Instituto de Física de Cantabria
(CSIC-Universidad de Cantabria)
y
Departamento de Física Moderna
(Universidad de Cantabria)

Medida de la Sección Eficaz de Producción del Quark "Single" Top y del Elemento $|V_{tb}|$ de la Matriz CKM en CDF Run II

Memoria de tesis presentada por

Bruno Casal Laraña

para optar al grado de Doctor

Dirigida por

Dr. Alberto Ruiz Jimeno

Santander, Enero de 2010

Instituto de Física de Cantabria
(CSIC-Universidad de Cantabria)
y
Departamento de Física Moderna
(Universidad de Cantabria)

Measurement of the Electroweak Single Top Quark Production Cross Section and the CKM Matrix Element $|V_{tb}|$ at CDF Run II

Memoria de tesis presentada por

Bruno Casal Laraña

para optar al grado de Doctor

Dirigida por

Dr. Alberto Ruiz Jimeno

Santander, Enero de 2010

Dr. Alberto Ruiz Jimeno, Catedrático de Universidad del área de Física Atómica, Molecular y Nuclear de la Facultad de Ciencias de la Universidad de Cantabria

Certifica:

Que la presente memoria: **"Medida de la Sección Eficaz de Producción del Quark "Single" Top y del Elemento $|V_{tb}|$ de la Matriz CKM en CDF Run II"**, ha sido realizada bajo mi dirección en el Departamento de Física Moderna de la Facultad de Ciencias de la Universidad de Cantabria por *Bruno Casal Laraña*, para optar al grado de Doctor en Ciencias Físicas.

Y para que así conste, en cumplimiento de la legislación vigente, firmo el presente certificado:

Santander, Enero de 2010

CONTENTS

List of Figures XIII

List of Tables XIX

1. Introduction 1

2. Theoretical Overview 5

2.1. The Standard Model 5

2.1.1. Fermions 5

2.1.2. Gauge Bosons 6

2.1.3. Electroweak Interactions 7

The Higgs Mechanism 9

CKM Matrix 11

2.1.4. Strong Interactions 13

2.1.5. Cross Section Calculation 14

2.1.6. Parton Distribution Functions 16

2.2. Physics Beyond the Standard Model 17

2.3. Top Quark Physics 18

2.3.1. Top Quark Pair Production 19

Top Quark Discovery 19

2.3.2. Electroweak Single Top Quark Production 22

CKM Matrix Element $|V_{tb}|$ 23

Single Top Quark Polarization 24
Single Top as a Test for New Physics 27

3. Experimental Apparatus 29

3.1. The Tevatron Collider and the Fermilab Accelerator Complex 29
3.1.1. Proton Production and Boosting 30
3.1.2. Main Injector 31
3.1.3. Antiproton Production 32
3.1.4. Recycler Ring 32
3.1.5. Tevatron 33
3.1.6. Luminosity 34
3.1.7. Beam Monitors 34
3.2. The CDF II Detector 35
3.2.1. Standard Definitions in CDF 37
3.2.2. Tracking Systems 39
Silicon Tracking Detectors 40
Central Outer Tracker 43
Pattern Recognition Algorithms 46
3.2.3. Time of Flight 47
3.2.4. The Solenoid 48
3.2.5. Calorimeters 48
Overview 48
Central Calorimeter 49
Plug Calorimeter 52
3.2.6. Muon Systems 53
3.2.7. The Cherenkov Luminosity Counter 56
3.2.8. Trigger 57
Level 1 Trigger 59
Level 2 Trigger 60
Level 3 Trigger 60
Online Monitoring 62

4. Event Simulation and Reconstruction 65

4.1. Event Simulation . . . 65

4.1.1. Monte Carlo Generators . . . 66

PYTHIA . . . 66

MADEVENT . . . 66

ALPGEN . . . 67

4.1.2. Parton Showering and Hadronization . . . 67

4.2. Detector Simulation . . . 68

4.3. Event Reconstruction . . . 68

4.3.1. Primary Vertex Reconstruction . . . 69

4.3.2. Lepton Identification . . . 69

Central Electrons . . . 70

Plug Electrons . . . 72

Muons . . . 73

ID efficiency . . . 76

4.3.3. Jet Reconstruction . . . 76

Jet Energy Corrections . . . 77

4.3.4. Missing Transverse Energy . . . 78

Neutrino Reconstruction . . . 79

4.3.5. B Jet Identification . . . 79

SecVtx Algorithm . . . 80

4.3.6. Neural Net Flavor Separator . . . 84

5. Data Sample and Candidate Event Selection 87

5.1. Trigger Data Samples . . . 88

5.1.1. High-p_T Lepton Triggers . . . 88

Trigger Efficiencies . . . 90

5.1.2. $\not\!\!E_T$+Jets Trigger . . . 91

Trigger Efficiencies . . . 92

5.2. Candidate Even Selection . . . 93

5.2.1. Jet Requirements . . . 94

5.2.2. Lepton Requirements . . . 94
5.2.3. Missing Transverse Energy Requirements . . . 95
5.2.4. Event Vetoes . . . 95
Z Boson Veto . . . 95
Cosmic Ray Veto . . . 95
QCD Veto . . . 96

6. Modeling of Processes **99**
6.1. Signal Event Modeling . . . 99
6.1.1. s-channel Single Top . . . 100
6.1.2. t-channel Single Top . . . 100
6.2. Background Event Modeling . . . 101
6.2.1. Top Pair Production . . . 102
6.2.2. Diboson Production . . . 102
6.2.3. W+Jets . . . 103
Parton-Jet Matching . . . 103
Heavy Flavor Removal . . . 104
Mistags . . . 105
6.2.4. Z+Jets . . . 105
6.2.5. QCD Multijet . . . 106
Jet-Electrons and Anti-Electrons . . . 106
Non-Isolated Muons . . . 108
b-tagging . . . 108

7. Predicted Event Yield **109**
7.1. MC-based Predictions . . . 109
7.2. Data-Driven Predictions . . . 111
7.2.1. QCD-Multijet Estimate . . . 112
7.2.2. W+Jets Estimate . . . 112
Heavy Flavor Fraction Calibration . . . 114
7.3. Candidate Event Yield Prediction . . . 116

7.4. Validation of Monte Carlo Simulation 117

8. Multivariate Classifier **129**

8.1. Discriminating Variables 130

8.2. The Boosted Decision Tree Technique 133

8.2.1. Building a DT 135

8.2.2. Boosting Algorithm 136

8.2.3. Pruning Method 137

8.3. Training of the BDTs 138

8.3.1. Training Samples 138

8.3.2. Input Variables 139

8.4. BDT Output Distributions 140

8.4.1. Validation of the Discriminant 144

9. Analysis **147**

9.1. Systematic Uncertainties 147

9.1.1. Rate Uncertainties 148

9.1.2. Shape Uncertainties 151

9.2. Statistical Interpretation 153

9.2.1. Likelihood Function 155

9.2.2. Cross Section Measurement 158

Extraction of $|V_{tb}|$ 159

9.2.3. Hypothesis Testing and Significance Calculation 159

9.3. Results 162

9.3.1. Cross Section Measurement 163

9.3.2. Significance of the Result 164

10. Observation of Single Top Quark Production **165**

10.1. Lepton + Jets Analyses 166

10.1.1. Boosted Decision Trees 166

10.1.2. Artificial Neural Networks 166

10.1.3. Matrix Element Method 167

10.1.4. Multivariate Likelihood Funtion 168
10.1.5. Super Discriminant Analysis 170
10.2. $\not\!\!E_T$ + Jets Analysis 171
10.3. Combination Results 172
10.3.1. Extraction of the CKM Matrix Element $|V_{tb}|$ 174

11. Conclusions **177**

A. Resumen en Castellano **179**

B. Validation of Input Variables **199**

C. Search for the Higgs Boson in Association with a W. **209**

D. Muon Gap Triggers **213**

E. Measurement of $|V_{td}|$ and $|V_{ts}|$ at CDF **221**

Bibliography **229**

LIST OF FIGURES

2.1. Sketch of the unitarity triangle. 13

2.2. 95% CL constraints on the $(\bar{\rho}, \bar{\eta})$ plane. 14

2.3. Examples of Feynman diagrams for electron-positron scattering. 15

2.4. The CTEQ5L parton distribution functions at $Q^2 = (175\,\mathrm{GeV})^2$. 17

2.5. Comparison of masses of SM quarks 19

2.6. Feynman diagrams of the leading order processes for $t\bar{t}$ production 20

2.7. Reconstructed top mass distributions as published in the CDF evidence and discovery papers 21

2.8. Leading-order Feynman diagrams for the single top quark production modes. . . . 23

2.9. Cosine of the angle between the charged lepton and the non-tagged jet in the top quark rest frame. 26

3.1. Layout of the Fermilab accelerator complex 30

3.2. Total luminosity gathered by the CDF detector as of December 2009 36

3.3. The CDF II Detector with quadrant cut 37

3.4. Elevation view of the Collider Detector at Fermilab (CDF) 38

3.5. Transversal view of the Silicon Vertex Detector at CDF 40

3.6. The CDF II tracker layout showing the different subdetector systems 41

3.7. Detailed view of the Silicon L00 42

3.8. Coverage of the different silicon subdetectors projected into the $r-z$ plane 44

3.9. Layout of wire planes on a COT endplate and wires in a COT supercell 45

3.10. Wedge of the Central Electromagnetic Calorimeter . . . 51
3.11. View of the Plug Calorimeter . . . 52
3.12. Segmentation of the Plug Calorimeter . . . 53
3.13. Coverage (in the $\eta \times \phi$ plane) of the upgraded CDF muon system . . . 54
3.14. Location of the CDF Cherenkov Luminosity Counter . . . 57
3.15. The CLC assembly diagram . . . 58
3.16. Diagram of the CDF II Detector trigger system . . . 59
3.17. Block diagram of the Level 1 and Level 2 trigger paths . . . 61
3.18. Principle of Event Building and Level 3 Filtering . . . 62
3.19. Design of the CDF online consumer framework . . . 63

4.1. Lepton coverage . . . 70
4.2. Diagram of a secondary vertex. . . . 80
4.3. Tagging efficiency of the SecVtx b-tagger. . . . 82
4.4. SecVtx mistag rate as a function of the jet E_T. . . . 83
4.5. Templates of the output of the neural net flavor separator. . . . 85

5.1. Trigger turn-on curve applied to the untriggered muons. . . . 93
5.2. Illustration of the QCD veto for CEM events. . . . 97

6.1. The two different t-channel processes considered in our signal model. . . . 100
6.2. Feynman diagrams of the $t\bar{t}$ background to single top quark production. . . . 102
6.3. Feynman diagrams for diboson production. . . . 103
6.4. Some representative diagrams of W+jets production. . . . 104
6.5. Representative Feynman diagram for a Z+jets production. . . . 106
6.6. Representative Feynman diagram for a QCD multijet event. . . . 107

7.1. Fits to $\not\!\!E_{\rm T}$ distributions in the W+two jet pretag sample. . . . 113
7.2. Fits to $\not\!\!E_{\rm T}$ distributions in the W+2 jet single-tagged sample. . . . 114
7.3. Distributions of the jet flavor separator b_{NN}. . . . 115
7.4. Predicted and observed W+jets events as a function of the jet multiplicity. . . . 117
7.5. Validation plots comparing data and MC for the lepton transverse momentum in W+2 jet events with zero b-tags. . . . 118

7.6. Validation plots comparing data and MC for the lepton transverse momentum in W+2 jet events with at least one b-tag. 119

7.7. Validation plots comparing data and MC for the lepton transverse momentum in W+3 jet events with zero b-tags. 119

7.8. Validation plots comparing data and MC for the lepton transverse momentum in W+3 jet events with at least one b-tag. 120

7.9. Validation plots comparing data and MC for the lepton pseudorapidity in W+2 jet events with zero b-tags. 120

7.10. Validation plots comparing data and MC for the lepton pseudorapidity in W+2 jet events with at least one b-tag. 121

7.11. Validation plots comparing data and MC for the lepton pseudorapidity in W+3 jet events with zero b-tags. 121

7.12. Validation plots comparing data and MC for the lepton pseudorapidity in W+3 jet events with at least one b-tag. 122

7.13. Validation plots comparing data and MC for the leading jet transverse energy in W+2 jet events with zero b-tags. 122

7.14. Validation plots comparing data and MC for the leading jet transverse energy in W+2 jet events with at least one b-tag. 123

7.15. Validation plots comparing data and MC for the leading jet transverse energy in W+3 jet events with zero b-tags. 123

7.16. Validation plots comparing data and MC for the leading jet transverse energy in W+3 jet events with at least one b-tag. 124

7.17. Validation plots comparing data and MC for the leading jet pseudorapidity in W+2 jet events with zero b-tags. 124

7.18. Validation plots comparing data and MC for the leading jet pseudorapidity in W+2 jet events with at least one b-tag. 125

7.19. Validation plots comparing data and MC for the leading jet pseudorapidity in W+3 jet events with zero b-tags. 125

7.20. Validation plots comparing data and MC for the leading jet pseudorapidity in W+3 jet events with at least one b-tag. 126

7.21. Validation plots comparing data and MC for the missing transverse energy in W+2 jet events with zero b-tags. 126

7.22. Validation plots comparing data and MC for the missing transverse energy in W+2 jet events with at least one b-tag. . . . 127
7.23. Validation plots comparing data and MC for the missing transverse energy in W+3 jet events with zero b-tags. . . . 127
7.24. Validation plots comparing data and MC for the missing transverse energy in W+3 jet events with at least one b-tag. . . . 128

8.1. Templates (left) and validation plots comparing data and MC for the discriminating variable $Q \times \eta$. . . . 131
8.2. Templates (left) and validation plots comparing data and MC for the discriminating variable $M_{\ell\nu b}$. . . . 131
8.3. Templates (left) and validation plots comparing data and MC for the discriminating variable H_T. . . . 132
8.4. Templates (left) and validation plots comparing data and MC for the discriminating variable M_{jj}. . . . 132
8.5. Templates (left) and validation plots comparing data and MC for the discriminating variable M_T^W. . . . 133
8.6. Schematic view of a Decision Tree. . . . 134
8.7. Example of an overtrained and a non-overtrained BDT. . . . 137
8.8. Templates of the predictions of the BDT outputs. . . . 143
8.9. Predicted distributions of the BDT outputs. . . . 144
8.10. Distribution of the BDT outputs applied to the untagged control sample with two and three identified jets. . . . 145
8.11. Distribution of the BDT outputs applied to the control sample with four identified jets and at least one b-tag. . . . 146

9.1. The ratio of transverse momentum between data and MC as a function of the invariant mass of the leptons in the event . . . 150
9.2. Example of two shape systematic templates. . . . 151
9.3. Plots showing the mismodeling of the second jet pseudorapidity and the distance between the two jets in the η-ϕ plane. . . . 153
9.4. Distributions of the test statistic using simulated pseudo-experiments for the test H_1 and the null H_0 hypotheses. . . . 161

9.5. Distributions of the output of the BDT applied to the data compared to the predictions from Monte Carlo. 162

9.6. The posterior probability density as a function of the single top quark cross section; and results of the cross section measurement in each signal sample. 163

9.7. Observed test statistic compared to the distributions for the null hypothesis H_0 and the test hypothesis H_1. 164

10.1. Predicted templates, and data to Monte Carlo comparison for the BDT discriminant. 166

10.2. Predicted templates, and data to Monte Carlo comparison for the NN discriminant. 167

10.3. Predicted templates, and data to Monte Carlo comparison for the ME discriminant. 169

10.4. Predicted templates, and data to Monte Carlo comparison for the LF discriminant. . 170

10.5. Predicted templates, and data to Monte Carlo comparison for the SD discriminant. . 171

10.6. Predicted templates, and data to Monte Carlo comparison for the MJ discriminant. . 172

10.7. The posterior probability density for the combination; and results of the cross section measurement in each individual analysis and combination. 173

10.8. Observed test statistic compared to the distributions for the null hypothesis H_0 and the test hypothesis H_1 for the combination. 174

10.9. Posterior probability curve for the $|V_{tb}|$ limit calculation. 175

B.1. Templates and validation plots comparing data and MC for the input variable $E_T^{j_2}$. . 199

B.2. Templates and validation plots comparing data and MC for the input variable η_{j_2}. . 200

B.3. Templates and validation plots comparing data and MC for the input variable η_W. . 200

B.4. Templates and validation plots comparing data and MC for the input variable $M_{\ell\nu j_1 j_2}$.201

B.5. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{\ell\nu}$. 201

B.6. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_1\nu}$.202

B.7. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_2\nu}$.202

B.8. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_1\ell}$. 203

B.9. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_2\ell}$. 203

B.10. Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j_1\ell}$.204

B.11. Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j_2\ell}$.204

B.12. Templates and validation plots comparing data and MC for the input variable $E_T^{j_3}$. . 205

B.13. Templates and validation plots comparing data and MC for the input variable $M_{j_1 j_3}$. 205

B.14. Templates and validation plots comparing data and MC for the input variable $M_{j_2 j_3}$. 205
B.15. Templates and validation plots comparing data and MC for the input variable $M_{j_1 j_2 j_3}$. 206
B.16. Templates and validation plots comparing data and MC for the input variable $M_{\ell\nu j_1 j_2 j_3}$. 206
B.17. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_3\nu}$. 206
B.18. Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_3\ell}$. 207
B.19. Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j_3\ell}$. 207

C.2. The 95% C.L. upper limits on the Higgs boson production cross section relative to the SM expectation 211

D.1. The η–ϕ scatter of CMP-only muons. 214
D.2. CSP efficiencies. The figures correspond to north, top, south and bottom walls from top to bottom; and for east (left) and west (right) sides. 215
D.3. Observed rate reduction for the Phi-Gap test trigger due to the CSP matching requirement and the ratio with to without the matching 217
D.4. The η–ϕ scatter of CMU-only muons. 218
D.5. Efficiencies of the Hadron TDC confirmation 219

E.1. Comparison between the NN-based selection and the cut-based selection, and mass spectrum for the golden hadronic decay mode. 224
E.2. Combined amplitude scan and the logarithm of the ratio of likelihoods for $A = 1$ and $A = 0$. 226

LIST OF TABLES

2.1. Charges and masses of the three generations of quarks and leptons. 6
2.2. Charges and masses of the gauge bosons. 7

3.1. Accelerator parameters for Run I and Run II configurations 35
3.2. Relevant parameters for the layout of the sensors of the SVX-II layers 43
3.3. Calorimeter segmentation 49
3.4. Some parameters for the different calorimeter subdetectors 50
3.5. Parameters of the Muon Detectors at CDF 55

4.1. Selection requirements for central (CEM) electrons. 72
4.2. Selection requirements for plug (PHX) electrons. 73
4.3. Common selection cuts for muon identification. 76
4.4. Lepton identification data/MC scale factors. 77

5.1. Trigger efficiency and total integrated luminosity for each trigger path. 88
5.2. Selection requirements for the CEM electron trigger. 89
5.3. Selection requirements for the PHX electron trigger. 89
5.4. Selection requirements for the CMUP muon trigger. 90
5.5. Selection requirements for the CMX muon trigger. 90
5.6. Selection requirements for the $\not\!\!E_T$+jets trigger. 92
5.7. Fraction of Monte Carlo single top events as a function of the jet multiplicity. . . . 94

7.1. Cross sections used for the yield estimate of the MC-based samples. 110

7.2. Expected signal and background event yield and observed number of events in data. 116

8.1. Variable ranking result for the BDT training in the 2 jet event samples. 141

8.2. Variable ranking result for the BDT training in the 3 jet event samples. 142

9.1. Sources of systematic uncertainty considered in this analysis. 154

10.1. A summary of the results of each individual analysis, with their measured cross-sections, observed significance, and sensitivity. 173

Chapter 1

INTRODUCTION

The top quark is the most massive known elementary particle; its mass is $173.1 \pm 1.3\ \mathrm{GeV}/c^2$ [1], about forty times larger than that of the bottom quark, the second-most massive standard model (SM) fermion. The top quark's large mass, at the scale of electroweak symmetry breaking, hints that it may play a role in the mechanism of mass generation [2–4]. Up to now, the Tevatron is the only place to produce and study top quarks. The cross section for $t\bar{t}$ production via the strong interaction at the Tevatron is approximately 7 pb [5], and the decay signature of two top quarks decaying to two pairs of a W boson and a b quark is very distinct from the background processes. The presence of the top quark was established in 1995 by the CDF and DØ collaborations with approximately 100 pb^{-1} of $p\bar{p}$ data collected per collaboration at $\sqrt{s} = 1.8$ TeV [6, 7].

Since then, larger data samples have enabled detailed study of top quarks. The $t\bar{t}$ production cross section [8], the top quark's mass [1], the top quark decay branching fraction to Wb [9], and the polarization of W bosons in top decay [10] have been measured precisely. Nonetheless, many properties of the top quark have not yet been tested as precisely. In particular, the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{tb} remains poorly constrained by direct measurements [11]. The coupling V_{tb} governs the decay rate of the top quark and its partial width into Wb; other decays are expected to have much smaller branching fractions. In the three-generation SM, this matrix element is expected to be very close to unity, using the measurements of the other CKM matrix elements and the assumption that the 3×3 CKM matrix is unitary.

Top quarks are expected to be produced singly in $p\bar{p}$ collisions via weak, charged-current interactions. The dominant processes at the Tevatron are the s- and t-channel exchange of a virtual W boson. The next-to-leading-order (NLO) cross sections for these two processes are

$\sigma_s = 0.88 \pm 0.11$ pb and $\sigma_t = 1.98 \pm 0.25$ pb, respectively [12, 13]. A third process, the associated production of a W boson and a top quark, has a comparatively negligible small expected cross section at the Tevatron.

The reasons for studying single top quarks are compelling. The production cross section is directly proportional to the square of the CKM matrix element $|V_{tb}|$, allowing an overconstrained test of the unitarity of the CKM matrix. A broad class of extensions to the SM predict measurable deviations of σ_s or σ_t from their SM values, and thus a measurement of the single top quark production rates could give constraints to fourth-generation models, flavor-changing neutral currents models, or other new phenomena [14]. On the other hand, measuring single top quark production also implies establishing all SM background processes to the search for the Higgs boson production in the $WH \to \ell\nu b\bar{b}$ channel, since both processes share the same final state. Furthermore, all methods developed in the search for single top quark production can be used for the search of the Higgs boson.

Measuring the single top quark cross section is well motivated but it is also extremely challenging at the Tevatron. The total production cross section is expected to be smaller than half of that of $t\bar{t}$, and with only one top quark in the final state, the signal is far less distinct from the dominant background processes. The rate at which a W boson is produced along with jets, at least one of which must be b-tagged, is approximately 12 times the signal rate. The *a priori* uncertainties on the background processes are about a factor of three larger than the expected signal rate. In order to expect to observe single top quark production, the background rates must be small and well constrained, and the expected signal must be much larger than the uncertainty on the background. A much more pure sample of signal events therefore must be separated from the background processes in order to make observation possible.

Single top quark production is characterized by a number of kinematic features. The top quark mass is known, and precise predictions of the distributions of the top quarks and the recoil products are also available. Top quarks produced singly via the weak interaction are also expected to be nearly 100% polarized [15, 16]. The background W+jets and $t\bar{t}$ processes also have distinct kinematic features which differ from those of single top quark production. These kinematic features, coupled with the b-tagging requirement, provide the keys to further purification of the signal. This thesis describe a search of the combined s- and t-channel single top quark production, employing a multivariate technique based on Boosted Decision Trees, in order to achieve maximum signal-to-background separation.

In the Tevatron Run I, limits on the single top quark production cross section were set by both the CDF [17] and the DØ [18] collaborations, either by counting experiments or by performing a maximum likelihood fit to the distribution of a kinematic variable. These limits were updated by both collaborations using multivariate methods based on neural networks [19, 20]. Stronger

limits were set in Tevatron Run II with early datasets of 162 pb^{-1} by CDF [21] and 230 pb^{-1} by DØ [22, 23]. Evidence for the production of single top quarks was first reported by the DØ collaboration using 0.9 fb^{-1} of data [24, 25], and then confirmed by the CDF collaboration using 2.2 fb^{-1} of data [26]. Electroweak single top quark production is finally observed simultaneously by the CDF collaboration using 3.2 fb^{-1} of data [27], and by the DØ collaboration using 2.3 fb^{-1} of data [28].

Chapter 2

THEORETICAL OVERVIEW

2.1. The Standard Model

The Standard Model (SM) of elementary particles is a theory that describes the fundamental constituents of the universe and their interactions. The SM provides a very elegant theoretical framework based on a quantum field theory that explains the nuclear and electromagnetic interactions between particles as resulting from the introduction of local symmetries. It has been very successful in predicting a variety of properties of particles and their interactions, describing to an unprecedented level of precision many experimental results [29]. A complete description of the theory can be easily found in the scientific literature [30–33].

Based on several group symmetries, the SM includes the electromagnetic, weak and strong interaction. The building blocks of Nature, according to the SM, are divided by their spin into two closed sets of particles: fermions and gauge bosons. Fermions are the constituents of matter, half integer spin particles which follow Fermi statistics. Bosons are the carriers of the forces, which have integer spin and follow Bose statistics.

2.1.1. Fermions

The SM postulates that all known matter is composed of a few basic, point-like and structureless constituents: the fermions. One distinguishes two groups: quarks and leptons. The quarks come in six different flavors: up, down, charm, strange, top and bottom; formally described by assigning flavor quantum numbers. The SM incorporates six leptons: the electron (e^-) and the electron-neutrino (ν_e), the muon (μ^-) and the muon-neutrino (ν_μ), the tau (τ^-) and the tau-neutrino

(ν_τ). They carry electron, muon and tau quantum numbers. Quarks and leptons can be grouped into three generations (or families) as shown in Table 2.1 which also contains the charges and masses of the particles.

Generation	First		Second		Third	
Quarks (spin = 1/2)						
Particle	down	up	strange	charm	bottom	top
Symbol	d	u	s	c	b	t
Charge (e)	-1/3	+2/3	-1/3	+2/3	-1/3	+2/3
Mass ($\mathrm{MeV/c^2}$)	3.5-6	1.5-3.3	104^{+26}_{-34}	$(1.27^{+0.07}_{-0.11}) \cdot 10^3$	$(4.20^{+0.17}_{-0.07}) \cdot 10^3$	$(171.2 \pm 2.1) \cdot 10^3$
Interaction	EM, Weak, Strong					
Leptons (spin = 1/2)						
Particle	e-neutrino	electron	μ-neutrino	muon	τ-neutrino	tau
Symbol	ν_e	e	ν_μ	μ	ν_τ	τ^-
Charge (e)	0	-1	0	-1	0	-1
Mass ($\mathrm{MeV/c^2}$)	$< 2 \cdot 10^{-6}$	0.511	<0.19	105.658	<18.2	1776.84 ± 0.17
Interaction	Weak	EM, Weak	Weak	EM, Weak	Weak	EM, Weak

Table 2.1: Charges and masses of the three generations of quarks and leptons [11].

The three generations exhibit a striking mass hierarchy, the top quark having by far the highest mass. Understanding the deeper reason behind the hierarchy and generation structure is one of the open questions of particle physics. Each quark, q, and each lepton, ℓ^-, has an associated antiparticle with the same mass but opposite charge, $\bar{q}$ and ℓ^+ respectively.

2.1.2. Gauge Bosons

The forces of nature acting between quarks and leptons are described by quantized fields. The interactions between elementary particles are due to the exchange of field quanta which are said to mediate the forces. The SM incorporates the electromagnetic force, responsible for the emission of light from excited atoms, the weak force, which for instance causes nuclear beta decay, and the strong force which keeps nuclei stable. Gravitation is not included in the framework of the SM but rather described by the theory of general relativity. All particles with mass or energy feel the gravitational force. However, due to the weakness of gravitation with respect to the other forces, it does not play an important role in elementary particle reactions.

The electromagnetic, weak and strong forces are described by so called quantum gauge field theories. The quanta of these fields carry spin 1 and are therefore called gauge bosons. Table 2.2 shows the charges and masses of the gauge bosons. The electromagnetic force is mediated by the massless photon (γ), the weak force by the massive $W^\pm$ and Z^0 bosons, and the strong force by

Gauge Bosons (spin = 1)							
Symbol	Charge	Mass (GeV/c^2)	Force	Coupling	Range	Typical lifetime (m)	Typical cross section (mb)
photon, γ	0	0	EM	$\alpha \approx 10^{-2}$	∞	$10^{-20} \sim 10^{-16}$	10^{-3}
gluon, g	0	0	Strong	$\alpha_s \approx 1$	10^{-15}	10^{-23}	10
$W^{\pm}$	± 1	80.398 ± 0.025	Weak	$G_F \approx 10^{-5}$	10^{-18}	10^{-8}	10^{-11}
Z^0	0	91.1876 ± 0.0021					

Table 2.2: Charges and masses of the gauge bosons [11].

eight massless gluons (g).

Quarks participate in electromagnetic, weak and strong interactions. All leptons experience the weak force, the charged ones also feel the electromagnetic force. But leptons do not take part in strong interactions.

2.1.3. Electroweak Interactions

In quantum field theory quarks and leptons are represented by spinor fields Ψ which are functions of the continuous space-time coordinates x_μ. To take into account that the weak interaction only couples to the left-handed particles, left- and right-handed fields $\Psi_{\mathrm{L}} = \frac{1}{2}(1 - \gamma^5)\,\Psi$ and $\Psi_{\mathrm{R}} = \frac{1}{2}(1 + \gamma^5)\,\Psi$ are introduced. Here, $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the chirality operator, where γ^μ are the γ-matrices, and $(1 \pm \gamma^5)$ are helicity projectors. The left-handed states of one generation are grouped into weak-isospin doublets, the right-handed states form singlets:

$$Q^i_{\mathrm{L}} = \begin{pmatrix} u^i \\ d^i \end{pmatrix}_{\mathrm{L}} = \left(\begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} c \\ s \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} t \\ b \end{pmatrix}_{\mathrm{L}} \right), \tag{2.1}$$

$$L^i_{\mathrm{L}} = \begin{pmatrix} \nu^i_\ell \\ \ell^i \end{pmatrix}_{\mathrm{L}} = \left(\begin{pmatrix} \nu_e \\ e \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_{\mathrm{L}} \right), \tag{2.2}$$

$$u^i_{\mathrm{R}} = (u_{\mathrm{R}},\, c_{\mathrm{R}},\, t_{\mathrm{R}})\,, \qquad d^i_{\mathrm{R}} = (d_{\mathrm{R}},\, s_{\mathrm{R}},\, b_{\mathrm{R}})\,, \qquad l^i_{\mathrm{R}} = (e_{\mathrm{R}},\, \mu_{\mathrm{R}},\, \tau_{\mathrm{R}})\,, \tag{2.3}$$

In the original SM the right-handed neutrino states are omitted, since neutrinos are assumed to be massless. Recent experimental evidence [34–38], however, strongly indicates that neutrinos have mass and the SM needs to be extended in this respect.

The dynamics of the electromagnetic and weak forces follow from the free particle Lagrangian density

$$\mathcal{L}_0 = i\,\overline{\Psi}\,\gamma^\mu \partial_\mu\,\Psi, \tag{2.4}$$

by demanding the invariance of $\mathcal{L}_0$ under local phase transformations:

$$\Psi_{\rm L} \longrightarrow \Psi'_{\rm L} = e^{ig\boldsymbol{\alpha}(x)\cdot\boldsymbol{T}+ig'\beta(x)Y}\ \Psi_{\rm L} \quad \text{and} \quad \Psi_{\rm R} \longrightarrow \Psi'_{\rm R} = e^{ig'\beta(x)Y}\ \Psi_{\rm R}\ . \tag{2.5}$$

For historical reasons these transformations are also referred to as gauge transformations. In Eq. 2.5, g and g' are real coupling constants, $\boldsymbol{\alpha}(x)$ is an arbitrary three-component vector, and $\boldsymbol{T} = (T_1,\, T_2,\, T_3)^t$, where t denotes the tronspose, is the weak-isospin operator whose components T_i are the generators of $SU(2)_L$ symmetry transformations. The index L indicates that the phase transformations act only on left-handed states. The matrix representations are given by $T_i = \frac{1}{2}\,\tau_i$ where the τ_i are the Pauli matrices. The T_i do not commute: $[T_i, T_j] = i\,\epsilon_{ijk}\,T_k$. That is why the $SU(2)_L$ gauge group is said to be non-Abelian. $\beta(x)$ is a one-dimensional function of x. Y is the weak hypercharge which satisfies the relation $Q = T_3 + Y/2$, where Q is the electromagnetic charge. Y is the generator of the symmetry group $U(1)_Y$. The weak-isospin assignment for the doublet is: up-type quarks u^i and neutrinos carry $T_3 = +\frac{1}{2}$; down-type quarks d^i, electron, muon and tau lepton have $T_3 = -\frac{1}{2}$. Since the right-handed states are isoscalars, they carry $T_3 = 0$. The weak hypercharge associated to left-handed leptons and neutrinos is $Y = -\frac{1}{2}$, and $Y = -1$ for right-haded leptons. The assigment for quarks is $Y = \frac{1}{6},\ \frac{2}{3},\ \frac{-1}{3}$, respectively for left-handed quarks, right-handed up-type quarks, and right-handed down-type quarks.

Demanding the Lagrangian $\mathcal{L}_0$ to be invariant under the combined gauge transformations of $SU(2)_L \times U(1)_Y$, see Eq. 2.5, requires the addition of terms to the free Lagrangian which involve four additional vector (spin 1) fields: the isotriplet $\boldsymbol{W}^\mu = (W_1^\mu\,, W_2^\mu\,, W_3^\mu)^t$ for $SU(2)_L$ and the singlet B^μ for $U(1)_Y$, which transform as:

$$\boldsymbol{W}^\mu \longrightarrow \boldsymbol{W'}^\mu = \boldsymbol{W}^\mu - \partial^\mu\boldsymbol{\alpha}(x) - g\boldsymbol{\alpha}(x) \times \boldsymbol{W}^\mu, \tag{2.6}$$

$$B^\mu \longrightarrow B'^\mu = B^\mu - \partial^\mu\beta(x). \tag{2.7}$$

This is technically done by replacing the derivative ∂_μ in $\mathcal{L}_0$ by the covariant derivative

$$\partial^\mu \longrightarrow D^\mu = \partial^\mu + i\,g\,\boldsymbol{W}^\mu\cdot\boldsymbol{T} + i\,g'\frac{1}{2}B^\mu Y, \tag{2.8}$$

and adding the kinetic energy terms of the gauge fields: $-\frac{1}{4}\boldsymbol{W}_{\mu\nu}\cdot\boldsymbol{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}\cdot B^{\mu\nu}$. The field tensors $\boldsymbol{W}_{\mu\nu}$ and $B_{\mu\nu}$ are given by $\boldsymbol{W}_{\mu\nu} = \partial_\mu\boldsymbol{W}_\nu - \partial_\nu\boldsymbol{W}_\mu - g\cdot\boldsymbol{W}_\mu\times\boldsymbol{W}_\nu$ and $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$. Since the vector fields $\boldsymbol{W}^\mu$ and B^μ are introduced via gauge transformations they are called gauge fields and the quanta of these fields are named gauge bosons. For an electron-neutrino pair, for example, the resulting Lagrangian is:

$$\begin{aligned}\mathcal{L}_1 \;=\; & i\,\overline{\begin{pmatrix}\nu_e\\ e\end{pmatrix}}_L\,\gamma^\mu\left[\partial_\mu + i\,g\,\boldsymbol{W}_\mu\cdot\boldsymbol{T} + i\,g'\,Y_L\,\frac{1}{2}\,B_\mu\right]\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L + \\ & i\,\bar{e}_R\,\gamma^\mu\left[\partial_\mu - g'\,Y_R\,\frac{1}{2}\,B_\mu\right]e_R - \frac{1}{4}\,\boldsymbol{W}_{\mu\nu}\cdot\boldsymbol{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}\cdot B^{\mu\nu}.\end{aligned} \tag{2.9}$$

This model developed by Glashow [39], Weinberg and Salam [40, 41] in the 1960s allows to describe electromagnetic and weak interactions in one framework. One therefore refers to it as unified electroweak theory.

The Higgs Mechanism

One has to note, however, that $\mathcal{L}_1$ describes only massless gauge bosons and massless fermions. Mass-terms such as $\frac{1}{2}M^2 B_\mu B^\mu$ or $-m\overline{\Psi}\Psi$ are not gauge invariant and therefore cannot be added. To include massive particles into the model in a gauge invariant way the Higgs mechanism is used [42–48]. Four scalar fields are added to the theory in form of the isospin doublet $\mathbf{\Phi} = (\phi^+,\ \phi^0)^t$ where ϕ^+ and ϕ^0 are complex fields. This is the minimal choice. The term $\mathcal{L}_H = (D_\mu \mathbf{\Phi})^\dagger\,(D^\mu \mathbf{\Phi}) - V(\Phi^\dagger\Phi)$ is added to $\mathcal{L}_1$, where the scalar potential takes the form $V(\Phi^\dagger\Phi) = \mu^2\ \mathbf{\Phi}^\dagger\mathbf{\Phi} + \lambda\ (\Phi^\dagger\Phi)^2$.

In most cases particle reactions cannot be calculated from first principles. One rather has to use perturbation theory and expand a solution starting from the ground state of the system which is in particle physics called the vacuum expectation value. The parameters μ and λ can be chosen ($\lambda > 0$ and $\mu^2 < 0$) such that the vacuum state of the Higgs potential V is degenerate with an expectation value different from zero:

$$|\Phi_{\text{vac}}| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}} \qquad (v > 0). \tag{2.10}$$

Without loss of generality, one can choose the ground state $\mathbf{\Phi}_0 = \left(0,\ v/\sqrt{2}\right)^t$. This state is not invariant under $SU(2)_L \times U(1)_Y$ gauge transformations, however, it is invariant under $U(1)$ electromagnetic gauge transformations, so that the photon remains massless and the electric charge is conserved. This mechanism where the ground states do not share the symmetry of the Lagrangian is called spontaneous symmetry breaking.

Applying spontaneous symmetry breaking as described above to the combined Lagrangian $\mathcal{L}_2 = \mathcal{L}_1 + \mathcal{L}_H$ and enforcing local gauge invariance of $\mathcal{L}_2$, makes the three electroweak gauge bosons acquire mass. The massive bosons are, however, not the original fields $\boldsymbol{W}_\mu$ and B_μ but rather mixtures of those: the $W^\pm_\mu = (W^1_\mu \mp i\ W^2_\mu)/\sqrt{2}$, the Z^0 and the photon field A_μ:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix}, \tag{2.11}$$

where the mixing angle θ_W is the Weinberg angle defined by the coupling constants $g'/g = \tan\theta_W$.

Having started from the original four massless vector fields ($\boldsymbol{W}_\mu$ and B_μ, i.e. eight degrees of

freedom[1]) and a doublet of complex scalar fields ($\mathbf{\Phi}$, i.e. four degrees of freedom), we have ended up with three massive vector fields (the $W^{\pm}$ and Z^0 bosons, i.e. nine degrees of freedom) and the massless photon field (two degrees of freedom). The remaining degree of freedom corresponds to the scalar real field associated with the observable massive spin 0 particle called the Higgs boson.

The boson masses are related to the basic parameters as follows:

$$m_W = \frac{1}{2} v g \ , \qquad m_Z = m_W / \cos\theta_W \ , \qquad m_H = \sqrt{-2\mu^2} \ , \tag{2.12}$$

Rather remarkably, the masses of the $W^{\pm}$ and Z^0 bosons can be determined in terms of three experimentally well known quantities. Neglecting radiative corrections:

$$m_W = \left(\frac{\alpha\pi}{G_F\sqrt{2}} \right)^{\frac{1}{2}} \frac{1}{\sin\theta_W} \approx 77\,\mathrm{GeV} \ , \qquad m_Z = \left(\frac{\alpha\pi}{G_F\sqrt{2}} \right)^{\frac{1}{2}} \frac{2}{\sin 2\theta_W} \approx 88\,\mathrm{GeV} \tag{2.13}$$

where α is the fine structure constant and G_F is the Fermi coupling constant. Historically, the above was used to predict the masses of the $W^{\pm}$ and Z^0 bosons until they were discovered in 1983 by UA1 and UA2 collaborations at the CERN SPS[49, 50]. Nowadays, the high precision of the experimental measurements of the $W^{\pm}$ and Z^0 boson masses are instead used as inputs to calculate the weak mixing angle θ_W.

The only missing parameter to determine is the Higgs boson mass. The Higgs particle has not yet been observed because of its small coupling, leaving it as the last missing piece of the electroweak theory to be experimentally tested.

Spontaneous symmetry breaking also generates lepton masses if Yukawa interaction terms of the lepton and Higgs fields are added to the Lagrangian:

$$\mathcal{L}^{\mathrm{lepton}}_{\mathrm{Yukawa}} = -G_e \left[\bar{e}_R \mathbf{\Phi}^{\dagger} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L + \overline{\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L} \mathbf{\Phi} \ e_R \right] . \tag{2.14}$$

Here the Yukawa terms for the electron-neutrino doublet are given as an example. G_e is a further coupling constant describing the coupling of the electron and electron-neutrino to the Higgs field. In this formalism neutrinos are assumed to be massless.

Quark masses are also generated by adding Yukawa terms to the Lagrangian. However, for the quarks, both the upper and the lower member of the weak-isospin doublet need to acquire mass. For this to happen an additional conjugate Higgs multiplet has to be constructed: $\mathbf{\Phi}_c = i\tau_2 \mathbf{\Phi}^* = (\phi^{0^*}, -\phi^-)^t$. The Yukawa terms for the quarks are given by:

$$\mathcal{L}^{\mathrm{quark}}_{\mathrm{Yukawa}} = \sum_{i=1}^{3} \sum_{j=1}^{3} \ G^d_{ij} \ \overline{Q}^{\,i}_L \mathbf{\Phi} \ d^j_R + G^u_{ij} \ Q^i_L \mathbf{\Phi}_c \ u^j_R \ + \ \mathrm{hermitian\ conjugate}, \tag{2.15}$$

where $G^{d,u}_{ij}$ are quark coupling constants.

[1] For massless vector fields there are only two independent polarization states, the third is eliminated by gauge invariance.

CKM Matrix

Couplings between quarks of different generations are allowed by this ansatz. After spontaneous symmetry breaking the Yukawa terms produce mass terms for the quarks which can be described by mass matrices in generation space: $\bar{u}_R^{\,i}\,\mathcal{M}_{ij}^u\,u_L^j$ and $\bar{d}_R^{\,i}\,\mathcal{M}_{ij}^d\,d_L^j$ with $\mathcal{M}_{ij}^u = |\Phi_{\rm vac}| \cdot G_{ij}^u$ and $\mathcal{M}_{ij}^d = |\Phi_{\rm vac}| \cdot G_{ij}^d$. The mass matrices are non-diagonal but can be diagonalized by unitary transformations

$$u_L'^i = U_u^{ij} u_L^i\,, \qquad d_L'^i = U_d^{ij} d_L^i\,, \tag{2.16}$$

which essentially means to change from the original basis to the basis that diagonalizes their Higgs couplings. This latter basis is the physical one, since it gives the mass eigenstates. In this basis the quark coupling to the $W^\pm$ boson field takes the form

$$\mathcal{L}_{QW^\pm} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L^i \gamma^\mu (U_u^\dagger U_d)^{ij} d_L^j + \text{hermitian conjugate}, \tag{2.17}$$

This means that in charged-current interactions ($W^\pm$ exchange), transitions between mass eigenstates of different quark flavors are possible, this is referred to as generation mixing. The charged weak interaction link the three u_L^i quarks with a unitary rotation of the triplet of d_L^i quarks, with this rotation given by the unitary matrix

$$V_{\rm CKM} = U_u^\dagger U_d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.18}$$

The matrix $V_{\rm CKM}$ is known as the Cabbibo-Kobayashi-Maskawa (CKM) mixing matrix [51, 52]. The CKM element $|V_{q_1 q_2}|$ is proportional to the coupling strength of two quarks q_1 and q_2 to a W boson.

The matrix $V_{\rm CKM}$ is a general 3×3 unitary matrix. Such a matrix requires 9 real parameters to be specified. Of these, 3 are rotating angles, and the remaining 6 parameters are phases. We can remove these phases by making phase rotations of quark fields, but the overall phase is redundant, so we can remove only 5 of these phases. The final form of $V_{\rm CKM}$ contains three mixing angles, and a phase which is responsible of all CP-violating phenomena in flavor changing processes in the SM. The standard parameterization of the $V_{\rm CKM}$ is given by

$$V_{\rm CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.19}$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$, θ_{ij} are the mixing angles, and δ is the CP-violating phase.

It is known experimentally that $s_{13} \ll s_{23} \ll s_{12} \ll 1$, and it is convenient to exhibit this hierarchy using the Wolfenstein parameterization [53–55], with parameters $\lambda,\ A,\ \rho,\ \eta$ defined as:

$$\lambda = s_{12}\,, \qquad A\lambda^2 = s_{23}\,, \qquad A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta}\,. \tag{2.20}$$

These parameterization choice ensures that the CKM matrix written in terms of $\lambda,\ A,\ \rho,\ \eta$ is unitary to all orders in λ. To order λ^3:

$$V_{\rm CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.21}$$

Unitarity of the CKM matrix implies

$$\Sigma_i V_{ij}V_{ik}^* = \delta_{ij} \quad \text{and} \quad \Sigma_j V_{ij}V_{kj}^* = \delta_{ik}. \tag{2.22}$$

The non-vanishing relation are written as $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$, and similarly for all other rows and columns. This implies that the sum of all couplings of any of the up-type quarks to all the down-type quarks (and vice-verse) is the same for all generations. This relation is called "weak universality".

Even more interesting constraints are given by the six vanishing combination in Equation 2.22. These expressions can be represented as triangles in a complex plane. The most commonly used unitarity triangle arises from

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^*, \tag{2.23}$$

by dividing each side by $V_{cd}V_{cb}^*$ (see Figure 2.1). The vertices are exactly (0, 0), (1, 0) and, due to the definition in Equation 2.20, $(\bar{\rho},\ \bar{\eta})$. The angles are given by

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \qquad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \qquad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.24}$$

An important goal of flavor physics is to overconstrain the CKM elements. The unitarity triangle(s) are useful because they provide a simple, vivid summary of the CKM mechanism. Separate measurements of lengths, through decay and mixing rates, and angles, through CP asymmetries, should fit together. Furthermore, when one combines measurements, from the B, B_s, K, and D systems, as well as from hadronic W decays, all triangles should have the same area and orientation. If there are non-CKM contributions to flavor or CP violation, however, the interpretation of rates and asymmetries as measurements of the sides and angles no longer holds; the triangle built from experimentally defined sides and angles will not fit with the CKM picture. As for today, all direct and indirect experimental measurements of the CKM elements are consistent with the SM expectation.

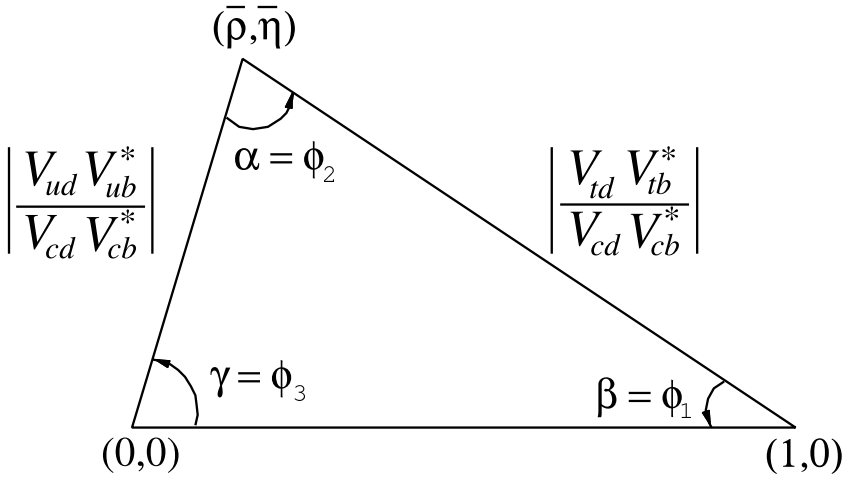


Figure 2.1: Sketch of the unitarity triangle [11].

The CKM matrix elements can be precisely determined by a global fit that uses all available measurements and imposes the SM constraints. There are several approaches to combining the experimental data [55–58], which provide similar results. The results for the Wolfenstein parameters are [11]:

$$\lambda = 0.2257^{+0.0009}_{-0.0010}, \quad A = 0.814^{+0.021}_{-0.022}, \quad \rho = 0.135^{+0.031}_{-0.016}, \quad \eta = 0.349^{+0.015}_{-0.017}. \tag{2.25}$$

The allowed ranges of the magnitudes of all nine CKM elements are [11]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{pmatrix}. \tag{2.26}$$

Figure 2.2 illustrates the constraints on the $(\bar{\rho},\, \bar{\eta})$ plane from various measurements and the global fit result. The shaded 95% CL regions all overlap consistently around the global fit region.

2.1.4. Strong Interactions

The theory of strong interactions is called quantum chromodynamics (QCD) since it attributes a color charge to the quarks. There are three different types of strong charges (colors): “red”, “green” and “blue”. Strong interactions conserve the flavor of quarks, thus, there cannot be quark flavor transitions via the strong interaction in contrast to the weak interaction where these transitions can exist via a $W^{\pm}$ boson exchange. Leptons do not carry color at all, they are inert with respect to strong interactions. QCD is a quantum field theory based on the non-Abelian gauge group $SU(3)_C$ of phase transformations on the quark color fields. Invoking local gauge invariance of the Lagrangian yields eight massless gauge bosons: the gluons. The gauge symmetry is exact and not

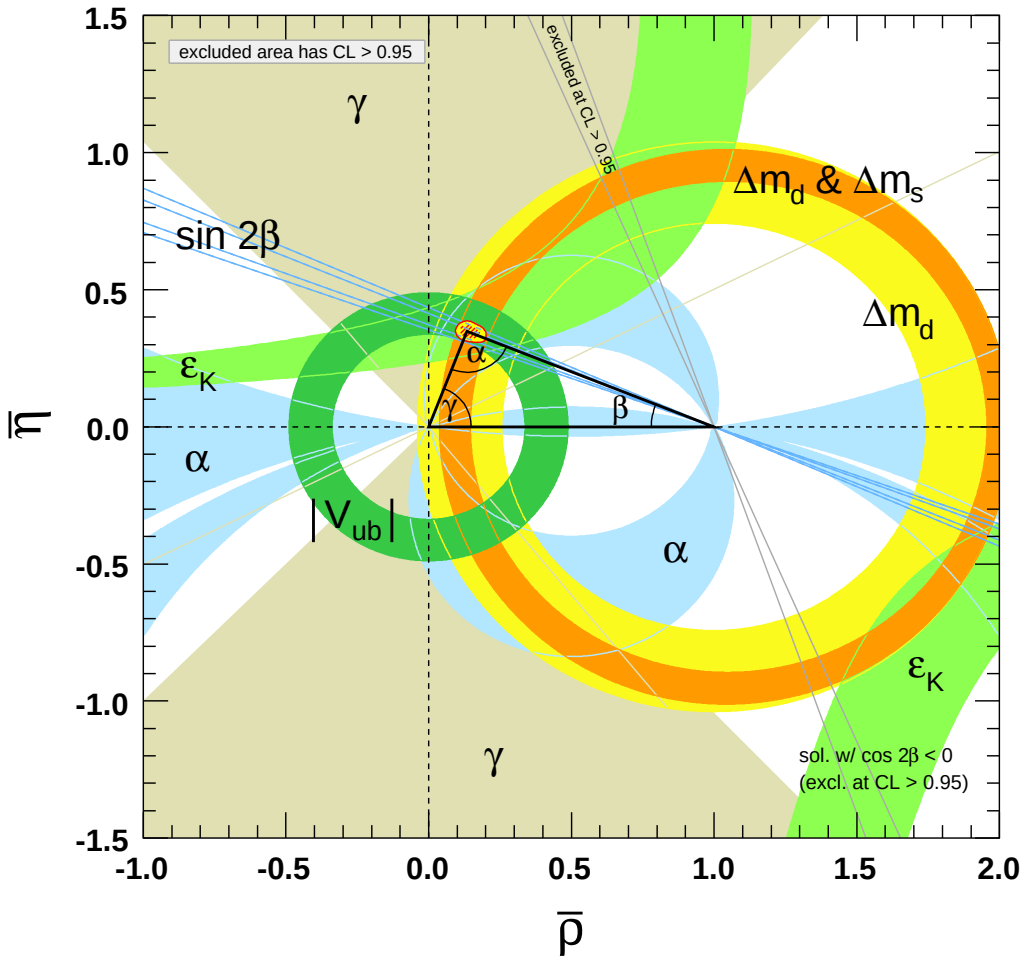


Figure 2.2: 95% CL constraints on the $(\bar{\rho},\ \bar{\eta})$ plane [11].

broken as in the case of weak interactions. Each gluon carries one unit of color and one unit of anticolor. The strong force binds quarks together to form bound-states called hadrons. There are two groups of hadrons: mesons consisting of a quark and an antiquark and baryons built of either three quarks or three antiquarks. All hadrons are color-singlet states. Quarks cannot exist as free particles. This experimental fact is summarized in the notion of quark confinement: quarks are confined to exist in hadrons.

2.1.5. Cross Section Calculation

The probability of a given interaction occurring is related to the cross section of the interaction, which is measured in units of barns (b), where one barn is $10^{-24}\,\mathrm{cm}^2$. Current theory is unable to calculate these cross sections exactly; instead, a perturbative expansion must be made in powers of a coupling constant. These perturbative terms can be conveniently represented by Feynman diagrams, which are graphical representations of each term in the expansion. Feynman diagrams consist of lines, representing fields, and vertices, representing the interactions of the fields (see

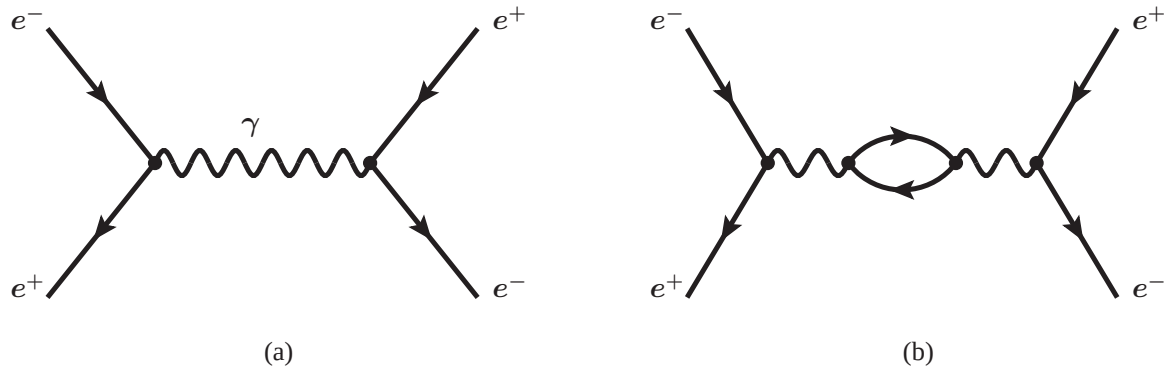


Figure 2.3: Examples of Feynman diagrams for electron-positron scattering (Bhabha scattering): (a) a tree-level diagram for the pair annihilation process, and (b) a loop diagram representing the photon self energy correction to the leading order diagram (a). In the following, Feynman diagrams are always drawn with time on the horizontal axis, increasing from left to right.

Figure 2.3 for an example). These combine according to a simple set of rules, and the lines and vertices of a Feynman diagram can be converted directly into a calculation of the term in the cross section associated with that diagram. The sum of all terms gives the amplitude $\mathcal{M}$ of the process.

For a given set of initial- and final-state particles, the Feynman diagrams with the fewest possible number of vertices represent the leading-order term of the perturbative expansion, and often constitute a good approximation of the underlying physics. One pleasant feature of most leading-order diagrams is that they contain no closed loops; these are referred to as "tree-level" diagrams. Next-to-leading-order diagrams have at least one more vertex and represent the next term in the expansion. Calculating these is much more difficult because of the properties of loop diagrams. Whenever the topological feature of a loop appears, the calculations require an integral to be performed over the momenta of the particles in the loop, and the integral often diverges. This is not, of course, a problem with reality, but an artifact of perturbation theory: the next-to-leading-order term of a perturbative expansion is a theoretical construct and cannot be measured. These divergences, usually called "ultraviolet divergences" because they occur for very large momentum scales, can be dealt with by a process called "renormalization", in which a renormalization scale is introduced to truncate the integral before it diverges. This gives a finite result to the calculation that agrees well with experiment for many interactions.

Another kind of divergence is called an "infrared divergence", because it arises for small momentum scales. The source of these divergences is the perturbative nature of quantum field theory calculations. In calculations involving the strong force, at low energies the coupling constant becomes larger than one. In this case, each successive term in the perturbation is larger than the one before it, and perturbation will no longer give a valid answer. Such divergences are dealt

with by introduction of a factorization scale which truncates the integral before it reaches the non-perturbative region. Fortunately, the energies of particles in this analysis are well above the cutoff for valid perturbation; unfortunately, the choice of renormalization and factorization scale affects the cross section calculation, so it sometimes requires the addition of a systematic uncertainty.

Because the cross section of a particle with given exact values of momentum and energy is infinitesimally small, it is more sensible to calculate the differential cross section of the interaction. This quantity is defined for an infinitesimal slice of the momentum space of all final state particles. For a cross section calculation, it is given by Fermi's Golden Rule. For scattering of two particles with four-momenta q_1 and q_2 into n particles with four-momenta p_i, the differential cross section is given by

$$\mathrm{d}\sigma = |\mathcal{M}|^2 \frac{\hbar^2 S}{4\sqrt{(q_1 q_2)^2 - (m_1 m_2 c^2)^2}} \mathrm{d}\Phi_n, \tag{2.27}$$

where S is a combinatorial factor for identical particles, m_i is the particle mass, $\mathcal{M}$ is the matrix element for the interaction, and $\mathrm{d}\Phi_n$ is the phase space factor given by

$$\mathrm{d}\Phi_n = \delta\left(q_1 + q_2 - \sum_{i=1}^{n} p_i\right) \prod_{i=1}^{n} \frac{c\,\mathrm{d}^3 p_i}{(2\pi)^3\, 2E_i}. \tag{2.28}$$

Integrating this expression for all final-state momenta gives the total cross section of the interaction.

2.1.6. Parton Distribution Functions

The calculation of a measurable cross section at a hadron collider has to deal with an extra complication, and it is that the colliding particles are composed particles (protons and antiprotons at the Tevatron). Therefore, it has to be taken into account that the momentum of the proton (or antiproton) is shared among all the elementary constituent particles. The proton consists of two u and one d quark, the antiproton of the respective antiparticles. Those constituents, called valence quarks, are bound by virtual gluons which can split into quark-antiquark pairs, the so-called sea quarks. This leads to the situation, that the momentum of the proton is shared by all three valence quarks, sea-quarks and gluons. The fraction of the momentum x_i, carried by each quark or gluon (generically called partons), is described by the parton distribution function (PDF) $f_{i,p}\left(x_i, Q^2\right)$. It depends on the scale Q^2, describing the typical energy scale of the considered interaction, which for top quark production is usually set to the order of the top quark mass.

Because PDFs rely on non-perturbative QCD effects, they require input from experimental data. PDFs at a given scale Q_0^2 are extracted from fits to data and DGLAP [59–61] equations are used to predict PDFs to a higher scale Q^2. Different groups provide parameterizations of parton

densities. Among others, PDFs come from the "Coordinated-Theoretical-Experimental Project on QCD" (CTEQ collaboration) [62] and from the Martin, Roberts, Stirling and Thorne (MRST) group [63, 64]. The central PDFs used in this analysis are the CTEQ5L set, although sets from the MRST group are used to evaluate systematics (see Section 9.1).

Figure 2.4 shows an example of the CTEQ5L parton distribution function for the scale $Q^2 = (175\,\mathrm{GeV})^2$. These PDFs have to be folded with the partonic cross sections to calculate the measurable cross section in $p\bar{p}$ collisions.

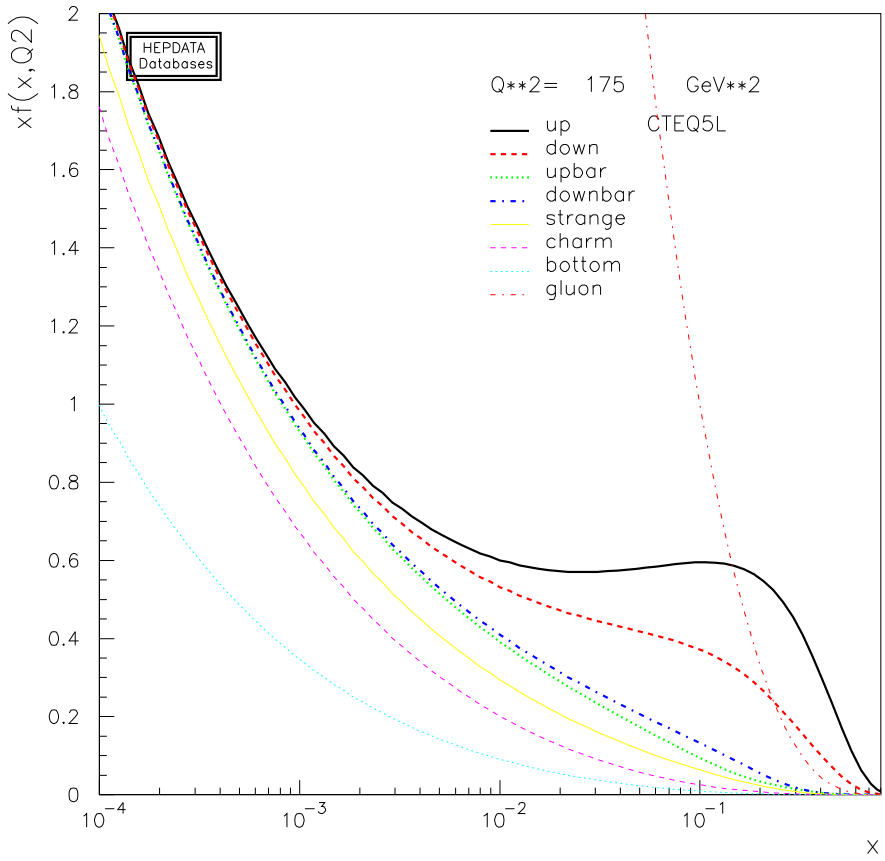


Figure 2.4: The CTEQ5L parton distribution functions at $Q^2 = (175\,\mathrm{GeV})^2$.

2.2. Physics Beyond the Standard Model

In spite of all its success, the SM is not complete, many mysteries remain. The already mentioned non null mass of the neutrinos calls for an extension of the SM. Another caveat of the SM that needs to be solved: the Higgs mass is subject to divergent quadratic radiative corrections that need to be somehow controlled in order for the Higgs mass to remain at the electroweak scale. If the SM is to remain valid up to the Planck scale without extension, such a cancellation requires some fine-tuned cancellation: this is referred to as the hierarchy problem. Several ways of solv-

ing the hierarchy problem have been proposed. For example, new strong dynamics could appear around 1 TeV (technicolor theories).

Another possibility is that the radiative corrections are canceled by a new spectrum of particles at the electroweak scale: supersymmetric (SUSY) theories propose that to every SM particle corresponds a supersymmetric partner with different spin, so that radiative correction contributions to the Higgs mass from a particle is canceled by the contribution from its supersymmetric partner. To SM fermions (bosons) correspond bosonic (fermionic) superpartners. For example, the superpartner of the top quark is called stop, the superpartner of the gluon is the gluino $\tilde{g}$, and the superpartner of the gauge bosons W and Z are the gauginos $\tilde{\chi}^0$, $\tilde{\chi}^{\pm}$. SUSY requires additional Higgs fields in order to provide mass to both up and down families. In the minimal supersymmetric extension of the SM (MSSM), there are five Higgs bosons: h, H, A, and $\mathrm{H}^{\pm}$.

Furthermore, the SM is unable to describe gravity, and cannot explain the existence of dark matter and dark energy without some extension.

2.3. Top Quark Physics

At present, top quarks can only be directly produced at the Tevatron. The top quark is, by far, the heaviest of the six fundamental quarks in the Standard Model of particle physics. Its large mass made the search for the top quark a long process, since accelerators with high center-of-mass energies are needed. In 1977 the discovery of the bottom quark indicated the existence of a third quark generation, and shortly thereafter the quest for the top quark began. Searches were conducted in electron-positron (e^+e^-) and proton-antiproton ($p\bar{p}$) collisions during the 1980s and early 1990s. Finally, in 1995 the top quark was discovered at the Fermilab Tevatron $p\bar{p}$ collider by the CDF and DØ collaborations [6, 7]. Subsequently, its mass has been precisely measured to be $m_{\mathrm{top}} = (173.1 \pm 1.3)\,\mathrm{GeV}/c^2$ [1] . The relative precision of this measurement, less than 1%, is better than our knowledge of any other quark mass. As is shown in Fig. 2.5, the top quark is about 40 times heavier than the second-heaviest quark, the bottom quark. Its huge mass, at the scale of the electroweak symmetry breaking, hints that it may play a role in the mechanism of mass generation, and makes it an ideal probe for new physics beyond the SM. Furthermore, its large mass implies an extremely short lifetime of $\tau_{\mathrm{top}} \approx 4 \cdot 10^{-25}$ s which is smaller than the characteristic formation time of hadrons $\tau_{\mathrm{form}} \approx 1/\Lambda_{\mathrm{QCD}} \approx 2 \cdot 10^{-24}$ s. Consequently no top-flavored hadrons can form and the top quark provides a unique opportunity to study a bare quark which passes all its properties, including spin information, to its decay products. It decays almost exclusively into a b quark and a W boson ($|V_{tb}| \approx 1$). Decays into quarks of the first and second generation are strongly suppressed by small CKM matrix elements.

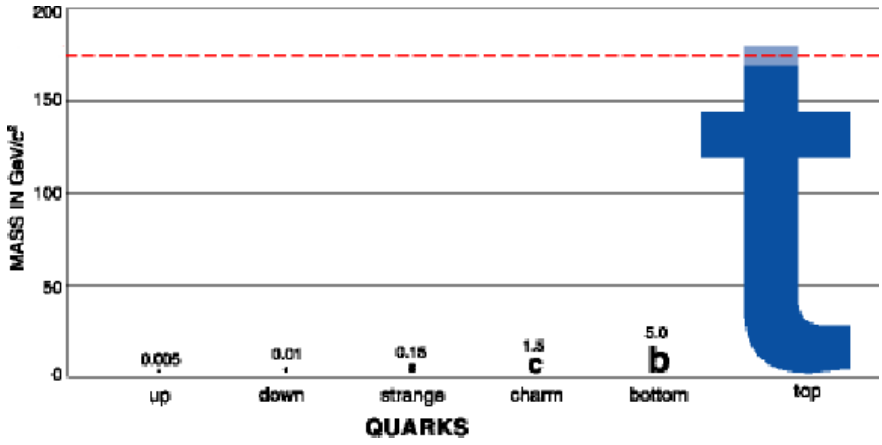


Figure 2.5: Comparison of masses of SM quarks.

2.3.1. Top Quark Pair Production

In proton-antiproton ($p\bar{p}$) collisions at Fermilab's Tevatron, the dominant top quark production mode is the top pair ($t\bar{t}$) production via the strong interaction. Even though $t\bar{t}$ pairs can be produced via the electroweak interaction through an exchange of a Z boson or photon, this contribution is negligible compared to the strong QCD cross section.

Figure 2.6 shows the corresponding tree level Feynman diagrams for the $t\bar{t}$ production processes. The production process shown on the left is called $q\bar{q}$ annihilation and the others are called gg fusion. Calculation at the next-to-leading order predicts that the relative contributions to the $t\bar{t}$ production from these two processes at Tevatron Run II are 85% and 15% respectively.

Calculations at next-to-leading order (NLO), including initial- and final-state gluon bremsstrahlung, gluon splitting, and virtual additions to the LO processes, contribute with α_s^3 to the perturbation series. At the Tevatron, the corrections to the cross section are dominated by initial-state gluon radiation. Further calculations of these soft radiative corrections at higher orders lead to an overall enhancement compared to NLO. The current approximate next-to-next-to-leading-order (NNLO) prediction of the $t\bar{t}$ production cross-section at the Tevatron is $6.9^{+0.5}_{-0.6}$ pb [5], for a top-quark mass of $m_{top} = 175\,\mathrm{GeV}/c^2$.

Top Quark Discovery

Immediately after the discovery of the b quark in 1977 the existence of a weak isospin doublet partner, the top quark, was hypothesized. The mass of the sixth quark was unknown and a wealth of predictions appeared based on many different speculative ideas, see for example references [65–67]. Typical expectations were in the mass range of about $20\,\mathrm{GeV}/c^2$, which became accessible two years later with measurements at the PETRA e^+e^- collider.

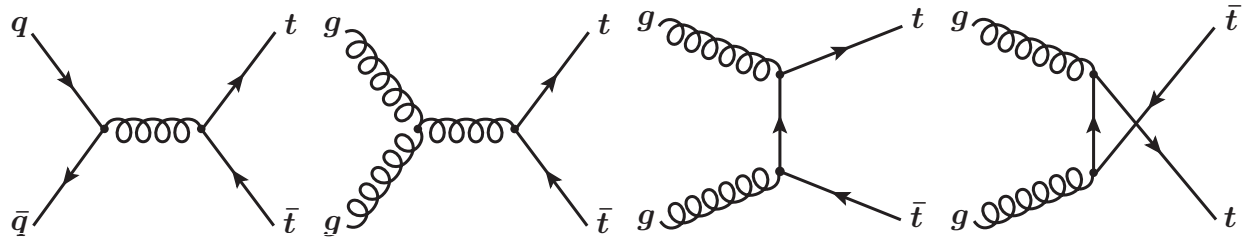


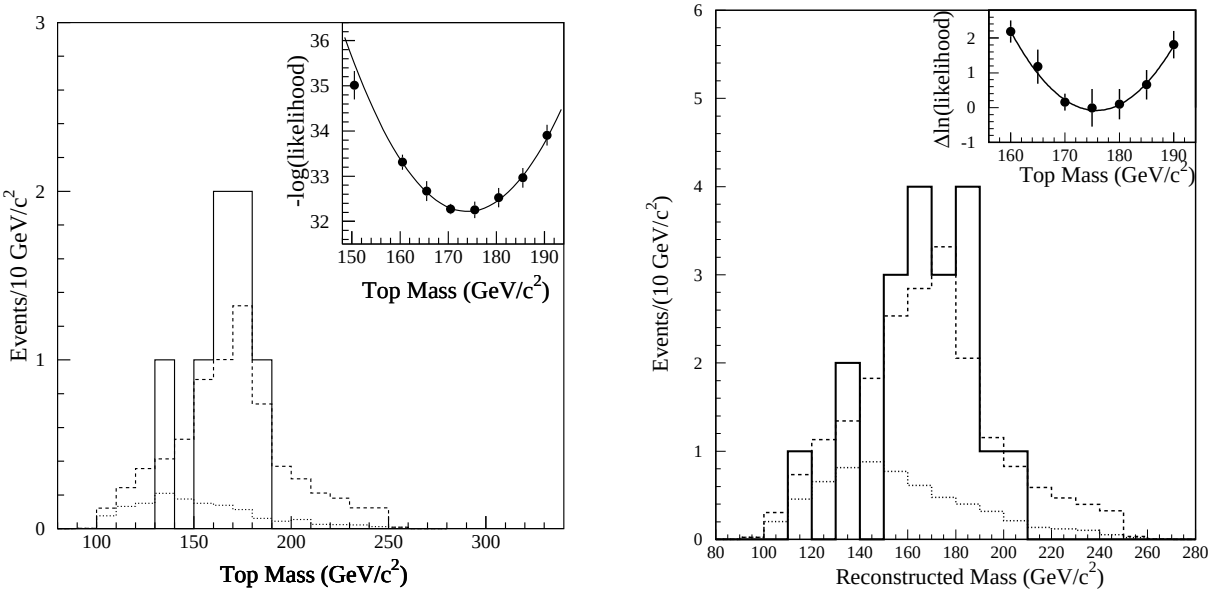
Figure 2.6: Feynman diagrams of the leading order processes for $t\bar{t}$ production: quark-antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$) and gluon fusion ($gg \rightarrow t\bar{t}$). The $q\bar{q}$ annihilation mode is dominant at the Tevatron energies.

The CDF experiment at the Tevatron joined the race for discovery of the top quark in 1988. Due to the higher center-of-mass energy at the Tevatron of $\sqrt{s} = 1.8$ TeV, top quarks are predominantly produced as $t\bar{t}$ pairs. The first CDF top quark search uses a data sample with an integrated luminosity of 4.4 pb^{-1} accumulated in Run 0 which lasted from 1988 to 1989 [68, 69]. This search pushed the lower limit on the top quark mass to $m_{\mathrm{top}} > 91\,\mathrm{GeV}/c^2$ at the 95% confidence level [70].

In 1992, with the start of Tevatron Run I, the DØ experiment joined the hunt for the top quark. In April of 1994, DØ published its first top quark analysis setting the last lower limit on the top quark mass before its discovery [71]. The data sample was recorded in 1992/93 and corresponds to an integrated luminosity of $15\ \mathrm{pb}^{-1}$. The intersection of the derived upper limit on the $t\bar{t}$ cross section with the theoretical prediction [72] yields a lower limit on the top quark mass of $131\,\mathrm{GeV}/c^2$ [71].

In 1993 and 1994, CDF saw mounting evidence for a top quark signal. The detector upgrade for Run I, mainly the addition of a silicon vertex detector, was the keystone for the discovery of the top quark at CDF. The new silicon detector allowed for the reconstruction of secondary vertices of b hadrons and a measurement of the transverse decay length L_{xy} with a typical precision of $130\,\mu\mathrm{m}$. Secondary vertex b tagging proved to be a very powerful tool to discriminate the top quark signal against the W+jets background and increase the sensitivity of the lepton-plus-jets $t\bar{t}$ analysis. In July 1994, CDF published a paper announcing first evidence for $t\bar{t}$ production at the Tevatron based on events in the dilepton and the lepton-plus-jets channel [73, 74]. The analysis uses a data sample with an integrated luminosity of $(19.3 \pm 0.7)\ \mathrm{pb}^{-1}$. The resulting top mass distribution, shown in Fig. 2.7(a), is fitted to a sum of the expected distributions from W+jets and $t\bar{t}$ production for different top quark masses. The fit yields a value of $m_{\mathrm{top}} = (174 \pm 10^{+13}_{-12})\ \mathrm{GeV}/c^2$.

In November 1994 the DØ collaboration confirmed the evidence seen at CDF. An update of the previous DØ analysis, now with an integrated luminosity of $(13.5 \pm 1.6)\ \mathrm{pb}^{-1}$, added soft muon b



(a) (b)

Figure 2.7: Reconstructed top mass distributions as published (a) in the CDF evidence paper of 1994 [73] and (b) in the CDF discovery paper of 1995 [6]. The solid histogram shows CDF data. The dotted line shows the shape of the expected background, the dashed line the sum of background plus $t\bar{t}$ Monte Carlo events for $M_{\rm top} = 175\ {\rm GeV}/c^2$. In both plots, the inset shows the likelihood curve used to determine the top quark mass.

tagging [75, 76]. In total, DØ observed nine events over a background of 3.8 ± 0.9.

As Run I continued more data were accumulated and finally, in April 1995, CDF and DØ were able to claim discovery of the top quark [6, 7]. CDF used a data sample corresponding to $67\ {\rm pb}^{-1}$ and significantly improved its background estimate and its b tagging techniques. Three separate analyses were performed at CDF as in the previous searches, a secondary vertex and a soft lepton b-taggers in the lepton+jets sample and a dilepton analysis. There are 27 jets with a secondary vertex b tag in 21 $W+ \geq 3$ jets events, with an estimated background of 6.7 ± 2.1 b tags. Six dilepton events are observed over a background of 1.3 ± 0.3. And 23 soft lepton tags are observed in 22 events, with 15.4 ± 2.0 b tags expected from background sources. Six events contain both a jet with a secondary vertex and a soft lepton tag. The probability for all CDF data events to be due to a background fluctuation alone is $1 \cdot 10^{-6}$, which is equivalent to a 4.8 σ deviation in a Gaussian distribution. Again the top quark mass is kinematically reconstructed for $W+ \geq 4$ jets events as described above. The mass distribution is shown in Fig. 2.7(a) (right). The best fit is obtained for $m_{\rm top} = (176 \pm 8 \pm 10)\ {\rm GeV}/c^2$.

Simultaneously to CDF, the DØ collaboration updated its top quark analyses based on data with an integrated luminosity of $50\ {\rm pb}^{-1}$ [7]. The updated analysis is very similar to the previ-

ous searches, involving the dilepton channel, soft muon b tagging and the topological analysis. From all channels, DØ observes 17 events with an expected background of 3.8 ± 0.6 events. The probability for this measurement to be an upward fluctuation of the background is $2 \cdot 10^{-6}$, which corresponds to 4.6 standard deviations for a Gaussian probability distribution. To measure the top quark mass, lepton+4 jets events are subjected to a constrained kinematic fit. A likelihood fit to the observed mass distribution yields a central value for the top quark mass of $m_{top} = 199^{+19}_{-21}\ (\text{stat.}) \pm 22\ (\text{syst.})\ \text{GeV}/c^2$.

Finally, 17 years after the discovery of the b quark, its weak isospin partner, the top quark, was firmly established. The good agreement between the measured top quark mass and the prediction obtained from electroweak precision measurements constituted a major success of the Standard Model.

2.3.2. Electroweak Single Top Quark Production

Besides the strong production of top quark pairs, the production of single top-quarks via electroweak interaction is predicted by the SM as well. Two electroweak production modes are dominating at the Fermilab Tevatron: the t-channel process (Figure 2.8(a)) and the s-channel process (Figure 2.8(b)). In $p\bar{p}$ collisions the third electroweak production mode, the Wt associated production (Figure 2.8(c)) has by comparison a negligible small predicted cross section. Since electroweak top quark production proceeds via a Wtb vertex, it provides the unique opportunity of the direct measurement of the CKM matrix element $|V_{tb}|$.

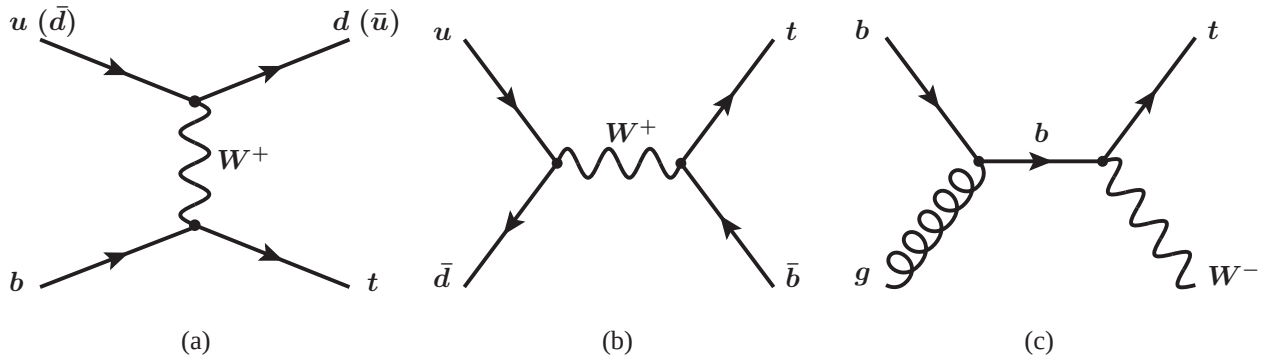
All three production modes are distinguished by the virtuality Q^2 of the participating W boson ($Q^2 = -q^2$, where q is the four-momentum of the W):

1. The ***t*-channel** ($q^2 = \hat{t}$): A virtual space-like W boson ($q^2 < 0$) strikes a sea b quark inside the proton or antiproton. The Feynman diagram representing this process at leading-order is shown in Figure 2.8(a). This mode is also known as *W-gluon fusion*, since the process actually involves a virtual gluon splitting into a $b\bar{b}$ pair, with one of the bottom quarks participating in the hard scattering. The predicted NLO cross-section at the Tevatron is $\sigma_{t-\text{chan}} = 1.98^{+0.28}_{-0.22}$ pb assuming $m_{top} = 175\ \text{GeV}/c^2$ [12, 13]. The overall uncertainty includes the choice of the factorization scale ($\pm 4\%$), the choice of PDF parameterization ($^{+11.3}_{-8.1}\%$), and the uncertainty in the top quark mass ($^{+1.6}_{-1.75}\%$). The mass of the b quark and the error in α_s play an insignificant role in the uncertainty [12].

2. The ***s*-channel** ($q^2 = \hat{s}$): This production mode, also called W^* production, is of Drell-Yan type. A time-like W boson ($q^2 \geq (m_{\text{top}} + m_b)^2$) is produced by the fusion of two quarks. Figure 2.8(b) shows the leading-order Feynman diagram for this process. The

cross-section prediction at NLO at the Tevatron energies is $\sigma_{s-\mathrm{chan}} = 0.88^{+0.12}_{-0.11}$ pb for $m_{top} = 175\ \mathrm{GeV}/c^2$ [12, 13]. The uncertainty includes $(^{+5.7}_{-5.70}\%)$ due to the factorization scale, $(^{+4.7}_{-3.9}\%)$ due to the PDF parameterization and $(^{+2.33}_{-2.71}\%)$ due to the uncertainty in the top quark mass, and a small contribution from the b quark and the error in α_s.

3. **Associated production**: The top quark is produced in association with an on-shell W boson ($q^2 = m_W^2$). The initial b quark is a sea quark inside the proton or antiproton. Figure 2.8(c) shows the leading order Feynman diagram for this process. The cross section for this mode is negligible at the Tevatron, therefore it will be further ignored in this document.



(a) (b) (c)

Figure 2.8: Leading-order Feynman diagrams for the single top quark production modes. Figures a) and b) are t- and s-channel processes respectively, while c) is associated Wt production which contribution is negligible at the Tevatron. For antitop quark production, the charge conjugate processes are implied.

In $p\bar{p}$ collisions the cross section is dominated by contributions from up and down quarks coupling to the W boson on one side of the Feynman diagrams as shown in Figure 2.8. There is, of course, also a small contribution from s or c sea-quarks in the initial state; an effect of about 2% for s- and 6% for t-channel production [77, 78].

There are also other production modes other than those involving the Wtb vertex, however, the production channels via a Wtd or a Wts vertex are strongly suppressed due to small CKM matrix elements. Their contribution to the total cross section is $\sim 0.1\%$ and $\sim 1\%$ respectively, and thus negligible at the Tevatron.

CKM Matrix Element $|V_{tb}|$

Measurement of the electroweak single top production provides the unique opportunity to directly measure the value of the CKM matrix element $|V_{tb}|$. The single top quark production cross-section is directly proportional to $|V_{tb}|^2$, hence its measurement yields a direct extraction of the

$|V_{tb}|$ value, relying only the assumption that $|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$ and that new physics contributions affects only $|V_{tb}|$. Since it does not depend on the unitarity of the CKM matrix, such a measurement would be a good test of this unitarity, and could also provide a useful test for a fourth quark generation, or other non-Standard Model behavior.

Other measurements of $|V_{tb}|$ are possible indirectly, by studying the rate of top quark decays [9]. It can be extracted from the ratio of branching fraction $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq) = |V_{tb}|^2/(\sum_q |V_{tq}|) = |V_{tb}|^2$, where $q = b, s, d$. However, this assumes exactly three quark families and unitarity of the CKM family, in contrast, single top quark production allows a direct extraction that makes neither of these assumptions.

The SM predicts a $|V_{tb}|$ value very close to the unity (see Equation 2.26), however, theoretical models beyond the SM exist predicting a significant deviation of $|V_{tb}|$ from unity while explaining the current constraints on the CKM matrix [79]. Several minimal extensions of the SM, such as a vector-like top singlet t', or a complete new fourth generation of fermions, can introduce additional elements to the CKM matrix, preserving the unitarity of the first two rows but allowing a value of $|V_{tb}|$ considerably different from unity. Parameters for these models are constrained by various experimental results and theoretical considerations. However, a direct measurement from single top quark production would provide a far more stringent test on such models.

Single Top Quark Polarization

The polarization of top quarks represents and interesting way to probe the properties of the top quark interaction and its study allows for searches for physics beyond the standard model as it is sensitive to anomalous couplings of the top quark [14, 78, 80]. In the SM, the Wtb vertex is entirely left-handed, which means that the top quark polarization is passed on to the W boson and b quark into which the top quark decays. Since the W interaction with fermions into which it decays is also left-handed, the W polarization information is thus also reflected in the kinematics of its decay products. The same weak interaction is also responsible for single top quark production, which has the consequence that single top quarks also show a large degree of polarization.

