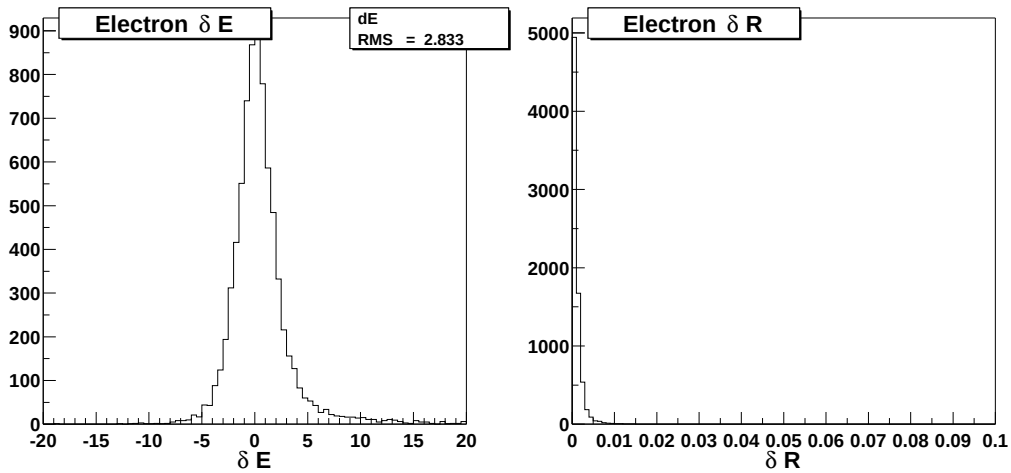


Figure 5.1: The δE and the δR distributions for electrons in Monte Carlo events. Here δE is the difference between the reconstructed energy and the original energy of the electron, and δR is the difference between the reconstructed position and the original position. $\delta R = \sqrt{(\delta\phi)^2 + (\delta\eta)^2}$.

events are generated using VECBOS [52] interfaced to HERWIG [56] to simulate underlying events, additional initial state radiation, parton hadronization, and the decay of the W boson. All signal and background processes are generated with a center-of-mass energy $\sqrt{s} = 1.96$ TeV. CTEQ4L [53] is used to model the parton distribution functions in the proton and the antiproton.

5.5.2 Detector Simulation

All the Monte Carlo samples used in this analysis, after the hadronization stage, have additional minimum bias proton-antiproton events superposed. The number of added events is Poisson-distributed with a mean of 0.5, as expected for an average instantaneous luminosity of 2×10^{31} cm^{-2}s^{-1} at which the data used in this analysis are taken. Each event is then processed with a GEANT [17]-based simulation program to simulate detector response, and with a DØ signal digitization simula-

tion program to produce simulated raw data, and finally with the DØ reconstruction program (RECO) for full event reconstruction.

5.5.3 Event Selection

We apply a set of selection cuts on both the signal and background samples. These cuts are identical to those applied to real data samples in the analysis, and are discussed in detail in Chapter VI. The key elements of the cuts are listed here:

- one high-p_T electron in the central calorimeter fiducial region, with $|\eta| < 1.1$, and $E_T > 20$;
- exactly four high-p_T jets with $|\eta| < 2.0$, and $E_T > 15$;
- $\not\!\!E_T$ greater than 20 GeV;
- the azimuthal separation between the electron and the neutrino (whose momentum is calculated as described in Sec. 3.4) is greater than 0.5.
- the W boson which decays into the electron and the neutrino is required to have $E_T > 60$ GeV, and $|\eta| < 2.0$.

The average acceptances for signal and background processes, denoted $\overline{\mathcal{A}}_s$ and $\overline{\mathcal{A}}_b$ respectively, depend on the selection cuts. They are calculated from Monte Carlo samples as the fraction of events that pass the selection cuts.

5.6 Production Probability Calculation

Although we do not know the intrinsic momenta of all the final-state particles in a real data event, we can get such information for Monte Carlo events. So we can calculate the production probability density for Monte Carlo events using Equation 5.5. The key component in the equation is the matrix element, whose formula is already

coded in some Monte Carlo generator programs. The W+jets event generator, VECBOS, contains subroutines to calculate leading order (LO) matrix element values for W+jets processes, and we use these subroutines to calculate the background production probability density $P_b^{prd}(y)$. The top event generator, ONETOP [51], contains subroutines to calculate leading order matrix element values for $t\bar{t}$ production processes, including $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$. We use them to obtain the signal production probability density $P_s^{prd}(y)$. The parton distribution function set CTEQ4L [53] is used in the calculation of both $P_b^{prd}(y)$ and $P_s^{prd}(y)$ to match what is used in the Monte Carlo generation.

Figure 5.2 shows the signal production probability density distributions for $t\bar{t}$ and W+4jets Monte Carlo events. As we expect, signal events have larger signal production probabilities.

5.7 Observation Probability Calculation

Our next step is to calculate the observation probability densities, $P_s^{obs}(x)$ and $P_b^{obs}(x)$, using Equation 5.11. The integration over jet energy transfer functions of the four jets in each event is done using the Monte Carlo integration technique. Given energies of the four jets, we randomly generate N sets of the final-state parton energies such that they are distributed according to the jet energy transfer functions $f(\delta E)$. For the ith set of final-state parton energies, we calculate the observation probability $P_{Ai}^{obs}(x)$. The result of the integration is given as:

$$P_A^{obs}(x) = \frac{1}{N}\sum_{i=1}^{N} P_{Ai}^{obs}(x). \tag{5.20}$$

While p_x and p_y of the neutrino is inferred by the transverse momentum conservation, the longitudinal momentum of the neutrino is obtained by solving Equation Set 5.18, and we sum the probabilities for the two solutions of the equation set. The

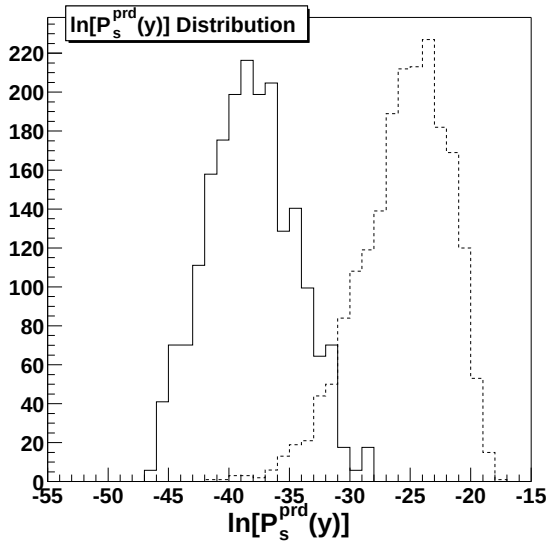


Figure 5.2: Signal production probability density ($P_s^{prd}(y)$) distributions for the signal (dashed line) and background (solid line) Monte Carlo events. In calculating $P_s^{prd}(y)$ for background events, we iterate over 24 permutations of parton assignments, and select the largest value of $P_s^{prd}(y)$. Both samples are required to pass the selection cuts.

four jets lead to $4! = 24$ different ways of assignment to the partons, and we take the average of the 24 corresponding $P_A^{obs}(x)$ values.

The Monte Carlo integration requires proper choice of the number of sampling points N. While accuracy increases as N gets larger, so does computing time. Figure 5.3 shows the uncertainties of the $P_s^{obs}(x)$ integration for a given event with different number of sampling points. Shown on each plot is the distribution of 100 Monte Carlo integrations with different seeds for the same event. The uncertainty of the integration is around 10% when N is 10000. In the Monte Carlo tests, we choose $N = 10000$ in the integration for $P_s^{obs}(x)$; in the data analysis, we choose $N = 20000$. Figure 5.4 shows the uncertainty of the $P_b^{obs}(x)$ integration for a given event with different number of sampling points. In the Monte Carlo tests, N is chosen to be 500 for the $P_b^{obs}(x)$ calculation, where the uncertainty is shown to be less than 3%; in the data analysis, we choose $N = 1000$.

As discussed before, the four jets in each event does not always match the final-state partons. Monte Carlo simulation shows that of all the signal events that pass the selection cuts, around 25% of them have each of the four jets matched to a parton, with their angular separation δR less than 0.1. Figure 5.5 compares the $P_s^{obs}(x)$ distributions for events whose jets are all matched to partons and for events without the matching requirement. Those with parton-jet matching have higher signal probability since their final-state kinematic configuration better resembles that of the signal process. Figure 5.5 also shows the $P_s^{prd}(y)$ distribution. The distance between the $P_s^{prd}(y)$ distribution and the other two is mainly due to the process of averaging the 24 possible jet assignment to partons with the two event samples, since most combinations yield small $P_s^{obs}(x)$ values.

Figure 5.6(a) shows the distributions of $P_s^{obs}(x)$ and $P_b^{obs}(x)$ for Monte Carlo

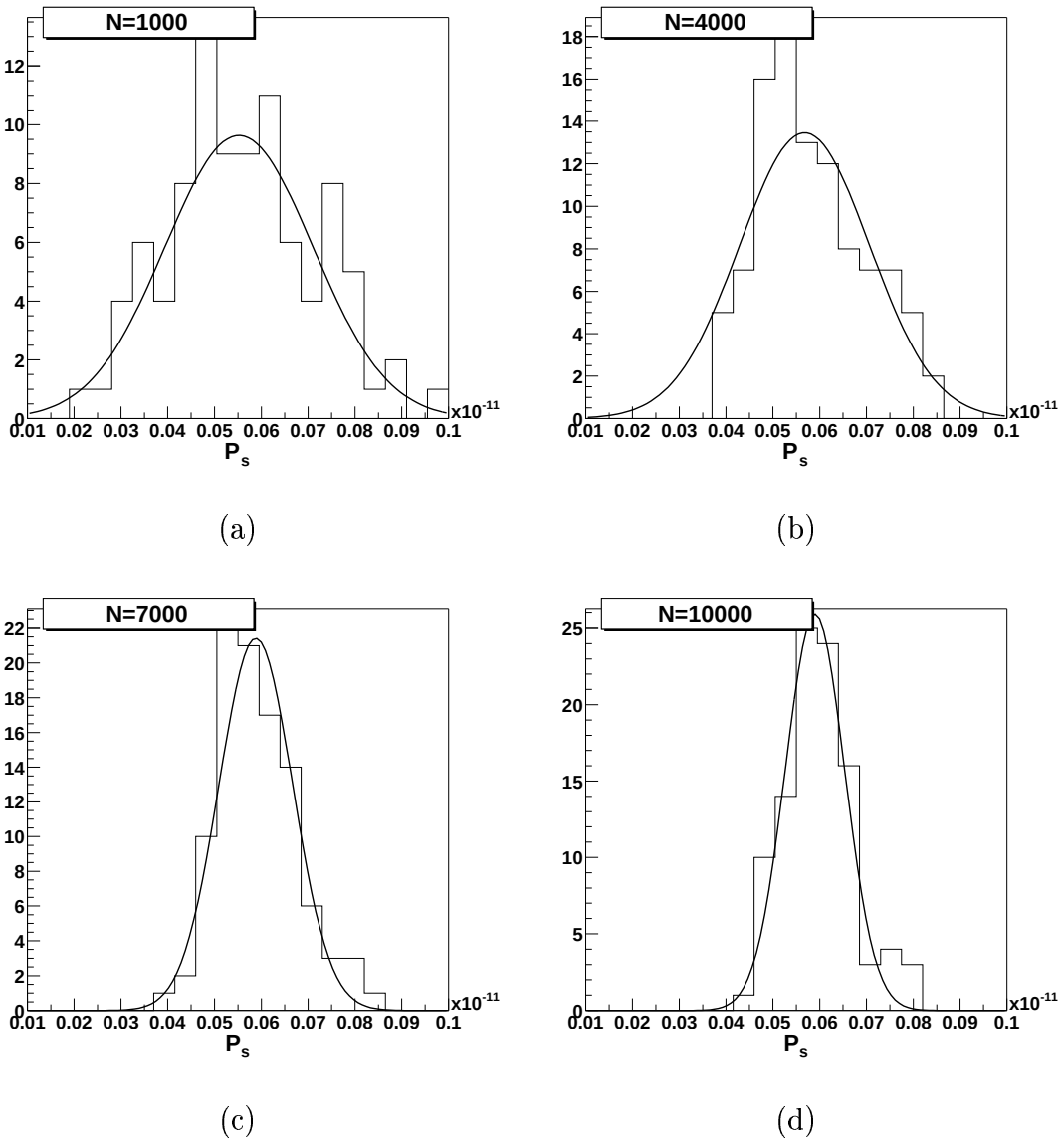


Figure 5.3: The uncertainty of Monte Carlo integrations in the calculation of $P_s^{obs}(x)$. Each plot shows the integration result distribution for a randomly selected event. Different number of sampling points N is used in each plot: (a) $N = 1000$, (b) $N = 4000$, (c) $N = 7000$, (d) $N = 10000$.

