Measurement of Single Top Quark Production at D0 Using a Matrix Element Method

Jovan Mitrevski

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2007

©2007

Jovan Mitrevski

All Rights Reserved

ABSTRACT

Measurement of Single Top Quark Production at D0 Using a Matrix Element Method

Jovan Mitrevski

Until now, the top quark has only been observed produced in pairs, by the strong force. According to the standard model, it can also be produced singly, via an electroweak interaction. Top quarks produced this way provide powerful ways to test the charged-current electroweak interactions of the top quark, to measure $|V_{tb}|$, and to search for physics beyond the standard model.

This thesis describes the application of the matrix element analysis technique to the search for single top quark production with the D0 detector using $0.9\,\mathrm{fb}^{-1}$ of Run II data. From a comparison of the matrix element discriminants between data and the background model, assuming a Standard Model s-channel to t-channel cross section ratio of $\sigma_s/\sigma_t = 0.44$, we measure the single top quark production cross section:

$$\sigma\left(p\bar{p} \to tb + X,\ tqb + X\right) = 4.8^{+1.6}_{-1.4}\ \mathrm{pb}.$$

This result has a p-value of 0.08%, corresponding to a 3.2 standard deviation Gaussian equivalent significance.

Contents

List of Tables **v**

List of Figures **viii**

Keys for Figures **xiv**

Acknowledgements **xv**

1 Introduction **1**

2 The Standard Model **4**

2.1 The Constituents 4

2.2 Symmetries and Gauge Theories, QED, and QCD 6

2.3 Electroweak Interactions 10

3 Single Top Production **22**

3.1 Single Top Event Signature 24

3.2 Polarization Effects 31

3.3 Single Top and the Search for Physics Beyond the Standard Model . 36

4 Experimental Apparatus **40**

4.1 Accelerator Chain 40

4.2 The D0 Detector 45

5 Object Reconstruction **69**

5.1 Tracking 69

5.2 Primary Vertices 71

5.3 Calorimeter Preprocessing 73

5.4 Jets 74

5.5 Electrons 84

5.6 Muons 93

5.7 Missing Transverse Energy 96

5.8 b-Tagging Jets 97

6 The Analysis: Event Selection **102**

6.1 Background Processes 103

6.2 Signal and Background Modeling 107

6.3 Data and Triggers 111

6.4 Event Selection 117

6.5 Correcting the Monte Carlo Model Efficiencies and Resolutions 121

6.6 Modeling the Trigger for Monte Carlo Events 126

6.7 Normalizing the W+jets and Multijets Backgrounds to Data 127

6.8 b-Tagging 134

6.9 Event Yields 141

6.10 Cross Checks of the Background Model 141

7 The Analysis: Matrix Element Technique 147

7.1 Matrix Element Technique Overview 148

7.2 Calculation of the Event Probability Density Functions 149

7.3 Normalization of the Probabilities 152

7.4 The Matrix Elements 153

7.5 Integration Details 155

7.6 Assignment Permutations 160

7.7 Object Transfer Functions 163

7.8 Single Top Discriminant 169

7.9 Cross-Check Samples 176

8 The Analysis: Extracting a Result 182

8.1 Systematic Uncertainties 183

8.2 Extracting a Measurement Using a Bayesian Approach 186

8.3 Generating Ensembles 188

8.4 Calibration of the Method 189

8.5 Expected Results 191

8.6 Observed Results 196

8.7 Event Characteristics 199

9 Summary and Conclusions **205**

9.1 The Current D0 Single Top Results 206

9.2 Outlook 211

A Systematic Uncertainties **234**

B Plots After Selection **241**

C Discriminant Output Plots **254**

List of Tables

2.1 The three generations of leptons 5

2.2 The three generations of quarks 6

2.3 The weak isospin and hypercharge 14

3.1 Single top cross sections 24

4.1 SMT sensor specifications 52

5.1 The variables used by the NN tagger 100

6.1 The Monte Carlo cross sections and numbers of events 110

6.2 Triggers and integrated luminosities for the electron channel 112

6.3 Triggers and integrated luminosities for the muon channel 114

6.4 Trigger details for electron channel 115

6.5 Trigger details for muon channel 115

6.6 Electron channel trigger efficiencies 116

6.7 Muon channel trigger efficiencies 117

6.8 The electron fake rates used in the matrix method 131

6.9 The matrix method numbers 132

6.10 Yields after selection and before b tagging 142

6.11 Yields after selection for events with no b-tagged jets 143

6.12 Yields after selection for events with exactly one b-tagged jet 144

6.13 Yields after selection for events with exactly two b-tagged jets 145

7.1 Factorization scales 151

7.2 The Matrix Elements used in the analysis 154

7.3 Background weights 171

8.1 Electron channel uncertainties for one b-tag and two jets 187

8.2 Expected cross sections 195

8.3 Measured cross sections 197

9.1 Published single top results 206

9.2 Improved single top results 206

9.3 Systematic uncertainties required to extract $|V_{tb}|$ 210

A.1 Electron channel uncertainties for one b-tag and three jets 234

A.2 Muon channel uncertainties for one b-tag and two jets 235

A.3 Muon channel uncertainties for one b-tag and three jets 236

A.4 Electron channel uncertainties for two b-tags and two jets 237

A.5 Electron channel uncertainties for two b-tags and three jets 238

A.6 Muon channel uncertainties for two b-tags and two jets 239

A.7 Muon channel uncertainties for two b-tags and three jets 240

List of Figures

1.1 The preferred mass for the Higgs boson 2

2.1 The CTEQ6M PDF at $Q = 100\,\mathrm{GeV}$ 11

2.2 $Z \to b\bar{b}$ vertex corrections containing a top quark 21