In the semileptonic decay chain of a polarized top quark, the decay angular distributions are simple linear in the cosine of these decay angles [81]:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}\left(1 + \alpha_i \cos\theta_i\right), \tag{2.29}$$

where θ_i is the angle between the i-th decay product and the top quark spin vector in the top quark rest frame, and Γ is the partial width of the top quark decay in the SM. The degree to which each decay product is correlated with the spin is encoded in the value of α_i. In a situation where a

mixture of spin up and spin down top quarks is present Equation 2.29 becomes

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}\left(1 + \mathcal{A}_{\uparrow\downarrow}\alpha_i \cos\theta_i\right), \tag{2.30}$$

where $A_{\uparrow\downarrow} = \frac{N_\uparrow - N_\downarrow}{N_\uparrow + N_\downarrow}$ is the spin asymmetry.

For the charged lepton $\alpha_\ell = 1$ [81, 82], thus the charged lepton is maximally correlated with the top spin direction independently of the top and W masses. The charged lepton possesses a stronger correlation than its parent, the W boson ($\alpha_W \approx 0.4$). This is due to the the significant interference between the polarization states of the intermediate W boson [83].

Single top quark production provides a clean source of polarized top quarks in which Equation 2.30 can be exploited. The usual basis to study spin-related observables in high energy physics is the helicity basis, however, because of the large top quark mass the ultrarelativistic limit is not valid at the Tevatron and the helicity is not a Lorenz invariant quantity. Therefore, there is no *a priori* reason to believe that the helicity basis will give the best description of the spin of the top quarks at the Tevatron. The so-called "optimized basis" relies on the SM dynamics responsible for single top quark production which predicts the spin of the top quark to be 100% polarized in the direction of the d-type quark in the event [15, 84]. However, the exact direction of the d-type quark is unknown experimentally, so it is necessary to choose the direction which is most likely to be correct.

In the s-channel process, $u\bar{d} \to t\bar{b}$, the largest contribution to the total cross section comes from the case where the $\bar{d}$ is donated by the antiproton. In the antiproton basis (proton for the antitop production), 98% of the top quarks produced in the s-channel have spin-up ($A_{\uparrow\downarrow} = 0.96$).

For the t-channel process, the largest portion of the total cross section comes from $ub \to td$ ($ug \to t\bar{b}d$ in the $2 \to 3$ process), with the spectator jet containing the d-quark [83]. Furthermore, in those events where the d-type quark is in the initial state, the fact that the spectator jet in the final state tends to be produced in the forward direction [78] means that it is still not a bad choice for the spin quantization axis: it is "almost" in the ideal direction. Thus, we define the "spectator basis" as the basis in which we choose the spin axis to be aligned with the momentum of the spectator jet in the final state. In this basis, the top quark is produced in the spin up state more than 96% of the time ($A_{\uparrow\downarrow} = 0.93$) [15, 83].

Experimentally, the correlation expected from Equation 2.30 is smeared by several effects. In particular, for the t-channel production the optimal variable is the angle between the charged lepton and the untagged jet in the reconstructed top quark rest frame $\theta_{l,j}$. The momenta and angles of the final state objects is not perfect, in particular the neutrino z component is unknown and must be inferred from constraints to the W boson mass. Therefore, the top quark reconstruction will suffer large uncertainties, mainly for events where more than one jet is b-tagged and a choice

must be made for the top quark reconstruction. Furthermore, the lepton identification requires an isolation cut which depletes the angular distribution near $\cos\theta_{l,j} = 1$. After all selection cuts described in Chapter 5, the angular distribution of $\cos\theta_{l,j}$ for MC simulated events, as described in Chapter 6, are shown in Figure 2.9. The figure shows how the single top simulation follows the expected angular correlation taking into account the lepton isolation effect near $\cos\theta_{l,j} = 1$ and other resolution effects. A modified single top model where the top quark decays as in the SM but is produced through right-handed interaction (RRLL), is also included in the figure. This exotic model presents an angular correlation that is anticorrelated with that of the SM single top production. Background events are however flat. Hence, a study of the single top quark polarization using this angular distribution could distinguish SM single top production from other possible non-SM single top production.

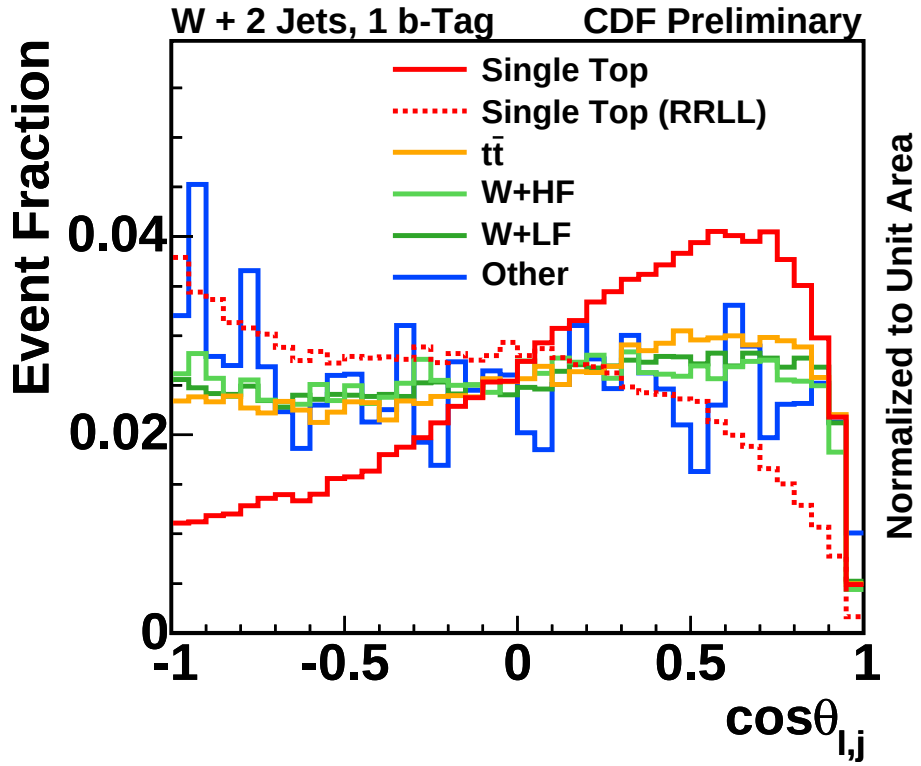


Figure 2.9: Cosine of the angle between the charged lepton and the non-tagged jet in the reconstructed top quark rest frame. Templates are shown for MC simulation of single top signal and backgrounds, as well as for an exotic signal model where the top quark is produced through right-handed interaction (RRLL).

Single Top as a Test for New Physics

The two modes of single top quark production are sensitive to quite different manifestations of physics beyond the Standard Model. Many such models have been proposed and can be classified as to whether they involve the effects of a new particle (either fundamental or composite) that couples to the top quark, or the effect of a modification of the SM coupling between the top and other known particles [14].

Examples of additional non-standard model particles include extra quark with fourth generation scenario (such as a b' quark that couples to W boson and a top quark), extra gauge boson (such as a W' vector boson [85, 86] that couples to top and bottom), extra scalar boson (such as a $H^{\pm}$ boson or charged top-quark-pion [87] that couples to top and bottom). Additional top-color [88–90], and top-flavor [91, 92] models, and gauged-flavor symmetry models [93], which give special dynamics to the third family in order to explain the large top mass, also lead to modified s- and t-channel cross section ratio.

Example of modified top quark interactions are anomalous Wtb couplings [78, 94, 95], and flavor changing neutral currents [96–100]. Extra dimension theories, such as Kaluza-Klein excitations of the W boson also predict altered single-top cross sections, in this case a decrease in the s-channel production rate [101].

Chapter 3

EXPERIMENTAL APPARATUS

The analysis presented in this thesis uses data collected between February 2002 and August 2008 from proton-antiproton collisions produced by the Tevatron at a center-of-mass energy of 1.96 TeV and observed by the Collider Detector at Fermilab (CDF). This chapter describes the collider apparatus and the detector.

Between 1997 and 2001, both the accelerator complex and the collider detectors underwent major upgrades, mainly aimed to increase the luminosity of the accelerator, and gather data samples of $2\ \mathrm{fb}^{-1}$ or more. The upgraded machine accelerates 36 bunches of protons and antiprotons, whereas the previous version of the accelerator operated with only 6. Consequently, the time between bunch crossings has been decreased from 3.5 μs for the previous version to 396 ns for the current collider. The center of mass energy was also increased from 1.8 to 1.96 TeV.

The new configuration required detector upgrades at CDF II to ensure a maximum response time shorter than the time between beam crossings.

In the subsequent sections, we describe how the proton and antiproton beams are produced, accelerated to their final center of mass energy of 1.96 TeV, and collided. We then describe the components used to identify and measure properties of the particles produced in the collision.

3.1. The Tevatron Collider and the Fermilab Accelerator Complex

The Fermilab's Tevatron Collider represents the high energy frontier in particle physics. It is a proton-antiproton storage ring system located at Fermilab (Fermi National Laboratory) in Batavia,

Illinois (USA). With a center-of-mass energy of $\sqrt{s}$ = 1.96 TeV it is currently the source of the highest energy proton-antiproton ($p\bar{p}$) collisions and the only apparatus capable of producing and directly studying top quarks. The collisions occur at two points on an underground ring, which has a radius of about 1 km. At these collision points there are two detectors: the Collider Detector at Fermilab (CDF II) and DØ. The 2 km diameter storage ring is the last step of a complex chain of accelerators that produce and accelerate the proton and antiproton beams (see Fig. 3.1).

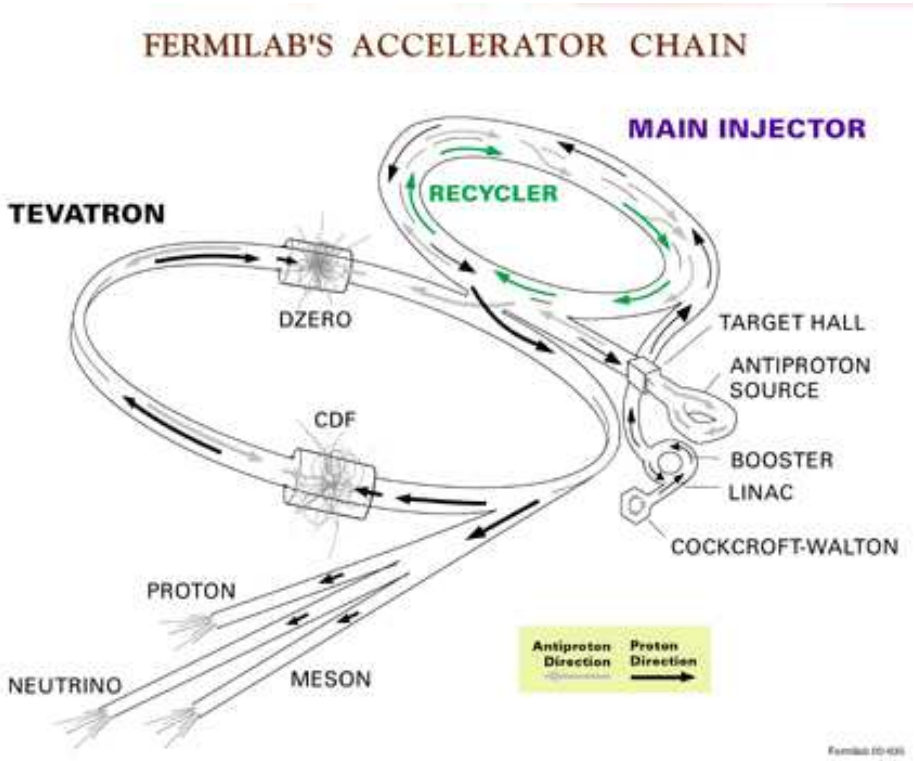


Figure 3.1: Layout of the Fermilab accelerator complex.

Production and acceleration of protons and antiprotons at Fermilab require a chain of accelerators, each one boosting particles to higher energies. Each step will be described in the following pages.

3.1.1. Proton Production and Boosting

The Cockcroft-Walton [102] pre-accelerator provides the first stage of acceleration. Inside this device, hydrogen gas is ionized to create H^- ions, which are accelerated to 750 keV of kinetic energy. The pre-accelerator produces 750 keV hydrogen ions every 66 ms.

Next, the H^- ions enter a linear accelerator (Linac) [103] , approximately 150 m long, where they are accelerated to 400 MeV. This acceleration is also done every 66 ms (with an offset to

catch the ions from the preacc). The Linac itself was upgraded in 1993, increasing its energy from 200 to 400 MeV; this made it possible, during Run IIb, to double the number of protons per bunch and to increase by about 50% the production rate of antiprotons. The acceleration in the Linac is done by a series of "kicks" from Radio Frequency (RF) cavities. The oscillating electric field of the RF cavities groups the ions into bunches.

The 400 MeV H^- ions are then injected into the Booster [103] , a circular synchrotron (circular accelerator) $74.5\ \mathrm{m}$ in diameter. A carbon foil strips the electrons from the H^- ions at injection, leaving bare protons. The intensity of the proton beam is increased by injecting new protons into the same orbit as the circulating ones. The protons are accelerated from 400 MeV to 8 GeV by a series of "kicks" applied by RF cavities. Each turn around the Booster, the protons accumulate about 500 keV of kinetic energy. The Booster is the first synchrotron in the Tevatron complex. It is composed of a series of 75 magnets arranged around a 74.5 m radius circle, with 18 RF cavities inside. This stage of production is also operated at 66 ms, with sufficient phase offset to catch the ions from the Linac.

Together, Linac and Booster are able to provide pulses of $5 \cdot 10^{12}$ protons for antiproton production every 1.5 s, or $6 \cdot 10^{10}$ protons per bunch in series of 5 to 7 bunches, repeated 36 times every 4 s.

At this point, protons are transferred to the Main Injector, a newly built circular accelerator that replaced the older Main Ring.

3.1.2. Main Injector

The Main Ring was originally built to provide 400 GeV protons to Fermilab's fixed target experiment; later on, it was converted to act as an injector to the Tevatron. The new operational requirements for the Main Ring did not match its original design. Therefore, during Run I, the Main Ring was a performance bottleneck, and the situation would be even worse in Run II.

The Main Injector was designed to solve this problem while providing further benefits, being capable of containing larger proton currents than its predecessor, which results in a higher rate of antiproton production. It is a 3 km long circular accelerator. It is composed of 18 accelerating RF cavities and can accelerate protons from a kinetic energy of 8 GeV to a total energy of up to 150 GeV every 2.2 s. The Main Injector can be used in other different operation modes:

- Antiproton production: it produces 120 GeV protons which are then used to strike the antiproton source and create antiprotons. This process is called "stacking pbars"
- Proton and antiproton boosting, before injection into the Tevatron in collider mode

- Antiproton deceleration, in order to recover antiprotons after a Tevatron collision run
- Proton and antiproton acceleration for fixed target experiment, either directly or as a booster for the Tevatron

3.1.3. Antiproton Production

In order to produce antiprotons, a pulse of $5 \cdot 10^{12}$ protons at 120 GeV is extracted from the Main Injector and focused into a nickel target. In the collisions, about 20 antiprotons are produced for each million protons, with a mean kinetic energy of 8 GeV. The antiprotons produced by the collision are collected by a lithium lens and separated from other particle species by a pulsed magnet.

Before the antiprotons can be used in the narrow beams needed in the collider, the differences in kinetic energy between the different particles need to be reduced. Since this process reduces the spread of the kinetic energy spectrum of the beam, it is referred to as "cooling" the beam. New batches of antiprotons are initially cooled in the Debuncher Ring (rounded triangular synchrotron with a mean radius of 90 m), collected and further cooled using stochastic cooling [104] in the 8 GeV Accumulator (also a rounded triangular synchrotron). The principle of stochastic cooling is to sample a particles motion with a pickup sensor and correct its trajectory later with a kicker magnet. In reality, the pickup sensor samples the average motion of particles in the beam and corrects for the average. Integrated over a long period of time, this manifests itself as a damping force applied onto individual particles which evens out their kinetic energies. It takes between 10 and 20 hours to build up a "stack" of antiprotons which is then used in collisions in the Tevatron. Antiproton availability is the most limiting factor for attaining high luminosities, assuming there are no technical problems with the accelerator (assuming, for example, perfect transfer efficiencies between accelerator subsystems) [103, 105] .

Roughly once a day, the stacked antiprotons (36 bunches of about 3×10^{10} antiprotons per bunch) are injected back into the Main Injector. They are accelerated to 150 GeV together with 36 bunches of roughly 3×10^{11} protons. Both the protons and antiprotons are transferred to the Tevatron.

3.1.4. Recycler Ring

Not all antiprotons in a given store are used up by the collisions. Recycling the unused antiprotons and reusing them in the next store significantly reduces the stacking time. The task of the Recycler is to receive antiprotons from a Tevatron store, cool them and re-integrate them into the stack, so that they can be used in the next store.

Antiproton production is one of the limiting factors in the efficiency in the Tevatron; by recycling 2/3 of these antiprotons, the average luminosity can be increased by a factor of two.

The Recycler Ring lies in the same enclosure as the Main Injector and, contrarily to the other rings at Fermilab, it is built with permanent magnets. During Run I the antiproton accumulation ring was found to suffer some kind of failure approximately once a week; this led to the loss of the entire store. Permanent magnets, not being prone to the most common causes of failure (such as power loss and lightning) provide a very stable repository for up to $3 \cdot 10^{12}$ antiprotons at a time.

In principle, the Recycler Ring can accept antiprotons from the Tevatron after a store is ended, however this functionality has not been proved possible. The Recycler also maintains the antiprotons momenta at 8 GeV. It can transfer these antiprotons back into the Tevatron for shots. Right now, while the Recycler is not capable of accepting recycled protons from the Tevatron, it is being used to pull antiprotons off the Accumulator, "stacking" the ones which can then be injected into the Tevatron.

3.1.5. Tevatron

The Tevatron is the last stage of Fermilab's accelerator chain. It is a circular synchrotron with a 1 km radius. It is composed of eight accelerating cavities, quadrupole and dipole focusing magnets. The Tevatron is also cryogenically cooled to 4 K, and the accelerating cavities are made of superconducting materials. It is desirable to use superconducting magnets because the very large fields necessary to maintain TeV-scale energies would require currents so large that it is more cost effective to use superconducting magnets than ordinary resistive magnets.

The Tevatron is not a perfect circle. There are six sectors (A-F) and each one has five service buildings (0-4). The "0" sections have large straight sections. A0 is where the Tevatron tunnel connects to the injection point. It also contains one of two beam aborts. B0 contains CDF (which will be described below), and the D0 detector is aptly named for it's place along the ring. At B0 and D0, the colliding beams are focused into very narrow beamlines of order 32 $\mu\mathrm{m}$, and the beams then collide. C0 is the location of the other beam abort (protons only). E0 used to be the site of the old Main Ring transfer to the Tevatron, but now it is unused. F0 houses the RF stations which "kick" the beam back into position if it has wandered off its axis. It is also where the transfer lines from the Main Injector connect with the Tevatron. It also houses the transfer line to the antiproton source.

The Tevatron receives 150 GeV protons and antiprotons from the Main Injector or the Recycler (for antiprotons) and accelerates them to 980 GeV in 85 s. Since the antiprotons and protons are oppositely charged, they circle in opposite directions in the magnetic field, and are housed in the same ring. The beams are brought to collision at two "collision points", B0 and D0. The

two collider detectors, the Collider Detector at Fermilab (CDF II) and DØ, are built around the respective collision points. The Tevatron can then sustain both beams for hours at a time (called a "store").

The Tevatron can also be used in fixed-target mode: it can accelerate up to $3 \cdot 10^{13}$ protons at a time to an energy of 800 GeV and deliver single bunches to be used in proton, meson and neutrino experiments.

3.1.6. Luminosity

The number of collisions per second is described by the "luminosity", $\mathcal{L}$. Making use of the upgrades in the rest of the accelerator chain, the Tevatron can provide an initial luminosity of $10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. During a collider store, instant luminosity slowly decreases. In the early stage of the store, the most important cause for this decrease is intrabeam scattering; some hours later, the depletion of antiprotons during collisions becomes more relevant. Luminosity is expected to decrease to 50% in about seven hours, and to 1/*e* in twelve hours. After a typical store duration of eight hours, 75% of the antiprotons are still available; they are decelerated in the Tevatron and in the Main Injector, and then stored in the Recycler Ring.

The luminosity of collisions can be expressed as:

$$\mathcal{L} = \frac{f N_B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right), \tag{3.1}$$

where f is the revolution frequency in Hz, N_B is the number of bunches, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch, and $\sigma_{p(\bar{p})}$ is the protons (antiprotons) RMS beam size at the interaction point. This is multiplied by a form factor, F, that depends on the ratio of the bunch longitudinal RMS size, σ_l, and the beta function at the interaction point, β^*, which is a measure of the transverse beam width and it is proportional to the beam's x and y extent in phase space. Table 3.1 shows a comparison of Run I and design Run II [105] accelerator parameters. Figure 3.2 shows the total luminosity collected by CDF as of December 2009.

However, the luminosity is not determined from this formula, but from the measured rate of some reference physical processes. The measurement of the luminosity delivered by the Tevatron to the CDF experiment is described in Sec. 3.2.7.

3.1.7. Beam Monitors

Operation of colliders at the Tevatron requires a constant monitoring of the beam position and luminosity. From a conceptual point of view, this is done in Run II as it was done in Run I.

Parameter	Run Ib	Run II
Number of bunches (N_B)	6	36
Protons/bunch (N_p)	2.3×10^{11}	2.7×10^{11}
Antiprotons/bunch ($N_{\bar{p}}$)	5.5×10^{10}	3.0×10^{10}
Total antiprotons	3.3×10^{11}	1.1×10^{12}
β^* [cm]	35	35
Bunch length [m]	0.6	0.37
Bunch spacing [ns]	3500	396
Interactions/crossing	2.5	2.3
Energy [GeV/particle]	900	980
Integrated luminosity [pb^{-1}]	112	1800
Peak luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	2×10^{31}	$\sim 2 \times 10^{32}$

Table 3.1: Accelerator parameters for Run I and Run II configurations.

The luminosity monitor consists of two arrays of scintillators, placed on both sides of the interaction region. A coincidence of particles moving away from the interaction point, both in the p and $\bar{p}$ direction, is interpreted as a contribution to luminosity. Bunches of particles moving in a single direction, without a coincident bunch in the opposite direction, are flagged as beam losses.

The beam position, on the other hand, is measured by the collider detectors themselves. During Run I, the detector was able to locate the beam within 5 μm in about five minutes; other beam parameters, such as slope and transverse profile, were calculated over longer time intervals (about two hours). In Run II, the same operations are performed but more quickly.

3.2. The CDF II Detector

The CDF II Detector [106, 107] is a substantial upgrade of the original CDF Detector [108] . It is located at the B0 collision point of the Tevatron Collider. The detector is designed to detect and measure properties of particles emanating from $p\bar{p}$ collisions. The design is not geared toward one particular physics measurement, but rather optimized toward extracting a number of different properties about all particle species created in the $p\bar{p}$ collision. Such particle detectors are often called multi-purpose detectors.

A diagram of the CDF II Detector is shown in Fig. 3.3. A quadrant of the detector is cut out to expose the different subdetectors. The detector consists of 3 primary subsystems: the tracking, the calorimetry and the muon systems. All these systems surround the “beam pipe”, a vacuum tube of

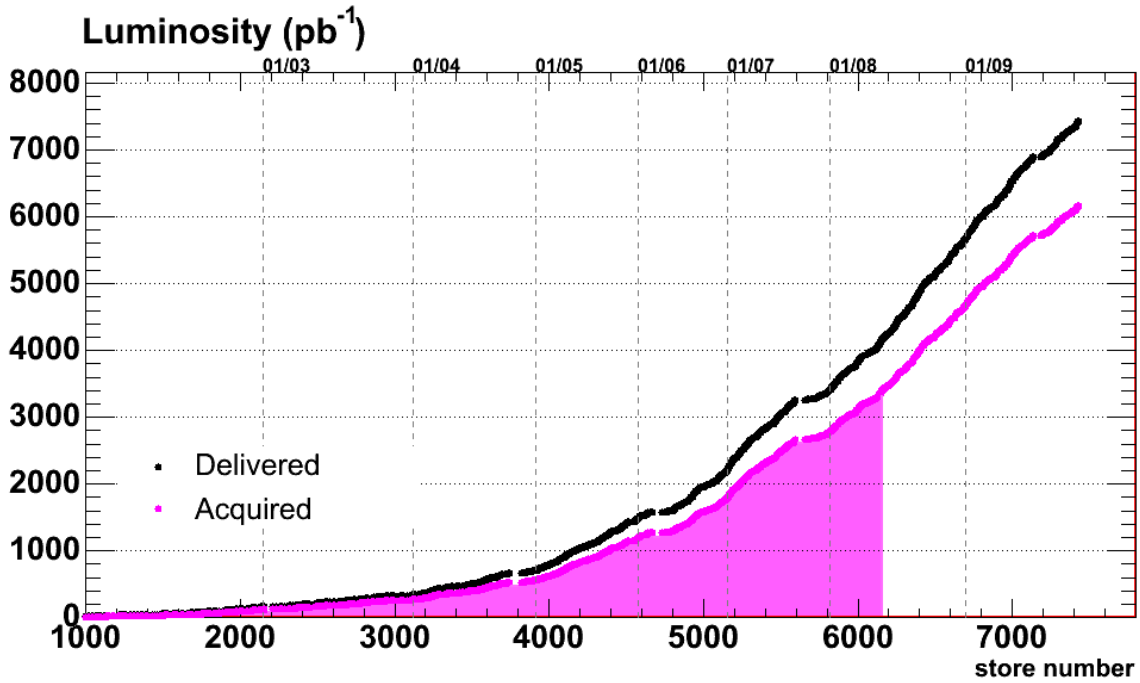


Figure 3.2: Total luminosity gathered by the CDF detector as of December 2009. The black curve is luminosity delivered and the purple curve is luminosity written to tape by CDF. The filled area indicates the integrated luminosity used in this analysis.

diameter 2.2 cm located at the innermost part of the detector and where the proton and antiproton beams travel and collide. The beam pipe is made of beryllium because this metal has the best mechanical qualities, yet lowest nuclear interaction cross section of all materials.

The detector subsystems can be grouped as follows. The innermost system is the integrated tracking system. The tracking system is barrel-shaped and consists of cylindrical subsystems which are concentric with the beam. It is designed to detect charged particles, measure their momenta and displacements from the point of collision (primary interaction vertex). The tracking system is surrounded by the Time of Flight system, designed to provide particle identification for low-momentum charged particles. Both the tracking and Time of Flight systems are placed inside a superconducting coil, which generates a 1.4 T solenoidal magnetic field. The coil is surrounded by calorimetry systems, which measure the energy of particles that shower when interacting with matter. The calorimetry systems are surrounded by muon detector systems. When interacting with matter, muons act as “minimally ionizing particles” - they only deposit small amounts of ionization energy in the material. Therefore, they are able to penetrate both the tracking and calorimeter systems. The integrated material of the tracking system, TOF, solenoid and calorimetry systems serves as a particle filter. Particles which penetrate through all that material are mostly muons, and they are detected by leaving tracks in the muon detection system, located outside of the calorimeter.

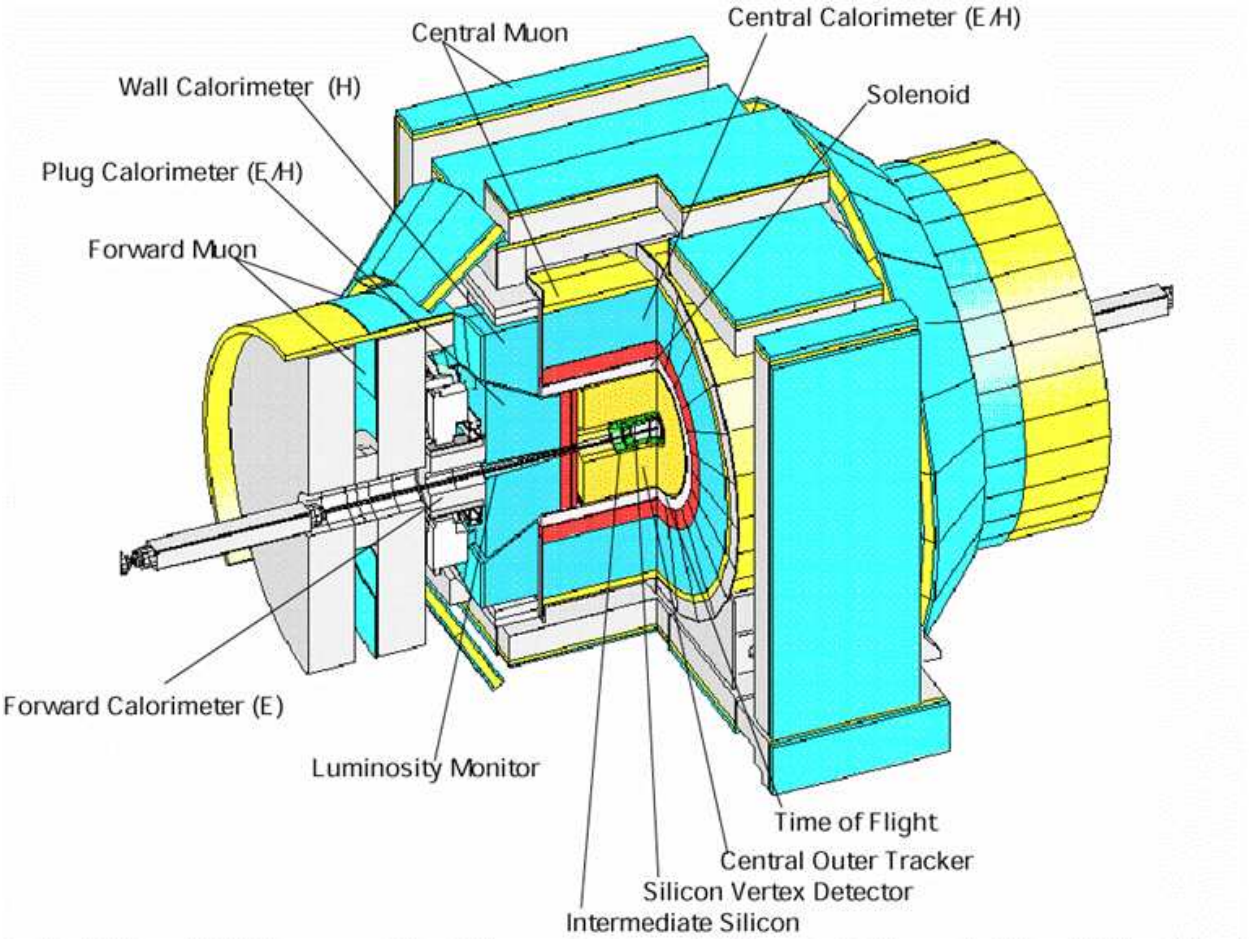


Figure 3.3: The CDF II Detector with quadrant cut to expose the different subdetectors.

The Cherenkov Luminosity Counter (CLC) measures the rate of interactions near the beam used for luminosity measurements.

The rest of this chapter will provide a short description of each detector subsystem, with an emphasis on the upgrades since Run I. More detailed information on each system can be found in the Technical Design Report of the CDF II Detector [106] .

3.2.1. Standard Definitions in CDF

Figure 3.4 shows an elevation view of the detector. Protons enter the detector from the west side and antiprotons enter from the east side.

Because of its barrel-like detector shape, the CDF II Detector uses a cylindrical coordinate system (r, ϕ, z) with the origin at the center of the detector and the z-axis along the nominal direction of the proton beam (toward east). The y-axis points upwards. Since the coordinate

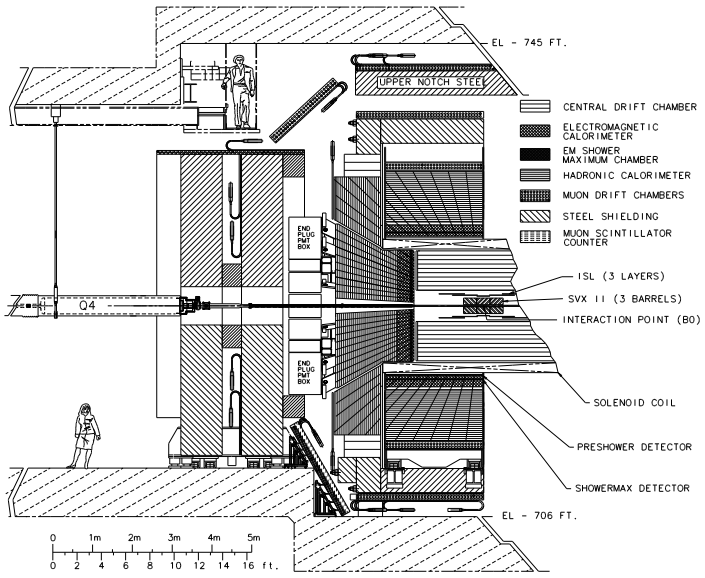


Figure 3.4: Elevation view of the Collider Detector at Fermilab (CDF). West (East) corresponds to the right (left) side of the picture.

system is right-handed, this also defines the direction of the x-axis; it is horizontal pointing north (outward with respect to the center of the Tevatron).

Spherical coordinates are also commonly used: the polar angle θ is defined with respect to the proton beam and the azimuthal angle ϕ is defined with respect to the x-axis. However, θ is not a good variable to use in this case because it is not a Lorentz invariant. Due to the fact that the proton (and antiproton) is an extended object, the actual constituent partons will not be traveling at 980 GeV. Thus, the number of particles per unit angle ($dN/d\theta$) will not be the same for particles with different velocity.

Instead, we use the concept of the *rapidity*, defined as

$$Y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{3.2}$$

where E is the energy and p_z is the z component of the momentum of the particle. For the high energy particles, $p \gg m$ so $E \sim p$ and the rapidity is approximated by the *pseudo-rapidity*, defined as

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

In this case, the number of particles per unit rapidity ($dN/d\eta$) is invariant under boosts in the z direction.

Particles moving through a homogeneous solenoidal magnetic field follow helical trajectories.

Reconstructed charged particle trajectories are referred to as "tracks". The plane perpendicular to the beam is referred to as the "transverse plane", and the transverse momentum of the track is referred to as p_T. As opposed to e^+e^- collisions, in $p\bar{p}$ collisions not all of the center of mass energy of the $p\bar{p}$ system is absorbed in the collision. The colliding partons inside the proton carry only a fraction of the kinetic energy of the proton. As a result, the center of mass system of the parton collisions is boosted along the beam direction (the "longitudinal" direction) by an unknown amount, but quantities defined in the transverse plane are conserved in the collisions. For instance, the sum of all transverse momenta of particles in a collision is zero, $\sum \vec{p_T} = 0$.

To uniquely parameterize a helix in three dimensions, five parameters are needed. The CDF II coordinate system chooses three of these parameters to describe a position, and two more to describe the momentum vector at that position. The three parameters which describe a position describe the point of closest approach of the helix to the beam line. These parameters are d_0, ϕ_0, and z_0, which are the ρ, ϕ and z cylindrical coordinates of the point of closest approach of the helix to the beam. The momentum vector is described by the track curvature (c) and the angle of the momentum in the $r-z$ plane ($\cot\theta$). From the track curvature we can calculate the transverse momentum. The curvature is signed so that the charge of the particle matches the charge of the curvature. From $\cot\theta$, we can calculate $p_z = p_T \cdot \cot\theta$. At any given point of the helix, the track momentum is a tangent to the helix. This basically means that the angle ϕ_0 implicitly defines the direction of the transverse momentum vector at the point of closest approach.

The impact parameter (d_0) of a track is another signed variable; its absolute value corresponds to the distance of closest approach of the track to the beamline. The sign of d_0 is taken to be that of $\hat{p} \times \hat{d} \cdot \hat{z}$, where $\hat{p}$, $\hat{d}$ and $\hat{z}$ are unit vectors in the directions of $\vec{p}$, $\vec{d_0}$ and $\vec{z}$, respectively.

For decaying particles, we often define the displacement L_{xy},

$$L_{xy} = \vec{d} \cdot \hat{p_T}\,, \tag{3.4}$$

where $\vec{d}$ is the displacement of the decay vertex in the transverse plane, and $\hat{p_T}$ is the unit vector in the direction of $\vec{p_T}$.

3.2.2. Tracking Systems

The detector has a cylindrical tracking system immersed in a 1.4 T solenoidal magnetic field for the measurement of charged-particle momenta. We will describe this system starting from the devices closest to the beam and moving outwards. The innermost tracking device is a silicon strip vertex detector, which consists of three subdetectors: Layer 00 (L00), the Silicon Vertex Detector (SVX-II) and the Intermediate Silicon Layers (ISL). Fig. 3.5 shows a view in the $r-\phi$ plane of the three subsystems.

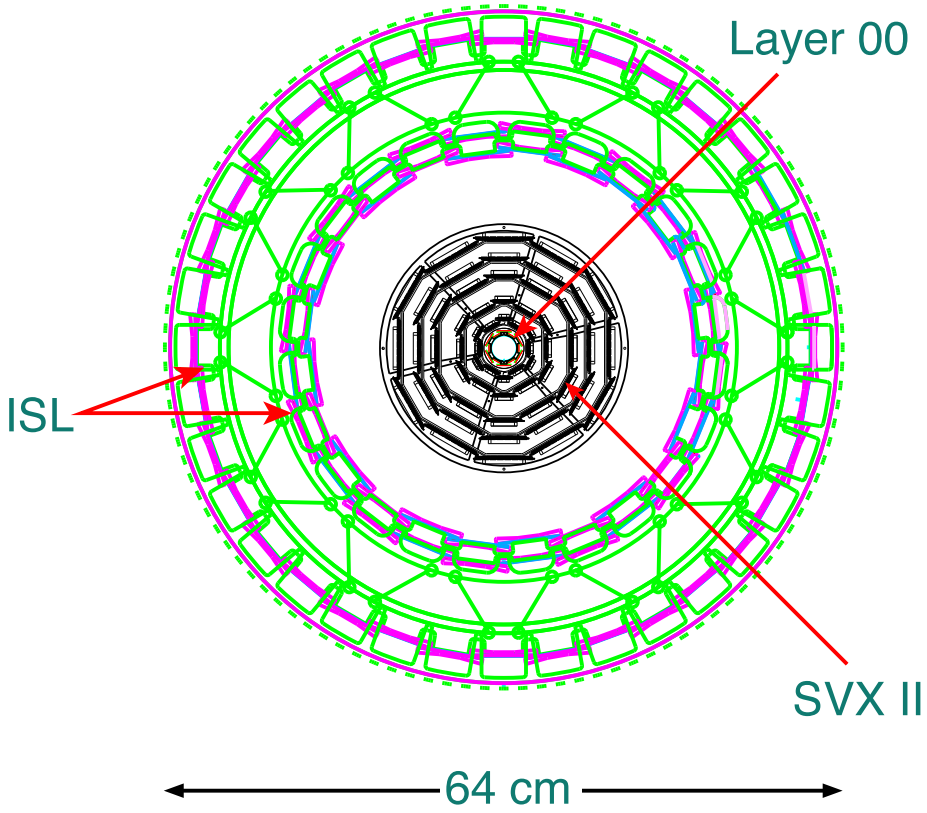


Figure 3.5: Transversal view of the Silicon Vertex Detector at CDF showing the different layers and parts of the detector.

Surrounding the silicon detector is the Central Outer Tracker (COT), a $3.1\ \mathrm{m}$-long cylindrical open-cell drift chamber covering radii from 43.4 to $132.3\ \mathrm{cm}$. Figure 3.6 shows the coverage of the whole tracking system.

The silicon detectors provide excellent impact parameter, azimuthal angle and z resolution. They are also instrumental in vertexing. The COT provides excellent resolution of the curvature, ϕ and η. Together they provide very accurate measurements of the helical paths of charged particles.

Silicon Tracking Detectors

Silicon tracking detectors are used to obtain precise position measurements of the path of a charged particle. A silicon tracking detector is fundamentally a reverse-biased p-n junction. When a charged particle passes through the detector material, it causes ionization. In the case of a semiconductor material, this means that electron-hole pairs will be produced. Electrons drift towards the anode, and holes drift toward the cathode, where the charge is gathered. The amount of charge is, to first order, proportional to the path length traversed in the detector material by the charged

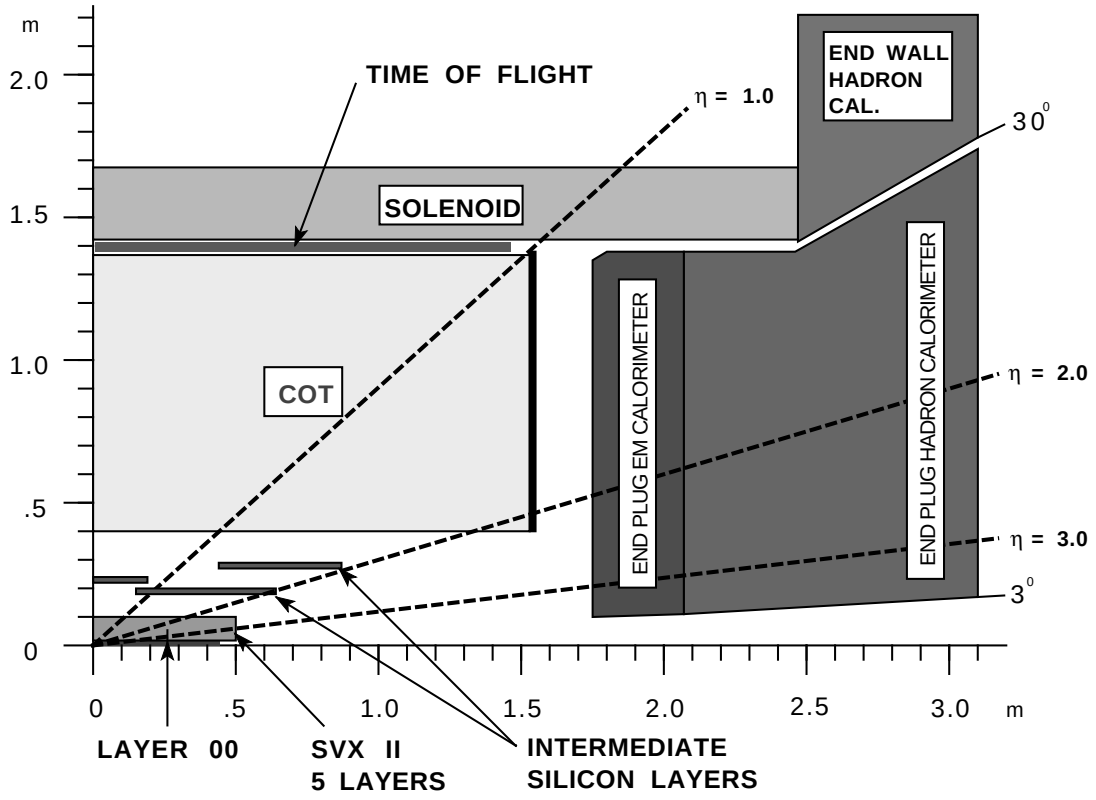


Figure 3.6: The CDF II tracker layout showing the different subdetector systems.

particle.

By segmenting the p or n side of the junction into "strips" and reading out the charge deposition separately on every strip, we obtain sensitivity to the position of the charged particle. All the CDF II silicon tracking detectors are implemented as microstrip detectors. The typical distance between two strips is about $60\ \mu\mathrm{m}$. Charge deposition from a single particle passing through the silicon sensor will be read out on one or more strips. This charge deposition is called a "cluster". There are two types of microstrip detectors: single and double-sided. In single-sided detectors only one (p) side of the junction is segmented into strips. Double-sided detectors have both sides of the junction segmented into strips. The benefit of double-sided detectors is that while one (p) side has strips parallel to the z direction, providing $r-\phi$ position measurements, the other (n) side can have strips at an angle (stereo angle) with respect to the z direction, which will give z position information.

The innermost layer, L00 [109] , is a radiation-hard, single-sided silicon detector installed directly onto the beryllium vacuum beam pipe. L00 is the most recent addition to the CDF II

tracker. The geometry of L00 is such that there are two overlapping hexagonal structures at radii between 1.35 and 1.62 cm from the beam. Figure 3.7 shows a detailed view of the L00. It only provides $r-\phi$ measurements. Being so close to the interaction point, L00 improves noticeably the spacial resolution up to $\approx$15 $\mu\mathrm{m}$ per hit.

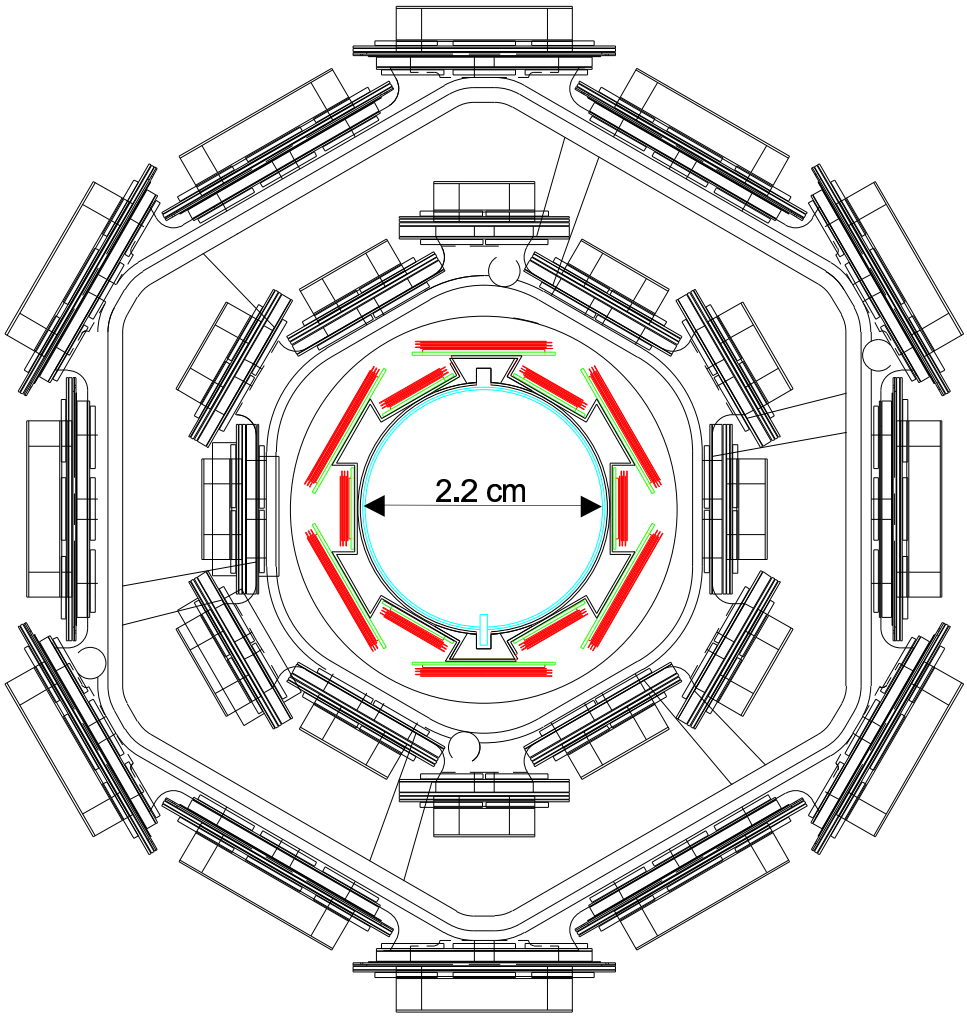


Figure 3.7: Detailed view of the Silicon L00 along with the two innermost layers of the SVX.

The layer of silicon on the beam pipe is followed by the SVX-II [110] . It consists of five concentric layers of double-sided silicon sensors. One side of each sensor provides measurements in the transverse plane (axial strips); the other side's strips deliver 3D information. SVX-II extends radially from 2.5 to $11\ \mathrm{cm}$, and along z up to 45 cm on either side of the interaction point. The spacial resolution of the SVX-II is $\approx$20 $\mu\mathrm{m}$.

The Intermediate Silicon Layers (ISL) [111] are the outermost silicon subdetector systems, consisting of one double-sided silicon, similar to those on SVX-II, placed at a radius of $22\ \mathrm{cm}$ in the central region ($|\eta| < 1$), and two forward layers ($1 < |\eta| < 2$) at radii 20 and $28\ \mathrm{cm}$ from the

beam line. Together with SVX-II, the ISL makes it possible to reconstruct tracks in the forward regions, which lie beyond the acceptance region of the outer tracker.

The SVX-II and ISL are made of double-sided silicon sensors. As shown in Table 3.2, the SVX-II layers have different stereo angles. Two layers have a 1.2° stereo angle and three have a 90° stereo angle. The ISL detector provides small angle (1.2°) stereo information.

Property	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4
Number of ϕ strips	256	384	640	768	869
Number of z strips	512	576	640	512	869
Stereo angle	90°	90°	-1.2°	90°	+1.2°
ϕ strip pitch $[\mu\mathrm{m}]$	60	62	60	60	65
z strip pitch $[\mu\mathrm{m}]$	141	125.5	60	141	65
Active width $[\mathrm{mm}]$	15.30	23.75	38.34	46.02	58.18
Active length $[\mathrm{mm}]$	72.43	72.43	72.38	72.43	72.43

Table 3.2: Relevant parameters for the layout of the sensors of the SVX-II layers.

Four silicon sensors are stacked length-wise into a "ladder" structure which is $29\ \mathrm{cm}$ long. The readout electronics are mounted onto the ends of the ladders. The ladders are organized in an approximately cylindrical configuration, creating "barrels". A SVX-II barrel is segmented into 12 wedges, each covering approximately 30° in ϕ with a small overlap at the edges, allowing for several silicon hits per track. There are three SVX-II barrels, adjacent to each other along the z-axis, covering the nominal interaction point in the center of the CDF II Detector. The coverage of the silicon detector subsystems is shown in Fig. 3.8. The silicon tracking system is used in stand-alone mode to provide an extension of tracking down to 2.8 in pseudorapidity.

Compared to the shorter, 4-layer, single-sided vertex detector of Run I, the new silicon tracker provides a much wider acceptance, better resolution, three-dimensional reconstruction and, as stated above, can be used in stand-alone mode without input from the Central Outer Tracker (described hereafter).

Central Outer Tracker

The Central Outer Tracker (COT) [112] is a multiwire drift chamber built to replace the one used in Run I (CTC). The active volume of the COT begins at a radius of $43.4\ \mathrm{cm}$ from the nominal beamline and extends out to a radius of $132.3\ \mathrm{cm}$. The chamber is $310\ \mathrm{cm}$ long. The COT contains 96 sense wire layers, which are radially grouped into eight "superlayers", as inferred

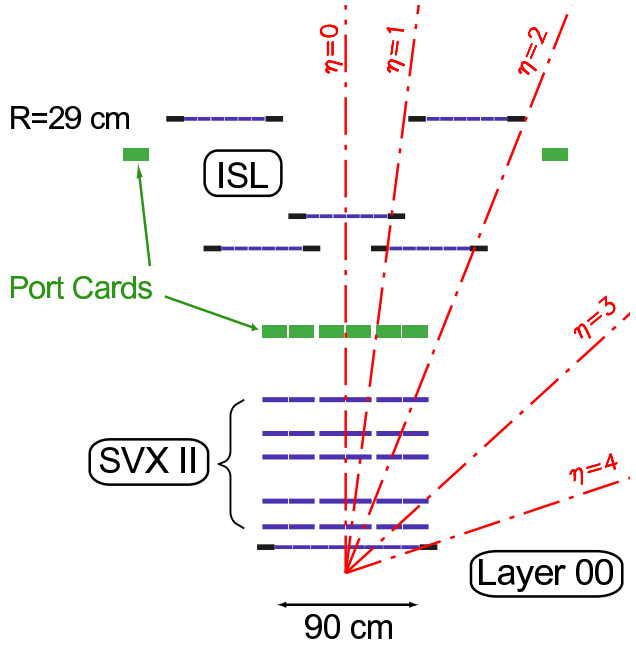


Figure 3.8: Coverage of the different silicon subdetectors projected into the $r-z$ plane. The r and z axes have different scales.

from the end plate section shown in Fig. 3.9(a). Four superlayers (axial superlayers) provide $r-\phi$ measurements and are alternated with the remaining four that provide 2° stereo measurements (stereo superlayers). Each superlayer is divided in ϕ into "supercells", and each supercell has 12 sense wires and a maximum drift distance that is approximately the same for all superlayers. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30,240 sense wires. Approximately half the wires run along the z direction ("axial"). The other half are strung at a small angle (2°) with respect to the z direction ("stereo"). Particles originating from the interaction point, which have $|\eta| < 1$, pass through all 8 superlayers of the COT. Particles which have $|\eta| < 1.3$ pass through 4 or more superlayers.

The COT drift chamber provides accurate information in the $r-\phi$ plane for the measurement of transverse momentum, p_T, and substantially less accurate information in the $r-z$ plane for the measurement of the z component of the momentum, p_z.

The supercell layout, shown in Fig. 3.9(b) for superlayer 2, consists of a wire plane containing sense, potential and shaper (for field shaping) wires and a field (or cathode) sheet on either side. Both the sense and potential wires are $40\ \mu\mathrm{m}$ diameter gold plated Tungsten. The field sheet is $6.35\ \mu\mathrm{m}$ thick Mylar with vapor-deposited gold on both sides. Each field sheet is shared with the neighboring supercell.

The COT is filled with an Argon-Ethane gas mixture and Isopropyl alcohol (49.5:49.5:1). The

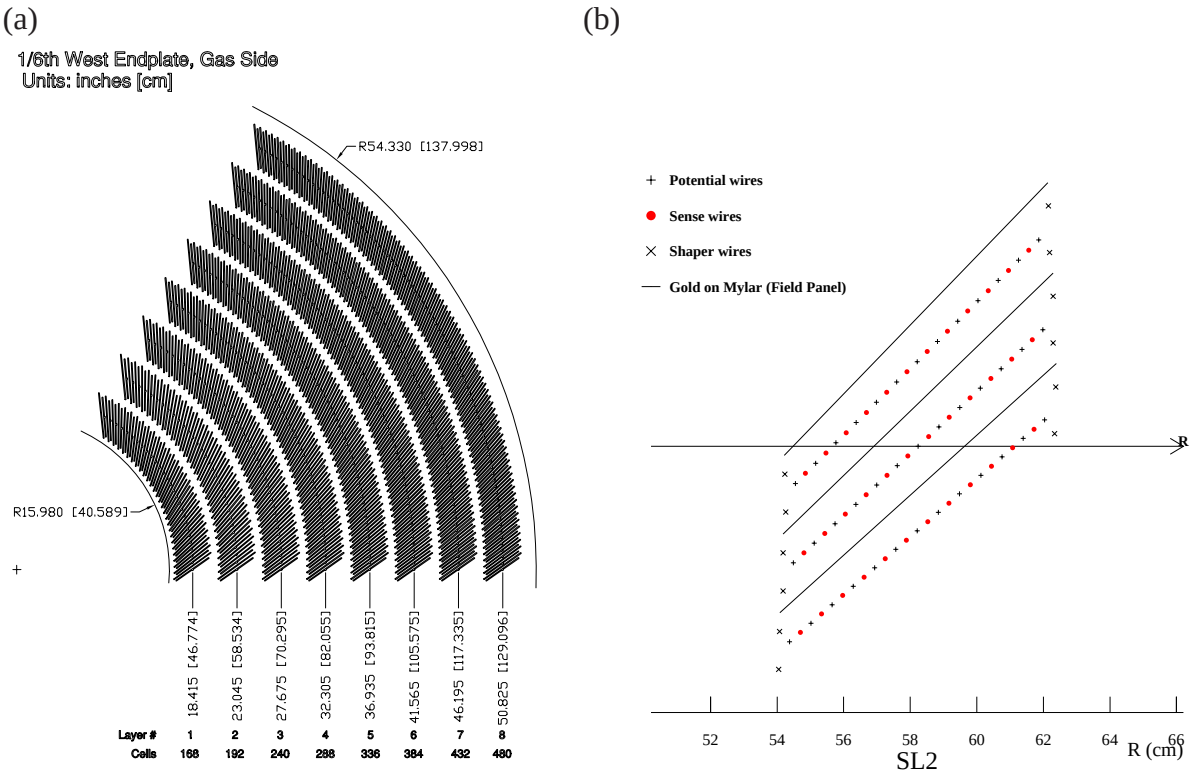


Figure 3.9: (a) Layout of wire planes on a COT endplate. (b) Layout of wires in a COT supercell.

mixture is chosen to have a constant drift velocity across the cell width. This allows a maximum drift time of 177 ns with a drift velocity of 100 $\mu\mathrm{m/ns}$. This prevents pileup of events in the drift chamber from the previous event.

When a charged particle passes through, the gas is ionized. Electrons drift towards the sense wires. The electric field in a cylindrical system grows exponentially with decreasing radius. As a result, the electric field very close to the sense wire is large, resulting in an avalanche discharge when the charge drifts close to the wire surface. This effect provides a gain of $\sim 10^4$. Due to the magnetic field that the COT is immersed in, electrons drift at a Lorentz angle of $\sim 35°$. The supercell is tilted by $35°$ with respect to the radial direction to compensate for this effect.

Signals on the sense wires are processed by the ASDQ (Amplifier, Shaper, Discriminator with charge encoding) chip, which provides input protection, amplification, pulse shaping, baseline restoration, discrimination and charge measurement [113]. The charge measurement is encoded in the width of the discriminator output pulse, and is used for particle identification by measuring the ionization along the trail of the charged particle (dE/dx). The pulse is sent through $\sim$ 11 m of micro-coaxial cable, via repeater cards to Time to Digital Converter (TDC) boards in the collision hall. Hit times are later processed by pattern recognition (tracking) software to form helical tracks. The hit resolution of the COT is about 140 $\mu\mathrm{m}$. The transverse momentum resolution has been

measured using cosmic ray events to be

$$\frac{\sigma_{p_T}}{p_T^2} = 0.17\% \, [\mathrm{GeV}/c]^{-1}. \tag{3.5}$$

Pattern Recognition Algorithms

As explained in the previous sections, charged particles leave small charge depositions as they pass through the tracking system. By following, or "tracking", these depositions, pattern recognition algorithms can reconstruct the charged particle track.

There are several pattern recognition algorithms used to reconstruct tracks in the CDF II tracking system. Most of the tracks are reconstructed using "Outside-In" (OI) algorithms which we will describe here. The name of this group of algorithms suggests that the track is followed from the outside of the tracking system inwards.

The track is first reconstructed using only COT information. The COT electronics report hit time and integrated charge for every wire in an event. The hit time corresponds to the time that an avalanche occurred at a sense wire. The hit time can be interpreted as the drift time of the charge in the gas, but first it has to be corrected for time of flight. The hit timing resolution is of the order of a few ns; this roughly corresponds to the average spread in collision times. It is assumed that the collision times always happen at the same time in a cycle during a store. An average of collision times is done for many previous events and this is used as the event collision time. Hit times corrected for the collision time are interpreted as drift times and used in pattern recognition. To perform the final track fit, an additional time of flight correction is performed assuming massless particles.

The helical track, when projected into the two dimensional $r-\phi$ plane, is a circle. This simplifies pattern recognition, so the first step of pattern recognition in the COT looks for circular paths in radial superlayers of the COT. Supercells in the radial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called "segments". The straight-line fit for a segment gives sufficient information to extrapolate rough measurements of curvature and ϕ_0. Once segments are found, there are two approaches to track finding. One approach is to link together segments for which the measurements of curvature and ϕ_0 are consistent. The other approach is to improve the curvature and ϕ_0 measurement of a segment reconstructed in superlayer 8 by constraining its circular fit to the beamline, and then adding hits which are consistent with this path. Once a circular path is found in the $r-\phi$ plane, segments and hits in the stereo superlayers are added by their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. This results in a high track reconstruction efficiency ($\sim 95\%$) in the COT for tracks which pass through all 8 superlayers

($p_T \geq 400\ \mathrm{MeV/c}$). The track reconstruction efficiency mostly depends on how many tracks there are to be reconstructed in the event. If there are many tracks present close to each other, hits from one track can shadow hits from the other track, resulting in efficiency loss.

Once a track is reconstructed in the COT, it is extrapolated into the SVX-II. Based on the estimated errors on the track parameters, a three-dimensional "road" is formed around the extrapolated track. Starting from the outermost layer, and working inwards, silicon clusters found inside the road are added to the track. As a cluster gets added, the road gets narrowed according to the knowledge of the updated track parameters. Reducing the width of the road reduces the chance of adding a wrong hit to the track, and also reduces computation time. In the first pass of this algorithm, $r-\phi$ clusters are added. In the second pass, clusters with stereo information are added to the track.

For the identification of electrons in the forward region, a special algorithm, called Phoenix (PHX) tracking, is used. This forward tracking algorithm is analog to the outside-in tracking where an energy cluster in the PEM, instead of a COT track, and the primary vertex are used to construct seed tracks. For each seed, two hypotheses about the charge of the particle are considered by computing the curvature for both an electron and a positron corresponding to the deposited energy. The extrapolation of those seed helices into the silicon sub-detector works similarly to the outside-in tracking algorithm.

3.2.3. Time of Flight

Outside the tracking system, still inside the superconducting magnetic coil, CDF II has a Time of Flight (TOF) [114] system. The TOF system is designed to distinguish low momentum pions, kaons and protons by measuring the time it takes these particles to travel from the primary vertex of the $p\bar{p}$ collision to the TOF system. The system consists of 216 bars of scintillating material, roughly $300\ \mathrm{cm}$ in length and with a cross section of $4 \times 4\ \mathrm{cm}^2$. The bars are arranged into a barrel around the COT cylinder, at a radius of $\sim$140 cm. They are surrounded by the superconducting solenoid on the outside. The scintillating material is Bicron 408, which has a short rise time and a long (380 cm) attenuation length.

Particles passing through the scintillating material of the bars deposit energy causing small flashes of visible light. This light is detected by photomultiplier (PMT) tubes which are attached at both ends of each bar and provide time and pulse height measurements. The signal from the photomultiplier tube is processed by a pre-amplifier circuit mounted directly onto the tube. The readout electronics perform both time and amplitude digitization of the signal. The TDC information is a digitization of the time when the signal pulse reaches a fixed discriminator threshold. This time depends on the amplitude of the pulse, since a large pulse crosses the threshold earlier (time walk).

The digitization of the pulse amplitude is needed to correct for this effect. After correcting for time walk effects, the timing resolution of the TOF system is about 110 ps for particles crossing the bar exactly in front of one of the photomultiplier tubes. The timing resolution varies with displacement from the photomultiplier tube. Large pulses give better timing resolution, as light attenuates while traveling through the scintillator material. Therefore, particles passing through the bar near the photomultiplier tube have better timing resolution than those which are farther away. A more detailed description can be found in [115] .

3.2.4. The Solenoid

The tracking and the TOF systems are enclosed in a superconducting solenoid which provides a nearly uniform magnetic field of up to 1.4 T along the detector axis, over a cylindrical fiducial volume 3.5 m long and 2.8 m in diameter.

The coil itself is 4.8 m long and $\approx$25 cm thick, with an inner radius of 1.4 m. It is built of an aluminum-stabilized Nb Ti superconductor, able to withstand currents up to 5000 A, and operating at liquid helium temperature. During most of Run I, the magnet operated at 4650 A, corresponding to a current density of 1115 $\mathrm{A/m}$ and a central field of 1.4 T.

Although the design lifetime of the solenoid was only ten years, it is possible to reuse the magnet during Run II. The cool-down procedures that were used during Run I limited mechanical stress to the coil, avoiding fatigue damage.

3.2.5. Calorimeters

The main effort of the Run II upgrade of the CDF II calorimeter system dealt with upgrading the electronics to handle the faster bunch crossings. The active detector parts were taken over from Run I without modification. We will describe shortly this system in the next subsections. A more detailed description can be found in the CDF II Technical Design Report [106] .

Overview

The basic structure of the CDF calorimeters is based on scintillating sampling. That is, the detector after the absorbing material is a scintillating sheet, guided into a fiber, where the light produced from the incoming particles is passed through a wavelength shifting fiber to a photomultiplier tube, and then on to an amplifier. The calorimeter is divided into separate electromagnetic (large number of radiation lengths X_0 and small number of interaction lengths λ for photon and electron identification and energy measurement) and hadronic (large number of interaction lengths

for hadron energy measurement) sections.

The entire calorimeter is segmented into "projective towers", whose geometry is summarized in Table 3.3. This means that it is segmented in η and ϕ "towers" that point to the interaction region. The coverage of the calorimetry system is 2π in ϕ and $|\eta| < 3.6$ in pseudorapidity.

η range	$\Delta\phi$	$\Delta\eta$
0 - 1.1 (1.2 had)	15°	0.1
1.1 (1.2 had) - 1.8	7.5°	0.1
1.8 - 2.1	7.5°	0.16
2.1 - 3.6	15°	0.2 - 0.6

Table 3.3: Calorimeter segmentation.

The calorimeter system is divided into three regions: central, plug and forward. Corresponding to these regions, the subsystems will have one of the letters C, P and F in their acronym. Each calorimeter tower consists of an electromagnetic shower counter followed by a hadron calorimeter. This allows for comparison of the electromagnetic and hadronic energies deposited in each tower, and therefore separation of electrons and photons from hadrons.

There are three subdetectors for the electromagnetic calorimeter: CEM, PEM and FEM. These correspond to the central, plug and forward regions of $|\eta|$, respectively. The hadron calorimeters in the central region are the central (CHA) and the endwall (WHA). The plug and forward regions are covered by the PHA and FHA calorimeters, respectively.

The central region of the detector is covered by the Central Electromagnetic (CEM) [116] and Central Hadronic (CHA) [117] calorimeters, in the pseudorapidity ranges $|\eta| <$1.1 and $|\eta| <$0.9, respectively. In the forward region, the plug electromagnetic (PEM) [118] and hadronic (PHA) calorimeters cover the regions 1.1$< |\eta| <$3.6 and 1.3$< |\eta| <$3.6 respectively. The Wall Hadronic Calorimeter (WHA) [117] fills the gap between the CHA and the PHA in the pseudorapidity range 0.7$< |\eta| <$1.3.

The pseudorapidity coverage, resolutions, thickness and absorber material for the different electromagnetic and hadron calorimeters are given in Table 3.4. The details of each calorimeter are based on the specific physics needs and are discussed below.

Central Calorimeter

Apart from the electronics, the central calorimeter in the CDF II detector is the same used during Run I.

System	η coverage	Energy Resolution (%)	Thickness	Absorber
CEM	$\lvert\eta\rvert < 1.1$	$13.5/\sqrt{E_T} \oplus 2$	$18X_0$	3.18 mm lead
PEM	$1.1 < \lvert\eta\rvert < 3.6$	$16/\sqrt{E_T} \oplus 1$	$21X_0$	4.5 mm lead
CHA	$\lvert\eta\rvert < 0.9$	$50/\sqrt{E_T} \oplus 3$	$4.5\lambda_0$	2.5 cm steel
WHA	$0.7 < \lvert\eta\rvert < 1.3$	$75/\sqrt{E_T} \oplus 4$	$4.5\lambda_0$	5 cm steel
PHA	$1.3 < \lvert\eta\rvert < 3.6$	$80/\sqrt{E_T} \oplus 5$	$7.0\lambda_0$	5.08 cm steel

Table 3.4: Pseudorapidity coverage, energy resolution and thickness for the different calorimeter subdetectors of the CDF II Detector. The $\oplus$ symbol means that the constant term is added in quadrature to the resolution. λ_0 signifies interaction lengths and X_0 radiation lengths.

The CEM is a sampling device made of 31.5 mm thick layers of polystyrene scintillator, alternated with 3.18 mm thick layers of aluminum-clad lead. In order to maintain a constant number of radiation lengths as a function of θ, some lead layers are replaced by acrylic (Plexiglas), so that the actual number of absorber layers varies from 30 near the center to 20 at $\eta \approx$1.1. The CEM is divided into four arches (Noert-West, South-West, North-East and South-East) made of identical 15° modules, each of them being segmented into 10 projective towers. Thus each tower covers a solid angle of 0.1 by 15° in $\eta \times \phi$ space. The blue light emitted by the scintillators is collected on each side of the two towers by acrylic wavelength shifters that convert it to green light and guide the light toward two photomultipliers (Hamamatsu R580) outside the CHA (see Fig. 3.10). The two most forward towers of one of the CEM and CHA modules are not instrumented (the so called "chimney"), in order to provide access for cryogenics to the solenoid. Based on test beam data, the CEM energy resolution for an electron going through the center of a tower is found to be $\frac{13.7\%}{\sqrt{E}} \oplus 2\%$.

The Central EM Max Detector (CES) [116] is a strip chamber designed to provide a measurement of charged tracks very close to the calorimeter, with very little material in between. This is done to distinguish electrons from photons, which otherwise look very similar in the detector. They are located between the 8^{th} lead layer and the 9^{th} scintillator layer (counting outward), which is the expected position of shower maximum ($\approx 6X_0$, including tracking and solenoid material). In each CEM module, a CES module is a multi-wire proportional chamber with 64 anode wires parallel to the beam axis, spaced 0.73 cm apart and split at $|z|$ = 121 cm. The spatial resolution achieved is $\approx$2 mm.

The CEM is also equipped with a pre-shower detector (CPR) , useful in discriminating between hadrons and photons/electrons. The CPR is a set of multi-wire proportional chambers with wires

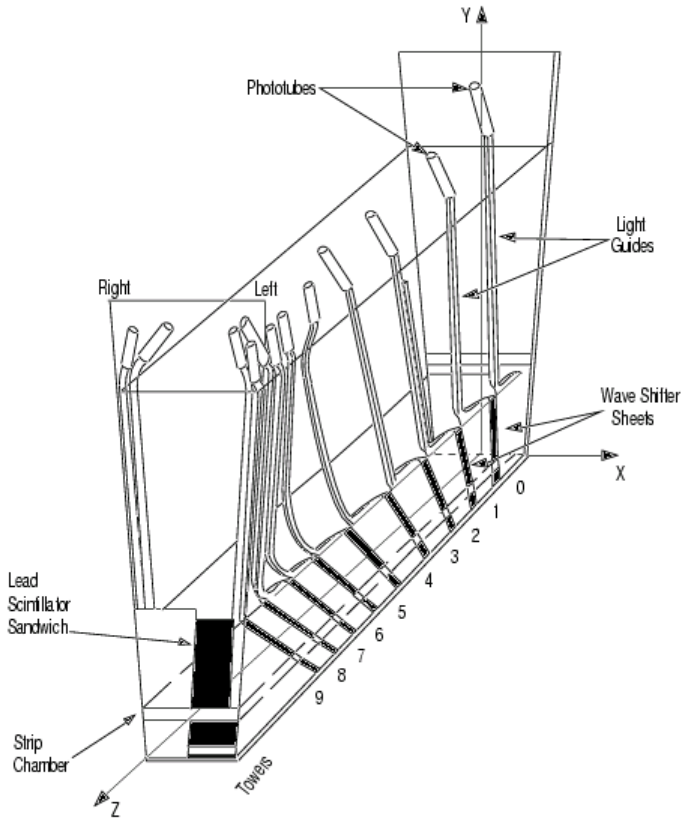


Figure 3.10: Wedge of the Central Electromagnetic Calorimeter.

parallel to the beam providing transverse measurements and strip cathodes providing z information, with a resolution of the order of a few millimeters.

The CHA is a sampling hadronic calorimeter surrounding the CEM, following the same segmentation (0.1 by 15° in $\eta \times \phi$). The WHA extends the CHA coverage and uses the same technology as the CHA. Altogether, a wedge contains 12 towers, 6 of which are fully in the CHA, 3 in the WHA and 3 are shared between the two. The number of interaction lengths is constant through the entire range of pseudorapidity and is equal to 4.5. The CHA is made of 32 layers of 2.5 cm thick steel absorber and 1.0 cm thick scintillator. The WHA is made of 15 layers of 5.0 cm thick steel absorber and 1.0 cm thick scintillator. Two PMT's per tower are linked to the scintillators by a wavelength shifter and a light guide. The CHA and WHA single pion energy resolutions are $\frac{50\%}{\sqrt{E}} \oplus 3\%$ and $\frac{75\%}{\sqrt{E}} \oplus 4\%$, respectively.

Plug Calorimeter

The Plug Calorimeter, shown in Fig. 3.11, covers the η region between 1.1 and 3.64, corresponding to polar angles between 37° and 3°. It was designed and built to replace the CDF I forward calorimeters, and to cope with the Run II requirements (higher luminosity and 132 ns bunch spacing).

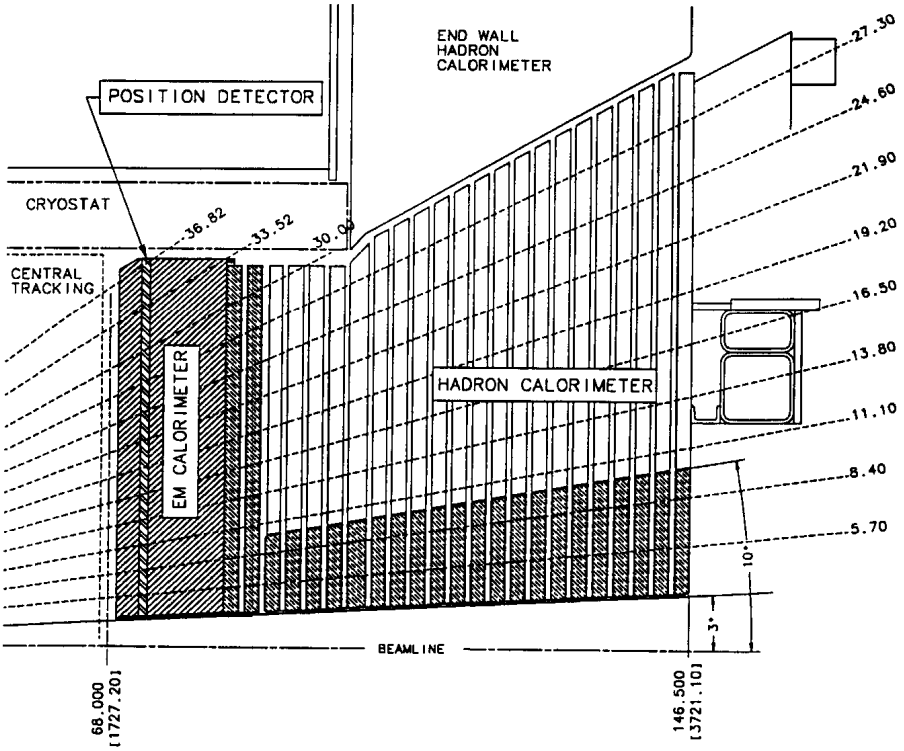


Figure 3.11: View of the Plug Calorimeter (PEM and PHA).

The Plug Calorimeter consists of an electromagnetic (PEM) and hadronic (PHA) calorimeter with the same projection segmentation. Figure 3.12 shows the segmentation pattern of a 15° module: towers cover an azimuthal angle of 7.5° down to η = 2.22 and 15° further; similarly, the segmentation in η becomes coarser as one moves closer to the beam. Figure 3.12 also shows how towers are combined for the purpose of being used by the trigger system.

The PEM is made of 22 layers of 4.5 mm lead and 4 mm thick scintillator tiles. Each scintillator tile is read by a single PMT. In front of the 22 sampling layers is a 1 cm thick scintillator tile read out by a multi-anode photomultiplier (MAPMT) which is used as a pre-shower detector. The PEM energy resolution is $\frac{16\%}{\sqrt{E}} \oplus 1\%$.

As in the Central Calorimeter, a shower maximum detector (PES) is also embedded in the PEM.

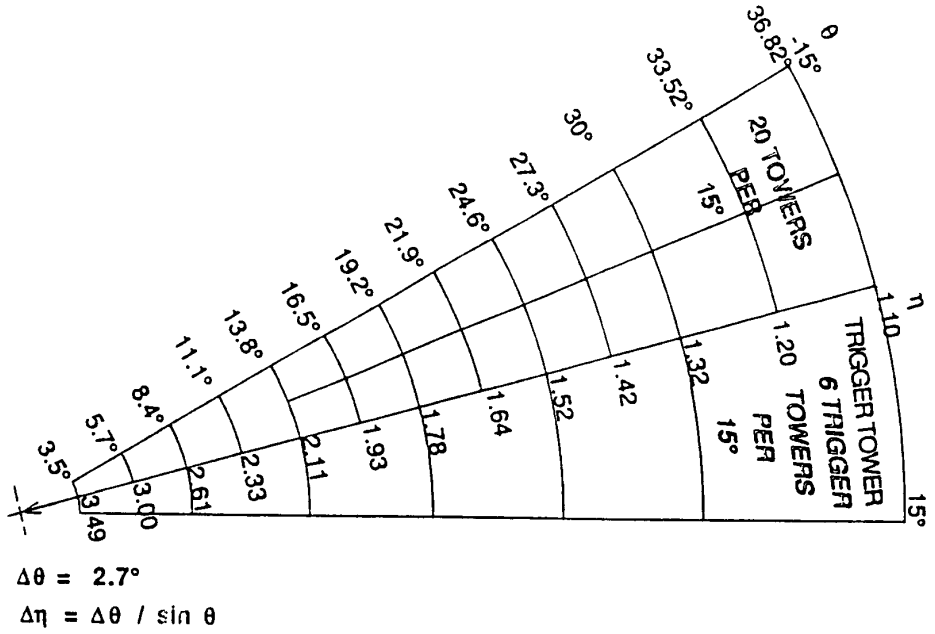


Figure 3.12: Segmentation of the Plug Calorimeter (PEM and PHA).

It is made of two sets of scintillating strips that provide precise 2D shower position measurement (resolution $\approx 1\ \mathrm{mm}$).

The PHA is made of 23 layers of 5.08 cm thick steel absorber and 6 mm thick scintillator. Its resolution is $\frac{80\%}{\sqrt{E}} \oplus 5\%$.

3.2.6. Muon Systems

Muons are particles which interact with matter only by ionization. For energies relevant to this experiment, they do not cause showers in the electromagnetic or hadronic calorimeters. As a result, if a muon is created in the collision and has enough momentum, it will pass through the calorimeter with minimal interaction with the material inside. Therefore, the calorimeter can be considered as a filter which retains particles that shower when interacting with matter and muons, which do not. Muon detection systems are therefore placed radially outside the calorimeters, being the outermost component of CDF.

The muon detectors at CDF make use of single wire drift chambers as well as scintillator counters for fast timing. The various subsystems are the Central Muon Detector (CMU), the Central Muon uPgrade Detector (CMP), the Central Scintillator uPgrade (CSP), the Central Muon eXtension Detector (CMX), the Central Scintillator eXtension (CSX), the Toroid Scintillator Upgrade

(TSU), the Barrel Muon Upgrade (BMU) and the Barrel Scintillator Upgrade (BSU). The CMU, CMP and CSP cover $|\eta| < 0.6$, the CMX and CSX cover $0.6 < |\eta| < 1.0$ and the TSU, BMU and BSU cover $1.0 < |\eta| < 1.4$. Figure 3.13 shows the coverage of each subdetector in the $\eta \times \phi$ plane.

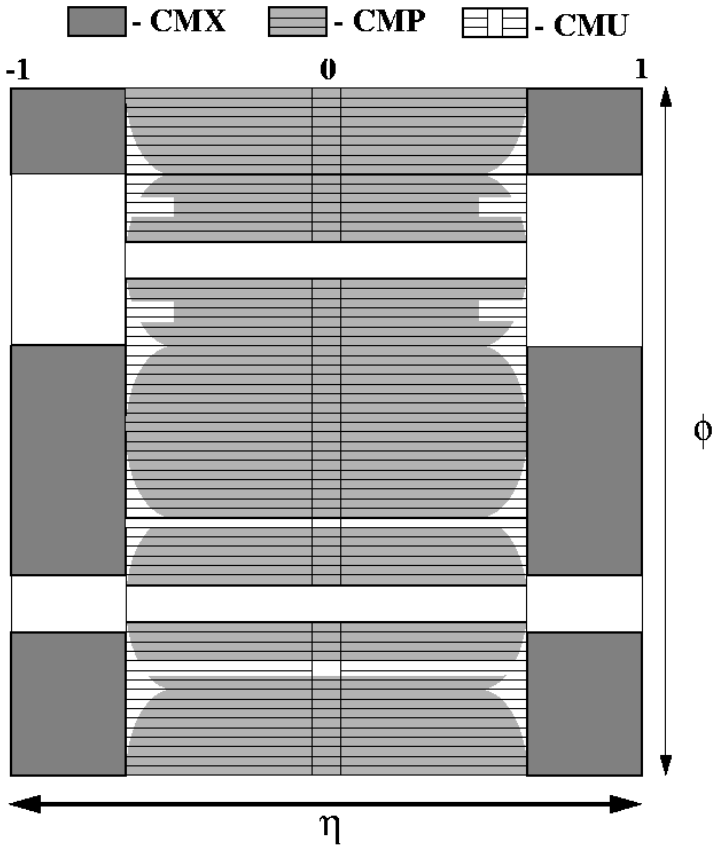


Figure 3.13: Coverage (in the $\eta \times \phi$ plane) of the upgraded CDF muon system.

Table 3.5 summarizes the information on the muon subsystem.

The first muon system built at CDF, the CMU [119], is placed just outside the CHA.mIt provides roughly 5.5 interaction lengths for pions, absorbing more than 99% of the outgoing charged hadrons. The p_T threshold of the CMU is 1.4 GeV/c. It is cylindrical in geometry with a radius of 350 cm, arranged into 24 12.6° wedges, which means that there is a gap of 2.4° between adjacent wedges. Each wedge contains three modules (stacks) with four layers of four rectangular drift cells. The cells have 50 $\mu\mathrm{m}$ sense wires at the center of the cell, parallel to the z direction. The system is filled with an Argon-Ethane gas mixture and alcohol (49.5:49.5:1) as the COT.

A second set of chambers, the CMP, is situated outside an additional layer of 60 cm thick steel to act as an absorber, which is 3.5 additional interaction lengths (for a total of $9.0\lambda_0$). The

	CMU	CMP/CSP	CMX/CSX
η coverage	0 - 0.6	0 - 0.6	0.6 - 1.0
Min p_T [GeV/c]	1.4	2.2	1.4
Drift Tubes			
Thickness [cm]	2.68	2.5	2.5
Width [cm]	6.35	15	15
Length [cm]	226	640	180
Max drift time [μs]	0.8	1.4	1.4
Scintillators			
Thickness [cm]	N/A	2.5	1.5
Width [cm]	N/A	30	30 - 40
Length [cm]	N/A	320	180

Table 3.5: Parameters of the Muon Detectors at CDF.

p_T threshold of the CMP is 2.2 GeV/c. It is rectangular in geometry, consisting in four layers of single-wire drift cells, staggered by one half cell per layer.

On the other surface of the CMP lies the CSP [120], a single layer of rectangular scintillator tiles, with a waveguide to move the scintillated light into a PMT. This provides a fast detection mechanism used in triggering muons.

The CMX is located on either side of the detector straddling the beamline. It is a conical geometry of drift tubes with drift chambers, similar to the CMP, and scintillators on both sides. The CSX is another scintillator array similar to the CSP. The CMX cover 360° in ϕ. The segmentation is in 15° wedges in azimuthal angle. Each wedge consists of eight layers of rectangular tubes in the radial direction, also offset to provide better resolution.

Using the timing information from the drift cells of the muon systems, short tracks (called "stubs") are reconstructed. Tracks reconstructed in the COT are extrapolated to the muon systems. Based on the projected track trajectory in the muon system, the estimated errors on the tracking parameters and the position of the muon stub, a χ^2 value of the track-stub match is computed. To ensure good quality of muons, an upper limit is placed on the value of χ^2_ϕ, the χ^2 of the track-stub match in the ϕ coordinate.

3.2.7. The Cherenkov Luminosity Counter

The Cherenkov Luminosity Counter (CLC) [121, 122] was designed for the Tevatron Run II in order to achieve a precision measurement of the instantaneous luminosity up to $\approx 2{\cdot}10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ and to cope with the 132 ns bunch-spacing that was originally envisioned. Since luminosity measurement is critical to the cross section measurement presented in this document, it is explained here in some detail.

The detector, located in the 3° gap between the plug calorimeter and the beam pipe as shown in Fig. 3.14, is made of two identical CLC modules installed at small angles, inside the Plug Calorimeter, on each side of the interaction point. Figure 3.15 shows two views of such a module. Each module is composed of 48 thin, long, conical, gaseous Cherenkov counters pointing toward the interaction point and covering the pseudorapidity range $3.7 < |\eta| < 4.7$. The counters are arranged around the beam pipe in three concentric layers, with 16 counters each. The cones in the two outer layers are about $180\ \mathrm{cm}$ long and the inner layer counters (closer to the beam pipe) have a length of $110\ \mathrm{cm}$; their diameter varies from 2 to 6 cm. At the widest end of each one (the furthest away from the interaction point), a conical mirror collects the Cherenkov light into $2.5\ \mathrm{cm}$ diameter photomultiplier tubes (Hamamatsu R5800Q). The tubes have a concave-convex, $1\ \mathrm{mm}$ thick, quartz window for efficient collection of the ultraviolet part of Cherenkov spectra and operate at a gain of $2 \cdot 10^6$. The modules are filled with isobutane at atmospheric pressure; it is however possible to increase the pressure up to 2 atm, in order to increase the yield of Cherenkov light. Isobutane was chosen because of its large refractive index at atmospheric pressure and its good transparency to ultraviolet light. The Cherenkov angle is 3.1° and the momentum threshold for light emission is $9.3\ \mathrm{MeV/c}$ for electrons and $2.6\ \mathrm{GeV}/c$ for pions.