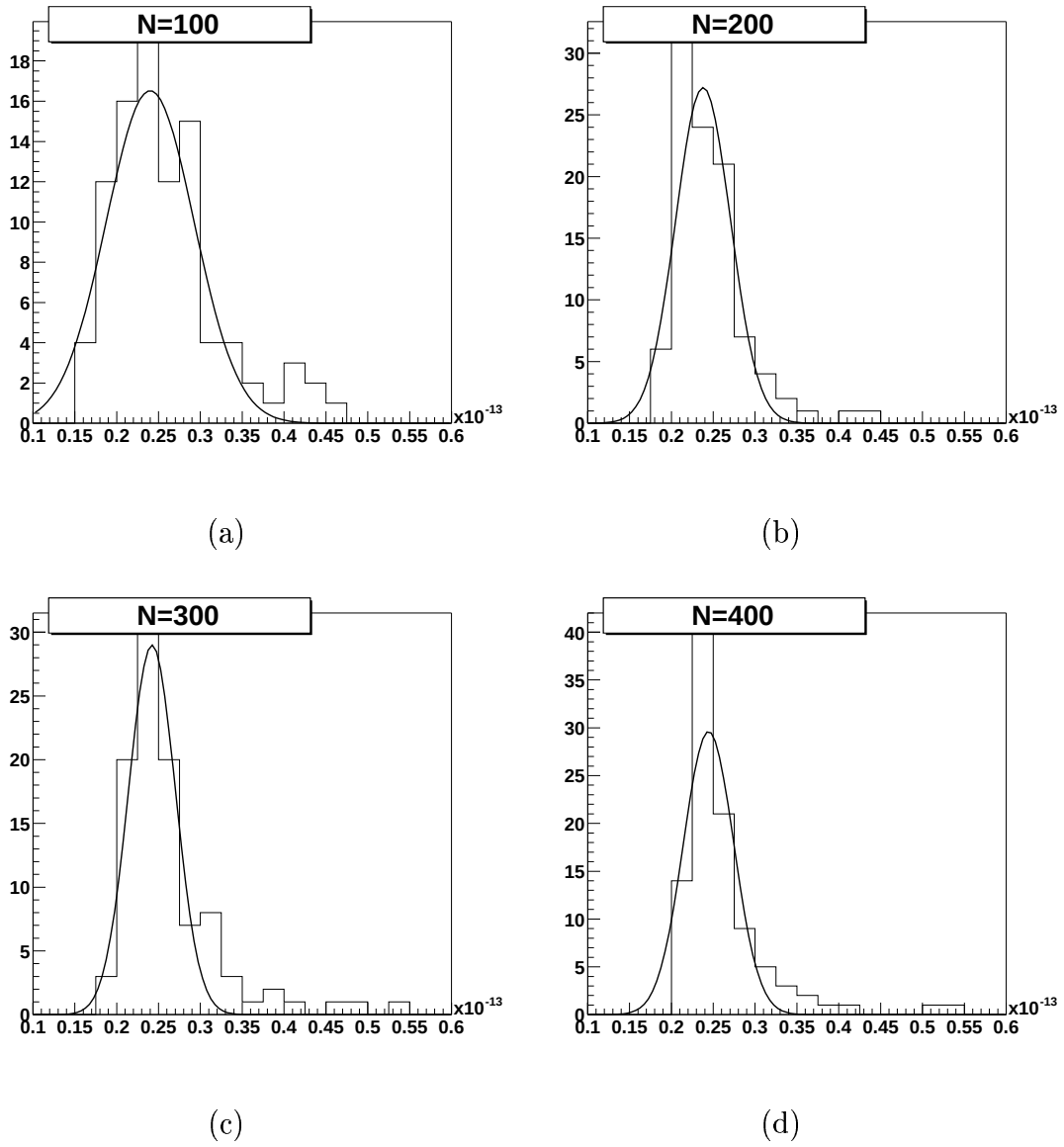


Figure 5.4: The uncertainty of Monte Carlo integrations in the calculation of $P_b^{obs}(x)$. Each plot shows the integration result distribution for a randomly selected event. Different number of sampling points N is used in each plot: (a): $N = 100$, (b): $N = 200$, (c): $N = 300$, (d) $N = 400$.

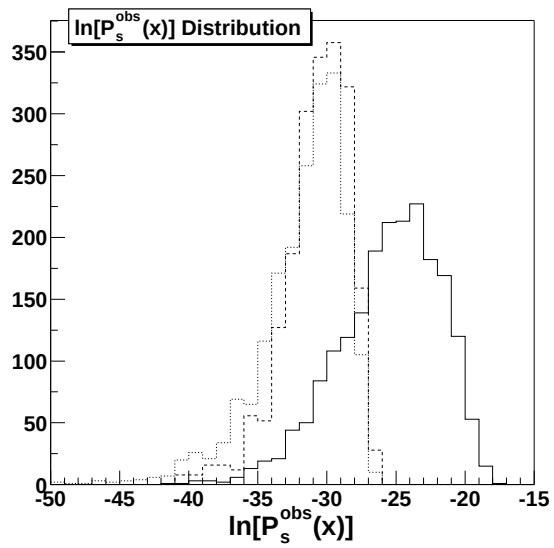


Figure 5.5: Comparison of observation probability ($P_s^{obs}(x)$) distributions. The dashed line plots the $P_s^{obs}(x)$ distribution for signal events whose jets are all matched to partons. The dotted line plots the $P_s^{obs}(x)$ distribution for all signal events that pass the selection cuts. The solid line plots the $P_s^{prd}(y)$ distribution for all signal events that pass the selection cuts.

signal and background samples together in a scatter plot. Figure 5.6(b) and (c) show the projected one-dimensional distributions in $P_s^{obs}(x)$ and $P_b^{obs}(x)$ dimension, respectively. As expected, signal events have larger signal production probability, while background events have large background production probability.

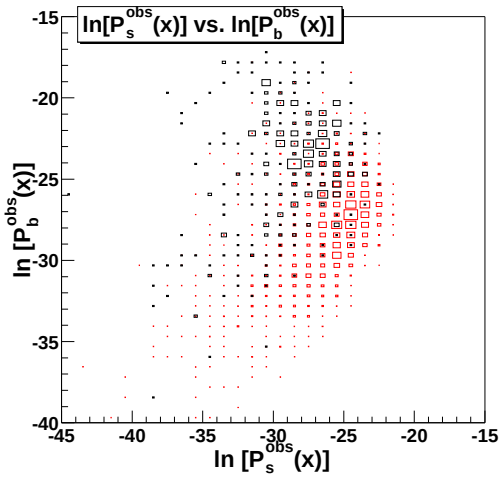
5.8 f_s Extraction

5.8.1 Direct Likelihood Analysis

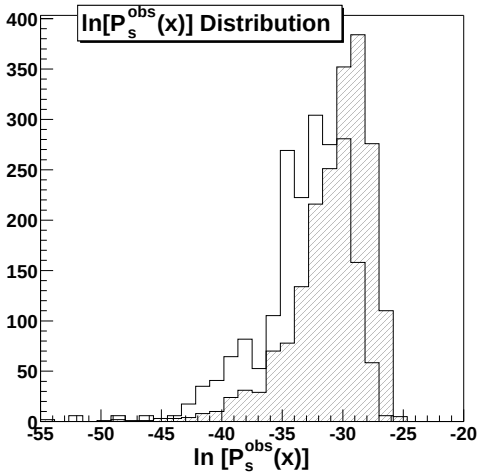
First we demonstrate the Monte Carlo test result when we do analysis with Equation 5.13, the Direct Likelihood Function.

We fix the sum of the number of signal events (n_s) and the number of background events (n_b) in all the test datasets, so that $n_s + n_b = 50$. Figure 5.7 shows shapes of the likelihood function for randomly selected datasets with different number of signal events. Figure 5.8 shows the uncertainty of the extracted f_s with different input n_s. To determine how f_s is related to the true fraction of signals in the dataset, we do a linearity test as follows. We select N random datasets with each pair of $< n_s, n_b >$ values, extract f_s for each dataset, and calculate the mean value of f_s in the N experiments. Figure 5.9(a) shows the the average of the extracted f_s as a function of the true fraction of signal events in the datasets, $n_s/(n_s + n_b)$. The average f_s value deviates from the actual values by a factor of around 0.78. This is not surprising since in a large fraction of signal events, the reconstructed jets do not map into the four final-state partons.

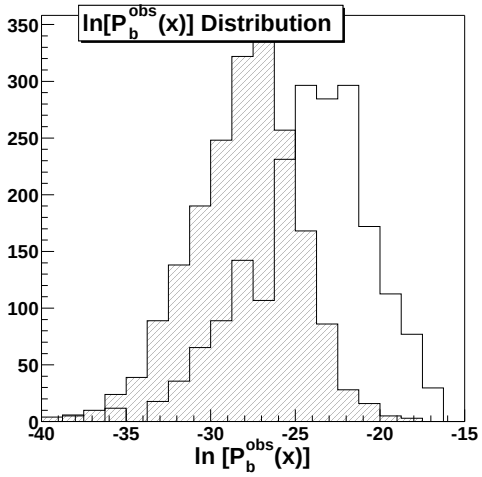
To test the validity of this argument, we select a subsample of Monte Carlo signal events, where the events are required to have all the four jets match the final-state partons with a δR smaller than 0.1 (jet-parton match requirement). We carry out the same linearity test with the Monte Carlo background sample and this signal



(a)



(b)



(c)

Figure 5.6: Observation probability ($P_s^{obs}(x)$ and $P_b^{obs}(x)$) distributions for signal and background Monte Carlo samples. Figure (a) shows the two-dimensional distributions for signal events (in red) and background events (in black). (b) and (c) are the distributions projected to the $P_s^{obs}(x)$ and $P_b^{obs}(x)$ coordinates, respectively.

subsample, and the result is shown in Figure 5.9(b). In this case, the average f_s is close to the real fraction of signal events.

5.8.2 Discriminator Likelihood Analysis

The discriminator variable defined in Equation 5.14 reflects the relative strength of $P_s^{obs}(x)$ and $P_b^{obs}(x)$ for each event. The discriminator distributions for signal and background Monte Carlo samples are shown in Figure 5.10. In Figure 5.10(a), the distribution for signal events is made for events passing the selection cuts. The discriminator values for a significant portion of the signal events do not approach 1, mainly due to the jet-parton mismatch issues discussed in the previous section. In Figure 5.10(b), only signal events that satisfy jet-parton match requirement are used to plot the discriminator distribution, and we have much better separation between signal and background events. In both cases, the separations between signal and background events are much more distinct compared to previous discriminants used in DØ RunI mass analysis [49], demonstrating the superior strength of the Matrix Element Method to discriminate between signal and background events.

For each event i in the test dataset, we calculate its discriminator $D(x_i)$, and extract $P_s'(D_i)$ and $P_b'(D_i)$ from the discriminator distributions in Figure 5.10(a). f_s for each test dataset is then extracted by minimizing the negative logarithm of the Discriminator Likelihood Function. Figures 5.12 and 5.13 show the typical shape of the likelihood function and the f_s distribution for different number of signal events in the dataset. Just as in the Direct Likelihood Analysis, we do a linearity test. Figure 5.11 shows the test result. Since the discriminator distribution for signal events in Figure 5.10(a) includes all the signal events passing the selection cuts, events in the test datasets have the same $P_s'(D)$ distribution as in the plot. Therefore,

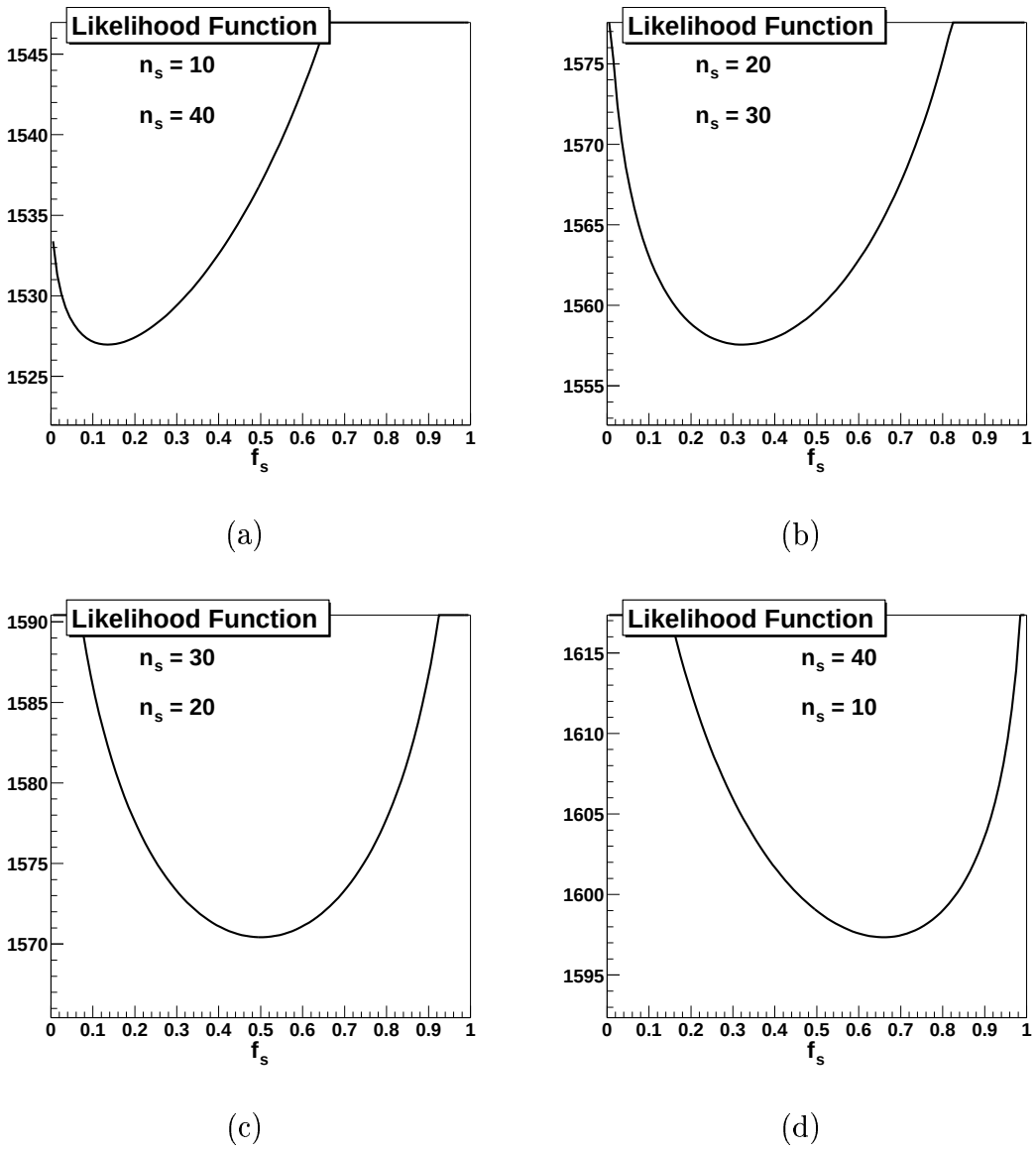


Figure 5.7: The Direct Likelihood Function shapes with different n_s values. Each test dataset contains 50 Monte Carlo events.