3.1 LO Feynman diagrams for single top quark production via the s-channel and t-channel processes 23

3.2 LO Feynman diagrams for tW associated production 23

3.3 LO Feynman diagrams for s-channel production, with decay 25

3.4 Real emission corrections for the s-channel process 26

3.5 s-channel NLO jet distributions 27

3.6 s-channel NLO lepton distributions 28

3.7 LO Feynman diagrams for t-channel production, with decay 29

3.8 Real emission corrections for the t-channel process 30

3.9 t-channel NLO jet distributions 32

3.10 t-channel NLO lepton distributions 33

3.11 Correlation between the spin of the top quark and the direction of the charged lepton 35

3.12 Feynman diagrams for anomalous single top quark production via s-channel processes 36

3.13 Examples of BSM effects to the single top cross sections 37

3.14 FCNC diagrams for single top quark production via the s-channel and t-channel processes 38

4.1 Integrated luminosity delivered by the Tevatron Collider 41

4.2 Fermilab's accelerator chain 42

4.3 Schematic of the D0 detector 45

4.4 Overview of the trigger and DAQ systems 48

4.5 Overview of the L1 and L2 trigger systems 48

4.6 Schematic of the inner tracker 50

4.7 The silicon microstrip tracker 53

4.8 Cross-section of the scintillator strips used in the preshower detectors 56

4.9 A schematic of a calorimeter cell 59

4.10 The calorimeters 59

4.11 Schematic view of the calorimeter showing pseudo-projective towers . 61

4.12 Muon system 64

4.13 The magnetic field in the toroidal magnets 65

5.1 Diagram showing possible infrared sensitivity in jet algorithms 75

5.2 The jet f_{EM} distributions 79

5.3 JES correction terms 81

5.4 Jet energy scale uncertainties 82

5.5 Difference between data and MC imbalances in γ+jet events 83

5.6 The electron f_{EM} and f_{iso} distributions 85

5.7 Selected electron discrimination variables 90

5.8 Efficiencies for electrons with quality *top_loose* 91

5.9 Efficiencies for electrons with quality *top_tight* 92

5.10 Muon reconstruction efficiencies as a function of η and ϕ 95

5.11 Muon isolation efficiencies 96

5.12 $\not{E}_{\mathrm{T}}$ resolution 97

5.13 b-tagging efficiencies 101

5.14 b-tagging fake rates 101

6.1 Percentage of selected signal events and signal:background 104

6.2 Combining the "soft" with the "hard" t-channel MC diagrams 108

6.3 The triangle cuts 119

6.4 Electron correction factors 123

6.5 Electron trigger turn-on curves 127

6.6 The electron efficiencies used in the matrix method 130

6.7 The muon efficiencies used in the matrix method 130

6.8 The muon fake rates used in the matrix method 132

6.9 The values of α for various zero-tag bins 134

6.10 Electron channel taggability rate function cross checks 136

6.11 Electron channel taggability rate function cross checks 137

6.12 The b-flavor and c-flavor TRFs 139

6.13 The W transverse mass distribution 146

7.1 The p_{T} and η distribution of unmatched quarks 158

7.2 The p_{T} distribution of b-quarks and light quarks 158

7.3 The definition of $p_{\mathrm{T}_{\mathrm{rel}}}$ 162

7.4 The muonic direct decay probability of b quarks 162

7.5 Jet transfer functions for $\left|\eta^{\mathrm{det}}\right| < 0.5$ and $0.5 < \left|\eta^{\mathrm{det}}\right| < 1.0$ 166

7.6 Jet transfer functions for $1.0 < \left|\eta^{\mathrm{det}}\right| < 1.5$ and $1.5 < \left|\eta^{\mathrm{det}}\right| < 1.5$ . . . 167

7.7 One-tag e+jets discriminant performance 172

7.8 One-tag μ+jets discriminant performance 173

7.9 Two-tag e+jets discriminant performance 174

7.10 Two-tag μ+jets discriminant performance 175

7.11 The H_{T} distributions for single-top MC events 177

7.12 $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ cross-check plots in two-jet events 178

7.13 $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ cross-check plots in three-jet events 179

7.14 $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ cross-check plots in two-jet events 180

7.15 $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ cross-check plots in two-jet events 181

8.1 Calibration ensembles 190

8.2 Calibration fit 191

8.3 Expected posteriors for two-jet events 192

8.4 Expected posteriors for three-jet events 193

8.5 Expected posteriors for all events combined 193

8.6 Summary of expected cross sections 194

8.7 Expected p-value 195

8.8 Discriminant result for two-jet events 197

8.9 Discriminant result for three-jet events 198

8.10 Measured posteriors for all events combined 199

8.11 Summary of measured cross sections 200

8.12 Measured zero-signal p-value 201

8.13 Measured SM p-value 201

8.14 t-channel event characteristics 203

8.15 s-channel event characteristics 204

9.1 The single top combined cross section measurements 208

9.2 Combined zero-signal p-value 209

9.3 Posterior probability density for $|V_{tb}f_1^L|^2$ and for $|V_{tb}|^2$ 211

9.4 The single top measurement projections from D0 212

B.1 The transverse momentum of the leading jet, electron channel 242

B.2 The transverse momentum of the leading jet, muon channel 243

B.3 The transverse momentum of the second jet, electron channel 244

B.4 The transverse momentum of the second jet, muon channel 245

B.5 The transverse momentum of the electron 246

B.6 The transverse momentum of the muon 247

B.7 The $\not{E}_{\mathrm{T}}$, electron channel 248

B.8 The $\not{E}_{\mathrm{T}}$, muon channel 249

B.9 The opening angle $\Delta R(\mathrm{jet1}, \mathrm{jet2})$, electron channel 250