Because of the narrow shape and the orientation of the cones, particles produced by $p\bar{p}$ interactions close to the center of the detector are likely to go through a large portion of the CLC, producing an important light yield (several hundred photo-electrons), while particles from the beam halo or from secondary interactions traverse the detector at large angle, and have lower energy, hence producing a much smaller light signal. Thus the background is easily rejected by requiring a certain minimal light yield threshold in each channel; the number of particles is measured from the total yield in the module. Thanks to the CLC's excellent time resolution (less than 100 ps), it is also possible to select hits from prompt particles by requiring time coincidence between hits in the two different modules.

At hadron collider experiments the beam luminosity can be expressed as a function of the number of hits per bunch-crossing as follows:

$$L = \frac{f_{bc}}{\sigma_{in} \cdot \epsilon} \cdot \mu\,, \tag{3.6}$$

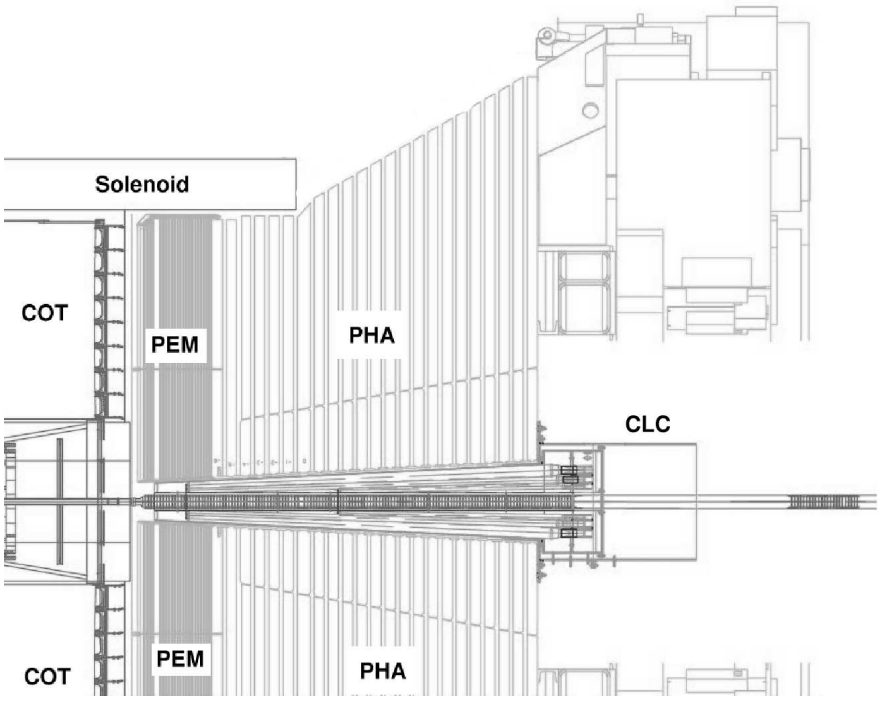


Figure 3.14: Location of the CDF Cherenkov Luminosity Counter in the 3° gap between the plug calorimeter and the beam pipe.

where L is the instantaneous luminosity, f_{bc} is the rate of bunch-crossings in the Tevatron, σ_{in} is the inelastic scattering cross section, ϵ is the acceptance times efficiency of the CLC for inelastic scattering events and μ is the (measured) average number of interactions per bunch-crossing.

In Eq. 3.6, f_{bc} and ϵ are known and the total inelastic cross section was measured in several experiments. CDF Run I and E811 measurements were combined, giving $\sigma_{in} = 60.4 \pm 2.3\ \mathrm{mb}$ at 1.8 TeV, which can be extrapolated to $61.7 \pm 2.4\ \mathrm{mb}$ at 1.96 TeV. Therefore, we just need to know the number of hits per bunch-crossing in order to calculate the luminosity. And this is what the CLC was designed for by measuring the number of particles and their arrival time in each bunch-crossing. A precision of 5.9% [123] on the luminosity is achieved with the CLC; 4.4% comes from the CLC acceptance and operation of the luminosity monitor and 4% from the calculation of the inelastic cross section. The luminosity measured by the CLC is used to monitor the Tevatron's performance.

3.2.8. Trigger

Triggering systems are necessary because it is not physically possible to store information about every single $p\bar{p}$ collision. Collisions happen roughly at a rate of 2.5 MHz, and the readout of the full detector produces an event roughly the size of 250 kB. There is no medium available

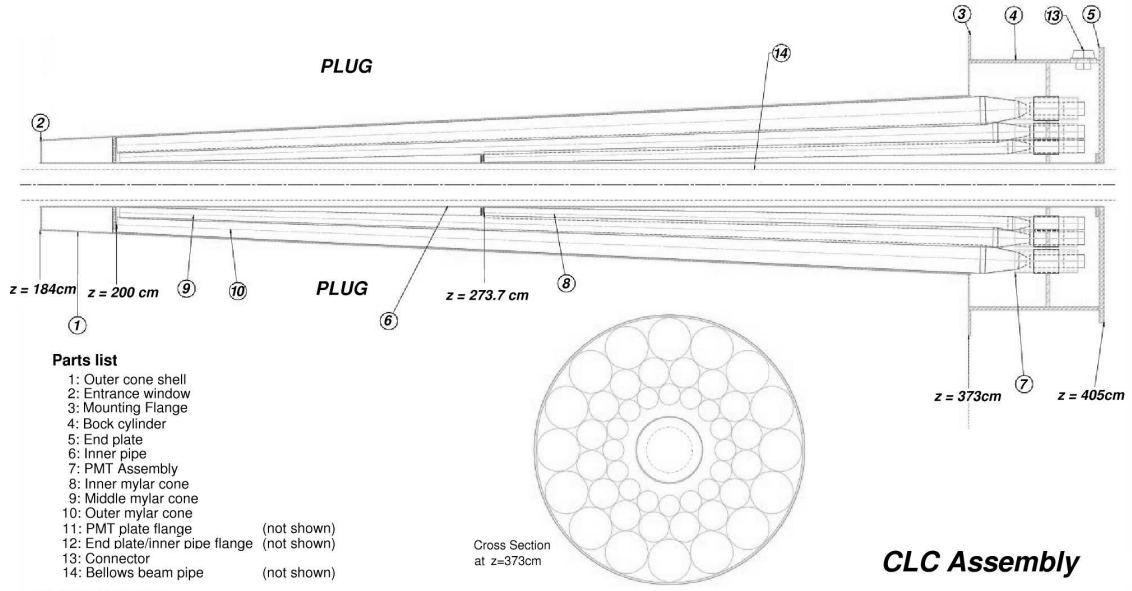


Figure 3.15: The CLC assembly diagram. The cross section view at z=373 cm is also shown.

which is capable of recording data this quickly, nor would it be practical to analyze all these data later on. The trigger system is a pre-filter, which reduces data rates and volumes to manageable levels, according to all possible or foreseen physics prescriptions.

The CDF II triggering system is designed based on three conditions. The first condition is that the trigger has to be dead-timeless. This means that the trigger system has to be quick enough to make a decision for every single event, before the next event occurs. The second condition is imposed by the Tevatron upgrade for Run II, and it is the time between collisions, 132 ns. The last condition is that the data logging system can write about 75 events per second to tape, because of limited resources. In short, the trigger has to be fast enough to analyze every collision, and it has to figure out which 75 of 2.5 million events it should save in a given second. This is achieved by staging trigger decisions in three levels, as shown in Fig. 3.16. This new architecture is fully capable of withstanding a 132 ns bunch separation, while keeping dead time as short as possible.

Each level of the trigger is given a certain amount of time to reach a decision about accepting or rejecting an event. By increasing the time allowed for triggering at different levels of the trigger, the complexity of reconstruction tasks can be increased at every level. At the first level of the trigger, only very rough and quick pattern recognition and filtering algorithms are used. In order to do this in time, the Level 1 and Level 2 triggering mechanisms are implemented with custom electronics. The third level of the trigger is implemented with a PC farm with about 300 CPUs.

The delay necessary to make a trigger decision is achieved by storing detector readout infor-

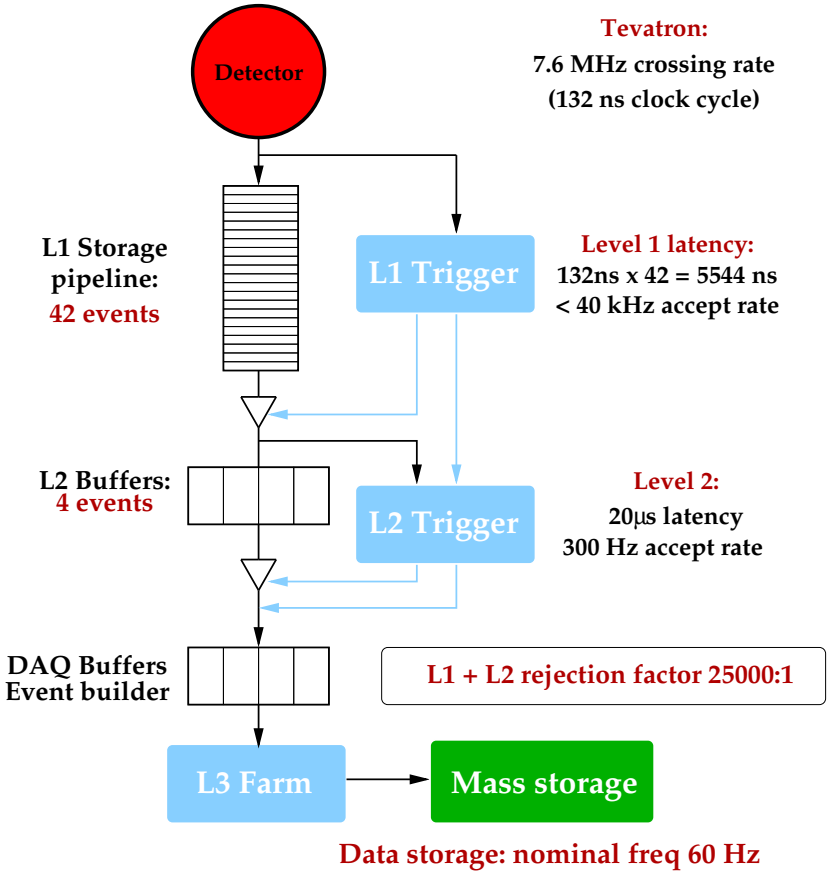


Figure 3.16: Diagram of the CDF II Detector trigger system.

mation in a storage pipeline.

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a trigger path. Requiring that an event be accepted through a well defined trigger path eliminates volunteer events. A volunteer event is an event which passed a higher level (L2, L3) trigger requirement but did not pass the preceding lower level (L1, L1/L2) trigger requirement. The CDF II trigger system implements about 100 trigger paths. An event will be accepted if it passes the requirements of any one of these paths.

Level 1 Trigger

At Level 1, for every Tevatron clock cycle, the event is moved up one slot in the pipeline. By the time it reaches the end of the pipeline, the trigger will have reached a decision whether to accept or reject this event. If the event is accepted, its information will be sent to the higher level of the trigger. Otherwise, the event is simply ignored.

The front-end electronics of all detectors is fitted with a synchronous pipeline, 42 events deep,

where the entire data regarding each event is stored for 5544 ns. Meanwhile, part of the data is examined in a first layer of dedicated, synchronous, highly parallel hardware processors:

- XFT [124] , the eXtremely Fast Tracker, which reconstructs tracks on the transverse plane of the COT to propagate them to the calorimeters and muon chambers;
- the Calorimeter Trigger, which detects electron and photon candidates, jets, total transverse energy, and missing transverse energy;
- the Muon Trigger, which matches XTRP (eXTRaPolation module) tracks [125] to stubs in the muon chambers.

Since the Level 1 buffer has 42 slots, the time allocated for making a trigger decision is about 5 μs. The rejection factor after Level 1 is about 150, so the Level 1 accept rate is below 40 kHz.

Level 2 Trigger

Events matching the requirements of the Level 1 are downloaded into one of four asynchronous event buffers, and further analyzed by a second set of hardware processors. This allows for 20 μs for the trigger decision. The Level 2 rejection factor is again around 150, and the accept rate is around 300 Hz.

The Level 2 is able to reconstruct calorimeter clusters, and to use the maximum shower detector information. A novelty in hadronic physics, it is also able to use the Silicon Vertex Detector: the Silicon Vertex Trigger (SVT) [126] uses XFT tracks as an input and tries to reconstruct tracks based on silicon hits in the neighborhood of an XFT track. This technique significantly reduces the number of candidate hits, hence allows very fast reconstruction, while achieving a resolution comparable with the full tracking reconstruction. The SVT is able to identify tracks that are significantly displaced from the beam location, hence selecting heavy flavor enriched events. Figure 3.17 shows what information is available to Level 1 and Level 2.

Level 3 Trigger

Finally, after being accepted by the Level 2, the entire event data is read out and loaded into a Linux PC farm, where the event is fully reconstructed in software. The Level 3 reconstruction program is almost fully written in C++, using object-oriented techniques. After an event is reconstructed, it is sent to an event counter, where its characteristics are histogrammed; if the event passes the Level 3 cuts, it is also permanently stored to tape.

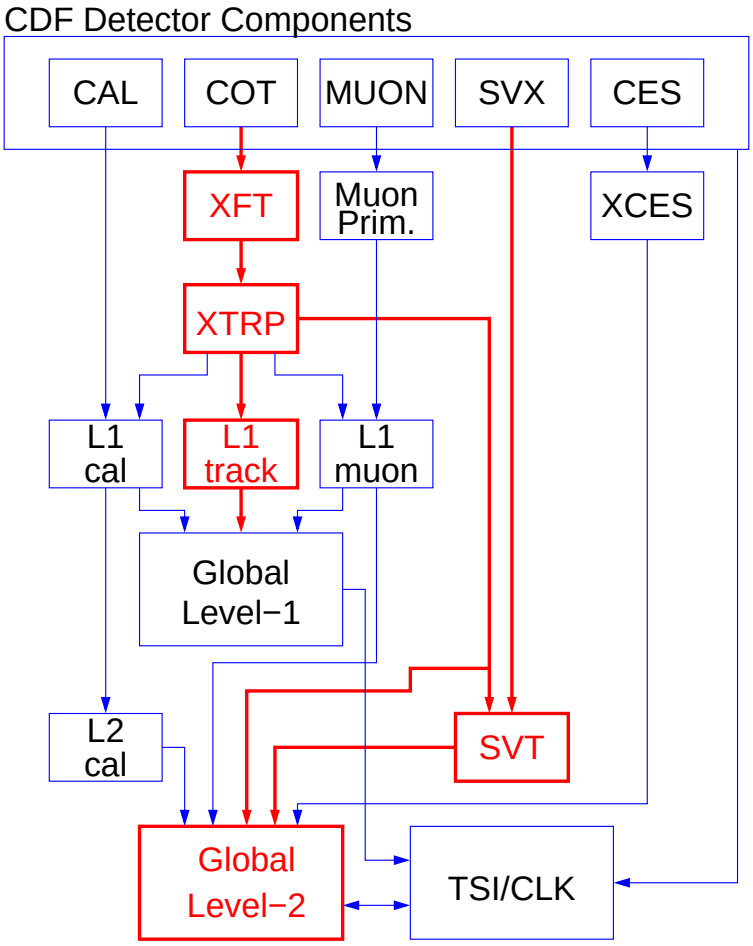


Figure 3.17: Block diagram of the Level 1 and Level 2 trigger paths.

Every CPU in the farm provides a processing slot for one event. With roughly 300 CPUs, and an input rate of $\sim$ 300 Hz, this allocates approximately 1 second to do event reconstruction and reach a trigger decision. As a result, nearly offline quality event reconstruction is available at the third level of triggering. The Level 3 rejection rate is about 4, resulting in about 75 events/sec being accepted by the Level 3 trigger and written to tape.

Fig. 3.18 shows the implementation of the Level 3 farm. The detector readout from the Level 2 buffers is received via an Asynchronous Transfer Mode (ATM) switch and distributed to 16 "converter" node PCs , shown in Fig. 3.18 in light blue. The main task of these nodes is to assemble all the pieces of the same event as they are delivered from different subdetector systems through the ATM switch. The event is then passed via an Ethernet connection to a "processor" node, of which there are about 150 in the farm and are shown in Fig. 3.18. Each processor node is a separate dual-processor PC. Each of the two CPUs on the node process a single event at a time. The Level 3 decision is based on near-final quality reconstruction performed by a "filter" executable. If the executable decides to accept an event, it is then passed to the "output" nodes of

the farm. These nodes send the event onward to the Consumer Server / Data Logger (CSL) system for storage first on disk, and later on tape.

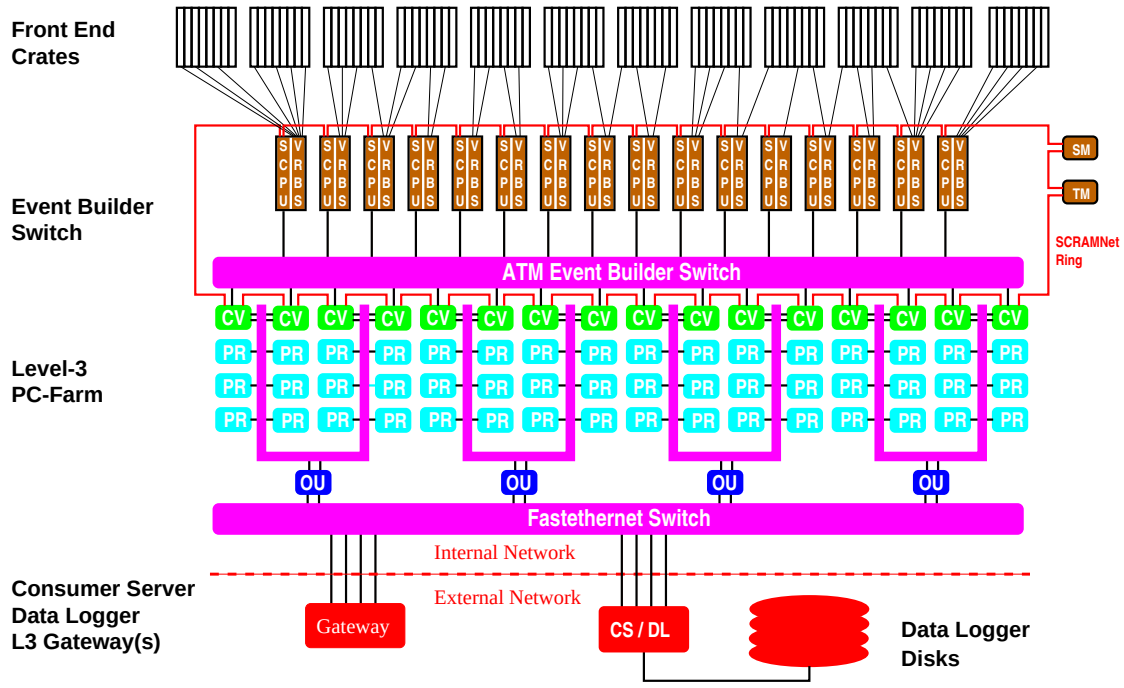


Figure 3.18: Principle of Event Building and Level 3 Filtering. Data from the front end crates is prepared by Scanner CPUs (SCPU) and fed into the ATM switch. On the other side of the switch, converter nodes (CV) assemble events and pass them to processor nodes (PR). Accepted events are passed to output nodes (OU) which send them to the Consumer Server and Data Logging systems (CS/DL).

Online Monitoring

The CDF detector consists of many detector subsystems and runs a high rate large bandwidth data transfer environment. To take data with high efficiency and high quality, it is necessary to quickly spot problems with one of these subdetectors in real time. Multiple event monitor programs are attached to the DAQ system [127, 128] . The online monitoring programs are called Consumers, where a Consumer is defined as a process which receives events from Consumer Server Logger (CSL) in real time. CSL sends the data to the computer center where they are written to tape and forwards copies of a subset of the data to the online monitoring programs. Figure 3.19 shows a schematic view of the CDF online monitoring system (Consumer Framework). The task of the Consumers is to analyze and monitor the event data and to make histograms and tables. These results could be viewed by the display browser via a server in real time. Results of the monitor are also stored as data files periodically during a run, and also archived systematically.

The display browser provides a GUI to view the online monitored results, while also providing some basic utilities to do comparisons with previously stored results. By separating the two tasks of monitoring and displaying, we remove CPU bound associated with displaying graphics from the machine which runs the consumers. During the data taking, multiple consumer processes run in parallel, receiving event data with the desired trigger types from the CSL. Communication between consumers and run control, which controls the overall CDF DAQ system, is handled by the error Receiver. Severe errors detected by a consumer monitor program are forwarded to run control to take necessary actions. The state manager watches the state of consumers.

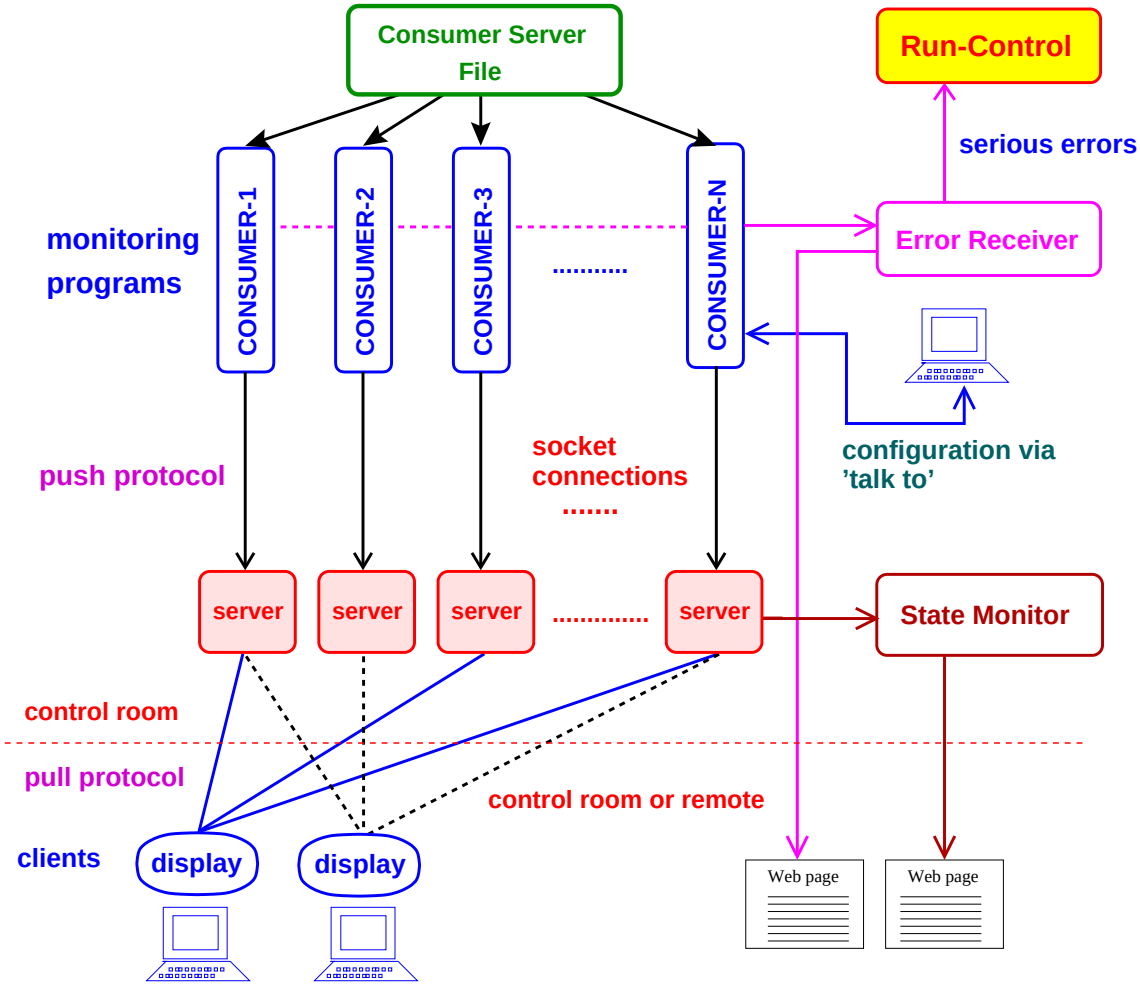


Figure 3.19: Design of the CDF online consumer framework.

Chapter 4

EVENT SIMULATION AND RECONSTRUCTION

The understanding of efficiencies, acceptances and kinematic properties of collision events is essential for data analyses and thus requires deep knowledge of the physics processes and the detector response. To evaluate measured data, it is therefore useful to simulate all physics processes expected to contribute to the corresponding data sample and to emulate the detector response. Both measured and simulated objects are subject to the same event reconstruction algorithms, which allow for a direct comparison between simulated processes and observed data.

4.1. Event Simulation

The hard interaction of the incoming beams results in the production of up to hundreds of outgoing particles. Unfortunately, a full theoretical quantum-mechanical treatment is unfeasible due to two main reasons: first of all, the number of particles involved gives rise to a tremendous number of interfering contributions that grows factorially with the number of particles. Furthermore, perturbation theory is not able to account for the transition of partons to hadrons. This failure of perturbation theory necessitates other strategies to obtain a detailed description of the production of multiple particles. This is realized by Monte Carlo (MC) event generators which randomly produce collision data according to the probability density of phase space and the matrix element of a given process.

Any theoretical model describing an elementary process starts from the knowledge of its cross section and must both contain a way to compute or to estimate the effects of higher-order perturba-

tion theory and a way to describe hadronization effects. Including finite higher-order corrections through the exact computation of a given number of emissions is performed by matrix element event generators. The common approach to estimate the effects due to emissions at all orders in perturbation theory is done by the parton showering technique.

4.1.1. Monte Carlo Generators

Effects of higher-order corrections in perturbation theory can be taken into account by exact computation of the result of a given and usually small number of emissions. This can be realized by considering only those diagrams corresponding to the emission of real particles. Basically, the number of emissions coincides with the perturbative order in α_s. This approach forms the core of the parton-level generators, which compute tree-level matrix elements for a fixed number of partons in the final state.

A Monte Carlo generator uses an "unweighting" method to simulate the relative rate of different event kinematics.First, it creates a large number of events with randomly assigned kinematic properties. It calculates a weight for each event based on the differential cross section for the event's kinematic properties. Then it converts each weight to a probability, taking the highest weight to be unity. This gives the relative contribution of each region of phase space. The generator then examines each event again, choosing a random number between 0 and 1 for each event and keeping only events for which the random number is less than the probability for that event. This results in a set of discrete, unit-weight events whose kinematics, for a large number of events, properly reflect the differential cross section of the process.

PYTHIA

The most used event generator at CDF is called PYTHIA [129, 130], This program contains showering routines as well as an event generator and it takes little effort to pass events between the two. The event generator in PYTHIA can handle simple Feynman diagrams; however, it does not include spin correlations of polarized top quarks, and it uses a parton shower approximation to account for the effects of initial- and final-state radiation which does not include color information. When color and polarization effect are not significant, however, it performs very well.

MADEVENT

MADEVENT [131] is a Monte Carlo generator that can calculate arbitrary tree-level diagrams with full color and spin polarization information included. In this analysis it is used for the simu-

lation of the signal diagrams in which the polarization of the top quark is an important part of the event kinematics. MADEVENT is powered by the matrix element generator MADGRAPH [132]. Given a standard model process, MadGraph automatically generates the amplitudes for all relevant sub-processes and produces the mappings for the integration over the phase space. This process dependent information is passed to MADEVENT and a stand-alone code is produced that allows the user to calculate cross sections and to obtain unweighted events. Once the events have been generated, they may be passed to showering MC programs.

ALPGEN

Processes with an electroweak boson and radiated gluons are difficult to deal with because of the large amount of radiation they produce. The showering approximation used by PYTHIA, being based only on the tree-level diagram, does not include effects of color flow. However, a full calculation of the matrix elements involved is difficult because the number of distinct diagrams grows as the factorial of the number of jets.

ALPGEN [133] is a Monte Carlo generator designed specifically for processes whose final state contains an electroweak boson and several radiated quarks and gluons, a major background in this analysis. ALPGEN calculates the matrix elements for processes with gluon radiation and passes the color information to the showering algorithm. This should give a more accurate modeling of the kinematics of the process than PYTHIA's showering approximation, since it includes proper matrix element calculations of the event. ALPGEN also calculates the leading-order cross section of each interaction it generates, which is useful for combining different processes.

4.1.2. Parton Showering and Hadronization

All events, regardless of how they were generated, are passed to PYTHIA for parton showering [130] and hadronization. The showering procedure generates initial- and final-state gluon radiation for each event and allows them to decay to quark pairs, increasing the number of particles in the final state of the event. More particles may be added from effects of beam remnants or multiple interactions. This gives the final set of particles that are passed to the hadronization routine.

Since the hadronization of quarks and gluons, which describes the formation of jets, takes place at low Q^2 and large α_s, perturbation theory cannot be applied. The phenomenologic models, used to describe hadronization in the absence of any firm theoretical understanding are different for distinct Monte Carlo generators. PYTHIA performs its hadronization using the Lund color string model [134, 135]. Each pair of quarks is modeled as though connected by a relativistic string

which increases linearly in energy as separation increases. As the energy in the string increases, it becomes increasingly more probable that it will form a new quark-antiquark pair. These new particles can then be grouped with the original ones to form mesons and baryons.

Most of the particles resulting from hadronization are unstable, so PYTHIA causes them to decay into relatively stable particles (electrons and muons, protons and neutrons, pions and kaons) that can actually be detected. This step uses branching ratios and lifetimes measured in various experiments to calculate the final decay products. In this procedure, PYTHIA ignores spin information and uses a simplified algorithm for B mesons and tau leptons.

4.2. Detector Simulation

Once the final long-lived particles have been generated, it is important to determine how the detector will respond to them. This requires a full detector simulation which simulates the response of the different subcomponents of the detector, including resolution effects, inherent inefficiencies in the detector, and the behavior of the particles as they pass through passive material (such as cables or support structures) in the detector. When this is done, the Monte Carlo events can be put into a data structure identical to that obtained from collision data, thus allowing reconstruction algorithms to work exactly the same way on data and Monte Carlo events.

The modeling of the CDF detector response is based on a detailed simulation using the GEANT3 package [136]. The charge deposition in the silicon layers is calculated using a simple geometrical model based on the path length of the ionizing particle. The drift model used in the COT simulation is based on the GARFIELD package [137], a general drift chamber simulation program. To speed up the simulation, the charged particle ionization and drift properties in the COT are parameterized and tuned to data. The calorimeter simulation based on the shower development package GFLASH [138] was also tuned using test-beam data for electrons and high-p_T pions. A detailed description of the CDF II detector simulation can be found in reference [139].

4.3. Event Reconstruction

Once the detector data is obtained, either from Tevatron collisions or simulated MC events, it needs to be converted from the raw data in the detector to reconstructed physics quantities in order to be analyzed. First, information from subdetectors is combined to form high-level detector objects: tracks in the tracking detectors, stubs in the muon chambers, and clusters of energy towers in the calorimeters (see Chapter 3). Then these objects are analyzed to associate them with candidates of physical objects: electrons, muons, jets, or neutrinos. These can finally be used in a

physics analysis.

4.3.1. Primary Vertex Reconstruction

Accurate determination of the primary interaction point of the hard scattering is essential for any high precision analysis. Because the interaction region of the particle beams has a substantial volume, the knowledge of the position of the primary vertex affects the measured kinematic properties of the event objects. Primary vertices are reconstructed by fitting prompt tracks fulfilling certain quality requirements to a common vertex. Tracks contributing a large χ^2 to the fitted vertex are iteratively removed if the χ^2 exceeds a given threshold. The iteration stops either if no track fails the χ^2 cut or the number of tracks associated to the vertex falls below a minimum quantity. This gives the final position of the primary vertex.

4.3.2. Lepton Identification

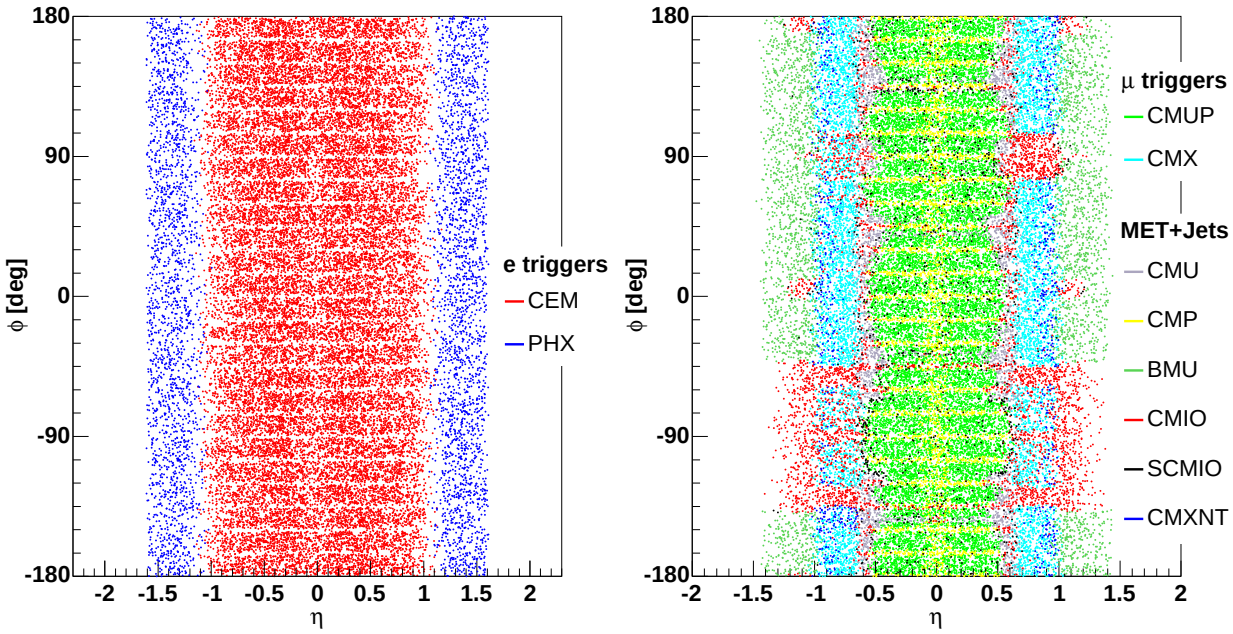
In CDF, an electron is identified as an isolated track matched to an electromagnetic calorimeter cluster. A muon requires an isolated track matched to a stub in a muon detector. Since a muon is a minimum ionizing particle (MIP), it is further required that the muon candidate object leaves only minimum ionizing energy as it passes through the calorimeter in order to reduce fake muons from energetic particles that make it through the calorimeter.

Good lepton identification is vital to purify the sample by removing fake leptons, making it easier to understand and estimate the background to the single top signal. One important variable for lepton identification is isolation. This quantity allows discrimination against leptons coming from jets and do not originate from hard scattering events. Isolation is defined by computing the transverse energy in a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$. The isolation is the ratio of the transverse energy that is not in the lepton cluster to the transverse energy in the cluster. A small number indicates that there is little extra activity in the calorimeter near the lepton, so it is unlikely to come from a jet. If this quantity is less than 0.1, the lepton is said to be isolated or tight; otherwise, it is non-isolated or loose. All leptons in this analysis are required to be tight, however, loose leptons are still used to remove dilepton events (see Section 5.2.2) and to model fake leptons in some muon categories (see Section 6.2.5).

Another common cut for all lepton types is the primary vertex requirement. The z coordinate of the reconstructed vertex of the track, z_0, must be less than 60 cm from the center of the detector. This requires the tracks to come from a hard scattering process and not elastic scattering or cosmic rays. This affects the luminosity calculation because some true hard scattering events occur outside this region. Measurements in data, using a trigger on events with hits in the CLC, show that the

efficiency of this cut in data is $\epsilon_{z0}^{\text{data}} = 96.4 \pm 0.4\%$.

The geometrical coverage of all considered types of leptons, described in the next subsections, is shown in Figure 4.1.



(a) (b)

Figure 4.1: Geometrical coverage of all considered types of electrons (a) and muons (b).

Central Electrons

The largest acceptance in the single top event selection comes from the Central Electromagnetic Calorimeter (CEM). These electron candidates consist of an electromagnetic cluster measured using the CEM and CES detectors matched to an extrapolated track in the COT. Identification cuts are summarized in table 4.1 using a variety of energy and track variables as follows:

- Geometry: The candidate must be fiducial to the central calorimeter region.
- E_T: The transverse energy of the calorimeter cluster. Since electrons emit bremsstrahlung photons easily, the energy, rather than the momentum, is typically used to characterize electrons.
- p_T: Transverse momentum of the associated track.

- E_{Had}/E_{EM}: This requires that the shower is mostly electromagnetic to suppress hadronic objects. The second term accounts for some leakage from the CEM when the electron is especially energetic

- E/p: The ratio of the cluster's energy and the track's momentum. This ratio should be nearly one for a true electron, so this requirement removes many fake electrons. For sufficiently high momentum tracks, fakes of this sort are unlikely, so the requirement is relaxed.

- $|\Delta z|$ and $Q \times \Delta x$: Δz and Δx are the signed distances between the extrapolated track and the center of the electromagnetic cluster measured with the CES in the r-z and r-ϕ planes respectively. Q is the measured charge of the particle. The asymmetry in the second requirement results from the trajectory of particles in the detector, if the sign of the charge and Δx are opposite, the particle traverses a larger part of the calorimeter in adjacent towers, which results in more radiation and a less precise final position. This means the cut must be looser in this case to preserve signal efficiency.

- χ^2_{strip}: The transverse profile of the electromagnetic cluster in the CES strip chamber must be consistent with an electron. The measured shape is compared, strip by strip, to the predicted profile, derived from theoretical parameterizations and test beam studies.

- L_{shr} characterizes the lateral sharing of the electromagnetic shower among calorimeter towers. The value of L_{shr} is the likelihood for the sum over towers of the difference between the expected and measured energy deposits divided by the root-mean-squared uncertainty.

- Track z_0: primary vertex requirement.

- Good COT track: The tracks in the COT must be of high quality. The track must have at least five hits in each of at least three axial superlayers and two stereo superlayers. This ensures that the track is cleanly reconstructed.

- Conversion: Photon conversions are an important background for electrons. A photon traveling through material can convert into an electron-positron pair, and the electron, though a true electron, is not meaningful to the analysis because it comes from a photon and not a hard scattering event. The conversion veto looks for a track with the opposite charge of the electron track that is separated from it by less than 2 mm in the $r - \phi$ plane at the point at which they are parallel. It also requires the cotangent of the polar angle between the two tracks to be less than 0.04. If such a track can be found, the electron is likely to come from a photon conversion and is not accepted.

Selection cuts
Fiducial in CEM
$E_T > 20$ GeV
$p_T > 10$ GeV/c
$E_{Had}/E_{EM} < 0.055 + 0.00045E$
$E/p < 2$ unless $p_T \geq 50$ GeV/c
$\lvert\Delta z\rvert < 3.0$ cm
-3.0 cm $< Q \times \Delta x < 1.5$
$\chi^2_{strip} < 10$
$L_{shr} < 0.2$
Track $\lvert z_0\rvert < 60.0$ cm
Good COT track
Isolation < 0.1
Not a conversion

Table 4.1: Selection requirements for central (CEM) electrons.

Plug Electrons

In addition to the central electrons, we also trigger on high η electrons called 'plug electrons' or 'phoenix electrons' (PHX). These plug electrons consist of a PEM and a PES calorimeter cluster matched to a silicon track using the Phoenix tracking algorithm. There is no trigger directly made on clusters in the plug calorimeter, instead, this sample relies on the presence of large missing transverse energy from W decay in the final state to trigger events (see Section 5.1.1). The lack of a tracking requirement for this trigger makes it less pure having more fake leptons than any of the other samples, thus requiring additional selection cuts to purify the sample. The selection requirements are summarized in Table 4.2 and discussed thereafter.

- Geometry: The candidate electron must be found in the plug calorimeter fiducial region. Outside the $1.2 < |\eta| < 2.0$ region the reconstruction is inefficient and has a large background from elastic proton-antiproton collision, candidates in that η region are then rejected.
- p_T: Although this trigger contains leptons with lower momenta, they have a large background from fake electrons and are difficult to simulate properly.
- E_{Had}/E_{EM}: This requires that the shower is mostly electromagnetic and suppresses hadronic objects.
- : χ^2_{PEM}: The shower profile in the PEM detector must be consistent to the predicted Lorentz

distribution.

- E_5/E_9: Clusters resulting from true electrons tend to have a narrow energy distribution. This can be seen in the PES shower maximum detector, where the energy of the five strips at the center of the cluster (E_5) is compared to the energy of all nine strips in the cluster (E_9) for both layers of the PES. For an electron, most of the energy will be contained near the center of the cluster, and this ratio will be large.

- $\Delta R(PEM, PES)$: The distance in the r-ϕ plane between the position of the reconstructed cluster in the PEM calorimeter and the PES shower maximum detector. This removes a background from poorly reconstructed clusters that might otherwise fake an electron.

- N_{hits}^{Si}: The number of hits in the silicon detector of the associated track. This requirement improves the quality of the track and reduces the background from poorly reconstructed tracks.

Selection cuts
Fiducial in PEM
$1.2 < \|\eta\| < 2.0$
$p_T > 20$ GeV/c
$E_{Had}/E_{EM} < 0.05$
$\chi^2_{PEM} < 10.0$
$E_5/E_9 < 0.65$ for both layers
$\Delta R(PEM, PES) < 3.0$ cm
$N_{hits}^{Si} \geq 3$
Track $\|z_0\| < 60.0$ cm
Isolation < 0.1

Table 4.2: Selection requirements for plug (PHX) electrons.

Muons

Muon candidates consist of a high p_T COT track consistent with a minimum ionizing particle that (usually) extrapolates to a track segment (stub) in one of the muon chambers. The large amount of material to absorb other particles and the requirement of minimum ionizing energy in the calorimeter make the muon samples very pure. We define the following orthogonal muon types (see Figure 4.1(b)) based on the muon chamber they are associated with:

- CMUP: the primary muon sample, requires stubs in both the CMU and CMP muon chambers. These chambers are the most central and the redundancy of the two systems allows for a very pure sample with high efficiency. The stubs are required not to fall within 3 cm of the most forward part of the CMP detector because the detector is not considered trustworthy in that region.

- CMX: requires stubs in the CMX subdetector. However, the forward region of the CMX, where the radius is less than 140 cm, is too far forward to get good tracking and has too high fake rate to use in the trigger. As with the CMP, a CMX stub is required to be 3 cm from the forward edge of the detector to ensure its quality.

- CMU: require stubs in the CMU detector but not the CMP. These are often called η-gap muons because they recover muons in the higher η regions left by the CMP due to its particular rectangular geometry.

- CMP: require stubs in the CMP detector but not the CMU. These are also called ϕ-gap muons because they recover muons in the ϕ-gaps of the CMU wedges.

- BMU: require stubs in the BMU subdetector. Because it is the most forward of the muon detectors, the BMU's tracking is less precise, and it needs looser tracking cuts. The stub must be more than 3 cm from the forward edge of each chamber, and more than 13 cm if $|\eta| < 1.25$, which puts it in a less reliable region of the subdetector.

- CMXNT: muons leave stubs in the non-triggerable region ($\rho > 140$ cm) of the CMX detector. They have the same quality requirements as the normal CMX sample.

- CMIO: isolated tracks matched to calorimeter clusters that do not point toward a muon detector. An additional energy requirement $E_{EM} + E_{Had} > 0.1$ GeV strengthens the requirement of minimum ionizing energy in the calorimeter.

- SCMIO: muons are similar to CMIO muons but are matched to a stub in a non-fiducial region of the detector. These stubs are still required to be matched to a track in the COT and a low-energy calorimeter cluster. The same quality requirements as to CMIO muons apply to these muons.

CMUP and CMX muon types are directly triggered and form the two primary muon samples. The remainder of the muon types are included through the missing transverse energy plus jets triggers (improving these triggers is discussed in appendix D). All these muon types are collected as a single muon category and referred as "untriggered muons" or "EMC muons" (standing for Extra Muon Coverage).

For the muon identification a common set of cuts is applied to all muon, while additional cuts are required for each type. The common muon cuts are presented in Table 4.3 and described as follows:

- p_T: The transverse momentum of the track.
- E_{EM} and E_{Had}: The energy deposited in the electromagnetic and the hadronic calorimeters respectively. This requirement removes particles, especially pions, with enough energy to "punch through" the calorimeter, since they will leave more energy in the calorimeter. The second term, dependent on the particle's momentum p, accounts for the natural rise in ionization energy that a true muon will leave if its momentum is large, in accordance with the Bethe-Bloch equation [32]. The scaling is different for the electromagnetic and hadronic calorimeters because of the different material and thickness between them.
- Track d_0: The impact parameter of the track. This quantity is the distance between the beamline and the position of the track's reconstructed vertex in the r-ϕ plane. This cut removes a background from in-flight decays of long-lived particles into pions or kaons by requiring that the track point back to the beamline. If there are no hits in the silicon detector, which indicates a lower quality track, the cut must be looser.
- χ^2_{track}: The track is compared to the position of the hits in the COT tracking chamber and a χ^2 is calculated. This reduces the background from poorly reconstructed tracks, primarily from kaons that decay in flight.
- COT tracking requirements: Tracks for muons in all subdetectors except the BMU must have at least three axial and at least two stereo COT superlayers with at least five hits each. This ensures that a good track is reconstructed. The BMU, because it is so far forward that tracks do not pass through as much of the COT, cannot use as strict a tracking requirement, and instead requires that 60% of the COT wires along the track have hits in them.

In addition to the common set of cuts, each muon type has slightly different cuts to account for the specific characteristics and geometry of the subdetectors. These particular cuts includes fidutiality (or non-fidutiality) to certain subdetectors including a ρ cut for the CMX and CMXNT types, in addition to a ΔX cut for the stubbed types and $E_{EM} + E_{Had} > 0.1$ GeV for the CMIO and SCMIO types. ΔX is the distance between the stub and the extrapolated track in the direction perpendicular to the beamline and to the radial vector to the cluster. ΔX is required to be less than 7, 5, 6 and 9 cm for the CMU, CMP, CMX and BMU detectors respectively.

Selection cuts
$p_T > 20$ GeV/c
$E_{EM} \leq 2.0 + \max(0, 0.0115(p - 100))$ GeV
$E_{Had} \leq 6.0 + \max(0, 0.0028(p - 100))$ GeV
Track $\lvert z_0 \rvert < 60.0$ cm
Track $\lvert d_0 \rvert < 0.02$ cm (0.2 if no silicon hits)
$\chi^2_{track} < 2.3$ (2.75 if run number < 190697 and no BMU)
Pass COT tracking requirements
Isolation < 0.1

Table 4.3: Common selection cuts for muon identification.

ID efficiency

While all of the cuts shown here remove mostly background, they will also cut out some true leptons. Thus, it is important to estimate the rate at which these events can be reconstructed. This can be estimated by processing Monte Carlo events; however, due to the imperfect modeling of the detector, the efficiency will not be perfectly modeled. This necessitates the addition of a Monte Carlo scale factor, which is the ratio of the data and Monte Carlo efficiencies. Applying this factor to the Monte Carlo acceptance gives a proper estimate of the efficiency of lepton identification.

The identification efficiency of lepton identification cuts is measured in data using $Z \rightarrow e^- e^+$ and $Z \rightarrow \mu^- \mu^+$ samples. A cut around the Z boson mass makes this sample very pure; counting leptons with the same sign gives an estimate of the background. One of the leptons is required to pass the trigger and tight identification cuts, and the other one is examined to see if it also passed the identification cuts. The fraction of identified leptons is the data efficiency. For the central electrons the efficiency is about 80%, while for the plug electrons it is about 70%, lower because of the higher backgrounds and less efficient tracking in the forward region. The efficiency for the muon identification is about 90%.

The same procedure is done in a Monte Carlo simulation of the same process to measure the Monte Carlo efficiency. Because these two are not identical, their ratio is taken as an efficiency scale factor. The scale factors, averaged over all run ranges are quoted in Table 4.4.

4.3.3. Jet Reconstruction

Jets are broad streams of particles resulting from quark or gluon hadronization. The energy of the jet is calculated from the energy deposited in the electromagnetic and hadronic calorime-

Lepton type	ID Scale Factor (%)
CEM	97.9 ± 0.5
PEM	91.4 ± 1.4
CMUP	92.6 ± 0.6
CMX	99.3 ± 0.7
CMU	89.1 ± 1.4
CMP	92.0 ± 1.2
BMU	112.0 ± 1.3
CMXNT	99.3 ± 0.7
CMIO	$104..0 \pm 1.3$
SCMIO	98.6 ± 1.6

Table 4.4: Lepton identification data/MC scale factors.

ter towers using a fixed-cone jet clustering algorithm called JETCLU [140] which relies only on calorimetry. The algorithm is seeded by a local maximum in the total tower energy with at least 1 GeV of deposited energy that is not already associated with a lepton candidate. The energy deposits of all adjacent towers within a fixed $\Delta R \leq 0.4$ cone are considered to build clusters of energy. The algorithm recalculates the center of the cluster with an energy-weighted mean and the jet is reclustered. This procedure is reiterated until the jet energy and its center is stable. Jets overlapping by more than 50% of their energy are merged together, otherwise, the towers in the overlap region are assigned to the nearest jet.

Jet Energy Corrections

The estimation of the energy of the parton associated to a reconstructed jet is one of the most difficult tasks at a hadron collider detector. The reconstructed jet must be corrected for several effects [141] either physical or instrumental, in order to get an accurate estimation of the energy of the parton.

- Level 0 of jet energy corrections sets the energy scales of the calorimeters through different calibrations.

- Jet energy corrections level 1 also referred to as the "η-dependent or relative correction", is applied to raw jet energies measured in the calorimeter to make jet energy uniform along η. The transverse energy of the two jets in a $2 \rightarrow 2$ process should be equal. This property is used to scale jets outside the $0.2 < |\eta| < 0.6$ region to jets inside the region. This is done

because CEM and CHA are the best understood calorimeters in CDF and the selected region is far away of the cracks.

- Jet energy correction levels 2 and 3 are no longer used. The numbering scheme is kept for historical reasons.

- Level 4 is the multiple interaction correction. The energy from different $p\bar{p}$ interactions during the same bunch crossing falls inside the jet cluster, increasing the energy of the measured jet. This correction subtracts this contribution in average. The correction is derived from minimum bias data and it is parameterized as a function of the number of vertices in the event.

- Jet energy correction level 5, or the "absolute" correction, aims to transform the jet energy measured in the calorimeter into the energy corresponding to the underlying particle jet. It corrects the jet energy measured in the calorimeter for any non-linearity and energy loss in the uninstrumented regions of each calorimeter. The jet energy measured is corrected to the sum of the particles' p_T within the cone of same size around the parton direction which matched the jet direction with $\Delta R < 0.4$.

- Level 6 corrections takes into account effects from the underlying event. The underlying event is defined as the energy associated with the spectator partons in a hard collision event. Depending on the details of the particular analysis, this energy needs to be subtracted from the particle-level jet energy. The underlying event energy was measured from minimum bias data requiring events only one vertex.

- Level 7 corrections is also called "out-of-cone" correction. It corrects the particle-level energy for leakage of radiation outside the clustering cone used for jet definition, taking the "jet energy" back to "parent parton energy". The level 7 cone-size-dependent systematic uncertainties are evaluated by looking at energy leakage from the jet clustering cone up to radius $R = 1.3$. The uncertainty for the leakage outside radius $R = 1.3$ is referred to as level 8 "splash-out" uncertainty.

4.3.4. Missing Transverse Energy

Missing transverse energy, $\not\!\!E_T$, is a signature for neutrinos or other exotic particles that do not interact with detectors. Since the longitudinal component of the energy of the colliding initial state partons is not known, while the initial transverse component is assumed to be zero, the transverse energetic balance is used for the neutrino reconstruction. The raw value of the transverse missing energy vector $\vec{\not\!\!E}_T$ is calculated by summing energy deposits in the calorimeter towers, each

projected on a unit vector from the beam axis to the tower:

$$\vec{\not\!\!E}_T = -\sum_i E_i \cdot \sin\theta_i \hat{n}_i, \tag{4.1}$$

where E_i is the energy of the i-th tower, $\hat{n}_i$ is the radial direction unit vector from the beam axis to the i-th tower, and θ_i is the polar angle pointing from the center of the detector to the i-th tower. This sum extends up to $|\eta| < 3.6$.

After $\vec{\not\!\!E}_T$ is calculated it needs to be corrected for the position of the primary vertex of the event, which will cause an adjustment in the direction of the vector associated with each tower. If muons are present, another correction is necessary since muons pass through the calorimeter as minimum ionizing particles. This correction is applied by adding the transverse momenta of the escaping muons to the sum and subtracting the energy deposit by muons in the calorimeters as it has been already counted in the raw calculation. A final correction is applied to account for the difference between the transverse energy of level 0 and level 5 corrected jets.

Neutrino Reconstruction

The missing transverse energy does not give enough information to fully reconstruct the neutrino. The z component of the neutrino momentum in unknown, leaving some ambiguity in the event reconstruction. We can take advantage of the assumption that the neutrino and the lepton come from the semileptonic decay of a real W boson in order to reconstruct the longitudinal component of the neutrino. This result in the following quadratic equation:

$$m^W = \sqrt{(p^\ell + p^\nu)^2} \tag{4.2}$$

where $m^W = 80.4\ \mathrm{GeV}/c^2$ [11]. The smallest $|p_z^\nu|$ solution is chosen, and when the solution is complex[1] the real part is taken as the solution. This choice is correct approximately 77% of the time in the t-channel Monte Carlo.

4.3.5. B Jet Identification

Identifying jets originated from bottom quarks, or "b-jets", is critical for many CDF analysis. Top quarks promptly decay to a W boson and b quarks, the latter hadronizes almost immediately, regardless of its production mode, forming a b-jet. On the other hand, the majority of single top background events contain only light quark or gluon jets in their final states, therefore distinguishing b-jets is a very useful tool to remove backgrounds. This procedure called b-tagging, takes advantage of the fact that B hadrons can only decay through weak interactions and thus have a

[1] Complex solutions can occur due to detector resolution effects.

relatively long lifetime of about 1.6 ps. Furthermore, these B hadrons usually carry most of the large transverse momentum of the original b quark, and thus a large Lorentz boost, causing the B hadrons to travel an average transverse distance, L_{xy}, of several mm before they decay. Hence, B hadron decay products give rise to a secondary vertex displaced from the primary interaction point as shown in Figure 4.2. All b-tagging algorithms at CDF exploit this feature, in this analysis the "SecVtx" [142] algorithm is used.

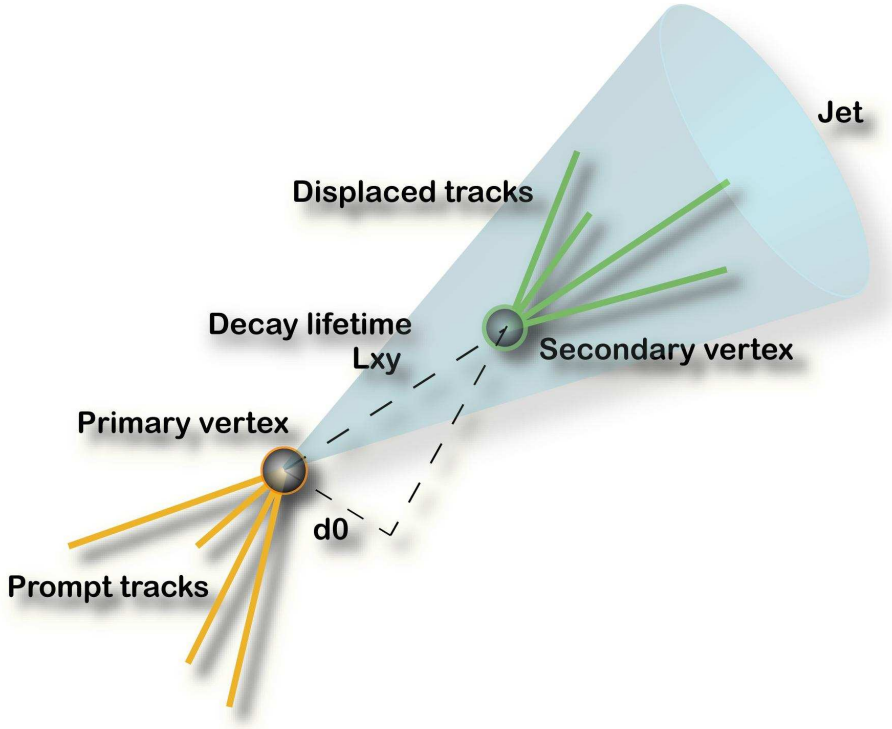


Figure 4.2: Diagram of a secondary vertex with its associated displaced tracks.

SecVtx Algorithm

The SecVtx is operated on an per-jet basis, where silicon tracks within the jet cone are considered for each jet in the event. The selected tracks have to pass certain quality requirements in order to protect against poorly reconstructed tracks. These cuts include the transverse momentum of the tracks, the χ^2/ndf of the final track fit, and the number of silicon hits attached to the track to ensure enough resolution from the silicon detector to attempt b-tagging. To reject long-lived particles that can also lead to displaced vertices, any pair of tracks whose invariant mass is consistent with a K^0 or a Λ particles are removed. Another set of cuts, such as to the impact parameter d_0 with respect to the primary vertex, is applied requiring these tracks to be displaced from the primary interaction. If at least two tracks in a jet pass both the quality and the displaced selection requirements, the jet is

said to be "taggable". As will be demonstrated, taggability is a useful requirement in constructing control regions.

Having selected candidates for a displaced vertex, the SecVtx algorithm performs an iterative fit to the selected tracks to reconstruct a secondary vertex. Once the secondary vertex candidate is formed, it is defined to be b-tagged if it passes a series of vertex quality requirements. SecVtx uses two passes of a vertex finder approach, the first one with at least three tracks emerging from a joint vertex, the second with only two tracks yet satisfying more stringent track quality requirements. Three versions of the SecVtx tagger with different set of selection requirements are built, the so called "tight", "loose", and "ultratight"; the "tight" version is used in this analysis.

A b-tagged jet is said to be positively or negatively tagged based on the sign of L_{xy}, defined to be positive when the secondary vertex is in the same direction of the jet. A positive value of L_{xy} (positive tag) is consistent with a decay of a long-lived particle which results in a jet; a negative value of L_{xy} (negative tag) indicates an improper assignment of a b-tag dues to improperly reconstructed tracks. These negative tags are useful for estimating the rate of false b-tags.

Efficiency and scale factors

While the efficiency of tagging b-jets is only about 40%, due to limited tracking resolution, the efficiency of tagging light jets is less than 1%, so the b-tagged sample is highly purified in b-jet content.

Unfortunately, the extreme complexity of the tagging technique means that the detector simulation is not able to accurately simulate all details of b-tagging: simulated Monte Carlo events typically have more tagged b-jets than the actual data. However, this overestimated efficiency seems to have no kinematic dependence; it affects the rate of tagging but not the kinematic distributions of tagged events. This means that it can be compensated for with a scale factor on the tagging efficiency.

This scale factor is estimated using two methods giving consistent results, one is based on electrons and the other on muons. Both methods start with a sample of two jets, each with transverse energy larger than 15 GeV. One of the jets (the "away jet") is required to be b-tagged while the other is require to contain an electron or a muon with transverse momentum of at least 9 or 8 GeV respectively. The high-momentum electron in the electron/muon jet makes it likely that it comes from a semileptonic decay of the b quark. Requiring the away jet to be tagged increases the purity of the sample, since b quarks often come in pairs and double-tagged events are very rarely faked. The tagging rate of the electron/muon jet allows an estimate of the tagging efficiency. The ratio of this quantity in data and Monte Carlo is the scale factor. Combining both the electron and the muon method, an overall scale factor for the SecVtx of 95 ± 4 % is obtained. Figure 4.3 shows the

SecVtx b-tagging efficiency as a function of the jet E_T for the “tight” and “loose” versions of the algorithm.

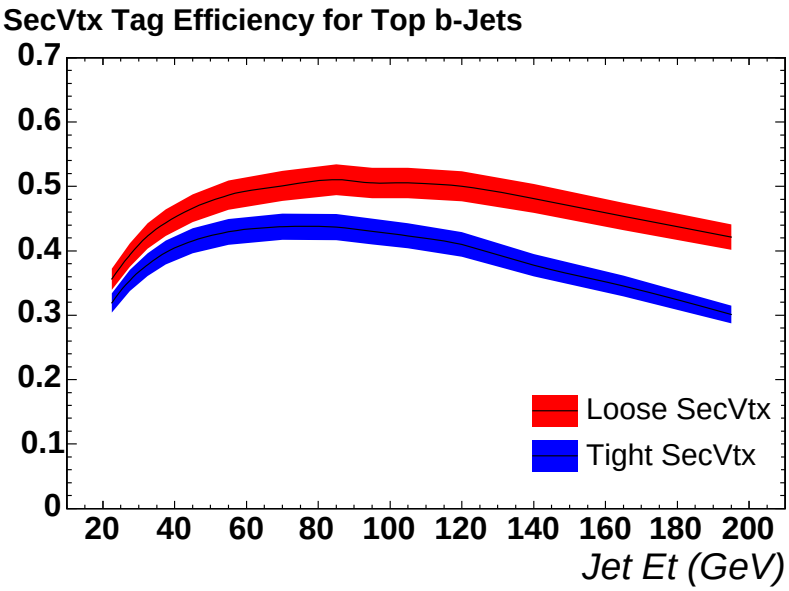


Figure 4.3: Tagging efficiency as a function of the jet E_T to for both the ”tight” and ”loose” tunes of the SecVtx b-tagger. The bands represent the systematic error on the data/MC scale factors. The decrease in efficiency at high jet E_T is due to declining yield of good silicon tracks passing the quality cuts.

Mistag Matrix

An important part of b-tagging is the accidental tagging of jets that do not contain true bottom quarks, called “mistags”. Even though the fraction of mistag is very low, there are far more events with light jets than with b-jets, causing a substantial contamination in the tagged sample.

There are a variety of reasons that the SecVtx algorithm could cause mistags: false positive tags that come from an incorrect identification of a secondary vertex in a jet which does not contain a heavy flavor quark. Tracks in light quark jets could have spuriously large impact parameters because of limited detector resolution, long-lived light particle decays, or material interactions. For many of these mistags, the rate of positively and negatively tagged jets are the same. To a good approximation a secondary vertex from a mistag is equally likely to have a positive or negative decay length, and most vertices with negative decay lengths are spurious. Thus, the number of negatively tagged jets is a good estimate of the light-flavor contribution to the positive sample, although corrections are applied to account for differences between the tagging rates for positively and negatively tagged jets.

Unfortunately, the detector simulation does not properly model the rate of this contribution, and unlike the true tags, the mistag rate modeling has a strong dependence on kinematic variables. To properly understand mistag requires a "mistag matrix", which is a parameterization of the mistag rate as a function of several variables: transverse energy of the jet, the number of tracks in the jet, the sum of the transverse energies of all jets in the event, pseudorapidity, the number of reconstructed vertices in the event, and the z position of the primary vertex. Each variable is divided into four to eleven bins and used to construct a matrix of the rate of negative tags as a function of these six variables from an inclusive sample of jet trigger data. Since the negative tag rate does not fully reflect the positive mistags due to the decays of long-lived particles and interactions with the detector material, a correction factor for the mistag asymmetry $\alpha\beta$ is applied. The factor α corrects for the asymmetry between the positive and negative tag rates of light-flavor jets, and β corrects for the presence of b-jets in the inclusive jet sample in which was used to derive the mistag rate. These are derived from fits to distributions of the invariant mass of the reconstructed secondary vertex in tagged jets in an inclusive jet sample. A systematic uncertainty is derived from fits to templates of pseudo-$c\tau$, which is defined as $L_{xy}\frac{m}{p_{\mathrm{T}}}$ [142], where m is the invariant mass of the tracks in the displaced vertex, and p_T is the magnitude of the vector sum of the transverse momenta of the tracks in the displaced vertex. Figure 4.4 shows the mistag rate as a function of the jet transverse energy.

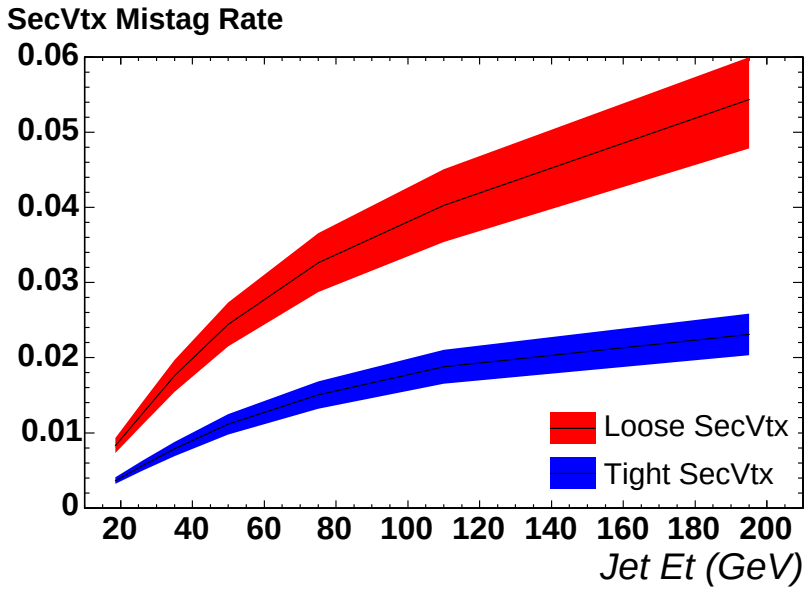


Figure 4.4: The rate of mistags for the loose and tight SecVtx b-taggers as a function of jet E_T. These have been measured from inclusive jet data and includes the mistags asymmetry corrections.

4.3.6. Neural Net Flavor Separator

Due to the long lifetime of charm hadrons, the false reconstruction of secondary vertices in light jets, and the small fraction of events in the pretag sample containing B hadrons, only 52% of the expected background events contain real b-quarks even after imposing the SecVtx b-tagging requirement. Tagged jets without B hadrons in them can be further separated from those containing B hadrons by exploiting further information from other variables related to the secondary vertex and its tracks. On average, vertices from B hadrons tend to have a larger invariant mass of the tracks and a larger track multiplicity, with higher impact parameter and transverse momenta relative to the jet axis per track.

To make full use of all discriminating quantities and their correlations, these variables are used as inputs to a neural network which is applied to jets which are selected by the secondary vertex tagger [143]. The NeuroBayes® package [144] is used for the training of the neural network jet flavor separator. The network is trained with simulated events of single top quark production and the main background processes, mixed according to the background estimation. Processes with secondary vertices due to B hadron decays are treated as signal events, namely single top quark, $t\bar{t}$, and $Wb\bar{b}$ production. Physical processes containing no b-quarks but charm and light flavors are treated as background: $Wc\bar{c}$, Wcj, and W+light jets.

While the SecVtx algorithm only gives a binary decision whether a jet contains a reconstructable secondary vertex, the jet flavor separator output, b_{NN}, can be treated as a measure for the probability of a true b-quark being present within the SecVtx tagged jet. The output of the jet flavor separator is shown in figure 4.5. For jets containing a b quark, the output of the network accumulates at +1, whereas jets without any heavy quark produce an output close to -1. The very similar shapes for different b-quark-producing processes indicates that it is sensitive to the properties of b-quark jets and does not depend on the underlying processes that produce them.

The b_{NN} output not only will be a valuable variable as input to the multivariate analysis to separate single top signal from backgrounds (as will be shown in Chapter 8), but also it will be a crucial tool for estimating the background composition of the different flavors of W+jet events (as will be described in Section 7.2.2).

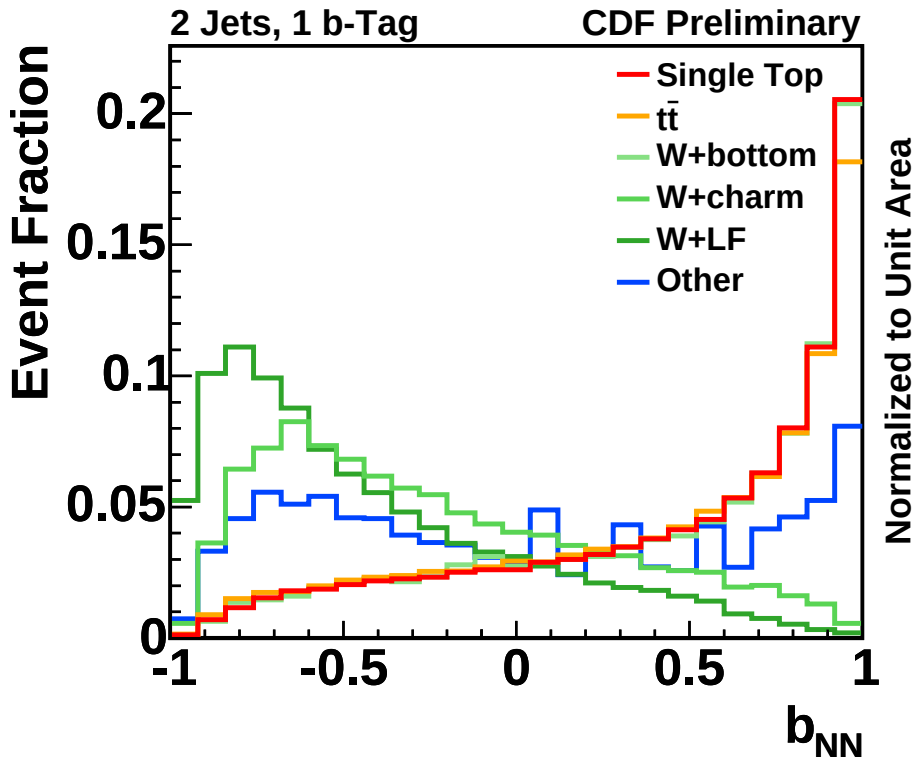


Figure 4.5: Templates of the output of the neural net flavor separator b_{NN} for W+light, W+charm, and W+bottom events. The strong discrimination between them makes this a powerful variable in multivariate analysis. Note that the W+bottom template (in lightest green) is under the red line, which is a good sign as it indicates that all b-flavored jets have the same shape independently of the precess they come from.

Chapter 5

DATA SAMPLE AND CANDIDATE EVENT SELECTION

The common CDF event selection for single top quark candidates exploits the kinematic features of the s- and t-channel signature. Since top quarks promptly decay to a W boson and b quarks, the final state of a single top production event has a W boson and two quarks, at least one of which is a bottom quark. Both s- and t-channel single top diagrams can easily radiate an extra gluon, so some events may have an additional gluon in the final state, thus the desired signature is a W plus two or three jets with at least one being tagged as a b-jet. The W boson can decay into either two quarks or a charged lepton and a neutrino. The branching ratio to quarks is twice as large as that to leptons, however, the corresponding signature is not practical due to the overwhelming QCD multijet background. By contrast, the leptonically decay of the W into an electron, e, or a muon, μ, and its corresponding anti-neutrino $\nu_{e/\mu}$ provides a much cleaner signature. The non-detectable neutrino is required to manifest itself as large missing transverse energy in the detector's energy balance. Because of the difficulty identifying taus, the W boson leptonic decay into a tau and a tau-neutrino only enters the event selection when the tau further decays into detectable electrons or muons.

This finally gives a signal final state with a charged lepton (electron or muon), large missing transverse energy from the neutrino, and two or three jets, at least one of them originated from a b-quark. Based on this signature, this chapter describes the event selection used in this analysis designed to maximize the amount of signal and minimize the amount of background in the final data sample.

5.1. Trigger Data Samples

The first step of an analysis is to choose a trigger to examine. Data is written out by CDF along a variety of "trigger paths", and choosing the right trigger is key to optimizing the event selection. To collect collision data in the W+jets sample, it is sensible to trigger on its distinct features. The most prominent features in leptonically decaying W bosons are the appearance of an energetic charged electron or muon and the large missing transverse energy $\not\!\!E_T$. Thus, high-p_T lepton triggers and $\not\!\!E_T$ triggers are used. Each trigger is a composite of the three levels in the CDF trigger system, with stricter quality requirements imposed at each level.

Because some triggers have prescales, and some have not been active for as long as others, the collected luminosity is different for each trigger. The integrated luminosity recorded for each trigger sample is summarized in Table 5.1.

Trigger	Efficiency (%)	Integrated luminosity (pb^{-1})
CEM trigger	96.5 ± 0.4	3190 ± 190
PHX trigger	98.6 ± 0.6 + turn-on	3190 ± 190
CMUP trigger	91.5 ± 0.5	3180 ± 190
CMX trigger	95.3 ± 0.7	3130 ± 190
$\not\!\!E_T$+jets trigger	turn-on	2990 ± 180

Table 5.1: Trigger efficiency and total integrated luminosity used in this analysis for each trigger path.

5.1.1. High-p_T Lepton Triggers

Central electrons (CEM) have to pass the ELECTRON_CENTRAL_18 trigger that requires at Level 3 a COT track with $p_T > 9$ GeV/c matching an energy deposition in the CEM with $E_T > 18$ GeV. The shower profile of this cluster has to be consistent with the expectation obtained by measurements with test-beam electrons.

Plug electron (PHX) candidates have to pass the MET_PEM trigger, which does not trigger solely on clusters in the PEM calorimeter. This is because of higher background energy depositions from elastic collisions and beam remnants in the forward region. Additionally, due to its central geometry, COT tracking is not available for PHX electron candidates. Instead, this trigger relies on the presence of large missing transverse energy, as well as requiring an energy deposition of at least 20 GeV in the PEM. The lack of tracking requirements for this trigger makes it less pure, thus requiring additional quality selection cuts to purify the sample.

	Selection Requirement		
Level 1	E_T	$>$	8 GeV
	E_{Had}/E_{EM}	$<$	0.125
	p_T	$>$	8.34 GeV$/c$
	$N_{\text{COT hits}}$	$\geq$	4
Level 2	E_T	$>$	18 GeV
	$\lvert\eta\rvert$	$<$	1.317
Level 3	L_{shr}	$<$	0.4
	$\lvert\Delta z\rvert$	$<$	2.0 cm
	E_{Had}/E_{EM}	$<$	0.125 for three towers
	p_T	$>$	9 GeV$/c$

Table 5.2: Selection requirements for the CEM electron trigger.

	Selection Requirement		
Level 1	E_T	$>$	8 GeV
	E_{Had}/E_{EM}	$<$	0.0625
	$\not\!\!E_T$	$>$	15 GeV
Level 2	E_T	$>$	20 GeV
	E_{Had}/E_{EM}	$<$	0.125
	$1.1 <$	$\lvert\eta\rvert$	< 3.6

Table 5.3: Selection requirements for the PHX electron trigger. The Level 3 requirements are the same as Level 2, but with fully reconstructed clusters and tracks.

The central muon (CMUP) trigger MUON_CMUP18 requires at Level 3 a track in the COT with $p_T > 18$ GeV$/c$ matched to track segments in both central muon chambers CMU and CMP simultaneously.

Forward muon (CMX) candidates collected by the MUON_CMX18 trigger have a COT track matched to hits in the CMX muon chambers. Compared to the coincidence of CMUP muon hits in both sub-detectors, CMX candidates are reconstructed from hits in only one sub-detector, which furthermore is less shielded. Even by requiring a timing signal from CSX scintillators consistent with particles coming from Tevatron collisions, this trigger leads to a higher background rates of non-muon entries. Therefore, additional prescaling are imposed during different run ranges in order to control trigger rates with increasing instantaneous luminosity.

A summary of the selection cuts required at each trigger level for all the high-p_T triggers is shown in Tables 5.2- 5.5. In later run ranges (run>226196), a high p_T stereo XFT confirmation for

	Selection Requirement		
	Stub in CMU and CMP		
Level 1	p_T^{CMU}	$>$	6 GeV
	p_T	$>$	4.09 GeV/c
Level 2	Minimum ionizing energy		
	p_T	$>$	14.77 GeV/c
Level 3	$\lvert\Delta X_{CMU}\rvert$	$<$	10 cm
	$\lvert\Delta X_{CMP}\rvert$	$<$	20 cm
	p_T	$>$	18 GeV/c

Table 5.4: Selection requirements for the CMUP muon trigger.

the COT track was added to the Level 2 requirements of both muon triggers.

	Selection Requirement		
	Stub in CMX		
	p_T^{CMX}	$>$	6 GeV
Level 1	p_T	$>$	4.09 GeV/c
	$N_{\text{COT hits}}$	$\geq$	4
	Pass CSX timing requirement		
Level 2	p_T	$>$	14.77 GeV/c
Level 3	$\lvert\Delta X_{CMX}\rvert$	$<$	10 cm
	p_T	$>$	18 GeV/c

Table 5.5: Selection requirements for the CMX muon trigger.

Trigger Efficiencies

The triggers, while very efficient, will not trigger every event that they ought to. Thus, each trigger's efficiency must be estimated to properly predict the rate of each physics process. The basic approach to this is to examine a pure sample obtained through a different trigger, apply the trigger's selection cuts, and see how often the trigger for such events actually fired. In particular, the CEM trigger efficiency is estimated by examining events in a pure sample of W bosons which decay to an electron and a neutrino, selected from a trigger that uses a single electron and large missing transverse energy.

For the PEM trigger efficiency determination, a sample of Z bosons that decay to electrons is

used, with one electron triggered in the central electron trigger and the other detected in the plug calorimeter. Requiring that these electrons come from a Z boson makes this sample very pure. Missing transverse energy is calculated more carefully when processed offline than at trigger level, so the final value may lie below the value that the trigger uses. Because of this, Monte Carlo events must be corrected by a turn-on function to sculpt their kinematics to match the data. This function is parameterized in a W sample triggered with only a calorimeter energy requirement, modeled by the function $\epsilon(x) = \frac{1}{1+e^{-\beta(x-\alpha)}}$. This parameterization gives a weight that is applied to Monte Carlo events in this region. Because additional corrections are applied to plug energy measurements in offline processing which are not included in the Level 2 trigger calculation, the distribution of transverse energy in this sample also needs a turn-on curve. The curve is derived in the same way as the missing transverse energy curve and corrects the electron energy in this sample to match the data.

The efficiencies of the muon triggers are measured in events with Z bosons that decay to two muons, requiring the muons to have opposite charges and an invariant mass near the Z mass. With these events, one muon activates the trigger and the other one is examined to see if it also activated the trigger. Much information can be gained by examining the events in which one muon is detected in the CMUP and the other is detected in the CMX. This comparison allows a simultaneous extraction of efficiencies for both muon triggers.

5.1.2. $\not\!\!E_T$+Jets Trigger

CDF's triggered muons lead to fewer fake events compared to electrons, but at the expense of a lower muon acceptance. The η-ϕ coverage of the triggered muons, which is limited by the fiducial volume of the detectors supplying the muon stubs required by the triggers, is shown in Figure 4.1(b). This figure also shows the coverage of other muon categories that can be reconstructed offline but are not triggerable. Thus, many muon events missed by the muon triggers are potentially recoverable at a later stage during offline reconstruction if the events are collected by a non-muon trigger. Since the most prominent feature of W+jets events besides the high-p_T leptons are large missing transverse energy and several jets, it is feasible to trigger on this signature. Furthermore, since muons are minimum ionizing in the calorimeter, the missing transverse energy at trigger level can be enhanced by the presence of muons in the event over the true missing transverse energy due the escaping neutrino.

The MET35_&_TWO_JETS trigger path requires $\not\!\!E_T > 35$ GeV and two jets with uncorrected $E_T > 10$ GeV. As instantaneous luminosity increased at the Tevatron, it was necessary to modify this trigger in order to reduce the trigger rates to acceptable values. The resulting MET35_&_CJET_&_JET trigger additionally requires one of the jets to be central, $|\eta| < 1.1$. More

details of the trigger requirements at different levels is shown in Table 5.6. Even with the central jet requirement this path had to be dynamically prescaled with increasing instantaneous luminosity. As a result, the integrated luminosity acquired with this trigger is smaller than that of the other discussed triggers.

	Selection Requirement		
Level 1	$\not\!\!E_T$	$>$	25 GeV (15 GeV for later run ranges)
Level 2	Two jets E_T	$>$	10 GeV
	One central jet $\|\eta\|$	$<$	1.1 (only MET35_&_CJET_&_JET)
Level 1	$\not\!\!E_T$	$>$	35 GeV

Table 5.6: Selection requirements for the $\not\!\!E_T$+jets trigger.

Trigger Efficiencies

The study of the efficiency of this trigger is done in two steps, in order to satisfy both the two energetic jets required at Level 2 and the missing transverse energy cut at Level 3. The result of studies using independent samples from generic jets and muon triggers show that the Level 2 trigger requirement on the jets is fully efficient if simple offline jet requirements are made:

- Two jets with $E_T > 25$ GeV corrected by level 5 jet energy corrections. Because the offline transverse energy differs from the energy reconstructed at Level 2 trigger, it is necessary to apply a stronger cut on this variable to ensure fully efficiency.
- One central jet with $|\eta| < 0.9$. Again, a stronger cut than the one applied at trigger level is necessary to ensure fully efficiency.
- Distance between the jets $\Delta R > 1.0$. This is applied to ensure that the trigger jet reconstruction algorithm really detects two jets.

With these cuts we lose about 30-40% of the potential muon signal acceptance. Fortunately, the ΔR cut is very effective removing some of the backgrounds, in particular only about 40% of the W+heavy flavor jets pass the above requirements. As a result, the $\not\!\!E_T$+jets trigger end up in a very pure sample even though it is not as large as the main CEM or CMUP samples.

The second step is to study the trigger efficiency with respect to the missing transverse energy. A stringent cut to the $\not\!\!E_T$ could be applied making the trigger efficiency to be close to 100%, however, in order to increase the signal acceptance, a trigger turn-on curve parameterized as a function the vertex $\not\!\!E_T$ is derived and applied as an event-by-event weight to the MC samples to

reproduce the efficiency of the trigger in data. The vertex $\not\!\!E_T$ is the raw $\not\!\!E_T$ corrected only by the primary vertex position (see Section 4.3.4), it is a well defined offline observable that closely models the $\not\!\!E_T$ reconstructed at Level 3 of the trigger. The $\not\!\!E_T$ trigger turn-on was studied with a sample of CMUP triggered events. The efficiency is defined as the number of events passing both the CMUP and $\not\!\!E_T$+jets trigger out of all CMUP events which pass the Level 2 jet requirements defined above. The trigger turn-on is derived by fitting the measured efficiency as a function of the vertex $\not\!\!E_T$ to the following function:

$$\epsilon(x) = \frac{p_0}{1 + e^{-(x-p_1)/p_2}}, \tag{5.1}$$

where p_0, p_1 and p_2 are the parameters to fit, and x is the vertex $\not\!\!E_T$. The result of the fit is shown in Figure 5.1.

After the trigger Level 2 jet requirements are satisfied and the trigger turn-on is applied to the MC, this trigger adds an effective signal acceptance equal to about 47% of the total acceptance recorded by the triggered muons. While the additional acceptance gain for the sum of the backgrounds is only about 25%.

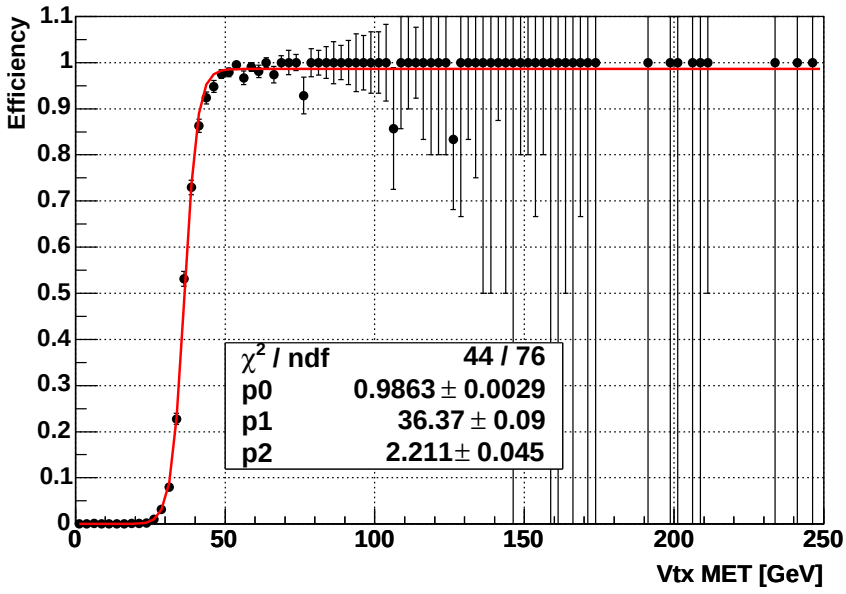


Figure 5.1: Trigger turn-on curve applied to the untriggered muons parameterized as a function of the missing transverse energy corrected only by the primary vertex position.