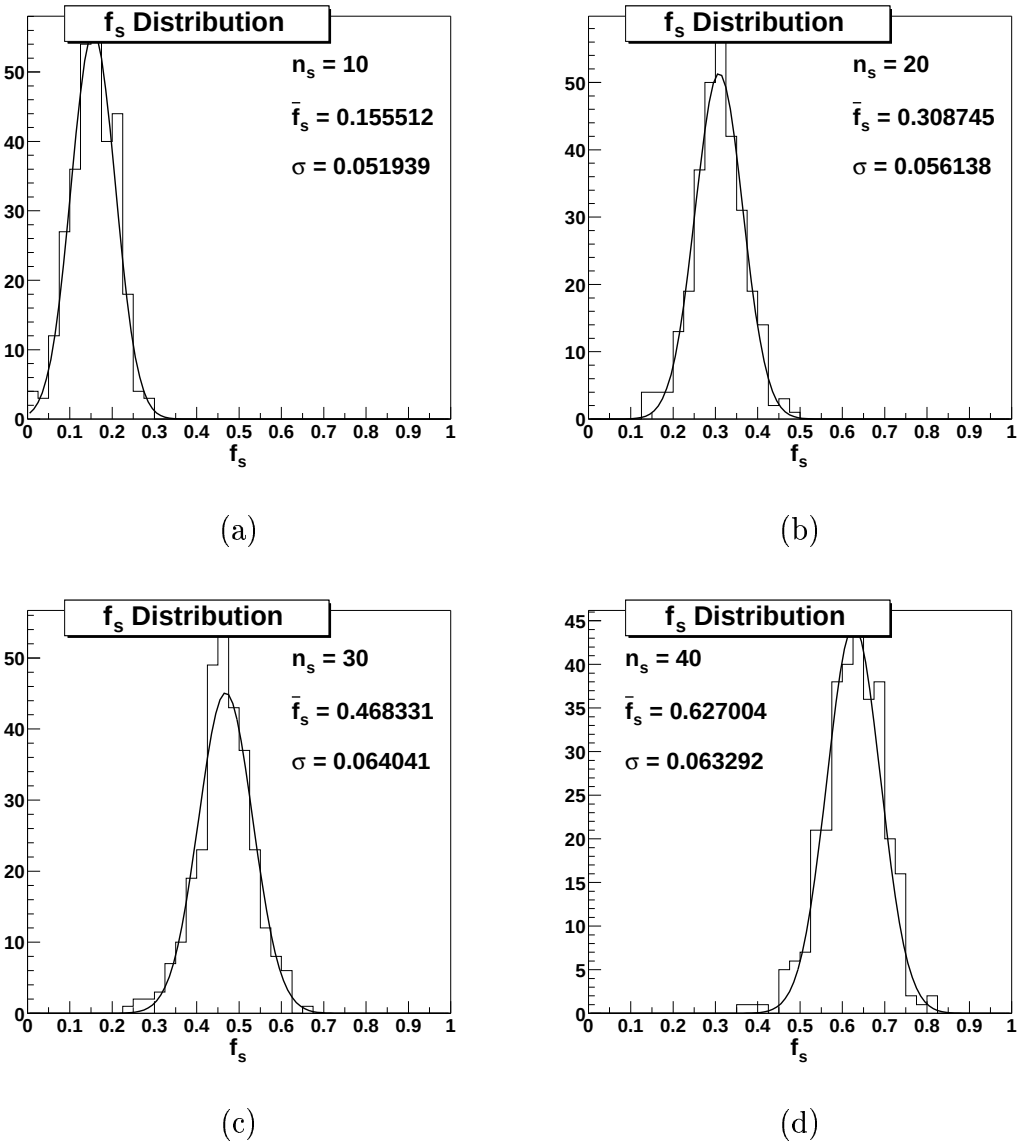
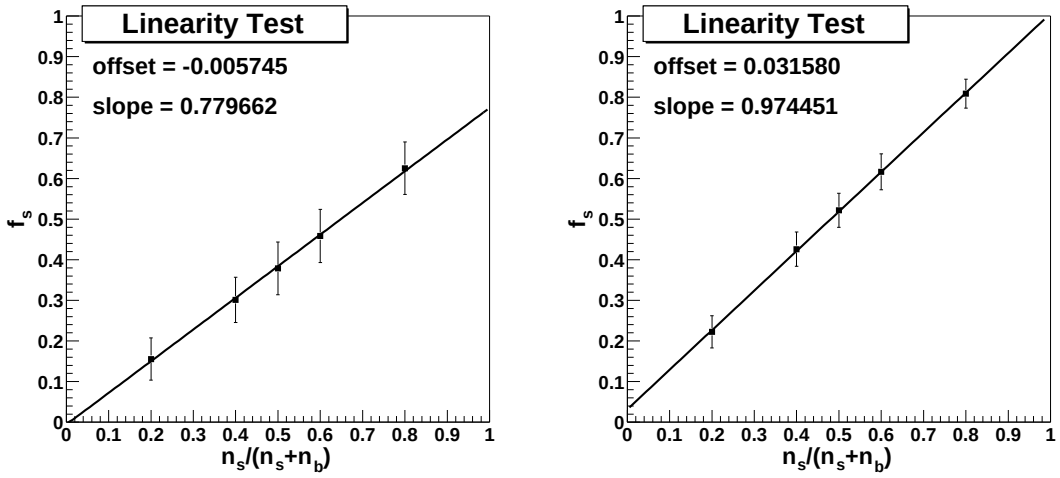


Figure 5.8: Distributions of the extracted f_s in Monte Carlo tests of the Direct Likelihood Analysis.



(a) (b)

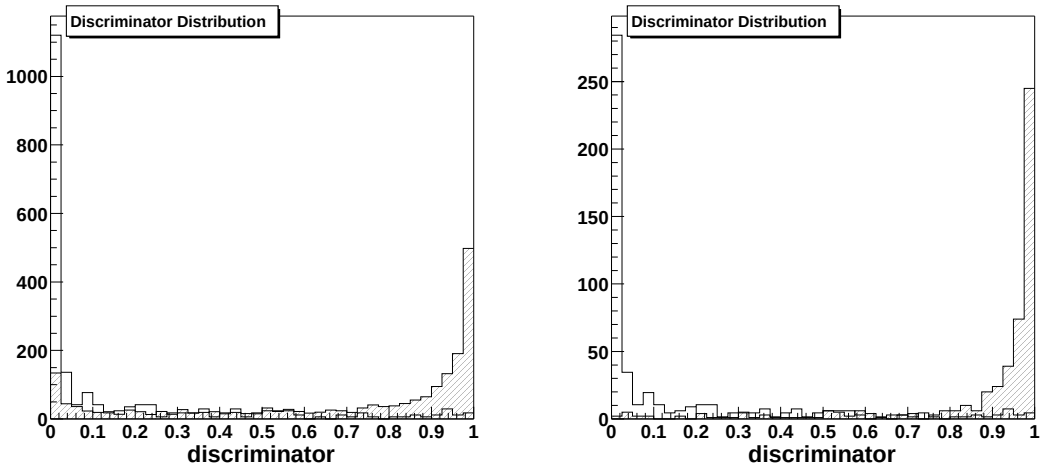
Figure 5.9: The linearity tests in the Direct Likelihood Analysis. In (a) the Monte Carlo signal sample consists of all the events which pass selection cuts. In (b) the Monte Carlo signal sample consists of events that satisfy jet-parton match requirement.

the extracted f_s is expected to be equal to the true fraction of signal events.

Both the Direct Likelihood Function and the Discriminator Likelihood Function can be used to do data analysis. However, each method has its associated systematic uncertainties that need to be assessed separately. In this thesis, the Discriminator Likelihood Function is used. The analysis using the Direct Likelihood Function is left for future work.

5.9 QCD Background Effects

QCD multijet events constitute the fake background in the event sample. At this early stage of the DØ RunII experiment, different aspects of the hardware and software have not been fully understood, so a significant fraction of QCD events are included after preselection cuts are applied to the data sample. It is important to



(a) (b)

Figure 5.10: The discriminator distributions for signal (shaded histogram) and background (open histogram) Monte Carlo events. (a): All the signal events which pass selection cuts are included; (b): Only those signal events that pass jet-parton match requirement are included. In each plot, the two histograms are properly scaled so that they contain the same number of entries.

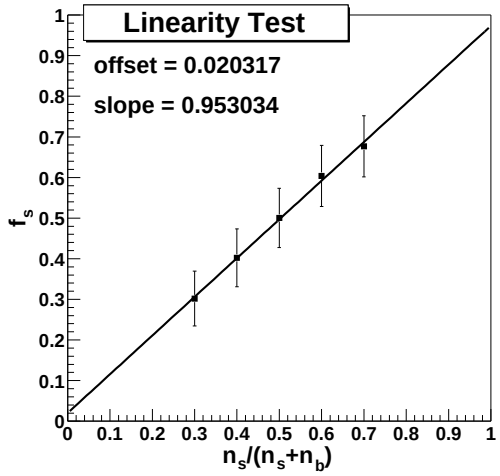


Figure 5.11: The linearity test in the Discriminator Likelihood Analysis.

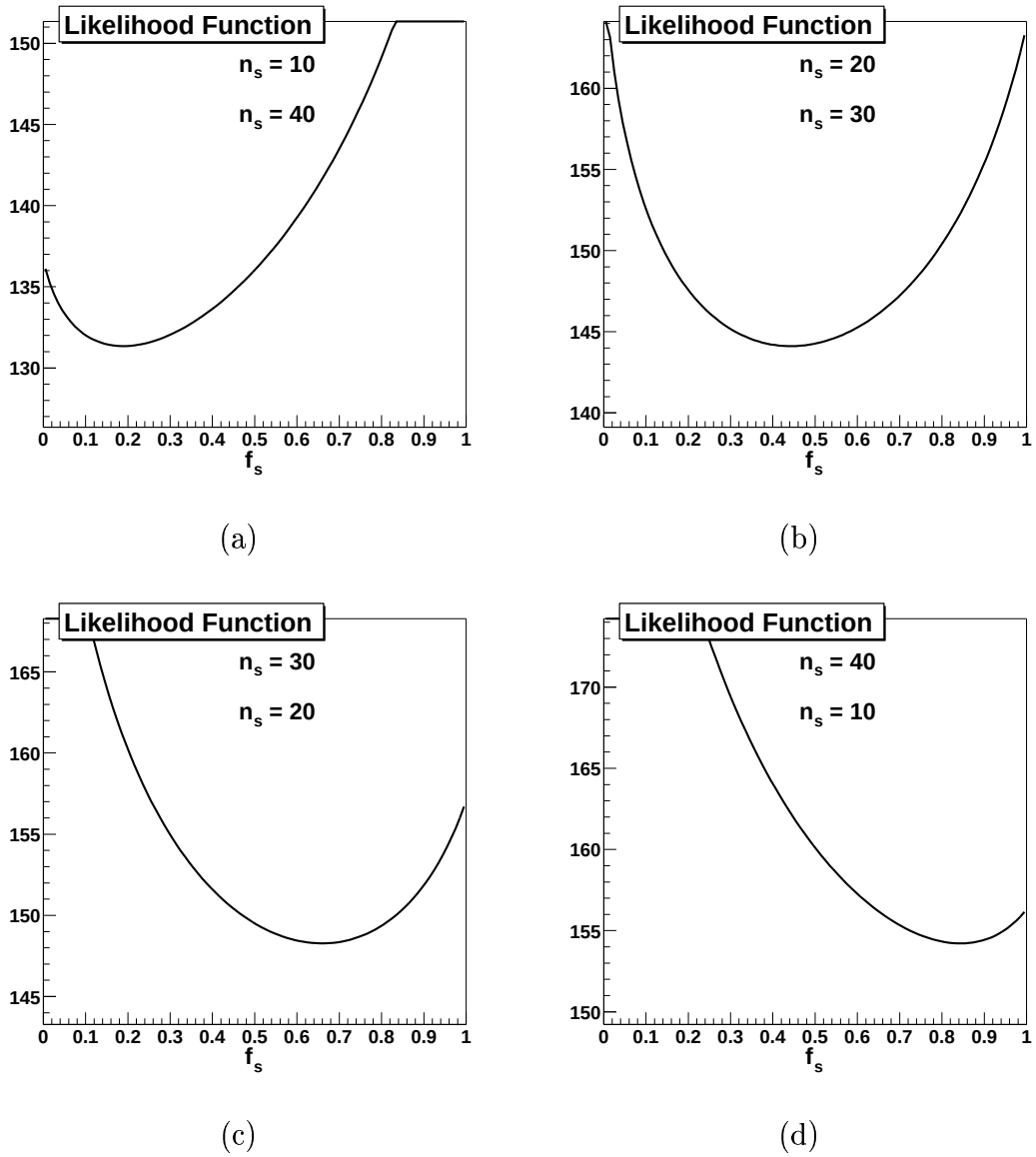


Figure 5.12: Shapes of the Discriminator Likelihood Function for randomly selected Monte Carlo datasets with different number of signal events.