B.10 The opening angle $\Delta R(\mathrm{jet1}, \mathrm{jet2})$, muon channel 251

B.11 The W transverse mass, electron channel 252

B.12 The W transverse mass, muon channel 253

C.1 1D discriminant projections for the electron channel with one b tag . 255

C.2 1D discriminant projections for the electron channel with two b tags . 256

C.3 1D discriminant projections for the muon channel with one b tag . . . 257

C.4 1D discriminant projections for the muon channel with two b tags . . 258

C.5 1D discriminant projections for the electron channel with one b tag . 259

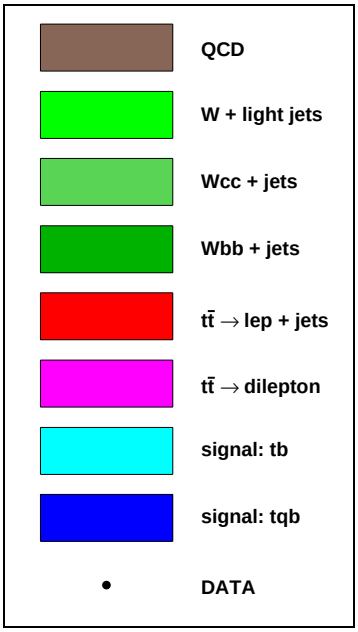
C.6 1D Discriminant projection for the electron channel with two b tags . 260

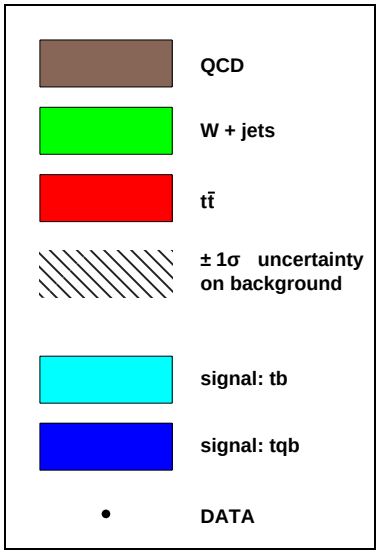
C.7 1D Discriminant projection for the muon channel with one b tag . . . 261

C.8 1D Discriminant projection for the muon channel with two b tags . . 262

Keys for Figures

Below are two keys for figures that contain the standard signals and backgrounds, with the one on the left making finer distinctions between the various backgrounds, and the one on the right showing the uncertainties in the backgrounds, though with coarser distinctions on the background.





ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Prof. John Parsons, for his guidance and support during my graduate studies. His help was invaluable. I would also like to thank the former and current Columbia professors on the D0 Experiment, Mike Tuts, Hal Evans, and Gustaaf Brooijmans, for their help and guidance.

I would like to thank my committee members for reviewing my thesis: Profs. Bob Mawhinney, Emlyn Hughes, Reshmi Mukherjee, and Dr. Howard Gordon.

I would also like to thank Jaroslav Ban and Bill Sippach, with whom I worked on the Level 1 calorimeter trigger upgrade while at Nevis Laboratories. It was my first real work for D0, and it nicely made use of my electrical engineering background. I would like to thank Jon Hays, Christian Schwanenberger, and the rest of the electron ID group, with whom I worked after moving to Fermilab. It was a great introduction to the D0 software environment, the reconstruction algorithms and the detector, and in general working on-site. Finally, the top group, especially the single top subgroup, deserve a very special thanks. Arán García-Bellido, Ann Heinson, and Reinhard Schwienhorst made working on single top a wonderful experience.

I would also like to thank Raphael Galea, Dominik Dannheim, Tulika Bose, Kevin Black, Yannis Katsanos, Paul Russo, Lorenzo Feligioni, and Michel Sorel for making my time at Columbia and Fermilab enjoyable. An extra special thank you goes to Tracy O'Connor, who has been wonderful.

Finally, I would like to thank my mother and father, Pandora and Pavle Mitrevski, for their love and support.

Chapter 1

Introduction

The standard model (SM) of particle physics was developed in the 1960s and 70s, building upon much earlier work. The quark model was proposed, leading to the development of QCD for the strong force, and the weak and electromagnetic forces were unified in the Glashow-Weinberg-Salam (GWS) model. The SM has had remarkable success, and predictions of the model have been confirmed with greater and greater precision [1].

Nevertheless, there is a cause for concern, the primary one being that the Higgs boson, the particle responsible for the electroweak symmetry breaking that is central to the GWS model, has yet to be discovered, and its most favored mass according to electro-weak fits has been excluded, as can be seen in Fig. 1.1. Furthermore, there are theoretical prejudices against the SM concerning the large number of input parameters it has and fact that the Higgs mass would naturally be much higher than needed for its role in the SM if it were not for the very fine tuning of certain parameters. These concerns suggest that studying the predictions of the SM, testing it where it has not been yet tested, remains very worthwhile.

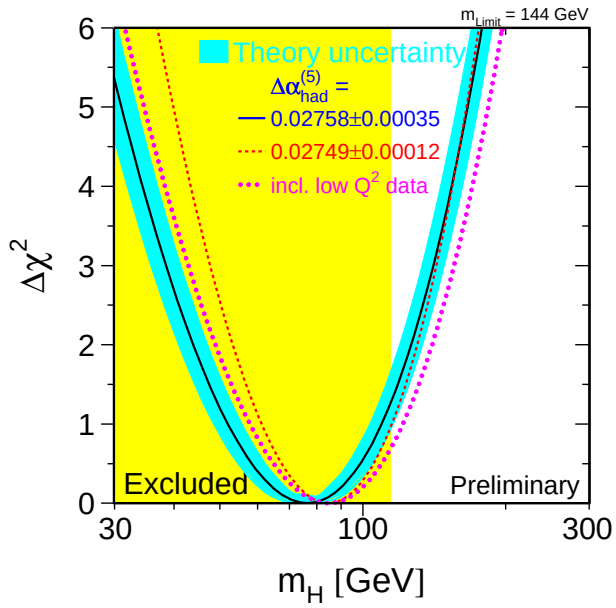


Figure 1.1: The preferred mass for the Higgs boson. (Figure from Ref. [1].)