5.2. Candidate Even Selection

Once an event is selected by the trigger, it is recorded and passed through the offline reconstruction as described in Chapter 4. Further selection requirements are applied to the trigger samples

in order to create a final candidate event sample with a maximized single top signal content while keeping the backgrounds as low as possible.

5.2.1. Jet Requirements

Since single top quark signal events feature forward jets, the jet acceptance is enlarged up to $|\eta| < 2.8$, while CDF's standard lepton+jets selection is limited to $|\eta| < 2.0$. All jet candidates have to fulfill $E_T > 20$ GeV, where the jet energies are corrected up to Level 5, which transforms them into an absolute scale.

For the untriggered muons further jet requirements are applied to make the $\not\!\!E_T$+jets trigger at Level 2 fully efficient as described in Section 5.1.2. At least two jets with $E_T > 25$ GeV, where at least one is central $|\eta| < 0.9$, and separated by a distance $\Delta R > 1.0$.

The distribution of Monte Carlo signal events as a function of the jet multiplicity before any selection cuts further than the trigger and the jet identification requirements is shown in Table 5.7. This analysis examines events with two or three, thus selecting most of the single top events. Events with only one jet have a very large background from W+jets processes that make them of little use for measuring a single top production signal.

Finally, in order to reduce the amount of background contamination, in particular events from light-flavor processes, at least one of these jets is required to be tagged as a b-jet using the SecVtx tagger algorithm. Because the composition of signal and background events is very different among the jet multiplicities and for single and double tagged events, we usually divide our sample into four separate samples: W+2 jets and 1 tag, W+2 jets and 2 tags, W+3 jets and 1 tag, W+3 jets and $\geq$2 tags.

Sample	0 jets	1 jet	2 jets	3 jet	4 jets	$\geq$5 jets
s-channel	1%	14%	46%	29%	8%	2%
t-channel	1%	18%	48%	26%	6%	1%

Table 5.7: Fraction of Monte Carlo single top events as a function of the jet multiplicity.

5.2.2. Lepton Requirements

The basic lepton selection cuts made by the trigger still leave a large number of fake leptons. Additionally, some triggers have no lepton requirement at all. Therefore, further identification cuts are applied in order to select good offline isolated electrons and muons as described in Section 4.3.2.

To ensure that there is exactly one lepton per event, the "dilepton veto" is applied. This veto looks for any leptons in an event, including non-isolated leptons, and rejects the event if there is more than one lepton. This requirement is very efficient removing background events from top pair production in which both final state W bosons decay to leptons.

5.2.3. Missing Transverse Energy Requirements

Missing transverse energy, the signature of a neutrino, is an important part of event selection cuts because it removes events from many QCD multijet processes that have no final-state neutrino. Combined with the lepton identification requirement, a missing transverse energy cut makes it very likely that an event included a W boson in its final state that decayed leptonically. Because of this, it is important that missing transverse energy be calculated as accurately as possible, therefore primary vertex position corrections, muon corrections and jet energy corrections are applied to the raw $\not\!\!E_T$ (see Section 4.3.4).

This analysis requires the missing transverse energy to be greater than 25 GeV, which removes a large portion of the QCD multijet background. However, events that do not pass this selection requirement are useful for estimating the remaining QCD multijet contribution (see Section 7.2.1).

To simulate the PEM and the $\not\!\!E_T$+jets trigger efficiencies, event-by-event weights from the corresponding $\not\!\!E_T$ trigger turn-on functions are applied to the MC samples.

5.2.4. Event Vetoes

Several additional selection requirements are required to remove specific backgrounds to increase the purity of the final sample and to improve the background modeling.

Z Boson Veto

To remove events from Z boson production, events are rejected in which the charged lepton can be paired with any more loosely defined jet or lepton to form an invariant mass in a range consistent with a Z boson mass (76–106 GeV). This leaves very little residual contamination from Z+jets events.

Cosmic Ray Veto

Muons coming from decays of cosmic-ray pions in the upper atmosphere pass through the detector frequently. They pass the muon identification requirements, as they are real muons, although

they must be remove since they do not originate from collisions in the Tevatron. Thus, events with a pair of back-to-back tracks consistent with the characteristic timing and topology of that of a cosmic ray are rejected.

QCD Veto

The safest way to deal with the difficult QCD multijet sample is to remove as much of it as possible. These events often have very strong kinematic features not described by other Monte Carlo samples that can be observed as a difference between data and Monte Carlo prediction. Because these events do not contain a W boson, we apply additional selection requirements which are based on the assumption that these events do not have a large $\not\!\!E_T$ from an escaping neutrino, but rather the $\not\!\!E_T$ that is observed comes from lost or mismeasured jets. In events lacking a W boson, one would expect small values of the transverse mass, defined as

$$M_T^W = \sqrt{2\left(p_T^\ell \not\!\!E_T - p_x^\ell \not\!\!E_x - p_y^\ell \not\!\!E_y\right)}. \tag{5.2}$$

Because the $\not\!\!E_T$ in events that do not contain W bosons often comes from jets which are erroneously identified as leptons, it often points opposite to the lepton, giving the event a low transverse mass. Thus, applying a cut of 10 GeV on the transverse mass removes a large part of the QCD multijet background. This purifies the muon samples well enough since it is very hard to fake a muon stub and a calorimeter cluster with minimum ionizing energy. However, for the SCMIO muon category a cut of 20 GeV is required since this sample is more sensitive to QCD multijet contamination.

The electron samples, lacking the advantages of the muon detectors, have a more significant contamination from QCD multijet events. A transverse mass cut at 20 GeV is applied, nonetheless further purification in the electron samples is needed to achieve a fraction of the QCD multijet background similar to that of the muon sample. This can be provided by a variable called MET significance, defined as

$$METsig = \sqrt{\not\!\!E_T / \hat{\not\!\!E}_T \cdot \vec{E}_T^{\,\text{unclustered}}}, \tag{5.3}$$

where unclustered energy refers to energy not included in any reconstructed jets. Carefully chosen triangular cuts (see Figure 5.2 for an example) are applied to reduce the amount of QCD multijet events. Central electron events must have $METsig > 0.05 m_T + 3.5$ and $METsig > 2.5 - 3.125\Delta\phi_{\not\!\!E_T,jet2}$. Plug electron, due to the lower quality of its tracks, has more contamination from QCD events and requires tighter cuts such as $METsig > 2$ and $\not\!\!E_T > 45 - 30\Delta\phi_{\not\!\!E_T,jet}$ for all jets in the event. These cuts improve the modeling and substantially reduce the amount of contamination from non-W events.

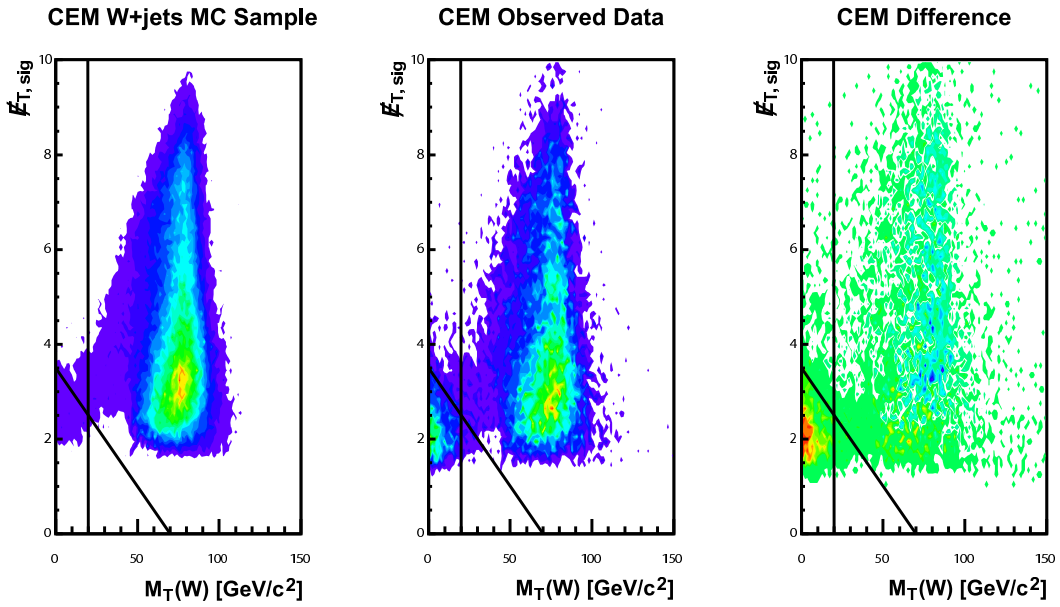


Figure 5.2: Illustration of the QCD veto for CEM events. MET significance and transverse mass for Monte Carlo, data, and the two plots subtracted. The black line indicates the cut which is applied.

Chapter 6

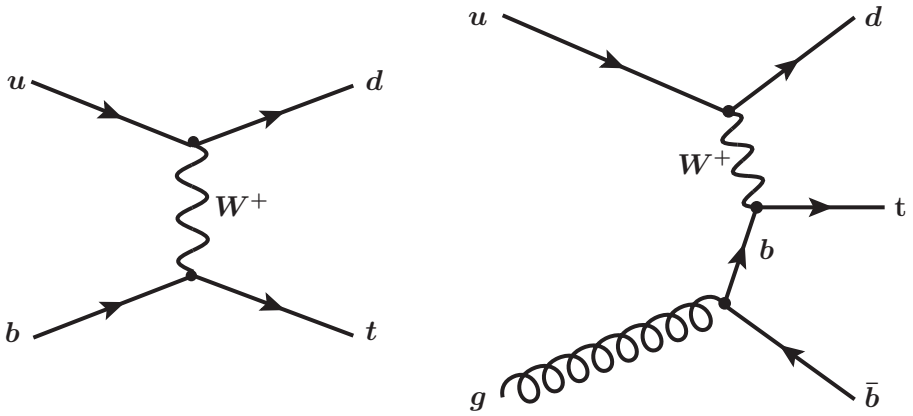
MODELING OF PROCESSES

In order to perform a search for a previously undetected signal such as single top quark production, accurate models predicting the characteristics of the expected data are needed for both the signal being tested and the SM background processes. Special care must be taken for analysis using multivariate techniques, as the final result is highly sensitive to kinematic mismodeling. Most of the processes involved in this analysis are described using Monte Carlo simulation, while some background processes are derived from data.

6.1. Signal Event Modeling

Top quarks are expected to be produced singly in $p\bar{p}$ collisions via weak, charged-current interactions. The dominant processes at the Tevatron are the t-channel and the s-channel. The next-to-leading-order (NLO) cross sections for these two processes are $\sigma_{s-\mathrm{chan}} = 0.88^{+0.12}_{-0.11}$ pb and $\sigma_{t-\mathrm{chan}} = 1.98^{+0.28}_{-0.22}$ pb, respectively, assuming $m_{top} = 175\ \mathrm{GeV}/c^2$ [12, 13]. A third process, associated production of a W boson and a top quark, is not expected to be observable at the Tevatron, due to the massive final state, a W boson and a top quark that must originate from a bottom sea quark.

The matrix element generator MADEVENT [131] is used to produce simulated events for the single top samples. This generator fully incorporates the spin of the top quarks in contrast to the PYTHIA generator, thus preserving the information from the polarization of the top quark which is an interesting feature of the electroweak top quark production. The generator is interfaced to the CTEQ5L [62] parameterization of the parton distribution functions (PDF). The MADEVENT output is passed to the PYTHIA [129, 130] program to perform the parton shower and hadronization to



(a) (b)

Figure 6.1: The two different t-channel processes considered in our signal model: (a) the $2 \to 2$ process and (b) the $2 \to 3$ process.

generate all expected final-state particles.

6.1.1. s-channel Single Top

Although the MADEVENT program only calculates events at leading-order (LO) accuracy, studies [12] indicate that the kinematic distributions of s-channel events are not affected by Next-to-Leading-Order (NLO) corrections. Therefore s-channel events are generated at leading-order and scaled to the Next-to-Leading-Order cross section.

6.1.2. t-channel Single Top

The t-channel process is more complicated. Several authors point out [12, 145, 146] that the leading order contribution to t-channel single top quark production as modeled in parton-shower Monte Carlo programs does not adequately represent the expected distributions of observable jets, which are better predicted by NLO calculations.

The leading-order process is a $2 \to 2$ process with a b quark in the initial state: $b + u \to d + t$, as shown in Figure 6.1(a). For antitop quark production, the charge conjugate processes are implied. A parton distribution function for the initial state b quark is used for the calculation. Since flavor is conserved in the strong interaction, a $\bar{b}$ quark must be present in the event as well. In what follows, this $\bar{b}$ quark is called the spectator b quark. Leading-order parton shower programs create the spectator b quark through backward evolution following the DGLAP scheme [59–61]. Only the low-p_{T} portion of the transverse momentum distribution of the spectator b quark is modeled well,

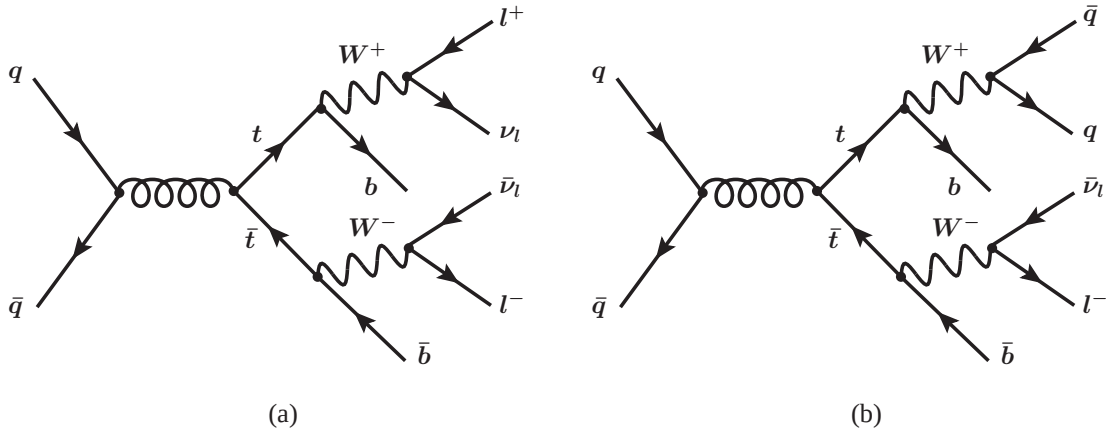
while the high-p_{T} tail is not estimated adequately [12]. In addition, the pseudorapidity distribution of the spectator b quark as simulated by the leading-order process is biased towards higher pseudorapidities than predicted by NLO calculations.

We improve the modeling of the t-channel single top quark process by using two samples: one for the leading $2 \to 2$ process $b + q \to q' + t$, and a second one for the $2 \to 3$ process in which an initial-state gluon splits into $b\bar{b}$, $g + q \to q' + t + \bar{b}$. In the second process the spectator b quark is produced directly in the hard scattering described by the matrix element (Figure 6.1(b)). This sample describes the most important next-to-leading order (NLO) contribution to t-channel production and is therefore suitable to describe the high-p_{T} tail of the spectator b quark p_{T} distribution. This sample, however, does not adequately describe the low-p_{T} portion of the spectrum of the spectator b quark. In order to construct a Monte Carlo sample which closely follows NLO predictions, the $2 \to 2$ process and the $2 \to 3$ process must be combined.

A joint event sample was created by matching at generator level (before any showering or hadronization) the p_{T} spectrum of the spectator b quark to the differential cross section predicted by the ZTOP program [12] which operates at NLO. The matched t-channel sample is created in such a manner that the p_{T} distribution of spectator b quarks in the matched t-channel sample consists of $2 \to 2$ events for transverse momenta below a certain cutoff K_{T} and of $2 \to 3$ events for transverse momenta above K_{T}. The ratio R between the two processes is varied until the rate of events with a detectable spectator b quark jet, with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.8$, matches the NLO prediction by ZTOP. This gives a continuous distribution in the transverse momentum that matches the theoretically predicted distribution [147].

6.2. Background Event Modeling

The major experimental challenge of a single top measurement is the large amount of background events that mimic the signal events. Top pair production, in contrast, has a clean signature of a single lepton, four energetic jets, and a neutrino that shows up as large missing transverse energy, and there is far more top pair production than anything else of events with this signature. On the other hand, the single top production has a smaller cross section and a final state consisting of a lepton, missing transverse energy, and only two or three jets with at least one being originated from a b-quark. This is also the final state of the $Wb\bar{b}$ process which has a much larger cross section. Other processes with similar final states, such as $Wc\bar{c}$ and $t\bar{t}$, also mimic the single top signature because of misreconstruction or loss of one or more components of the expected final state. A detailed understanding of the rate and kinematic features of the background processes is necessary in order to measure the single top quark production cross section precisely.



(a) (b)

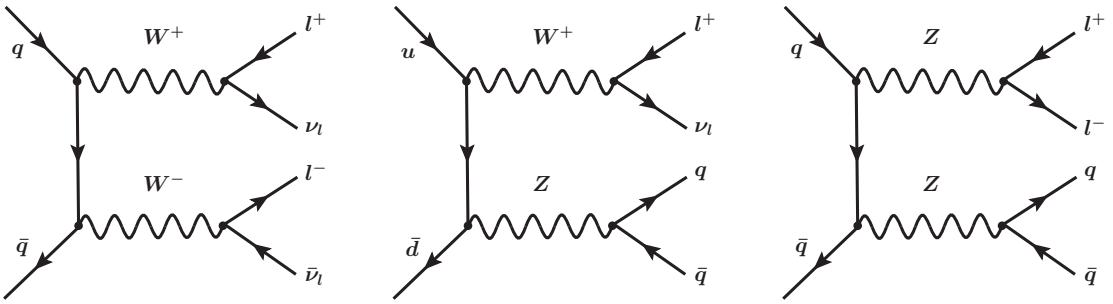
Figure 6.2: Feynman diagrams of the $t\bar{t}$ background to single top quark production. To pass the event selection, these events must have one charged lepton or two hadronic jets that go undetected.

6.2.1. Top Pair Production

Top pair production can be a substantial background to the single top production, mainly in the three jet sample or in the double tag samples since top pair production final state contains many jets with two real b-jets. While no top pair decay has the same final state as single top, it can be misidentified if particles fail to be detected. This happens in dilepton top decays in which one of the leptons is missed, or in lepton + jets decays in which two jets are not detected (Figure 6.2). The resulting events, since they also come from top decays, are very similar kinematically to single top events, making this background difficult to deal with. The top pair production events are generated by PYTHIA [130] and scaled to the NLO theoretical cross section prediction [148, 149] of $\sigma_{t\bar{t}} = (6.70 \pm 0.83)$ pb, assuming $m_t = 175$ GeV/c^2. The systematic uncertainty contains a component which covers the differences between the calculation chosen and others [5, 150].

6.2.2. Diboson Production

Electroweak diboson production, including WW, WZ, or ZZ production (Figure 6.3), also creates a small background, especially WW and WZ, which have the same final state as single top production. However, since their cross section is small, they constitute a small contribution to the total background. The diboson production samples are generated using PYTHIA Monte Carlo and normalized to the NLO theoretical cross section predicted for a center of mass energy of $\sqrt{s} = 2.00$ TeV [151] and extrapolate the value to $\sqrt{s} = 1.96$ TeV. This leads to $\sigma_{WW} = (13.30 \pm 0.40)$ pb, $\sigma_{WZ} = (3.96 \pm 0.12)$ pb, and $\sigma_{ZZ} = (1.57 \pm 0.05)$ pb.



(a) (b) (c)

Figure 6.3: Feynman diagrams for diboson production, which provides a small background for single top quark production.

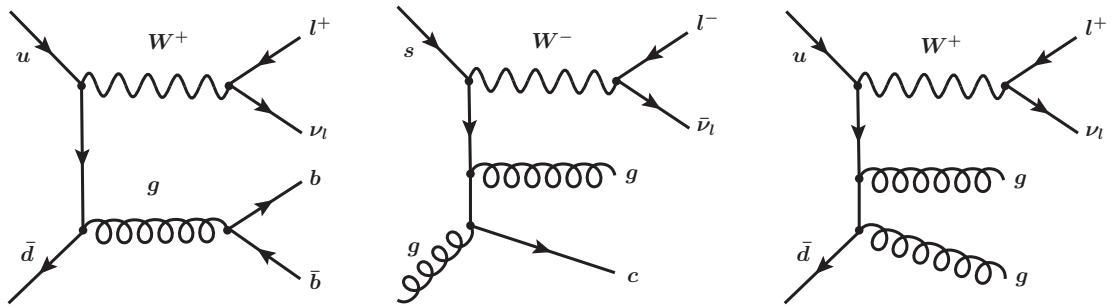
6.2.3. W+Jets

The largest background to single top production comes from quark interactions that radiate a W boson in association with two jets, which has the same final state as single top and a much larger cross section. Because most of the quarks in the final state are light quarks (up, down, or strange quarks), the level of this background can be reduced by b-tagging , but their cross section is very large, so they still form a substantial background. However, this fails to help in the case where a radiated gluon splits into at least one heavy flavor quarks (charm or bottom). These W+heavy flavor (W+HF) processes are $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$, each with possibly more jets, and constitute the largest background processes in this analysis.

The W+jets production is described by a huge amount of possible Feynman diagrams (some are shown in Figure 6.4) which describe the color and kinematic characteristics of the radiated gluons. ALPGEN [133] generator is used to generate these events because it properly calculates all tree level matrix elements with full color and spin correlation information. The generated events are passed to the PYTHIA showering routine.

Parton-Jet Matching

The PYTHIA showering algorithm gives rise to a difficulty when used with ALPGEN because there is an overlap in their generation of events. ALPGEN generates events at the matrix element level with initial- and final-state radiation, while PYTHIA approximates the effects of radiation by its showering. PYTHIA performs much more showering than ALPGEN does, but the initial stages of showering overlaps: ALPGEN might produce a diagram with a W boson and two radiated gluons, or it could produce a diagram with a W boson and one radiated gluon, while PYTHIA adds another gluon through parton showering. Because both these cases can occur, these events will appear with



(a) (b) (c)

Figure 6.4: Some representative diagrams of W+jets production: (a) $W + b\bar{b}$, (b) $W + c + \mathrm{jet}$, (c) W+light flavor jets. The productions cross sections of these processes are much larger than that of single top quark production.

too large a rate.

In order to solve this we follow the so-called "MLM matching prescription" [133] . In this method, after parton showering, the final-state particles are grouped into jets by a jet-cone clustering algorithm, which groups all particles within a certain region of $\eta - \phi$ space. Each jet is then matched to a parton if the parton lies within the cone of the jet. Only one parton can be matched to each jet. An event is rejected if it cannot match every parton to a jet.

To get the counting right, the number of jets is required to be the same as the number of partons. Exclusive samples are generated with different numbers of jets, then added together after matching is performed. This removes double-counting in the showering.

Heavy Flavor Removal

There is another problem of double-counting when using PYTHIA with ALPGEN, and it involves quarks with non-trivial masses, i.e. heavy flavor quarks. It is important to separate events with these quarks because their kinematic behavior is different from the lighter quarks. However, they can arise in two different ways: they can be created at the matrix-element level in a $W + b\bar{b}$ event, or they can arise from gluon splitting in the parton shower from a W+light flavor event. Because there is no difference between these two cases (they have the same Feynman diagram), combining ALPGEN and PYTHIA will overestimate the heavy flavor rate by counting the same events in both $W + b\bar{b}$ and W+light flavor samples.

The scheme for removing the overlap of heavy flavor divides heavy flavor events into two disjoint sets based on matching to fully reconstructed jets. Heavy flavor events generated by the matrix element are kept only if the heavy quarks lie in two different jets, while events generated by

the parton shower are kept only if the heavy quarks lie in the same jet. This division is motivated by the expectation that quarks from showering will usually be close to their parents, while quarks from the matrix element are more likely to be well separated. This prescription removes the overlap between these events.

Mistags

Modeling the contribution of W+light flavor jet events that are nonetheless b-tagged (W+LF, or the so-called mistag sample) is difficult because, while the Monte Carlo simulation does a good job predicting general W+jets kinematic shapes, some events are more kinematically disposed to mistagging, and the Monte Carlo does not model their kinematic features well. Therefore, the W+light flavor sample is not tagged directly; rather, the tagging requirement is relaxed to require only one taggable jet, and each event is weighted by the product of the mistag probabilities of its taggable jets. This results in a kinematic distribution that closely models the kinematic shapes of the mistag sample.

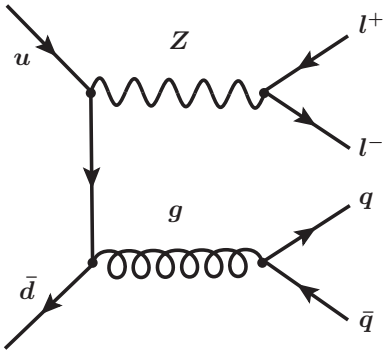
Because events with one and two tagged jets are treated separately, it is necessary to have a different mistag model in each region. The single-tagged model is constructed as described above; the double-tagged sample uses only events with at least two taggable jets weighted by their mistag probability, and taking into account the combinatorics associated to multiple jets.

The problem with the method described above is that the non-tagged jets do not have a neural net flavor separator output (see Section 4.3.6). To simulate the output of the jet flavor separator, a random value is chosen from the distribution in light-flavor jets shown in Figure 4.5. If a W+LF event has more than one taggable jet, then random values are assigned to both jets. These events are used for both the single-mistag prediction and the double-mistag prediction with appropriate weights.

6.2.4. Z+Jets

Since Z bosons decay to either two leptons or two neutrinos, Z+jets events do not often fake a single top signature and the background is not large. However, because of its large inclusive cross section, some background remains from events in which a lepton is lost and its energy is counted as missing transverse energy (Figure 6.5).

The Z+jets process has the same difficulties as W+jets and is generated by ALPGEN in the same way, with PYTHIA is used to model the parton shower and hadronization. The Z+jets cross section is normalized to that measured by CDF [152]: $\sigma_Z \cdot \mathrm{Br}(Z \to \ell^+\ell^-) = (336.0 \pm 8.0)$ pb, where $\mathrm{Br}(Z \to \ell^+\ell^-)$ is the branching ratio of events with leptonically decaying Z bosons.



(a)

Figure 6.5: Representative Feynman diagram for a Z+jets production. To pass the event selection, these events must have one undetected charged lepton.

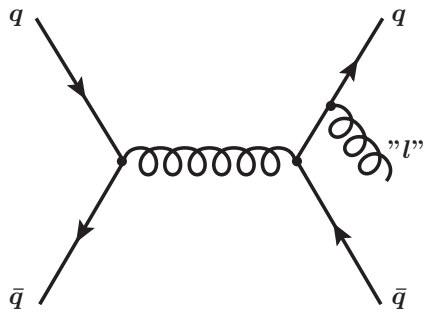
6.2.5. QCD Multijet

The most difficult events to model come from QCD multijet events also referred to as "non-W" events. In order to be confused with a single top event, one of the jets must fake a lepton and a mismeasurement has to create a large missing transverse energy (Figure 6.6). The extremely large cross section of QCD multijet events means that even kinematically unlikely configurations can contaminate the signal sample significantly. Because of the extremely small probability of these events occurring, and because they come from several different QCD processes, all of them difficult to calculate or model, it is impossible to simulate these events with Monte Carlo events and a data-based sample and estimation must be constructed.

A data-based model of this sample is complicated as well, since different lepton types may have vastly different rates and shapes of QCD multijet events, and each needs to be examined separately. This analysis uses three different models for QCD multijet events. All of them are based on the principle that QCD multijet events must contain a jet that is falsely identified as a lepton. Thus, by looking at jets that are not leptons but come close to passing electron cuts, it is possible to create a model of this background.

Jet-Electrons and Anti-Electrons

One strategy for creating a model of this sample uses a sample of inclusive jets, triggered through a generic jet trigger which simply looks for clusters of energy in the calorimeter. Since QCD multijet events must involve a jet that is falsely identified as an electron, this sample is examined for jets that look similar to electrons. Specifically, they must be energetic, having a



(a)

Figure 6.6: Representative Feynman diagram for QCD multijet event, in which a jet has to be misidentified as a lepton and $\not\!\!E_T$ must be mismeasured to pass the event selection. Because the cross section of non-W events is so large, it still appears as a background process to the single top signature.

transverse energy of at least 20 GeV; they must have a high fraction of energy deposited in the electromagnetic calorimeter, between 80% and 95%; and they must have fewer than four tracks, since a true electron has only a single track. This jet is then assumed to be an electron and all the other event selection cuts are applied. The electron charge is assigned randomly since jets have multiple tracks and their charge cannot be determined. Because they come from generic jets similar to electrons, these events are referred to as "jet-electrons".

Another sample takes advantage of the fact that fake leptons from non-W events have difficulty passing the lepton selection requirements. We look at lepton candidates in the central electron trigger that fail at least two of five identification cuts that do not depend on the kinematic properties of the event, such as the fraction of energy in the hadronic calorimeter. These objects are treated as leptons and all other selection cuts are applied. Because these events are similar to electrons with some selection cuts inverted, they are called "anti-electrons".

The drawback of these two samples is their small size. Because of the similarities in the kinematic properties of the anti-electron and the jet-electron events, we combine the two samples for our non-W model for triggered central electrons. For plug electrons, we use only the jet-electron sample because the other sample does not include electron candidates with high pseudorapidities. Remarkably, the same samples also simulate the kinematics of events with misidentified triggered muon candidates (CMUP and CMX); we use the samples again to model those events.

Non-Isolated Muons

Misidentified, untriggered muons cannot be modeled well by either of these samples because the kinematic properties of this sample are sculpted by the substantially different trigger requirements. For this sample, we use muon events that fail the lepton isolation requirement, referred to as "non-isolated" muons; such events come from the same trigger and model the non-W contribution well for this sample. This is based on the rationale that non-isolated events are typically leptons contained in jets, and jets that contain energetic leptons are more likely to pass lepton identification cuts. This sample has the advantage of a large size, unlike the other models; however it does not reproduce kinematic variables for the triggered leptons as good as the other models and therefore this model is only used for untriggered muons.

b-tagging

The statistics of the QCD multijet samples are too small for direct tagging; only a handful of jet-electron and anti-electrons events contain a tagged jet, making them useless for modeling kinematic distributions. In order to increase sensitivity, the b-tagging requirement is relaxed to requiring a taggable jet instead of a tagged jet to increase the size of the selected samples. This matches the missing transverse energy distribution in the data of the QCD-enriched sample of low missing transverse energy. In addition, the measured tagging rate in data shows no significant dependence on missing transverse energy, indicating that there is no large change in the kinematic distributions of this sample after tagging.

The distributions of b_{NN} (see Section 4.3.6) for QCD multijet events are more difficult to predict because the flavor composition of the jets is poorly known. The fraction of each flavor: bottom, charm, and light-flavored jets (originating from light quarks or gluons), is estimated by fitting the jet flavor separator templates (shown in Figure 4.5) to the $15 < \not\!\!E_T < 25$ GeV sideband of the data. In this sample, we find a flavor composition of 45% b quark jets, 40% c quark jets, and 15% light-flavored jets. Each event in the QCD modeling samples is randomly assigned a flavor according to the fraction given above and then assigned a jet flavor separator value chosen at random from the appropriate flavor distribution. The most b-like possibility of the errors on the flavor measurement gives an alternative "worst-case" flavor composition of 60% b quark jets, 30% c quark jets, and 10% light-flavored jets. This alternative flavor composition affects the shapes of the final discriminant distribution through the different flavor-separator neural network values and therefore a shape systematic uncertainty will be associated to the flavor model of the QCD background as described in Section 9.1.

Chapter 7

PREDICTED EVENT YIELD

Properly estimating the signal and background composition of the sample of candidate events after applying the event selection is an essential and elaborate process. First, the background estimate method assumes that all processes contributing significantly to the lepton+jets sample are known. In case of the search for single top quark production analyses, these are single top quark, top quark pair and diboson production, W+jets and Z+jets production, and QCD-multijet events. Precise NLO calculations exist for the production cross section of some of the processes, thereby making the estimation of their contributions is a relatively straightforward process. However, because the Monte Carlo simulation is not a perfect model for some samples, the background estimate relies heavily on data.

The single top sample is broken up into events with exactly two or three jets, and with one or more b-tags. This gives a total of four categories with a separate estimate of the sample composition. Treating each sample separately improves the sensitivity of the analysis by combining regions with different signal purities. In addition, backgrounds are calculated for other orthogonal control regions. The sample of one jet events is an important control region of the W+jets background. The "untagged" samples, which requires at least one taggable jet but no tagged jets, are important W+light flavor enriched control samples. Finally, four jet events are useful for validating the top pair production background.

7.1. MC-based Predictions

For any process with a well understood cross section, the yield estimate is derived from the MC simulation. These processes include the single top samples, top pair and diboson production, and

Z+jets events. The cross sections for these processes are described in Chapter 6 and summarized in Table 7.1. The expected number of events for these samples is given by:

$$N = \sigma \cdot \varepsilon_{\rm evt} \cdot \mathcal{L}_{\rm int}, \tag{7.1}$$

where σ is the predicted cross section times the branching ratio of the respective process, $\varepsilon_{\rm evt}$ is the event detection efficiency, and $\mathcal{L}_{\rm int}$ is the integrated luminosity. The predicted production cross sections for each process will be discussed in the next subsections. The integrated luminosity used in this analyzed is quoted in Table 5.1 for each trigger sample.

The determination of event detection efficiency for each process starts from the calculation of the acceptance in MC of the given process $\epsilon_{\rm MC}$. The acceptance is derived from the simulated MC samples as the fraction of events passing all the event selection cuts. This number needs to be corrected in order to take into account the differences between the simulation and the real experimental setup. Thus, the event detection efficiency can be decomposed in the following factors:

$$\epsilon_{\rm evt} = \epsilon_{\rm MC} \cdot \epsilon_{\rm trig} \cdot \epsilon_{\rm tag} \cdot SF_{\rm z_0} \cdot SF_{\ell_{\rm ID}}, \tag{7.2}$$

where $\epsilon_{\rm trig}$ is the corresponding trigger efficiency described in Sections 5.1.1 and 5.1.2; the factor $\epsilon_{\rm tag}$ corrects by the SecVtx b-tagging scale factor and mistag matrix described in Section 4.3.5; $SF_{\rm z_0} = \epsilon_{\rm z_0}^{\rm data}/\epsilon_{\rm z_0}^{\rm MC} = 1.09$ is the scale factor that takes into account the observed differences between data and MC in the efficiency of the primary vertex cut (see Section 4.3.2); and $SF_{\ell_{\rm ID}}$ is the lepton identification scale factor described in Section 4.3.2.

For the PEM and the $\not\!\!E_T$+jets trigger samples, the trigger efficiencies corresponding to the turn-on curves are absorbed by the MC acceptance by applying an event-by-event weight derived from the turn-on functions to each MC event.

Process	Cross section (pb)
s-channel	$0.88^{+0.12}_{-0.11}$ [12, 13]
t-channel	$1.98^{+0.28}_{-0.22}$ [12, 13]
$t\bar{t}$	6.70 ± 0.83 [148, 149]
WW	13.30 ± 0.40 [151]
WZ	3.96 ± 0.12 [151]
ZZ	1.57 ± 0.05 [151]
Z+jets	336.0 ± 8.0 [152]

Table 7.1: Cross sections used for the yield estimate of the MC-based samples. For the single top and top pair samples a top mass $m_{top} = 175\ {\rm GeV}/c^2$ is assumed.

The b-tagging correction requires special attention. Each MC jet is assigned a tagging probability weight based on whether it can be matched (before detector simulation) to a heavy flavor quark. If a jet is matched to a heavy flavor hadron and tagged, it is given a weight equal to the b-tag scale factor. If it is matched to heavy flavor but not tagged, it is given a weight of zero. If the jet is not matched to heavy flavor, it is assigned a weight equal to its mistag probability, regardless of whether or not it was tagged, because the Monte Carlo simulation does not properly model mistagging. Untaggable jets always have a weight of zero.

After a tagging weight is assigned to every jet in an event, the event is given a tagging probability for every assumption of number of tags in the event. This probability must take into account all possible combination of jets for the given assumption. For example, the tagging probability of a two jet event to be single-tagged is $w_1(1-w_2)+(1-w_1)w_2$, where w_i is the tagging weight of the jet i or the probability of the jet to be tagged (whereas $1-w_i$ is the probability of the jet not to be tagged). Thus, every MC event is assigned a probability of no tags, exactly one tag, and two or more tags. This probability is used as an event weight in the drawing of the histograms and in the calculation of the event yield estimate (with this weighting method, the factor $\epsilon_{\rm tag}$ is absorbed by $\epsilon_{\rm MC}$).

Including all trigger and identification efficiencies, we find event detection efficiencies for the signal processes of the order of $\varepsilon_{\rm evt}(t\text{-channel}) = (1.2\pm0.1)\%$ and $\varepsilon_{\rm evt}(s\text{-channel}) = (1.8\pm0.1)\%$.

7.2. Data-Driven Predictions

Estimating the QCD multijet contribution to the final sample is difficult because of the difficulty simulating these events. Having a data-based model for this contribution (see Section 6.2.5) allows a method for the estimation of its rate by fitting to a kinematic distribution in a QCD multijet enriched region.

Events with a W boson accompanied by heavy flavor production (W+HF) constitute the majority of the b-tagged lepton+jets sample. These processes are $Wb\bar{b}$, $Wc\bar{c}$ and Wc, each with possibly more jets. While these events can be simulated using the ALPGEN generator, the theory uncertainties on the cross sections of these processes remain large compared with the size of the single top quark signal [153–160]. It is because of these large *a priori* uncertainties on the background predictions and the small signal-to-background ratio in the selected data samples that we must use advanced analysis techniques to purify further the signal, and also we must use the data itself to constrain the background rates.

The W+HF and the W+light flavor (W+LF) contributions in the tagged samples are derived from the total W+jets contribution in the pretag sideband region, where all selection cuts are

applied except the SecVtx tag requirement. For the former, the heavy flavor fraction calibrated in the W+1 jet sample and the tagging rate are applied to estimate the contribution in the tagged sample. For the latter, the mistag matrix probability is applied in order to estimate its contribution in the final sample.

7.2.1. QCD-Multijet Estimate

The expected number of QCD multijet events is first calculated in the pretag sample. The QCD multijet events are expected to fall at small values of the $\not\!\!E_T$ spectrum, since they have no true neutrino. Thus, removing the missing transverse energy cut creates a sample with a large QCD multijet component which can be fit to the data. The MC-based samples are fixed to their expected value while the normalization of the QCD multijet and W+jets samples is fit to the data. As Figure 7.1 shows, the resulting templates describe the data quite well. The fraction of QCD multijet events in the pretag sample, $F_{\rm QCD}^{\rm pretag}$, is then calculated counting the events that pass the missing transverse energy cut $\not\!\!E_T > 25$:

$$F_{\rm QCD}^{\rm pretag} = N_{\rm QCD}^{\rm pretag}/N_{\rm total}^{\rm pretag}. \tag{7.3}$$

The number of W+jets events in the pretag region, $N_{W+{\rm jets}}^{\rm pretag}$, is also extracted from the pretag fit. This number is used to extract the W+HF and W+LF contribution in the tagged region as described next. This provides a second fit to the $\not\!\!E_T$ distribution in the final tagged sample, where only the QCD multijet contribution is allowed to float. The resulting shape is shown in Figure 7.2 and is used to derive the QCD multijet fraction in the final sample. Because of the uncertainties in the tagging rate, template shape, and method estimations, the rate is given a generous systematic uncertainty of 40% in single-tagged events and 80% in double-tagged events.

7.2.2. W+Jets Estimate

The number of W+jets events is estimated by assuming that if the contribution of all other backgrounds (i.e. MC-based and QCD multijet) is accounted for, any remaining events must be from W+jets. It is safer to use the pretag sample to ensure that this estimate is statistically independent of the signal events. This sample is ten to twenty times larger than the final sample, and it is dominated by light jets. This gives an estimate for the overall normalization of the W+jets sample using pretag data:

$$N_{W+{\rm jets}}^{\rm pretag} = N_{\rm total}^{\rm pretag} \cdot (1 - F_{\rm QCD}^{\rm pretag}) - N_{\rm MC-fixed}^{\rm pretag}. \tag{7.4}$$

It is necessary now to differentiate between the W+HF and the W+LF components of the sample. The heavy flavor fraction, $F_{\rm HF} = \frac{N_{\rm HF}}{N_{W+{\rm jets}}}$, is calculated by looking directly at Monte Carlo

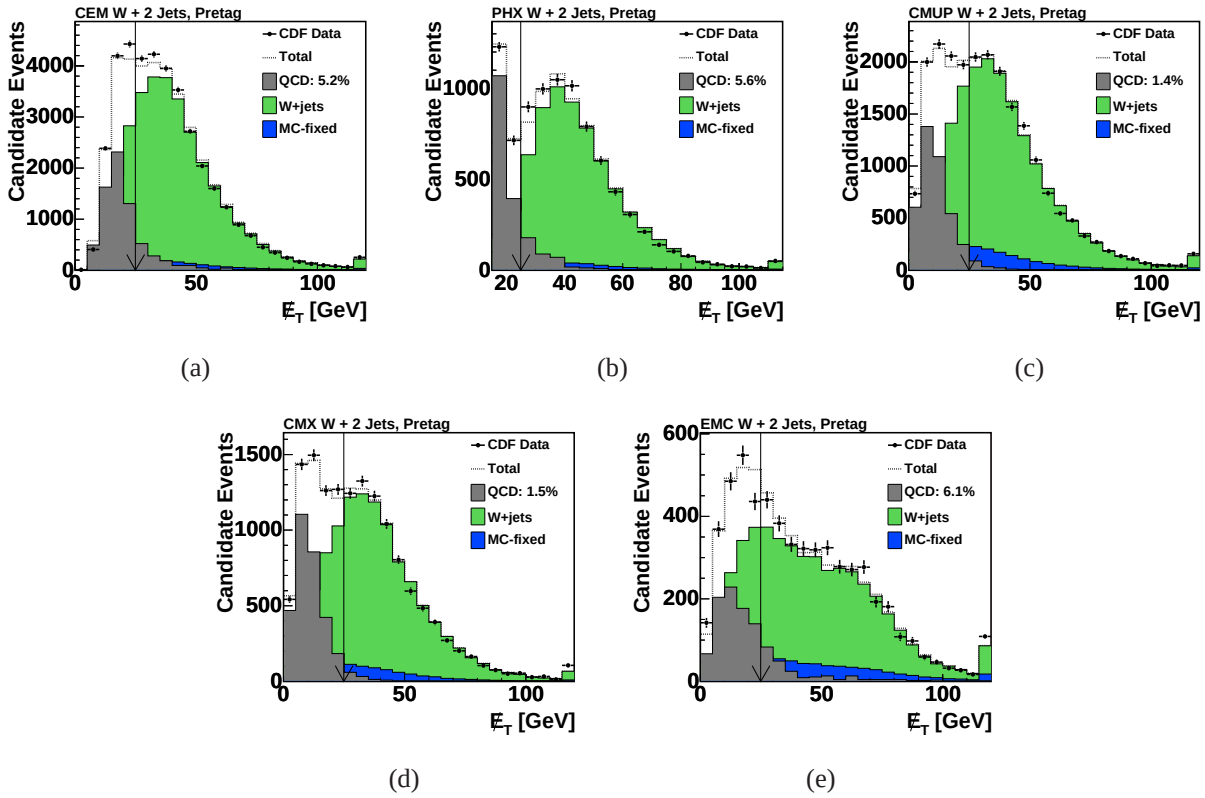


Figure 7.1: Fits to $\not\!\!E_{\rm T}$ distributions in the W+two jet pretag sample for the five different lepton types. The fraction of QCD multijet events is estimated from the fraction of the template above the $\not\!\!E_{\rm T}$ threshold shown.

events. However, the Monte Carlo simulation does not properly predict the heavy flavor fraction in this sample, therefore we apply an additional factor, $K_{\rm HF}$, to match the observed fraction in control data region (see next Section). This gives an estimate of the rate of W+HF contribution in the final tagged sample as follows:

$$N^{\rm tag}_{W+{\rm HF}} = \left(N^{\rm pretag}_{\rm total} \cdot \left(1 - F^{\rm pretag}_{\rm QCD}\right) - N^{\rm pretag}_{\rm MC-fixed}\right) \cdot F_{\rm HF} \cdot K_{\rm HF} \cdot \epsilon_{\rm tag}, \tag{7.5}$$

where $\epsilon_{\rm tag}$ is the corresponding tagging rate calculated using the same method as described in Section 7.1.

Once the heavy flavor has been properly estimated, the remainder of the pretag sample is assumed to be W+LF. This is by far the largest contribution to the pretag sample, which means that a substantial number of events will be mistakenly tagged. The estimation of the W+LF contribution in the final tagged sample uses the mistag matrix described in Section 4.3.5. The number of mistags, $N_{\rm mistag}$, from the pretag samples is obtained by applying the per-jet mistag rate corrected by the mistag asymmetry factors. This gives the following expression for the W+LF final estimate:

$$N^{\rm tag}_{W+{\rm LF}} = \left(N^{\rm pretag}_{\rm total} \cdot \left(1 - F^{\rm pretag}_{\rm QCD}\right) - N^{\rm pretag}_{\rm MC-fixed} - N^{\rm pretag}_{W+{\rm HF}}\right) \cdot \frac{N_{\rm mistag}}{N^{\rm pretag}_{\rm total}}. \tag{7.6}$$

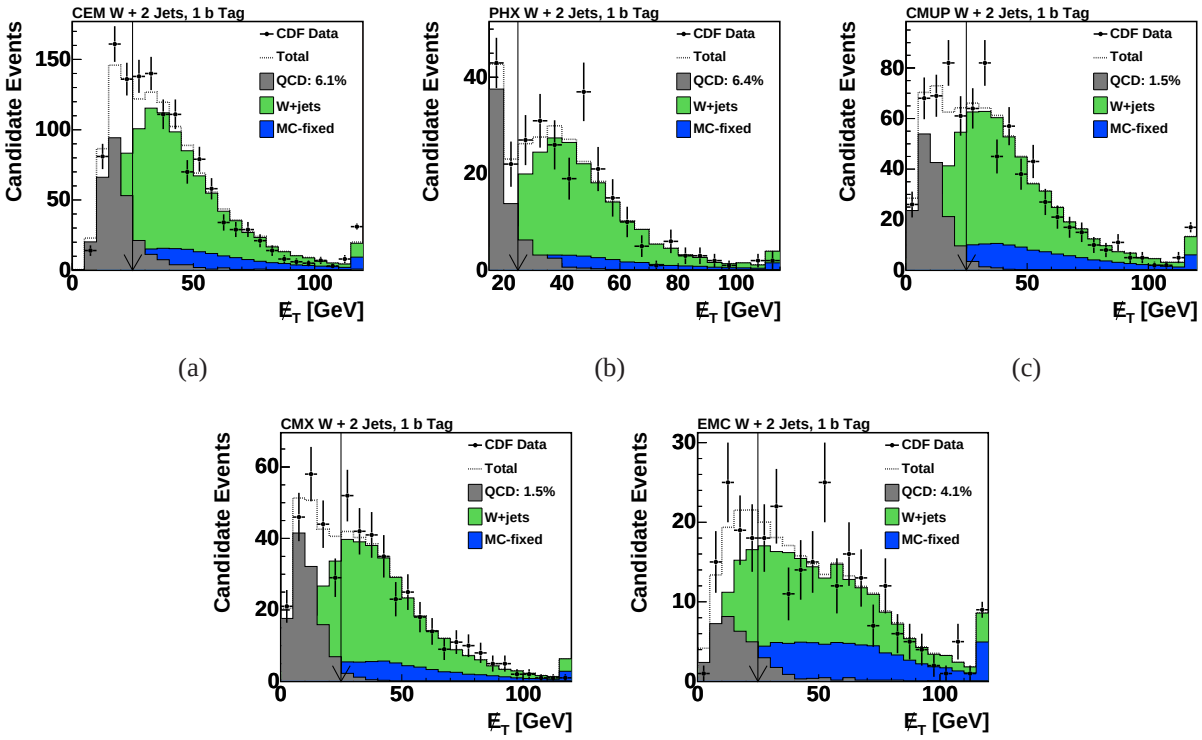


Figure 7.2: Fits to $\not\!\!E_{\mathrm{T}}$ distributions in the W+2 jet single-tagged sample for the five different lepton types. The fraction of QCD multijet events is estimated from the fraction of the template above the $\not\!\!E_{\mathrm{T}}$ threshold shown.

Heavy Flavor Fraction Calibration

As mentioned above, the fraction of W+jets events that contain heavy flavor quarks is not well predicted by our ALPGEN+PYTHIA model. In order to improve our modeling of these fractions, we perform fits to templates of flavor-separating variables in the b-tagged W+1 jet data sample, which contains a vanishingly small component of single top quark signal events and is not otherwise used in the final signal extraction procedure. This sample has high statistics and is almost entirely composed of W+jets events, making it ideal for the estimation of the heavy flavor content. We include the contribution of the MC-based backgrounds as separate templates, normalized to their SM expected rates, in the fits to the data. Care must be exercised in the estimation of the W+heavy flavor fractions, because fitting in the W+1 jet sample and using the fit values for the W+2 jet and W+3 jet samples involves an unavoidable extrapolation. We seek to estimate the bottom and charm fraction in these events with as many independent methods as possible and we assign generous uncertainties that cover the differences in the different estimations of the rates.

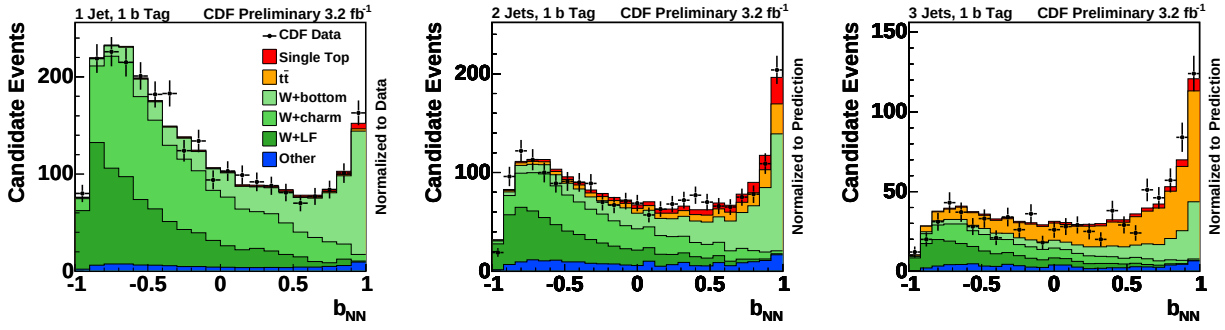
We use the jet-flavor separator b_{NN} described in Section 4.3.6 and ALPGEN+PYTHIA Monte Carlo samples to produce template distributions for W+light, $W + c$, and $W + b$ events. The

W+light template's rate is constrained by the data-derived mistag estimate within its uncertainty; the other W+jets templates' rates are not constrained. The Wbb and Wcc components are scaled together, as corresponding diagrams, with b and c quarks interchanged, contribute in the ALPGEN model, and we expect a similar correspondence for the leading processes in the data. We also let the Wc fraction float in the fit. The best fit in the $W+1$ jet sample is shown in Figure 7.3(a).

The fit indicates that the ALPGEN-predicted $Wbb+Wcc$ fraction must be multiplied by 1.4 ± 0.4 in order for the templates to match the data. In addition to the fit to the b_{NN} distribution, we also fit the W+heavy flavor fractions in the b-tagged W+1 jet sample with another variable, the reconstructed invariant mass of the secondary vertex. We perform this alternate fit in our standard b-tagged sample as well as in one with loosened b-tag requirements. The results are found to be consistent with the result using the b_{NN} variable.

The problem with the heavy flavor estimate is assumed to have to do with the difficult theoretical problem of the infrared divergence in the case of gluon splitting. Thus, the contribution of the $W + c$ + jets sample, which has no gluon splitting, is not scaled; only events with two heavy quarks in the final state are scaled by this factor. A recent measurement of the $W + c$ + jets cross section at CDF [161] is consistent with the ALPGEN calculation, reinforcing this view. Thus, the multiplicative factor of the Wc component is set to 1.0 ± 0.3 for the use in the two- and three-jet bins.

The 30% uncertainties assessed on the $Wbb + Wcc$ and Wc yields cover the differences in the measured fit values and also approximates our uncertainty in extrapolating this fraction to W+2 and 3 jet events. We check these extrapolations in the $W+2$ and 3 jet events as shown in Figures 7.3(b) and 7.3(c) where no additional fit is performed for this comparison. The rates and flavor compositions match very well with the observed data in these samples.



(a) (b) (c)

Figure 7.3: Distributions of the jet flavor separator b_{NN}. Panel (a) shows the fit to the W+1 jet data sample allowing the b, c and light-flavor components float. Panels (b) and (c) compare the data and the corresponding predictions in the W+2 jet and W+3 jet samples.

7.3. Candidate Event Yield Prediction

The results of the candidate event yield estimate in the different signal regions, as well as the observed number of events in data, are summarized in Table 7.2. All quoted uncertainties include

Process	W + 2 jets		W + 3 jets	
	1 b-tag	2 b-tag	1 b-tag	$\geq$2 b-tag
$Wb\bar{b}$	581.1 ± 175.1	75.9 ± 23.6	173.9 ± 52.5	27.4 ± 8.5
$Wc\bar{c}$	288.5 ± 89.0	3.7 ± 1.2	95.7 ± 29.4	2.4 ± 0.8
Wc+jet	247.3 ± 76.2	3.2 ± 1.0	50.8 ± 15.6	1.3 ± 0.4
Mistags	499.1 ± 69.1	2.2 ± 0.6	150.3 ± 21.0	1.6 ± 0.4
QCD multijet	88.4 ± 35.4	2.3 ± 0.9	35.4 ± 14.1	0.2 ± 0.1
$t\bar{t}$	167.6 ± 24.0	36.4 ± 6.0	377.3 ± 54.8	104.7 ± 17.3
Diboson	83.3 ± 8.5	5.0 ± 0.6	28.1 ± 2.9	2.0 ± 0.3
Z+jets	34.8 ± 5.3	1.7 ± 0.3	14.6 ± 2.2	1.0 ± 0.2
Total Background	1990.1 ± 349.6	130.4 ± 26.8	926.1 ± 113.4	140.6 ± 19.7
s-channel	45.3 ± 6.4	12.8 ± 2.1	14.7 ± 2.1	4.5 ± 0.7
t-channel	85.3 ± 12.6	2.4 ± 0.4	22.7 ± 3.3	3.5 ± 0.6
Total Prediction	2120.5 ± 350.1	145.6 ± 26.9	963.4 ± 113.5	148.6 ± 19.7
Observation	2090	139	920	166

Table 7.2: Expected signal and background event yield and observed number of events in data. The background prediction is consistent with the number of data events; however, the systematic uncertainty on the background prediction is far larger than the expected single-top signal.

the systematic uncertainties on the theoretical cross section calculations, the trigger efficiency, the lepton ID efficiency, the b-tagging scale factor, the mistag matrix uncertainties, the heavy-flavor K-factor, the QCD multijet estimate, the primary vertex position scale factor, and the luminosity, where appropriate.

Figure7.4 compares the final result of the candidate event yield estimate in terms of jet multiplicity with the number of observed collision events. Even though events with one or four jets are not considered for the search of single top-quark production, the outcome in those data sidebands is shown. The uncertainty corresponds to the overall sum of all contributing processes derived in the procedure of the background estimate method . The signal expectation is much smaller than the overall uncertainty on all background processes, making a simple counting experiment impossible.

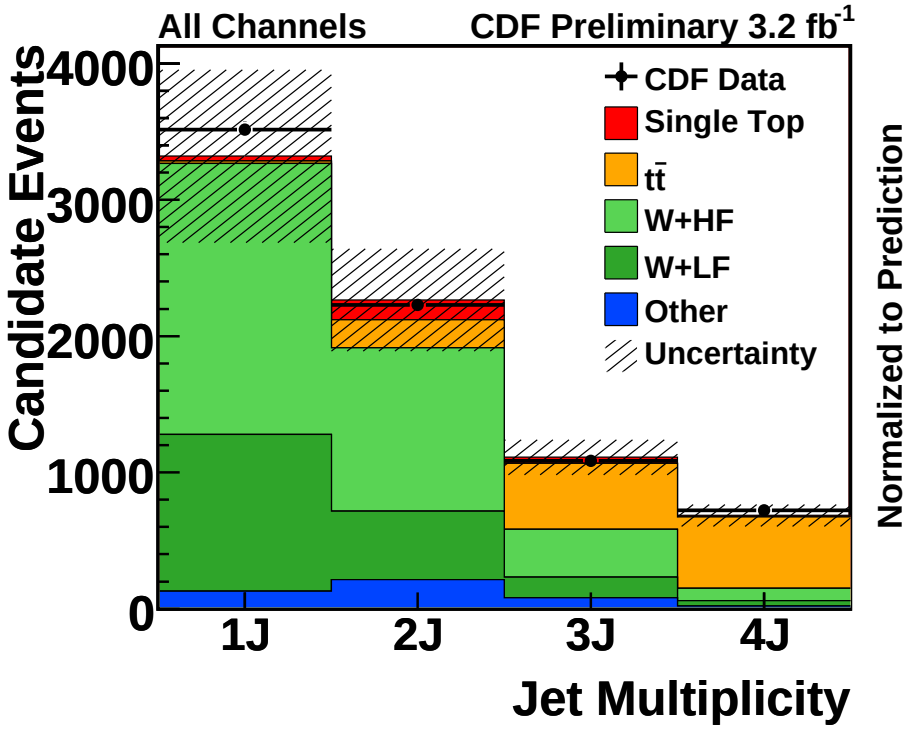


Figure 7.4: Predicted and observed W+jets events as a function of the jet multiplicity. The data are indicated with points, and the colored histograms show the signal and background predictions which are stacked to form the total prediction. The systematic uncertainty is far too large to use a simple counting experiment to measure the single top quark cross section.

7.4. Validation of Monte Carlo Simulation

In the previous section it has been shown the good agreement between the expected and observed number of events passing the candidate event selection. Because multivariate analyses depend so heavily on properly simulating events, it is very important to further validate the modeling of the distributions in Monte Carlo by checking them with the data. Thousands of data and Monte Carlo distributions are compared for several variables in the signal samples and in many control regions. These control regions include samples in which no jets have been b-tagged to test the W+light jets shapes, W+1 jet events to examine W+heavy flavor fraction and shapes, and W+4 jet events where $t\bar{t}$ is dominant and thus can be checked.

A sample of the validation plots we examine is shown in Figures 7.5-7.24. In these plots, data is compared to the signal and background contributions which are stacked to form the total prediction. Some basic kinematic quantities are shown in W+2 or 3 jets with zero or at least one b-tag for each lepton type. The close match of the distributions gives confidence in the results. Other distribution will be shown in the next Chapter.

Out of the thousands of distributions checked for discrepancies, only two distributions in the untagged W+jets data were found to be poorly simulated by our Monte Carlo model: the pseudo-rapidity of the lowest energy jet in both W+2 jet and W+3 jet events and the distance $\Delta R(j_1,\ j_2)$ between the two jets in the η-ϕ space in W+2 jet events. These discrepancies are used to estimate systematic uncertainties on the shapes of our final discriminant variables. These distributions and the discussion of associated systematic uncertainties are given in Section 9.1.

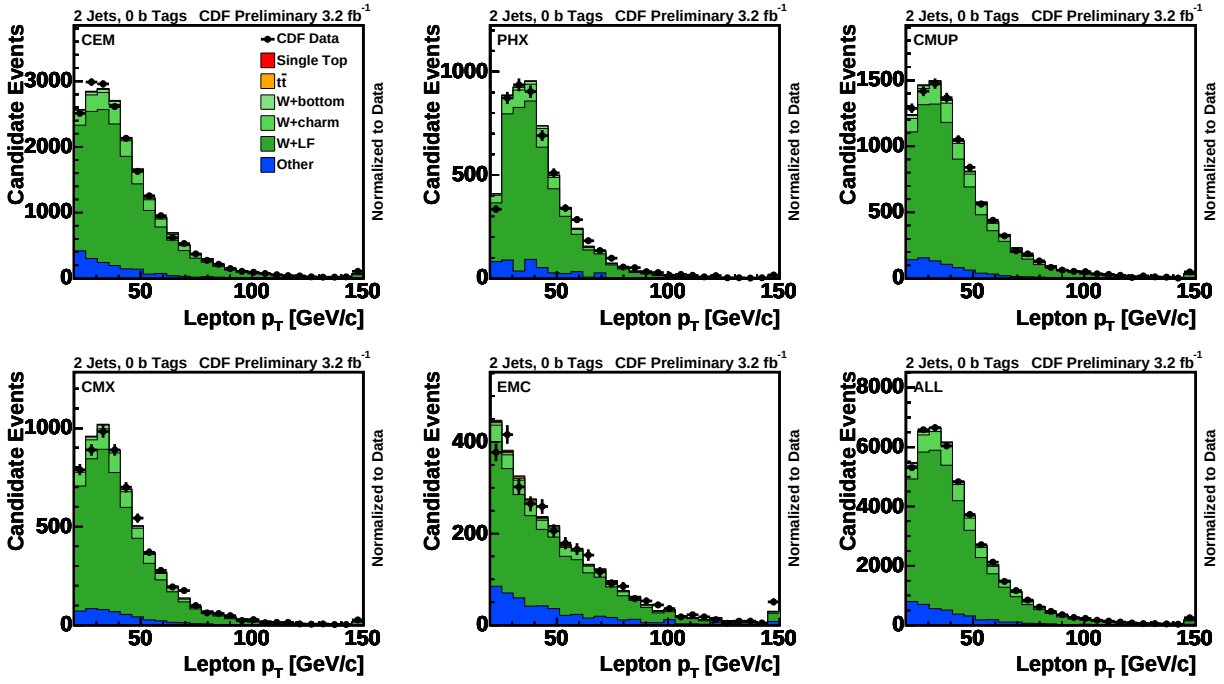


Figure 7.5: Validation plots comparing data and MC for the lepton transverse momentum in W+2 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

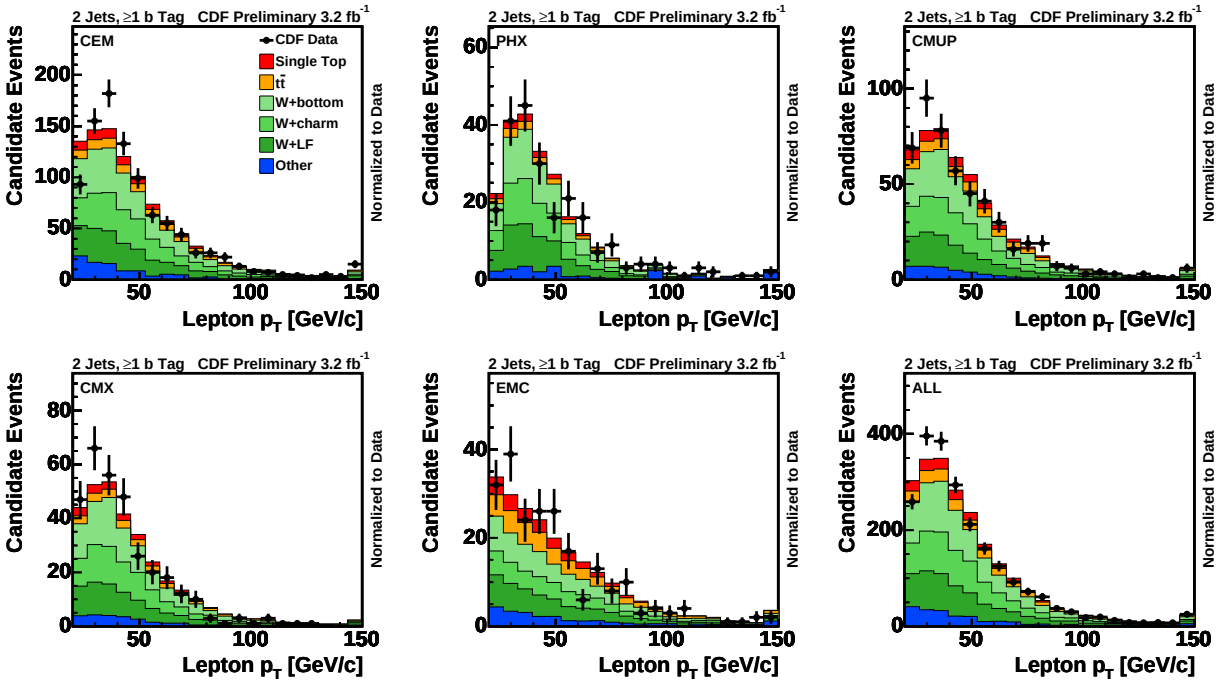


Figure 7.6: Validation plots comparing data and MC for the lepton transverse momentum in W+2 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

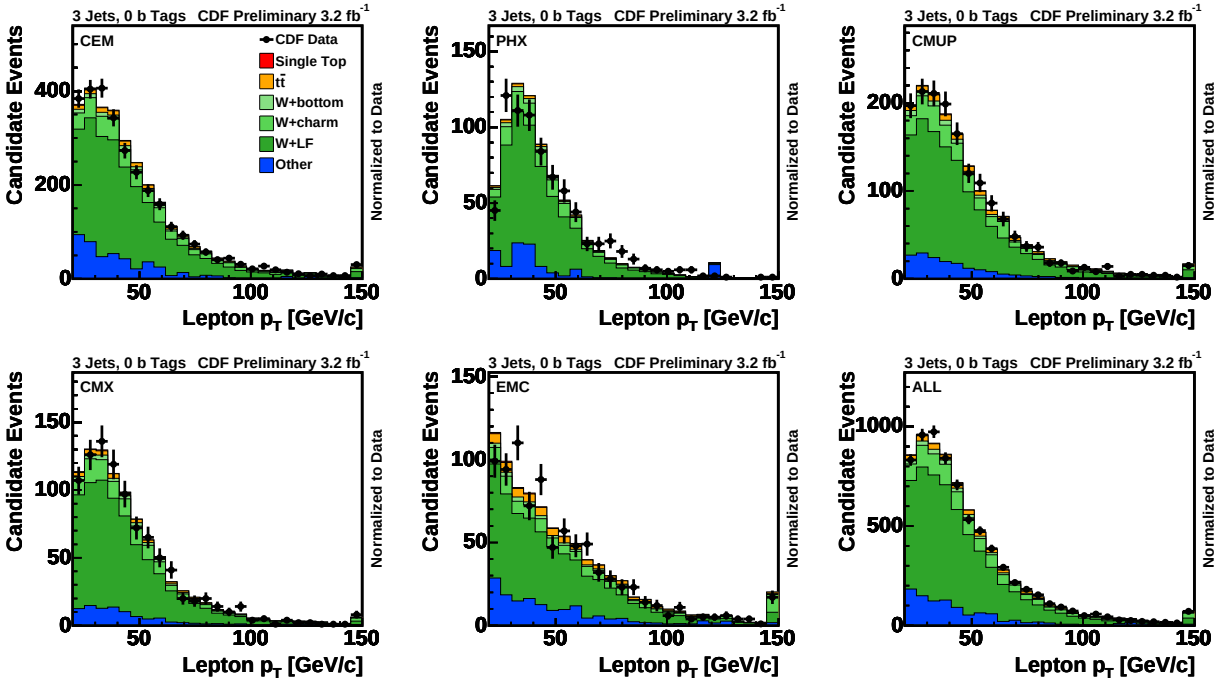


Figure 7.7: Validation plots comparing data and MC for the lepton transverse momentum in W+3 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

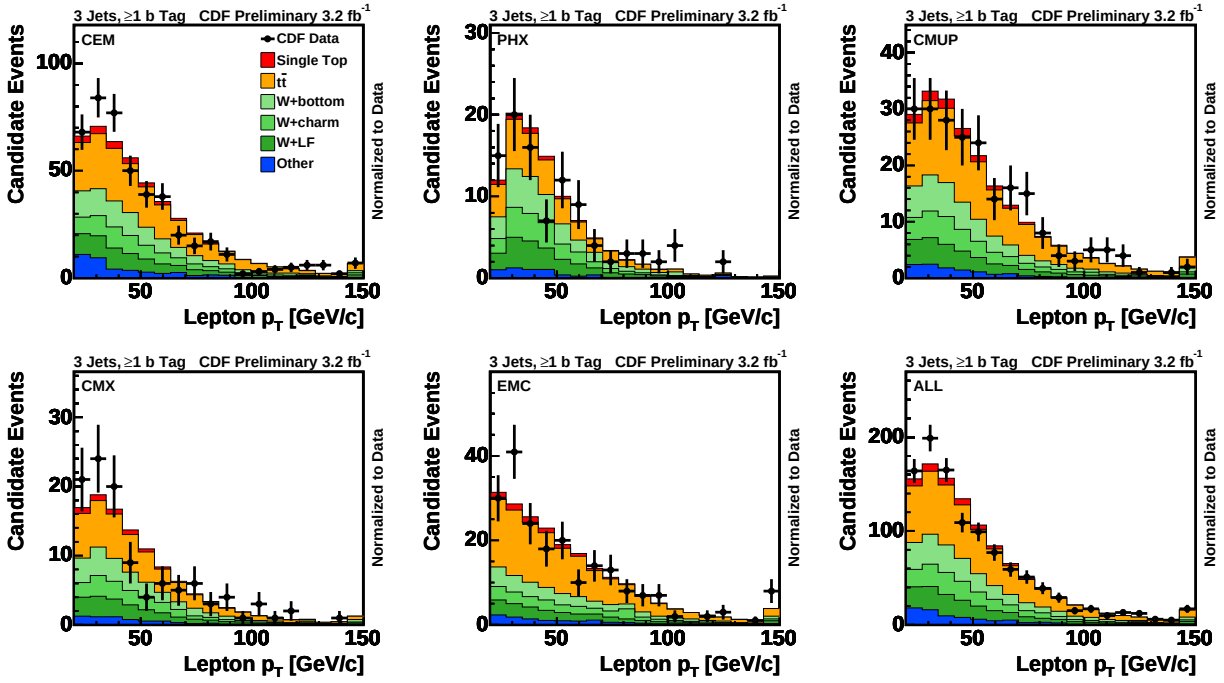


Figure 7.8: Validation plots comparing data and MC for the lepton transverse momentum in W+3 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

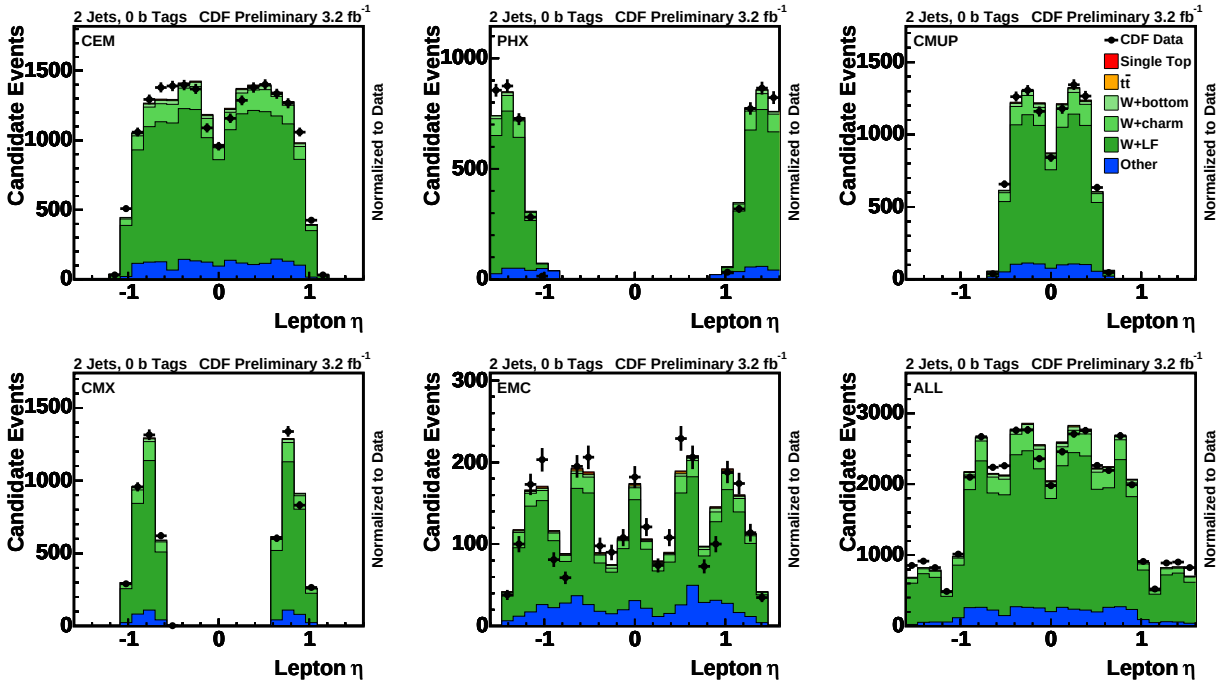


Figure 7.9: Validation plots comparing data and MC for the lepton pseudorapidity in W+2 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

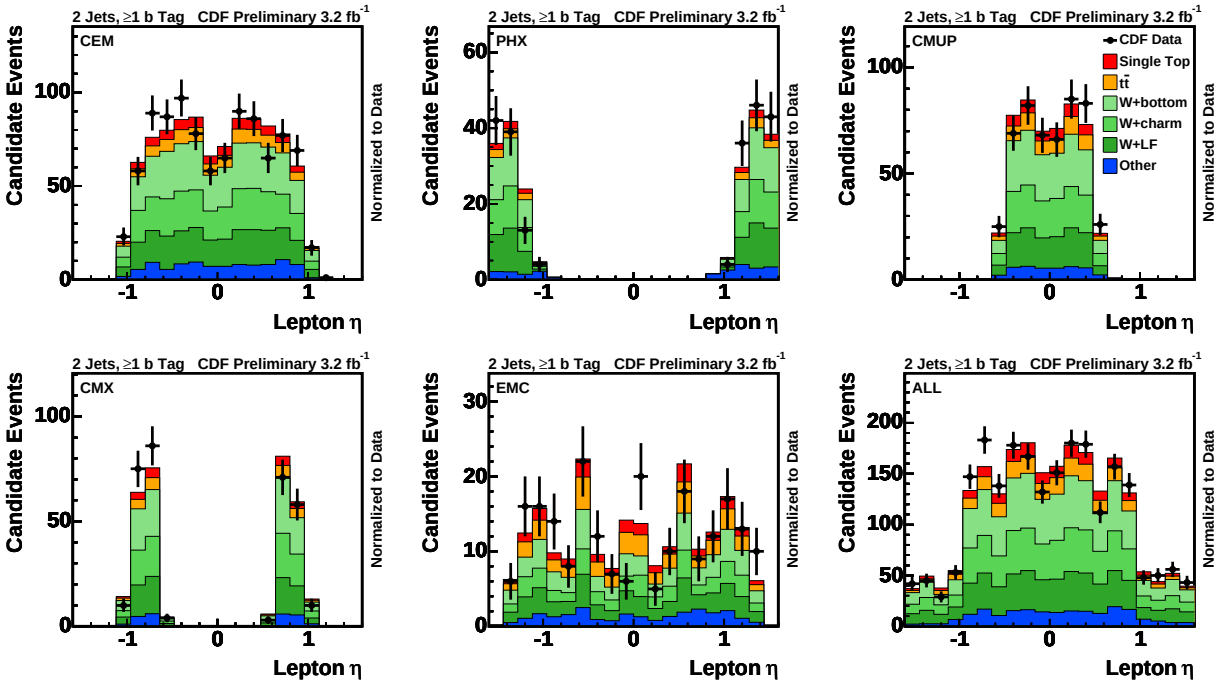


Figure 7.10: Validation plots comparing data and MC for the lepton pseudorapidity in W+2 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

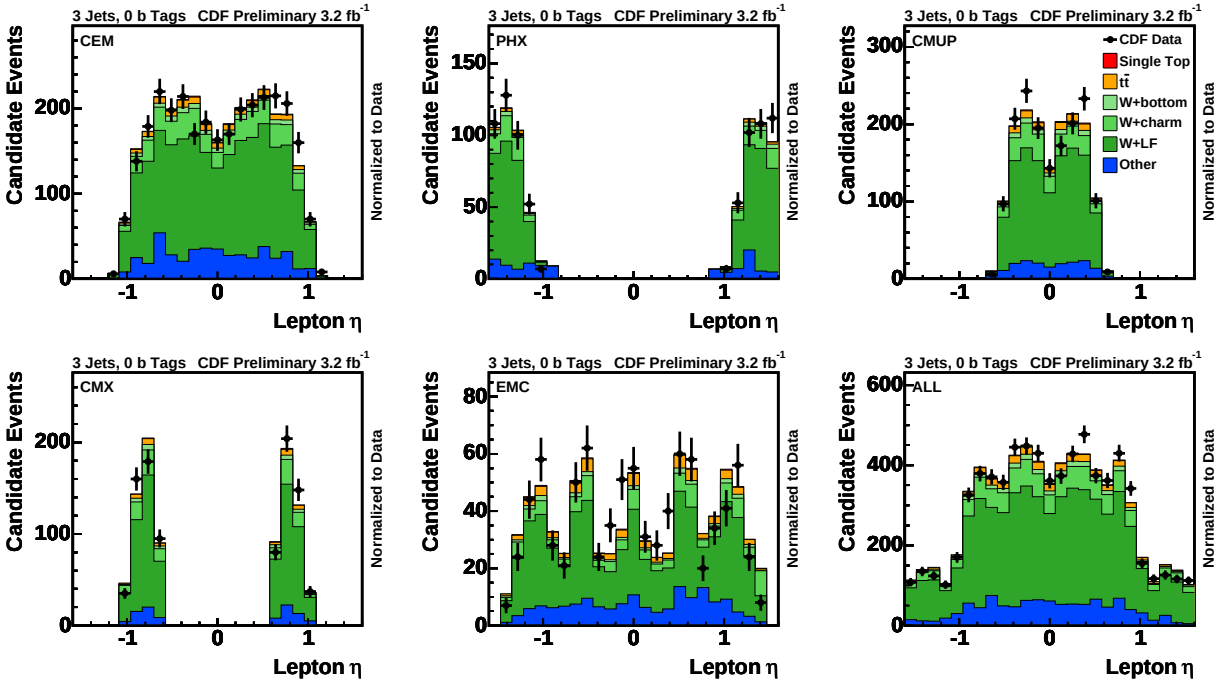


Figure 7.11: Validation plots comparing data and MC for the lepton pseudorapidity in W+3 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

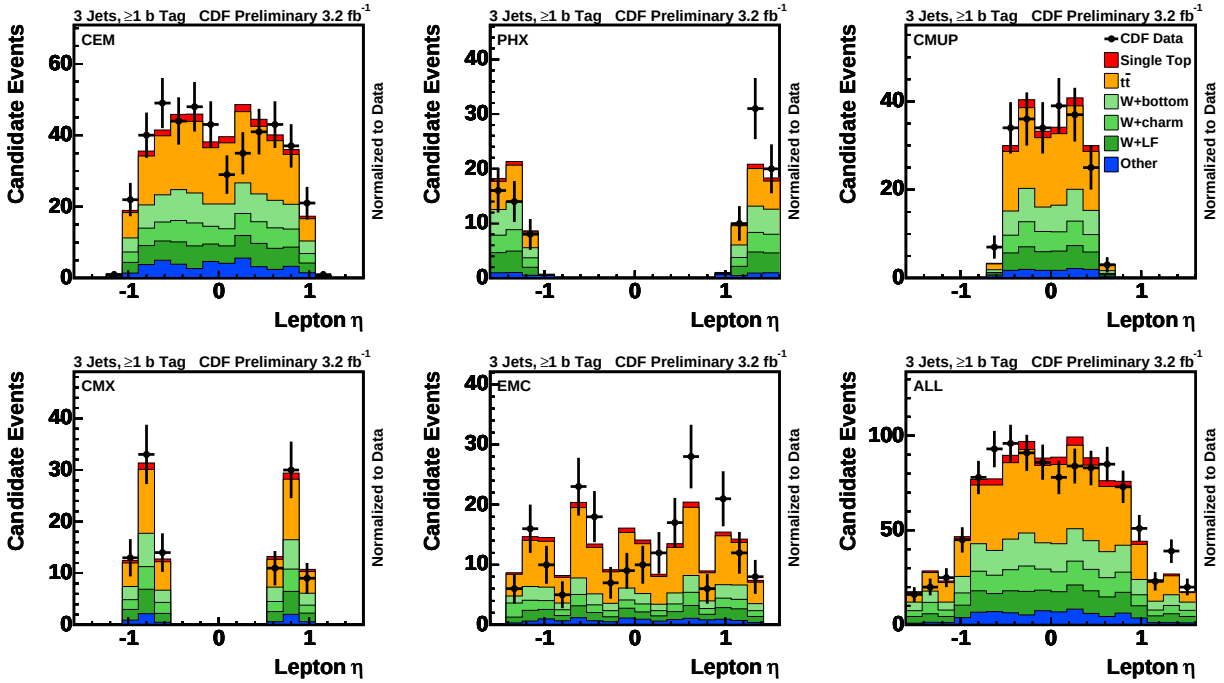


Figure 7.12: Validation plots comparing data and MC for the lepton pseudorapidity in W+3 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

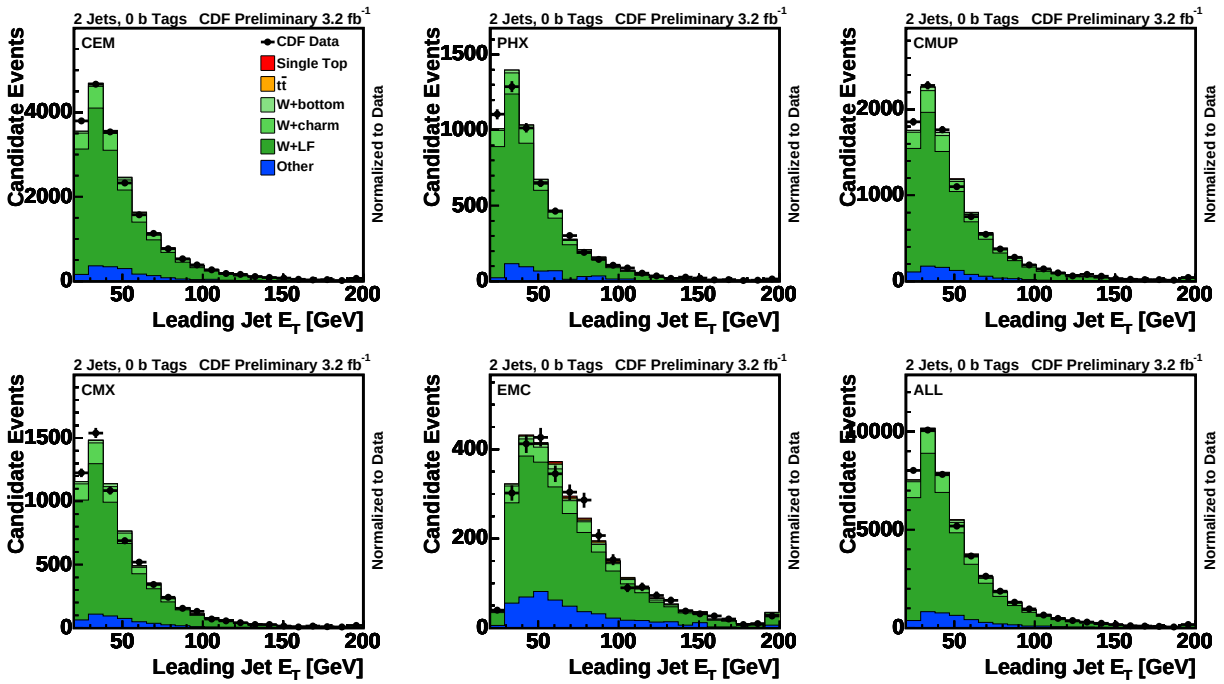


Figure 7.13: Validation plots comparing data and MC for the leading jet transverse energy in W+2 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