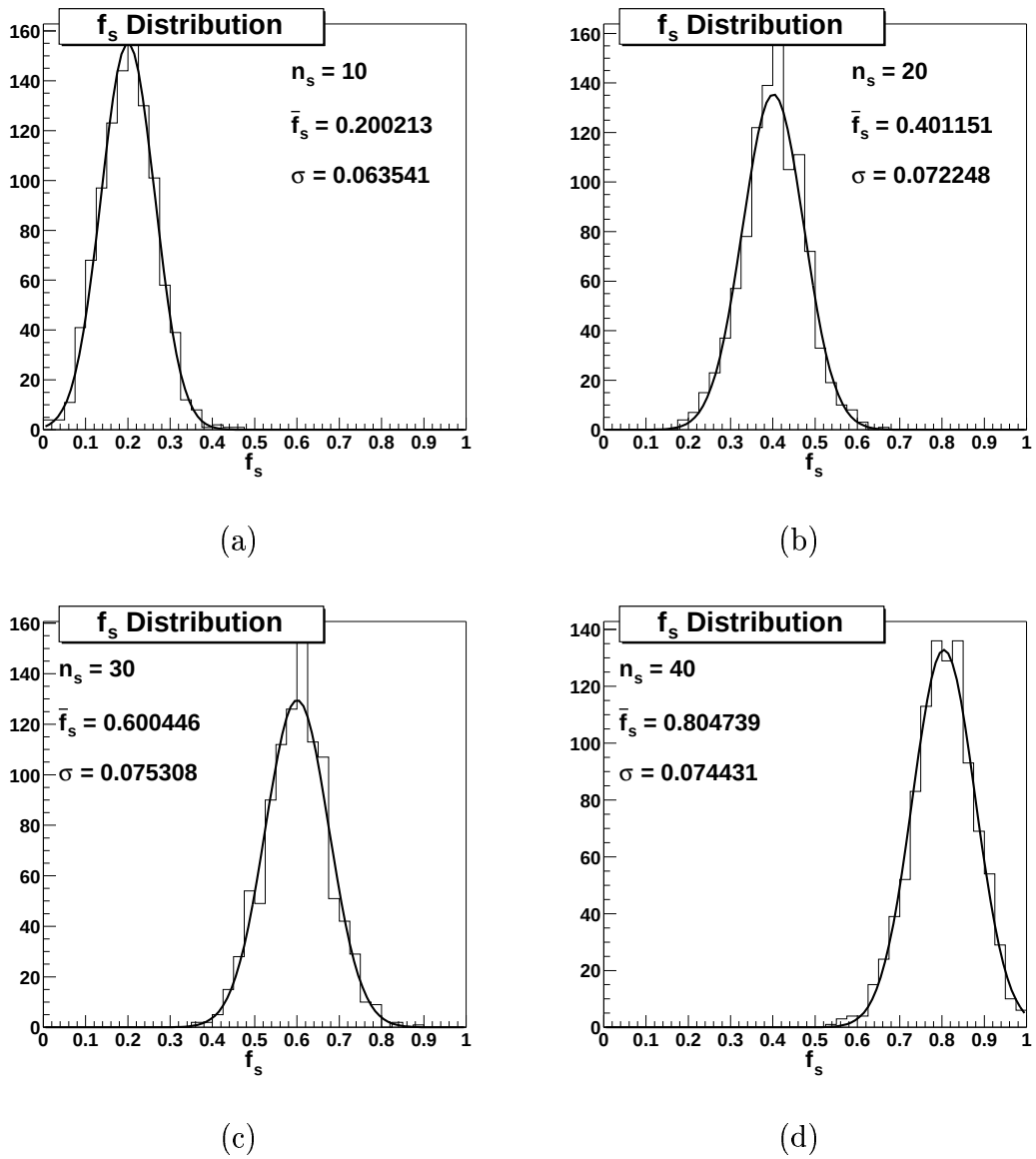
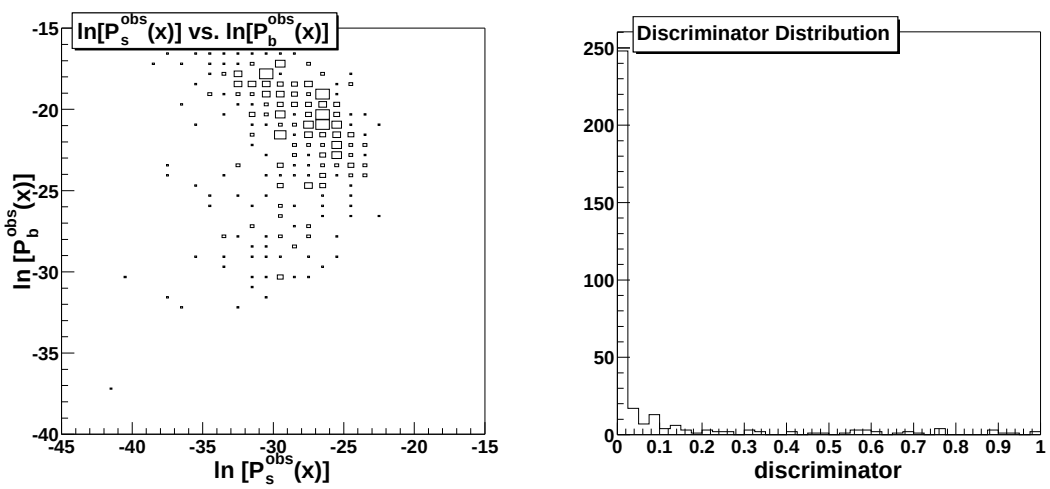


Figure 5.13: Distributions of the extracted f_s in Monte Carlo tests of the Discriminator Likelihood Analysis. For each plot, the test dataset is randomly selected with different number of signal events, and the total number of events in each test dataset is 50.

study the effect of QCD events present in our dataset.

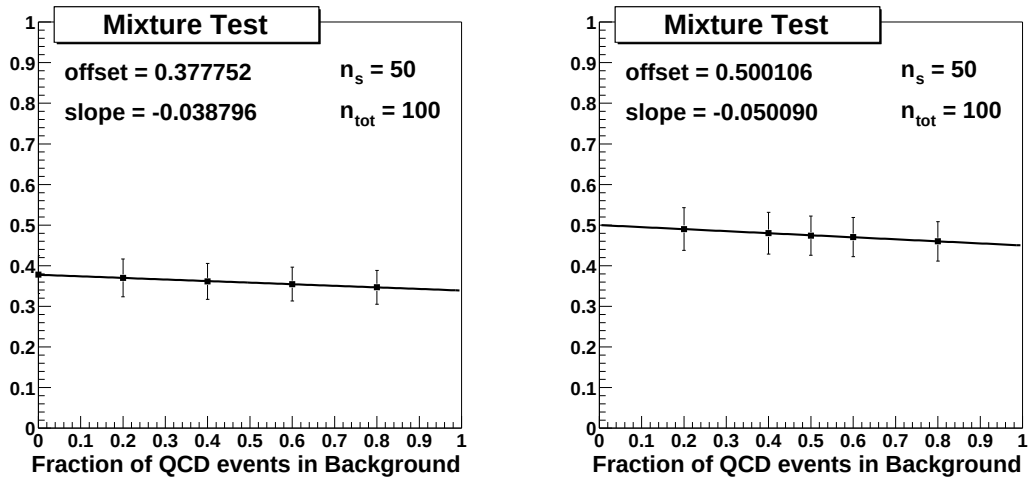
The QCD control sample for the tests is selected from data. Each of those events does not contain electrons that pass all the quality cuts, but contains one electron candidate that does not have an asosciated track. All the other selection cuts are applied to obtain the QCD events. The observation probability density distributions are shown in Figure 5.14(a). The discriminator distribution is shown in Figure 5.14(b).



(a) (b)

Figure 5.14: (a): The observation probability density distribution for the QCD multijet data events. (b): The discriminator distribution for the QCD multijet data events.

We did a mixture test to see the effect of QCD events in the dataset. In this test, we put certain number of QCD events into the Monte Carlo test dataset, and calculate the extracted f_s as a function of the fraction of QCD events. Figure 5.15(a) shows the mixture test result with the Direct Likelihood Analysis, and Figure 5.15(b) shows the mixture test result with the Discriminator Likelihood Analysis.



(a) (b)

Figure 5.15: (a): The result of the mixture test. The Direct Likelihood Function is used to extract f_s. (b): The result of the mixture test. The Discriminator Likelihood Function is used to extract f_s.

5.10 Method Comparison

Both the Direct Likelihood Analysis method and the Discriminator Likelihood Analysis method can be used to do analysis. However, one closer look at the Discriminator Likelihood Analysis method reveals one problem. The principle of this method is to fit the discriminator distribution of the dataset with the D_s and D_b histograms from Monte Carlo events. Given limited number of Monte Carlo events currently available, the D_s and D_b histograms are not well populated, leading to a statistical uncertainty term to the result that has to be estimated with care.

Because of this reason, we do a complete data analysis with the Direct Likelihood Analysis method, and only do a preliminary data analysis with the Discriminator Likelihood Analysis method.

CHAPTER VI

DATA ANALYSIS

This chapter first describes the selection cuts used to obtain the dataset for analysis, and the associated efficiencies for the cuts, then it presents the data analysis results using the Matrix Element Method.

6.1 Trigger Selection

We only select events that pass the trigger `EM15_2JT15`, which is designed specifically to capture the top production process events in the electron + jets channel. The corresponding Level1, Level2, and Level3 trigger criteria are listed in Table 6.1.

In each of the three trigger levels, this trigger looks for an EM object and two jet objects. The definiton of the jet object is inclusive in each level: one of the jet requirements is fulfilled by the EM object.

6.2 Data Sets

We use the DØ data accumulated during the data taking period Aug. 2002 - Jun. 2003, which are processed by the DØ reconstruction program `RECO` version P13. The sample consists of two subsets, each of which is reconstructed by a subversion of the P13 RECO release. The integrated luminosities and the approximate number

Trigger Level	Specifications
Level 1	one Calorimeter EM trigger tower with $E_T > 10$ GeV, and two towers with $E_T > 5$ GeV.
Level 2	one EM object with $E_T > 10$ GeV and EM fraction > 0.85, and two jet candidates with $E_T > 10$ GeV.
Level 3	One loose EM object with $E_T > 15$ GeV and shower-shape cut, and two jet candidates with $E_T > 15$ GeV.

Table 6.1: Details of the DØ top trigger EM15_2JT15. This trigger only uses calorimeter information.

of events for each subset are listed in Table 6.2. A conservative 10% uncertainty is assigned to the estimated integrated luminosity.

6.3 Offline Selection Cuts

The purpose of the offline selection cuts is twofold: (a) use physics object quality cuts to remove fake physics objects in an event, and (b) use kinematic cuts to remove background events while maintaining reasonable acceptance for signal events. In traditional cross section analysis, the cuts used here are referred to as "pre-selection cuts", after which additional topological cuts are applied. Details of the cuts are listed in Table 6.3. After the selection cuts, major background in the sample are W+4jets events, and QCD multijet events.

We require exactly four jets in the final state. Figure 6.1 shows the jet multiplicity

RECO Version	Integrated Luminosity (pb^{-1})	Number of events
p13.05.00	26.51	$\approx 2.85 \times 10^6$
p13.06.01	80.28	$\approx 8.63 \times 10^6$
Total	106.79±10.7	$\approx 11.48 \times 10^6$

Table 6.2: Data samples used in this analysis. The integrated luminosities and the number of events are calculated for the trigger EM15_2JT15. Only events in good runs are included in the calculation.

Physics Objects	Cuts
Electron	Electron ID cuts $E_T > 20$ second electron veto
Jets	jet ID cuts $E_T > 15$ GeV number of jets = 4
Missing ET	MET > 20 GeV $\Delta\phi$(electron, ν_e) > 0.5
W	$\lvert\eta\rvert < 2.0$ $E_T < 60$ GeV

Table 6.3: Details of the selection cuts.

distribution for signal events selected with at least four jets. Over 80% of those signal events are retained with the requirement of four jets.

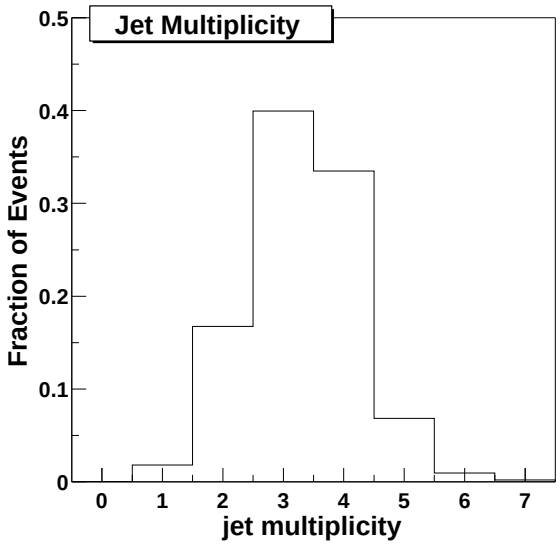


Figure 6.1: The jet multiplicity distribution for signal events. These events pass all the selection cuts, except the jet multiplicity cut.

6.4 Selection Efficiency

We divide the total selection efficiency for $t\bar{t} \rightarrow e + jets$ events into the following parts:

- trigger efficiency
- acceptance and reconstruction efficiency
- electron ID efficiency
- four jet efficiency
- kinematic cut efficiency

The following sections explain the details to calculate each efficiency term.

6.4.1 Trigger Efficiency

As described in Section 2.11, the EM15_2JT15 trigger is used for this analysis. It requires one high p_T electron, and two high p_T jets. Since in the three levels of triggers, a jet is defined as a cluster of energy with minimum quality cuts, an electron which fires the EM trigger term will be identified as a jet, and the trigger criteria is equivalent to one high p_T electron and one additional high p_T jet. The trigger efficiency is calculated from data by factorizing the trigger efficiency into terms for single physics objects. The single physics object trigger efficiency for the electron and the jet is calculated individually, and combined to get the trigger efficiency.

The per-electron trigger efficiency is calculated using the sample of events that pass a muon trigger to ensure an unbiased sample. In this unbiased sample offline electrons are identified to yield an unbiased reference sample with a p_T spectrum. The fraction of these offline electrons which fire the electron trigger is a measure of the trigger efficiency with respect to offline p_T. Folding this efficiency with the electron p_T distribution from the Monte Carlo signal sample, the electron trigger term efficiency (ε_{EM}) is found to be (88.1±5)%.

The per-jet trigger efficiency as a function of jet p_T is derived following the same procedure as for the electron. Folding the trigger efficiency with the jet p_T distribution in the Monte Carlo signal sample gives a jet trigger term efficiency (ε_{Jet}) of (85.1±4.9)%.

In the folding processes to calculate ε_{EM} and ε_{Jet}, the electron and jet p_T distributions are obtained using thosee Monte Carlo signal events that pass the offline selection cuts. Since offline selection efficiencies are calculated with samples that are

not biased with respect to the trigger EM15_2JT15, the trigger efficiency calculated from ε_{EM}, and ε_{Jet} forms a multiplicative term to the offline efficiency.

Events passing our offline selection cuts contain one electron and four jets in the final state. So the total trigger efficiency can be calculated from the following formula:

$$\varepsilon_{TRIG} = \varepsilon_{EM} \times [1 - (1 - \varepsilon_{Jet})^4]. \tag{6.1}$$

The term $(1 - \varepsilon_{Jet})^4$ denotes the probability that none of the four jets fires the jet trigger term. Due to the presence of four jets, the trigger efficiency with respect to the jet requirement is close to 100%. The total efficiency is mainly determined by the electron trigger term ε_{EM}, and the value is give as $(88 \pm 1.0 \pm 5)\%$.