The top quark was discovered [2, 3] in 1995 by the D0 and CDF collaborations at the Fermilab Tevatron Collider, with its most distinctive feature being its high mass, the highest of any currently known fundamental particle. Until now, the top quark has only been observed via pair production. The work leading to this thesis found first evidence supporting the standard model's prediction that the top quark can also be produced singly.

The outline of thesis is as follows. The next two chapters give a theoretical introduction, first to the SM in general, and then to single top production. Then Chapter 4 discusses the experimental apparatus: the Fermilab Tevatron Collider and the D0 detector. Chapter 5 discusses how the data from the detector is processed to identify and measure physics objects. The three subsequent chapters discuss the analysis. Chapter 6 explains how events are selected and the background models

built. Chapter 7 explains the multivariate method used to discriminate the signal from background: the matrix element method. Chapter 8 explains how a cross section is measured and gives the results. Finally, Chapter 9 gives a summary and future outlook.

Chapter 2

The Standard Model

2.1 The Constituents

The standard model (SM) is a quantum field theory of the fundamental particles and interactions between them, other than gravity.[1]. In the SM there are two categories of spin-$\frac{1}{2}$ particles: leptons and quarks. Leptons interact by the weak and, if charged, the electromagnetic force, but not the strong force. Table 2.1 lists the types of leptons, along with their electric charges and masses [9]. Until recently the standard model assumed that neutrinos were massless, but recent experiments imply that the neutrinos do in fact have nonzero masses [10, 11]. The leptons can be divided into three generations, with corresponding particles across generations having similar properties except for the mass. There also exist the charge conjugates to the leptons: the antineutrinos, the positron, the antimuon, and the antitau. They have the opposite quantum numbers, including charge, of the particles. Whether the neutrino is its own antiparticle is not yet known, and is the subject of investigation [12].

[1]This discussion is based on the descriptions in Ref. [4, 5, 6, 7, 8].

Name	Symbol	Charge (e)	Mass (MeV)
electron neutrino	ν_e	0	< 0.000002
electron	e	-1	0.511
muon neutrino	μ_ν	0	< 0.19
muon	μ	-1	105.7
tau neutrino	ν_τ	0	< 18.2
tau	τ	-1	1777

Table 2.1: The three generations of leptons.

Quarks interact via the strong force, in addition to the weak and electromagnetic forces. They come in three varieties, called "colors": red, blue, and green. These are the charges for quantum chromodynamics (QCD), the theory of the strong force. Table 2.2 lists the types of quarks, along with their electric charges and masses [9] (except for top quark mass, which is from [13]). The quarks too can be divided into three generations, with corresponding particles across the generations sharing similar properties except for the mass. The mass is quite different, though: the top quark mass is roughly 10^4 times greater than the up quark mass. The generation of masses for the fermions is discussed more in Sec. 2.3.1. There also exist the charge conjugates to the quarks: the antiquark, which have the opposite quantum numbers, including change and color (antired, antiblue and antigreen).

The standard model also contains spin-1 particles that mediate the forces, and one spin-0 particle, the Higgs boson, responsible for giving masses. The Higgs boson has not yet been observed. The properties of these particles will be discussed below, where the form of the SM is motivated.

Name	Symbol	Charge (e)	Mass (MeV)
up	u	$2/3$	1.5 to 3.0
down	d	$-1/3$	3 to 7
charm	c	$2/3$	1250
strange	s	$-1/3$	95
top	t	$2/3$	170 900
bottom	b	$-1/3$	4200

Table 2.2: The three generations of quarks.

2.2 Symmetries and Gauge Theories, QED, and QCD

Symmetries have played a fundamental role in the development of the standard model. Noether's theorem showed that for each continuous global symmetry there is a conservation law [14]. Invariance to spatial translation implies conservation of momentum, invariance to translation in time implies energy conservation, and invariance to rotation implies angular momentum conservation, for example. More central to this motivation, though, are internal symmetries, and the conservation laws they imply. For example, take the Lagrangian for a free spin-$\frac{1}{2}$ field ψ:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi. \tag{2.1}$$

This equation is invariant under the transformation, $\psi \to e^{-i\alpha}\psi$, where α is a constant. Noether's theorem then gives $j^\mu = q\bar{\psi}\gamma^\mu\psi$ as the conserved current, which is the electrical current density.