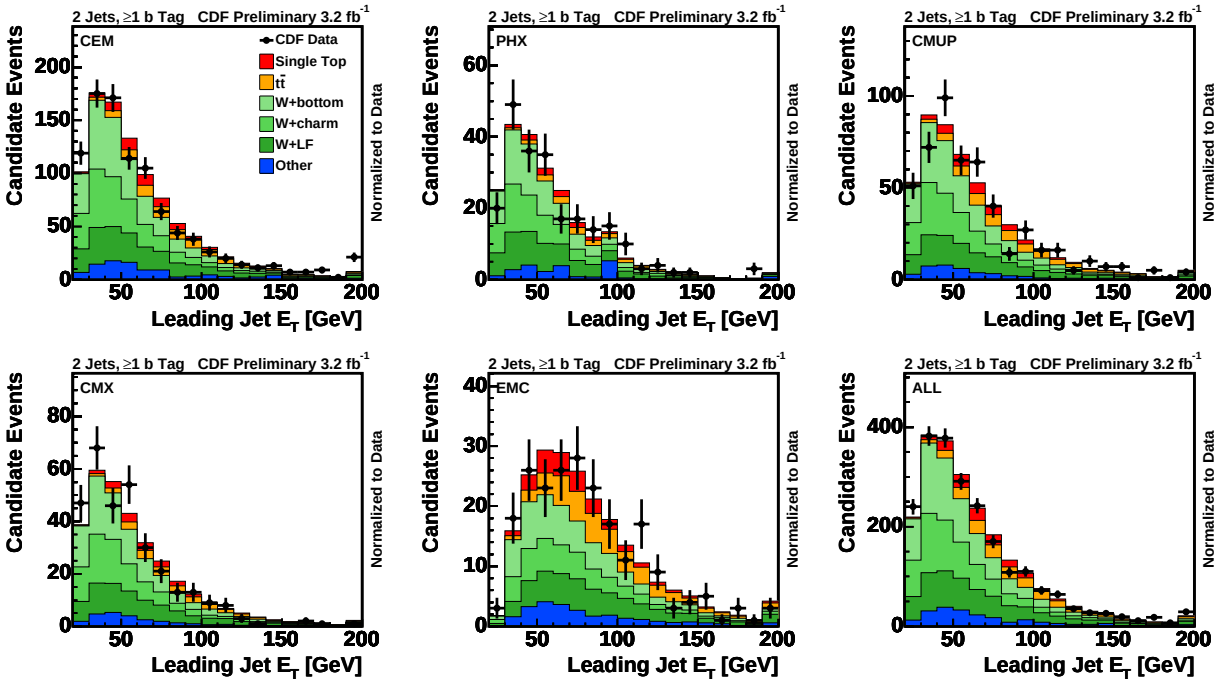


Figure 7.14: Validation plots comparing data and MC for the leading jet transverse energy in W+2 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

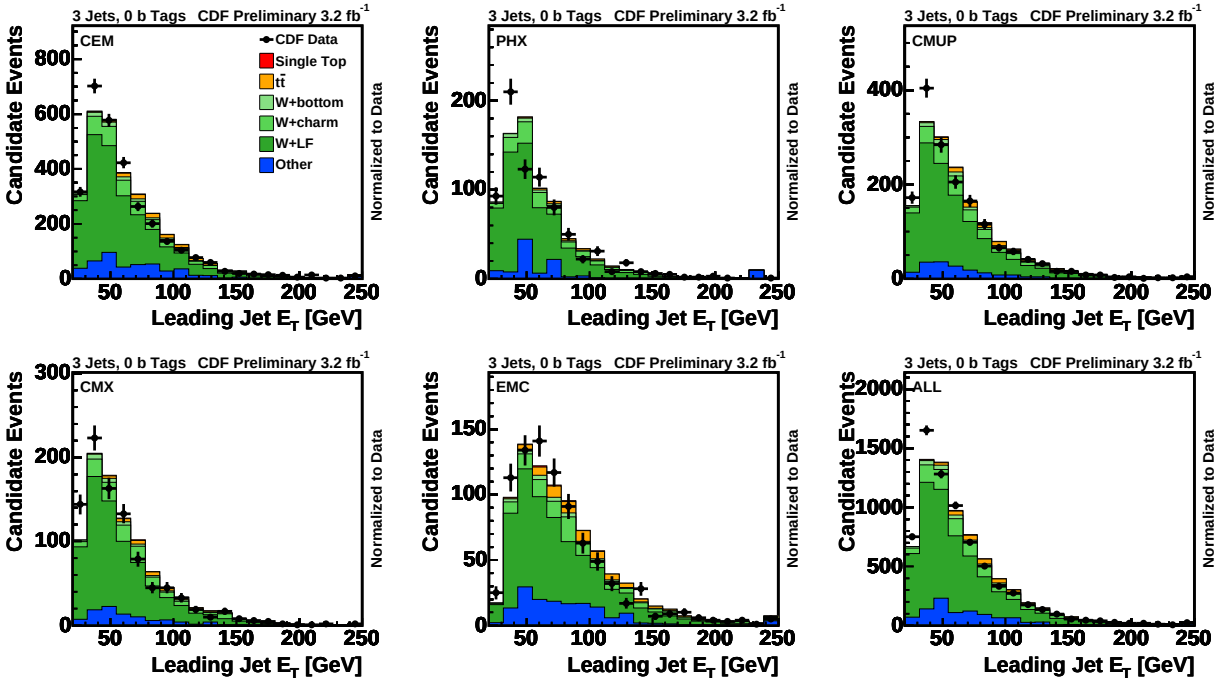


Figure 7.15: Validation plots comparing data and MC for the leading jet transverse energy in W+3 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

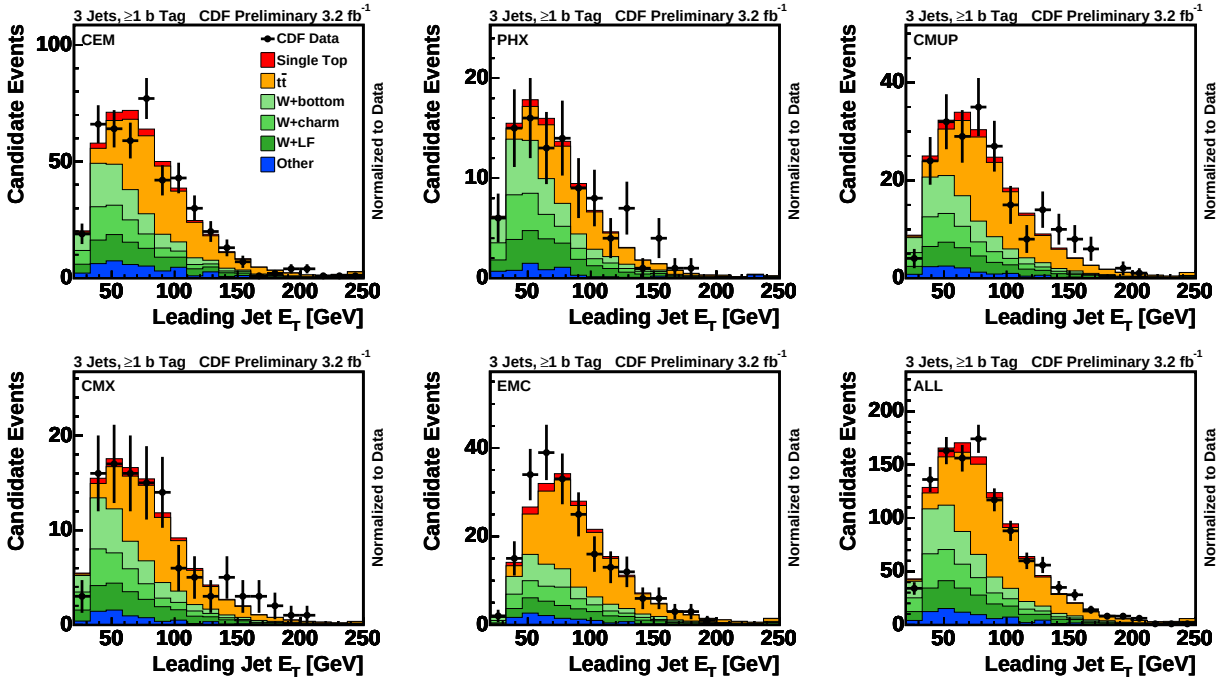


Figure 7.16: Validation plots comparing data and MC for the leading jet transverse energy in W+3 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

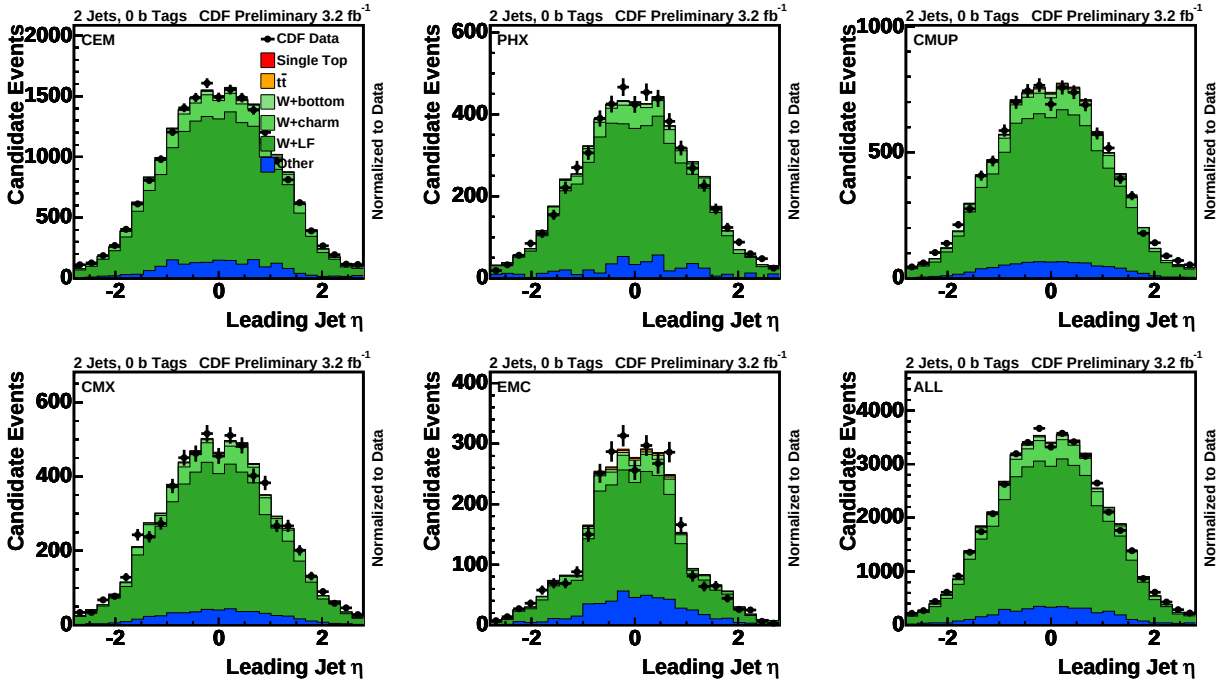


Figure 7.17: Validation plots comparing data and MC for the leading jet pseudorapidity in W+2 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

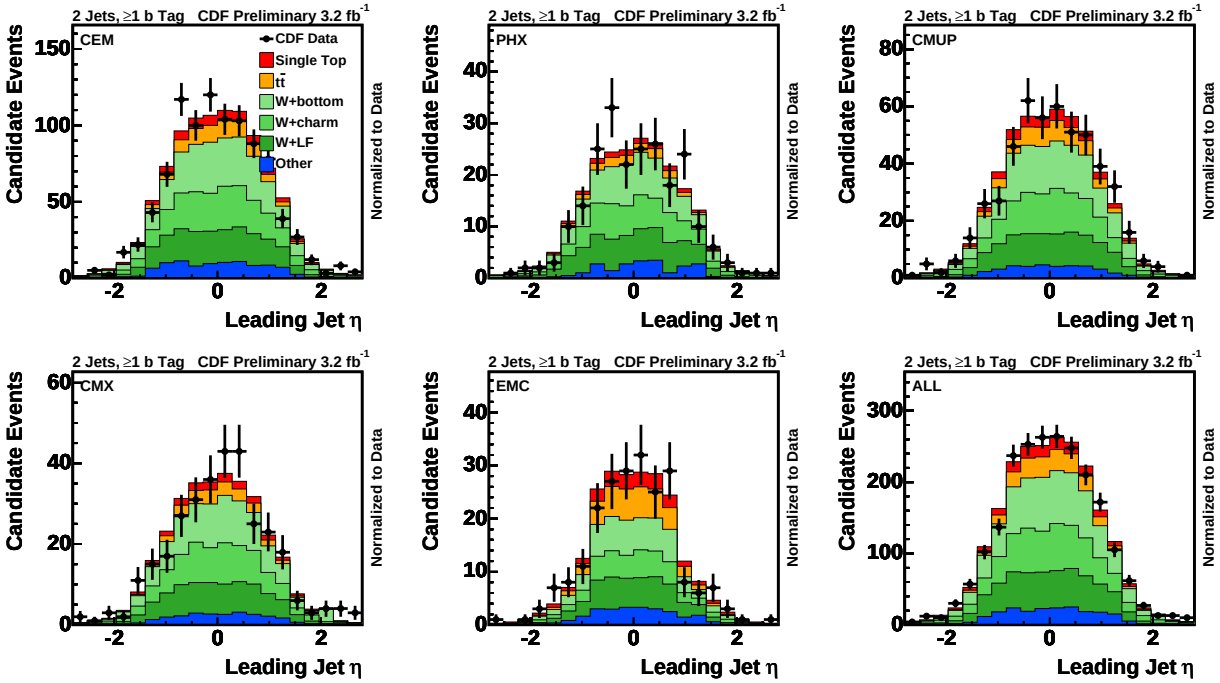


Figure 7.18: Validation plots comparing data and MC for the leading jet pseudorapidity in W+2 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

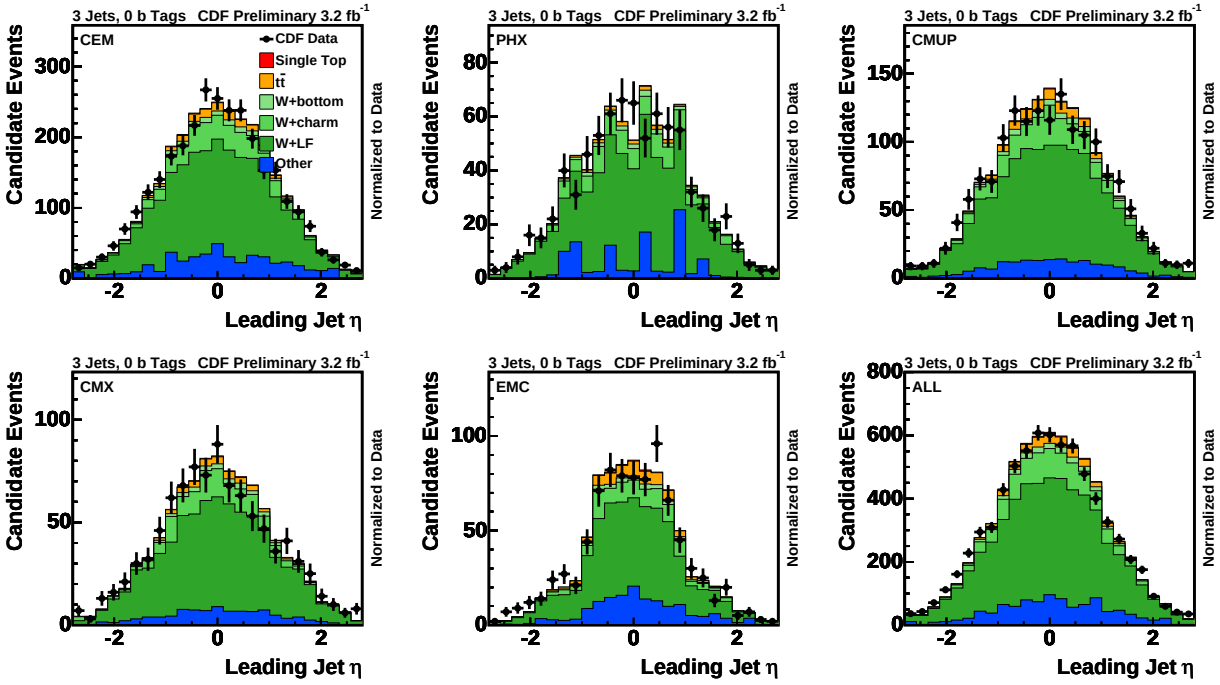


Figure 7.19: Validation plots comparing data and MC for the leading jet pseudorapidity in W+3 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

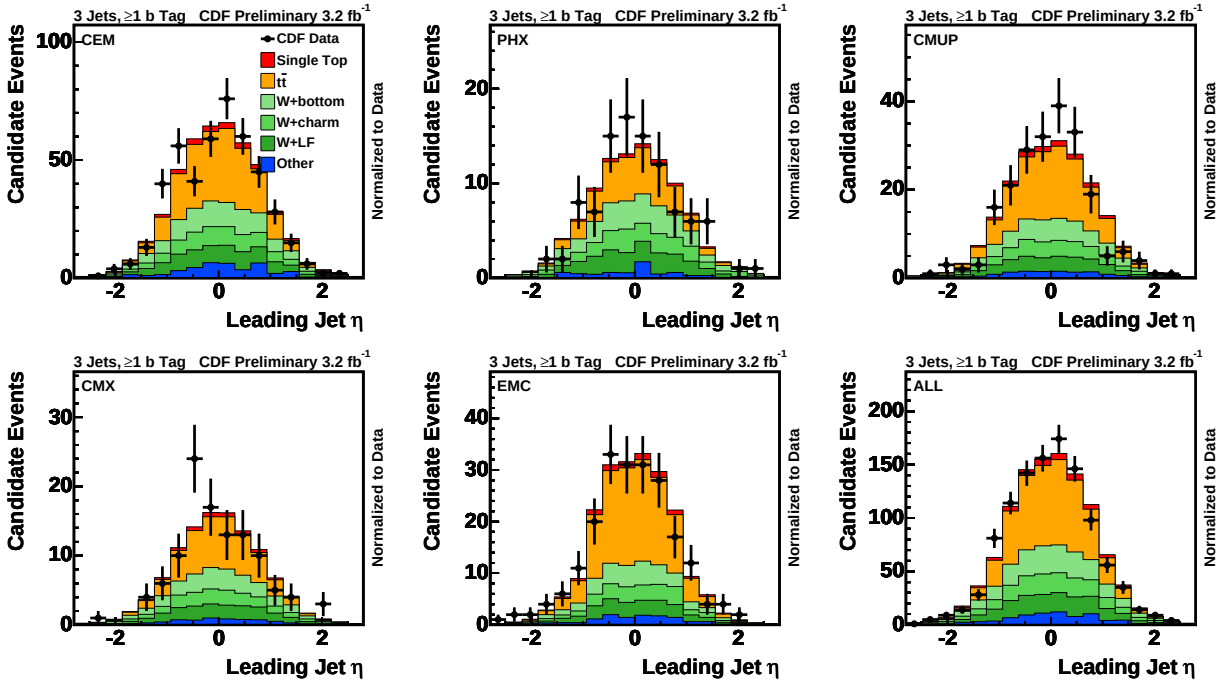


Figure 7.20: Validation plots comparing data and MC for the leading jet pseudorapidity in W+3 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

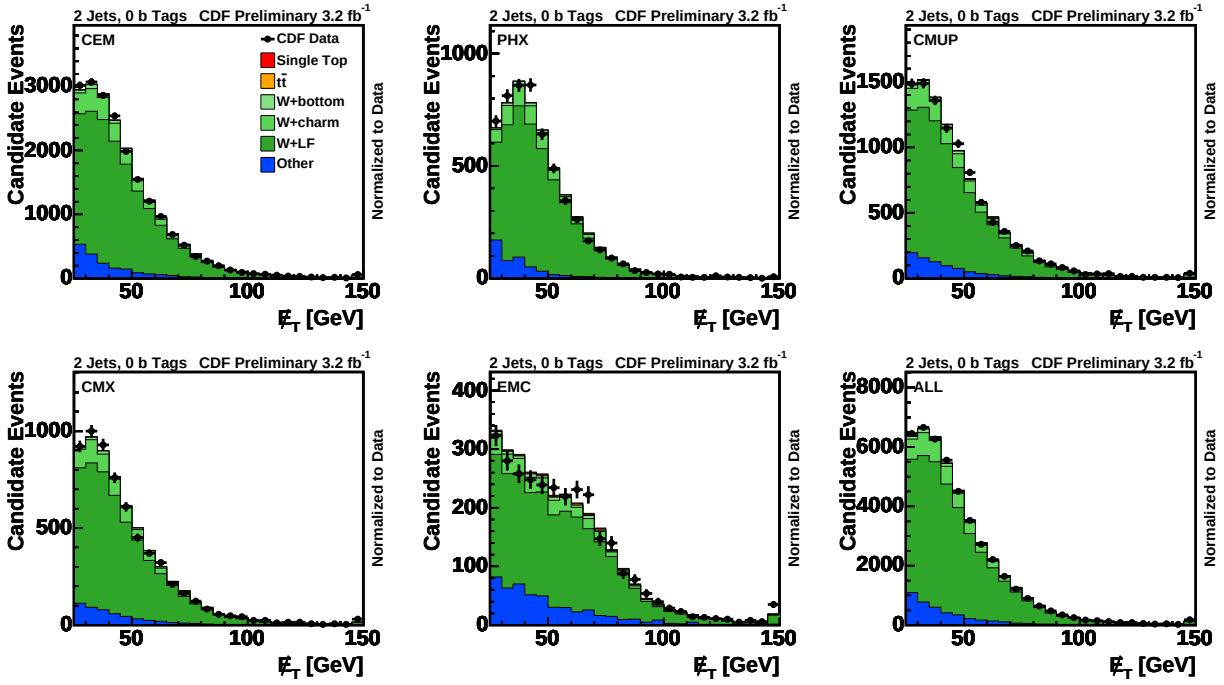


Figure 7.21: Validation plots comparing data and MC for the missing transverse energy in W+2 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

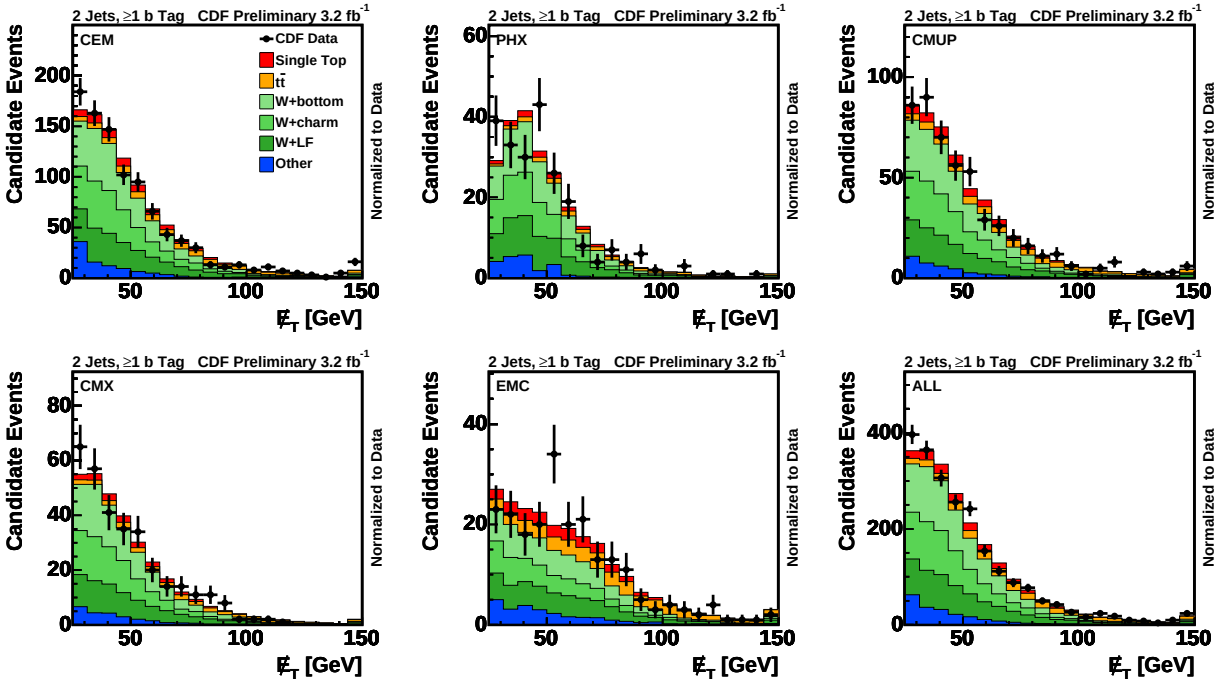


Figure 7.22: Validation plots comparing data and MC for the missing transverse energy in W+2 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

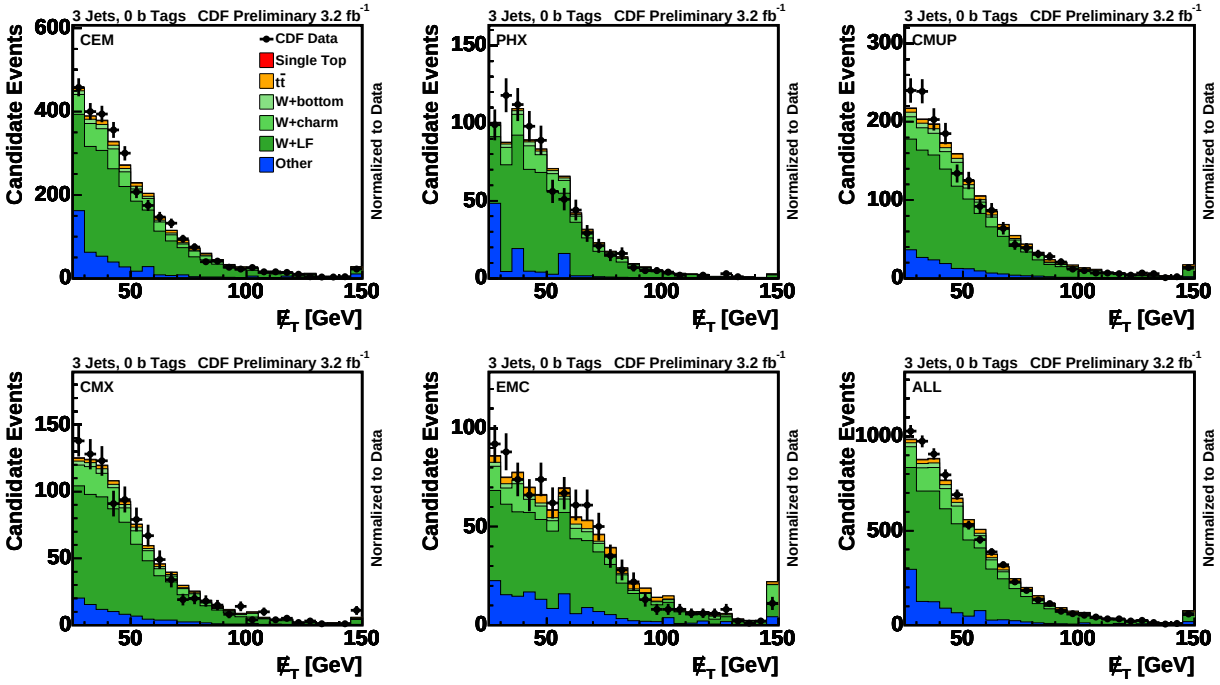


Figure 7.23: Validation plots comparing data and MC for the missing transverse energy in W+3 jet events with zero b-tags. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

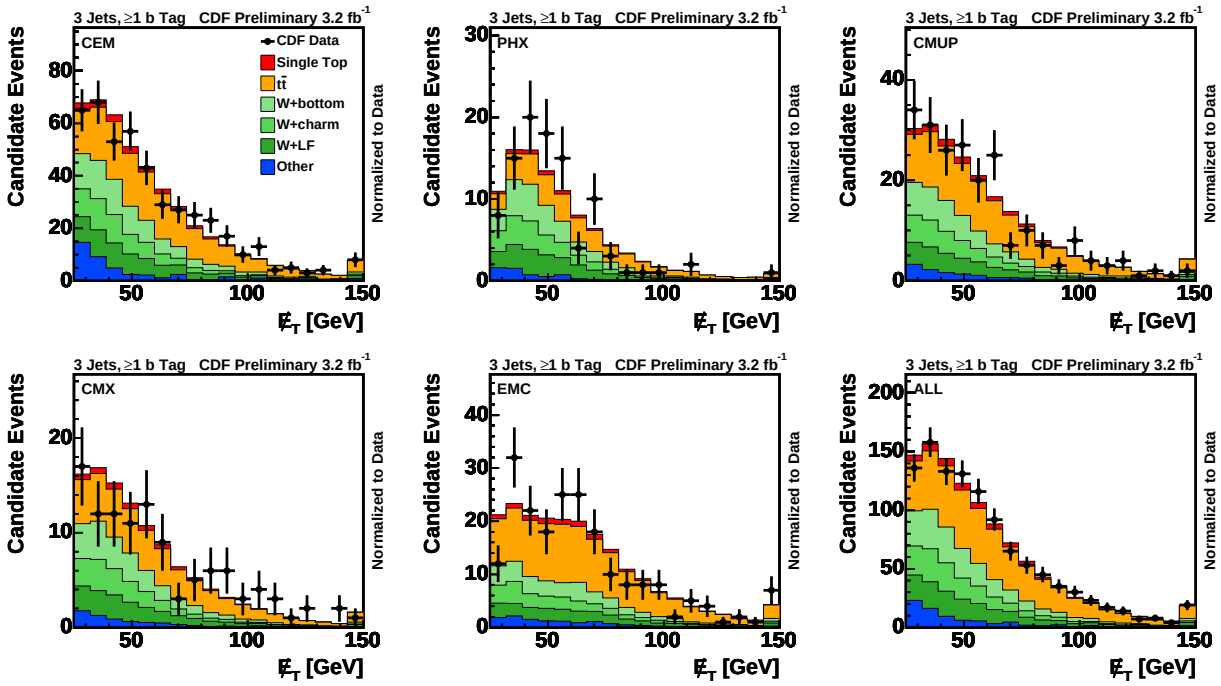


Figure 7.24: Validation plots comparing data and MC for the missing transverse energy in W+3 jet events with at least one b-tag. Events are shown for CEM, PHX, CMUP, CMX, EMC, and for all lepton types together, respectively, from left to right, top to bottom.

Chapter 8

MULTIVARIATE CLASSIFIER

The search for single top production and the measurement of its cross section present substantial experimental challenges. Compared with the search for $t\bar{t}$ production, the search for single top suffers from a lower SM production rate and a larger background. The most serious challenge arises from the systematic uncertainty on the background prediction, which is approximately three times the size of the expected signal. Simply counting events which pass our event selection will not yield a measurement of the single top quark cross section no matter how much data are accumulated because the systematic uncertainty on the background is so large. In fact, in order to have sufficient sensitivity to expect to observe a signal at the 5σ level, the systematic uncertainty on the background must be less than one-fifth of the expected signal rate.

Further separation of the signal from the background is required. Events that are classified as being more signal-like are used to test for the presence of single top quark production and measure the cross section, and more background-like events are used to improve our knowledge of the rates of background processes. In order to optimize the sensitivity to single top quark, we construct a discriminant function based on kinematic and b-tag properties of the events in order to classify them in a continuous spectrum that runs from very signal-like events to very background-like events. These distributions are then fit to the background and signal+background predictions, allowing uncertain parameters to float, as described in Section 9.2.1.

While there are many distinctive features of a single-top signal, no single variable is sufficiently sensitive to extract the signal with our current statistics. This requires the use of a more powerful technique that combines the discrimination power of many variables. In this thesis, a Boosted Decision Tree (BDT) technique is utilized, constructed and trained using the TMVA package [162].

8.1. Discriminating Variables

To separate signal events from background events, we look for features of the event that differ between signal and background. Events from single-top production have distinctive energy and angular features. Some of the most sensitive variables include:

- $Q \times \eta$: the charge of the lepton times the pseudorapidity of the untagged jet. If there is more than one untagged jet, the most energetic is chosen (see Figure 8.1). Large $Q \times \eta$ is characteristic of t-channel single-top events, because the light quark in the initial state usually carries most of the momentum of the collision and therefore ends up going far forward in the detector in a direction that is correlated with the charge of the lepton. This correlation comes from the fact that the lepton charge determines whether in the event a top or an antitop quark was produced, and single top quark production is most likely to be initiated by a u quark in the proton direction, while single top antiquark is most likely to be initiated by a $\bar{u}$ quark in the antiproton direction.

- $M_{\ell\nu b}$: the reconstructed top mass, i.e., the invariant mass of the lepton, neutrino, and quark from from the top decay (see Figure 8.2). The neutrino four-momentum is reconstructed as described in Section 4.3.4. The SecVtx tagged jet is assigned to be the b-jet from the top quark decay, in case of more than one as the b-tag jet, if there is more than one b-tagged jet in the event, the one with largest $Q \times \eta$ is picked for the top reconstruction. This distribution is expected to peak at the mass of the top quark for events containing one real top quark.

- H_{T}: the scalar sum of the transverse energies of the jets, charged lepton, and $\not\!\!E_T$ in the event (see Figure 8.3). This quantity is much larger for top events than for W+jet events.

- M_{jj}: the dijet mass, i.e., the invariant mass of the two jets (see Figure 8.4). If more than two jets are identified in the event, the two leading jets are used. This is substantially higher for events coming from top quarks.

- M_T^W: the transverse mass of the reconstructed W boson (see Figure 8.5). This variable is defined as $M_T^W = \sqrt{2p_T^\ell p_T^\nu(1-\cos\Delta\phi)}$, where p_T^ℓ and p_T^ν are the transverse momenta of the lepton and neutrino, and $\Delta\phi$ is the azimuthal angle between them. The missing energy is assumed to come from the undetected neutrino, therefore $p_T^\nu = \not\!\!E_T$. For events not containing a real W boson or events in which the missing transverse energy is not due to the neutrino, the shape of this distribution is expected to be different from the real W transverse mass distribution.

- b_{NN}: the jet flavor separator described in Section 4.3.6.

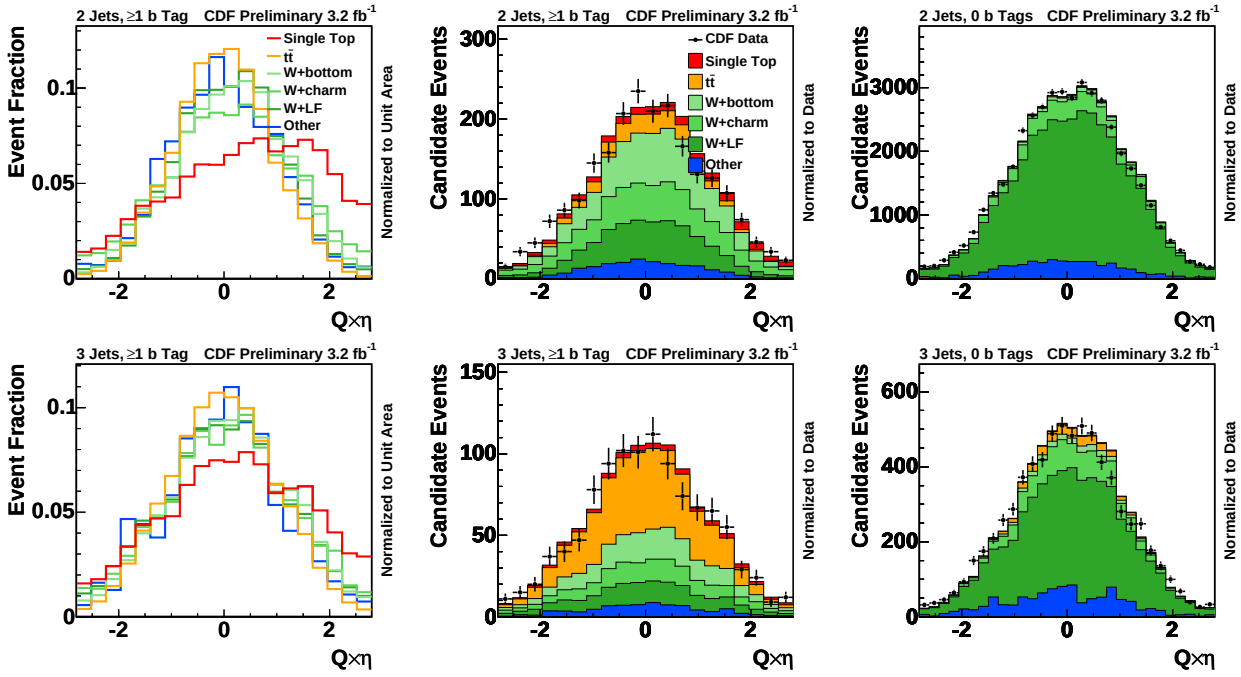


Figure 8.1: Templates (left) and validation plots comparing data and MC for the discriminating variable $Q \times \eta$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle corresponds to events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

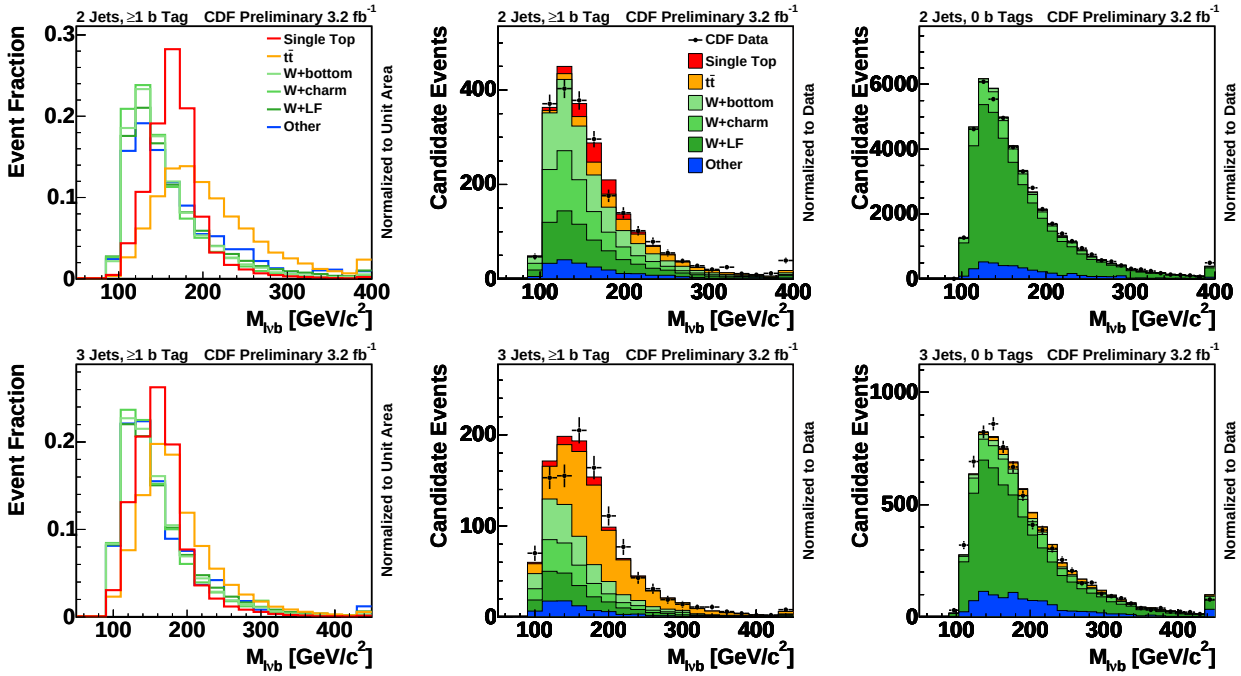


Figure 8.2: Templates (left) and validation plots comparing data and MC for the discriminating variable $M_{\ell\nu b}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle corresponds to events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

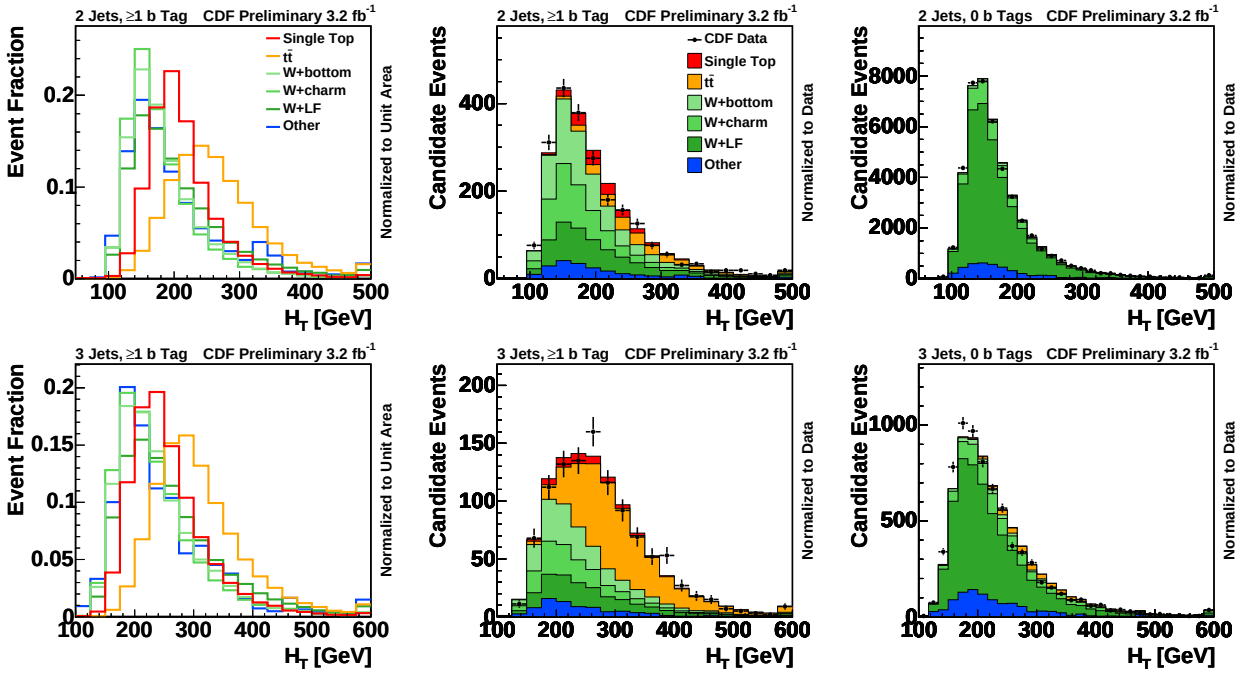


Figure 8.3: Templates (left) and validation plots comparing data and MC for the discriminating variable H_T. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle corresponds to events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

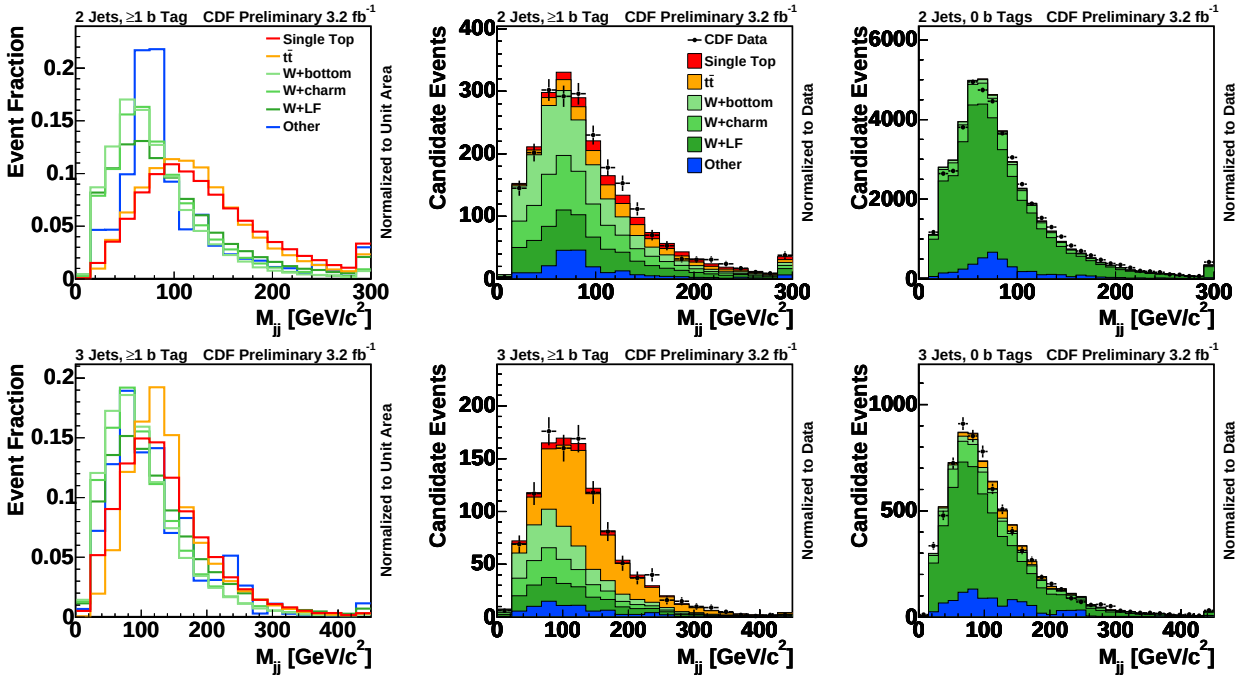


Figure 8.4: Templates (left) and validation plots comparing data and MC for the discriminating variable M_{jj}. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle corresponds to events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

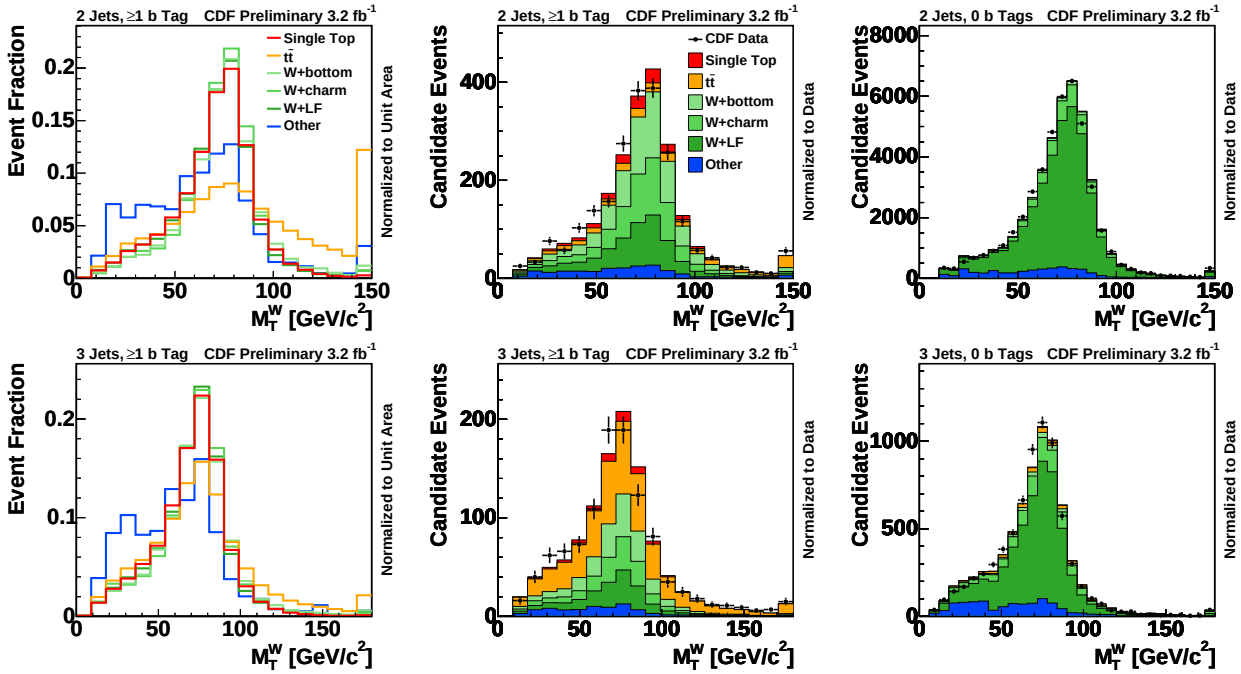


Figure 8.5: Templates (left) and validation plots comparing data and MC for the discriminating variable M_T^W. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle corresponds to events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

These variables are useful to distinguish signal from background events, however, individually they are not powerful enough to extract the small signal from the large backgrounds. Therefore, they will be used, together with other variables, as inputs to a Boosted Decision Tree classifier that will exploit all their features and correlation.

It is important to make sure that the Monte Carlo properly models the data for all these variables. Figures 8.1-8.5 show the MC templates of these variables (the plots for b_{NN} can be seen in Figures 4.5 and 7.3) for each process and the comparison between data and MC for the signal samples and for untagged control regions. The templates demonstrates the separation power of these variables, while the good agreement between the data and the MC predictions validates the use of these variables into the multivariate analysis and increases our confidence in the result.

8.2. The Boosted Decision Tree Technique

A Decision Tree (DT) [163, 164] is an algorithm that classifies events with a series of binary decisions, each one based on a single variable. It is built by recursively splitting the initial sample into two disjointed subsets until the number of events in a "node" drops below a set minimum. Each node in the tree splits the sample based on a cut adjusted to provide optimal separation

between signal and background events. The same variable may be used in multiple nodes, and some variables may not be used at all. The phase space is split into regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final "leaf" nodes. A schematic view of a DT is shown in Figure 8.6.

A single DT is very similar to rectangular cuts. However, whereas a cut-based analysis is able to select only one hypercube as region of phase space, the Decision Tree is able to split the phase space into a large number of hypercubes, each of which is identified as either signal-like or background-like. The path down the tree to each leaf node represents an individual cut sequence that selects signal or background depending on the type of the leaf node.

A shortcoming of Decision Trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. For example, if two input variables exhibit similar separation power, a fluctuation in the training sample may cause the tree growing algorithm to decide to split on one variable, while the other variable could have been selected without that fluctuation. In such a case the whole tree structure is altered below this node, possibly resulting also in a substantially different classifier response.

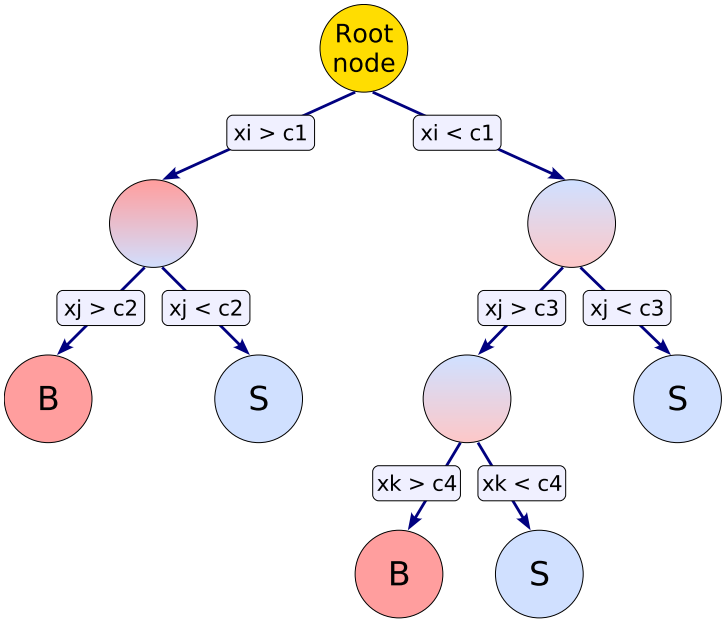


Figure 8.6: Schematic view of a Decision Tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is performed. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled S for signal and B for background depending on the majority of events that end up in the respective nodes.

This problem is overcome by a boosting [165] procedure that extends this concept from one tree to several trees which form a "forest" of Decision Trees. The trees are derived from the same training ensemble by reweighting misclassified events for the training of the next tree. The trees in the forest are finally combined into a single classifier which is given by a weighted average of the individual Decision Trees. Boosting stabilizes the response of the Decision Trees with respect to fluctuations in the training sample and is able to considerably enhance the performance with respect to a single Decision Tree. However, the advantage of the straightforward interpretation of the Decision Tree is lost. While one can of course still look at a limited number of trees trying to interprete the training result, one will hardly be able to do so for hundreds of trees in a forest. Nevertheless, the general structure of the selection can already be understood by looking at a limited number of individual trees.

As other multivariate techniques, Boosted Decision Trees are sensitive to overtraining, i.e., over-performing the training data while poorly generalizing to new samples. To reduce this effect, a "pruning" algorithm, that removes statistically insignificant nodes, is employed.

8.2.1. Building a DT

The training or building of a DT is the process that defines the splitting criteria for each node. At each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. The node splitting is stopped once it has reached a minimum number of events. The leaf nodes are classified as signal or background according to the class the majority of events belongs to.

A variety of separation criteria can be configured to assess the performance of a variable and a specific cut requirement. Because a cut that selects predominantly background is as valuable as one that selects signal, the criteria are symmetric with respect to the event classes. All separation criteria have a maximum where the samples are fully mixed, i.e., at purity $p = 0.5$, and fall off to zero when the sample consists of one event class only. The purity p of a node is defined as $p = s/(s + b)$, where s and b are the weighted sum of signal and background events, respectively, in the node; hence pure signal nodes have $p = 1$, whereas pure background nodes have $p = 0$. Since the splitting criterion is always a cut on a single variable, the training procedure selects the variable and cut value that optimizes the increase in the separation index between the parent node, I_{parent}, and the sum of the indices of the two daughter nodes, I_{left} and I_{right}, weighted by their relative fraction of events, this is

$$\Delta I = (s + b) \cdot I_{\mathrm{parent}} - (s_{\mathrm{left}} + b_{\mathrm{left}}) \cdot I_{\mathrm{left}} - (s_{\mathrm{right}} + b_{\mathrm{right}}) \cdot I_{\mathrm{right}}. \tag{8.1}$$

The separation criteria used in this analysis is the *Gini Index* [166], defined as $I_G = p \cdot (1 - p) =$

$s \cdot b/(s+b)^2$. Other splitting criteria quantities have been investigated, however, the results based on the optimization of the Gini Index has been found to perform slightly better than the results from other methods.

The cut values are optimized by scanning over the variable range with a granularity given by the parameter $nCuts$. The value of $nCuts = 25$ proved to be a good compromise between computing time and step size. Finer stepping values did not increase noticeably the performance of the BDTs.

In principle, the splitting could continue until each leaf node contains only signal or only background events, which could suggest that perfect discrimination is achievable. However, such a Decision Tree would be strongly overtrained.

8.2.2. Boosting Algorithm

Boosting is a general procedure in which the same classifier is trained several times using a successively boosted (reweighted) training event sample. The final classifier is then derived from the combination of all the individual classifiers. This analysis uses the ADABOOST[165, 167] (adaptive boost) algorithm, where events that were misclassified during the training of a tree are given a higher event weight in the training of the next following tree. Starting with the original event weights when training the first Decision Tree, the subsequent tree is trained using a modified event sample where the weights of previously misclassified events[1] are multiplied by a common boost weight α. The boost weight used for the training of the i-th tree is derived from the weighted fraction of misclassified events, r, of the previous tree,

$$\alpha_i = \frac{1 - r_{i-1}}{r_{i-1}}. \tag{8.2}$$

The weights of the entire event sample are then renormalized such that the sum of weights remains constant.

The resulting event classification $y_{BDT}(x)$ for the Boosted Decision Tree is given by

$$y_{BDT}(x) = \sum_{i \in forest} \ln(\alpha_i) \cdot h_i(x), \tag{8.3}$$

where the sum is over all trees in the forest. Large (small) values for $y_{BDT}(x)$ indicate a signal-like (background-like) event. $h(x)$ (x being the tuple of input variables) is the result of an individual tree that can either be defined to be $+1$ (-1) for events ending up in signal-like (background-like) leaf nodes according to the class the majority of training events belong to in that leaf, or it can be defined as the purity of the leaf node in which the event is found. We found that the latter option

[1] A given event x is considered as misclassified if $h(x) < 0$ ($h(x) > 0$) for signal (background) events, where $h(x)$ is the classification result of the individual tree as will be defined later in this section.

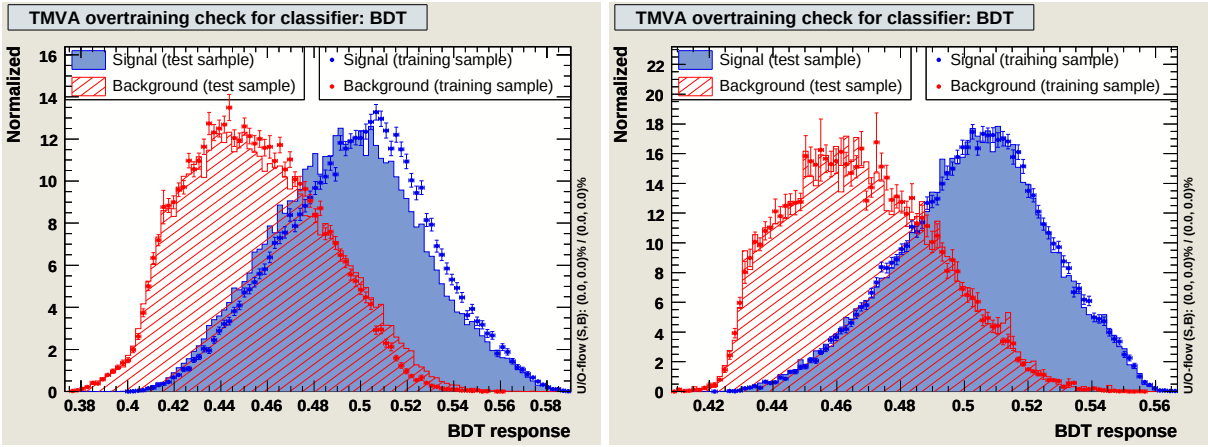
performs better for the training of the single-tag samples, while the double tag samples –which have fewer events– perform better when trained with the former option. For this analysis, it has been found that the performance of the classifiers stops increasing after about 350-500 boosting cycles, depending on the trained sample.

8.2.3. Pruning Method

Overtraining occurs when a machine learning problem has too few degrees of freedom, because too many model parameters of an algorithm were adjusted to too few data points. Overtraining leads to a seeming increase in the classification performance, if measured on the training sample, over the objectively achievable one and to an effective performance decrease when measured with an independent test sample. Boosted Decision Trees usually suffer from at least partial overtraining due to their large number of nodes. To avoid overtraining a Decision Tree must be pruned.

Pruning is the process of cutting back a tree from the bottom up after it has been built to its maximum size. Its purpose is to remove statistically insignificant nodes and thus reduce the overtraining of the tree. It has been found to be beneficial to first grow the tree to its maximum size and then cut back, rather than interrupting the node splitting at an earlier stage. This is because apparently insignificant splits can nevertheless lead to good splits further down the tree.

In this analysis we use the *Cost Complexity* [163] pruning algorithm. This algorithm relates the gain, in terms of misclassified training events, obtained by the subtree below a node compared to



(a) (b)

Figure 8.7: Example of an overtrained BDT (a) and a BDT trained with enough $PruneStrength$ to avoid overtraining (b). Blue are signal events and red are background events. Points are events from the training sample, and histogram are events from the test sample.

the node itself with no further splitting. The cost estimate in a node is given by the misclassification error in the node $R = 1 - \max(p, 1-p)$. The cost complexity for this node is then defined by

$$\rho = \left((s+b) \cdot R_{\text{node}} - \sum_{i \in \text{leaves}}^{\text{subtree}} (s_i + b_i) \cdot R_i \right) / \left(N_{\text{leaves}}^{\text{subtree}} - 1 \right), \tag{8.4}$$

where the sum is over all leaves in the subtree below the given node.

The node with the smallest ρ value in the tree is recursively pruned away as long as $\rho < PruneStrength$, where $PruneStrength$ is a parameter which has to be tuned for each analysis until overtraining is completely avoided. A validation of the overtraining is provided by comparing the shape of the BDT output between the sample used for the training and an independent test sample. An example of an overtrained and a non-overtrained BDT is shown in Figure 8.7.

8.3. Training of the BDTs

For the search of electroweak single top quark production, four different Boosted Decision Trees are optimized in each of the four signal regions: two jet events with one b-tag, two jet events with both being b-tagged, three jet events with one b-tag, and three jet events with at least two b-tagged jets. In this way, the training of each BDT exploits the different background composition in each sample and the different features to separate the single top signal from the backgrounds in each sample.

8.3.1. Training Samples

A natural way of choosing the sample composition for the training of the Boosted Decision Trees is to use a mixture of all the expected processes with the correct estimated composition. The simulated samples have been generated with different MC generators (see Chapter 6), therefore one should expect different sample sizes for each process. It has been found that the performance of the classifier for the single top search improves significantly by increasing the training sample[2], as long as there is no single process with a sample size considerably larger than the others, in which case the performance of the trained BDT could be degraded. As the mistag sample typically has ~ 10 times more events than the other samples, this sample is truncated to approximately the size of the next larger sample for the training, to avoid a degradation of the BDT. The TMVA package has the ability of introducing event by event weights into the training sample, in this way we can use as much statistics as we have available in our MC samples, while keeping the correct predicted

[2]With an infinite sample, a training algorithm is expected to behave ideally.

composition between processes. To build a BDT, half of the events are used for the training and the other half is used for overtraining tests as described in Section 8.2.3.

The signal events are given a global weight such as the weighted sum of signal events in the training equals the weighted sum of background events in the training. The signal training sample is a mixture of t- and s-channel events weighted according to the expected composition, except for the BDT training of the two jet two tag sample where it has been found a better performance when using only the dominant s-channel events. The backgrounds processes included in the training are $t\bar{t}$, W+bottom, W+charm and W+light; all of them are used for the training of the single b-tagged samples, while for the double b-tagged samples only the former two are used as they are by far the dominant contributions to those samples. The inclusion of all backgrounds (except QCD multijet) into the training has been investigated; since no improvement has been found, the simpler case has been chosen.

8.3.2. Input Variables

One advantage of Decision Trees compared to other multivariate techniques is that they are insensitive to the inclusion of poorly discriminating input variables (as long as the variable is well modeled), an extra variable can only do good to the classification problem. The Decision Tree training algorithm will basically ignore non discriminating variables as for each node splitting only the best discriminating variable is used.

22 variables are used for the training of the Boosted Decision Trees in the two jet event samples. For the three jet event samples, 30 variables are used due to the presence of a third jet. The input variables are:

- The transverse energy and pseudorapidity of each jet.
- The transverse momentum and pseudorapidity of the lepton.
- $\not\!\!E_T$: the missing transverse energy in the event.
- H_T: The scalar sum of the transverse energies of all the objects in the event.
- M_{jj}: the invariant mass of the di-jet system. In three jet events, all possible pairs of jets are used, as well as the invariant mass of the tri-jet system.
- $M_{\ell\nu b}$: the invariant mass of the reconstructed top quark.
- $M_{\ell\nu jj}$: the invariant mass of the lepton, neutrino and both jets. In three jet events, the two leading jets are used, as well as $M_{\ell\nu jj}$, where all jets are used.

- $Q \times \eta$: the charge of the lepton times the pseudorapidity of the untagged jet.
- b_{NN}: the neural net jet flavor separator.
- m_T^W: the transverse mass of the reconstructed W boson.
- η_W: the pseudorapidity of the reconstructed W boson.
- $\Delta\phi$ between each jet and the missing transverse energy.
- $\Delta\phi$ between each jet and the lepton.
- $\Delta\phi$ between the lepton and the missing transverse energy.
- The cosine of the angle between the lepton and each jet.

All energies involved are level-5 corrected. When a variables requires a b-quark jet, it is chosen as the b-tagged jet for the single-b-tag events, and as the jet with largest $Q \times \eta$ for the double-b-tag events. The neutrino is reconstructed as described in Section 4.3.4.

Validation plots of all input variables are shown in Appendix B

A ranking of the BDT input variables is derived by counting how often the variables are used to split Decision Tree nodes, and by weighting each split occurrence by the separation gain-squared (see Equation 8.1) it has achieved and by the number of events in the node. Tables 8.1-8.2 show the variable ranking for each trained channel.

8.4. BDT Output Distributions

The outputs of the Boosted Decision Trees trained in each of the four signal samples are used as the discriminant for the combined search of s- and t-channel single top quark production. These output are use to create templates of each process, which finally are to be fitted to the output distribution of the observed data events. By definition, the raw output of the Boosted Decision Trees lie always inside the range [-1, 1], however, due to the complicate topology of the final tress, the output may fall in a small subrange inside the allowed range [-1, 1] (see for example Figure 8.7). Therefore, a linear transformation of the output is applied in such a way that the final output always goes from -1 (for very background-like events) to +1 (for very signal like-events). To prevent problems with insufficient Monte Carlo statistics at the most signal-like extreme end of the BDT output range, bins in the highest output regions are grouped together until it has a nonzero background prediction.

Figure 8.8 shows the final templates used in the fit for the cross section measurement. For ease of display, different processes are grouped in the templates labeled as W+HF (W+bottom and W+charm), $t\bar{t}$ (lepton+jets and dilepton top quark production) and "Other" (WW, WZ and ZZ diboson production, Z+jets, and QCD multijet), however, in the actual likelihood fits every process has its own template.

Figure 8.9 shows the predicted distributions of the BDT output in each of the four signal samples. Here is where the separation power of the BDT discriminant is shown, for instance, we started

	2 jets, 1 b-tag		2 jets, 2 b-tags	
Rank	Variable	Importance	Variable	Importance
1	b_{NN}	4.597e-01	M_{jj}	7.453e-02
2	M_{jj}	1.799e-01	$M_{\ell\nu b}$	6.599e-02
3	$Q \times \eta$	1.077e-01	m_T^W	6.453e-02
4	$M_{\ell\nu b}$	4.684e-02	$\Delta\phi_{j_1\nu}$	5.142e-02
5	$E_T^{j_1}$	4.570e-02	$M_{\ell\nu jj}$	4.805e-02
6	m_T^W	1.890e-02	b_{NN}	4.656e-02
7	$E_T^{j_2}$	1.874e-02	$\cos\theta_{j_1\ell}$	4.645e-02
8	p_T^ℓ	1.639e-02	η_{j_1}	4.563e-02
9	$\cos\theta_{j_1\ell}$	1.572e-02	$E_T^{j_1}$	4.522e-02
10	$\Delta\phi_{\ell\nu}$	1.481e-02	$E_T^{j_2}$	4.424e-02
11	η_{j_2}	1.063e-02	η_ℓ	4.366e-02
12	η_ℓ	8.744e-03	$\Delta\phi_{j_2\nu}$	4.339e-02
13	$M_{\ell\nu jj}$	8.106e-03	$\Delta\phi_{j_2\ell}$	4.252e-02
14	H_T	7.755e-03	$\cos\theta_{j_2\ell}$	4.249e-02
15	$\Delta\phi_{j_1\nu}$	7.589e-03	$\Delta\phi_{\ell\nu}$	4.098e-02
16	$\Delta\phi_{j_1\ell}$	6.232e-03	$\Delta\phi_{j_1\ell}$	3.886e-02
17	$\cos\theta_{j_2\ell}$	6.035e-03	p_T^ℓ	3.869e-02
18	$\not\!\!E_T$	4.676e-03	$Q \times \eta$	3.850e-02
19	$\Delta\phi_{j_2\nu}$	4.174e-03	$\not\!\!E_T$	3.840e-02
20	η_{j_1}	3.922e-03	η_{j_2}	3.573e-02
21	η_W	3.881e-03	H_T	3.227e-02
22	$\Delta\phi_{j_2\ell}$	3.839e-03	η_W	3.188e-02

Table 8.1: Variable ranking result for the BDT training in the 2 jet event samples in an arbitrary scale. Top variable is best ranked.

	3 jets, 1 b-tag		3 jets, 2 b-tags	
Rank	Variable	Importance	Variable	Importance
1	b_{NN}	3.846e-01	$Q \times \eta$	7.785e-02
2	$M_{j_2 j_3}$	7.124e-02	H_T	5.032e-02
3	m_T^W	5.449e-02	$M_{\ell\nu b}$	4.698e-02
4	$M_{j_2 j_3}$	5.358e-02	$M_{j_2 j_3}$	4.100e-02
5	H_T	4.978e-02	$M_{j_1 j_2}$	3.864e-02
6	$Q \times \eta$	4.729e-02	M_{jjj}	3.735e-02
7	$M_{\ell\nu b}$	3.924e-02	$\Delta\phi_{\ell\nu}$	3.677e-02
8	M_{jjj}	2.642e-02	m_T^W	3.632e-02
9	$E_T^{j_3}$	2.609e-02	η_{j_1}	3.573e-02
10	η_{j_2}	2.549e-02	$\cos\theta_{j_1\ell}$	3.447e-02
11	$M_{j_1 j_2}$	2.490e-02	$E_T^{j_2}$	3.400e-02
12	$\cos\theta_{j_1\ell}$	2.370e-02	$E_T^{j_3}$	3.324e-02
13	p_T^ℓ	1.910e-02	$M_{j_1 j_3}$	3.307e-02
14	η_ℓ	1.774e-02	η_{j_2}	3.275e-02
15	$\Delta\phi_{\ell\nu}$	1.382e-02	$\cos\theta_{j_3\ell}$	3.046e-02
16	η_{j_1}	1.373e-02	$E_T^{j_1}$	3.017e-02
17	$\Delta\phi_{j_1\ell}$	1.223e-02	η_W	2.993e-02
18	$\cos\theta_{j_2\ell}$	1.069e-02	$\Delta\phi_{j_2\nu}$	2.988e-02
19	$\cos\theta_{j_3\ell}$	9.661e-03	$\Delta\phi_{j_1\ell}$	2.898e-02
20	$E_T^{j_1}$	8.641e-03	b_{NN}	2.894e-02
21	$\Delta\phi_{j_3\nu}$	8.278e-03	η_ℓ	2.873e-02
22	$\Delta\phi_{j_2\nu}$	8.256e-03	$\Delta\phi_{j_3\ell}$	2.855e-02
23	$\not{E}_T$	8.141e-03	$M_{\ell\nu jjj}$	2.828e-02
24	$\Delta\phi_{j_1\nu}$	7.496e-03	$\Delta\phi_{j_3\nu}$	2.692e-02
25	$\Delta\phi_\ell$	6.743e-03	$\cos\theta_{j_2\ell}$	2.616e-02
26	$E_T^{j_2}$	6.133e-03	p_T^ℓ	2.518e-02
27	$\Delta\phi_{j_2\ell}$	6.035e-03	$\Delta\phi_{j1_\nu}$	2.508e-02
28	η_W	5.932e-03	$\Delta\phi_{j_2\ell}$	2.204e-02
29	$M_{\ell\nu jj}$	5.417e-03	$\not{E}_T$	2.160e-02
30	$M_{\ell\nu jjj}$	5.108e-03	$M_{\ell\nu jj}$	2.060e-02

Table 8.2: Variable ranking result for the BDT training in the 3 jet event samples in an arbitrary scale. Top variable is best ranked.

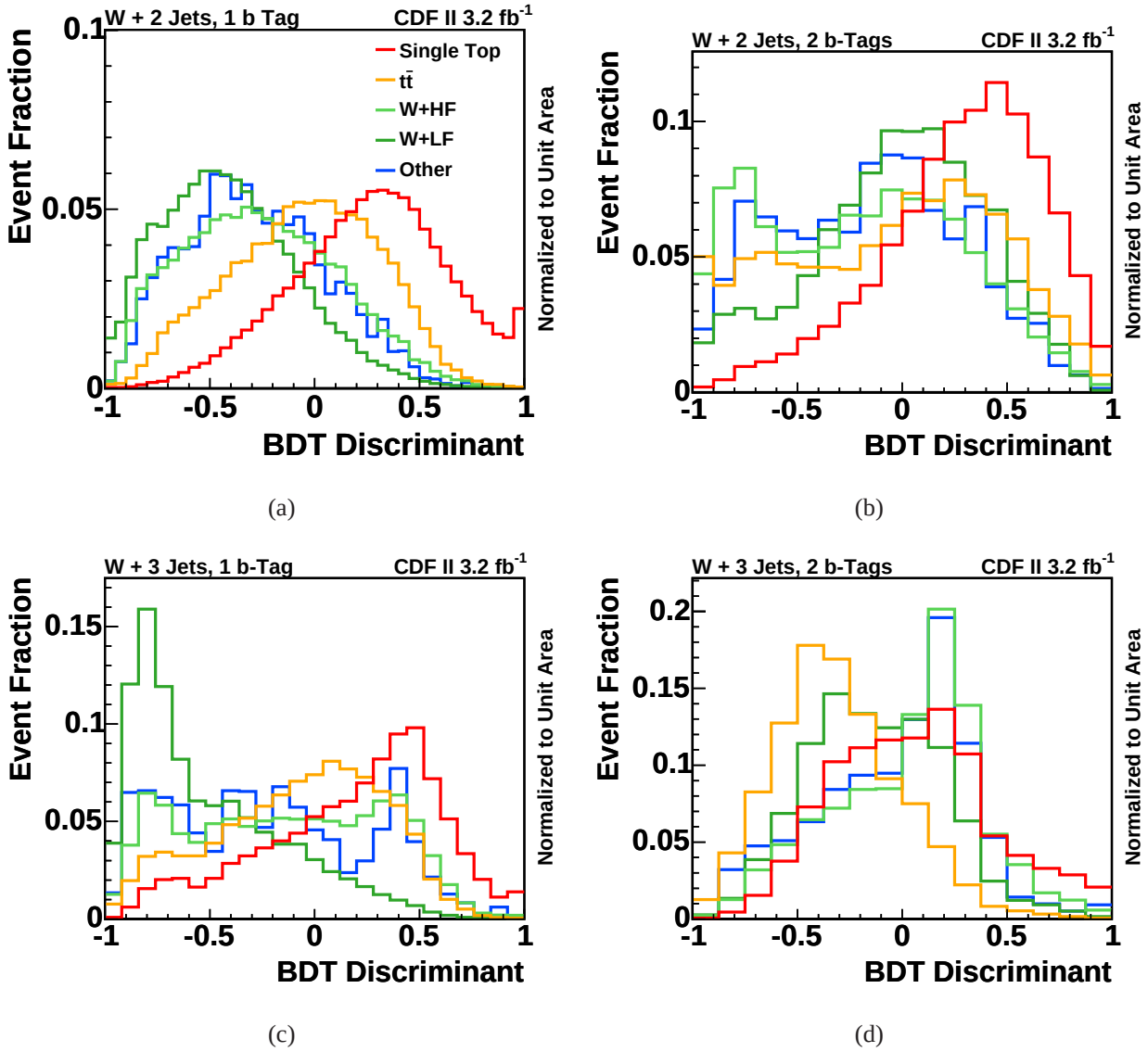


Figure 8.8: Templates of the predictions of the BDT outputs for the signal and background processes, each scaled to unit area. Single top quark events are predominantly found on the right-hand sides of the plots while background events are mostly found on the left-hand sides. (a) is the two jet one b-tag sample, (b) is the two jet two b-tag sample, (c) is the three jet one b-tag sample, and (d) is the three jet two b-tag sample.

from a signal to background ratio of $s/b \approx 1/15$ in the two jet one b-tag events (see Table 7.2), and we end up with a discriminant with a signal to background ratio of $s/b \approx 6.8$ in the most signal-like bin with an expected signal yield of about three single top events and less than 0.5 background events (in the five most signal-like events we expect $s/b \approx 2.3$ with about 12 signal events against about 5 signal events).

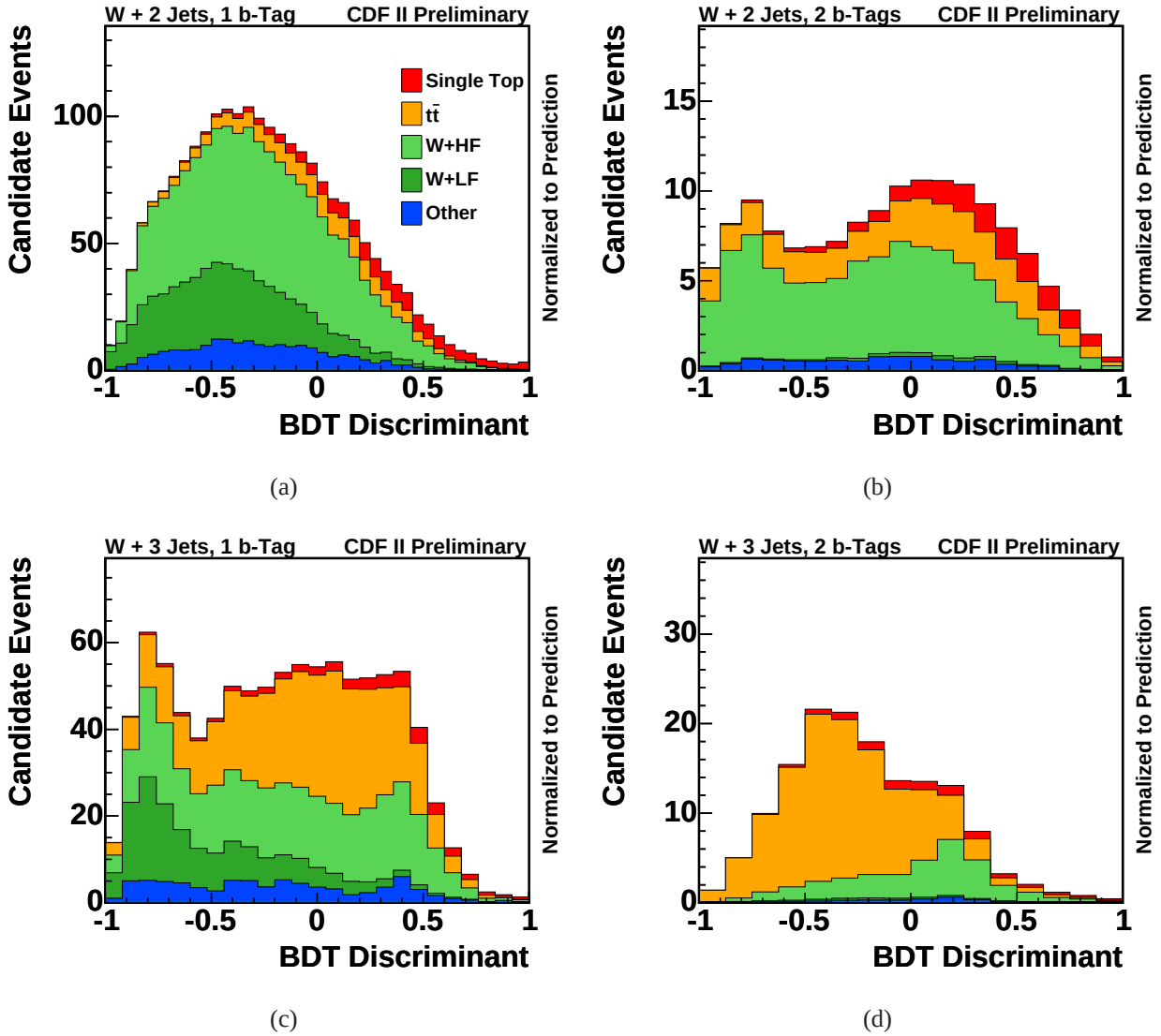


Figure 8.9: Predicted distributions of the BDT outputs. The four signal regions are shown: (a) two jets and one b-tag events, (b) two jets and two b-tag events, (c) three jets and one b-tag events, (d) three jets and at least two b-tag events.

8.4.1. Validation of the Discriminant

The distributions of all the input variables to each Boosted Decision Tree are checked in the zero, one, and two-tag samples for two- and three-jet events. Some figures have been shown in Sections 7.4 and 8.1, a complete set of validation plots is shown in Appendix B.

Further than checking the input variables to the Boosted Decision Trees, the output of the classifiers are validated in control regions before looking at the single top candidate events. Figures 8.10 show the output of the four BDT classifiers applied to the untagged samples with two or three identified jets. For the inputs variables that required a b-quark jet, the leading jet is chosen; and the neural net flavor separator (which is only defined for b-tagged jets) a randomized value

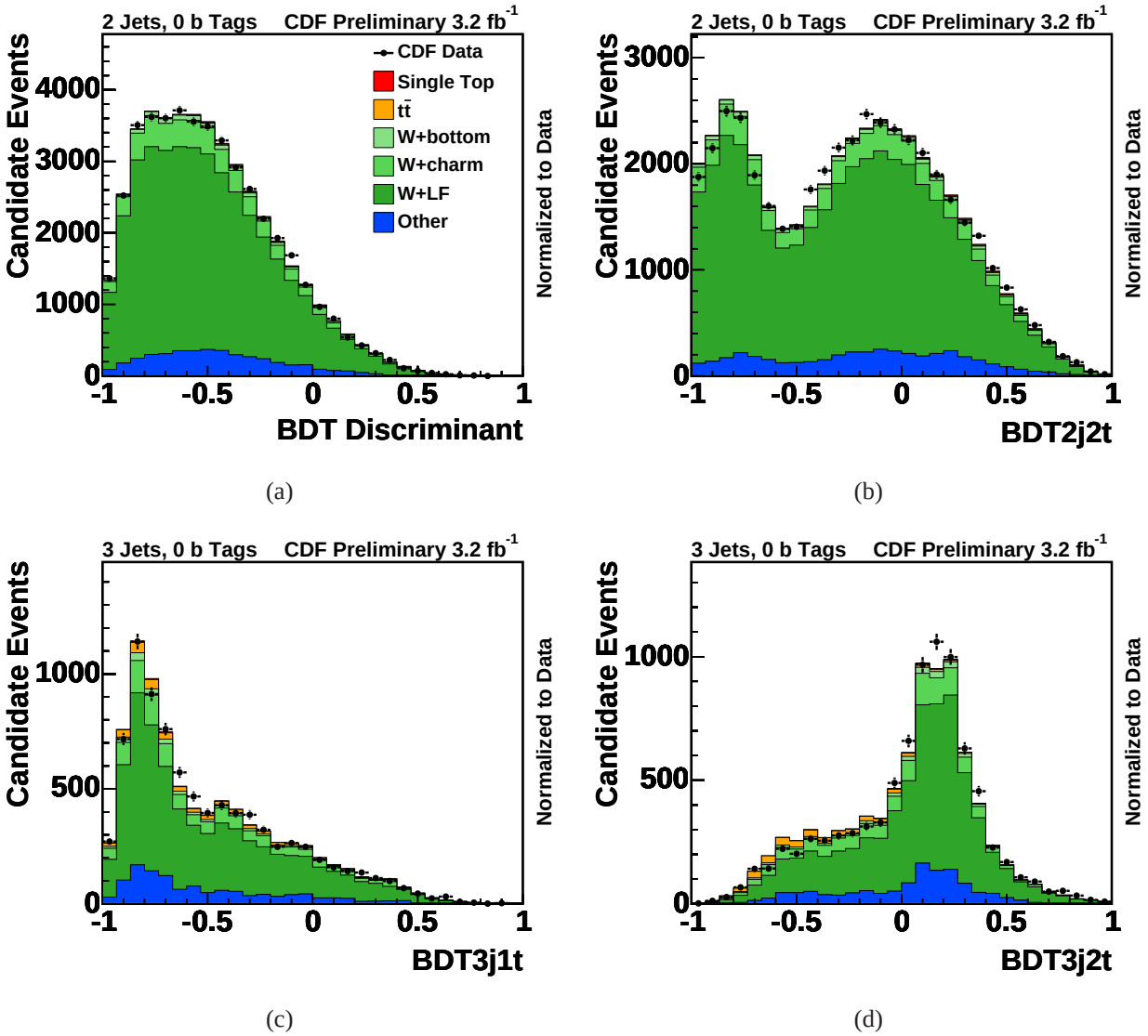


Figure 8.10: Distribution of the BDT outputs applied to the untagged control sample with two ((a) and (b)) and three ((c) and (d)) identified jets, showing close modeling of the data and good control of the W+light-flavor shape. The classifiers are optimized in the signal samples of (a) two jets with one b-tag, (b) two jets with two b-tags, (c) three jets with one b-tag,(d) three jets with at least two b-tags.

from the W+LF template is picked. This sample is orthogonal to the single top signal regions with very little contribution from top quark production ($< 0.5\%$). The dominant contribution in the untagged samples is W+light-flavor jets, therefore the good agreement found in the distributions indicates a good control over the W+LF shapes.

The distribution of the BDT outputs are also checked in the $t\bar{t}$-enriched control sample of W+4 jet events (Figure 8.11). The good agreement in this sample indicates a good control of the shape of the dominant $t\bar{t}$ contribution (about 75%). This cross-check is very important as

the $t\bar{t}$ process is one of the most difficult backgrounds to distinguish from the single top quark production, and its contribution in the discriminant typically lies closer to the signal-like region than other backgrounds.

The good agreement seen between the predictions and the observations in both the input variables and the output variables gives confidence in the validity of the technique.