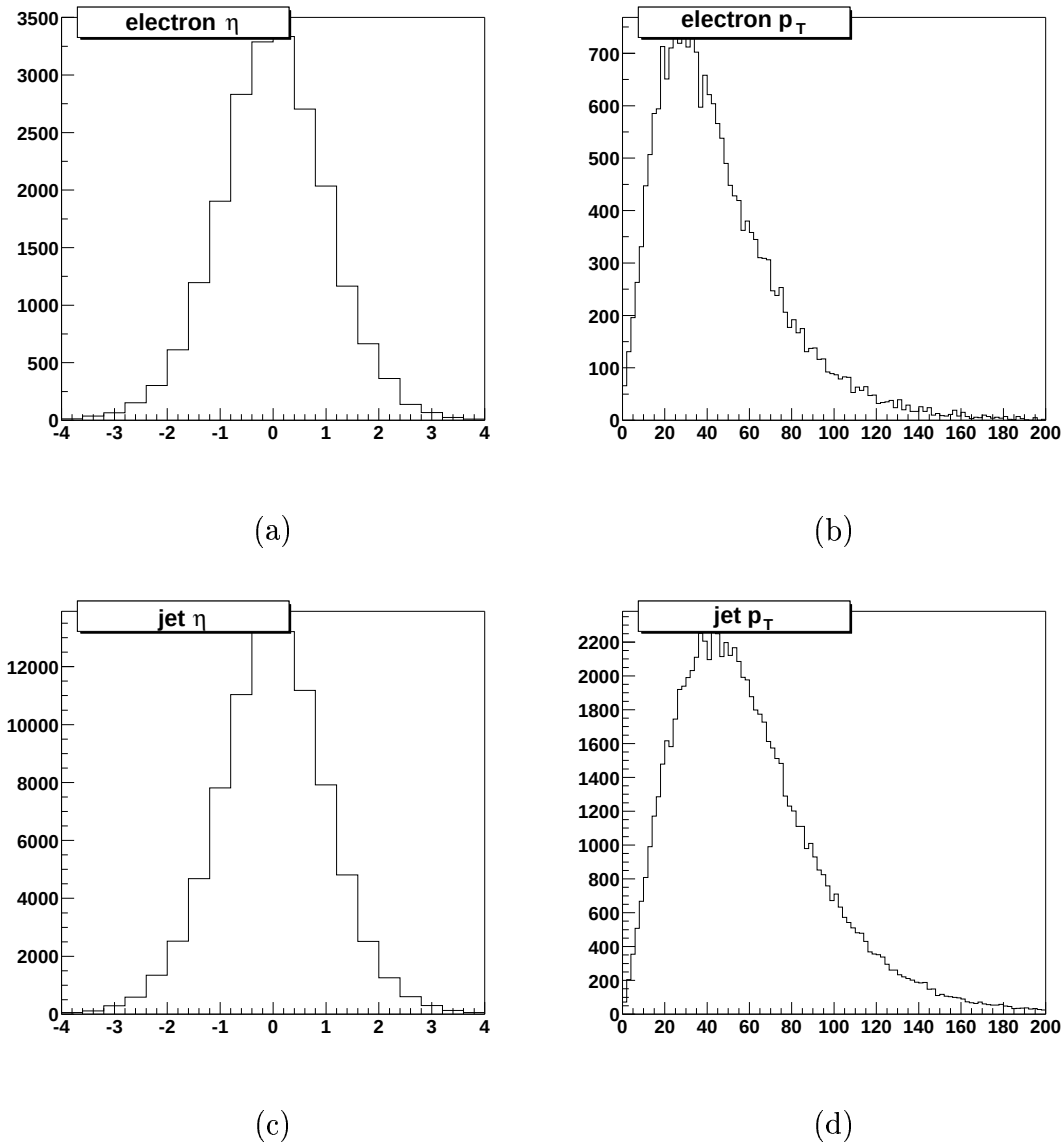
6.4.2 Acceptance And Reco. Efficiency

This efficiency term is calculated by applying the following loose cuts to the signal Monte Carlo sample:

- at least one reconstructed EM object in the central calorimeter fiducial region with $p_T > 20$ GeV;
- at least four reconstructed jets with $|\eta| < 2.0$, and $p_T > 15$ GeV.

The fiducial region requirement to EM objects rejects electrons in the 16 calorimeter module boundary regions in the $r - \phi$ plane of the central calorimeter, and those close to the edge of the central cryostat, due to poor energy resolution in those regions. These fiducial cuts reduces the acceptance by 20%.

No quality cuts are applied to electron or jets in this stage. Figure 6.2 shows the η and p_T distributions of electrons and jets in the signal events, depicting the high p_T and low $|\eta|$ features of the final-state particles. The efficiency for the acceptance and reconstruction cuts is given as $(46.18 \pm 0.47)\%$.



(a) (b)

(c) (d)

Figure 6.2: η and p_T distributions of electrons and jets of the signal process. Parton-level information is used.

6.4.3 Electron Identification Efficiency

Electron Identification

In the acceptance cuts in Section 6.4.2, EM objects are selected with no quality cuts, and contain a large fraction of fake electrons, which could be photons or jets out of the huge QCD background. We apply a set of quality cuts to reject fake electrons.

Since the electromagnetic particle loses most of its energy in the EM section of the calorimeter, we require that the fraction of cluster energy in the EM section is greater than 90%:

$$f_{EM} = E_{EM}/E_{TOT} > 0.9, \tag{6.2}$$

where E_{EM} is the cluster energy in the EM section of the calorimeter, and E_{TOT} is the total energy inside the cluster cone.

We also expect that an electron cluster should have a typical longitudinal and lateral development pattern compatible with Monte Carlo electrons. To this end, we study the value of the energy deposited in each layer of the calorimeter, and compare the pattern with the typical pattern of an Monte Carlo electron. A χ^2 probability value is assigned based on this comparison. We put a cut on this probability:

$$\chi^2 < 20 \tag{6.3}$$

We also require the EM cluster to be isolated. We find the leading p_T tower of the cluster as a centroid, and draw two cones of different sizes around the centroid. Then we sum up the EM section energy of all the towers within the small cone, and sum up the hadonic and EM section energy of all the towers within the large cone. The smaller cone has a radius of 0.2, and a total energy of $E_{EM}(R < 0.2)$; while the larger cone has a radius of 0.4, and a total energy of $E_{TOT}(R < 0.4)$. Then the

isolation requirement is formulated as:

$$f_{iso} = \frac{E_{TOT}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)} < 0.15 \tag{6.4}$$

Besides the selection requirements to the EM cluster above, we also require an associated track to the EM cluster. A track is defined as a sequence of hits in the central tracking detectors, including SMT and CFT, whose pattern matches the trajectory of a particle. Although different track finding algorithms build different types of tracks, the reconstructed tracks can be classfied into two groups:

- global (3-D) tracks: tracks built from CFT and SMT. These tracks are required to have at least 15 hits in CFT and 4 to 8 hits in SMT.
- 2-D tracks: tracks that are produced using only CFT axial information.

For a EM cluster and a reconstructed track a χ^2 probability is calculated. With a global track, the χ^2 probability is defined as:

$$\chi^2 = (\frac{\delta\phi}{\sigma_{\phi 3}})^2 + (\frac{\delta z}{\sigma_z})^2 + (\frac{E_T/p_T - 1}{\sigma_{E/p}})^2, \tag{6.5}$$

while with a 2-D track, the χ^2 probability is defined as:

$$\chi^2 = (\frac{\delta\phi}{\sigma_{\phi 2}})^2 + (\frac{E_T/p_T - 1}{\sigma_{E/p}})^2, \tag{6.6}$$

where $\delta\phi$ and δz are the differences in ϕ and z between the electron candidate and the track at the EM3 floor of the central calorimeter, and E_T/p_T is the ratio of the measured transverse energy of the electron candidate and the measured p_T of the track. The weights $\sigma_{\phi 3}$, $\sigma_{\phi 2}$, σ_z, and $\sigma_{E/p}$ are the root-mean-square of the experimental distributions of the associated quantities (ϕ with the global tracks, ϕ with the 2-D tracks, z, and E_T/p_T, respectively). The EM cluster is matched to a track if the χ^2 probability is larger than 1%.

Efficiency Calculation

We define the acceptance and reconstruction cuts on EM objects in Section 6.4.2 as preselection EM cuts, since they provide loose electron candidates. We define the EM quality cuts discussed above as tight EM cuts. The electron identification efficiency (ε_e) is defined as the probability that electrons selected by the preselection cuts pass the tight EM cuts.

ε_e is calculated from $Z \to e^+e^-$ events. First we select a loose $Z \to e^+e^-$ sample (Sample I), where each event contains exactly two electron candidates, one required to pass the tight cuts while the other is required to pass the preselection cuts. We tighten the cuts on the preselection electron candidate to obtain a tight $Z \to e^+e^-$ sample (Sample II), where both electron candidates pass the tight cuts. We then plot the invariant mass spectrums for the two samples, as shown in Figure 6.3.

Each distribution clearly shows a Z mass peak of 90 GeV. On both distributions we set a mass window of 20 GeV centered at the Z boson mass peak, where the majority events are $Z \to e^+e^-$ events. The small number of background events in each Z sample is estimated using the technique of side bands: Assuming that the number of background events decreases linearly with respect to the invariant mass around the mass window, we take the average number of events in two mass regions of 20 GeV on both sides of the mass window, and subtract it from the number of events in the mass window.

The number of estimated $Z \to e^+e^-$ events in the mass window in each sample is denoted n_1 and n_2 for Sample I and Sample II respectively. If n is the number of Zee events when only preselection cuts are applied to both electrons, then n_1 and

n_2 satisfy the equations:

$$\begin{cases} n_1 = n \cdot [1 - (1-\varepsilon)^2] \\ n_2 = n \cdot \varepsilon^2 \end{cases} \tag{6.7}$$

where $(1-\varepsilon)^2$ is the probability that none of the two preselection electrons pass the tight cuts. The electron identification efficiency can thus be derived:

$$\frac{n_2}{n_1} = \frac{\varepsilon_e^2}{1-(1-\varepsilon_e)^2} \quad \Rightarrow \quad \varepsilon_e = \frac{2n_2}{n_1+n_2} \tag{6.8}$$

The numbers of signal and background events in each sample is listed in Table 6.4. The efficiency obtained from this method is (65.49±1.73)%.

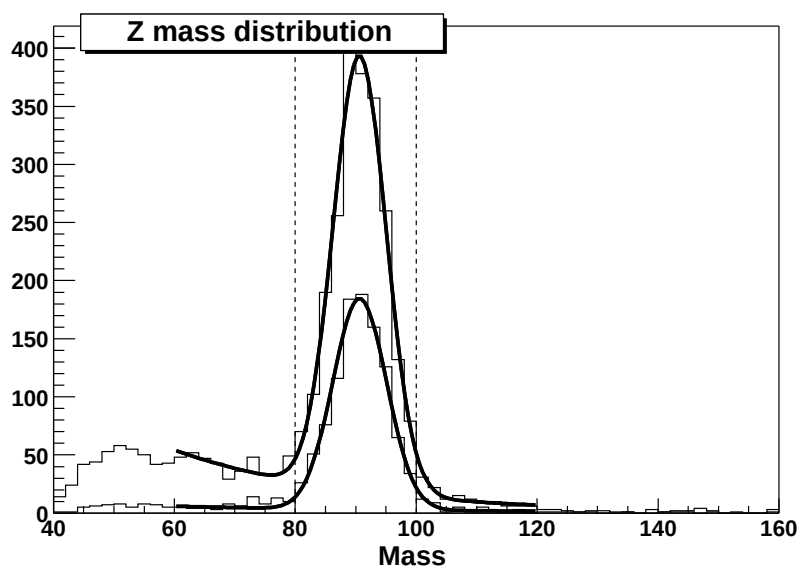


Figure 6.3: The invariant Z mass distributions in $Z \to e^+e^-$ events. The EM identification efficiency is obtained from $Z \to e^+e^-$ events. The dashed lines enclose the Z mass window used to calculate the electron identification efficiency.

The main systematic uncertainty comes from the estimation of background events in the mass window in each sample. To estimate the systematic uncertainty, we use

	Sample I	Sample II
Total Number of events	2953	1002
Num. of Signal Events	1986	967
Num. of Background Events	207	35
Fraction of Background Events	7.0%	3.5%

Table 6.4: Numbers of signal and background events in Sample I and Sample II in the calculation of the electron identification efficiency. All the numbers correspond to the mass window of (80, 100). The numbers of background events are estimated using side bands.

the following formula to fit the two distribution histograms:

$$f(M_Z) = a_0 \exp(-a_1 M_Z) + a_2 \exp[-\frac{(M_Z - a_3)^2}{2a_4^2}], \tag{6.9}$$

where M_Z is the invariant mass. The Gaussian function is used to fit the signal event distribution, and the exponential function is used to represent the background event distribution. The fitting is done in the mass window (60, 120). The fitted result is shown in Figure 6.3.

With the fitting functions we obtain a new set of n_1 and n_2. The efficiency is calculated to be 65.45%, consistent with the value obtained from the previous method.

Electrons from $t\bar{t} \to e$+jets events have different kinematic properties compared to those in $Z \to e^+e^-$ events. We use Monte Carlo samples to correct for this discrepancy. The electron identification efficiency is found to be 80.54% with the Monte Carlo $Z \to e^+e^-$ sample; while with Monte Carlo $t\bar{t} \to e$+jets events, it is found to be 77.65%. The ratio of the two values, 0.96, is multiplied to ε_e to give a final result of the electron identification efficiency as (62.85±1.67)%.

6.4.4 Jet Identification Efficiency

Jet quality cuts are necessary to remove fake jets due to electromagnetic particles and calorimeter noise. In addition to the E_T requirement in the acceptance and reconstruction cuts, we apply the following jet quality cuts in our analysis:

- The fraction of the total jet energy deposited in the electromagnetic section of the calorimeter (EMF) is in the range (0.05, 0.95). This cut is aimed at removing isolated electromagnetic clusters.

- The fraction of energy deposited in the coarse hadronic section of the calorimeter (CHF) is required to be less than 0.4. This cut is aimed at removing jets which are clustered around noise in the CH section.

- The number of towers containing 90% of the jet energy ($n90$) is required to be greater than 1, to removed jets clustered from a single hot tower.

- The ratio of the highest to the next-highest transverse energy cell in the calorimeter ($HotF$) is required to be less than 10, to remove jets clustered from hot cells, which are calorimeter cells with high occupancy noise.

Despite the numerous quality cuts applied to avoid clustering jets on noise in the calorimeter, a large number of jets seem to originate from low energy but high occupancy noise in certain regions of the calorimeter, especially the coarse hadronic section which has the largest calorimeter weights in the jet clustering scheme. Those noise jets appear to have numerous seed towers and undergo many merges in the jet clustering process. As these jets have their energy evenly spread over the towers, we define a discriminating variable, $f90$, which is the ratio of $n90$ and the total number of towers in the jet.