What happens, however, if we allow different space-time points to undergo different phase transformations: $\psi \to e^{-i\alpha(x)}\psi$. Is Eq. 2.1 still invariant to that transfor-

mation? It is not, because $\partial_\mu e^{-i\alpha(x)}\psi \neq e^{-i\alpha(x)}\partial_\mu\psi$. But what if we require it to be, as a fundamental principle, because the choice at locations not causally connected should not matter? Equation 2.1 can be fixed up to comply with this requirement by changing ∂_μ to the *covariant derivative*:

$$\mathcal{D}_\mu \equiv \partial_\mu + iqA_\mu \tag{2.2}$$

and requiring the full transformation to be:

$$\psi \rightarrow e^{-i\alpha(x)}\psi \tag{2.3}$$

$$A_\mu \rightarrow A_\mu + \frac{1}{q}\,\partial_\mu\alpha(x). \tag{2.4}$$

The Lagrangian then becomes

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\mathcal{D}_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.5}$$

where the term involving $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ has been added to make A^μ be a dynamic variable of the Lagrangian. Equation 2.5 is the Lagrangian for QED, where A^μ is the photon. Electromagnetic interactions came into being just by requiring invariance under local phase transformations. In the SM, all interactions come about from such local transformations. A mass term for the photon, $\frac{1}{2}m^2A_\mu A^\mu$, is not invariant to the transformation given in Eq. 2.4, so the photon is required to be massless. Also, the only interactions involving a photon are between the photon and the fermions. This is different from what we will find for other bosons.

Maxwell's equations are invariant to the transformation given in Eq. 2.4 for any function $\alpha(x)$, and Weyl, attempting to find a geometrical foundation for electromagnetism, motivated by Einstein's geometrical foundation for gravity, tried to associate

A^μ with a scale, or *gauge* transformation [15]. Though the effort was originally unsuccessful, with advances in quantum mechanics, including work by Vladimir Fock and Fritz London, A^μ became associated with the phase transformation described above, though the terminology, a gauge transformation, remained [16, 17, 18].

The gauge symmetry discussed above is called a U(1) symmetry, because the transformation, $\psi \to e^{-i\alpha(x)}\psi$, can be considered of the form $\psi \to U\psi$, where U is a one-dimensional unitary matrix. The symmetry is also called an Abelian symmetry because the members of the group, the one-dimensional unitary matrices, commute.

It is possible to have gauge transformations using non-Abelian groups [19]. For example, the set of 3×3 unitary matrices with a determinant equal to one forms the SU(3) group. Any matrix in that set can be written as $U[\alpha] = \exp\left(-i\frac{\lambda^a}{2}\alpha^a\right)$, with summation over a from 1 to 8 implied, where λ^a are the Gell-Mann matrices. Let

$$\psi = \begin{pmatrix} \psi_{\text{red}} \\ \psi_{\text{blue}} \\ \psi_{\text{green}} \end{pmatrix} \tag{2.6}$$

be a triplet for any particular quark flavor. The free Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{2.7}$$

though now ψ is a triplet. We immediately see that this equation is invariant to the global SU(3) gauge transformation, $\psi \to U[\alpha]\psi$. We proceed as before and require the Lagrangian to be invariant under a local gauge transformation: $\psi \to U[\alpha(x)]\psi$.

The solution is similar: ∂_μ needs to be replaced by

$$\mathcal{D}_\mu \equiv \partial_\mu - ig\frac{\lambda^a}{2}A^a_\mu. \tag{2.8}$$

If we define

$$A_\mu = \frac{\lambda^a}{2}A^a_\mu \tag{2.9}$$

$$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \tag{2.10}$$

where $G_{\mu\nu}$ is the field strength tensor, analogous to $F_{\mu\nu}$ in the Abelian case, we get the standard QCD Lagrangian:

$$\mathcal{L} = \sum_q^{n_f} \bar{\psi}_q(i\gamma^\mu \mathcal{D}_\mu - m_q)\psi_q - \frac{1}{2}\mathrm{tr}(G_{\mu\nu}G^{\mu\nu}), \tag{2.11}$$

where the summation is over the quark flavors q. The eight A^a_μ spin-1 particles are called gluons, which mediate the strong force. As with the photon, gluons are massless, but unlike the photon, gluons interact with each other. The interactions between the gluons arises because of the non-Abelian nature of the symmetry group, specifically because of the commutator in the definition of $G_{\mu\nu}$. The gluons carry a color charge, while the photons do not.

A special property of QCD, though not absolutely proven, is confinement. As the distances between quarks become greater, the energy between them increases, until there is enough energy to create a new quark-antiquark pair, for example. Therefore, free particles with a net color charge are never observed; only color singlets are observed. For this reason free quarks have never been observed, but only hadrons: mesons, which are made up of a quark-antiquark pair, and baryons, which are made

up of three quarks in a color singlet state. The proton is a baryon consisting mainly of two up quarks and a down quark, and a neutron is a baryon consisting mainly of one up quark and two down quarks. In addition to those quarks, there is a "sea" of other quarks and gluons that appear due to quantum fluctuations.

Another property of QCD is called asymptotic freedom. The greater the energy of the collision, the smaller the coupling constant is. Thus, in high-energy collisions, the quarks behave as almost free particles within the hadron. At the Tevatron, the collisions look as if they are between two quarks (or gluons), not between a proton and an antiproton.

The distribution of partons within a hadron with a given momentum fraction of the total momentum is parametrized by a parton distribution function (PDF). The PDF is needed whenever cross sections involving high-energy interactions with hadrons are calculated. Specifically, the PDF distribution, $f(x, Q)dx$, is the number of quarks or gluons of a particular type (u, d, etc.) in a particular type of hadron (proton, pion, etc) carrying a momentum fraction between x and $(x + dx)$ of the hadron's momentum in the infinite momentum frame, when probed by an interaction with Q momentum transfer. Figure 2.1 gives the PDF for protons for $Q = 100\,\mathrm{GeV}$ from the CTEQ6M NLO parametrization [20].