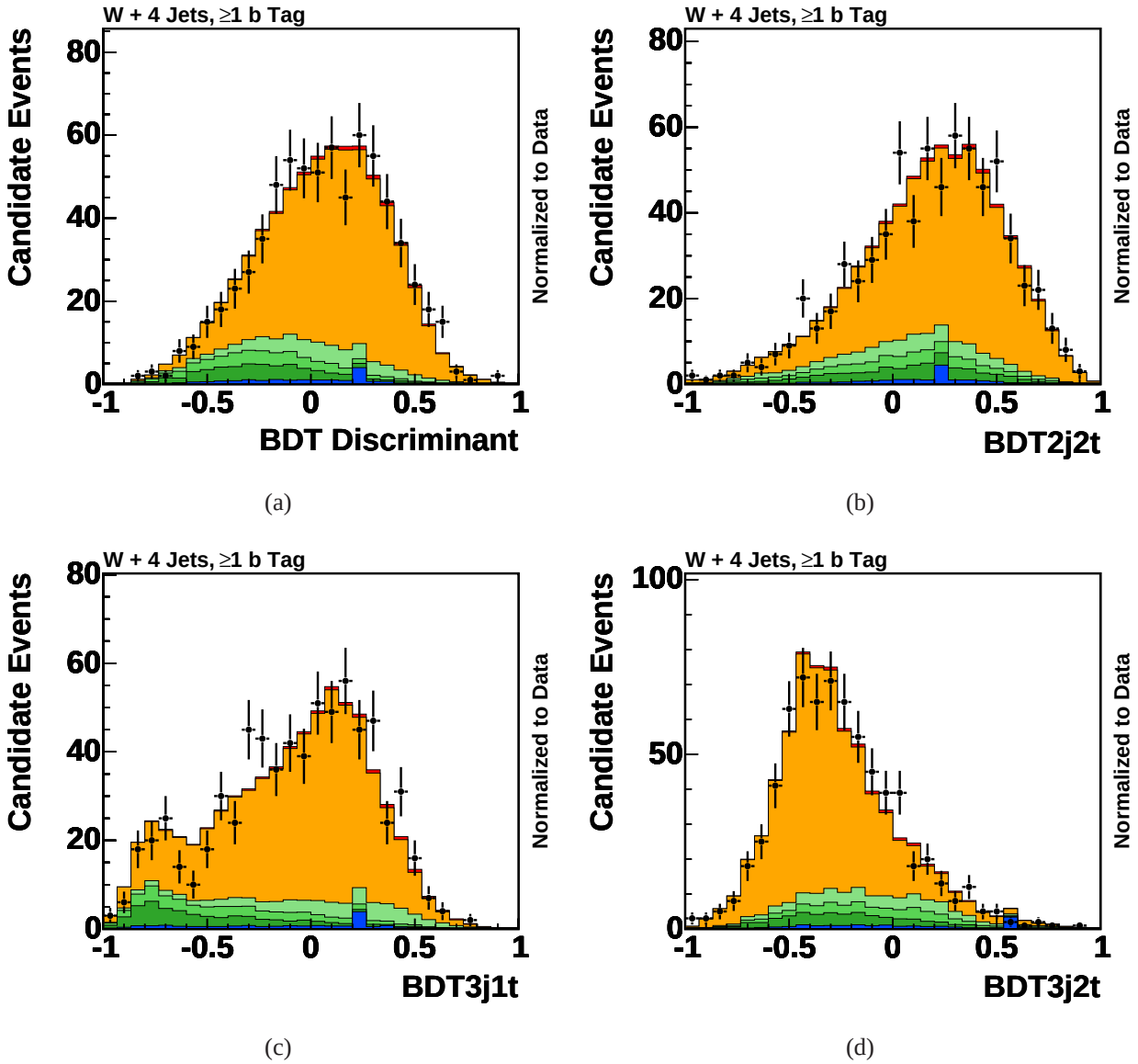


Figure 8.11: Distribution of the BDT outputs applied to the control sample with four identified jets and at least one b-tag, showing close modeling of the data and good control of the $t\bar{t}$ shape. The classifiers are optimized in the signal samples of (a) two jets with one b-tag, (b) two jets with two b-tags, (c) three jets with one b-tag,(d) three jets with at least two b-tags.

Chapter 9

ANALYSIS

9.1. Systematic Uncertainties

The search for single top quark production and the measurement of the cross section require substantial input from theoretical models, Monte Carlo simulations and extrapolations from control samples in data. We assign systematic uncertainties to our predictions and include the effects of these uncertainties on the measured cross sections as well as the significance of the signal.

We consider three categories of systematic uncertainty: uncertainty in the predicted rates of signal and background processes, uncertainty in the shapes of the distributions of the discriminant variable, and uncertainty arising from the limited number of Monte Carlo events used to predict the signal and background expectations in each bin of each discriminant distribution. Sources of uncertainty may affect multiple signal and background components. The effects of systematic uncertainty from the same source are considered to be fully correlated; for example, the integrated luminosity estimate affects the predictions of the Monte-Carlo-based background processes and the signal, so the uncertainty on the integrated luminosity affects all these processes in a correlated way. The effects of different sources of systematic uncertainty are considered to be uncorrelated.

The effects of all systematic uncertainties are included in the hypothesis tests and cross section measurement, as described in Section 9.2. Detailed descriptions of the sources of uncertainty and their estimation are given below.

9.1.1. Rate Uncertainties

Rate uncertainties affect only the expected contribution of the signal and background samples. Some sources have asymmetric uncertainties. All rate uncertainties are assigned truncated Gaussian priors, where the truncation prevents predictions from being negative for any source of signal or background. The sources of rate uncertainties in this analysis are listed below with their range of impact summarized in Table 9.1.

- **Integrated Luminosity** A symmetric uncertainty of $\pm$6% is applied to all Monte-Carlo based predictions. This uncertainty includes the uncertainty in the $p\bar{p}$ inelastic cross section as well as the uncertainty in the acceptance of CDF's luminosity monitor [168]. The requirement that the primary vertex position in z is within $\pm$60 cm of the origin causes a small acceptance uncertainty that is included as well.

- **Theoretical Cross Sections** Our MC-based background processes are scaled to theoretical predictions at NLO (or better). The associated theoretical uncertainties are applied (see Chapter 6 and Table 7.1). For the single top quark cross section fit, this uncertainty is not applied since this is the quantity being measured.

- **Monte Carlo Generator** The impact of using LO instead of NLO Monte Carlo event generators is evaluated for single top quark events (see Section 6.1). The uncertainty is quantified by taking the differences between the default LO Monte Carlo simulation with MADEVENT and the theoretical NLO calculation provided by ZTOP [12].

- **Acceptance and Efficiency Scale Factors** The predicted rates of Monte Carlo background processes and the signals are affected by trigger efficiency (Sections 5.1.1 and 5.1.2), lepton identification efficiency (Section 4.3.2), and the b-tagging efficiency (Section 4.3.5). Known differences between the data and the simulation are corrected for by scaling the prediction, and uncertainties on these scale factors are collected together in one source of uncertainty since they affect the predictions in the same way.

- **Heavy Flavor Fraction in W+jets** The prediction of the $W + b\bar{b}$, $W + c\bar{c}$, and $W + c$ fractions in the $W + 2$ jets and $W + 3$ jets samples are extrapolated from the $W + 1$ jet sample as described in Section 7.2.2. It is found that ALPGEN underpredicts the $W + b\bar{b}$ and $W + c\bar{c}$ fractions in the $W + 1$ jet sample by a factor of 1.4 ± 0.4. We assume that the $W + b\bar{b}$ and $W + c\bar{c}$ predictions are correlated. The uncertainty on this scale factor comes from the spread in the measured heavy-flavor fractions using different variables to fit the data, and in the difference between the $W + b\bar{b}$ and $W + c\bar{c}$ scale factors. The $W + c$ prediction from ALPGEN is compared with CDF's measurement [161] and is found not to

require scaling, but a separate, uncorrelated uncertainty is assigned to the $W + c$ prediction, with the same relative magnitude as the $W + b\bar{b}$+$W + c\bar{c}$ uncertainty. This is one of the largest systematic effects in this analysis.

- **Mistag Estimate** The total contribution of the mistag is calculated using mistag probabilities described on Section 4.3.5. The largest source of systematic uncertainty in the mistag estimate comes from extrapolating from the negative tag rate in the data to positive tags by estimating the asymmetry between positive light-flavor tags and negative light-flavor tags (see Section 4.3.5). Other sources of uncertainty come from differences in the negative tag rates of different data samples used to construct the mistag matrix.

- **QCD Multijet Estimate** The QCD multijet rate prediction varies when the $\not\!\!E_{\rm T}$ distribution is constructed with a different number of bins or if different models are used for the QCD templates (see Section 7.2.1). The $\not\!\!E_{\rm T}$ fits also suffer from small data samples, particularly in the double-tagged samples.

- **Initial State Radiation (ISR)** The model used for ISR is PYTHIA's "backwards evolution" method [129]. This uncertainty is evaluated by generating new Monte Carlo samples for $t\bar{t}$ and single top quark signals with Λ_{QCD} doubled or divided in half, for samples with more ISR and less ISR, respectively. Simultaneously, the initial transverse momentum scale is multiplied by four or divided by four, and the hard scattering scale of the shower is multiplied by four or divided by four, for more ISR and less ISR, respectively. These variations are checked with Drell-Yan Monte Carlo and data samples. The p_T distributions of dileptons are compared between data and Monte Carlo, and the ISR more/less prescriptions generously bracket the available data, as can be seen in Figure 9.1. ISR and FSR rate uncertainties are not evaluated for the W+jets Monte Carlo samples because they are scaled to the data and given an extrapolation uncertainties.

- **Final State Radiation (FSR)** PYTHIA's model of gluon radiation from partons emitted from the hard-scattering interaction has been tuned with high precision to LEP data [129]. Nonetheless, uncertainty remains in the radiation from beam remnants, and parameters analogous to those adjusted for ISR are adjusted in PYTHIA for the final-state showering, except for the hard-scattering scale parameter. The effects of variations in ISR and FSR are treated as 100% correlated with each other.

- **Jet Energy Scale (JES)** The calibration of the calorimeter response to jets is a multi-step process (see Section 4.3.3), and each step involves an uncertainty which is propagated to the final jet-energy scale. Raw jet energies are corrected for test-beam scales, detector non-uniformity, multiple interactions, and energy that is not assigned to the jet because it lies

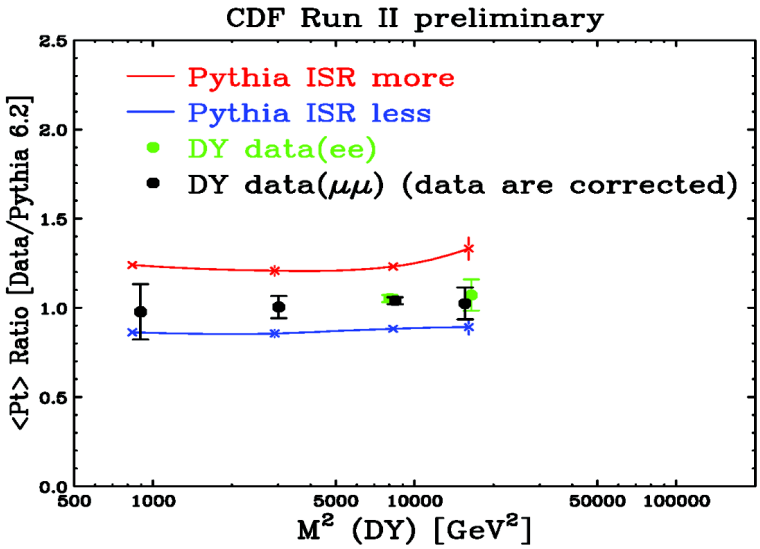
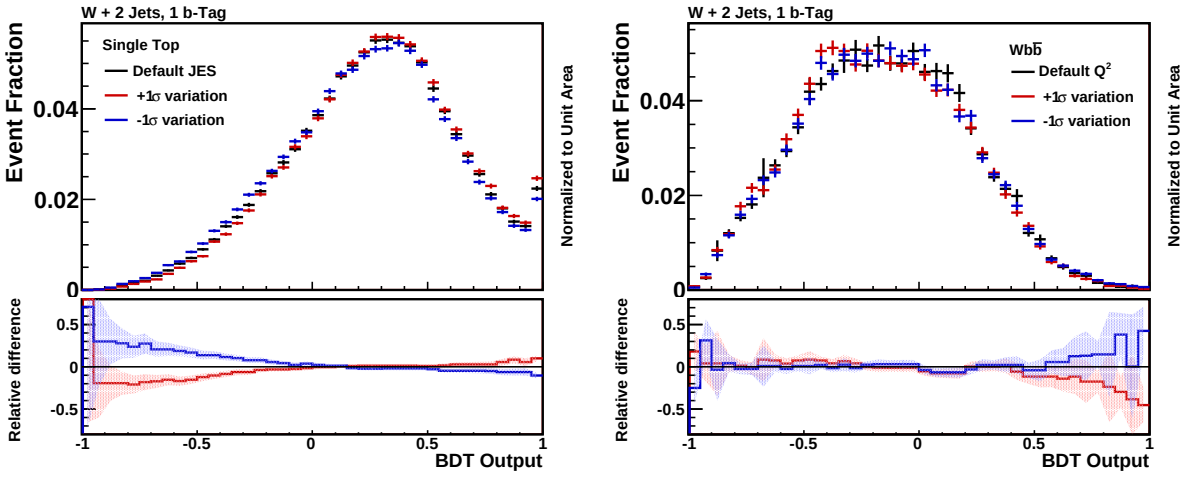


Figure 9.1: The ratio of transverse momentum between data and Monte Carlo as a function of the invariant mass of the leptons in the event. The systematic uncertainty covers the data points and their statistical uncertainty.

outside of the jet cone. The uncertainties in the jet energy scale are incorporated by processing all events in all Monte Carlo samples with the jet energy scale varied upwards and again downwards. The kinematic properties of each event are affected, and some events are re-categorized as having a different number of jets as jets change their $\not\!\!E_T$, inducing correlated rate uncertainties.

- **Parton Distribution Functions (PDF)** The central PDFs used in this analysis are the CTEQ5L set [62]. To evaluate the systematic uncertainties on the rates due to uncertainties in these PDF's, we add in quadrature the differences between the predictions of the following pairs of PDFs:

  - CTEQ5L and MRST72 [63], PDF sets computed by different groups. CTEQ5L is an NLO PDF set, and MRST72 is a LO PDF set.
  - MRST72 and MRST75, which differ in their value of α_s. The former uses 0.1125; the latter uses 0.1175.
  - CTEQ6L and CTEQ6L1, of which the former has a 1-loop α_s correction, and the latter has a 2-loop α_s correction.
  - The 20 signed eigenvectors of CTEQ6M, each compared with the central PDF.



(a) (b)

Figure 9.2: Example of two shape systematic templates in the two-jet one-b-tag sample. (a) shows the jet energy scale shifted histograms for single top events, while (b) shows the shifted histograms for the factorization and renormalization scale systematic for $Wb\bar{b}$ events.

9.1.2. Shape Uncertainties

Many of the sources of rate uncertainty listed above also induce distortions in the shapes of the templates for the signals and background processes used to model the data. These include ISR, FSR, JES, and PDF uncertainties. Here we list the sources of shape uncertainties which do not have associated rate uncertainties.

Shape uncertainty templates are all smoothed with a median smoothing algorithm. This procedure takes the ratio of the systematically shifted histograms to the central histograms and replaces the contents of each bin with the median of the ratios of a five-bin window around the bin. The first two bins and the last two bins are left unaffected by this procedure. The smoothed ratio histograms are then multiplied by the central histograms and normalize to the integral of the central histograms (since smoothing can introduce a rate change) to obtain the new varied template histograms. This procedure reduces the impact of limited Monte Carlo statistics in the bins of the central and varied templates. An example of two of the most important shape systematics for the two-jets one-b-tag sample is shown if Figure 9.2.

- **Jet Flavor Separator Modeling** Since no cut is applied on the output of the neural net jet flavor separator (see Section 4.3.6), the uncertainty associated with this quantity does not imply a rate uncertainty, but a shape uncertainty on the template distributions. The distribution of b_{NN} for light-flavor jets is found to require a small correction [143]. The full difference between the uncorrected light-flavor Monte Carlo prediction and the data-derived

corrected distribution is taken as a one-sided systematic. Since a pure sample of charm jets is not available in the data, a systematic is also assessed on the shape of the charm prediction, taking the difference between the distribution predicted by the Monte Carlo simulation and the Monte Carlo distribution altered by the light-flavor correction function.

- **Mistag Model:** Mistagged events are difficult to properly model, so their shape may not be totally trustworthy. This can be covered with a systematic shape uncertainty that uses untagged data, weighted by the mistag matrix, to construct an alternate shape template for the mistags. The untagged data largely consist of W+light flavored jets, but there is a contamination from $W + b\bar{b}$, $W + c\bar{c}$, $t\bar{t}$, and even single top quark signal events, making the systematic shape uncertainty conservative.

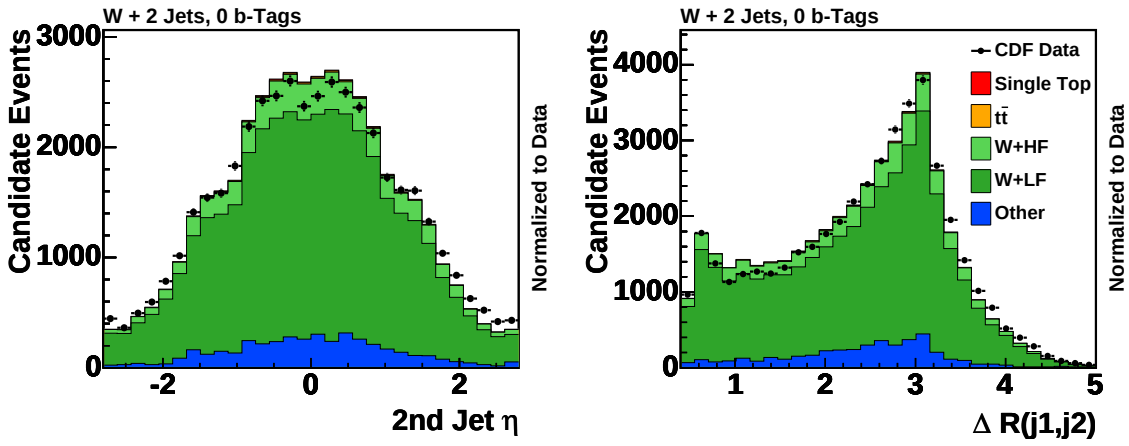
- **Factorization and Renormalization Scale:** Because ALPGEN performs fixed-order calculations to create W+jets diagrams, it requires renormalization and factorization scales as inputs. The factorization scale is set in our central samples as

$$\sqrt{M_W^2 + \sum_{\text{partons}} p_{\text{T}}^2}, \tag{9.1}$$

  The sum is over all final-state partons excluding the W boson decay products. The renormalization scale is set in the central samples to the transverse momentum of the hard-scattering vertex. Both parameters are halved and doubled in order to produce templates that cover the scale uncertainty. The heavy-flavor fraction in W+jets events is strongly dependent on the factorization scale in ALPGEN, and this rate is determined from data, so we need no additional rate uncertainty for this source of uncertainty.

- **QCD Multijet Flavor Composition:** The distribution of b_{NN} is used to fit the flavor fractions in the low-$\not\!\!E_T$ control samples (as discussed in Section 6.2.5). The limited statistical precision of these fits and the necessity of extrapolating to the higher-$\not\!\!E_T$ signal region motivates an uncertainty on the flavor composition. A systematic shape uncertainty is applied using the "worst-case" (because it makes the QCD multijet sample more signal-like, making it harder to discriminate from the signal) variation of the flavor composition: 60% b jets, 30% c jets, and 10% light-flavor jets.

- **Jet η Distribution:** Checks of the untagged $W + 2$ jet control region show a mismodeling of jets at high $|\eta|$ (Figure 9.3(a)). This mismodeling has a potentially significant impact on the analysis because of the sensitive variable $Q \times \eta$, which is highly discriminating for events with jets at large $|\eta|$. Three possible explanations for the mismodeling are possible—beam halo overlapping with real W+jets events, miscalibration of the jet energy scale in the forward calorimeters, and ALPGEN mismodeling. We could not distinguish between these



(a) (b)

Figure 9.3: Plots showing the mismodeling of the second jet pseudorapidity and the distance between the two jets in the η-ϕ plane. These must be accounted for with shape systematic uncertainties.

possibilities with the data, and thus chose to take as a systematic uncertainty the difference in the shape predictions of all Monte Carlo samples after reweighting them based based on the ratio of the data and Monte Carlo in the untagged sideband. No corresponding rate uncertainty is applied.

- **Jet ΔR Distribution:** Similarly, the distribution of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, a measure of the angular separation between two jets, is found to be mismodeled in the untagged control sample (Figure 9.3(b)). This mismodeling is believed to be due to the gluon splitting fraction in ALPGEN, but since this conclusion is not fully supported, we take as a systematic uncertainty the difference in predictions of all Monte Carlo based templates after reweighting them using the ratio of the untagged data to the prediction.

9.2. Statistical Interpretation

The goal of this analysis is to eventually discover single top quark production, and to make a precise measurement of its cross section. For the first, an evaluation of the significance of the excess of events compared with the background prediction is essential. These goals have much in common: better separation of signal events from background events and the reduction of uncertainties help improve both the cross section measurements and the expected significance if a signal is truly present. But there are also differences. For example, systematic uncertainty on the signal affects the precision of the cross section measurement, but has little effect on the observed

Source of Uncertainty	Rate	Shape	Processes affected
Jet energy scale	0–16%	✔	single top, $t\bar{t}$, diboson, Z+jets
Initial state radiation	0–11%	✔	single top, $t\bar{t}$
Final state radiation	0–15%	✔	single top, $t\bar{t}$
Parton distribution functions	1–3%	✔	single top, $t\bar{t}$
Acceptance and efficiency scale factors	0–14%		single top, $t\bar{t}$, diboson, Z+jets
Luminosity	6%		single top, $t\bar{t}$, diboson, Z+jets
Jet flavor separator		✔	W+charm, W+light
Mistag model		✔	W+light
QCD flavor model		✔	QCD multijet
Factorization and renormalization scale		✔	$Wb\bar{b}$
Jet η distribution		✔	all
Jet ΔR distribution		✔	all
QCD normalization	40%		QCD multijet
$Wb\bar{b}$ and $Wc\bar{c}$ normalization	30%		$Wb\bar{b}$, $Wc\bar{c}$
Wc normalization	30%		Wc
Mistag normalization	13–22%		W+light
Z+jets normalization	11%		Z+jets
$t\bar{t}$ cross section	12%		$t\bar{t}$
Diboson cross section	2%		diboson
Monte Carlo generator	1–2%		single top
Single top cross section	12%		single top
Top mass	2–12%		single top, $t\bar{t}$

Table 9.1: Sources of systematic uncertainty considered in this analysis. A range with the minimum and maximum observed effect on the normalization across all different processes and alanlysis imput channels is given in the rate column. Last three sources are used only in the calculation of the p-value and the calculation of the lower limit of the $|V_{tb}|$.

significance level, and only a minor effect on the predicted significance level. More importantly, a precision cross section measurement relies most on increasing acceptance and understanding the background in a larger sample. The significance of an excess, however, can be much larger if one bin in an analysis has a very low background and data in it that are inconsistent with that background, even though that bin may not contribute much information to the cross section measurement.

The contents of the low signal-to-background bins are important for the proper interpretation of the high signal-to-background bins. They serve as signal-depleted control samples which can be used to help constrain the background predictions. Not all bins are fully depleted in signal, and the signal-to-background ratio varies from very small to more than 2:1 for the most signal-like bins. Simultaneous use of all bins' contents, comparing the observations to the predictions, is needed to optimally measure the cross section and compute significance. Systematic uncertainties on the predicted rates and shapes of each component of the background and the two signals, and also bin-by-bin systematic uncertainties affect the extrapolation of the background fits to the signal regions.

These considerations are addressed below, and the procedures for measuring the cross section and the significance of the excess are performed separately. The handling of the systematic uncertainties is Bayesian, in that priors are assigned for the values of the uncertain nuisance parameters, the impacts of the nuisance parameters on the predictions are evaluated, and integrals are performed as described below over the values of the nuisance parameters.

9.2.1. Likelihood Function

The likelihood function we use in the extraction of the cross section and in the determination of the significance is the product of Poisson probabilities for each bin in each histogram of the discriminant output variable of each signal channel. The Poisson probabilities are functions of the number of observed events in each bin n_i and the predictions in each bin μ_i, where i ranges from 1 to $N_{\rm bins}$. The likelihood function is given by

$$L = \prod_{i=1}^{N_{\rm bins}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}. \tag{9.2}$$

The prediction in each bin is a sum over signal and background contributions:

$$\mu_i = \beta \cdot s_i + \sum_{k=1}^{N_{\rm bkg}} b_{ik}, \tag{9.3}$$

where b_{ik} is the background prediction in bin i for background source k. The signal prediction s_i in bin i is the sum of the s-channel and t-channel contributions based on the Standard Model rates, and scaled by β to test other values of the single top quark production cross section. A flat prior for non-negative values of β is assumed, while non-physical negative values of β are not allowed.

The likelihood is complicated by the presence of systematic uncertainties. Bayesian statistics provides a framework for incorporating systematic uncertainties by treating them as *nuisance parameters* [11, 169], whose values, known with some limited accuracy, affect the result but which are not themselves of interest to the analysis. The predictions b_{ik} and s_i, therefore, depend on

$N_{\rm syst}$ uncertain nuisance parameters θ_m, where $m = 1...N_{\rm syst}$, one for each independent source of systematic uncertainty. These nuisance parameters are given Gaussian priors centered on zero with unit width. The systematic uncertainties include overall rate uncertainties on the background estimate and uncertainties on the shape of the discriminant templates (see Section 9.1). When an uncertainty source affects both shape and rate, the effects are treated as 100% correlated. Finally, bin-by-bin Monte Carlo statistical uncertainties are taken into account introducing another set of nuisance parameters η_{ik}. This bin-by-bin uncertainties arise from limited Monte Carlo (or data control sample) statistics and are taken to be independent of each other and all other sources of systematic uncertainty.

The contribution to a bin's prediction from a single symmetric rate uncertainty parameterized as θ, which causes a percent change ρ_k in the background estimate of process k for a $\pm 1\sigma$ variation, is given by

$$b_{ik} = b^0_{ik} \cdot (1 + \rho_k \theta), \tag{9.4}$$

where b^0_{ik} is the central prediction for bin i. Usually, the shifts caused by systematic effects are not symmetrical. If the fractional uncertainty on the central value is ρ^+_k for a $+1\sigma$ variation and ρ^-_k for a negative variation, then a quadratic function [170] is determined to make a smooth application of the nuisance parameter to the predicted value:

$$b_{ik} = b^0_{ik} \cdot \left(1 + \frac{\rho^+_k + \rho^-_k}{2} \cdot \theta_m + \frac{\rho^+_k - \rho^-_k}{2} \cdot \theta^2_m\right), \tag{9.5}$$

For many systematic uncertainties

$$b_{ik} = b^0_{ik} \cdot \left\{ \prod_{m=1}^{N_{\rm syst}} \left(1 + \frac{\rho^+_{km} + \rho^-_{km}}{2} \cdot \theta_m + \frac{\rho^+_{km} - \rho^-_{km}}{2} \cdot \theta^2_m\right) \right\}. \tag{9.6}$$

The contribution to a bin's prediction from a given source of shape uncertainty is modified by linearly interpolating and extrapolating the difference between the central prediction b^0_{ik} and the prediction in a histogram corresponding to a $+1\sigma$ variation b^+_{ikm} if $\theta_m > 0$, and performing a similar operation using a -1σ varied histogram if $\theta_m < 0$:

$$b_{ik} = b^0_{ik} \cdot \left\{ 1 + \sum_{m=1}^{N_{\rm syst}} \left[\kappa^+_{ikm} H(\theta_m) + \kappa^-_{ikm} H(-\theta_m)\right] \cdot |\theta_m| \right\}, \tag{9.7}$$

where $H(\theta)$ is the Heaviside step function, and where we have introduced the relative shape uncertainties $\kappa^\pm_{ikm} = (b^0_{ik} - b^\pm_{ikm})/b^0_{ik}$. By definition, the shape uncertainties only affects the shape of the templates and not the overall rate of the background estimate, therefore $\sum_i \kappa_{ikm} = 0$. The application of shape uncertainties is not allowed to produce a negative prediction in any bin for any source of background or signal. The nuisance parameters are allowed to vary into regions which

would cause bins to go negative, but those particular bin contents for those contributions are set to zero in that case.

Each template histogram, including the systematically varied histograms, has a statistical uncertainty in each bin. These bin-by-bin uncertainties are linearly interpolated in each bin in the same way as the predicted values. If the bin-by-bin uncertainty on b^0_{ik} is δ^0_{ik}, and the bin-by-bin uncertainty on $b^\pm_{ikm}$ is $\delta^\pm_{ikm}$, then

$$b_{ik} = b^0_{ik} + \delta_{ik} \cdot \eta_{ik}, \tag{9.8}$$

with

$$\delta_{ik} = \delta^0_{ik} + \sum_{m=1}^{N_{\text{syst}}} \left[(\delta^+_{ikm} - \delta^0_{ik}) H(\theta_m) + (\delta^-_{ikm} - \delta^0_{ik}) H(-\theta_m) \right] \cdot |\theta_m|, \tag{9.9}$$

where η_{ik} are the associated nuisance parameters which are given Gaussian priors centered on zero with unit width.

Putting together all sources of uncertainties, by applying first the shape uncertainties, followed by the bin-by-bin MC statistical uncertainties and finally the overall rate uncertainties, the final expression takes the form:

$$\begin{aligned} b_{ik} = & \left\{ b^0_{ik} \cdot \left(1 + \sum_{m=1}^{N_{\text{syst}}} \left[\kappa^+_{ikm} H(\theta_m) + \kappa^-_{ikm} H(-\theta_m) \right] \cdot |\theta_m| \right) + \delta_{ik} \cdot \eta_{ik} \right\} \\ & \cdot \left\{ \prod_{m=1}^{N_{\text{syst}}} \left(1 + \frac{\rho^+_{km} + \rho^-_{km}}{2} \cdot \theta_m + \frac{\rho^+_{km} - \rho^-_{km}}{2} \cdot \theta^2_m \right) \right\}. \end{aligned} \tag{9.10}$$

The rate uncertainties are applied multiplicatively because most of them affect the rates by scale factors, such as luminosity uncertainty, or acceptance uncertainties, and they are applied last because they affect the distorted shapes in the same way as undistorted shapes. Multiple shape uncertainties are treated additively because most of them correspond to events migrating from one bin to another.

Because the same systematic uncertainties can show up in signal or background samples, and because signal template are also affected by statistical uncertainties in each bin, an expression similar to Equation 9.10 is obtained for the signal predictions in each bin s_i.

To summarize, the likelihood function depends on the observed data $\boldsymbol{n} = \{n_i\}$, the signal scale factor β, the nuisance parameters $\boldsymbol{\theta} = \{\theta_m\}$ and $\boldsymbol{\eta} = \{\eta_{s,i}, \eta_{b,ik}\}$, as well as the central values of the signal and background predictions $\boldsymbol{s} = \{s^0_i\}$ and $\boldsymbol{b} = \{b^0_{ik}\}$, and the rate, shape, and bin-by-bin uncertainties $\boldsymbol{\rho} = \{\rho^\pm_{s,m}, \rho^\pm_{b,km}\}$, $\boldsymbol{\kappa} = \{\kappa^\pm_{s,im}, \kappa^\pm_{b,ikm}\}$, $\boldsymbol{\delta} = \{\delta^0_{s,i}, \delta^\pm_{s,im}, \delta^0_{b,ik}, \delta^\pm_{b,ikm}\}$:

$$L = L(\boldsymbol{n}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\rho}, \boldsymbol{\kappa}, \boldsymbol{\delta}) = \prod_{i=1}^{N_{\text{bins}}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \tag{9.11}$$

where μ_i is given by Equation 9.3, b_{ik} is given by Equation 9.10 and similarly for s_i, and δ_{ik} is given by Equation 9.9.

9.2.2. Cross Section Measurement

The total cross section of single top quark production σ_{s+t} is measured assuming the SM ratio between s-channel and t-channel production. A Bayesian marginalization technique [11] is used to incorporate the effects of systematic uncertainty. This means that each nuisance parameter is assigned a prior that gives the degree of belief of different possible values and then marginalized, or integrated, to calculate the reduced likelihood as a function only of the parameter of interest (the cross section):

$$L'(\boldsymbol{n}|\beta) = \int L(\boldsymbol{n}|\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\eta},\boldsymbol{s},\boldsymbol{b},\boldsymbol{\rho},\boldsymbol{\kappa},\boldsymbol{\delta})\pi(\boldsymbol{\theta})\pi(\boldsymbol{\eta})\mathrm{d}\boldsymbol{\theta}\mathrm{d}\boldsymbol{\eta}, \tag{9.12}$$

where the π functions are the Bayesian priors assigned to each nuisance parameter. The priors are unit Gaussian functions centered on zero which are truncated whenever the value of a nuisance parameter would result in a non-physical prediction. In this case, zero represents the central value and ± 1 represents the result of a shift up or down by the amount quoted as the uncertainty. This is a conservative treatment because most uncertainties are chosen to include at least 95% of the possible range of values, whereas treating an uncertainty as one standard deviation on a Gaussian covers only 68%.

Assuming a flat prior for positive values of β

$$\pi(\beta) = \begin{cases} 1 & : \quad \beta \geq 0 \\ 0 & : \quad \beta < 0 \end{cases},$$

and using Bayes' theorem one can obtain the posterior probability density as a function of the parameter of interest:

$$p(\beta|\boldsymbol{n}) = \frac{L'(\boldsymbol{n}|\beta)\pi(\beta)}{\int_{-\infty}^{\infty} L'(\boldsymbol{n}|\beta)\pi(\beta)\mathrm{d}\beta}. \tag{9.13}$$

The most probable value of the single top quark cross section corresponds to the maximum of the posterior curve, which occurs at $\beta^{\max}$. Thus, the measured single top quark cross section is given by:

$$\sigma_{s+t}^{\mathrm{meas}} = \beta^{\max} \cdot \sigma_{s+t}^{\mathrm{SM}}. \tag{9.14}$$

The uncertainty corresponds to the shortest interval containing 68% of the integral of the posterior:

$$0.68 = \int_{\beta_{\mathrm{low}}}^{\beta_{\mathrm{high}}} p(\beta) d\beta. \tag{9.15}$$

This prescription has the property that the numerical value of the posterior on the low end of the interval is equal to that on the high end of the interval.

Because the signal template shapes and the $t\bar{t}$ background template rates and shapes are functions of m_t, the single top quark cross section is quoted assuming a top quark mass of $m_t = 175\ \mathrm{GeV}/c^2$; and $\partial\sigma_{s+t}/\partial m_t$ is also evaluated. Therefore the uncertainty on the top quark

mass is not included when measuring the cross section. This measurement is repeated with separate Monte Carlo samples and background estimates generated with masses of 170 GeV/c^2 and 180 GeV/c^2, and the result is used to find $\partial\sigma_{s+t}/\partial m_t$.

Extraction of $|V_{tb}|$

In the Standard Model, the single top quark cross section is proportional to the square of the CKM matrix elemtent V_{tb}, therefore $|V_{tb}|$ is identified as the square root of the parameter β

$$|V_{tb}| = \sqrt{\beta} = \sqrt{\frac{\sigma_{s+t}^{\mathrm{meas}}}{\sigma_{s+t}^{\mathrm{SM}}}}. \tag{9.16}$$

This is true only under the assumption that $|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$, and that new physics contributions affect only $|V_{tb}|$; neither a 3×3 structure of the CKM matrix nor its unitarity needs to be assumed for the extraction of $|V_{tb}|$. The theoretical uncertainty on $\sigma_{s+t}^{\mathrm{SM}}$ must be introduced for this calculation. The 95% confidence lower limit on $|V_{tb}|$ is calculated by requiring $0 \leq |V_{tb}|^2 \leq 1$ and finding the point at which 95% of the posterior probability curve (including the systematic uncertainties from the top quark mass and single top cross section) lies to the right of the point. This calculation uses a prior which is flat in $|V_{tb}|^2$.

9.2.3. Hypothesis Testing and Significance Calculation

In order to compute the expected and observed significance of the analysis, a modified frequentist approach is used, which extends a technique developed at LEP [171]. Except for the treatment of the systematic uncertainties, which is Bayesian-like, a frequentist approach is used to caracterize the significance by a p-value for single top quark production. The p-value is the probability of observing an outcome of an experiment at least as signal-like as the one observed, assuming that a signal is absent. We follow the convention that a p-value less than 1.35×10^{-3} constitutes evidence for a signal, and that a p-value less than 2.87×10^{-7} constitutes a discovery. These corresponds to the one-sided integrals of the tails of a unit Gaussian distribution beyond $+3\sigma$ and $+5\sigma$, respectively.

For this approach, two hypotheses are considered. The null hypothesis H_0, assumes a Standard Model production of all processes except for single top quark production ($\beta = 0$). The test hypothesis H_1, assumes background and single top quark signal production at a rate given by the Standard Model prediction ($\beta = 1$). The goal of the analysis is to observe single top quark production, that means to reject the null hypothesis H_0. Experimental outcomes are ranked based on

the test statistic given by the likelihood ratio

$$-2\ln Q = -2\ln\frac{L(\boldsymbol{n}|\beta=1,\hat{\boldsymbol{\theta}}_1,\hat{\boldsymbol{\eta}}_1)}{L(\boldsymbol{n}|\beta=0,\hat{\boldsymbol{\theta}}_0,\hat{\boldsymbol{\eta}}_0)}, \tag{9.17}$$

where $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\eta}}_1$ are the best-fit values of the nuisance parameters which maximize L given the data $\boldsymbol{n}$ for the test hypothesis H_1, and $\hat{\boldsymbol{\theta}}_0$ and $\hat{\boldsymbol{\eta}}_0$ are the best-fit values of the nuisance parameters which maximize L for the null hypothesis H_0.

The choice of a test statistic is arbitrary, however, Equation 9.17 is motivated by the Neyman-Pearson lemma [172], which indicates that a likelihood ratio is the most sensitive variable for separating hypotheses. Similarly, the fits to the nuisance parameters in the test statistic are not intended to incorporate systematic uncertainties, but are included to optimize the sensitivity of the hipothesis testing. It has been found that the sensitivity of the test statistic is improved by fitting for the most important nuisance parameters: the heavy-flavor fraction in $W+$jets events and the mistag rate. Fitting to other nuisance parameters do not appreciably improve the sensitivity and is computationally expensive. Furthermore, fitting for nuisance parameters for which piecewise linear interpolations are used (such as all shape uncertainties), yields discontinuities in the first derivatives of the likelihood function with respect to the nuisance parameters, resulting in posible incorrect solutions of the uncertainties quoted by the MINUIT [173] package used for the maximization. However, the heavy-flavor fraction in $W+$jets events and the mistag rate do not have such discontinuities since they only control the overall rates of the major backgrounds, and they have smooth quadratic dependences. Therefore only these two nuisance parameters are fit for.

The desired p-value is then

$$p = p(-2\ln Q \leq -2\ln Q_{\rm obs}|H_0), \tag{9.18}$$

since signal-like outcomes have smaller values of $-2\ln Q$ than background-like outcomes. Systematic uncertainties are included not in the definition of $-2\ln Q$, which is a known function of the observed data and is not uncertain, but rather in the expected distributions of $-2\ln Q$ assuming H_0 or H_1, since our expectation is what is uncertain. These uncertainties are included in a Bayesian fashion by fluctuating the values of the nuisance parameters in the generation of sets of pseudo-experiment ensembles, referred to as "prior-predictive ensembles". In practice, this is done by filling histograms of $-2\ln Q$, one for each hypothesis, with the results of simulated pseudo-experiments. The pseudo-data is randomly drawn from Poisson distributions of mean given by the predicted distributions (Eq. 9.10) after varying the nuisance parameters according to their Gaussian prior distributions.

Once the ensebles of pseudo-experiments for the two hypotheses are generated, the observed significance is determined from the distribution of $-2\ln Q$ for the H_0 hypothesis and the observed

test statistic in data using Equation 9.18. The sensitivity of the analysis is defined as the expected significance, where $-2\ln Q_{\rm obs}$ is replaced by the expected median from the distribution for the H_1 hypothesis (see Figure 9.4). Optimizations of the analysis were based on the median expected p-value, without reference to the observed data. Indeed, the data events passing the event selection requirements were hidden during the analysis optimization. Figure 9.4 shows that only 35 out of 403^6 backgound only pseudo-experiments fall below the median of the signal+background distribution, this corresponds to an expected p-value of 8.7^{-8}, or equivalently 5.2 standard deviations in Gaussian statistics.

In the computation of the observed and expected p-values, all sources of systematic uncertainties are included in the pseudoexperiments, including the theoretical uncertainty in the signal cross sections and the top quark mass. Because the observed p-value is the probability of an upward fluctuation of the background prediction to the observed data, it depends only weakly on the predicted signal model, and in particular, almost not at all on the predicted signal rate. Hence the inclusion of the signal rate systematic uncertainty in the observed p-value has practically no impact, and the shape uncertainties in the signal model also have little impact (the background shape

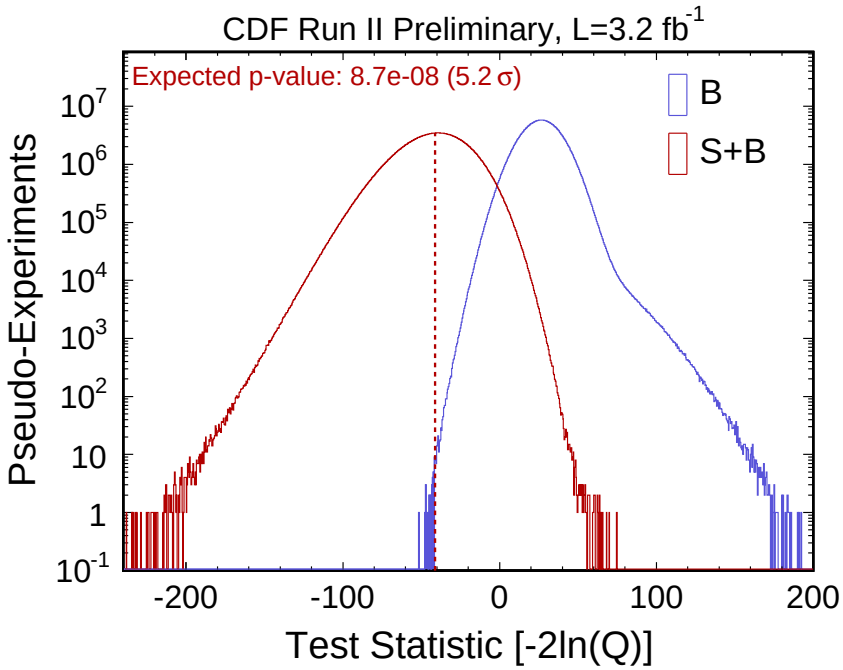


Figure 9.4: Distributions of the test statistic using 403 million of simulated pseudo-experiments for the test hypothesis H_1 (or the signal + background hypothesis, $s+b$) and the null hypothesis H_0 (or the background-only hypothesis, b). The dashed red line indicates the median of the distribution for the test hypothesis H_1. The expected significance is calculated as the area of the H_0 distribution below the H_1 meadian divided by its total area.

uncertainties are quite important though). On the other hand, the expected p-value and the cross section measurement depend on the signal model and its uncertainties.

9.3. Results

Once the Boosted Decision Tree classifiers are trained, templates created, cross checks completed, and systematic uncertainties accounted for, the discriminant distributions of the data is ready to be compared to the Monte Carlo templates. Figure 9.5 show the distribution of BDT

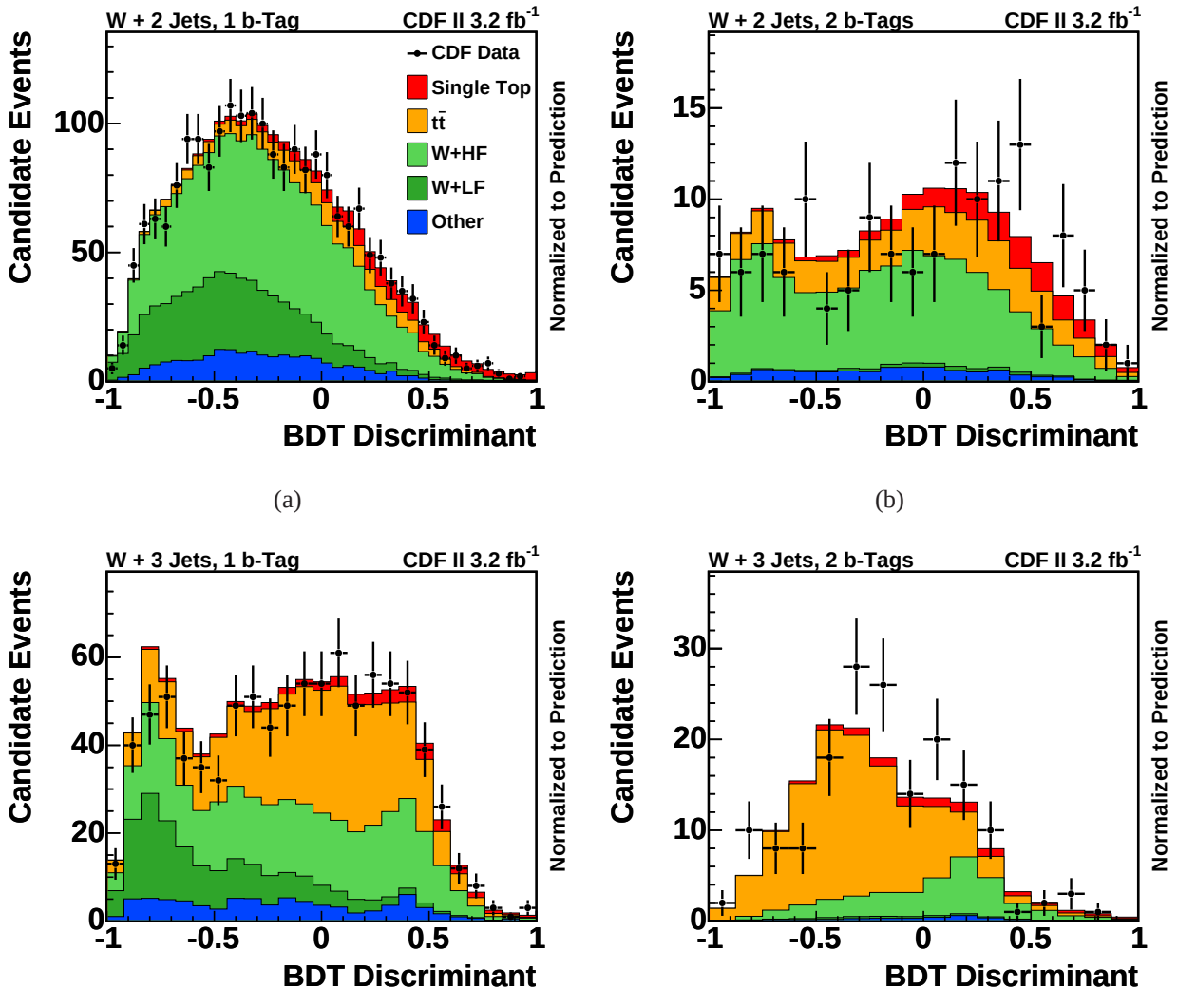
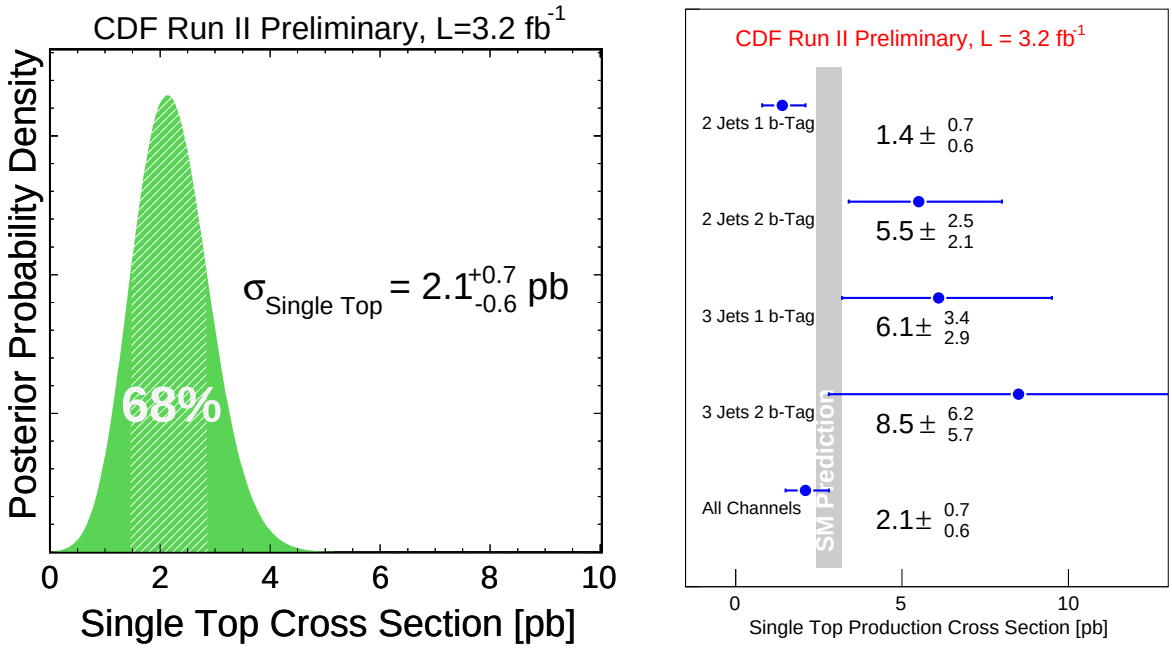


Figure 9.5: Distributions of the output of the BDT applied to the data compared to the predictions from Monte Carlo. The four signal regions are shown: (a) two jets and one b-tag events, (b) two jets and two b-tag events, (c) three jets and one b-tag events, (d) three jets and at least two b-tag events.

discriminants applied to the observed candidate events in the four signal regions, compared to the expected distributions. Figure 9.5(a) show the distribution for the most sensitive sample, it can be seen qualitatively that, in the most signal-like bins, the data fall a bit lower than the Standard Model expectation including the signal, however, it is evident that the data prefers the prediction that includes single top quark production. A quantitative analysis is given in the following sections.

9.3.1. Cross Section Measurement

To extract the s- and t-channel combined single top quark production cross section, the posterior probability density is constructed as a function of the cross section from the reduced likelihood including all rate and shape uncertainties by marginalization. Figure 9.6(a) shows the posterior probability function which yields a measurement of the combined single top quark cross section of $\sigma_{s+t} = 2.1^{+0.7}_{-0.6}$ pb assuming a top quark mass of 175 GeV/c^2. The dependence on the top quark mass is $+0.02\,\mathrm{pb}/(\mathrm{GeV}/c^2)$. This is below the theoretical NLO prediction of $\sigma_{s+t} = 2.86 \pm 0.36$ pb [12, 13], although still statistically consistent with it. A summary of



(a) (b)

Figure 9.6: (a) The posterior probability density as a function of the single top quark cross section. The measurement of the top quark cross section is extracted from as the most probable value of the curve. The uncertainty corresponds to the shortest interval containing 68% of the integral of the posterior. (b) Results of the cross section measurement in each signal sample.

the single top quark cross section measurements in each signal sample is shown in Figure 9.6(b)

9.3.2. Significance of the Result

Figure 9.7 shows the distributions of $-2\ln Q$ for both tested hypotheses, with n value $-2\ln Q_{\rm obs} = -18.8$ indicated by the arrow. This results in a significance given by the p-value of 0.0002 which corresponds to 3.5 standard deviations. The deficit in data in the most signal-like bins of the most sensitive sample (see Figure9.5(a)) made the significance to fluctuate downward from the expected $5.2\,\sigma$ to the observed $3.5\,\sigma$. This is not enough to claim observation, however it is a strong evidence that the data rejects the null hypothesis H_0, with no single top production, in favor of the test hypothesis H_1, with a Standard Model single top quark production.

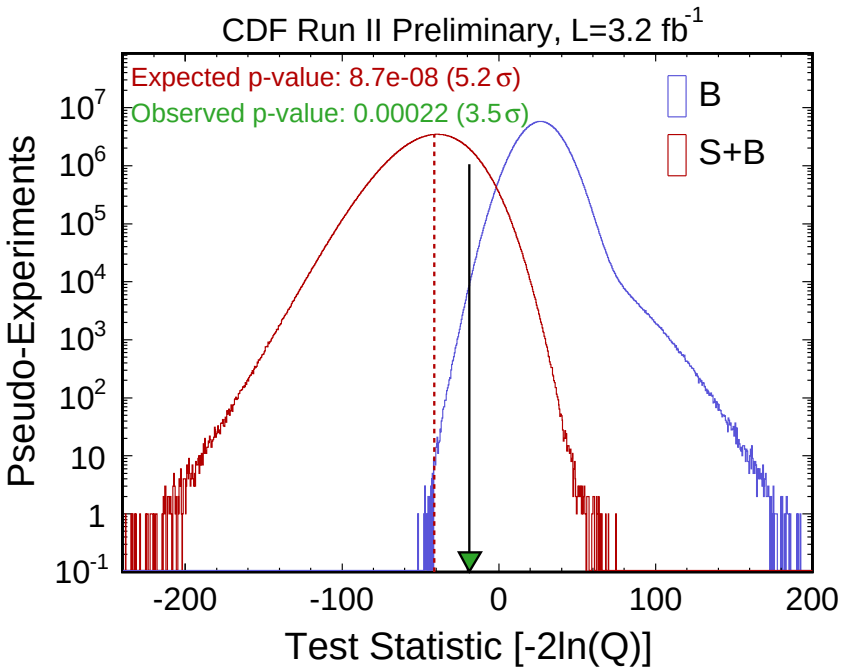


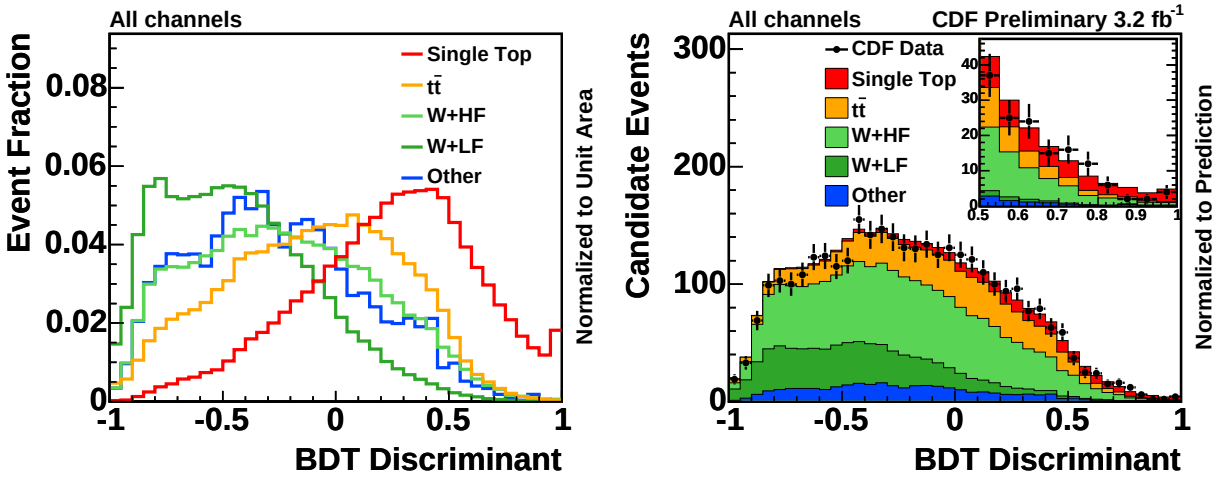
Figure 9.7: Observed test statistic compared to the distributions for the null hypothesis H_0, with no single top production, and the test hypothesis H_1, with signal at the SM rate. The arrow indicates the observed value.

Chapter 10

OBSERVATION OF SINGLE TOP QUARK PRODUCTION

Five separate analysis searching for single top quark production are developed at CDF. Four of them, including the Boosted Decision Tree analysis explained in this thesis, share the same lepton+jets candidate event selection described in Chapter 5 using an integrated luminosity of 3.2 fb^{-1}. These analyses ara based on Boosted Decision Trees (BDT), Neural Networks (NN) [174], Matrix Elements (ME) [175], and multivariate Likelihood Functions [176, 177]. The BDT, NN, and ME analyses use exactly the same events including data from the lepton triggers and the $\not\!\!E_T$+jets trigger (see Section 5.1). The LF analysis selects events using the same candidate event selection but using only data form the lepton triggers. Because there is 100% overlap in the data and Monte Carlo events selected by these analyses, and they are correlated among each other at the order of about 70%, additional gain in sensitivity is expected from a combination. A natural combination technique is to use the individual analysis discriminant outputs as inputs to a "Super Discriminant" (SD) multivariate analysis, based on a genetically evolved neural network.

In addition to the lepton+jets analysis, a fifth search is developed in the orthogonal $\not\!\!E_T$+jets sample where no lepton is identified. This analysis (MJ) [178] adds about 30% to the signal acceptance using an integrated luminosity of 2.1 fb^{-1}. The results of this analysis are combined with the results of the super discriminant analysis to yield the final results of the search for combined s- and t-channel single top quark production at CDF. With the combination of all analyses, CDF observes single top quark production with a significance of 5.0 standard deviations [27].



(a) (b)

Figure 10.1: Predicted templates (a), and data to Monte Carlo comparison (b) for the BDT discriminant.

10.1. Lepton + Jets Analyses

10.1.1. Boosted Decision Trees

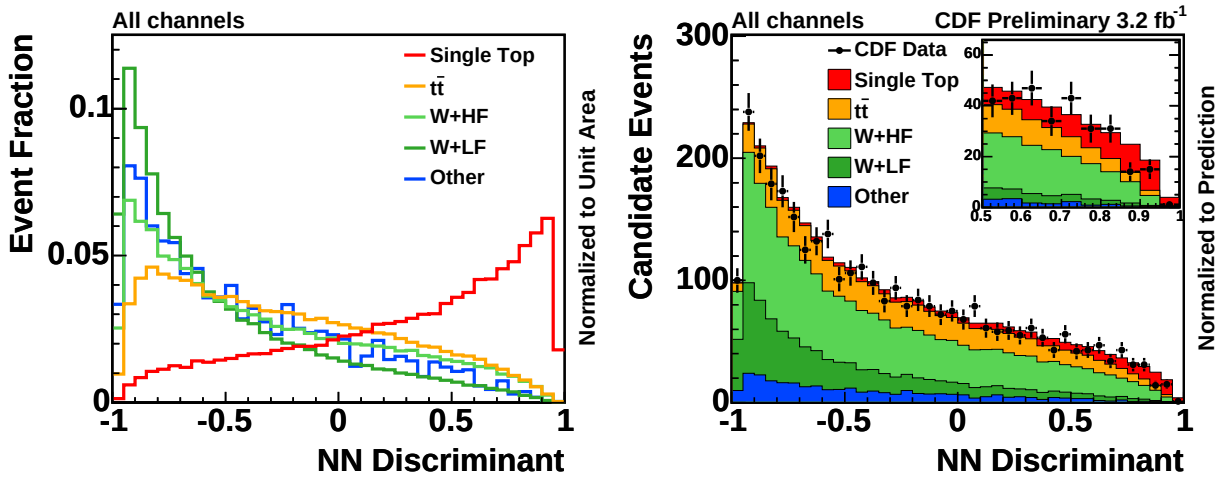
The BDT analysis is listed here for completeness, a detailed description can be found in Chapter 8. Figure 10.1 show the templates of the BDT discriminant for the signal and background processes, and the distributions comparing the observed BDT output in data to the predicted BDT output for all signal channels added toghether.

10.1.2. Artificial Neural Networks

A different approach uses artificial neural networks to combine sensitive variables to distinguish single top quark signal from background events. As with the neural network flavor separator b_{NN} described in Section 4.3.6, the networks are construct with the NeuroBayes® [144] package, which combines a three-layer feed-forward neural network with a complex and robust preprocessing of the input variables. Bayesian regularization techniques are utilized to avoid over-training.

Four different neural network are trained, one for each signal region, treating t-channel events as signal for all samples except the two-jet two-b-tag events, in which s-channel events are treated as signal. The background training sample is a mix of the predicted Standard Model process in the predicted ratios.

Each training starts with more than fifty variables, but the training procedure removes those with no significant discriminating power, reducing the number to 11–18 input variables. Among



(a) (b)

Figure 10.2: Predicted templates (a), and data to Monte Carlo comparison (b) for the NN discriminant.

the most important are: the invariant mass of the reconstructed top quark, the total scalar sum of transverse energy in the event, the dijet mass, the charge of the lepton times the pseudorapidity of the light-quark jet, the b_{NN} jet-flavor separator output, and the cosine of the angle between the lepton and the light-jet in the top quark rest frame. Each neural network has one hidden layer of 15 nodes and one output node which returns a value between -1 and 1, where events near 1 are very signal-like. More detailed information about this method can be found in [174].

Figure 10.2 shows the NN discriminant shapes and the comparison between the predicted NN discriminant and the observed LF discriminant in data for all considered channels added toghether.

10.1.3. Matrix Element Method

The matrix element method relies on the evaluation of event probabilities for signal and background processes based on calculations of the relevant SM differential cross sections. These probabilities are calculated on an event-by-event basis for the signal and background hypotheses and quantify how likely it is for the event to have originated from a given signal or background process. Rather than combine many complicated variables, the matrix element method uses only the measured energy-momentum four-vectors of each particle to perform its calculation.

The probability density of a given process is constructed by integrating over the parton-level differential cross-section $\mathrm{d}\sigma$; it can be reduced to the expression:

$$P(x) = \frac{1}{\sigma} \int 2\pi^4 |M|^2 \frac{f(x_1)}{|E_{q_1}|} \frac{f(x_2)}{|E_{q_2}|} W(y,x) d\Phi_4 dE_{q_1} dE_{q_2}, \tag{10.1}$$

where σ and $|M|$ are the cross section and the matrix element of the process, $f(x_i)$ and E_{q_i} are the

PDFs and energies of the incident partons, $d\Phi_4$ represents the 4-body phase, and $W(y,x)$ is the transfer function between partonic and measured quantities. The transfer functions $W(x,y)$ are used to include detector resolution effects. Lepton quantities and jet angles are considered to be well measured. However, jet energies are not, and their resolution is parameterized from Monte Carlo simulation to create a jet resolution transfer function. The integration is performed over the quark energies and over the unobserved z-component of the neutrino four momentum to create a final probability density. The matrix element for the event probability at tree-level is calculated using MADGRAPH [131]. Event probability densities are computed for all significant signal and background processes that can be easily modeled to first order: s-channel and t-channel single top quark production as well as the $Wb\bar{b}$, Wcg, Wgg and $t\bar{t}$ processes. In the specific case of the $t\bar{t}$ matrix element, additional integrations are performed over the momenta of particles not detected. More detailed information about this method can be found in [175].

The event probabilities are used as ingredients to build an event probability discriminant (EPD), a variable for which the distributions of signal events and background events are as different as possible. The EPD is defined to be $EPD = P_s/(P_s + P_b)$. This discriminant is close to zero if $P_b \gg P_s$ and close to unity if $P_s \gg P_b$. Several background processes in this analysis have no b-jet in the final state, and the matrix element probabilities do not include detector-level discrimination between b-jets and non-b-jets. In order to include this extra information, the neural net jet flavor separator is used to define the b-jet probability as $b = 0.5 \cdot (b_{\mathrm{NN}} + 1)$, and use it to weight each matrix element probability by the b flavor probability of its jets. Since single top quark production always has a b quark in the final state, we write the event-probability-discriminant as:

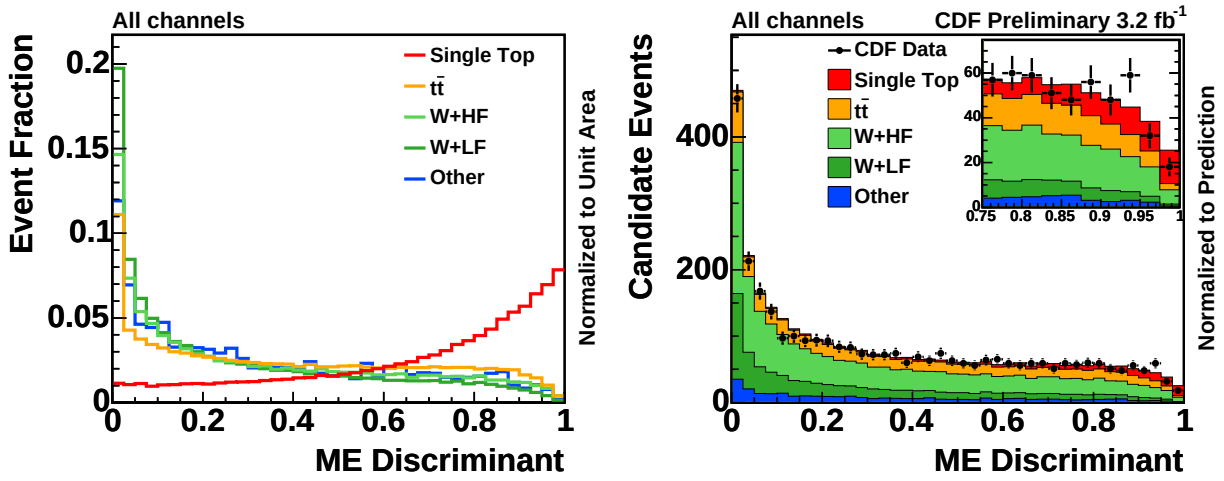
$$EPD = \frac{b \cdot P_{\mathrm{single-top}}}{b \cdot (P_{\mathrm{s}} + P_{Wb\bar{b}} + P_{t\bar{t}}) + (1-b) \cdot (P_{Wc\bar{c}} + P_{Wcg} + P_{Wgg})}. \tag{10.2}$$

The resulting templates and distributions for the ME discriminant are shown in Figure 10.3 with all signal channels added toghether.

10.1.4. Multivariate Likelihood Funtion

A multivariate likelihood function [179] is a method for combining several sensitive variables. This method makes use of the relative probabilities of finding an event in histograms of each input variable, compared between the signal and the background.

The likelihood function L_k for event class k is constructed using binned probability density functions for each input variable. The probability that an event from sample k will populate bin j of input variable i is defined to be f_{ijk}. The probabilities are normalized so that $\sum_j f_{ijk} = 1$ for all variables i and all samples k. For the signal, $k = 1$, and in this paper, four background classes are used to construct the likelihood function: $Wb\bar{b}$, $t\bar{t}$, $Wc\bar{c}/Wc$, and W+LF, which are event classes



(a) (b)

Figure 10.3: Predicted templates (a), and data to Monte Carlo comparison (b) for the ME discriminant.

$k = 2$, 3, 4, and 5, respectively. Histogram underflows and overflows are properly accounted for. The likelihood function for an event is computed in two steps. First, for each reconstructed variable i, the bin j in which the event falls is obtained, and the quantities

$$p_{ik} = \frac{f_{ijk}}{\sum_{m=1}^{5} f_{ijm}}, \tag{10.3}$$

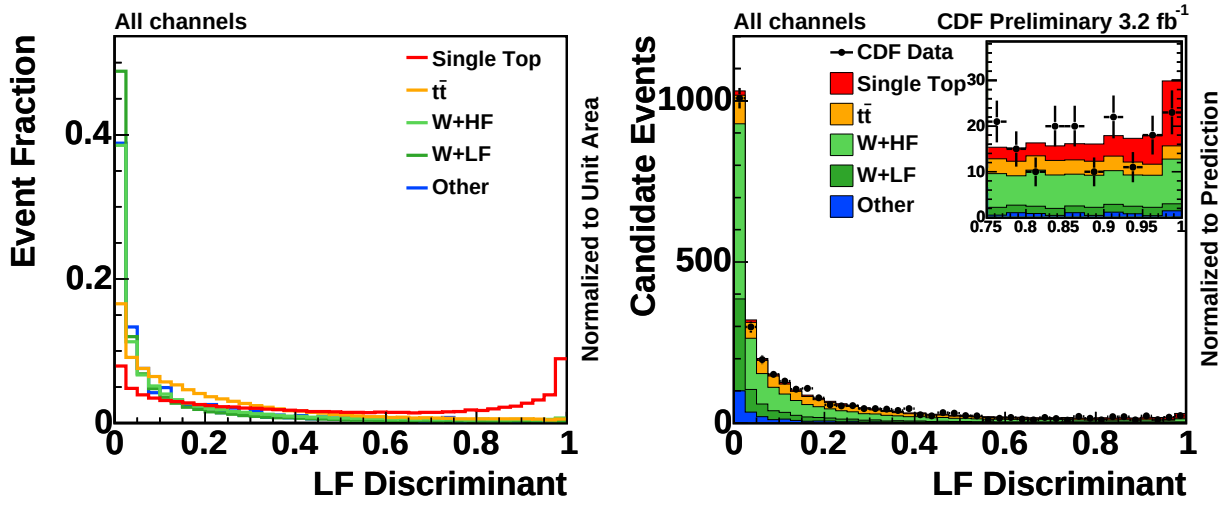
are computed for each variable i and each event class k. The p_{ik} are used to compute

$$\mathcal{L}_k = \frac{\prod_{i=1}^{n_{\rm var}} p_{ik}}{\sum_{m=1}^{5} \prod_{i=1}^{n_{\rm var}} p_{im}}, \tag{10.4}$$

where $n_{\rm var}$ is the number of input variables. The signal likelihood function is the one which corresponds to the signal class of events, $\mathcal{L}_1$. This method does not take advantage of the correlations between input variables, which may be different between the signal and the background processes. The predicted distributions of the likelihood functions are made from fully simulated Monte Carlo and data sets where appropriate, with all correlations in them, and so while correlations are not taken advantage of, they are included in the necessary modeling. More detailed information on this method can be found in [176] and [177].

Three likelihood functions are computed for use in the search for single top quark production: one using the t-channel as signal for events with two jets and one b tag (L_t); another optimized for the s-channel signal and which is applied to two-jet two-tag events (L_s); and a third the sum of both s- and t-channel single top quark production as the signal for events with three jets (L_{3j}). Even though the dedicated s-channel LF analysis is developed for further event samples[1], only the

[1]The original s-channel search include two and three jet events, with at least two identified as b-jets using the SecVtx algorithm and/or the JetProb [180] algorithm. Events from the $\not\!\!E_T$+jets trigger are also used.



(a) (b)

Figure 10.4: Predicted templates (a), and data to Monte Carlo comparison (b) for the LF discriminant.

subset of two jets with two b-tags from the lepton triggers is used in the lepton+jets combination.

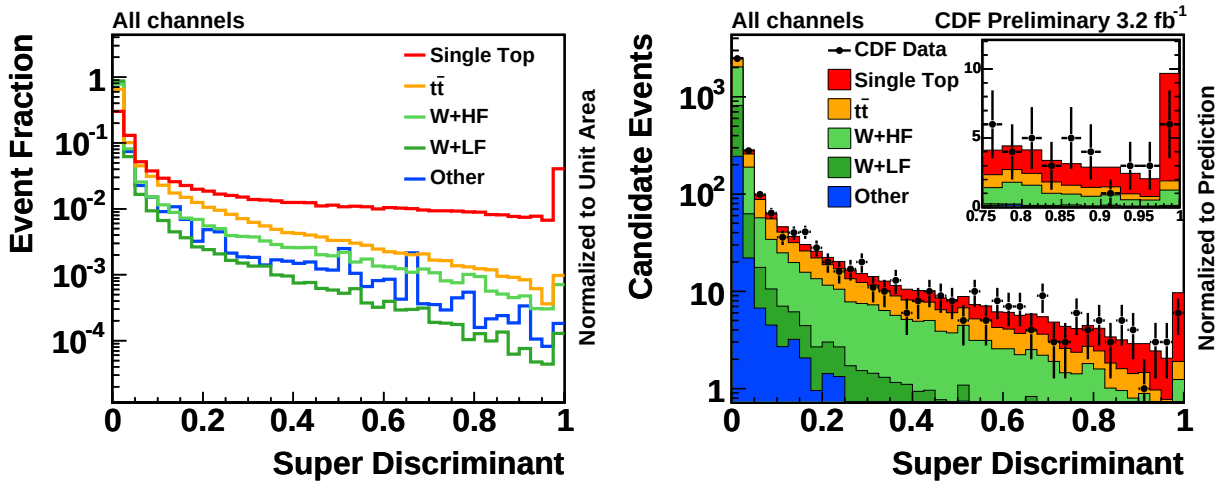
Among the most important input variables to the Likelihood function in the different channels are: the invariant mass of the reconstructed top quark, the total scalar sum of transverse energy in the event, the dijet mass, the charge of the lepton times the pseudorapidity of the light-quark jet, the χ^2 value of a kinematic solver used to find the most likely four-vector of the neutrino, the b_{NN} jet-flavor separator output, and the cosine of the angle between the lepton and the light-jet in the top quark rest frame.

Figure 10.4 shows the LF discriminant shapes and the comparison between the predicted LF discriminant and the observed LF discriminant in data for all considered channels.

10.1.5. Super Discriminant Analysis

The discriminant outputs of the individual lepton+jets analyses are combined into a single super discriminant (SD) using a genetically evolved neural network similar to the one used in CDF's published evidence for single top quark production [26]. The neural network weights and topology are optimized using a technique known as Neuro-Evolution of Augmenting Topologies (NEAT) [181].

For each of the eight considered channels, NEAT begins from a population of neural networks generated from a seed network by randomly varying the network weights. The evolution then proceeds in generations, where in each generation, the following three steps are completed: first, the fitness of each neural network classifier is evaluated by calculating its performance using a figure



(a) (b)

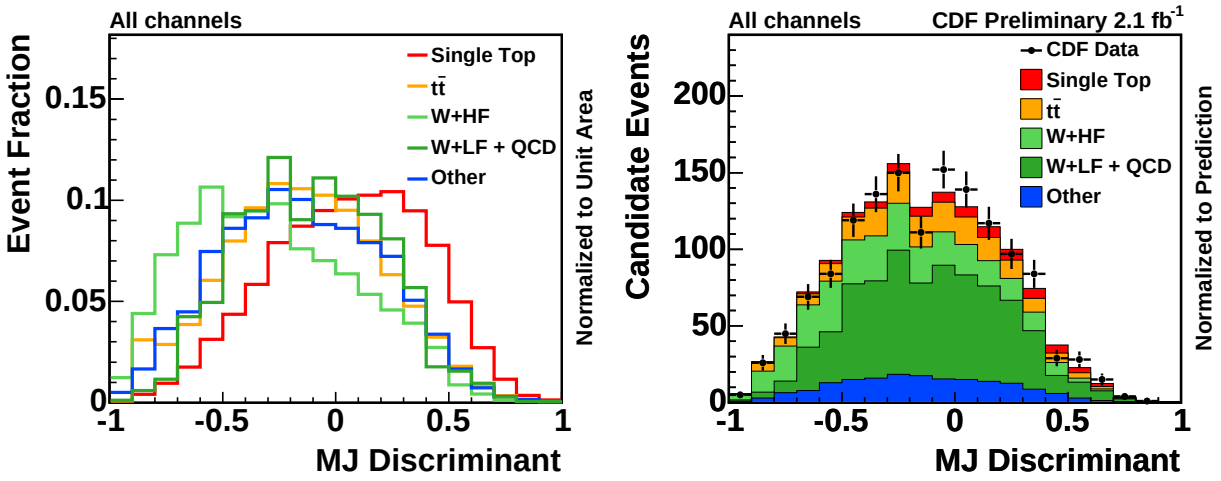
Figure 10.5: Predicted templates (a), and data to Monte Carlo comparison (b) for the SD combination.

of merit highly correlated to the expected significance. Second, neural networks with poor fitness are removed from the population. And last, the remaining classifiers are allowed to replenish the population through mutation and breeding. Possible mutations include randomly changing one or more NN weights, randomly adding a link between nodes, and randomly adding new nodes. Breeding involves blending randomly selected features from two neural networks. The population of neural networks remaining at the end of this process for one generation becomes the initial population for the next generation. At the end of the evolution, the classifiers with the highest fitness values are collected. A detailed estimation of the expected significance including all systematic uncertainties is subsequently performed in order to select the final neural network. With the SD analysis the a priori sensitivity improves by at least 13% over the best individual analysis, resulting in a median expected significance of $> 5.9\,\sigma$. The super discriminant of the sum of all eight considered channels is illustrated in figure 10.5.

10.2. $\not\!\!E_T$ + Jets Analysis

The MJ analysis [178] is designed to select events with $\not\!\!E_T$ and jets, while vetoing events selected by the lepton+jets analyses. It accepts events in which the W boson decays into τ leptons and those in which the electron or muon fails the lepton identification criteria.

The advantage of this analysis is that it is orthogonal to the lepton+jets analysis, increasing the signal acceptance by ∼30%. The disadvantage is the huge instrumental background due to QCD events in which mismeasured jet energies produce large $\vec{\not\!\!E}_T$ aligned in the same direction as jets.



(a) (b)

Figure 10.6: Predicted templates (a), and data to Monte Carlo comparison (b) for the MJ discriminant.

To reduce this background, a QCD removal neural network is used suppressing 77% of the QCD background while keeping 91% of the signal acceptance.

Finally, the MJ discriminant uses a neural network to combine information from several input variables. The most important variables are the invariant mass of the $\vec{\not\!\!E}_T$ and the second leading jet, the scalar sum of the jet energies, the $\not\!\!E_T$, and the azimuthal angle between the $\vec{\not\!\!E}_T$ and the jets. Figure 10.6 shows the MJ discriminant shapes and the comparison between the predicted MJ discriminant and the observed LF discriminant in data for all considered channels.

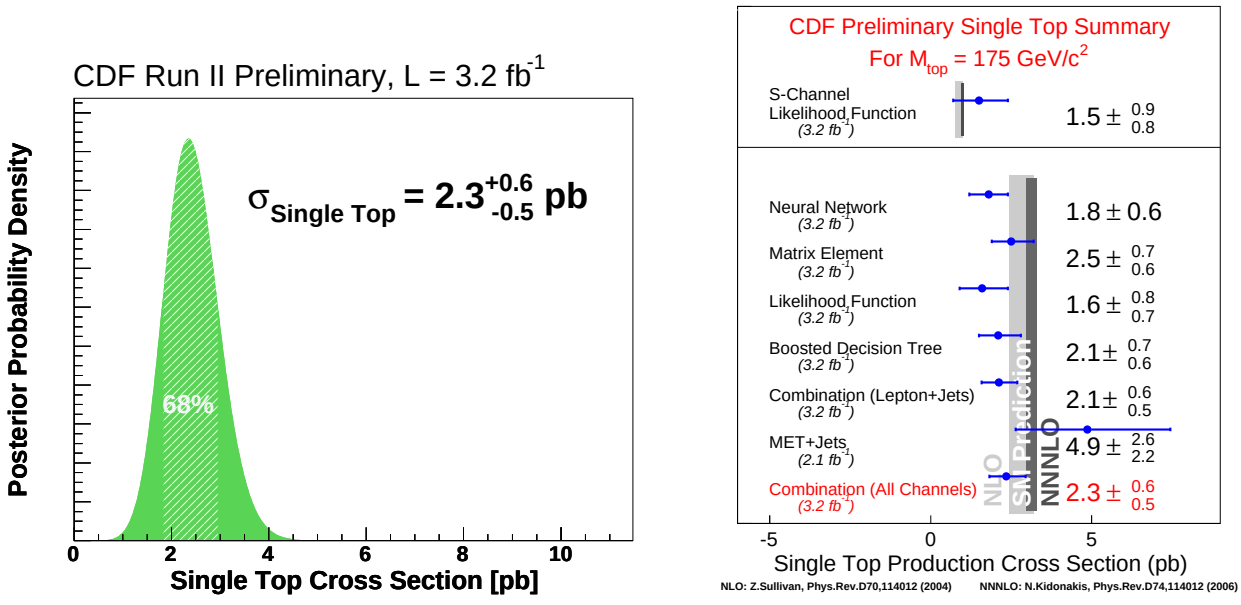
10.3. Combination Results

Each analysis separately measures the single quark top quark production cross section and calculates the signifficance of the observed excess using the same statistical method as described in Section 9.2. The cross sections and the observed and expected significance of each individual analysis is shown in Table 10.1.

The discriminants of the lepton+jets SD analysis is combined with the output of the MJ analysis in the ortoghonal $\not\!\!E_T$+jets sample. Because the samples have no overlapping events, they can be combined as independent channels using the same likelihood technique. Fitting all channels according to the procedure described in Section 9.2, we obtain a single top quark cross section of $\sigma_{s+t} = 2.3^{+0.6}_{-0.5}$ pb, assuming a top quark mass of 175 GeV/c^2. The dependence on the top quark mass is $+0.02$ pb/(GeV/c^2). The significance of this result is given by the p-value of 3.1×10^{-7} which corresponds to 5.0 standard deviations. The posterior probability used to extract the

Analysis	Cross Section (pb)	Significance (Std. Dev.)	Sensitivity (Std. Dev.)
LF	$1.6^{+0.8}_{-0.7}$	2.4	4.0
ME	$2.5^{+0.7}_{-0.6}$	4.3	4.9
NN	$1.8^{+0.6}_{-0.6}$	3.5	5.2
BDT	$2.1^{+0.7}_{-0.6}$	3.5	5.2
LFS	$1.5^{+0.9}_{-0.8}$	2.0	1.1
SD	$2.1^{+0.6}_{-0.5}$	4.8	> 5.9
MJ	$4.9^{+2.5}_{-2.2}$	2.1	1.4
Comb.	$2.3^{+0.6}_{-0.5}$	5.0	> 5.9

Table 10.1: A summary of the results of each individual analysis, with their measured cross-sections, observed significance, and sensitivity (expected p-value). These are combined into a super discriminant (SD), which is combined with the orthogonal $\not\!\!E_{\rm T}$+jets sample (MJ) to make the final CDF combination. The LFS analysis results shown here only measure the s-channel production cross section, while the other analyses measure the sum of the s- and t-channel cross sections.



(a) (b)

Figure 10.7: (a) The posterior probability density as a function of the single top quark cross section for the combination. (b) Results of the cross section measurement in each individual analysis andcombination.

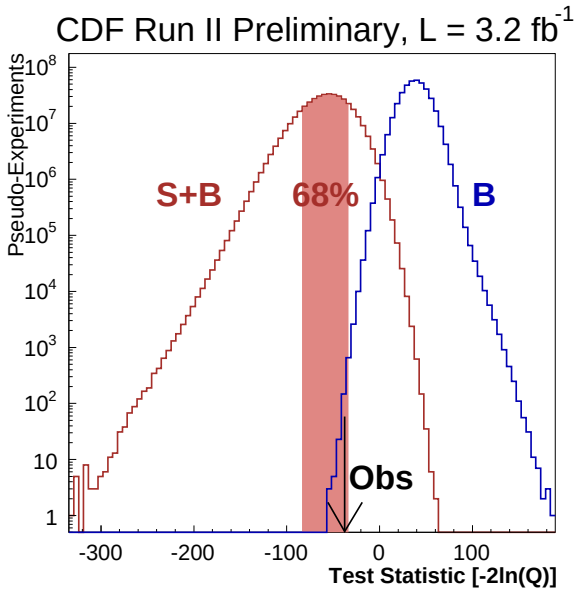


Figure 10.8: Observed test statistic compared to the distributions for the null hypothesis H_0, with no single top production (B), and the test hypothesis H_1, with signal at the SM rate (S+B) for the combination. The arrow indicates the observed value.

cross section, as well as a summary of each cross section measurement in each individual analysis is shown in Figure 10.7. Figure 10.8 compares the observed value of the test statistic with the corresponding distributions of both tested hypotheses.

10.3.1. Extraction of the CKM Matrix Element $|V_{tb}|$

As discussed in Section 9.2.2, the single top cross section is proportional to the CKM matrix element $|V_{tb}|$, therefore the $|V_{tb}|$ value can be extracted from the measured single top quark cross section and from $\sigma_{\mathrm{s+t}}^{\mathrm{theo}} = 2.86 \pm 0.36$ pb [12, 13] at $m_t = 175\ \mathrm{GeV}/c^2$. This corresponds to $|V_{tb}| = 0.91 \pm 0.11$(stat.+syst.)± 0.07(theory), where the uncertainty on the theoretical predicted cross section is propagated.

It is also possible to set a limit on $|V_{tb}|$. Figure 10.9 shows the posterior probability for $|V_{tb}|^2$ marginalizing over all systematic uncertainties (including the single top quark cross section and uncertainties on the top quark mass) and assuming a flat prior in $0 \leq |V_{tb}|^2 \leq 1$. The curve is integrated until 95% of the area is included resulting in a 95% confidence level lower limit of $|V_{tb}| > 0.71$.

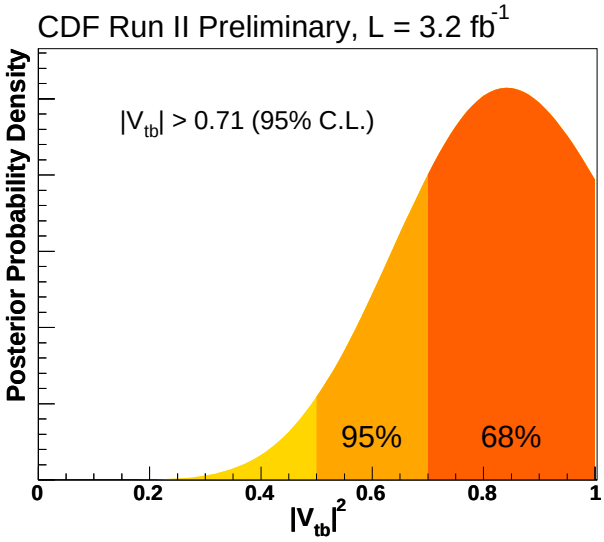


Figure 10.9: Posterior probability curve for the $|V_{tb}|$ limit calculation.