A proper cut on $f90$ is needed to avoid rejecting real jets which may undergo merging even in the absense of noise. The generation mechanism of this type of fake jets shows that the number of fake jets increases as the jet E_T decreases, or as the fraction of energy in the coarse hadronic section (CHF) increases. We adopt the following additional two-dimensional selection in two E_T bins to reject the fake background:

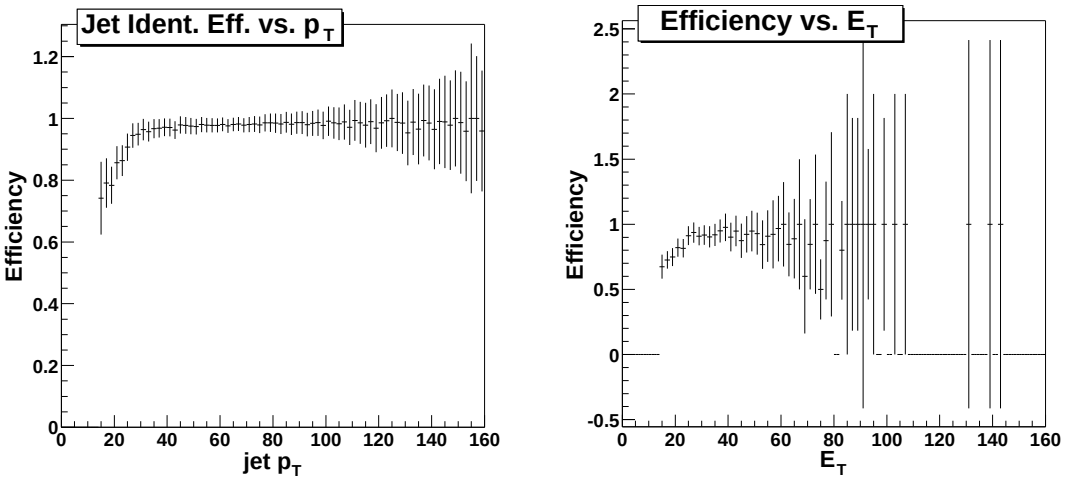
- if jet $E_T > 25$ GeV, require $f90 < 0.7 - 0.5 \times CHF$ or $CHF < 0.025$;
- if jet $E_T < 25$ GeV, require $f90 < 0.8 - 0.5 \times CHF$ or $CHF < 0.05$.

Identification Efficiency

The jet identification efficiency is defined as the probability that a jet selected by the acceptance and reconstruction cuts passes the jet quality cuts. It is measured from dijet events. An event is included in the dijet sample if it satisfies following requirements:

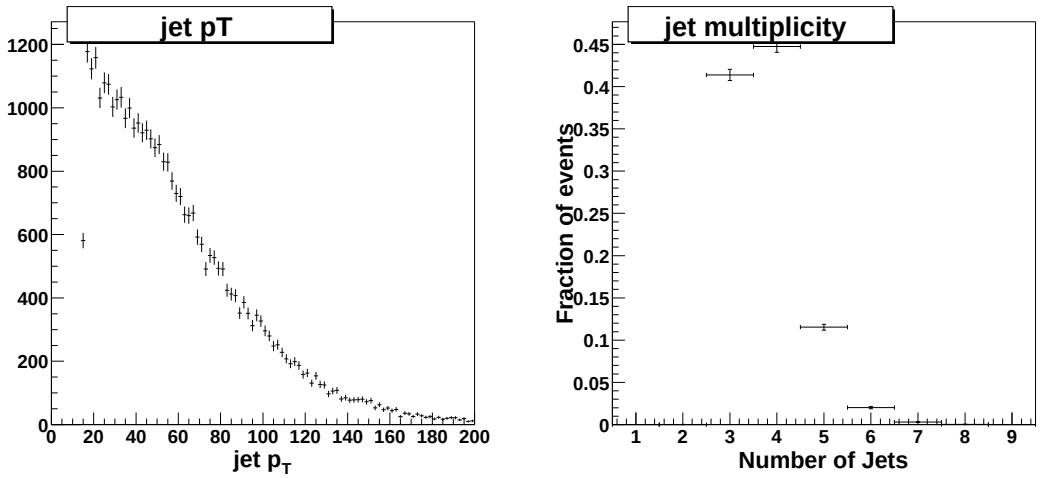
- the event does not contain electrons, photons, or muons;
- only two jet candidates in the event, one with a transverse energy greater than 15 GeV;
- the two jet candidates are back-to-back in ϕ: $\Delta\phi > 3.0$;
- one of the jet candiates passes the jet quality cuts, the other jet is called a probe jet.

The two jet candidates are real jets scattered away from each other. Then we apply the jet quality cuts on the probe jet. The fraction of events in which the probe



(a) (b)

Figure 6.4: The jet identification efficiency as a function of jet transverse energy. In (a), the efficiency is measured from dijet events. In (b), the efficiency is measured from photon+jet events.



(a) (b)

Figure 6.5: (a) shows the jet transverse energy distribution in Monte Carlo signal events. (b) shows the jet multiplicity distribution for Monte Carlo signal events. The jets are required to pass jet identification cuts.

jet passes the cuts is a measure of the jet identification efficiency. Figure 6.4(a) shows the resulting jet ID efficiency as a function of jet E_T.

Figure 6.5(a) shows the transverse momentum distribution of final-state jets in signal events. Convoluting the jet identification efficiency as a function of E_T with the E_T distribution gives an average jet identification efficiency (ε_{jetID}) of (93.81±1.00)%.

Our selection cuts requires exactly four jets in an event. Figure 6.5(b) shows the jet multiplicity distribution of Monte Carlo signal events which pass the acceptance and reconstruction cuts. The probability that exactly four jets pass the jet quality cuts(ε_{4jets}) is given by:

$$\varepsilon_{4jets} = \sum_{n\geq 4} \binom{4}{n} f(n)(1-\varepsilon_{jetID})^{n-4}\varepsilon_{jetID}^{4}, \tag{6.10}$$

where $f(n)$ is the fraction of events with exactly n jets, extracted from Figure 6.4. The value is given as (38.82±4.27)%.

To estimate the systematic uncertainty, we use photon+jet events to calculate the jet jet selection efficiency. The photon+jet events are selected with the following criteria:

- no electron or muon in the event;
- exactly one photon object which passes quality cuts;
- exactly one jet candidate object, and its azimuthal separation from the photon is close to π: $\Delta\phi > 3.0$;
- $\not{E}_T$ is smaller than 10 GeV.

Applying the jet quality cuts to the jet candidate, we obtain the jet identification efficiency as a function of jet E_T. The result is shown in Figure 6.4(b). Combining

this efficiency with the jet E_T distribution and the jet multiplicity distribution, we obtain the 4-jet selection efficiency as 33.05%.

We use the difference between the two calculation results as the systematic uncertainty, so the 4-jet selection efficiency is given as (38.82±4.27(stat.)±5.77(sys.))%.

6.4.5 Kinematic Cut Efficiency

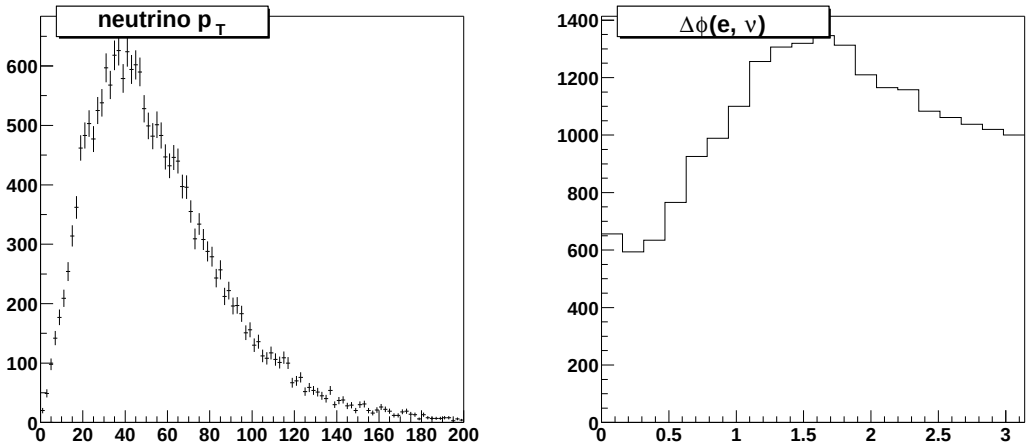
In addition to the electron and jet quality cuts, additional kinematic cuts are applied to reject background:

- The high p_T distribution of the neutrino in the final state of signal events, as shown in Figure 6.6(a), leads to a cut of $\not\!\!E_T > 20$ GeV.
- The electron and the neutrino tend to have large spatial separations in signal events, as shown in Figure 6.6(b). We apply a cut on their separation in ϕ: $\Delta\phi(\text{electron}, \not\!\!E_T) > 0.5$.
- We reject events that have more than one reconstructed electron.
- Based on the kinematic properties of the W boson which decays leptonically in signal events (shown in Figure 6.7), we require that the pseudorapidity of the W lies in the range (-2, 2), and that the transverse energy is greater than 60 GeV. The W momentum is deduced from the electron momentum and $\not\!\!E_T$.

The efficiency for these kinematic cuts is calculated using the Monte Carlo signal sample. The value is given as (73.88±1.70)%.

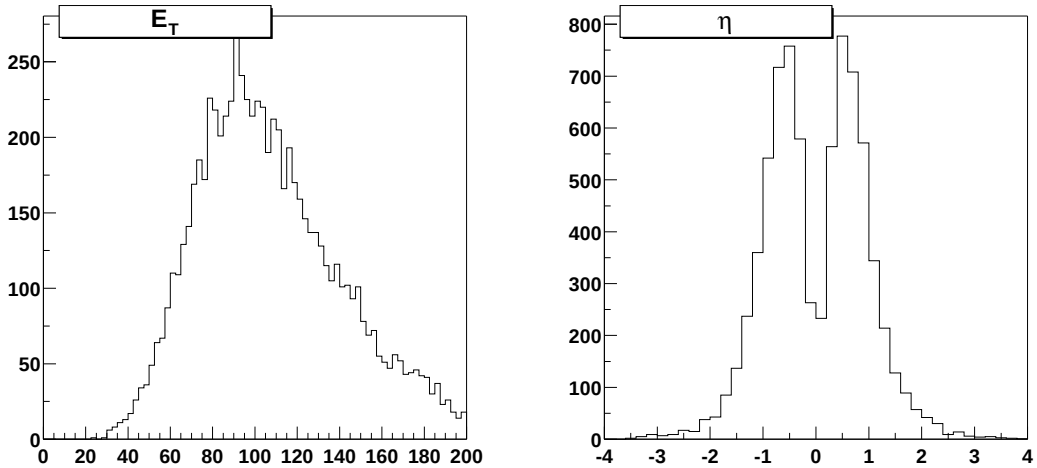
6.4.6 Total Efficiency

Each efficiency term is listed in Table 6.5. Combining all the efficiency terms together, the total efficiency is given as (7.74±0.90(stat.)±0.02(sys.))%.



(a) (b)

Figure 6.6: (a): The neutrino p_T distribution in Monte Carlo signal events. (b): The distribution of the azimuthal separation between the electron and the neutrino in Monte Carlo signal events.



(a) (b)

Figure 6.7: The kinematic distributions of the W boson which decays leptonically in Monte Carlo signal events. (a) shows the E_T distribution, and (b) shows the η distribution.

Efficiency term	Value
Acc. + Reco. Eff.	(46.18±0.47)%
Electron Selection Eff. (ε_e)	(62.85±1.67)%
Four Jets Selection Eff. (ε_{4jets})	(38.82±4.27±5.77)%
Kinematic Cut Eff.	(73.88±1.70)%
Trigger Eff. (ε_{TRIG})	(88±1±5)%
Total Eff. (ε)	(7.33±0.85±0.01)%

Table 6.5: A list of efficiency terms and the combined efficiency. Each efficiency value is followed by the statistical uncertainty. The systematic uncertainty, if available, is listed as the third percentage number in a row.

We obtain a dataset of 50 events after applying the selection cuts to the data sample. This corresponds to the "pre-selection" dataset in the traditional topological analysis, but is the final dataset in this analysis.

6.5 Direct Likelihood Analysis

We calculate the observation probability densities $P_s^{obs}(x)$ and $P_b^{obs}(x)$ for each event, and their distributions are shown in Figure 6.8. We feed the values into the Direct Likelihood Function of Equation 5.13. The shape of the negative logarithm of the likelihood function is shown in Figure 6.9(a).

To estimate the appropriate statistical uncertainty with this result, we plot the likelihood function in Figure 6.10(a). The shaded area covers 68% of the integrated area in the (0, 1) range, corresponding to a $(-\sigma, +\sigma)$ deviation from the peak point. So the maximum likelihood method gives the result:

$$f_s = 0.131 \pm 0.070 \tag{6.11}$$

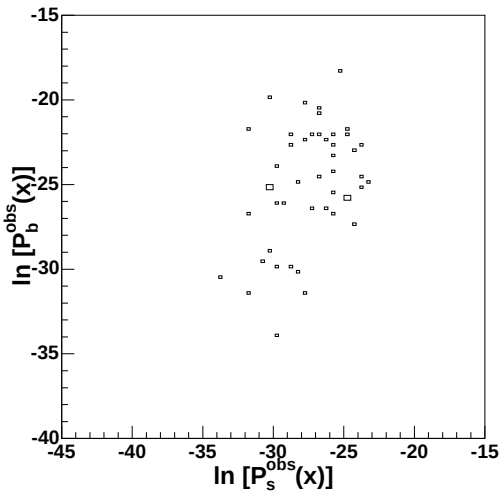
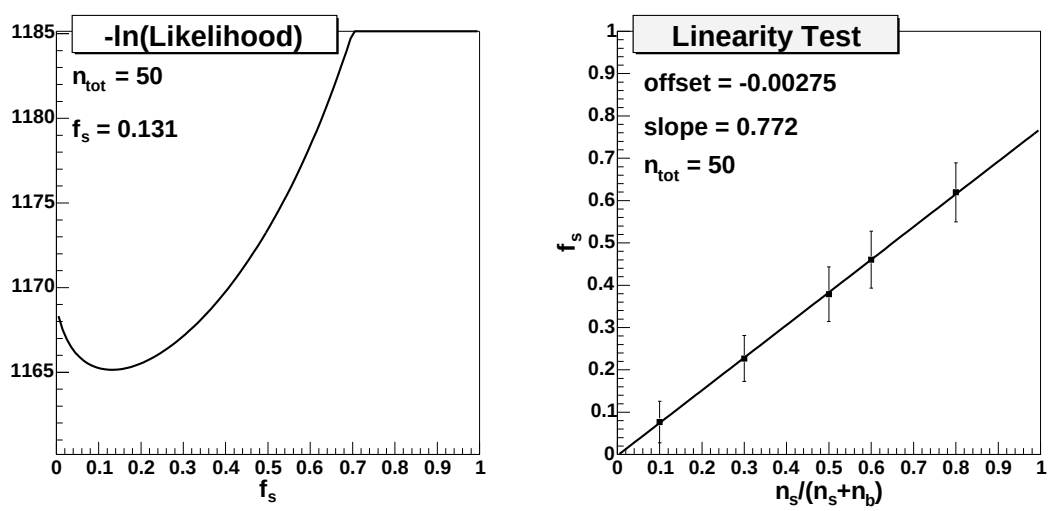


Figure 6.8: The signal and background observation probability distributions for the dataset are shown in a scatter histogram.