2.3 Electroweak Interactions

The weak force was first observed in β decay. Pauli postulated the existence of the neutrino to salvage the conservation of energy and momentum in such decays, and Fermi suggested a four-fermion interaction model to describe the interaction [21, 22]. Then there was a surprising discovery: the weak force does not conserve parity [23,

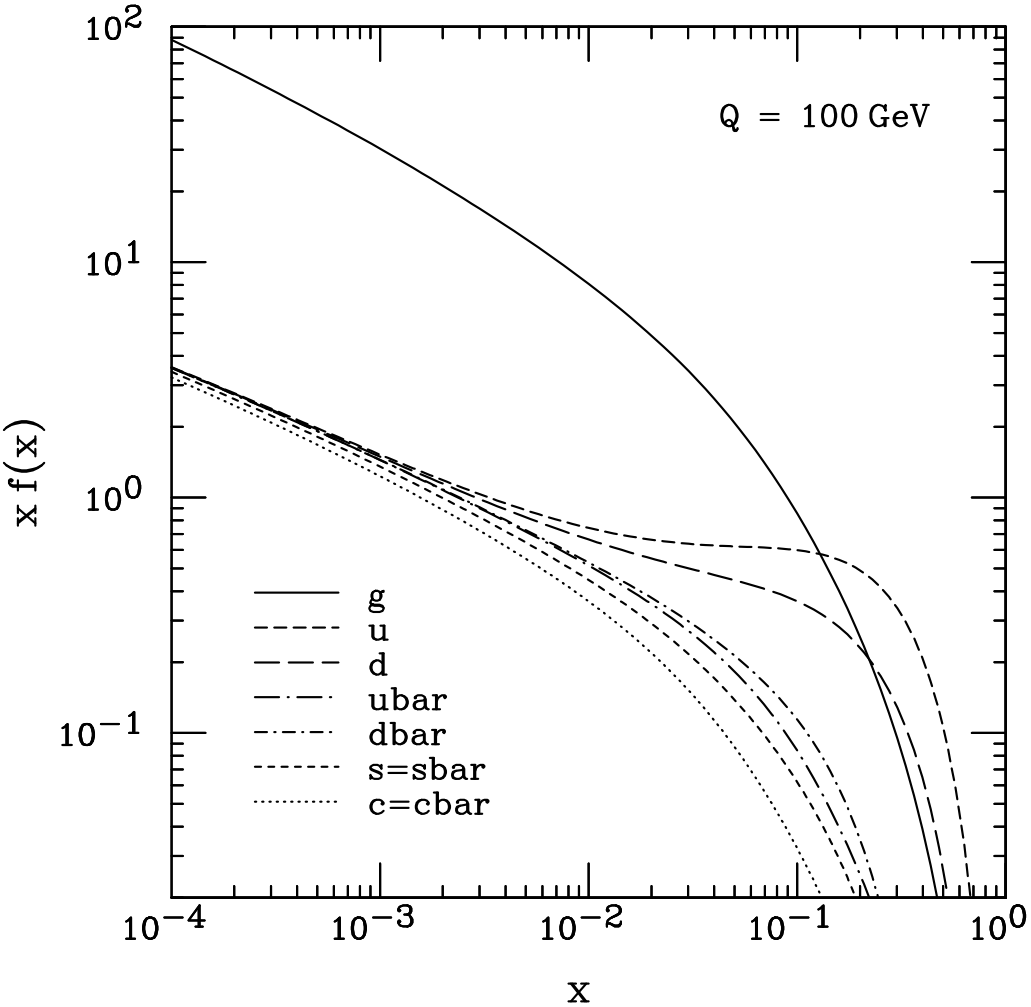


Figure 2.1: The CTEQ6M parton distribution function at $Q = 100\,\mathrm{GeV}$. The value $f(x)dx$ is the number of quarks or gluons of the particular type in a proton carrying a momentum fraction between x and $(x + dx)$ of the proton's momentum. (Figure from Ref. [20].)

24]. The theory was fixed up by making this interaction V−A, meaning that it involves the subtraction of a part that transforms like an axial vector from a part that transforms like a vector. More unexpected results followed when charge times parity (CP) violation was discovered in the K-meson system [25]. There was no easy way to include this feature in the theory at the time.

A problem with the four-fermion theory is that it is not renormalizable and violates unitarity. A modification was to include an intermediate vector boson to transmit the force, much like the photon in QED; however, since the weak force is weak and short-range, this boson is required to be massive. Making a renormalizable theory that has massive vector bosons massive, however, is difficult. In 1957, Schwinger suggested that the weak force and electromagnetism should be unified [26]. The solution is to make the boson acquire a mass by spontaneous symmetry breaking using the Higgs mechanism [27, 28, 29, 30, 31, 32, 33, 34, 35]. The current theory, called the Glashow-Weinberg-Salam (GWS) model, unifies the weak force and the electromagnetic force in a local gauge theory that has an $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry, broken by the Higgs mechanism, into a theory containing only $\mathrm{U}(1)_{\mathrm{EM}}$ [36, 37, 38]. In 1971 't Hooft proved that all gauge theories, even those with spontaneous symmetry breaking, are renormalizable, provided that anomalies cancel [39, 40]. Then in 1973 neutral current weak interactions, predicted by the GWS theory, were first observed with the Gargamelle detector [41]. The theory has been confirmed in all measurements so far, except that the Higgs boson has not been discovered.