Chapter 11

CONCLUSIONS

The establishment of the electroweak single top quark production at CDF is experimentally challenging. The small single top signal hidden under large uncertain background processes makes it necessary an excellent understanding of the detector and a detailed study of the processes involved. Moreover, simple counting experiments are not sufficient to extract enough information from the candidate event sample and multivariate analysis techniques are crucial to distinguish signal from background. This thesis presents the world's most sensitive individual search, together with CDF's Neural Network analysis, for the combined s- and t-channel single top production. This analysis uses a dataset that corresponds to an integrated luminosity of 3.2fb^{-1}, and is based on a Boosted Decision Tree method that combines information from several input variables to construct a final powerful discriminant, reaching a sensitivity to the combined single top quark production equivalent to 5.2σ. The measured combined single top quark production cross section is $2.1^{+0.7}_{-0.6}$ pb assuming a top quark mass of 175 GeV/c^2. The probability that this result comes from a background-only fluctuation (p-value) is 0.0002, which corresponds to 3.5σ.

This result of this analysis is combined with the other single top analysis at CDF to reach a sensitivity that exceeds 5.9σ. The observed signal has a p-value of 3.1×10^{-7} which corresponds to 5.0σ, thus electroweak single top quark production is conclusively observed [27]. The combination measures a cross section of $2.3^{+0.6}_{-0.5}$ pb which corresponds to a value of $|V_{tb}| = 0.91 \pm 0.11$(stat.+syst.)± 0.07(theory) and a 95% confidence level lower limit of $|V_{tb}| > 0.71$.

With the current integrated luminosity acquired by the CDF detector (above 6fb^{-1}), the systematic uncertainties will start to be the limiting factor for future single top analysis. In order to improve significantly the precision of the single top cross section measurement and consequently the parameter $|V_{tb}|$, the systematic uncertainties, which are currently conservative, will need to be

considerably reduced. On the other hand, now that the electroweak single top quark production is firmly established, doors are open for the study of the properties of the single top quark. For instance, single top quark production provides a unique opportunity for the study of the polarization of the top quarks, as the Standard Model expects top quarks to be nearly 100% polarized when produced singly via electroweak interaction, while other exotic models predict a significant deviation from the SM expectation.

Apéndice A

RESUMEN EN CASTELLANO

El Modelo Estándar

El Modelo Estándar (ME) de las partículas elementales es una teoría cuántica de campos que describe los componentes fundamentales del universo y sus interacciones. Mediante principios de simetría, el formalismo teórico del ME explica la interacción nuclear y electromagnética entre las partículas, describiendo satisfactoriamente y con un nivel de precisión sin precedentes una gran variedad de resultados experimentales acerca de las propiedades de las partículas y sus interacciones.

Según el ME, la Naturaleza está formada por unos pocos constituyentes, básicos, puntuales y sin estructura llamados partículas elementales, y que se pueden dividir en dos conjuntos cerrados de acuerdo a su spin: fermiones y bosones. Toda la materia conocida está compuesta de fermiones, partículas de spin semientero que siguen la estadística de Fermi-Dirac. La interacción entre los fermiones es debida al intercambio de bosones, partículas de spin entero que siguen la estadística de Bose-Einstein.

Se puede distinguir dos grupos de partículas fermiónicas: quarks y leptones. Los quarks se pueden presentar en seis sabores diferentes: up (u), down (d), charm (c), strange (s), top (t) y bottom (b); y están ligados por la interacción fuerte, formando hadrones de tres quarks de valencia (bariones), o de quark-antiquark (mesones). Por su parte, hay seis tipos distintos de leptones: el electrón (e^-) y el neutrino electrónico (ν_e), el muón (μ^-) y el neutrino muónico (ν_μ), el tau (τ^-) y el neutrino tauónico (ν_τ). Cada quark, q, y cada leptón, ℓ^-, lleva asociado una antipartícula con la misma masa pero carga opuesta, $\bar{q}$ y ℓ^+ respectivamente. Los quarks y leptones se pueden agrupar en tres generaciones como se muestra en la Tabla 2.1. En esta tabla se muestra también la carga y

masa de estas partículas. Estas tres generaciones presentan una jerarquía de masas muy llamativa, siendo el top, con mucho, el quark más pesado. Comprender la razón detrás de esta jerarquía y la estructura de las generaciones es una de las preguntas abiertas de la física de partículas a las escalas de energía de los actuales experimentos en física de partículas..

Las fuerzas de la naturaleza que actúan entre los quarks y los leptones se describen a través de campos cuantizados. Las interacciones entre partículas elementales se deben al intercambio de cuantos que son los mediadores de las fuerzas. El ME incorpora la fuerza electromagnética, responsable de la emisión de luz por parte de los átomos excitados, la fuerza débil, la causante, por ejemplo, de la desintegración nuclear beta, y la fuerza fuerte que mantiene a los núcleos estables. La gravitación no está incluida en el marco del ME sino por la teoría general de la relatividad. Todas las partículas con masa o energía sienten la fuerza gravitacional. Sin embargo, debido a la debilidad de la gravitación con respecto a las otras fuerzas elementales, ésta no juega un papel importante en las reacciones de partículas fundamentales.

Las fuerzas electromagnética, débil, y fuerte se describen mediante las llamadas teorías cuánticas de campos gauge. Los cuantos de estos campos tienen spin 1 y se llaman bosones gauge. La Tabla 2.2 muestra la carga y la masa de estos bosones. El mediador de la fuerza electromagnética es el fotón (γ) que es una partícula sin masa, los masivos $W^{\pm}$ y Z^0 son los mediadores de la fuerza débil, y los ocho gluones (g), sin masa, de la fuerza fuerte.

Los quarks pueden experimentar interacción electromagnética, débil, y fuerte. Todos los leptones experimentan la interacción débil y los cargados también la electromagnética, sin embargo, los leptones no toman parte en interacciones fuertes.

Para explicar la masa de fermiones y bosones en el ME, es necesario incluir un término adicional en el Lagrangiano de la teoría mediante un mecanismo de ruptura espontánea de simetría llamado mecanismo de Higgs. La introducción de ruptura de simetría en la teoría, no sólo dota de masa a fermiones y bosones, sino que también tiene como consecuencia la aparición de un nuevo campo real y escalar que está asociado con una partícula observable masiva y de spin 0, conocida como el bosón de Higgs. El último parámetro por determinar en la teoría electrodébil del ME es la masa del bosón de Higgs, partícula aún no observada debido a su pequeño acoplamiento, dejándolo como la última pieza en la teoría electrodébil por ser testada experimentalmente.

El acoplamiento entre quarks de diferentes generaciones está permitido en el ME, posibilitando transiciones entre autoestados de masa de quarks de diferentes sabores mediante interacciones de corrientes cargadas (intercambio de bosones $W^{\pm}$). A esto se le llama mezclado de generaciones y su descripción viene dada por la matriz unitaria 3×3 de Cabbibo-Kobayashi-Maskawa (CKM) (expresión 2.18), cuyos elementos de matriz $|V_{q_1 q_2}|$ son proporcionales al acoplamiento entre los quarks q_1 y q_2 y el bosón intermediario W. Uno de los parámetros libres de la matriz de unitariedad

CKM corresponde a un término de fase, siendo éste el responsable de los fenómenos de violación de la simetría CP en procesos con cambio de sabor en el Modelo Estándar. Uno de los retos de la física de altas energías es la determinación experimental de los valores de los elementos de la matriz CKM, ya que, por ejemplo, un desvío en la unitariedad de dicha matriz sería indicación de la existencia de física más allá del Modelo Estándar.

Física del Quark Top

El quark top es la más masiva de las partículas elementales conocidas, su masa es de $173{,}1 \pm 1{,}3\ \mathrm{GeV}/c^2$, unas cuarenta veces más pesado que el segundo quark más pesado, el quark bottom. Su gran masa, del orden de la escala de la ruptura de simetría, indica que puede jugar un papel importante en el mecanismo de generación de masa, haciendo del quark top un campo de estudio ideal como prueba de física más allá del Modelo Estándar. Además, también debido a su gran masa, el tiempo medio de desintegración del quark top, $\tau_{\mathrm{top}} \approx 4 \cdot 10^{-25}$ s, es extremadamente corto, más corto que el tiempo característico de formación de hadrones. Como consecuencia, no se forman hadrones top y todas las propiedades del quark top, incluyendo el spin, se pasan a sus productos de desintegración, proporcionando una oportunidad única para el estudio del quark top al desnudo. Por otra parte, debido a que $|V_{tb}| \approx 1$, el quark top se desintegra casi exclusivamente en un quark b y un bosón W. Desintegraciones en quarks de la primera y segunda generación están fuertemente suprimidas por los pequeños valores de los elementos $|V_{td}|$ y $|V_{ts}|$ de la matriz CKM.

La sección eficaz de producción de pares top-antitop ($t\bar{t}$) mediante interacción fuerte es de aproximadamente 7 pb, siendo su característica signatura de desintegración en dos bosones W y dos quarks b muy difícil de imitar por sucesos provenientes de procesos de fondo. Esto hizo posible que, tras el descubrimiento del quark bottom en 1977 que puso en evidencia la existencia de una tercera generación de quarks, la carrera por el descubrimiento del quark top finalmente viese su fin en 1995 por las colaboraciones CDF y DØ en el acelerador Tevatron de Fermilab con aproximadamente 100 pb^{-1} de datos de colisiones $p\bar{p}$ a energía centro de masas de $\sqrt{s} = 1{,}8$ TeV. El Tevatron es de hecho el único lugar en el que se pueden producir y estudiar directamente quarks top hasta la fecha.

El Quark Single Top

La producción de quarks top solitarios (single top) también es posible en el ME a través de la interacción electrodébil. Los procesos dominantes a energías del Tevatron son el intercambio de bosones W a través de los canales s y t mostrados en las Figuras 2.8(a) y 2.8(b). La sección eficaz de producción esperada a segundo orden (NLO) de estos dos procesos es $\sigma_s = 0{,}88 \pm 0{,}11$ pb

y $\sigma_t = 1{,}98 \pm 0{,}25$ pb, respectivamente. Existe un tercer proceso, la producción asociado de un quark top junto con un bosón W (Figura 2.8(c)), pero en comparación con los otros dos tiene una sección eficaz de producción despreciable a energías del Tevatron.

Las razones para el estudio del quark single top están bien motivadas. La sección eficaz de producción del quark single top es directamente proporcional al elemento $|V_{tb}|$ de la matriz CKM, permitiendo una oportunidad única para la medida directa de dicho elemento y proporcionando así una una prueba interesante de la unitariedad de la matriz CKM. Un espectro muy amplio de extensiones al ME predicen desviaciones observables de σ_s o σ_t con respecto a los valores del ME, por lo que dichas medidas podrían dar información hacia modelos de cuatro generaciones, modelos de corrientes neutras con cambio de sabor, u otros fenómenos más allá del ME. Debido a que la signatura del quark single top comparte estado final con la producción del bosón de Higgs en el canal $WH \rightarrow \ell\nu b\bar{b}$, una medida del primero es fundamental para tener bajo control todos los procesos de fondo del segundo, a la vez que las herramientas y técnicas desarrolladas para la búsqueda del primero pueden ser utilizadas para la búsqueda del segundo.

La medida de la sección eficaz de producción del quark single top queda bien motivada, pero cabe resaltar la dificultad que tiene la misma en el Tevatron. La sección eficaz de producción predicha es la mitad que la de la producción de pares $t\bar{t}$, y con sólo un quark top en el estado final, la señal es mucho menos nítida frente a los procesos de fondo. La tasa a la cual un bosón W es producido junto con jets, con al menos uno de ellos identificado como un jet proveniente de la fragmentación de un quark b, es unas 12 veces mayor que la tasa de producción de señal. Además, la incertidumbre en los procesos de fondo es aproximadamente un factor tres veces más grande que la cantidad de señal esperada. Para poder observar la producción del quark single top, una muestra pura de señal debe ser separada de los procesos de fondo, para ello se utiliza una técnica de análisis multivariable llamada *Boosted Decision Tree*.

Dispositivo Experimental

El análisis presentado en esta tesis utiliza los datos recogidos entre marzo del 2002 y agosto del 2008 de colisiones protón-antiprotón producidas por el Tevatron y observadas por el detector CDF II.

El Tevatron es un complejo acelerador-colisionador de protones y antiprotones situado en Fermilab (Fermi National Accelerator Laboratory) en Batavia, Illinois (EEUU). Con una energía de centro de masas de $\sqrt{s}$ = 1.96 TeV es la fuente de colisiones protón-antiprotón ($p\bar{p}$) a más alta energía y la única máquina capaz de producir y estudiar quarks top en la actualidad. Las colisiones tienen lugar cada 396 ns en dos puntos de un anillo bajo tierra de cerca de 1 km de radio , donde

hay situados sendos detectores: CDF II y DØ. Este anillo es el último eslabón de una compleja cadena de aceleradores que produce y acelera haces de protones y antiprotones. Esta cadena de aceleradores se muestra en la Fig. 3.1.

El detector CDF II es un detector de partículas de carácter general con forma de barril y compuesto por varias capas de subdetectores diseñadas para detectar y medir las propiedades de las partículas producto de colisiones $p\bar{p}$. Para expresar posiciones y ángulos se utiliza un sistema de coordenadas cilíndrico, con el eje z a lo largo de haz de protones. El ángulo azimutal ϕ alrededor del eje del haz se define con respecto un eje horizontal que apunta hacia fuera del centro del anillo del Tevatron, y el radio r se mide con respecto del eje del haz. El ángulo polar θ está definido con respecto de la dirección del haz de protones, y la pseudorapidez η está definida como $\eta = -\ln\tan(\theta/2)$. La energía y el momento transverso de una partícula se definen como $E_T = E\sin\theta$ y $p_T = p\sin\theta$, respectivamente.

La parte más interna del detector es el sistema de trazas, consistente en un detector de silicio y una cámara de deriva, y diseñado para medir el momento de las partículas cargadas. El detector de silicio consiste de tres subsistemas que cubriendo una región que va desde inmediatamente después del tubo del haz hasta 29 cm del eje del haz. La COT es una cámara de deriva de 3.1 m cuyo volumen activo cubre una región radial que va de 40 a 137 cm, proporcionando una cobertura en pseudorapidez de $|\eta| \leq 1$. El sistema de trazas permite una precisa reconstrucción tridimensional de la traza de las partículas cargadas, del vértice primario de interacción, y a su vez es utilizado para identificar vértices secundarios asociados con la desintegración de partículas de larga vida media.

El sistema de trazas está rodeado por el detector de tiempo de vuelo, compuesto de 216 barras centelladoras de unos 300 cm de largo, y diseñado para proporcionar una buena identificación de partículas a bajo momento transverso. Estos dos sistemas se encuentran inmersos dentro de un solenoide superconductor capaz de generar un campo magnético de $1{,}4$ T paralelo al eje del haz. El campo magnético es usado para curvar las trazas cargadas, permitiendo así medir su momento transverso.

Este imán está rodeado por el sistema de calorimetría, formado por calorímetros electromagnéticos y hadrónicos, que miden la energía de las partículas que interaccionan con la materia que lo forman. Éstos están divididos en segmentos con forma de torres proyectivas, cada una cubriendo un pequeño rango en azimut y pseudorapidez. La región central, $|\eta| < 1{,}1$, está formada por el CEM y el CHA. La región exterior, $1{,}1 < |\eta| < 3{,}6$, consta de los colorímetro PEM y PHA. Deposiciones de energía en los calorímetros electrónicos son usadas para la identificación de electrones y medida de su energía, mientras que jets son identificados y medidos a través de la energía que depositan tanto en torres de calorímetros electrónicos como hadrónicos.

En la parte más externa del detector se encuentran los sistemas de muones, compuestos de cámaras de deriva y centelladores. Detrás del calorímetro central se encuentran cuatro capas de cámaras de deriva (CMU), y más allá, detrás de una capa adicional de acero de 60 cm, se encuentran otras cuatro capas de cámaras de deriva (CMP) seguidas de centelleadores (CSP). El CMU, CMP y CSP cubren la misma región central $|\eta| < 0{,}6$. Las cámaras CMX, seguidas de los centelleadores CSX, extienden la cobertura en pseudorapidez del sistema de muones desde 0.6 a 1.0, cubriendo así la región completa de la COT. Muones con $1{,}0 < |\eta| < 1{,}5$ son detectados por un último subsistema de cámaras de muones (BMU).

No todas las colisiones producidas por el Tevatron son interesantes desde el punto de vista de la física de altas energías. Las colisiones ocurren con una frecuencia de 2.5MHz, y cada suceso ocupa unos 250 kB en disco. Dado que con la tecnología actual no hay modo práctico de almacenar tal cantidad de datos, es necesario diseñar un sistema de filtrado a tiempo real que reduzca la cantidad de datos a niveles manejables seleccionando sólo los sucesos relevantes para el estudio objetivo de estos experimentos. Este sistema, llamado *trigger*, tiene en CDF una arquitectura basada en tres niveles de selección como se muestra en la Fig. 3.16.

Muestra de Datos y Selección de Sucesos

La selección de sucesos candidatos a quark single top explota las características cinemáticas de las signaturas de sus procesos de producción en los canales s y t. Ya que los quarks top se desintegran casi inmediatamente en un bosón W y un quark b, el estado final de un suceso de quark single top presenta un bosón W junto con dos o tres jets (debido posible radiación de gluones extra en los diagramas de producción), con al menos uno de ellos provenientes de la fragmentación de un quark b. La desintegración del W en dos jets, aunque dominante, está sujeta a un fondo de QCD impracticable. Sin embargo, la desintegración semileptónica a un electrón o muón junto con su correspondiente neutrino proporciona una signatura mucho más limpia. El neutrino, que no es directamente detectable, se manifiesta como una gran cantidad de energía faltante ($\not\!\!E_T$) en el balance energético del detector. Debido a la dificultad en la identificación de taus, las desintegraciones tauónicas del bosón W sólo entran en la selección cuando el propio tau se desintegra en un electrón o un muón detectable. En definitiva, la selección debe estar optimizada para seleccionar estados finales que consisten en un leptón cargado (electrón o muón), gran cantidad de energía faltante debida al neutrino, y dos o tres jets, con al menos uno de ellos originado de un quark b.

El primer paso para seleccionar una muestra de sucesos candidatos es elegir los triggers pertinentes que seleccionen sucesos de acuerdo a sus características más distintivas. Las características más prominentes en una muestra de W+jets, donde el W se desintegra leptónicamente, son la

presencia de leptones de alto momento y de alta energía faltante junto con jets energéticos. Por tanto, los triggers elegidos son los triggers de electrones y muones de alto momento, y el trigger de $\not\!\!E_T$+jets. Cuatro triggers diferentes se distinguen entre los triggers de leptones de alto momento. El trigger de electrones CEM selecciona una traza en la COT con $p_T > 9$ GeV/c que apunta a una deposición energética en el calorímetro CEM de $E_T > 18$ GeV. El trigger de electrones PHX requiere deposiciones de energías en el colorímetro PEM de $E_T > 20$ GeV, junto con la presencia de energía faltante $\not\!\!E_T > 15$ GeV. El trigger de muones CMUP busca trazas en la COT con $p_T > 18$ GeV/c que apunten simultáneamente a las dos cámaras centrales de muones, CMU y CMP. El trigger de muones CMX busca trazas de $p_T > 18$ GeV/c que apunten a la cámara de muones CMX y que además sean consistentes con la señal temporal de los centelleadores CSX. El trigger de $\not\!\!E_T$+jets selecciona sucesos con $\not\!\!E_T > 35$ GeVy con dos jets de $E_T > 10$ GeVcon al menos uno de ellos central $|\eta| < 1{,}1$. Las eficiencias de estos triggers están resumidas en la Tabla 5.1 y la Figura 5.1. La muestra de datos que se utiliza en el análisis presentado en esta tesis ha sido recogida utilizando los mencionados triggers, y corresponde a una luminosidad integrada de aproximadamente 3.2 fb^{-1}(ver Tabla 5.1).

Una vez que un suceso es seleccionado por el trigger, se guarda en disco para su posterior análisis experimental. A estos sucesos se les aplica una selección *offline* para crear una muestra final de sucesos candidatos con un contenido de señal maximizado, manteniendo los fondos a niveles lo más bajo posible.

A los candidatos a electrón (CEM y PHX) y muón (CMUP y CMX), seleccionados por los triggers de leptones de alto momento, se les impone una serie de cortes *offline* para mejorar la pureza de su identificación como se muestra en las Tablas 4.1, 4.2 y 4.3. Se definen, además, varios tipos adicionales de muones (EMC), exclusivos entre si y de los muones CMUP y CMX, provenientes del trigger de $\not\!\!E_T$+jets. En la Figura 4.1 se muestra la cobertura en el plano (η, ϕ) de las diferentes categorías a candidato de electrón y muón. Se exige que en el suceso haya exactamente un candidato a leptón aislado con $|\eta| < 1{,}6$. Un leptón se considera aislado si la E_T no asignada al leptón en un cono de $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0{,}4$ centrado alrededor del leptón corresponde a menos del 10 % del E_T del electrón (p_T del muón). Con el objetivo de reducir la tasa de fondo proveniente de procesos de Z+jets, dibosones, y $t\bar{t}$ con estado final dileptónico, se rechazan sucesos con un candidato adicional a leptón, bien sea aislado o no aislado.

Los jets son reconstruidos utilizando un algoritmo de cono de radio $\Delta R \leq 0{,}4$. La energía de los jets es corregida por la dependencia con la pseudorapidez de la respuesta del calorímetro, por la dependencia temporal del calorímetro, y por extra E_T debida a interacciones múltiples. Los candidatos a jet deben tener energía corregida $E_T > 20$ GeV y pseudorapidez $|\eta| < 2{,}8$. La cobertura en pseudorapidez está extendida con respecto a los análisis estándar de quarks top en CDF ($|\eta| < 2{,}0$) debido a que la presencia de un jet con alta pseudorapidez es una propiedad

característica de los sucesos de quark single top. Los sucesos son aceptados si contienen exactamente uno o dos jets. Al menos uno de estos jets debe ser etiquetado como proveniente de la fragmentación de un quark b. Para ello se utiliza el algoritmo de etiquetado SecVtx, que se encarga de buscar vértices secundarias desplazados dentro del jet. La eficiencia de etiquetado del algoritmo SecVtx es aproximadamente del 40 % para jets que contienen quarks b, mientras que para jets que contienen quarks ligeros es menor que el 1 %, por tanto una muestra etiquetada por el algoritmo SecVtx es altamente pura en contenido de b-jets.

La energía faltante se define como $\vec{\not\!\!E}_T = -\sum_i E_T^i \hat{n}_i$, donde E_T^i es la energía transversa de la torre i del calorímetro, $\hat{n}_i$ es el vector unidad en la dirección radial que apunta desde el eje del haz a la torre i, y donde la suma se extiende hasta $|\eta| < 3{,}6$. Está expresión es corregida por varios factores. Primero, debe ser corregida por la posición del vértice primario de interacción, el cual causa un ajuste en la dirección del vector asociado con cada torre. Ya que el cálculo está basado en la energía de las torres de los calorímetros, también se debe ajustar por el efecto de las correcciones de todos los jets. Por último, se corrige por la presencia de muones en el suceso, ya que éstos apenas dejan energía en los calorímetros. Esta corrección se aplica añadiendo el momento transverso del muón en el sumatorio, y sustrayendo la deposición de energía del muón en las torres del calorímetro. Para purificar la muestra en contenido de bosones W con desintegración leptónica, se exige que la $\not\!\!E_T$ corregida del suceso sea mayor que 25 GeV.

Varios cortes adicionales son aplicados para rechazar fondos específicos y mejorando de esta manera tanto la pureza de la muestra final como el modelado de los fondos. Sucesos en los que el leptón y un segundo objeto forman una masa invariante consistente con la masa del bosón Z (76–106 GeV) son rechazados, dejando una contaminación residual de producción de bosones Z muy pequeña. Muones procedentes de rayos cósmicos son drásticamente reducidos rechazando sucesos cuya topología e información temporal del detector de tiempo de vuelo sea consistente con la de un rayo cósmico. Por otra parte, para asegurarnos una buena reconstrucción del suceso, requerimos que la posición z del vértice del suceso esté a menos de 60 cm del centro del detector. Finalmente, para reducir el fondo de QCD que no contiene bosones W, se aplica un corte en la masa transversa del W, definida como $M_T^W = \sqrt{2\left(p_T^\ell \not\!\!E_T - p_x^\ell \not\!\!E_x - p_y^\ell \not\!\!E_y\right)}$. La $\not\!\!E_T$ en los sucesos que no tienen bosones W es típicamente debida a jets erróneamente identificados como leptones, por lo que la $\not\!\!E_T$ apunta en la dirección del leptón resultando en una baja masa transversa. Se exige, pues, que la masa transversa sea mayor que 10 GeV para sucesos con muones, y mayor que 20 GeV para sucesos con electrones. Los sucesos con electrones son más sensibles a la presencia de fondo de QCD, por lo que a estos sucesos se les aplica cortes adicionales en la significancia de la $\not\!\!E_T$ y en correlaciones angulares entre la $\not\!\!E_T$ y la dirección de los jets.

Modelado y Estimación de Procesos

El estado final de un suceso de quark single top es también el estado final de varios procesos de fondo. Los procesos de fondo más difíciles de tratar son los de producción de W+jets, distinguiéndose los casos en que algún jet proviene realmente de quarks de sabor pesado (charm o bottom), o en los que los jets son de sabor ligero (up, down o strange) pero falsamente etiquetados como b-jets. Otros procesos, como producción de pares top, producción de procesos electrodébiles (dibosones, y Z+jets) y procesos de QCD también pueden imitar el mismo estado final de la señal. Un modelado preciso tanto de procesos de señal como de fondo, y una estimación apropiada de la composición de los mismos en la muestra de candidatos final es esencial para medir una señal tan pequeña como es la producción de quarks top solitarios. Muchos de estos procesos pueden simular y estimar utilizando simplemente simulaciones de Monte Carto (MC), mientras que para otros la simulación MC no proporciona un modelo completamente adecuado y necesitan apoyarse en datos en diferentes regiones de control.

Procesos de Señal

Como ya se ha dicho, los procesos dominantes de producción de quark top solitario en el Tevatron son los correspondientes al canal s y t. Para la simulación de estos procesos se utiliza el generador de Monte Carlo MADEVENT, el cual incorpora información completa del color y el spin de las partículas, preservando de esta manera propiedades interesantes de la producción de top quarks solitarios como es la polarización del quark top. Los sucesos generados por MADEVENT son pasados a la rutina de hadronización y showering del programa de MC PYTHIA.

Estudios de precisión han demostrado que las distribuciones cinemáticas del canal s a primer orden (LO) no vienen afectadas por correcciones a segundo orden (NLO). Por tanto la muestra de MC de la señal en el canal s es generada a LO y escalada a la sección eficaz esperada a NLO. Sin embargo, se ha observado que para el canal t las simulaciones a LO no reproducen adecuadamente las distribuciones esperadas de los jets observables, siendo éstas mejor predichas por cálculos a NLO. Para mejorar nuestro modelado del canal t, dos muestras de MC son generadas y combinadas: una para el proceso $2 \to 2$ dominante $b + q \to q' + t$ (Figura 6.1(a)), y otro para el proceso $2 \to 3$ donde un gluón en el estado inicial se divide en un par $b\bar{b}$, $g + q \to q' + t + \bar{b}$ (Figura 6.1(b)). Estas dos muestras se combinan de manera que reproduzcan las distribuciones cinemáticas predichas por el programa ZTOP que opera a NLO.

El número esperado de sucesos después de la selección de candidatos viene dada por

$$N = \sigma \cdot \varepsilon_{\mathrm{evt}} \cdot \mathcal{L}_{\mathrm{int}}, \tag{A.1}$$

donde σ es la sección eficaz predicha para el proceso correspondiente (ver Tabla 7.1), $\varepsilon_{\mathrm{evt}}$ es la

eficiencia de detección del suceso, y $\mathcal{L}_{\rm int}$ es la luminosidad integrada (ver Tabla 5.1). La eficiencia de detección del suceso se determina aplicando la selección de sucesos candidatos a las muestras simuladas de MC. Para ello, se tienen en cuenta tanto las eficiencias de trigger, de identificación de leptones y del etiquetado de jets por el algoritmo SecVtx, como las posibles diferencias observadas en las mismas entre datos y MC. Incluyendo todas las eficiencias de trigger y de identificación, se encuentran eficiencias de detección de sucesos de señal del orden de $\varepsilon_{\rm evt}(t\text{-channel}) = (1{,}2 \pm 0{,}1)\,\%$ y $\varepsilon_{\rm evt}(s\text{-channel}) = (1{,}8 \pm 0{,}1)\,\%$.

Procesos de Fondo $t\bar{t}$ y Electrodébiles

La producción de pares de quarks top $t\bar{t}$, puede ser un fondo sustancial a la producción de quarks single top. Aunque el estado final de este proceso es diferente al de la producción de single tops, este último puede ser imitado si alguna partícula no se detecta o es detectada erróneamente. Esto ocurre en desintegraciones dileptónicas de los quarks top cuando uno de los leptones escapa a la detección, o en desintegraciones de leptones+jets cuando uno o dos de los jets no son detectados (Figura 6.2).

La producción electrodébil de dibosones, incluyendo los procesos WW, WZ y ZZ (Figura 6.3), también constituyen un fondo a la señal del quark single top, sobre todo WW y WZ ya que poseen exactamente el mismo estado final que los tops solitarios. Sin embargo, debido a su pequeña sección eficaz, su contribución a la suma total de fondos es muy pequeña.

Los bosones Z se desintegran bien en dos leptones o bien en dos neutrinos, por tanto sucesos de Z+jets no suelen imitar la signatura de quarks single top. Sin embargo, debido a su gran sección eficaz inclusiva, la producción de Z+jets puede contaminar residualmente la muestra de señal en los casos en que un leptón no es detectado y su energía contribuye a la energía faltante del suceso (Figura 6.5).

Para la simulación MC de las muestras de procesos de $t\bar{t}$ y dibosones se utiliza el programa PYTHIA, que incluye tanto generación de sucesos como showering y hadronización de los mismos. Para la simulación de la muestra de Z+jets se utiliza el programa ALPGEN para la generación de sucesos junto con PYTHIA para una posterior simulación del showering de los mismos. ALPGEN está diseñado para la generación de sucesos con especial énfasis en procesos que contienen un bosón electrodébil junto con varios gluones o quarks radiados en el estado final. Al contrario que PYTHIA, ALPGEN incorpora completamente información del color y del spin de las partículas que intervienen en el diagrama.

El mismo método utilizado para la estimación de la contribución de sucesos de señal a la muestra final de candidatos a quark single top (EcuaciónA.1) se aplica para el cálculo de la correspondiente contribución a la muestra final proveniente de los procesos de $t\bar{t}$, dibosones y Z+jets.

Procesos de Fondo W+jets

El fondo más grande a la señal de quarks single top corresponde a las interacciones entre quarks en las que se radia un bosón W en asociación con múltiples jets (Figura 6.4), ya que puede tener el mismo estado final que el de la señal con una sección eficaz mucho más grande. La mayoría de los quarks en el estado final son de sabor ligero (W+LF), por tanto el nivel de este fondo puede ser reducido mediante el etiquetado de b-jets, sin embargo, debido a su enorme sección eficaz, el fondo W+LF aún contribuye de forma sustancial al fondo total. El etiquetado de jets no es útil en los casos en los que el bosón W es producido junto con un quark de sabor pesado (W+HF), por lo que los procesos $W + b\bar{b}$, $W + c\bar{c}$ y $W + c$ constituyen los fondos más grandes de este análisis.

La muestra de W+jets se simula utilizando un modelo ALPGEN +PYTHIA, al igual que para la muestra de Z+jets. Este modelo reproduce adecuadamente la cinemática de estos procesos, sin embargo, las incertidumbres teóricas en la sección eficaz de estos procesos es muy grande en comparación con el tamaño de la señal. Por tanto, una estimación del número de sucesos basada simplemente en MC (como se ha descrito en las secciones anteriores), no es suficiente para los propósitos de este análisis y un método de estimación basado en datos es necesaria para constreñir adecuadamente la contribución de estos procesos. Este método se detallará más adelante.

Fondo QCD

El fondo más difícil de modelar proviene de sucesos multijet de producción QCD. Para imitar la signatura de un suceso de quark top solitario, uno de los jets debe confundirse con un leptón y una alta energía faltante debe ser creada por malas medidas del suceso (Figura 6.6). Esto es altamente improbable pero la gran sección eficaz de producción QCD hace posible que este fondo pueda contaminar la señal. Debido a la gran cantidad de procesos QCD que pueden producir estas configuraciones, siendo éstos además difíciles de calcular y modelar, y debido a la pequeña posibilidad de que estos sucesos ocurra, es imposible simular estos sucesos con Monte Carlo y es necesario un modelo basado en datos.

Tres modelos diferentes de fondo QCD son usados, todos ellos basados en el principio de que estos sucesos contienen un jet que es falsamente identificado como un leptón. La muestra de *jet-electrons* se crea a partir de un trigger genérico de jets. Dentro de esta muestra se buscan jets que parecen electrones, específicamente se pide que el jet sea energético $E_T > 20$ GeV, que tenga una fracción alta de energía depositada en el calorímetro, entre 80 y 95 %, y que contenga menos de cuatro trazas. Jets con estas características son asumidos electrones asignándoles una carga aleatoriamente, y después el resto de cortes de selección de sucesos es aplicada para crear la muestra de jet-electrons. La muestra de *anti-electrons* se crea a partir del trigger de electrones

centrales. Se buscan candidatos a electrón que fallen al menos dos de los cortes de identificación que no dependen de propiedades cinemáticas. Estos objetos son tratados como electrones y el resto de cortes de selección son aplicados. Debido al pequeño tamaño de estas dos muestras, y ya que sus propiedades cinemáticas son similares, ambas son combinadas para crear la muestra de QCD para electrones CMU. Para electrones PHX, solo se utiliza la muestra de jet-electrons debido a que los anti-electrons no cubren la alta pseudorapidez de los electrones PHX. Remarcablemente, estas dos muestras también simulan adecuadamente el fondo QCD para los candidatos a muones CMUP y CMX, y por tanto también son utilizadas para simular estos sucesos. Para los muones EMC, estas muestras no describen adecuadamente su modelo de QCD ya que las propiedades cinemáticas de estos candidatos a muón son sustancialmente diferentes debido a los diferentes requisitos de trigger. Esta muestra se modela utilizando sucesos de muones provenientes del mismo trigger, pero que fallan el requisito de leptón aislado.

Predicciones Basadas en Datos

Para determinar el número esperado de sucesos de QCD y de W+jets se comienza por la muestra *pretag*, en la que todos los criterios de selección son aplicados excepto el etiquetado de jets. Esta muestra es entre diez y veinte veces más grande que la muestra de señal y contiene mayoritariamente sucesos de W+LF, de modo que nos aseguramos que la estimación es estadísticamente independiente de la muestra de señal. Se espera que los sucesos de fondo de QCD tenga valores pequeños de $\not\!\!E_T$ ya que no contienen un neutrino real. Por tanto, ignorando el corte de $\not\!\!E_T$ se puede crear una muestra con alto contenido en QCD que puede ser ajustada a los datos para obtener la normalización de QCD y W+jets mientras que el resto de fondos se fijan a sus valores esperados (ver Figura 7.1). Contando el número de sucesos de QCD que pasan el corte de $\not\!\!E_T$ se puede calcular la fracción de QCD en la muestra pretag, $F_{\mathrm{QCD}}^{\mathrm{pretag}}$.

Es necesario ahora distinguir entre las componentes de W+HF y de W+LF de la muestra total de W+jets. Para ello se calcula la fracción de sabor pesado mirando directamente a las simulaciones MC. Sin embargo, estas simulaciones no predicen correctamente esta fracción por lo que una calibración de la misma en la región de control de W+1 jet es necesaria. Esto permite una estimación de la contribución de W+HF y de W+LF en la muestra final de señal, aplicando la eficiencia de etiquetado para el primero, y la matriz de probabilidad de etiquetado erróneo para el segundo.

Una vez determinadas las contribuciones W+HF y de W+LF en la muestra final, se puede repetir el ajuste de $\not\!\!E_T$ en la muestra final fijando dichas contribuciones, para así determinar la fracción de QCD en la muestra final.

Los resultados de la estimación de sucesos candidatos en las muestras finales de señal para

los diferentes procesos, así como el número de sucesos observados en datos, son mostrados en la Tabla 7.2.

Boosted Decision Trees

De la tabla de estimación de fondos se deduce que la incertidumbre en los fondos es unas tres veces más grande que la señal, imposibilitando una medida de la sección eficaz de producción de quarks single top mediante un simple experimento de contaje. Es necesario, por tanto, construir un discriminante con una separación más optima entre señal y ruido de manera que sucesos clasificados como señal son utilizados para testar la presencia de señal y medir su sección eficaz, mientras que sucesos clasificados como fondo son usados para mejorar nuestro conocimiento de los procesos de fondo. Esto se consigue mediante técnicas de análisis multivariable, que explotan la información de varias propiedades cinemáticas y de etiquetado del suceso y la correlación entre ellas, para construir un discriminante que clasifica sucesos en un espectro continuo que va de más a menos parecido con sucesos de fondo y de menos a más parecido con sucesos de señal. Estas distribuciones son después ajustadas a las predicciones de señal y fondo para extraer la sección eficaz de producción de quarks single top.

Para el análisis presentado en esta tesis, la técnica de análisis multivariable elegida está basada en los llamados *Boosted Decision Trees* (BDT). Un Decision Tree (DT) es un algoritmo que clasifica sucesos de acuerdo a una secuencia de decisiones binarias, cada una basado en una única variable. Se construye dividiendo la muestra inicial en dos subconjuntos disjuntos que a su vez son recursivamente divididos hasta alcanzar un número mínimo de sucesos. Cada nodo del árbol divide la muestra basándose en la elección de una variable cuyo corte se ajusta para proporcionar una separación óptima entre sucesos de señal y de ruido. La misma variable puede ser utilizada en varios nodos, y alguna variable puede llegar a no ser utilizada. Este proceso resulta en una serie de nodos finales con separación máxima entre señal y ruido. En la Figura 8.6 se muestra un esquema de un Decision Tree. Un DT es muy similar a un análisis de cortes rectangulares. Sin embargo, mientras que los cortes rectangulares simplemente seleccionan un hipercubo en el espacio de fases, un DT es capaz de dividir el espacio de fases en una gran cantidad de hipercubos, cada uno de ellos identificado como de señal o de fondo.

Una deficiencia de un DT es su inestabilidad con respecto a fluctuaciones estadísticas de la muestra de entrenamiento con la que se ha construido el DT. Por ejemplo, si dos variables presentan un poder de separación similar, una fluctuación en la muestra de entrenamiento puede causar que el algoritmo elija una variable diferente a la elegida con la muestra inicial, posiblemente resultando en un DT sustancialmente diferente. Este problema es superado por un procedimiento de

boosting, que extiende el concepto de un árbol a varios árboles que forman un "bosque" de Decision Trees, llamado Boosted Decision Tree (BDT). Cada árbol es construido a partir de la misma muestra de entrenamiento, pero repesando los sucesos que fueron clasificados erróneamente en el entrenamiento del árbol anterior. Los árboles son después combinados en una única variable discriminante que viene dada por una media ponderada de cada DT individual. El algoritmo de boosting estabiliza la respuesta de los DT con respecto a fluctuaciones de la muestra de entrenamiento, a la vez que es capaz de mejorar considerablemente su poder de clasificación.

Entre las variables más importantes en el entrenamiento de los BDT se encuentran las siguientes:

- $Q \times \eta$: la carga del leptón por la pseudorapidez del jet no etiquetado. Valores altos de esta variable son característicos de sucesos de quark single top producidos en el canal t. Esto es debido a que el quark de sabor ligero del estado inicial suele tener la mayoría del momento de la colisión resultando en un jet con alto $|\eta|$. La carga del leptón esta correlacionado con el signo de esta η, ya que dicha carga determina si en el proceso se ha producido un quark o un antiquark top, y a su vez, la producción de quarks top son mayoritariamente iniciados por quarks u en el protón, mientras que antiquarks top son debidos a quarks $\bar{u}$ en el antiprotón.

- $M_{\ell\nu b}$: la masa reconstruida del quark top. Ésta es reconstruida a través de la masa invariante del leptón, el neutrino, y el jet etiquetado. Se espera que la distribución de esta variable se acumule en valores próximos a la masa del quark top para sucesos de procesos que contienen un quark top real.

- H_{T}: la suma escalar de las energías transversas de los jets, del leptón, y de la energía faltante en el suceso. Esta cantidad tiene valores mucho mayores para sucesos con quarks top que para sucesos de W+jets.

- M_{jj}: la masa invariante de los dos jets. Esta variable resulta en valores mas grandes para sucesos que contienen quarks top.

- M_T^W: la masa transversa del bosón W reconstruido. Para sucesos cuyo procesos no contiene un bosón W o en los que la energía faltante no es debida a la presencia de un neutrino, la forma de la distribución de esta variable no corresponde con la de la distribución de bosones W reales.

- b_{NN}: la salida de una red neuronal diseñada para separar entre jets de diferente sabor.

Una vez entrenados los BDTs, se consiguen distribuciones discriminantes cuyos bines más sensibles pueden alcanzar razones señal ruido de $s/b > 5$, en comparación con la muestras de candidatos con las que se comenzó que en el mejor de los casos poseía una razón $s/b \approx 1/15$.

Una serie de tests de validación son realizados para asegurarnos una respuesta adecuada y no sesgada de los resultantes BDTs. Por un lado, el buen acuerdo entre los datos y las predicciones de MC es chequeado en detalle para todas las variables utilizadas en el entrenamiento de los BDTs, tanto en las regiones de señal como en varias regiones de control cuyo contenido en señal es despreciable. Por otro lado, antes de mirar las distribuciones de los discriminantes de los BDT en las muestras de candidatos, se comprueba que dichas distribuciones muestran un buen acuerdo en las regiones de control. Dichas regiones de control corresponden a las muestras de W+2 jet y W+3 jets donde ningún jet es etiquetado, dominadas por W más jets de sabor ligero, y la muestra de W+4 jets con al menos un jet etiquetado, dominada por el proceso de producción de pares $t\bar{t}$.

Análisis Estadístico

Los objetivos de este análisis son observar la producción del quark top solitario y hacer una medida precisa de su sección eficaz. Para el primero, una evaluación de la significancia del exceso de sucesos comparado con la predicción de fondos es esencial. La realización de ambos objetivos tienen en común el logro de una mejor separación entre sucesos de señal y de fondo y una reducción de las incertidumbres de los procesos de fondo. Los contenidos de los bines con bajo cociente señal ruido del discriminante son importantes para una interpretación apropiada de los bines con alto cociente señal ruido, sirviendo como muestra de control, con un contenido reducido de señal, que pueden ser usadas para ayudar a constreñir las predicciones de fondo. Por tanto, un uso simultáneo del contenido de todos los bines del discriminante, comparando las observaciones con las predicciones, es necesario para medir de manera óptima la sección eficaz y computar su significancia.

La función de verosimilitud utilizada para la extracción de la sección eficaz y la determinación de la significancia viene dado por el producto de las probabilidades Poissonianas de cada bin de cada histograma de la variable discriminante en cada región de señal. Las probabilidades de Poisson son función del número de sucesos n_i observado y de las predicciones μ_i en cada bin i, de forma que la función de verosimilitud está dada por la expresión:

$$L = \prod_{i=1}^{N_{\text{bins}}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \tag{A.2}$$

donde la predicción en cada bin es la suma de las contribuciones de señal s_i y de cada fondo b_{ik}:

$$\mu_i = \beta \cdot s_i + \sum_{k=1}^{N_{\text{bkg}}} b_{ik}. \tag{A.3}$$

La predicción de señal viene dada por la suma de los procesos de producción en los canales s y t

asumiendo una razón entre ellas dada por el ME. Esta contribución de señal está escalada por el factor β, que es el parámetro a determinar para la extracción de la medida de la sección eficaz.

La expresión de la contribución de cada fondo viene dada por la Ecuación 9.10, y equivalentemente para la contribución de señal. Estas expresiones complican la función de verosimilitud debido a la presencia de incertidumbres sistemáticas. Las incertidumbres sistemáticas son incorporadas como parámetros *nuisance* mediante una interpretación estadística Bayesiana. Los errores sistemáticos incluyen incertidumbres en la normalización de la estimación de los fondos, e incertidumbres en las formas de las predicciones de las distribuciones discriminantes. Cuando una fuente de incertidumbre afecta tanto a la forma como a la normalización, ambos efectos son tratados como 100 % correlacionados. Por último, las incertidumbres estadísticas debido al limitado tamaño de las muestras de Monte Carlo son tenidas en cuenta bin a bin.

Las fuentes de incertidumbre sistemática están listadas en la Tabla 9.1, incluyendo: incertidumbre en la escala de energía de los jets, eficiencias de trigger, de etiquetado, y de identificación de leptones, incertidumbre en la cantidad de radiación en el estado inicial (ISR) y final (FSR), errores en las funciones de distribución de los partones (PDF), errores en las escalas de factorización y renormalización, e incertidumbres en el modelado MC de ciertas variables pobremente descritas.

Para la medida de la sección eficaz de producción de quarks single top, se utiliza una técnica de marginalización Bayesiana para incorporar el efecto de las incertidumbres sistemáticas. Para ello, a los parámetros nuisance se les asignan *priors* Gaussianos que representan el grado de conocimiento que tenemos acerca de sus valores, y después son integrados para crear una función de verosimilitud reducida en función sólo del parámetro de interés β. El valor más probable de la sección eficaz de producción de quarks single top viene dada por el máximo de la curva de la función de verosimilitud reducida, y su incertidumbre corresponde al intervalo más pequeño conteniendo el 68 % de su integral.

Para el computo de la significancia observada y esperada se utiliza una aproximación frecuentista modificada. Excepto para el tratamiento de las incertidumbres sistemáticas, que es de tipo Bayesiano, un método frecuentista es utilizado para caracterizar la significancia como la probabilidad (p-value) de que una fluctuación de ruido dé un resultado experimental como el observado o más parecido al que se esperaría en presencia de señal. Dos hipótesis son consideradas para ser testadas. La hipótesis nula H_0, asume producción a una tasa predicha por el ME de todos los procesos excepto el de señal ($\beta = 0$). La hipótesis test H_1, asume una producción de acuerdo al ME tanto de procesos de fondo como de señal ($\beta = 1$). Se define el test estadístico como $-2 \ln Q = -2 \ln \frac{L(\boldsymbol{n}|\beta=1,\hat{\boldsymbol{\theta}}_1)}{L(\boldsymbol{n}|\beta=0,\hat{\boldsymbol{\theta}}_0)}$, donde $\hat{\boldsymbol{\theta}}_1$ y $\hat{\boldsymbol{\theta}}_0$ son los valores del mejor ajuste de los parámetros nuisance que maximizan la función de verosimilitud, dados los datos observados $\boldsymbol{n}$, para las hipótesis H_1 y H_0 respectivamente. Para el cálculo de la significancia observada y esperada, es necesario generar un conjunto de pseudo-experimentos para cada hipótesis con cuyos resultados se llenan

histogramas de $-2\ln Q$ para cada hipótesis. Estos pseudo-experimentos son generados simulando pseudo-datos que son elegidos aleatoriamente de distribuciones Poissonianas con media dada por las predicciones de la Ecuación 9.10 una vez fluctuados los parámetros nuisance de acuerdo a distribuciones Gaussianas. Debido a que se esperan valores menores de $-2\ln Q$ para la hipótesis H_1 que para la hipótesis H_0, el p-value se define como la probabilidad $p(-2\ln Q \leq -2\ln Q_{\rm obs}|H_0)$. La significancia observada del análisis se determina calculando esta probabilidad a partir de la distribución de $-2\ln Q$ para la hipótesis H_0. La significancia esperada se obtiene sustituyendo el valor observado $Q_{\rm obs}$ por el valor mediano esperado de la distribución de la hipótesis H_1.

Resultados

La estimación final de sucesos candidatos de señal y fondo, así como el número observado de sucesos candidatos se muestra en la Tabla 7.2. La Figura 9.5 muestra las distribuciones del discriminante BDT aplicado a las muestras de sucesos candidatos observados comparado con las distribuciones esperadas. En la Figura 9.5(a), correspondiente a la muestra más sensible a señal, se ve cualitativamente que los datos caen ligeramente por debajo de los valores esperados por el ME, incluyendo señal, en los bines más sensibles, sin embargo, es evidente que los datos se ajustan mejor a la predicción con señal que a la predicción en ausencia de ella.

Cuantitativamente, se mide una sección eficaz de producción electrodébil de quarks single top de $\sigma_{s+t} = 2{,}1^{+0{,}7}_{-0{,}6}$ pb, asumiendo una masa del quark top de 175 GeV/c^2. Este valor está por debajo de la predicción teórica a NLO de $\sigma_{s+t} = 2{,}86 \pm 0{,}36$ pb, aunque es estadísticamente consistente con ella. El p-value esperado es de $8{,}7 \times 10^{-8}$ que equivale a una desviación de $5{,}2\,\sigma$ con respecto a la hipótesis nula. La significancia del exceso observado en datos corresponde a un p-value de 0.0002 o equivalentemente a una desviación de $3{,}5\,\sigma$ con respecto a la hipótesis de ausencia de señal. Esto no es suficiente reivindicar observación de producción de quarks single top, sin embargo, indica una fuerte evidencia de que los datos rechazan la hipótesis H_0, con ausencia de señal, en favor de la hipótesis H_1, con una producción electrodébil de quarks top de acuerdo al Modelo Estándar.

Observación de Producción Electrodébil de Quarks Top

En CDF existen cinco análisis para la búsqueda de producción electrodébil de quarks single top. Cuatro de ellos, incluyendo el basado en BDTs descrito en esta tesis, comparten la misma muestra de candidatos en la muestra de leptones+jets con 3.2 fb^{-1} ya descrita. Estos análisis están basados en Boosted Decision Trees (BDT), Redes Neuronales (NN), Elementos de Matriz (ME),

y funciones de verosimilitud multivariable (LF). Los análisis de BDT, NN y ME usan exactamente la mismos sucesos incluyendo datos de los triggers de leptones de alto momento y del trigger de $\not\!\!E_T$+jets. El análisis de LF selecciona sucesos utilizando la misma selección de candidatos pero incorporando solo datos del trigger de leptones. Ya que hay un solapamiento del 100 % en los sucesos de datos y de MC seleccionados por estos análisis, y que hay una correlación del orden del 70 % entre cada análisis, se espera obtener una ganancia adicional en sensibilidad de la combinación de ellos. La técnica de combinación consiste en utilizar los discriminantes de cada análisis como variables de entrada de un Super Discriminante (SD) basado en redes neuronales entrenadas mediante algoritmos de evolución genética en los que tanto los pesos como la topología de la red son optimizados. Con esta combinación se alcanza una significancia esperada que excede las $5{,}9\,\sigma$. La Figura 10.5 muestra la distribución observada y esperada del discriminante de la combinación SD.

Además de los análisis de leptones+jets, existe un quinto análisis utilizando la muestra independiente de $\not\!\!E_T$+jets en la que ningún leptón es identificado. Este análisis (MJ) añade alrededor de un 30 % en aceptancia de señal usando una luminosidad integrada de 2.1 fb^{-1}.

Cada análisis mide de manera independiente la sección eficaz de producción del quark single top y la significancia del exceso observado usando el mismo método estadístico ya descrito. Los resultados se muestran en la Tabla 10.1. Tanto el análisis de BDT como el de NN, esperan observar un exceso de señal de más de $5\,\sigma$, sin embargo, el exceso observado en ambos corresponde a $3.5\,\sigma$.

Ya que no hay solapamiento de sucesos entre las muestras de leptones+jets y la de $\not\!\!E_T$+jets, los discriminantes de SD y de MJ son combinados tratándolos como canales independientes en el método estadístico. De esta combinación se obtiene una sección eficaz de producción electrodébil del quark single top de $\sigma_{s+t} = 2{,}3^{+0{,}6}_{-0{,}5}$ pb, asumiendo una masa del quark top de 175 GeV/c^2. La dependencia de esta medida con la masa del quark top es de $+0{,}02\,\mathrm{pb}/(\mathrm{GeV}/c^2)$. La significancia de este resultado está dada por el p-value $3{,}1 \times 10^{-7}$, que equivale a $5{,}0\,\sigma$. Esta significancia se considera suficiente en la comunidad científica para reivindicar observación de la producción electrodébil del quark top.

La sección eficaz de producción de quarks single top es proporcional al elemento $|V_{tb}|$ de la matriz CKM, por tanto, el valor de dicho elemento de matriz puede se extraído de la medida de la sección y de su valor predicho por el ME $\sigma^{\mathrm{theo}}_{\mathrm{s+t}} = 2{,}86 \pm 0{,}36$ pb. El resultado corresponde a $|V_{tb}| = 0{,}91 \pm 0{,}11$(stat.+syst.)$\pm 0{,}07$(theory). También es posible fijar límites a los valores de $|V_{tb}|$. Para ello se marginaliza sobre todas los parámetros nuisance y se asume un *prior* plano en $0 \leq |V_{tb}|^2 \leq 1$. La curva de verosimilitud es integrada hasta cubrir un 95 % del área, para obtener un límite inferior de $|V_{tb}| > 0{,}71$ al 95 % de nivel de confianza.

Conclusiones

El establecimiento de la producción electrodébil del quark top solitario es un reto experimental en CDF. La pequeña señal escondida bajo la enorme cantidad de fondos con gran incertidumbre, hacen necesario una excelente comprensión tanto del detector como de los procesos involucrados para realizar dicho objetivo. Además, un simple experimento de contaje no es suficiente para extraer suficiente información de la muestra de sucesos candidatos, por lo que una técnica de análisis multivariable se vuelve crucial para distinguir entre sucesos de señal y de fondo. En esta tesis se presenta la búsqueda individual de producción combinada de quarks single top más sensible hasta la fecha, junto con el análisis de redes neuronales de CDF. El análisis presentado usa una muestra de datos de colisiones protón-antiprotón, a energía centro de masas de $\sqrt{s}$ = 1.96 TeV, que equivale a una luminosidad integrada de 3.2 fb^{-1}. Se utiliza una técnica de análisis multivariable basado en *Boosted Decision Trees*, combinando la información de varias variables para construir un potente discriminante final, alcanzando una sensibilidad a la producción de quarks single top equivalente a $5{,}2\,\sigma$. Se mide una sección eficaz de producción electrodébil de quarks top de $\sigma_{s+t} = 2{,}1^{+0{,}7}_{-0{,}6}$ pb, asumiendo una masa del quark top de 175 GeV/c^2. La dependencia de esta medida con la masa del quark top es de $+0{,}02\,\mathrm{pb}/(\mathrm{GeV}/c^2)$. La probabilidad de que el exceso observado provenga de una fluctuación de fondo es 0.0002, que corresponde a una significancia de 3.5 σ.

El resultado de este análisis es combinado con los otros búsquedas de quarks single top en CDF, alcanzando una sensibilidad que excede de $5{,}9\,\sigma$. La señal observada tiene una probabilidad de provenir de una fluctuación de fondo de $3{,}1 \times 10^{-7}$, que corresponde con $5{,}0\,\sigma$, por lo que la producción electrodébil de quarks top está concluyentemente observada. La combinación mide una sección eficaz de $2{,}3^{+0{,}6}_{-0{,}5}$ pb, que corresponde a un valor de $|V_{tb}| = 0{,}91 \pm 0{,}11$(stat.+syst.)$\pm 0{,}07$(theory) , o a un límite inferior al 95 % de nivel de confianza de $|V_{tb}| > 0{,}71$.

Con la actual luminosidad integrada adquirida por el detector CDF de más de 6 fb^{-1}, las incertidumbres sistemáticas empezarán a ser el factor limitante para futuros análisis del quark single top. Con el objetivo de mejorar significativamente la precisión de la medida de la sección eficaz de producción, y consecuentemente del elemento $|V_{tb}|$ de la matriz CKM, las incertidumbres sistemáticas, que actualmente son conservadoras, deberán ser considerablemente reducidas. Por otra parte, ahora que la producción electrodébil del quark top está firmemente establecida, es el momento de comenzar a estudiar propiedades interesantes de la producción de quarks single top. Por ejemplo, la producción de quarks single top proporciona una oportunidad única para el estudio de la polarización de los quark top, ya que el Modelo Estándar predice que éstos deben estar aproximadamente 100 % polarizados cuando son producidos mediante interacción electrodébiles, mientras que otros modelos más allá del Modelo Estándar predicen una desviación significante frente al ME.

Appendix B

VALIDATION OF INPUT VARIABLES

In Figures B.1 to B.19, the distributions comparing data and Monte Carlo predictions for all the BDT input variables are shown in the signal regions as well as in the untagged control regions. For the variables requiring a b-quark jet, the leading jet is chosen in the untagged samples.

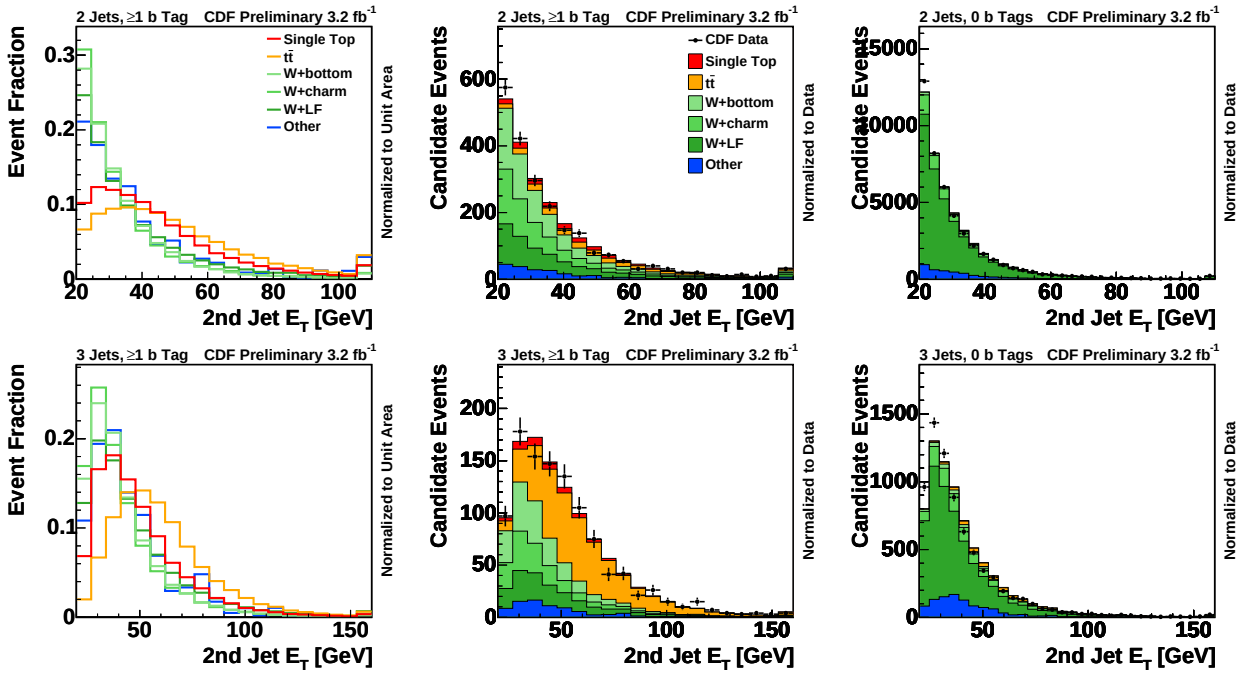


Figure B.1: Templates and validation plots comparing data and MC for the input variable $E_T^{j_2}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

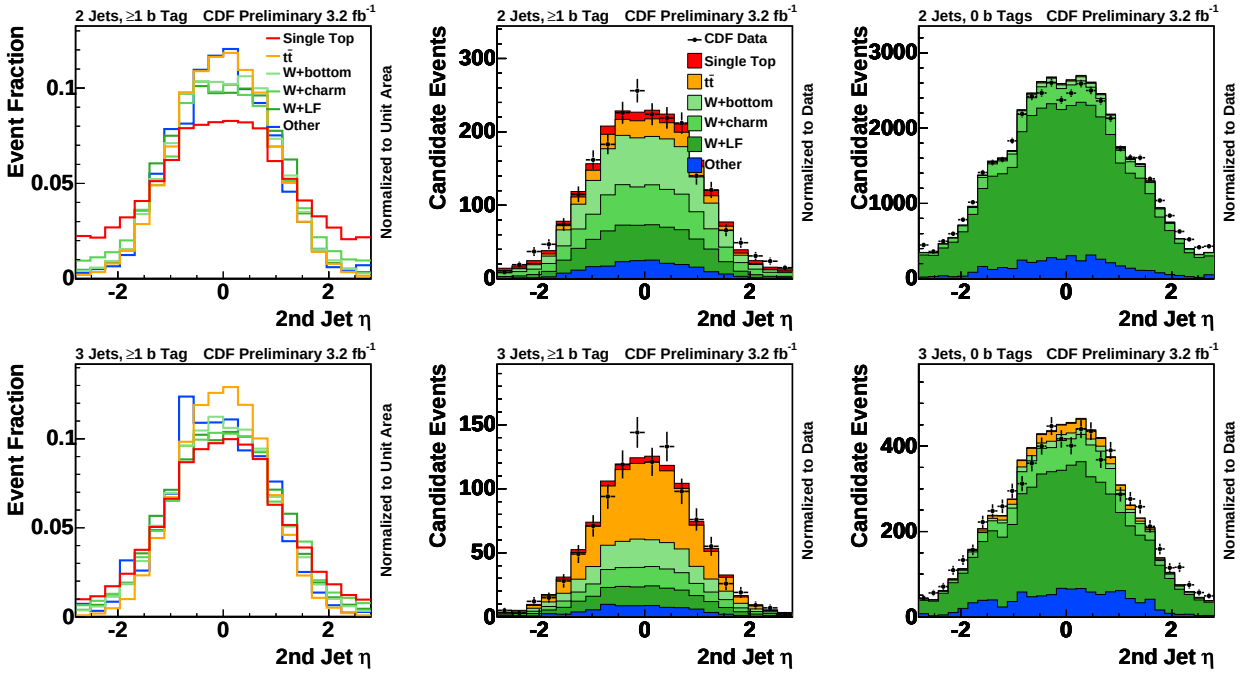


Figure B.2: Templates and validation plots comparing data and MC for the input variable η_{j_2}. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

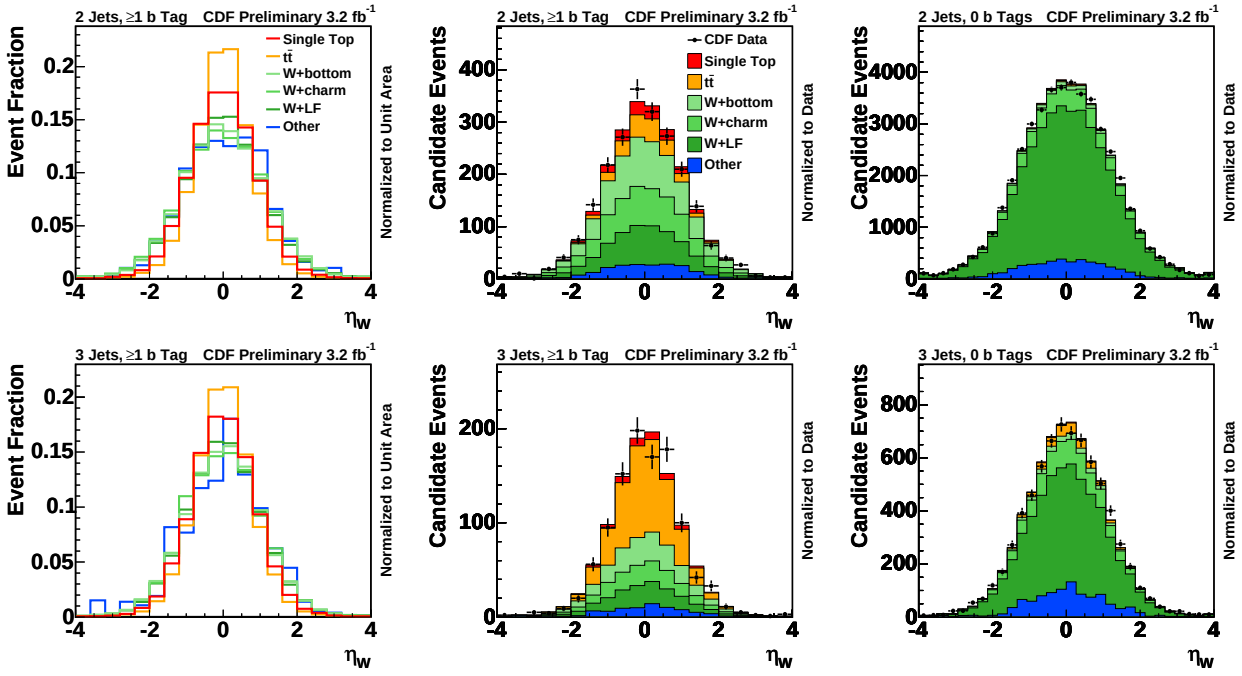


Figure B.3: Templates and validation plots comparing data and MC for the input variable η_W. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

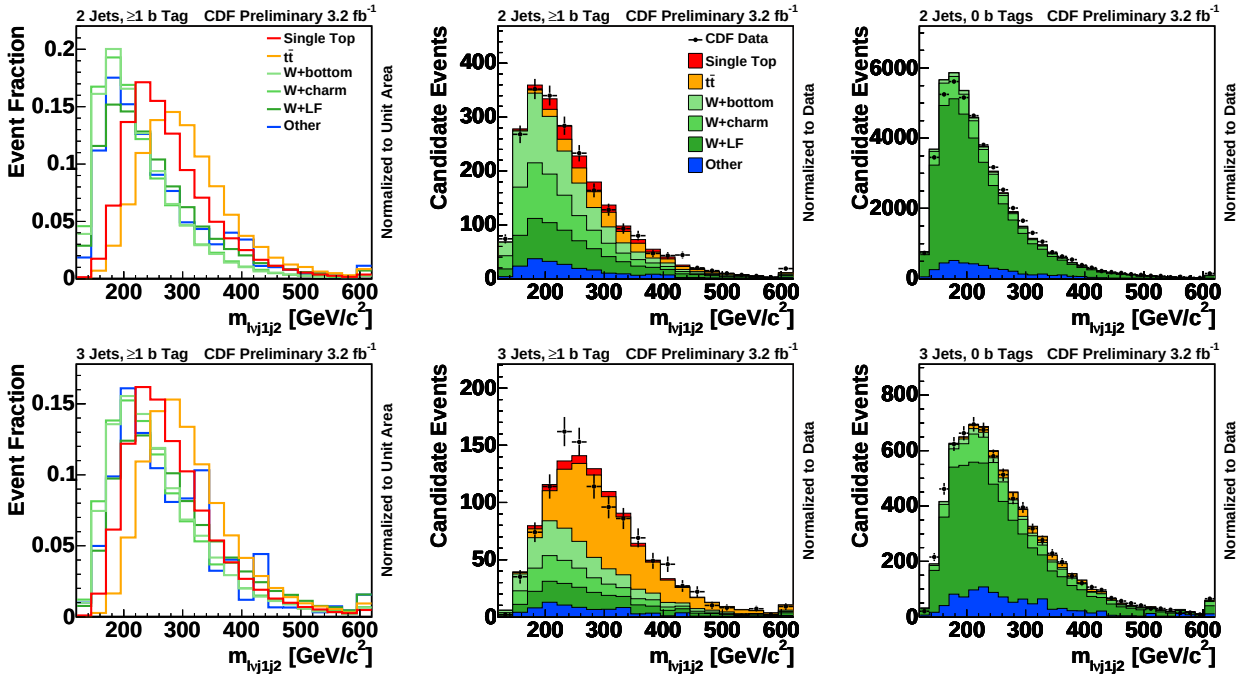


Figure B.4: Templates and validation plots comparing data and MC for the input variable $M_{\ell\nu j_1 j_2}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

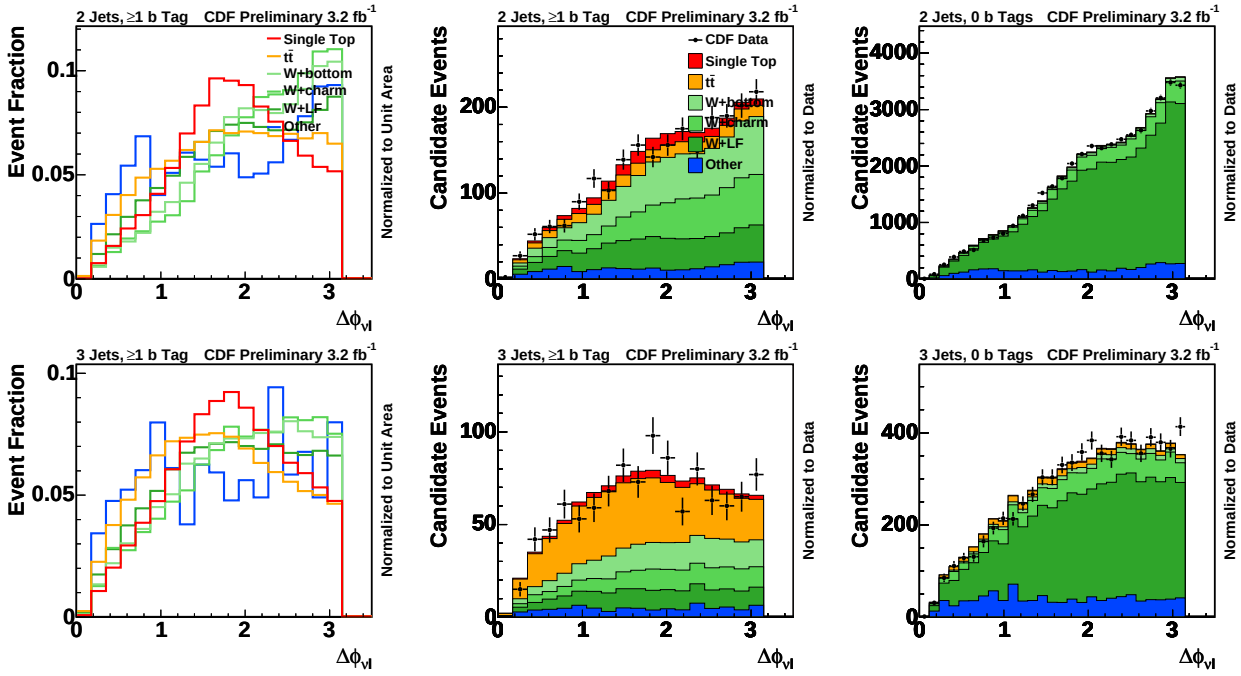


Figure B.5: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{\ell\nu}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

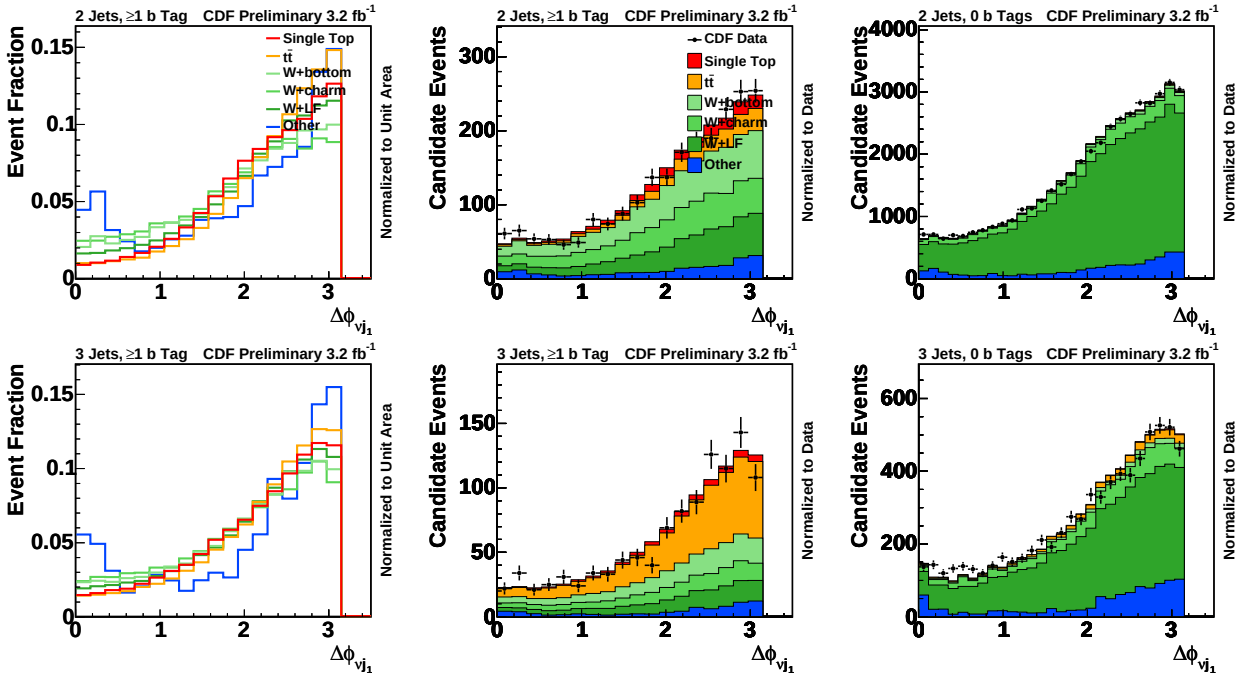


Figure B.6: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_1\nu}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

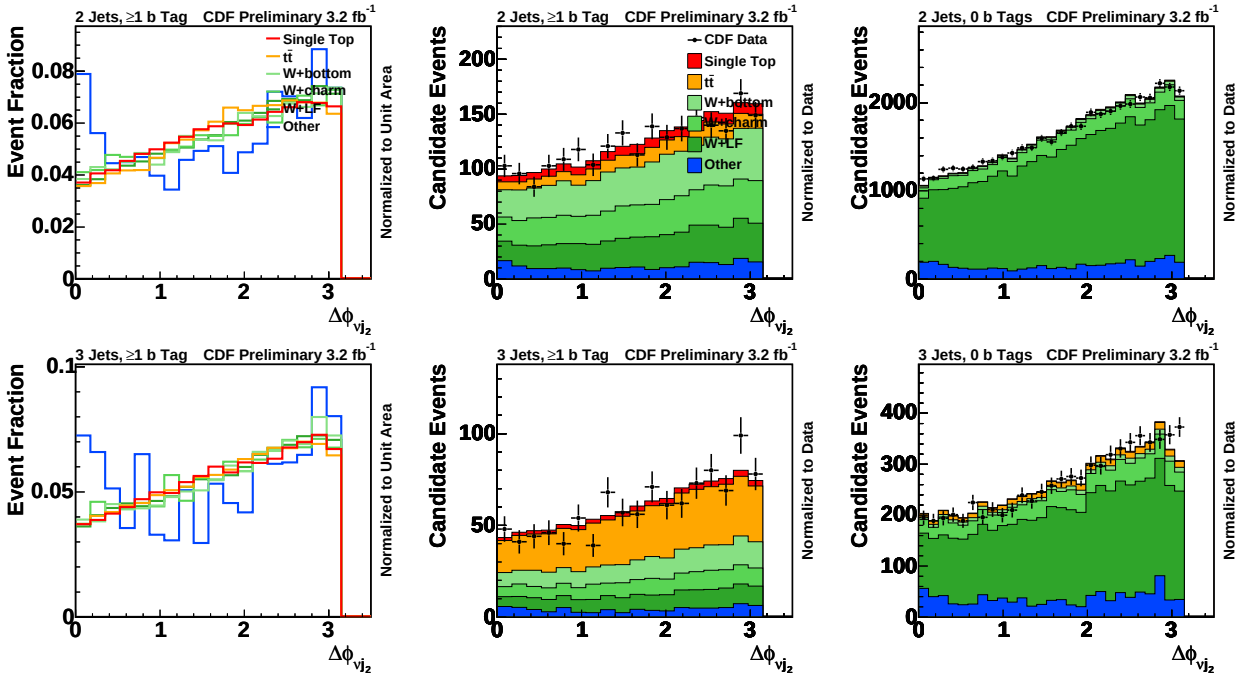


Figure B.7: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_2\nu}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

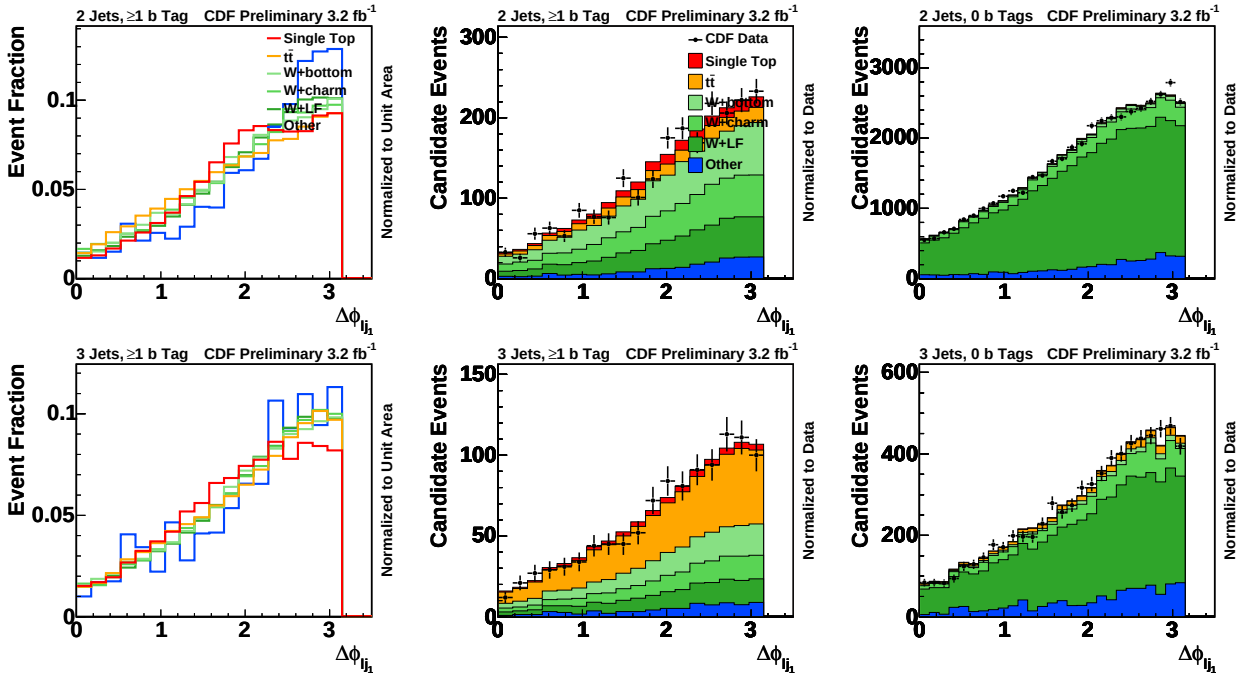


Figure B.8: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_1\ell}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

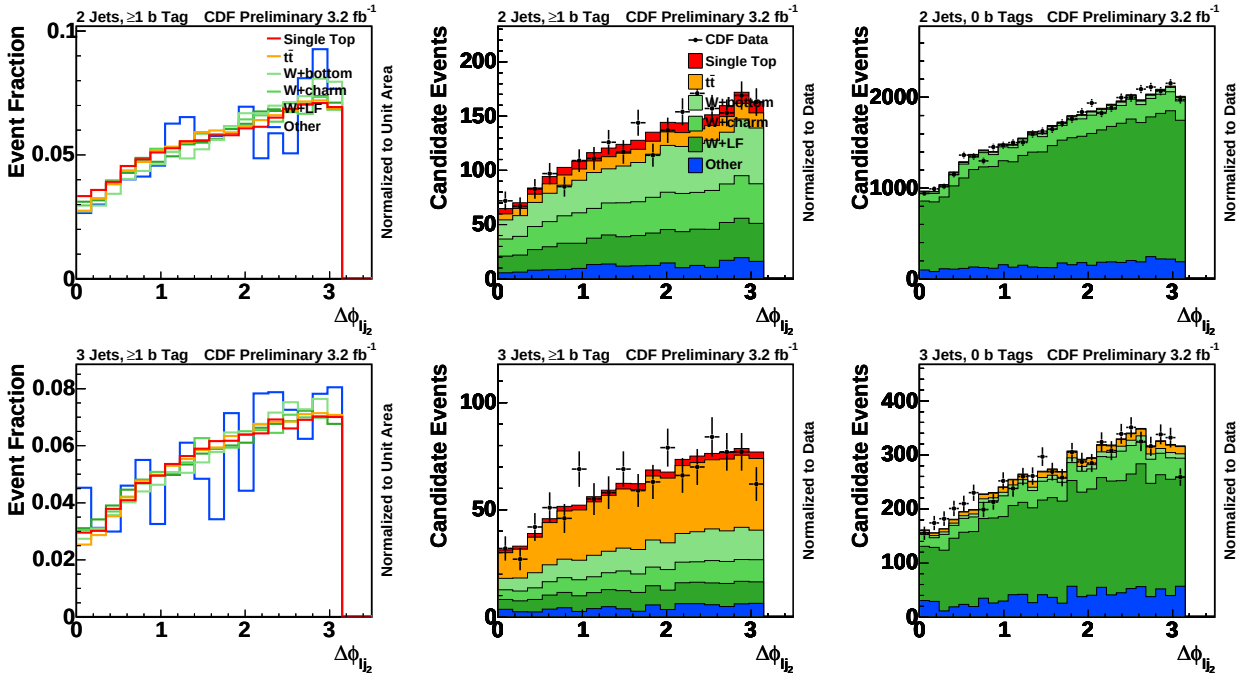


Figure B.9: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_2\ell}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

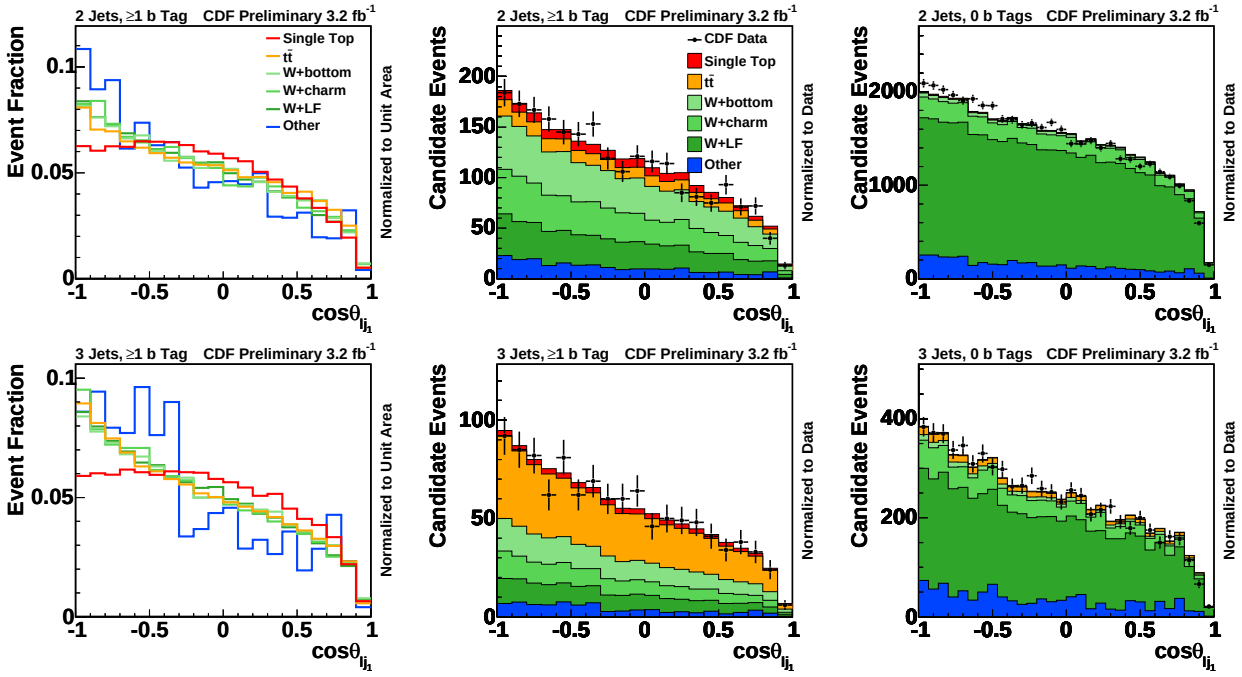


Figure B.10: Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j_1\ell}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