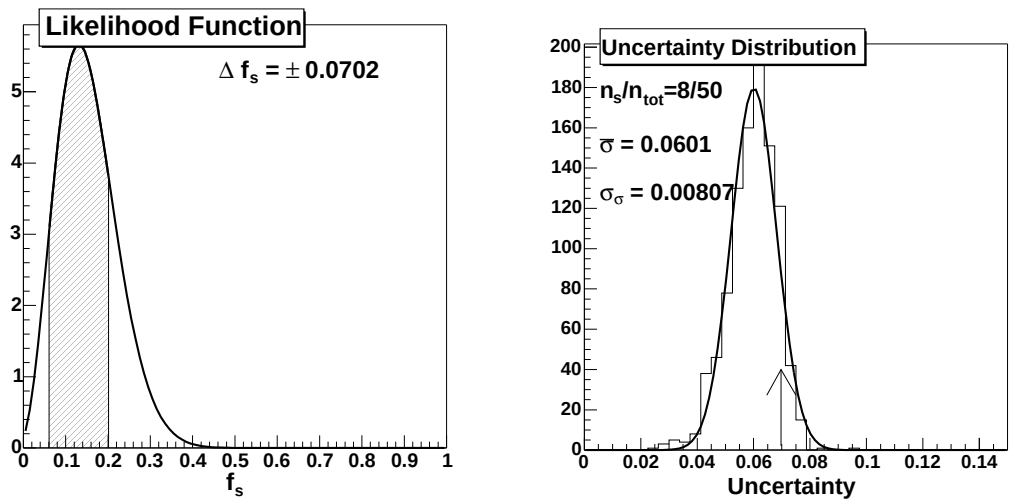
(a) (b)

Figure 6.9: (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the final dataset in this analysis. (b) shows the linearity test result using Monte Carlo events.

Figure 6.9(b) shows f_s as a function of $n_s/(n_s + n_b)$ in Monte Carlo tests. Using this linear relation, we correct f_s, and obtain:

$$f_s^{corr} = 0.170 \pm 0.091 \tag{6.12}$$

This corresponds to 8.5 signal events in the dataset.



(a) (b)

Figure 6.10: (a) show the shape of the Direct Likelihood Function for the dataset. The shaded area covers 68% of the total integrated area covered by the function. (b) shows the f_s statistical uncertainty distribution in Monte Carlo tests using the Direct Likelihood Analysis. The arrow indicates the f_s uncertainty in the data analysis.

6.5.1 Systematic Uncertainties

A number of systematic uncertainties affect the cross section measurement using the Matrix Element Method. Some of the systematic uncertainties are listed in Table 6.5, and others are calculated below.

First, we analyze the jet energy scale (JES) uncertainty, which is due to the uncertainty of the JES correction factors given by the DØ JES correction package

`jetcorr`. Since the jet energy transfer functions already take into account its effect on the matrix element calculations, we only estimate its effect on the efficiency of the selection cuts.

The uncertainty on each JES correction factor is provided by `jetcorr`. We estimate the JES uncertainty by varying the correction factor by $\pm\sigma$, and calculate the variation of the efficiency on Monte Carlo signal events.

When $+\sigma$ is applied to the correction factor, we obtaine a change of 2.91% to the total efficiency of our selection cuts. When $-\sigma$ is applied to the correction factor, a change of -6.29% is observed to the total efficiency. So the JES uncertainty is $^{+2.91}_{-6.29}\%$ for the total efficiency ε.

Then, we analyze the jet energy resolution uncertainty. This is the uncertainty caused by the choice of the jet energy transfer function in our data analysis. To estimate this uncertainty, we use the transfer functions obtained from Monte Carlo photon+jet events. The fraction of events is then extracted to be:

$$f_s^{res} = 0.161 \tag{6.13}$$

Comparing it with f_s in Equation 6.11, the jet energy resolution uncertainty to f_s is estimated to be $\Delta f_s/f_s = \pm 22.90\%$.

To estimate the systematic uncertainty caused by the choice of the parton distrubution function, we use CTEQ6L instead of CTEQ4L in the calculation of $P_s^{obs}(x)$ and P_b^{obs} for the dataset. With CTEQ6L we obtain the fraction of events:

$$f_s^{PDF} = 0.150 \tag{6.14}$$

The difference between f_s^{PDF} and f_s in Equation 6.11 gives a systematic uncertainty of $\pm 14.50\%$ to f_s.

Systematic Uncertainty Source	Affected Variable	Value
Jet Energy Scale	ε	$^{+2.91}_{-6.29}\%$
Jet Energy Resolution	f_s	$\pm 22.90\%$
Parton Distribution Function	f_s	$\pm 14.50\%$
QCD Background	f_s	$\pm 1.71\%$

Table 6.6: A list of additional systematic uncertainty sources and their values.

The uncertainty due to the presence of QCD events in the sample is estimated by the mixture test described in Section 5.9. Figure 6.11 shows the mixture test result with 2 signal events and 12 background events in the Monte Carlo test sample. This gives an uncertainty of $\pm 1.71\%$ to f_s.

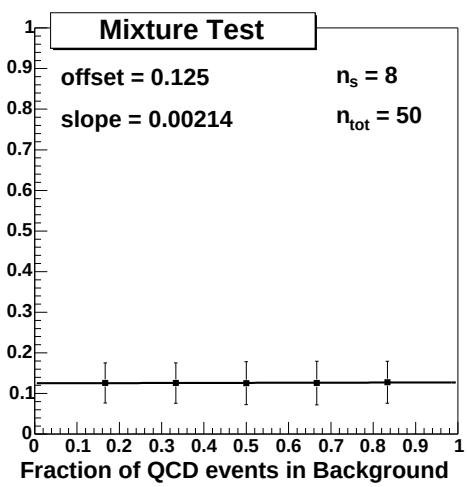


Figure 6.11: The mixture test with 8 signal events and 42 background events in the Monte Carlo test sample. This test is carried out with the Direct Likelihood Analysis.

The value for each systematic uncertainty term is listed in Table 6.6.

Combining all the systematic uncertainties together, we can calculated the cross

section for the process $t\bar{t} \rightarrow$ e+jets with Equation 5.2. Since the e+jets decay channel has a branching ratio of $BR = 12/81$, the total $t\bar{t}$ production cross section is given by:

$$\sigma = \frac{1}{BR}\sigma_{e+jets} = \frac{f_s \cdot N}{BR \cdot \varepsilon \cdot \mathcal{L}} = 7.33 \pm 4.01(\text{stat.})^{+2.13}_{-2.17}(\text{sys.})\ \text{pb} \tag{6.15}$$

6.6 Discriminator Likelihood Analysis

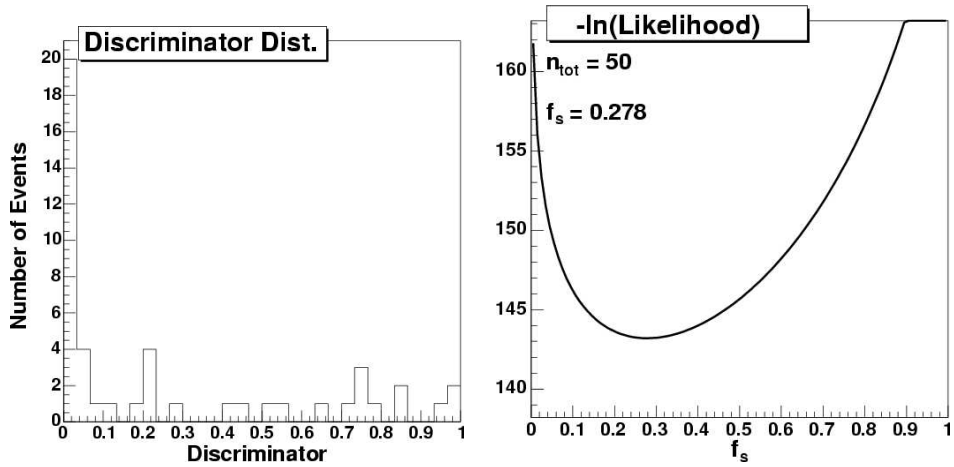
We calculate the observation probability densities $P_s^{obs}(x_i)$ and $P_b^{obs}(x_i)$ for each event, then derive the discriminator value $D(x_i)$. The discriminator distribution for the dataset is shown in Figure 6.12(a).

We then extract $P_s'(D_i)$ and $P_b'(D_i)$ for each event by using the Monte Carlo discriminator distributions obtained in Chapter V. The negative logarithm of the Discriminator Likelihood Function is shown in Figure 6.12(b). The function achieves the minimum value at $f_s = 0.278$, which corresponds to having 14 signal events in the dataset.

The statistical uncertainty for this value is estimated in the same way as in the Direct Likelihood Analysis. Figure 6.13(a) shows the shape of the Discriminator Likelihood Function. The shaded area covers 68% of the total integrated area covered by the function, corresponding to a $(-\sigma, +\sigma)$ deviation from the peak point. Figure 6.13(b) shows the statistical uncertainty distribution using Monte Carlo events. So the maximum likelihood method gives:

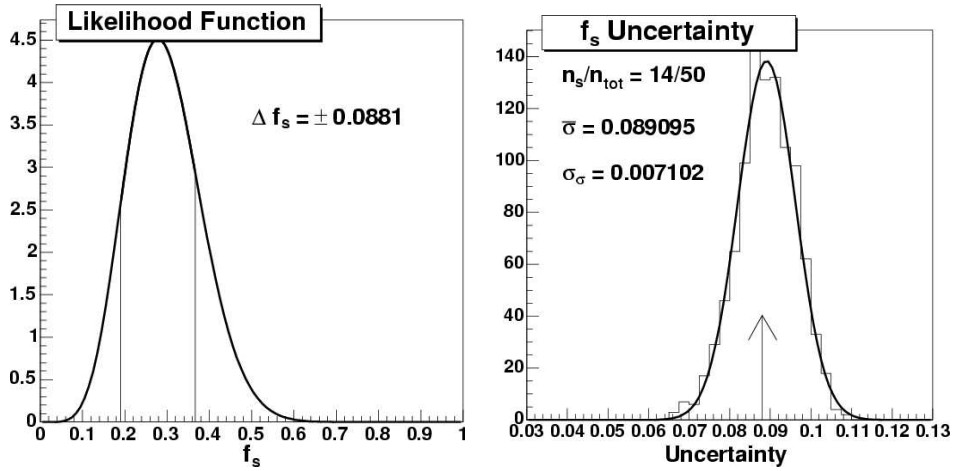
$$f_s = 0.278 \pm 0.088(\text{stat.}) \tag{6.16}$$

As discussed in Section 5.10, we will not do a complete analysis on the systematic uncertainties for this Discriminator Likelihood Method, due to its intrinsic problem with limited Monte Carlo events. As a result, we will not give a complete result to



(a) (b)

Figure 6.12: (a) shows the discriminator distribution for the dataset. (b) shows the shape of the negative logarithm of the Discriminator Likelihood Function for the dataset.



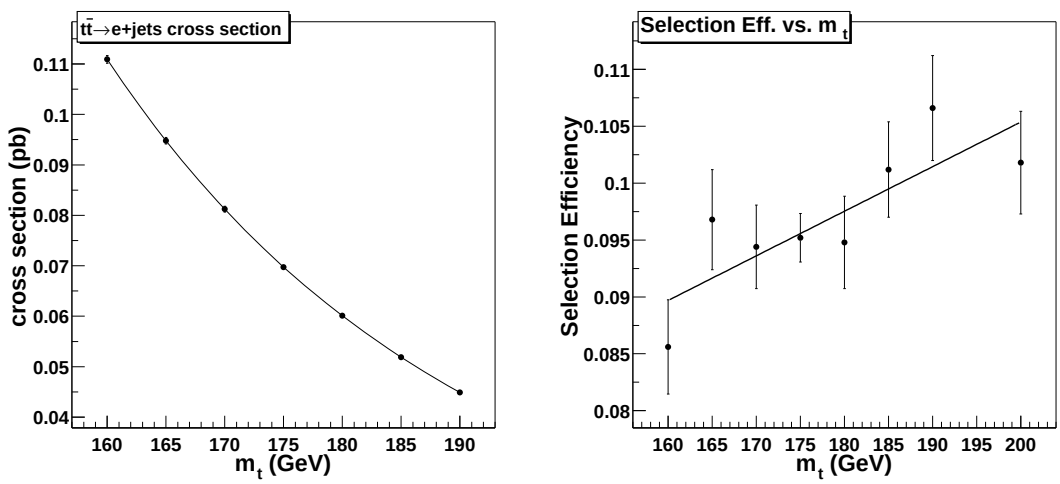
(a) (b)

Figure 6.13: (a) shows the Discriminator Likelihood Function for the data analysis. The shaded area covers 68% of the total integrated area covered by the function. (b) shows the f_s uncertainty distribution in Monte Carlo tests with the Discriminator Likelihood Analysis, where the number of signal events is 14, and the total number of events is 50 in the test sample. The arrow indicates the f_s uncertainty in the data analysis.

the cross section from this method. With f_s measured at 0.278, the expected total cross section is 11.99 pb.