One can construct two projection operators using the γ_5 matrix:

$$P_{\mathrm{R}} = \frac{1+\gamma_5}{2} \qquad\qquad P_{\mathrm{L}} = \frac{1-\gamma_5}{2} \tag{2.12}$$

which select the right-handed and left-handed spinors. Using the notation, $\psi_{\mathrm{L}} \equiv P_{\mathrm{L}}\psi$

and $\psi_{\mathrm{R}} \equiv P_{\mathrm{R}}\psi$, in the GWS model the leptons are arranged in the following SU(2) singlets and doublets:

$$L_1^\ell \equiv \begin{pmatrix} \nu_{e\mathrm{L}} \\ e_{\mathrm{L}} \end{pmatrix} \qquad L_2^\ell \equiv \begin{pmatrix} \nu_{\mu\mathrm{L}} \\ \mu_{\mathrm{L}} \end{pmatrix} \qquad L_3^\ell \equiv \begin{pmatrix} \nu_{\tau\mathrm{L}} \\ \tau_{\mathrm{L}} \end{pmatrix} \tag{2.13}$$

$$R_1^\ell \equiv e_{\mathrm{R}} \qquad R_2^\ell \equiv \mu_{\mathrm{R}} \qquad R_3^\ell \equiv \tau_{\mathrm{R}}.$$

Note that there are no right handed neutrinos in the model. Right handed neutrinos do not interact in the theory, and whether they exist or not is not known. In the SM as described here, the neutrinos will remain massless.

The quarks, replicated three times, once for each color index $a = (r, b, g)$, are arranged in the following SU(2) singlets and doublets:

$$L_1^{q,a} \equiv \begin{pmatrix} u_{\mathrm{L}}^a \\ d_{\mathrm{L}}^a \end{pmatrix} \qquad L_2^{q,a} \equiv \begin{pmatrix} c_{\mathrm{L}}^a \\ s_{\mathrm{L}}^a \end{pmatrix} \qquad L_3^{q,a} \equiv \begin{pmatrix} t_{\mathrm{L}}^a \\ b_{\mathrm{L}}^a \end{pmatrix} \tag{2.14}$$

$$R_1^{u,a} \equiv u_{\mathrm{R}}^a \qquad R_1^{d,a} \equiv d_{\mathrm{R}}^a \qquad R_2^{u,a} \equiv c_{\mathrm{R}}^a \qquad R_2^{d,a} \equiv s_{\mathrm{R}}^a \qquad R_3^{u,a} \equiv t_{\mathrm{R}}^a \qquad R_3^{d,a} \equiv b_{\mathrm{R}}^a.$$

The color index will often be suppressed in the equations to avoid clutter.

Each member of a given doublet (and trivially each member of a given singlet) has the same value for the charge of the fermion (Q) minus the weak isospin value T_3, so we define the weak hypercharge as $Y = 2(Q - T_3)$, using conventional scaling. Using this definition, Table 2.3 shows the hypercharge for each singlet and doublet.

Let W_μ^i be the three gauge fields associated with the $\mathrm{SU}(2)_L$ symmetry and B_μ be the gauge field associated with the $\mathrm{U}(1)_Y$ symmetry. Define the covariant derivative

Singlet or Doublet	Weak Isospin (T)	Hypercharge (Y)
L^ℓ	$1/2$	-1
R^ℓ	0	-2
L^q	$1/2$	$1/3$
R^u	0	$4/3$
R^d	0	$-2/3$

Table 2.3: The weak isospin and hypercharge of each singlet or doublet.

to be:

$$\mathcal{D}_\mu \equiv \partial_\mu - ig\mathbf{T}\cdot\mathbf{W}_\mu - ig'\frac{Y}{2}B_\mu \tag{2.15}$$

where $\mathbf{T} = \frac{\boldsymbol{\tau}}{2}$ in cases when $T = 1/2$ and $\mathbf{T} = \mathbf{0}$ for the $T = 0$ cases. With $F^i_{\mu\nu}$ and $G_{\mu\nu}$ being the associated field strength tensors, the Lagrangian contains the term:

$$\mathcal{L}_A = -\frac{1}{4}F^i_{\mu\nu}F^{i\,\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu}. \tag{2.16}$$

The terms in the Lagrangian for the interactions of each generation of quarks and leptons is:

$$\mathcal{L}_{\text{int}} = \bar{L}^\ell_j i\gamma^\mu\mathcal{D}_\mu L^\ell_j + \bar{R}^\ell_j i\gamma^\mu\mathcal{D}_\mu R^\ell_j + \bar{L}^{q,a}_j i\gamma^\mu\mathcal{D}_\mu L^{q,a}_j + \bar{R}^{u,a}_j i\gamma^\mu\mathcal{D}_\mu R^{u,a}_j + \bar{R}^{d,a}_j i\gamma^\mu\mathcal{D}_\mu R^{d,a}_j \tag{2.17}$$

with j being the generation index $(1, 2, 3)$ and a the color index (r, b, g).

At this point, however, everything is massless. To get massive particles the GWS theory uses the Higgs mechanism. Let

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.18}$$

be an SU(2) doublet of complex scalar particles, with the charges as specified by the superscripts (and hence $Y = 1$). The terms in the Lagrangian for this doublet are given by

$$\mathcal{L}_{\text{Higgs}} = (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) - V(\Phi) \tag{2.19}$$

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.20}$$

and the Yukawa couplings:

$$\mathcal{L}_{\text{Yukawa}} = f^\ell_{jk} \bar{L}^\ell_j \Phi R^\ell_k + f^u_{jk} \bar{L}^{q,a}_j \tilde{\Phi} R^{u,a}_k + f^d_{jk} \bar{L}^{q,a}_j \Phi R^{d,a}_k + h.c. \tag{2.21}$$

with $\tilde{\Phi} \equiv i\tau_2 \Phi^*$ and j and k being generation indices. If $\mu > 0$, the lowest potential energy is no longer at $\langle \Phi \rangle_0 = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. Instead, the lowest potential energy comes at $\langle \Phi^\dagger \Phi \rangle_0 = v^2/2$ with $v = \sqrt{\mu^2/\lambda}$. Because of the SU(2) symmetry, we can choose

$$\langle \Phi \rangle_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{2.22}$$