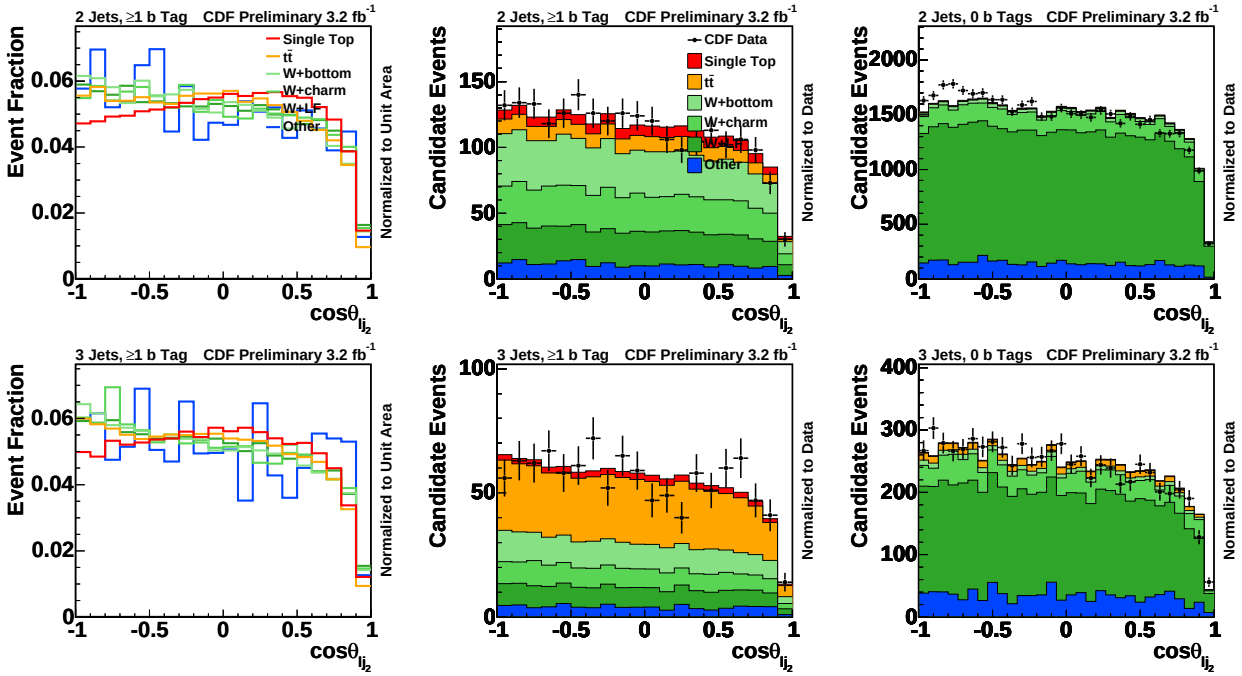


Figure B.11: Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j_2\ell}$. Plots in the top are for 2 jet events, while plots in the bottom are for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

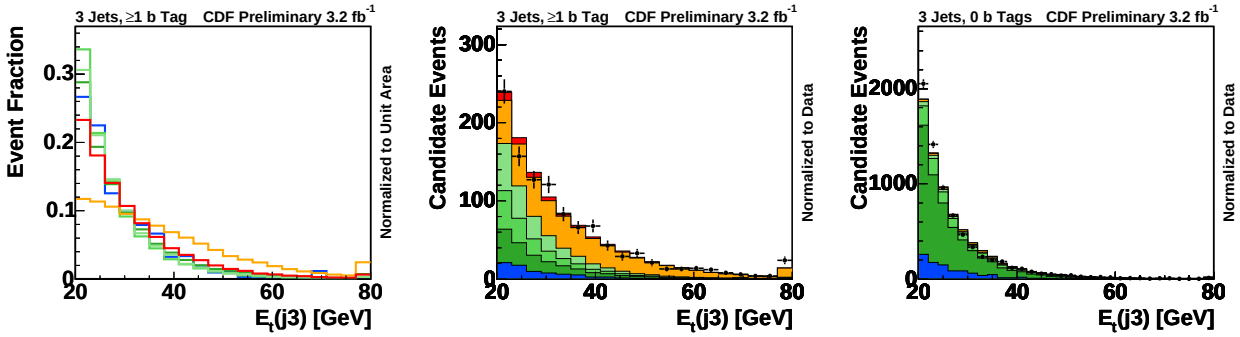


Figure B.12: Templates and validation plots comparing data and MC for the input variable $E_T^{j_3}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

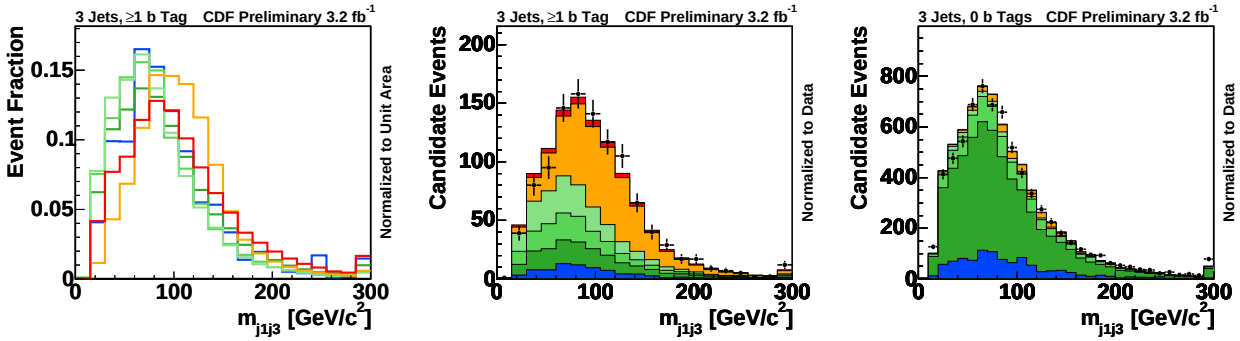


Figure B.13: Templates and validation plots comparing data and MC for the input variable $M_{j_1 j_3}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

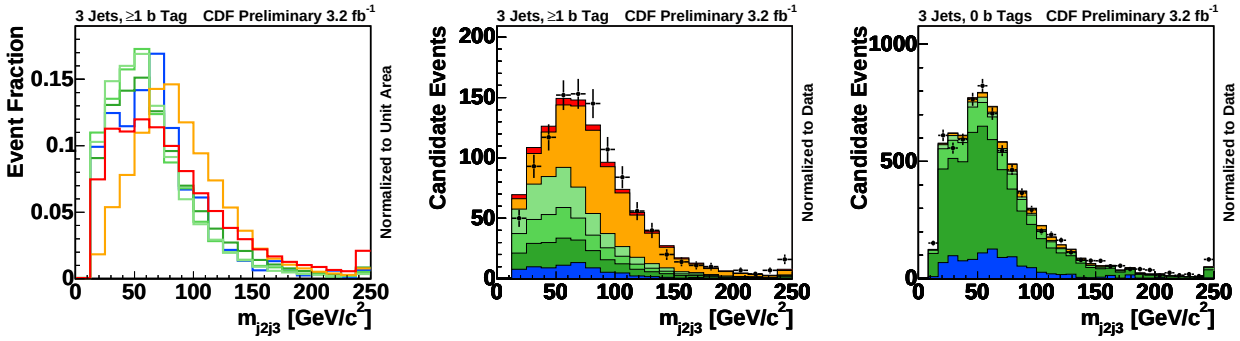


Figure B.14: Templates and validation plots comparing data and MC for the input variable $M_{j_2 j_3}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

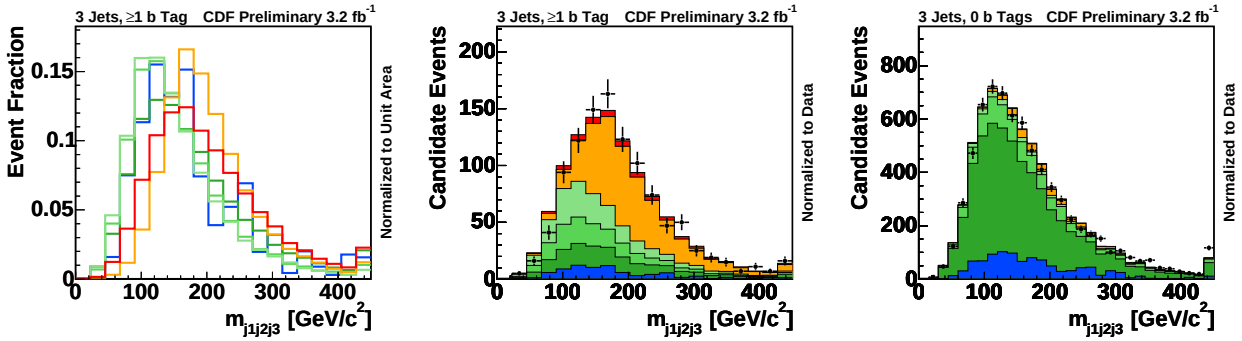


Figure B.15: Templates and validation plots comparing data and MC for the input variable $M_{j_1 j_2 j_3}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

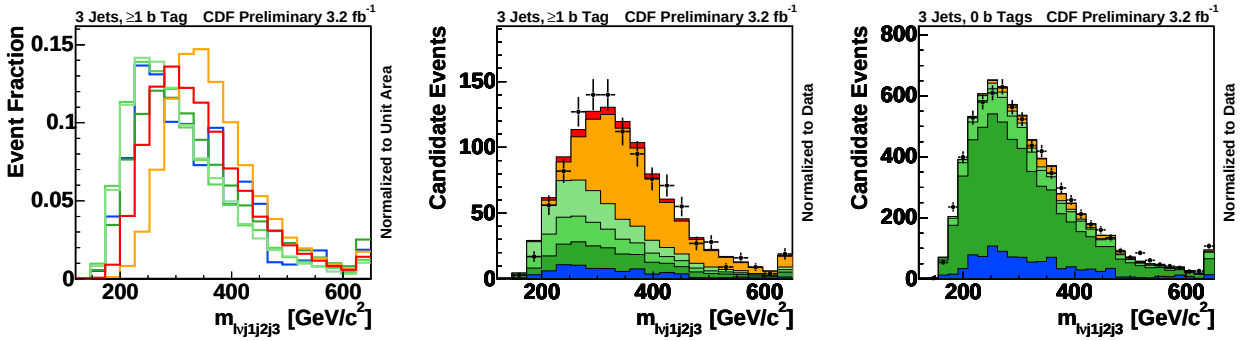


Figure B.16: Templates and validation plots comparing data and MC for the input variable $M_{\ell\nu j_1 j_2 j_3}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

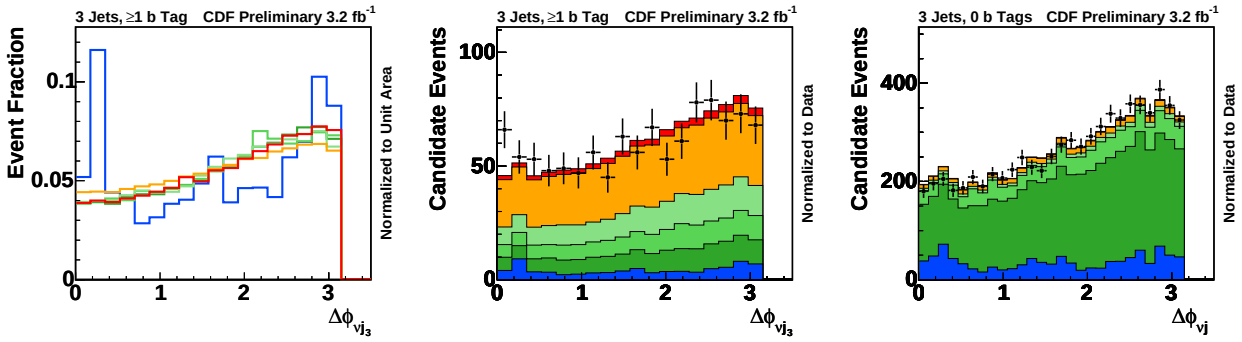


Figure B.17: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j_3\nu}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

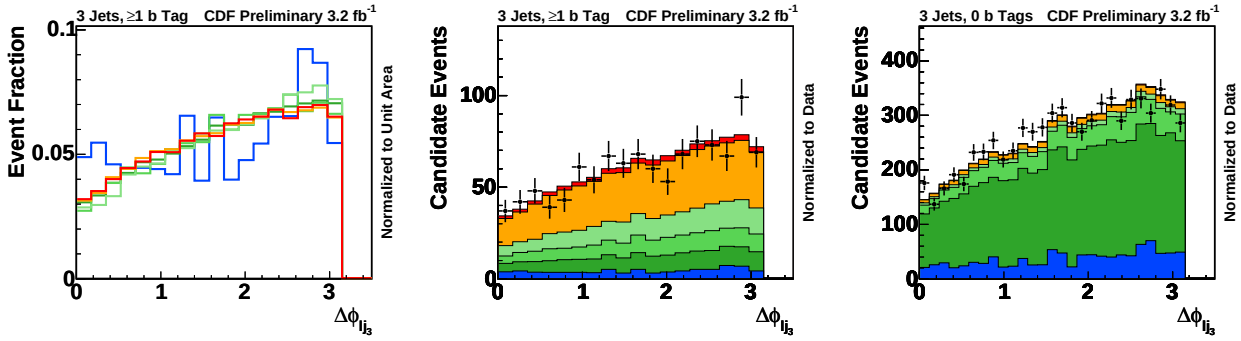


Figure B.18: Templates and validation plots comparing data and MC for the input variable $\Delta\phi_{j3\ell}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

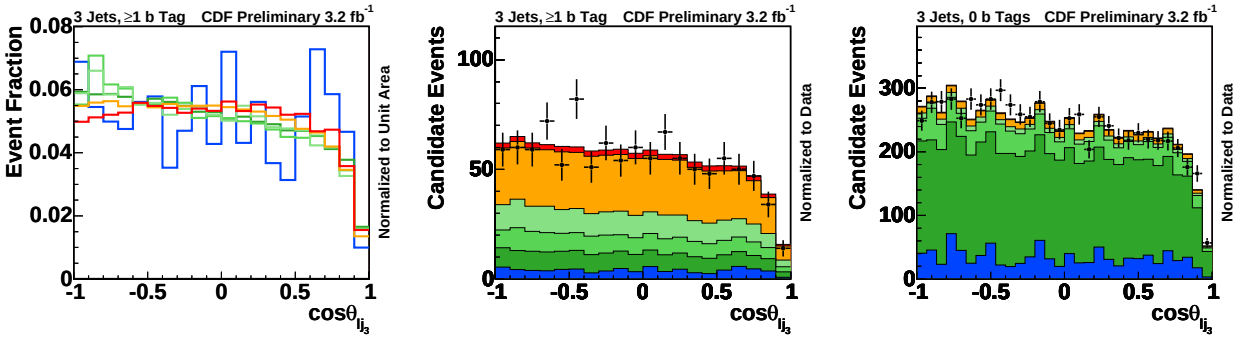


Figure B.19: Templates and validation plots comparing data and MC for the input variable $\cos\theta_{j3\ell}$ for 3 jet events. Left and middle show events with at least one b-tagged jet, and events in the untagged control sample are shown in the right.

Appendix C

SEARCH FOR THE HIGGS BOSON IN ASSOCIATION WITH A W.

The Higgs boson particle is associated to the scalar field predicted by the Higgs mechanism [42–48] in order to account for the masses of the W and Z bosons, as well as for the fundamental fermions through the spontaneous electroweak symmetry breaking (see Section 2.1.3). Although it has a strong theoretical grounding, it has not yet been observed because of its small coupling [182]. A low-mass Higgs boson (with a mass of 100–130 GeV) is most easily seen at the Tevatron in the associated WH production mode, with the W decaying semileptonically. This process has the same final state as single top, but its predicted cross section is ten to a hundred times smaller than that of the single top quark production cross section.

Because the final state of WH is the same as that of single top production, the same background estimate and analysis method as the single top analysis can be used for the search of the WH production. The analysis described in this Apendix was conducted with a data set of 2.7 fb^{-1} [183], using candidate event selection similar to the one described in Chapter 5 but slightly modified to be optimized to the Higgs signal. Because the two jets in the WH signature are expected to be central (as opposed to the t-channel signature where one of the jets is expected to be produced in the forward direction), the jet selection (see Section 5.2.1) was reduced up to $|\eta| < 2.0$. The Higgs Boson decays to a $b\bar{b}$ pair, therefore most of the sensitivity is expected to come from the double b-tagged sample. To increase sensitivity, a second b-tagging algorithm called JetProb [180] was introduced in such a way that the sample was divided in three tagging categories: the sample with two SecVtx tagged jets (SVSV), the sample with one SecVtx tagged jet and the other jet tagged by the JetProb algorithm (SVJP), and the sample with one SecVtx tagged jet with the second jet not being tagged neither by the SecVtx tagger nor by the JetProb tagger (SVnoJP). Furthermore,

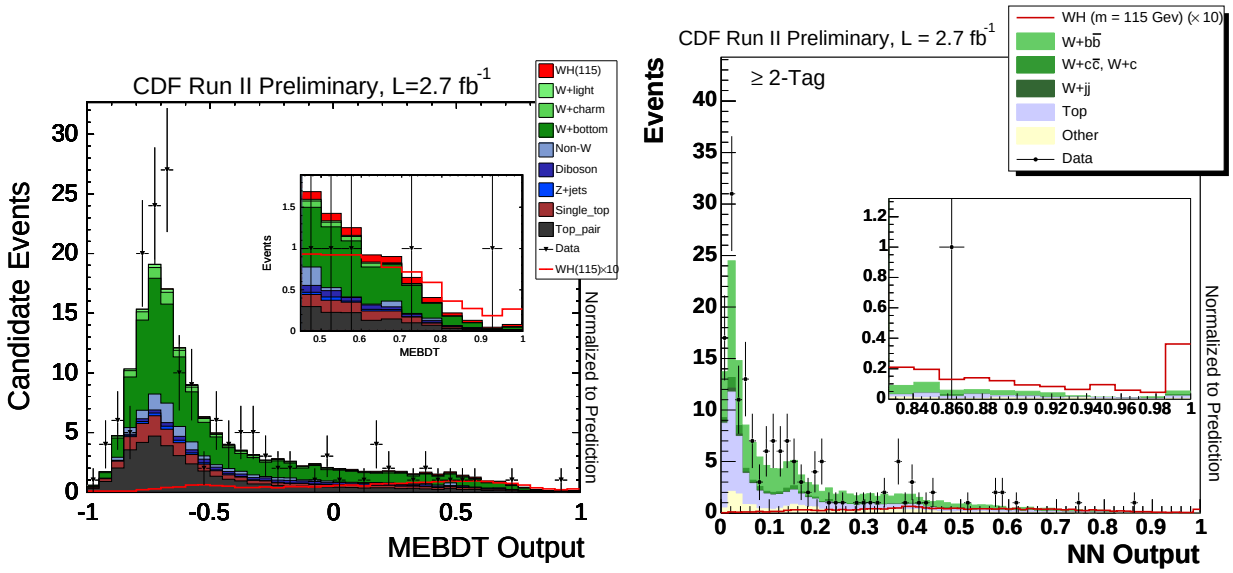


Figure C.1: MEBDT and combo discriminants for a Higgs mass of 115 GeV/c^2 in the SVSV sample.

in order to increase acceptance in the double b-tagged samples, the QCD veto (see Section 5.2.4) was not applied in the SVSV and the SVJP samples.

A Boosted Decision Tree (MEBDT) technique (see Chapter 8) was used to discriminate the WH signal from the backgrounds. The notation MEBDT underscores the use of inputs derived from the Matrix-Element approach developed in references [26, 175] (see Section 10.1.3). Matrix Element probabilities are calculated for the background hypotheses and for the WH production hypothesis for different Higgs mass points (from 100 GeV/c^2 to 150 GeV/c^2, in a step of 5 GeV/c^2). A BDT was trained for each Higgs mass using kinematic input variables (being the dijet mass the most important one) as well as the ratios of the signal event probabilities to various combinations of the background probabilities, and the event probability discriminant as described in Equation 10.2 (using the Higgs event probability as signal and the single top event probability as background). The MEBDT discriminant trained for a Higgs mass of 115 GeV/c^2 is shown in Figure C.1(a) for events with two SecVtx b-tagged jets.

An independent analysis based on Neural Networks (NN) was developed using exactly the same candidate event selection [184]. Both analyses were combined using a Super Discriminant (SD) technique similar to that described on Section 10.1.5; with this combination the sensitivity is improved by about 5–13% (depending on the Higgs mass point) over the best individual analysis. The SD discriminant is shown in Figure C.1(b) for events with two SecVtx tags and a Higgs mass of 115 GeV/c^2.

Finding no evidence for a Higgs boson signal, a Bayesian C.L. limit was calculated for each

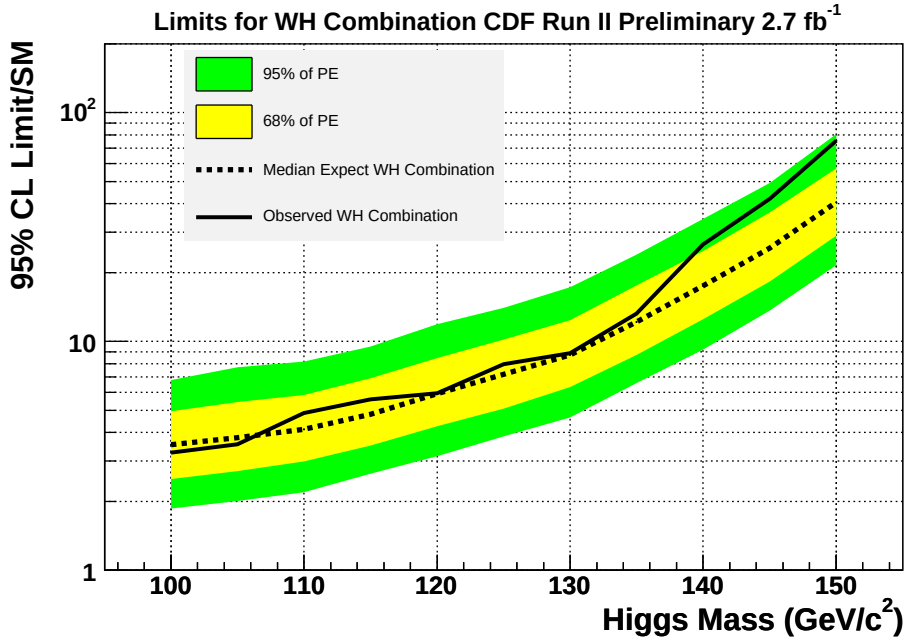


Figure C.2: The expected and observed 95% C.L. upper limits on the Higgs boson production cross section relative to the SM expectation as obtained from the SD combination as a function of the Higgs boson mass.

mass hypothesis based on the combined binned likelihood of the SD output distributions. A 95% C.L. limit was determined such that 95% of the posterior density for $\sigma \times \mathcal{B}(H \rightarrow b\bar{b})$ falls below the limit The limits are displayed graphically in Figure C.2. An observed (expected) 95% C.L. limit of 5.6 (4.8) times the SM prediction of the production cross section was observed for a Higgs boson mass of $115\mathrm{GeV}/c^2$ [185] (NLO theory predicts $\sigma \times \mathcal{B}(H \rightarrow b\bar{b}) = 136$ fb [186]).

Appendix D

MUON GAP TRIGGERS

High p_T muon triggers are essential for many interesting physics analysis at CDF, including single top quark and Higgs searches. In Section 5.1, high p_T muon triggers requiring a stub in both the CMU and the CMP muon detectors (CMUP), and a stub in the CMX detector (CMX) have been described. Additional muon categories were introduced in the analysis, including those with a stub only in the CMU detector or in the CMP detector, through a $\not\!\!E_T$+jets trigger. However, many interesting acceptance is lost due to the trigger requirements on the missing transverse energy and on the jets (the efficiency of these requirements has been measured to be 40-50% in single top Monte Carlo). Therefore, inclusive CMU-only and CMP-only muon triggers, which compliment the standard high p_T muon triggers, would be preferable to recover these events. Test triggers existed before the Tevatron 2007 shutdown, for both CMU-only and CMP-only triggers, but rates were too high at high luminosity and these trigger paths have not been useful for physics analysis so far. This appendix describes the work done to improve these triggers so that the trigger rates were manageable enough to allow them to be included during the Tevatron 2007 shutdown.

Phi–Gap Trigger

Gaps exist between each calorimeter wedge, and since the CMU chambers are mounted in the edge of these wedges these gaps also affect the CMU coverage. There is a 2.25 degree gap every 15 degrees in ϕ in the CMU coverage. The η–ϕ positions of CMP-only muons are shown in Figure D.1. The central crack as well as the crack between each calorimeter wedge are clearly visible in the figure.

Although, there is clearly a contribution to CMP-only muons from the central crack, it is less than one third of all CMP-only muons. Track ϕ information is available at Level-1 and Level-2,

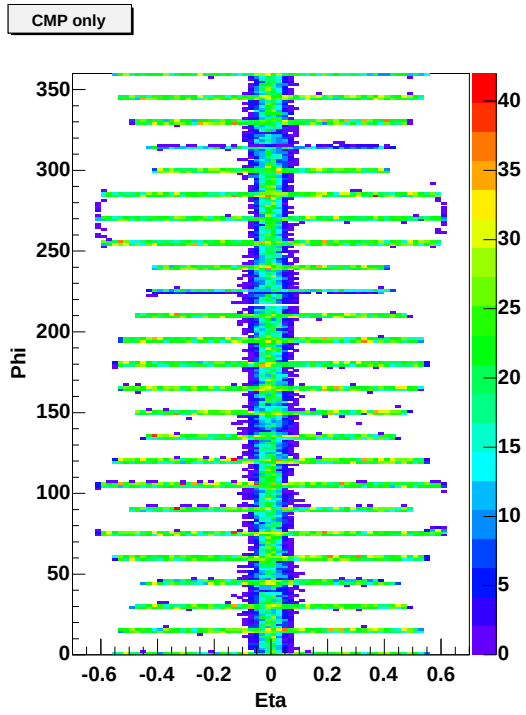


Figure D.1: The η–ϕ scatter of CMP-only muons. The CMU coverage has a large gap where the East and West Calorimeter modules come together, and 2.5 degree gaps every 15 degrees in ϕ which correspond the gaps between the calorimeter wedges.

which is useful to reduce the trigger rate by requiring that a track point at the gaps in ϕ. On the other hand, no track z information was available at Level-1 and Level-2 at the time of this work. For this reason the CMP-only trigger only reclaims the muons in the ϕ gaps and does not include the central crack.

A Phi–Gap trigger existed which required tracks to point towards the gaps in coincidence with a CMP stub. Rates are very high in this trigger due to fakes. In fact, at the time this work began, the Phi–Gap trigger was turned off for instantaneous luminosities above $120 \times 10^{30}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. Two ways of reducing the rates were considered: tightening the XFT Phi–Gap extrapolation window, and requiring a hit in the CSP scintillating system that matches the CMP stub.

Muon triggers work by matching XFT tracks to the muon chambers which report a stub in detector ϕ. This match is based on an extrapolation window in ϕ [187]:

$$\sigma = \sqrt{\left(\frac{3\sigma_k}{p_T}\right)^2 + \sigma_a^2}, \tag{D.1}$$

where $3\sigma_k$ is the 3 sigma multiple scattering term, and σ_a is the misalignment parameter. The σ_a parameter dominates at high p_T so it is where improvement can be made for high p_T leptons. The default value of the σ_a parameter was set to 1.5 degrees earlier at CDF when high luminosity and trigger rates were not a problem. A sample of Z bosons with tight offline cuts which require one

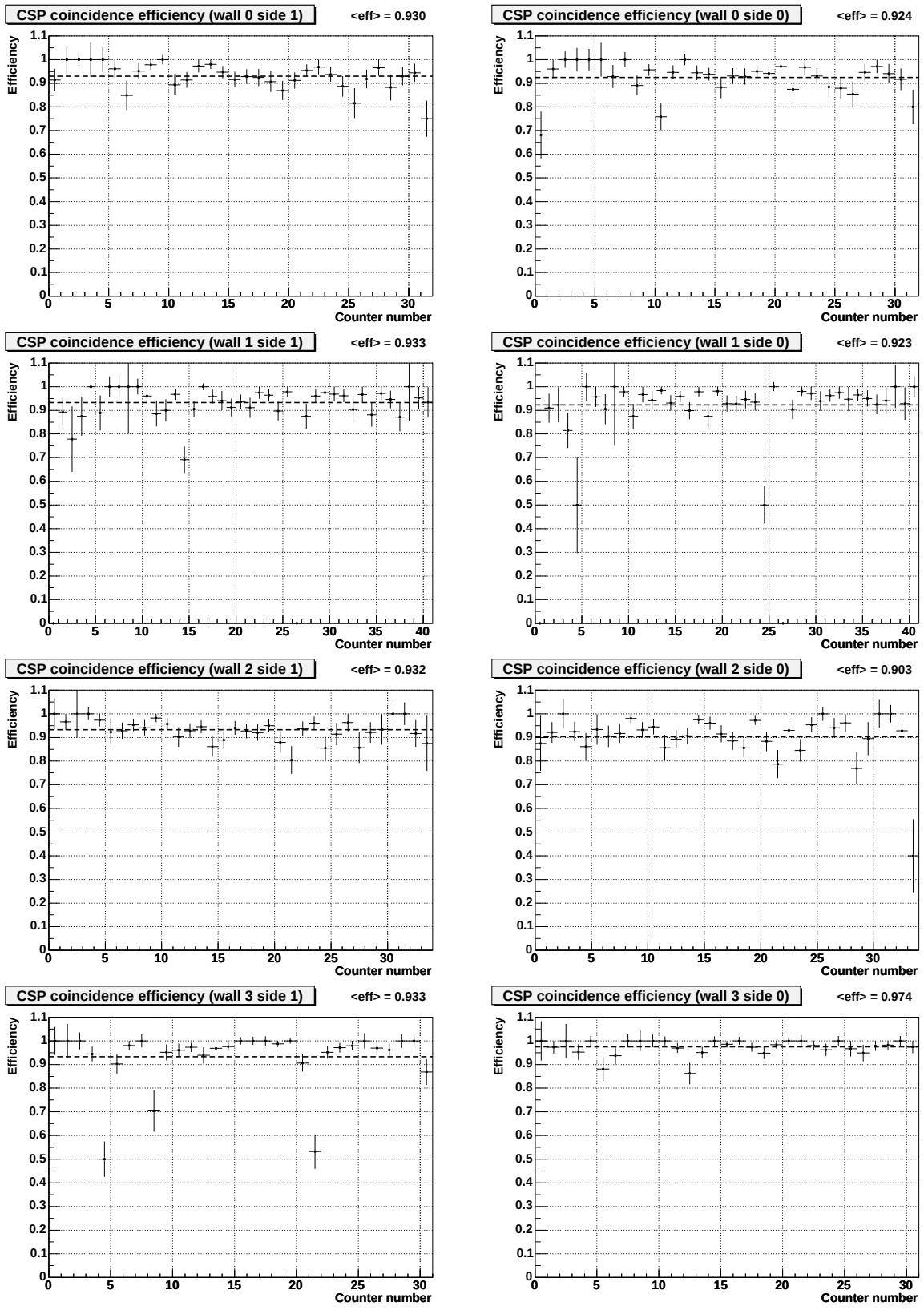


Figure D.2: CSP efficiencies. The figures correspond to north, top, south and bottom walls from top to bottom; and for east (left) and west (right) sides.

and only one good CMP muon was used for studying the efficiency of the CMP-only trigger as a function of σ_a. Based on this study, a new value of 0.8 was taken for σ_a in the Phi–Gap trigger keeping a trigger efficiency above 99%. This results in a rate reduction on the order of 30% at high luminosity.

The rate reduction obtained from the tighter matching requirement is not enough to make the Phi–Gap trigger safe to run at high luminosities. The CDF detector is equipped with a scintillating counter detector of the same geometry as the CMP which has never been used. The CSP has a smaller time window (< 100 ns or less than 1 bunch crossing) than the CMP (< 1700 ns or about 4 bunch crossings) and therefore may be useful to help reduce the rate of the CMP-only triggers. The CSP also provides east-west discrimination where as the CMP does not. This will be useful when stereo pointing is available at Level-2.

Before the CSP could seriously be considered as a valuable trigger requirement the timing gates and counter efficiencies had to be checked. It was found that the CSP timing window was unproperly set before the time of this work. Once the CSP timing gates were corrected, it was checked that the CSP trigger bit always fires when there is a CSP from a good muon.

Once the timing for the CSP is set its efficiency can be checked. There are 4 walls, in the north, the south, the top and the bottom sides of the detector. The efficiencies for each wall are shown in Figure D.2. The combined efficiency for all walls is about 93%.

With efficiency understood and CSP timing properly set the CSP is ready for use in the CDF trigger system. In order to use it, a matching algorithm had to be designed and incorporated in the Level-2 trigger code. There is no simple one-to-one correspondence between the CMP stacks and the CSP scintillation counters. This is especially clear in the corners of the CMP (for instance, where the top wall meets the north wall). Therefore, the matching was derived from data. For good muons, it was recorded how often each CSP scintillator was hit in coincidence with each CMP stack. The top four or five stack for each scintillator were included in the matching table. In the Level-2 code when a good CMP-XFT map is found, the hit map for the CSP is generated and converted to the corresponding CMP stack map. If the CMP stack is on the list of stacks corresponding to a CSP scintillator with a hit, then the event fires the L2 Phi-Gap trigger. The algorithm was found to be 93% efficient for good CMUP muons. This is consistent with the inefficiency of the CSP counters and means the matching algorithm functions properly.

The CSP-CMP matching algorithm was included in a Phi-Gap and CMUP18 test trigger starting in store 5562 (7/19/07). Therefore, data was taken with the test trigger before the 2007 shutdown. The rate reduction in the CMP triggers of about a factor of 6 is observed at high instantaneous luminosity, as shown in Figure D.3.

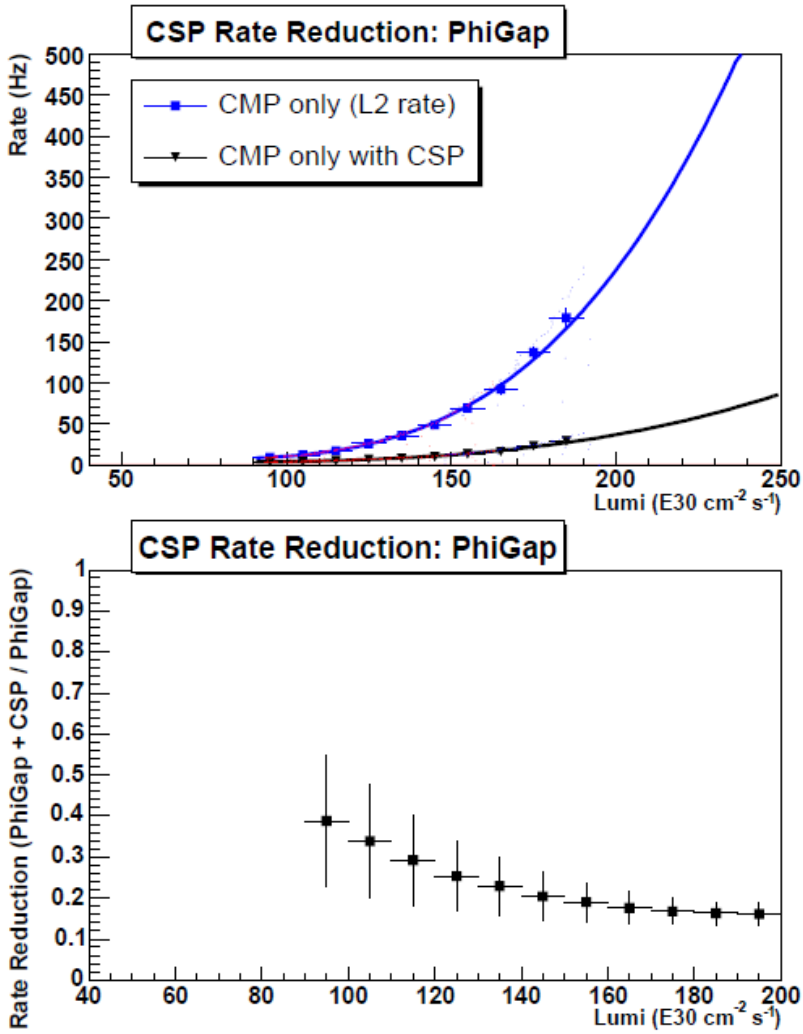


Figure D.3: Observed rate reduction for the Phi-Gap test trigger due to the CSP matching requirement (top) and the ratio with to without the matching (bottom). A rate reduction of about a factor of 6 is observed at high instantaneous luminosity.

Eta–Gap Trigger

Because of the rectangular geometry of the CMP detectors the overlap with the CMU coverage is not complete. The η–ϕ positions of CMU-only muons are shown in Figure D.4. Most of the CMU-only muons occur near the forward corners (large $|z|$) of the CMP where the radial distance from the beam axis is maximal.

A Eta–Gap trigger exists which requires a track to match a CMU stub only. Rates are high in this trigger due to fakes. Since the CMU muon chambers are mounted very close to the edge of the calorimeter wedges it is clear that punch through could be a problem. In fact, at the time this work began, the Eta–Gap was turned off for instantaneous luminosities above $150 \times 10^{30}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$.

One possible way to reduce the CMU rate is to require hadron TDC timing confirmation in the

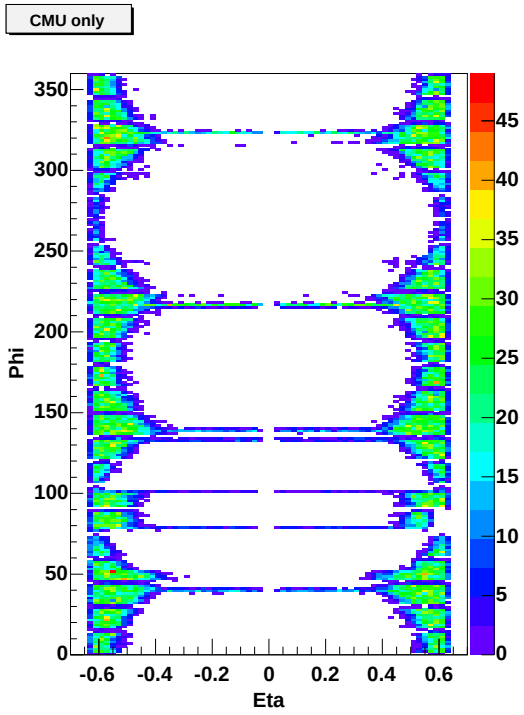


Figure D.4: The η–ϕ scatter of CMU-only muons. Due to the rectangular geometry of the CMP detector, coverage is not complete for the more forward regions with respect to the CMU.

tower where the CMU stub is reported. In preparation for adding this functionality the efficiencies of hadron TDC confirmation for good muons from Z bosons was studied. The results of Figure D.5 show that the Hadron TDC confirmation is approximately 100 good muons. This functionality would need to be added to the Level-2 trigger code to use the Hadron TDCs in the on-line trigger. It is not clear how much this will reduce the rate, and for now, the Eta-Gap triggers can be run with minimal prescale without any improvements.

The Phi–Gap trigger trigger has been running properly since the 2007 Tevatron shutdown, reaching an integrated luminosity that now exceeds 2fb^{-1}. This sample has been recently included in the search for the Higgs boson in the $H \rightarrow WW$ channel at CDF [188]. Recent studies have been performed in order to include this trigger in the WH channel as well, resulting in an increment in the CMU sample size of the order of 10 times the size of the CMU sample from the $\not\!\!E_T$+jets trigger.

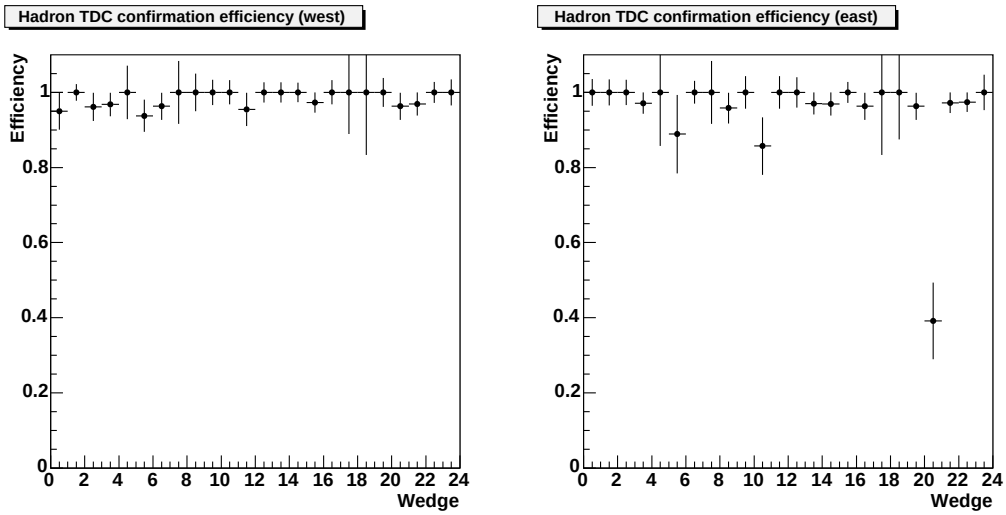


Figure D.5: Efficiencies of the Hadron TDC confirmation for good muons from Z boson decay. The efficiencies are close to 100%.

Appendix E

MEASUREMENT OF $|V_{td}|$ AND $|V_{ts}|$ AT CDF

The importance of overconstraining the CKM matrix elements in high energy physics has been introduced in Section 2.1.3. In this thesis, a measurement of the least experimentally know matrix element $|V_{tb}|$ has been presented. Other poorly constrained elements of the CKM matrix, that can be study at the Tevatron with an unprecedented level of precision, are the matrix elements $|V_{td}|$ and $|V_{ts}|$. This elements can be extracted from measurements of the neutral B mesons ($b\bar{q}$, with $q = d, s$ for $\overline{B}^0_d$ and $\overline{B}^0_s$) oscillations from particle to antiparticle due to flavor-changing weak interactions. The probability density P_+ (P_-) for a $\overline{B}^0_q$ meson produced at proper time $t = 0$ to decay as a $\overline{B}^0_q$ (B^0_q) at a time t is given by

$$P_{\pm}(t) = \frac{1}{2\tau_q} e^{-t/\tau_q} \left[1 + \cos(\Delta m_q t)\right], \tag{E.1}$$

where Δm_q is the mass difference between the heavy $B^0_{q,H}$ (CP odd) and light $B^0_{q,L}$ (CP even) weak eigenstates of the B^0_q mesons[1], and τ_q is the lifetime, which is assumed to be equal for the two mass eigenstates. The mass differences Δm_s and Δm_d are proportional to $|V_{ts}|^2$ and $|V_{td}|^2$, respectively, and can be used to extract these fundamental parameters. However, theoretical uncertainties limit the precision of $|V_{ts}|$ and $|V_{td}|$. A more reliable constraint is provided by the ratio $|V_{td}/V_{ts}|$ where several theoretical uncertainties are reduced. This ratio is given by

$$\left|\frac{V_{td}}{V_{ts}}\right| = \xi \sqrt{\frac{m_{B_s}}{m_{B_d}} \frac{\Delta m_d}{\Delta m_s}}, \tag{E.2}$$

where $m_{B_s}/m_{B_d} = 0.98390$ [189] is the precisely measured ratio of masses of the two neutral B mesons, and $\xi = 1.21^{+0.047}_{-0.035}$ [190] is determined from lattice QCD calculations. The mass differ-

[1] $\hbar$ and c is set to the unity and $\Delta m_q = m_{B^0_{q,H}} - m_{B^0_{q,L}}$ is reported in inverse picoseconds.

ence Δm_d is very well measured, $\Delta m_d = 0.507 \pm 0.005$ ps^{-1} [11]. Recently, CDF II presented the first precise measurement of Δm_s: $\Delta m_s = 17.31\ ^{+0.33}_{-0.18}\ (\mathrm{stat.})\ \pm 0.07\ (\mathrm{syst.})\,\mathrm{ps}^{-1}$ [191], based on the analysis of 1 fb^{-1}of data collected with the CDF II detector at the Fermilab Tevatron. The probability that random fluctuations would produce a comparable signal was 0.2% (3σ), which was too large to claim an observation. In a second Letter [192] CDF reported an update of [191], that uses the same data set and improved analysis techniques to reduce this probability to 8×10^{-8} ($> 5\sigma$), yielding the observation of time-dependent B_s oscillations. A more precise measurement of Δm_s was reported: $\Delta m_s = 17.77\ \pm 0.10\ (\mathrm{stat.})\ \pm 0.07\ (\mathrm{syst.})\,\mathrm{ps}^{-1}$. An overview of the improved analysis is given in this Appendix.

The production of B hadrons at the Tevatron is dominated by processes that produce $b\bar{b}$ pairs. The b quark and $\bar{b}$ antiquark are energetic enough that they are expected to fragment into B hadrons independently of one another. All B species (B^-, $\bar{B}^0$, $\bar{B}^0_s$, B^-_c, b-baryons) are produced, with $\sim 10\%$ of b quarks fragmenting into $\overline{B}^0_s$ [11]. The trigger used in this analysis to select the B_s sample is based on two displaced tracks with transverse momentum $p_T > 2$ GeV, $\sum p_T > 5.5$ GeV, and large impact parameter $120 < d_0 < 1000$ μm.

To measure time-dependent oscillations, three specific ingredients are required:

1. **Flavor at the time of production:** knowledge of whether the meson was produced as a B^0_s or a $\overline{B}^0_s$. We refer to this as "initial state flavor tagging" or simply "flavor tagging."
2. **Flavor at the time of decay:** knowledge of whether the meson was a B^0_s or $\overline{B}^0_s$ when it decayed. If the flavor of decay is different than (the same as) the flavor at production, the B_s is classified as "mixed" ("unmixed").
3. **Proper decay time:** The proper decay time is the decay time of the hadron in its rest frame. Since a B_s oscillates on average four times during its decay time, the time dependent observation of B_s oscillations requires excellent proper time resolution.

The first two items listed above refer to the flavor of the B_s at the time of production and decay. We begin by reconstructing B_s decays in fully reconstructed hadronic modes ($\overline{B}^0_s \to D^+_s\pi^-$ and $\overline{B}^0_s \to D^+_s\pi^-\pi^+\pi^-$), partially reconstructed hadronic modes ($\overline{B}^0_s \to D^{*+}_s\pi^-,\ D^{*+}_s \to D^+_s\gamma/\pi^0$ and $\overline{B}^0_s \to D^+_s\rho^-,\ \rho^- \to \pi^-\pi^0$) and semileptonic modes ($\overline{B}^0_s \to D^{+(*)}_s l^-\bar{\nu}_l$) using charged particles only [2]. In this analysis, the flavor at the time of decay is determined using charge correlations among the final state particles, for example $\overline{B}^0_s \to D^+_s\pi^-$ versus $B^0_s \to D^-_s\pi^+$. The flavor at the time of production is more difficult to ascertain, and several techniques have been developed to perform this function. The effectiveness of a flavor tag is quantified by the term $\varepsilon\mathcal{D}^2$, where the efficiency ε is the fraction of decay candidates that are tagged, and the dilution $\mathcal{D} = 1 - 2p_W$, where p_W is the probability that the flavor tag is incorrect.

[2] References to a particular process imply that the charge conjugate process is included as well

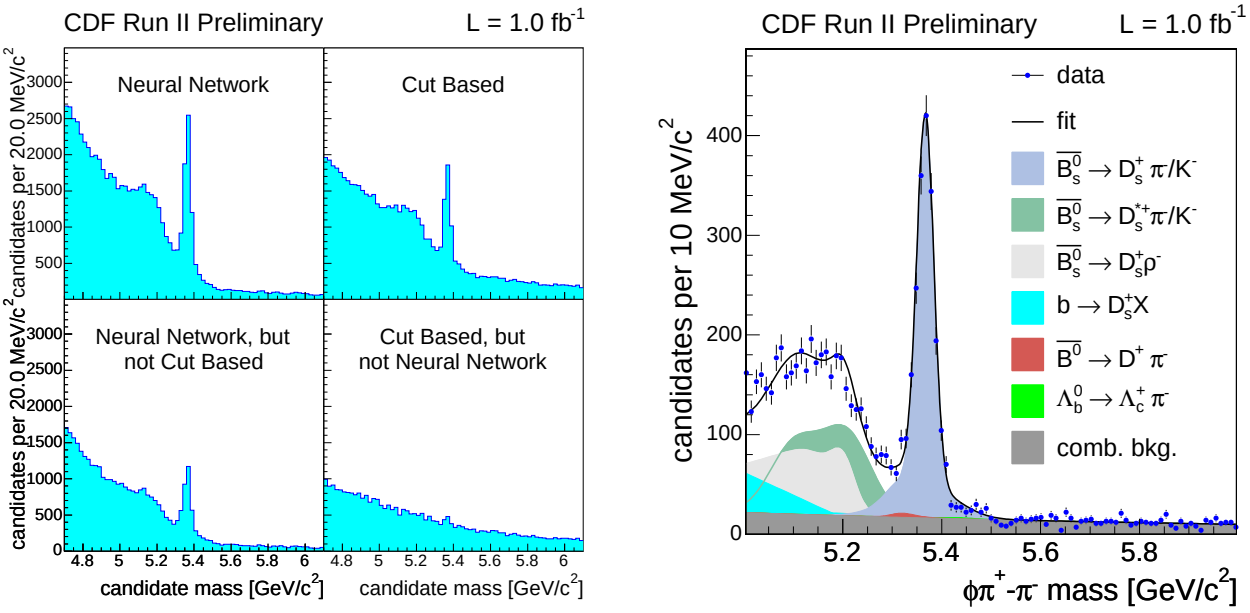
To perform this measurement, we must map out the oscillation probability as a function of decay time for at least a portion of the decay time spectrum. As each of the three items listed above has experimental limitations, this analysis requires large samples of B_s decays with a good signal-to-noise. To a good approximation, the significance of the analysis is given as

$$\mathcal{S} \sim \sqrt{\frac{\varepsilon \mathcal{D}^2}{2}} \times \exp\left(-\frac{\Delta m_s^2 \sigma_{ct}^2}{2}\right) \times \frac{S}{\sqrt{S+B}}, \tag{E.3}$$

with σ_{ct} the average resolution of the proper decay time measurement, and S and B the number of reconstructed signal and background events, respectively. This expression stress again the importance of the tagging performance, time resolution, and signal selection.

The B_s^0 meson candidates are reconstructed in reverse order of their decay chain, such that the event tracks are fit for B daughter resonances which are progressively combined in the expected topology of the B_s^0 signal. This is true for both hadronic and semileptonic modes. As an illustrative example, we briefly delineate the reconstruction of the main hadronic decay mode $\overline{B}_s^0 \to D_s^+\pi^-$, $D_s^- \to \phi\pi^+$, $\phi \to K^+K^-$. For a given event, every pair of oppositely-charged tracks is hypothesized to be a kaon pair and is fit in three dimensions with requirements on the kinematics and quality of the vertex. The resulting ϕ candidate vertices are then combined with an additional track, applying another set of selection criteria to form a higher-level collection of candidates. These candidates are assumed to be weakly decaying D_s^- mesons, and the vectors defined by their reconstructed momentum are combined with the remaining available tracks to form B_s^0 candidate vertices. In each step a new set of selection criteria, including kinematics and decay vertex fit quality, is applied.

For the hadronic decay event selection, six topologies of $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^-\pi^-\pi^+\pi^+$ decays are reconstructed using a progressions similar to the example described above. Each of the $B_s^0 \to D_s^-\pi^+(\pi^-\pi^+)$ channels has the charm meson reconstructed in one of the following final states: $D_s^- \to \phi\pi^-$ ($\phi \to K^+K^-$), $D_s^- \to K^*(892)^0K^-$ ($K^*(892)^0 \to K^+\pi^-$), or $D_s^- \to \pi^+\pi^-\pi^-$. First, candidates are pre-selected with loose selection requirements to a few powerful variables as are the quality of the vertices, the impact parameter of the decay, or the transverse momentum of the B_s^0 candidate. Further clean-up is applied to reduce contamination from B_d^0 decays. In the evidence paper, background from these decays were removed by a stringent cut on D_s candidates consistent with the D^- invariant mass. The use of particle identification, based on the time-of-flight from the TOF detector and on the $\mathrm{d}E/\mathrm{d}x$from the COT, allowed to relax this requirement in the observation paper leading to a substantial increase in signal efficiency. The final selection is done with a Neural Network trained separately for each decay mode using the ROOTSNNS package [193, 194]. The training samples are simulated events for the signal, and events from a upper mass sideband (region of the mass spectrum away from the fully reconstructed signal and dominated by combinatorial background) for background. Over 30 variables are used



(a) (b)

Figure E.1: Comparison between the NN-based selection and the cut-based selection (a), and mass spectrum for the golden hadronic decay mode (b).

in the training of the NN, including kinematics variables, vertex information, angular correlations and particle identification information. Special care is taken in order not to include variables with mass dependence. The selection is optimized with a cut on the NN output that maximizes the quantity $S/\sqrt{S+B}$. The NN-based selections corresponds to a big improvement with respect to the rectangular cut-based selection used in the evidence paper, as shown in Figure E.1(a). This figure show that events rejected by the cut-based selection but passing the NN-based selection are mostly events in the signal peak, while events accepted by the cut-based selection and rejected by the NN corresponds to background events, leading to a much larger and purer NN sample in comparison to the cut-based sample.

The cleanest mass spectrum is found for the decay channel $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$, and is shown over a wide mass range in Figure E.1(b). The main peak is quite narrow, with an excellent signal-to-background ratio. In addition to the fully reconstructed signal, the data sample contains incompletely reconstructed hadronic decays. For these modes, the reconstructed topology excludes one or more particles from the full B_s^0 decay. This signal is "nearly" fully reconstructed and is therefore a potential source of significant statistical power in the measurement. This statistical power is reduced by the increased level of background in the lower mass range, but the expected large size of signal content in these partially reconstructed modes makes them very valuable. The primary partially reconstructed modes in the $B_s^0 \to D_s^- \pi^+$ topology are $B_s^0 \to D_s^- \rho^+$ and $B_s^0 \to D_s^{*-} \pi^+$. In the first case, $\rho^+ \to \pi^+ \pi^0$, while in the second case $D_s^{*-} \to D_s^- \gamma$ or $D_s^- \pi^0$. In either

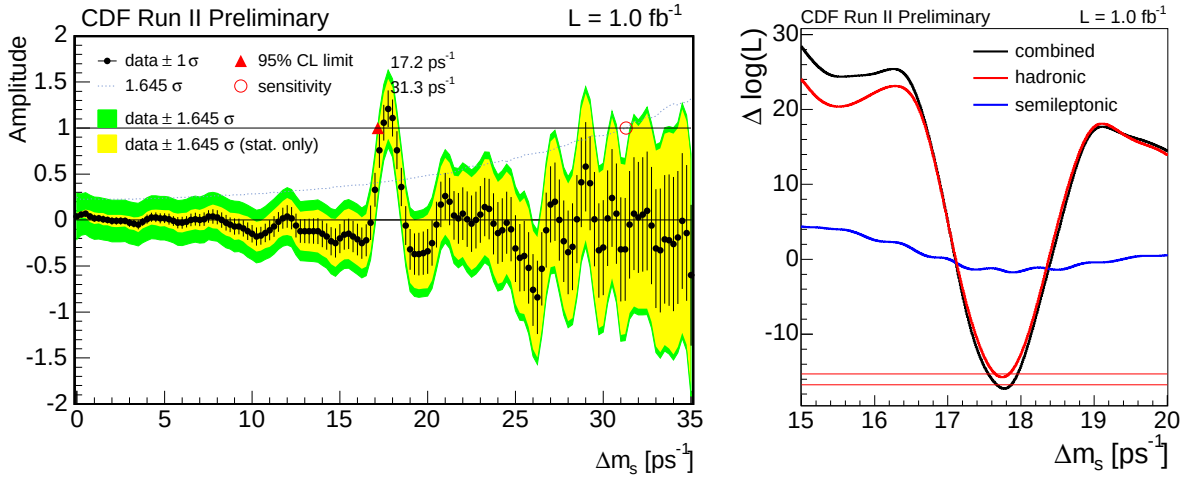
case, the soft neutral γ or π^0 leaves no track in the SVX or COT detectors and is neglected in the reconstruction.

Three semileptonic decay modes are reconstructed in the analysis: $\overline{B}_s^0 \rightarrow \ell^- X$, with $\rightarrow \phi\pi^+$ ($\phi \rightarrow K^+K^-$), $\rightarrow \overline{K}^*(892)^0 K^+$ ($\overline{K}^*(892)^0 \rightarrow K^-\pi^+$) and $\rightarrow \pi^+\pi^-\pi^+$. The hadronic and semileptonic decay modes are complementary. The overall signal sample corresponds to $61\,500$ B_s candidates in the semileptonic decay modes, and $8\,800$ B_s candidates in the partially and fully reconstructed hadronic decay modes. Due to the large branching ratio, the semileptonic decays provide a tenfold advantage in signal rate at the cost of significantly worsened decay-time resolution due to the unmeasured ν momentum. Semileptonic decays dominate the sensitivity to oscillations at lower values of Δm_s. The fully reconstructed hadronic B_s decays have superior decay time resolution, and our large sample of these decays is the unique feature that makes CDF sensitive to much larger values of Δm_s than other experiments.

The reconstructed decay time in the B_s rest frame is $t = m_{B_s} L_{xy}/p_T$, where L_{xy} is the B_s decay flight distance in the transverse plane, and p_T is the transverse momentum of the reconstructed B_s. For fully reconstructed decays, the only significant uncertainty is due to the uncertainty on L_{xy}. Semileptonic and partially reconstructed hadronic decays have an important additional uncertainty on the decay time due to the incomplete reconstruction of the B_s meson. The reconstructed momentum in data is corrected by a factor $\kappa = p_T^{recon}/p_T(B_s)$ which is determined from Monte Carlo simulation, and where p_T^{recon} is the combined p_T of the observed particles (e.g. the lepton and the D_s in the case of semileptonic decays) and the $p_T(B_s)$ is the true p_T of the B_s meson. The average decay time resolution for fully reconstructed decays is $< \sigma_t >= 87$ fs, which corresponds to one fourth of an oscillation period at $\Delta m_s = 17.8$ ps^{-1}. For semileptonic decays, σ_t is worse due to the large missing transverse momentum of the products that were not reconstructed, leading to an effective resolution of about 150 fs.

The flavor of the B_s at production is determined using both opposite-side and same-side flavor tagging techniques. At the Tevatron, the dominant b-quark production mechanisms produce $b\bar{b}$ pairs. Opposite-side tags infer the production flavor of the B_s from the decay products of the b hadron produced from the other b quark in the event. The charge from leptons, from jets, and from identified kaons are used as tags, and the information of these three tags are combined into a dedicated NN to construct a final opposite-side tag. The dilution is measured on a control data sample, resulting in a combined opposite-side tag effectiveness of $\varepsilon\mathcal{D}^2 = 1.8 \pm 0.1\%$.

Same-side flavor tags are based on the charges of associated particles produced in the fragmentation of the b quark that produces the reconstructed B_s. In the simplest picture of fragmentation, the $\overline{B}_s^0$ (B_s^0) is expected to be accompanied by a nearby K^- (K^+). $\mathrm{d}E/\mathrm{d}x$ and time-of-flight information is combined into a combined particle identification likelihood to identify the kaon associated with the B_s production. A NN is used to combine the kaon particle identification like-



(a) (b)

Figure E.2: (a) Combined amplitude scan and (b) the logarithm of the ratio of likelihoods for $A = 1$ and $A = 0$ ($\Lambda = \mathcal{L}(A = 1)/\mathcal{L}(A = 0)$).

lihood with kinematic quantities of the kaon candidate into a single tagging variable T. Tracks close in phase space to the B_s candidate are considered as same-side kaon tag candidates, and the track with the largest value of T is selected as the tagging track. The dilution of the same-side tag is predicted using MC simulated samples, and validated using B^- and $\overline{B}^0_d$ control samples. The effectiveness of this flavor tag is found to be $\varepsilon\mathcal{D}^2 = 3.7\%$ (4.8%) in the hadronic (semileptonic) decay sample. The fractional uncertainty on $\varepsilon\mathcal{D}^2$ is approximately 25%. If both a same-side tag and an opposite-side tag are present, we combine the information from both tags assuming they are independent.

To extract the maximum amount of information from the data, we perform an unbinned maximum likelihood fit. In this fit, decay candidates enter the likelihood with different weights. These weights are based upon the properties of the signal versus background, the proper decay time resolution and the probability that the flavor tag is correct. A time dependent measurement of the oscillation frequency is hard due to the fast oscillation and due to detector effects. Instead, a more sensitive unbinned maximum likelihood fit is performed in the frequency domain introducing an additional parameter, the amplitude $\mathcal{A}$, into the likelihood model. Following the method described in [195], we fit for the oscillation amplitude $\mathcal{A}$ while fixing Δm_s to a probe value. Schematically, in the description of the proper decay time model it enters as $\sim [1 \pm \mathcal{A} \cdot \mathcal{D} \cdot \cos(\Delta m_s t)]\ e^{-t/\tau}$. The oscillation amplitude is expected to be consistent with $\mathcal{A} = 1$ when the probe value is the true oscillation frequency, and consistent with $\mathcal{A} = 0$ when the probe value is far from the true oscillation frequency. The sensitivity of the measurement is defined by the maximum value of Δm_s where $\mathcal{A} = 1$ is excluded at 95% C.L. if the measured value of $\mathcal{A}$ were zero.

Figure E.2(a) shows the fitted value of the amplitude as a function of the oscillation frequency for all decay modes combined. The sensitivity is 31.3 ps^{-1}. At $\Delta m_s = 17.75$ ps^{-1}, the observed amplitude $\mathcal{A} = 1.21 \pm +0.20$ is consistent with unity, indicating that the data are compatible with B_s^0-$\overline{B}_s^0$ oscillations with that frequency, while the amplitude is inconsistent with zero: $\mathcal{A}/\sigma_{\mathcal{A}} = 6.05$ standard deviations from zero.

To measure Δm_s and to assess the significance of the signal, we use logarithm of the ratio of likelihoods $\Lambda \equiv \log[\mathcal{L}^{\mathcal{A}=0}/\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)]$, where $\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)$ is the likelihood of the data under the hypothesis that Δm_s is the true mixing frequency. The likelihood $\mathcal{L}^{\mathcal{A}=0}$ is independent of Δm_s and represents the likelihood for $\mathcal{A} = 0$, which is equivalent to oscillations with $\Delta m_s = \infty$. Figure E.2(b) shows the value of Λ as a function of Δm_s. The minimal observed value $\Lambda = -17.26$ is found at $\Delta m_s = 17.77$ ps^{-1}. The significance of this signal is quantified by the probability that the observed minimum value of $\Lambda = -17.26$, or a smaller one, would be produced at any value of Δm_s in case of random tags. This probability is evaluated by repeating 350 million times the likelihood scan on the data while using randomized flavor tag decisions, 28 of these scans have $\Lambda < -17.26$, corresponding to a probability of 8×10^{-8} (5.4σ). To measure Δm_s, a fit for the oscillation frequency is performed while fixing $\mathcal{A} = 1$. This results in a measured B_s^0-$\overline{B}_s^0$ oscillation frequency of $\Delta m_s = 17.77 \pm 0.10\,(\text{stat.}) \pm 0.07\,(\text{syst.})$ ps^{-1}. This measurement is used to extract the ratio of the CKM matrix elements $|V_{td}/V_{ts}| = 0.2060 \pm 0.0007(\text{exp.})\,^{+0.0081}_{-0.0060}(\text{theor.})$.

BIBLIOGRAPHY

[1] Tevatron Electroweak Working Group (2008) arXiv:0808.1089 [hep-ex]

[2] W. A. Bardeen, C. T. Hill, and M. Lindner, Phys. Rev. **D41**, 1647 (1990)

[3] V. A. Miransky, M. Tanabashi, and K. Yamawaki, Mod. Phys. Lett. **A4**, 1043 (1989)

[4] A. Hasenfratz, P. Hasenfratz, K. Jansen, J. Kuti, and Y. Shen, Nucl. Phys. **B365**, 79 (1991)

[5] N. Kidonakis and R. Vogt, Phys. Rev. **D68**, 114014 (2003), arXiv:hep-ph/0308222

[6] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995), arXiv:hep-ex/9503002

[7] S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995), arXiv:hep-ex/9503003

[8] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 082004 (2006), arXiv:hep-ex/0606017

[9] D. E. Acosta *et al.* (CDF), Phys. Rev. Lett. **95**, 102002 (2005), arXiv:hep-ex/0505091

[10] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. **D73**, 111103 (2006), arXiv:hep-ex/0511023

[11] C. Amsler *et al.* (Particle Data Group), Phys. Lett. **B667**, 1 (2008)

[12] Z. Sullivan, Phys. Rev. **D70**, 114012 (2004), arXiv:hep-ph/0408049

[13] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, Phys. Rev. **D66**, 054024 (2002), arXiv:hep-ph/0207055

[14] T. M. P. Tait and C. P. Yuan, Phys. Rev. **D63**, 014018 (2001), arXiv:hep-ph/0007298

[15] G. Mahlon and S. J. Parke, Phys. Rev. **D55**, 7249 (1997), arXiv:hep-ph/9611367

[16] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. **D58**, 094021 (1998), arXiv:hep-ph/9807340

[17] D. E. Acosta *et al.* (CDF), Phys. Rev. **D65**, 091102 (2002), arXiv:hep-ex/0110067

[18] B. Abbott *et al.* (D0), Phys. Rev. **D63**, 031101 (2000), arXiv:hep-ex/0008024

[19] D. E. Acosta *et al.* (CDF), Phys. Rev. **D69**, 052003 (2004)

[20] V. M. Abazov *et al.* (D0), Phys. Lett. **B517**, 282 (2001), arXiv:hep-ex/0106059

[21] D. E. Acosta *et al.* (CDF), Phys. Rev. **D71**, 012005 (2005), arXiv:hep-ex/0410058

[22] V. M. Abazov *et al.* (D0), Phys. Lett. **B622**, 265 (2005), arXiv:hep-ex/0505063

[23] V. M. Abazov *et al.* (D0), Phys. Rev. **D75**, 092007 (2007), arXiv:hep-ex/0604020

[24] V. M. Abazov *et al.* (D0), Phys. Rev. Lett. **98**, 181802 (2007), arXiv:hep-ex/0612052

[25] V. M. Abazov *et al.* (D0), Phys. Rev. **D78**, 012005 (2008), arXiv:0803.0739 [hep-ex]

[26] T. Aaltonen *et al.* (CDF), Phys. Rev. Lett. **101**, 252001 (2008), arXiv:0809.2581 [hep-ex]

[27] T. Aaltonen *et al.* (CDF), Phys. Rev. Lett. **103**, 092002 (2009), arXiv:0903.0885 [hep-ex]

[28] V. M. Abazov *et al.* (D0), Phys. Rev. Lett. **103**, 092001 (2009), arXiv:0903.0850 [hep-ex]

[29] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, Rev. Mod. Phys. **71**, S96 (Mar 1999)

[30] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (New York, Usa: Wiley, 1984)

[31] D. Griffiths, *Introduction to elementary particles* (Weinheim, USA: Wiley-VCH, 2008)

[32] D. h. Perkins, *Introduction to High-Energy Physics* (Reading, Usa: Addison-wesley, 1982)

[33] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory* (Reading, USA: Addison-Wesley, 1995)

[34] Y. Fukuda *et al.* (Super-Kamiokande), Phys. Rev. Lett. **81**, 1562 (1998), arXiv:hep-ex/9807003

[35] S. Fukuda *et al.* (Super-Kamiokande), Phys. Rev. Lett. **86**, 5651 (2001), arXiv:hep-ex/0103032

[36] S. Fukuda *et al.* (Super-Kamiokande), Phys. Lett. **B539**, 179 (2002), arXiv:hep-ex/0205075

[37] Q. R. Ahmad *et al.* (SNO), Phys. Rev. Lett. **87**, 071301 (2001), arXiv:nucl-ex/0106015

[38] Q. R. Ahmad *et al.* (SNO), Phys. Rev. Lett. **89**, 011301 (2002), arXiv:nucl-ex/0204008

[39] S. L. Glashow, Nucl. Phys. **22**, 579 (1961)

[40] A. Salam and J. C. Ward, Phys. Lett. **13**, 168 (1964)

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

[42] J. Goldstone, Nuovo Cim. **19**, 154 (1961)

[43] J. Goldstone, A. Salam, and S. Weinberg, Phys. Rev. **127**, 965 (1962)

[44] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964)

[45] P. W. Higgs, Phys. Lett. **12**, 132 (1964)

[46] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964)

[47] P. W. Higgs, Phys. Rev. **145**, 1156 (1966)

[48] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964)

[49] G. Arnison *et al.* (UA1), Phys. Lett. **B122**, 103 (1983)

[50] G. Arnison *et al.* (UA1), Phys. Lett. **B126**, 398 (1983)

[51] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963)

[52] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973)

[53] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983)

[54] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, Phys. Rev. **D50**, 3433 (1994), arXiv:hep-ph/9403384

[55] J. Charles *et al.* (CKMfitter Group), Eur. Phys. J. **C41**, 1 (2005), arXiv:hep-ph/0406184

[56] A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, Eur. Phys. J. **C21**, 225 (2001), arXiv:hep-ph/0104062, `http://ckmfitter.in2p3.fr/`

[57] M. Bona *et al.* (UTfit), JHEP **07**, 028 (2005), arXiv:hep-ph/0501199, `http://www.utfit.org/`

[58] M. Bona *et al.* (UTfit), JHEP **03**, 049 (2008), arXiv:0707.0636 [hep-ph]

[59] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977)

[60] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972)

[61] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977)

[62] H. L. Lai *et al.* (CTEQ), Eur. Phys. J. **C12**, 375 (2000), arXiv:hep-ph/9903282

[63] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C4**, 463 (1998), arXiv:hep-ph/9803445

[64] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Phys. Lett. **B652**, 292 (2007), arXiv:0706.0459 [hep-ph]

[65] S. Pakvasa and H. Sugawara, Phys. Lett. **B82**, 105 (1979)

[66] G. Preparata, Phys. Lett. **B82**, 398 (1979)

[67] K. T. Mahanthappa and M. A. Sher, Phys. Lett. **B86**, 294 (1979)

[68] F. Abe *et al.* (CDF), Phys. Rev. Lett. **64**, 142 (1990)

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

[70] F. Abe *et al.* (CDF), Phys. Rev. Lett. **68**, 447 (1992)

[71] S. Abachi *et al.* (D0), Phys. Rev. Lett. **72**, 2138 (1994)

[72] E. Laenen, J. Smith, and W. L. van Neerven, Nucl. Phys. **B369**, 543 (1992)

[73] F. Abe *et al.* (CDF), Phys. Rev. Lett. **73**, 225 (1994), arXiv:hep-ex/9405005

[74] F. Abe *et al.* (CDF), Phys. Rev. **D50**, 2966 (1994)

[75] S. Abachi *et al.* (D0), Phys. Rev. Lett. **74**, 2422 (1995), arXiv:hep-ex/9411001

[76] S. Abachi *et al.* (D0), Phys. Rev. **D52**, 4877 (1995)

[77] N. Kidonakis, Phys. Rev. **D74**, 114012 (2006), arXiv:hep-ph/0609287

[78] A. Heinson, A. S. Belyaev, and E. E. Boos, Phys. Rev. **D56**, 3114 (1997), arXiv:hep-ph/9612424

[79] J. Alwall *et al.*, Eur. Phys. J. **C49**, 791 (2007), arXiv:hep-ph/0607115

[80] D. O. Carlson and C. P. Yuan, Phys. Lett. **B306**, 386 (1993)

[81] M. Jezabek and J. H. Kuhn, Phys. Lett. **B329**, 317 (1994), arXiv:hep-ph/9403366

[82] G. Mahlon and S. J. Parke, Phys. Rev. **D53**, 4886 (1996), arXiv:hep-ph/9512264

[83] G. Mahlon(2000), arXiv:hep-ph/0011349

[84] G. Mahlon(1998), arXiv:hep-ph/9811219

[85] Z. Sullivan, Phys. Rev. **D66**, 075011 (2002), arXiv:hep-ph/0207290

[86] H.-J. He, T. M. P. Tait, and C. P. Yuan, Phys. Rev. **D62**, 011702 (2000), arXiv:hep-ph/9911266

[87] H.-J. He and C. P. Yuan, Phys. Rev. Lett. **83**, 28 (1999), arXiv:hep-ph/9810367

[88] B. A. Dobrescu and C. T. Hill, Phys. Rev. Lett. **81**, 2634 (1998), arXiv:hep-ph/9712319

[89] R. S. Chivukula, B. A. Dobrescu, H. Georgi, and C. T. Hill, Phys. Rev. **D59**, 075003 (1999), arXiv:hep-ph/9809470

[90] H. Collins, A. K. Grant, and H. Georgi, Phys. Rev. **D61**, 055002 (2000), arXiv:hep-ph/9908330

[91] J. C. Lee, K. Y. Lee, and J. K. Kim, Phys. Lett. **B424**, 133 (1998), arXiv:hep-ph/9711509

[92] E. Malkawi and C. P. Yuan, Phys. Rev. **D61**, 015007 (2000), arXiv:hep-ph/9906215

[93] G. Burdman, R. S. Chivukula, and N. J. Evans, Phys. Rev. **D62**, 075007 (2000), arXiv:hep-ph/0005098

[94] E. Boos, L. Dudko, and T. Ohl, Eur. Phys. J. **C11**, 473 (1999), arXiv:hep-ph/9903215

[95] J. A. Aguilar-Saavedra, Nucl. Phys. **B812**, 181 (2009), arXiv:0811.3842 [hep-ph]

[96] T. Han, M. Hosch, K. Whisnant, B.-L. Young, and X. Zhang, Phys. Rev. **D58**, 073008 (1998), arXiv:hep-ph/9806486

[97] T. Han and J. L. Hewett, Phys. Rev. **D60**, 074015 (1999), arXiv:hep-ph/9811237

[98] M. Hosch, K. Whisnant, and B. L. Young, Phys. Rev. **D56**, 5725 (1997), arXiv:hep-ph/9703450

[99] E. Malkawi and T. M. P. Tait, Phys. Rev. **D54**, 5758 (1996), arXiv:hep-ph/9511337

[100] T. M. P. Tait and C. P. Yuan, Phys. Rev. **D55**, 7300 (1997), arXiv:hep-ph/9611244

[101] A. Datta, P. J. O'Donnell, Z. H. Lin, X. Zhang, and T. Huang, Phys. Lett. **B483**, 203 (2000), arXiv:hep-ph/0001059

[102] J. Cockroft and E. Walton, Proc. of the Royal Soc. of London A **137**, 229 (1932)

[103] C. W. Schmidt. Presented at 1993 Particle Accelerator Conference (PAC 93), Washington, DC, 17-20 May 1993

[104] J. Marriner FERMILAB-PUB-04-092-AD

[105] D. MacGinis, *Run II Handbook* (Fermilab Beam Division, 2001) `http://www-bd.fnal.gov/runII/index.html`

[106] R. Blair *et al.* (CDF-II) FERMILAB-PUB-96-390-E

[107] D. Acosta *et al.*, Phys. Rev. D **71**, 032001 (Feb 2005)

[108] F. Abe *et al.* (CDF), Nucl. Instr. Meth. **A271**, 387 (1988)

[109] C. S. Hill (CDF), Nucl. Instrum. Meth. **A530**, 1 (2004)

[110] A. Sill (CDF), Nucl. Instrum. Meth. **A447**, 1 (2000)

[111] A. A. Affolder *et al.* (CDF), Nucl. Instrum. Meth. **A453**, 84 (2000)

[112] A. A. Affolder *et al.* (CDF), Nucl. Instrum. Meth. **A526**, 249 (2004)

[113] `http://penn01.fnal.gov/~cot/`

[114] D. Acosta *et al.* (CDF-II), Nucl. Instrum. Meth. **A518**, 605 (2004)

[115] Palencia, E., *Calibraciones del Detector de Tiempo de Vuelo del Experimento CDF en el Tevatron*, Master's thesis, Universidad de Cantabria (2005)

[116] L. Balka *et al.* (CDF), Nucl. Instrum. Meth. **A267**, 272 (1988)

[117] S. Bertolucci *et al.* (CDF), Nucl. Instrum. Meth. **A267**, 301 (1988)

[118] M. G. Albrow *et al.* (CDF), Nucl. Instrum. Meth. **A480**, 524 (2002)

[119] G. Ascoli *et al.*, Nucl. Instrum. Meth. **A268**, 33 (1988)

[120] T. Dorigo (CDF), Nucl. Instrum. Meth. **A461**, 560 (2001)

[121] D. Acosta *et al.* (CDF), Nucl. Instrum. Meth. **A461**, 540 (2001)

[122] D. Acosta *et al.* (CDF), Nucl. Instrum. Meth. **A494**, 57 (2002)

[123] S. Klimenko, J. Konigsberg, and T. M. Liss FERMILAB-FN-0741

[124] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002)

[125] `http://web.hep.uiuc.edu/Engin/CDF/XTRP/`

[126] B. Ashmanskas *et al.* (CDF-II), Nucl. Instrum. Meth. **A518**, 532 (2004), arXiv:physics/0306169

[127] W. Wagner *et al.* prepared for International Europhysics Conference on High- Energy Physics (HEP 2001), Budapest, Hungary, 12-18 Jul 2001

[128] B. Angelos *et al.* international conference on Computing in High-Energy Physics and Nuclear Physics (CHEP 2000), Padova Italy, 7-11 Feb 2000

[129] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001), arXiv:hep-ph/0010017

[130] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006), arXiv:hep-ph/0603175

[131] F. Maltoni and T. Stelzer, JHEP **02**, 027 (2003), arXiv:hep-ph/0208156

[132] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81**, 357 (1994), arXiv:hep-ph/9401258

[133] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003), arXiv:hep-ph/0206293

[134] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Phys. Rept. **97**, 31 (1983)

[135] T. Sjostrand, Nucl. Phys. **B248**, 469 (1984)

[136] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003)

[137] R. Veenhof, Nucl. Instrum. Meth. **A419**, 726 (1998)

[138] G. Grindhammer, M. Rudowicz, and S. Peters, Nucl. Instrum. Meth. **A290**, 469 (1990)

[139] E. Gerchtein and M. Paulini(2003), arXiv:physics/0306031

[140] F. Abe *et al.* (CDF), Phys. Rev. **D45**, 1448 (1992)

[141] A. Bhatti *et al.*, Nucl. Instrum. Meth. **A566**, 375 (2006), arXiv:hep-ex/0510047

[142] D. E. Acosta *et al.* (CDF), Phys. Rev. **D71**, 052003 (2005), arXiv:hep-ex/0410041

[143] S. Richter fERMILAB-THESIS-2007-35

[144] M. Feindt and U. Kerzel, Nucl. Instrum. Meth. **A559**, 190 (2006)

[145] E. E. Boos, V. E. Bunichev, L. V. Dudko, V. I. Savrin, and A. V. Sherstnev, Phys. Atom. Nucl. **69**, 1317 (2006)

[146] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano(2009), arXiv:0907.3933 [hep--ph]

[147] J. Lueck(2006), FERMILAB-MASTERS-2006-01

[148] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, Nucl. Phys. **B529**, 424 (1998), arXiv:hep-ph/9801375

[149] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, JHEP **04**, 068 (2004), arXiv:hep-ph/0303085

[150] E. L. Berger and H. Contopanagos(1997), arXiv:hep-ph/9706356

[151] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (Nov 1999)

[152] T. Aaltonen *et al.* (CDF - Run II), Phys. Rev. Lett. **100**, 102001 (2008), arXiv:0711.3717 [hep-ex]

[153] Z. Bern, L. J. Dixon, D. A. Kosower, and S. Weinzierl, Nucl. Phys. **B489**, 3 (1997), arXiv:hep-ph/9610370

[154] Z. Bern, L. J. Dixon, and D. A. Kosower, Nucl. Phys. **B513**, 3 (1998), arXiv:hep-ph/9708239

[155] W. T. Giele, S. Keller, and E. Laenen, Nucl. Phys. Proc. Suppl. **51C**, 255 (1996), arXiv:hep-ph/9606209

[156] R. K. Ellis and S. Veseli, Phys. Rev. **D60**, 011501 (1999), arXiv:hep-ph/9810489

[157] F. Febres Cordero, L. Reina, and D. Wackeroth, Phys. Rev. **D74**, 034007 (2006), arXiv:hep-ph/0606102

[158] J. M. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. **D75**, 054015 (2007), arXiv:hep-ph/0611348

[159] J. M. Campbell *et al.*, Phys. Rev. **D79**, 034023 (2009), arXiv:0809.3003 [hep-ph]

[160] F. F. Cordero, L. Reina, and D. Wackeroth, Phys. Rev. **D80**, 034015 (2009), arXiv:0906.1923 [hep-ph]

[161] T. Aaltonen *et al.* (CDF), Phys. Rev. Lett. **100**, 091803 (2008), arXiv:0711.2901 [hep-ex]

[162] A. Hocker *et al.*, PoS **ACAT**, 040 (2007), arXiv:physics/0703039

[163] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees* (Wadsworth and Brooks, Monterey, CA, 1984)

[164] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2003)

[165] J. Friedman, T. Hastie, and R. Tibshirani, Annals of Statistics **28**, 2000 (1998)

[166] C. Gini, "Variabilità e Mutabilità," (1912), reprinted in *Memorie di Metodologica Statistica*, edited by E. Pizetti and T. Savemini, Rome: Libreria Eredi Virgilio Veschi, 1955

[167] Y. Freund and R. E. Schapire, Journal of Computer and System Sciences **55**, 119 (1997)

[168] D. Acosta *et al.*, Nucl. Instrum. Meth. **A494**, 57 (2002)

[169] R. D. Cousins. Prepared for PHYSTATO5: Statistical Problems in Particle Physics, Astrophysics and Cosmology, Oxford, England, United Kingdom, 12-15 Sep 2005

[170] R. Barlow. In the Proceedings of PHYSTAT2003: Statistical Problems in Particle Physics, Astrophysics, and Cosmology, Menlo Park, California, 8-11 Sep 2003, arXiv:physics/0401042

[171] T. Junk, Nucl. Instrum. Meth. **A434**, 435 (1999), arXiv:hep-ex/9902006

[172] J. Neyman and E. Pearson, Phil. Trans. of the Royal Soc. of London **A31**, 289 (1933)

[173] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975)

[174] J. Lueck FERMILAB-THESIS-2009-33

[175] P. J. Dong(2008), FERMILAB-THESIS-2008-12

[176] S. R. Budd(2008), FERMILAB-THESIS-2008-41

[177] K. Nakamura(2009), FERMILAB-THESIS-2009-13

[178] A. Apresyan *et al.* Public CDF Note Number: CDF/ANAL/TOP/PUBLIC/9650

[179] K. Ackerstaff *et al.* (OPAL), Eur. Phys. J. **C1**, 45 (1998), arXiv:hep-ex/9709027

[180] A. Abulencia *et al.* (CDF), Phys. Rev. **D74**, 072006 (2006), arXiv:hep-ex/0607035

[181] S. Whiteson and D. Whiteson(2006), arXiv:hep-ex/0607012

[182] The LEP Electroweak Working Group, `http://lepewwg.web.cern.ch/LEPEWWG/`

[183] The CDF collaboration. Public CDF Note Number: 9463, http://www-cdf.fnal.gov/physics/new/hdg//Results_files/results/whlnubb%_081107/WH_ME+BDT/homepage.html

[184] The CDF collaboration. Public CDF Note Number: 9468, http://www-cdf.fnal.gov/physics/new/hdg//Results_files/results/whlnubb%_081107/WH_NN/

[185] T. Aaltonen *et al.* (CDF), Phys. Rev. Lett. **103**, 101802 (2009), arXiv:0906.5613 [hep-ex]

[186] T. Han and S. Willenbrock, Phys. Lett. **B273**, 167 (1991)

[187] R. Downing *et al.*, Nucl. Instrum. Meth. **A570**, 36 (2007), arXiv:physics/0606247

[188] The CDF collaboration. Public CDF Note Number: 9887, http://www-cdf.fnal.gov/physics/new/hdg//Results_files/results/hwwmenn%_090814/

[189] D. E. Acosta *et al.* (CDF), Phys. Rev. Lett. **96**, 202001 (2006), arXiv:hep-ex/0508022

[190] M. Okamoto, PoS **LAT2005**, 013 (2006), arXiv:hep-lat/0510113

[191] A. Abulencia *et al.* (CDF - Run II), Phys. Rev. Lett. **97**, 062003 (2006), arXiv:hep-ex/0606027

[192] A. Abulencia *et al.* (CDF), Phys. Rev. Lett. **97**, 242003 (2006), arXiv:hep-ex/0609040

[193] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997)

[194] A. Zell *et al.*, *Stuttgart Neural Network Simulator* (Univ. of Stuttgart, 1997)

[195] H. G. Moser and A. Roussarie, Nucl. Instrum. Meth. **A384**, 491 (1997)