6.7 Top Mass Effects

The signal process matrix element is a function of the top mass (m_t). So the kinematic properties of the signal process depend on m_t. Figure 6.14(a) shows the leading-order top pair production cross section as a function of m_t, calculated using the ONETOP Monte Carlo generator. In addition, the topology of the final-state particles depend on m_t, therefore the total efficiency of the selection cuts (ε) is a function of m_t, as shown in Figure 6.14(b).



(a) (b)

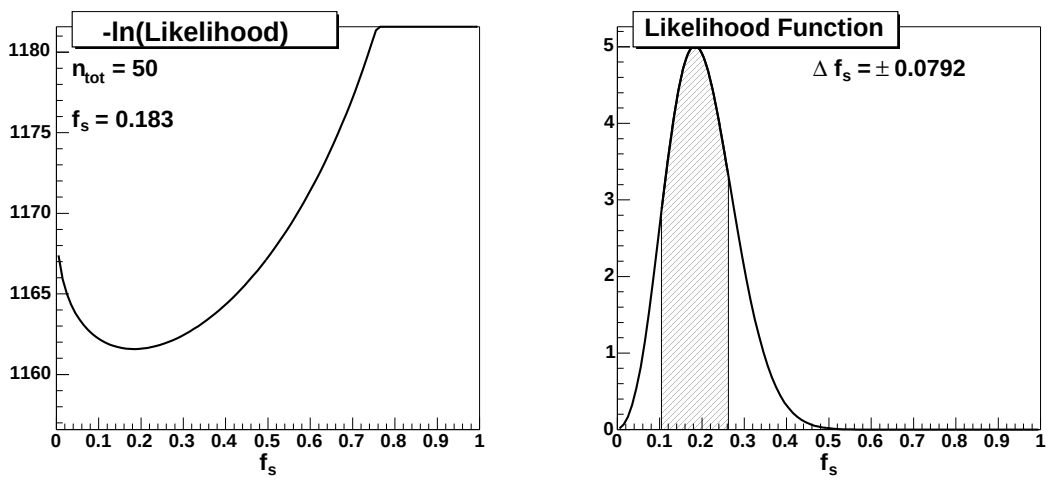
Figure 6.14: (a) shows the $t\bar{t}$ production cross section as a function of the top mass. This is the leading-order calculation by the Monte Carlo generator ONETOP. (b) shows the selection cut efficiency as a function of the top mass, obtained from Monte Carlo signal samples.

In this analysis, the top mass is assumed to be 175 GeV, consistent with the combined Tevatron RunI top mass measurement of 174.3±5.1 GeV [5]. In this section,

we present the effect of the top mass on our analysis results.

With $m_t = 170$ GeV in the matrix element calculations, the shapes of the Direct Likelihood Function and its negative logarithm are shown in Figure 6.15. The Direct Likelihood Analysis gives $f_s = 0.183 \pm 0.079$(stat.). We apply the linearity correction to f_s, and repeat the complete systematic analysis as done in the $m_t = 170$ GeV case. The total $t\bar{t}$ production cross section is given by:

$$\sigma_{170} = \frac{f_s \cdot N}{BR \cdot \varepsilon \cdot \mathcal{L}} = 10.25 \pm 4.58^{+2.98}_{-3.04} \text{ pb} \tag{6.17}$$



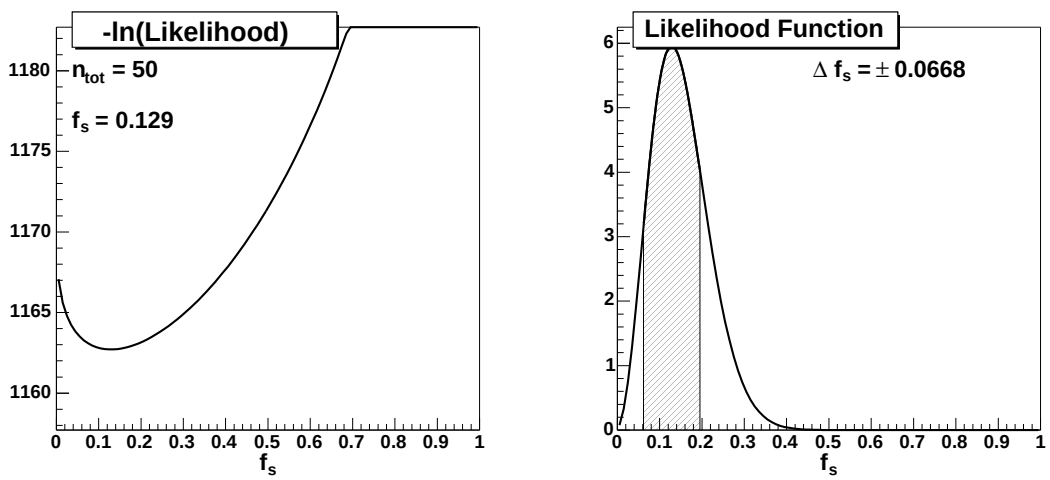
(a) (b)

Figure 6.15: (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the dataset assuming $m_t = 170$ GeV. (b) shows the shape of the Direct Likelihood Function for the dataset. The shaded area covers 68% of the total integrated area covered by the function, so its border is equivalent to a $(-\sigma, +\sigma)$ deviation from the peak point.

With $m_t = 180$ GeV in the calculation, the shapes of the Direct Likelihood Function and its negative logarithm are shown in Figure 6.16. The Direct Likelihood Analysis gives $f_s = 0.129 \pm 0.067$(stat.). We apply systematic analysis, and the total

$t\bar{t}$ cross section is given by:

$$\sigma_{180} = \frac{f_s \cdot N}{BR \cdot \varepsilon \cdot \mathcal{L}} = 7.22 \pm 3.84^{+2.10}_{-2.14} \text{ pb} \tag{6.18}$$



(a) (b)

Figure 6.16: (a) shows the shape of the negative logarithm of the Direct Likelihood Function for the dataset assuming $m_t = 180$ GeV. (b) shows the shape of the Direct Likelihood Function for the dataset. The shaded area covers 68% of the total integrated area covered by the function, so its border is equivalent to a $(-\sigma, +\sigma)$ deviation from the peak point.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, *Elementary Particle Physics* (Almquvist and Wiksells, Stockhom, 1968); S. Glashow, J.Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970); D. Gross and F. Wilczek, Phys. Rev. D **8**, 3633 (1973); *ibid*, Phys. Rev. Lett. **30**, 1343 (1973); H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973); LEP Electroweak Working Group, Report No. CERN-PPE/96-183 (1996).

[2] P. W. Higgs, Phys. Lett. **12**, 132 (1964); Phys. Rev. **145**, 1156 (1966); F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964); G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[3] E. Eichten and K. Lane, Phys. Lett. **327** 129 (1994).

[4] V. Barger and R.J.N. Phillips, MAD/PH/830 (1994).

[5] K. Hagiwara *et al.*, Phys. Rev. D **66** (2002) 01001

[6] D. Schail and PmM. Zerwas, Phys. Rev. D **45**, 3262 (1992).

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

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

[9] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[10] P. Bhat, H. Prosper, and S. Snyder, "Top Quark Physics at the Tevatron", hep-ex/9809011 (1998).

[11] M. C. Smith and S. Willenbrock, Phys. Rev. D **54**, 6696 (1996).

[12] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D **56**, 5919 (1997); **58**, 094021 (1998).

[13] S. Anderson *et al.*, "Measurement of the $t\bar{t}$ Cross Section at $sqrts = 1.96$ TeV", DØ Note 4116.

[14] CDF Collaboration, F. abe *et al.*, Phys. Rev. Lett. **84**, 216 (2000).

[15] DØ Collaboration, "ttbar production cross section in pbarp collisions at sqrt(s)=1.8 TeV", Phys. Rev. D **67**, 012004 (2003)

[16] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **64**, 03202 (2001).

[17] R. Brun *et al.*, "GEANT3 User's Guide", CERN-DD/EE 84-1, Geneva (1987).

[18] G. Jackson *et al.*, **Fermilab-TM-1991**, Fermilab, 1996.

[19] The D0 Detector, Nucl. Inst. and Methods, **A338**, 185(1994).

[20] The DØ Collaboration, The DØ Upgrade: The detector and Its Physics, (Fermilab Pub-96/357-E).

[21] DØ Upgrade Collaboration, DØ Silicon Tracker Technical Design Report, DØ Note 2169.

[22] M. Roco, The Silicon Microstrip Tracker for the DØ Upgrade, DØ Note 3553.

[23] DØ Run 2B Silicon Detector Upgrade Technical Design Report, October 11, 2001, Version 3.0.

[24] The DØ Upgrade Central Fiber Tracker Technical Design Report, (1997).

[25] D. Adams *et al.*, Fermilab-Conf-94/318-E, presented at DPF'94;
M. Atac *et al.*, Nucl. Instrum. and Meth. **A314**, 56 (1992);
M.D. Petroff and M.G. Stapelblock, IEEE Trans. Nucl. Sci. **NS-36**, 158 (1989);
M.D. Petroff and M. Atac, *ibid*, p. 163.

[26] M. Adams *et al.*, Design Report of the Central Preshower Detector for the DØ Upgrade, DØ Note 3014.

[27] Ken Del Signore *et al.*, Cabling Scheme of the Central Preshower Detector, DØ Note 3646.

[28] A. Gordeev *el al.*, Technical Design Report of the Forward Preshower Detector for the DØ Upgrade, DØ Note 3445.

[29] A. Patwa, The Forward Preshower System and a Study of the J/psi Trigger with the D0 Detector, Ph.D. thesis, SUNY at Stony Brook, 2002.

[30] J. Kotcher, Design, Performance and Upgrade of the D0 Calorimeter, DØ Note 2417 (FERMILAB-Conf-95/007-E).

[31] D. Schamberger, Calorimeter Electronics Upgrade, DØ Note 1116.

[32] Technical Design Report for the Upgrade of the ICD for DØ RunII, DØ Note 2686.

[33] DØ Collaboration, B. Baldin *et al.*, Technical Design of the Central Muon System, DØ Note 3365.

[34] DØ Collaboration, G. Alexeev *et al.*, Technical Design Report for the DØ Forward Muon Tracking Detector Based On Mini-drift Tubes,DØ Note 3366.

[35] DØ Collaboration, V. Abramov *et al.*, Technical Design Report for the DØ Forward Trigger Scintillation Counters, DØ Note 3237.

[36] DØ Collaboration, B. Jayanti *et al.*, The DØ Muon System Upgrade, DØ Note 2780.

[37] A. Lo, C. Miao, R. Partridge, "Luminosity Monitor Technical Design Report", DØ Note 3320.

[38] DØ Collaboration, Maris Abolis *et al.*, D0 RunII Level 1 Trigger Framework Technical Design Report. http://www.pa.msu.edu/hep/d0/11.framework/index.html

[39] Daniel Edmunds *et al.*, Technical Design Report for the Level 2 Global Processor, DØ Note 3402. reco

[40] F. Fleuret, The D0 Electron/Photon Analysis Package EMAnalyze, DØ Note 3888.

[41] S. Protopopescu, EMReco Algorithms, located at http://www-d0.fnal.gov/d0dist/dist/releases/current/emreco/doc/EMReco.ps.

[42] F. Fleuret, F. Beaudette, S. Protopopescu, The D0 Soft Electron Reconstruction Package SEMReco, DØ Note 3872.

[43] R. Engelmann *et al.*, Nucl. Inst. Meth., **A216**:45-56, 1983.

[44] Laurent Duflot, Simple Cone Preclustering For Cone Jets, DØ Note 3749.

[45] Gerald Blazey *et al.*, RunII Jet Physics, DØ Note 3750.

[46] Florncia Canelli *et al.*, Jet Energy Scale at DØ in Run II, DØ internal document (pending to be put into DØ Note).

[47] D. Whiteson, DØ Note 3996.

[48] Technical Design Report For an Ethernet-Based Data Acquisition System For the DØ Experiment.

[49] B. Abbott *et al.*, Phy. Rev. D **58**, 052001 (1998).

[50] Juan Estrada, Maximal Use of Kinematic Information for the Extraction of the Mass of the Top Quark in Single-lepton $t\bar{t}$ events at DØ Ph.D. thesis, University of Rochester, 2001.

[51] Douglas Olaf Carlson, Physics of Single-Top Quark Production at Hadron Colliders, Ph.D. thesis, Michigan State University, 1995.

[52] ON THE PRODUCTION OF A W AND JETS AT HADRON COLLIDERS, F.A. Berends, H. Kuijf, B. Tausk and W.T. Giele, Nucl. Phys. **B357**:32-64, 1991.

[53] CTEQ Collaboration, http://www.phys.psu.edu/ cteq/

[54] S. Anderson *et al.*, Measurement of the ttbar cross section at sqrt(s)=1.96TeV, DØ Note 4116.

[55] T. Sjostrand *et al.*, Computer Phys. Commun. **82**, 1707 (2001).

[56] G. Marchesini and B. Webber, Nucl. Phys. **B310**, 461 (1998); G. Marchesini *et al.*, Computer Phys. Commun. **67**, 465 (1992).

[57] DØ Top physics group p13 MC webpage: http://www-d0.fnal.gov/Run2Physics/top/d0_private/wg/mc/p13_p13.05.html

[58] The DØ Collaboration, ttbar production cross section in pbarp collisions at sqrt(s)=1.8TeV, hep-ex/0205019.

[59] B. Abott *et al.*, Single top search, Phys. Rev. D **63**, 031101 (2001).

[60] V. Abazov *et al.*, Single Top using NN., Phys. Lett. B **517**, 282 (2001).

ABSTRACT

MEASUREMENT OF TOP PAIR PRODUCTION CROSS SECTION IN ELECTRON + JETS CHANNEL IN PPBAR COLLISIONS AT $\sqrt{s} = 1.96$ TEV

by

Chunhui Han

Chairperson: Jianming Qian

We use a new method, the Matrix Element Method, to measure the $t\bar{t}$ production cross section at Tevatron RunII, in the electron + jets channel. We tested this method with Monte Carlo analysis with two approaches, using the Direct Likelihood Function and the Discriminator Likelihood Function, respectively. Data analysis is carried out using the Direct Likelihood Function, with 106.79 pb^{-1} of data collected by the DØ Experiment. Of the 50 events that pass our selection cuts, the method finds the fraction of signal events to be 0.170. The total cross section is determined to be $7.33 \pm 4.01(\text{stat.})^{+2.13}_{-2.17}(\text{sys.})$ pb.