with no loss in generality. $\langle \Phi \rangle_0$ is the vacuum expectation value (VEV), and after it is chosen, it breaks the $\text{SU}(2)_L \otimes \text{U}(1)_Y$ symmetry. We can reparametrize Φ as

$$\Phi = e^{-i\boldsymbol{\zeta}\cdot\boldsymbol{\tau}/v} \begin{pmatrix} 0 \\ (v+\eta)/\sqrt{2} \end{pmatrix} \tag{2.23}$$

and follow it with a gauge transformation, $\Phi' = U(\boldsymbol{\zeta})\Phi$, $U(\boldsymbol{\zeta}) = \exp(i\boldsymbol{\zeta}\cdot\boldsymbol{\tau}/v)$ to the *U-gauge* to get:

$$\Phi' = \begin{pmatrix} 0 \\ (v+\eta)/\sqrt{2} \end{pmatrix} \tag{2.24}$$

$$L' = U(\boldsymbol{\zeta})L \qquad R' = R \tag{2.25}$$

$$\frac{\boldsymbol{\tau}\cdot\mathbf{W}'_\mu}{2} = U(\boldsymbol{\zeta})\left(\frac{\boldsymbol{\tau}\cdot\mathbf{W}_\mu}{2}\right)U^{-1}(\boldsymbol{\zeta}) - \frac{i}{g}\left[\partial_m u U(\boldsymbol{\zeta})\right]U^{-1}(\boldsymbol{\zeta}) \tag{2.26}$$

$$B'_\mu = B_\mu. \tag{2.27}$$

The η field is associated with the Higgs particle, which is a neutral scalar particle. In this new gauge, the Higgs self interaction Lagrangian becomes:

$$\mathcal{L}_{\text{Higgs self}} = \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2 - \lambda v\eta^3 - \frac{\lambda}{4}\eta^4, \tag{2.28}$$

showing the η particle acquiring a mass of $\sqrt{2}\mu$. Using $\theta_{\mathrm{W}} \equiv \arctan(g'/g)$, the Weinberg weak mixing angle, four gauge fields are defined as:

$$W^\pm_\mu \equiv \frac{W'^{(1)}_\mu \mp W'^{(2)}_\mu}{\sqrt{2}} \tag{2.29}$$

$$Z_\mu \equiv \cos\theta_{\mathrm{W}} W'^{(3)}_\mu - \sin\theta_{\mathrm{W}} B'_\mu \tag{2.30}$$

$$A_\mu \equiv \sin\theta_{\mathrm{W}} W'^{(3)}_\mu + \cos\theta_{\mathrm{W}} B'_\mu. \tag{2.31}$$

The charged $W^\pm_\mu$ and neutral Z_μ fields are the gauge bosons that mediate the weak force, and A_μ is the photon. From Eq. 2.19, the mass terms for the vector bosons can be found:

$$\mathcal{L}_{\text{vector mass}} = \frac{g^2v^2}{4}W^+_\mu W^{-\mu} + \frac{1}{2}\frac{(g^2+g'^2)v^2}{4}Z_\mu Z^\mu \tag{2.32}$$

resulting in $M_W = gv/2$, $M_Z = M_W/\cos\theta_{\mathrm{W}}$, and a massless photon. The Higgs mechanism was able to provide the masses without spoiling renormalizability as manually-inserted masses would. The current measured value for the W mass is 80.398 ± 0.025 GeV [1] and for the Z mass is 91.1876 ± 0.0021 GeV [9].

Simplifying Eq. 2.17 with the new definitions, the *charged current* interactions,

i.e., interactions between fermions and W bosons, become

$$\begin{aligned}\mathcal{L}_{\mathrm{CC}} &= \frac{g}{2\sqrt{2}}\left[\bar{\nu}_j\gamma^\mu(1-\gamma_5)e_jW_\mu^+ + \bar{e}_j\gamma^\mu(1-\gamma_5)\nu_jW_\mu^-\right.\\ &\quad \left.+\,\bar{u}_j\gamma^\mu(1-\gamma_5)d_jW_\mu^+ + \bar{d}_j\gamma^\mu(1-\gamma_5)u_jW_\mu^-\right] \end{aligned} \tag{2.33}$$

with the notation, $e_j = (e, \mu, \tau)$, $\nu_j = (\nu_e, \nu_\mu, \nu_\tau)$, $u_j = (u, c, t)$, and $d_j = (d, s, b)$. The *neutral current* interactions, i.e., interactions between fermions and the neutral bosons, Z and A, become

$$\mathcal{L}_{\mathrm{NC}} = \sum_k^{\text{ferm.}}\left\{\mathsf{e}Q_k\bar{\psi}_k\gamma^\mu\psi_kA_\mu + \frac{g}{2\cos\theta_{\mathrm{W}}}\left[g_{\mathrm{L}}^k\bar{\psi}_k\gamma^\mu(1-\gamma_5)\psi_k + g_{\mathrm{R}}^k\bar{\psi}_k\gamma^\mu(1+\gamma_5)\psi_k\right]Z_\mu\right\} \tag{2.34}$$

where

$$\mathsf{e} \equiv g\sin\theta_{\mathrm{W}} \qquad g_{\mathrm{R|L}}^k \equiv T_3(\psi_{\mathrm{R|L},k}) - Q(\psi_{\mathrm{R|L},k})\sin^2\theta_{\mathrm{W}}. \tag{2.35}$$

The standard electromagnetic interaction is included in Eq. 2.34. Because the GWS theory is a non-Abelian gauge theory, like QCD, the vector bosons also interact between themselves. More information can be found in the references on which this derivation is based (especially [4, 5, 6, 8]).

2.3.1 Fermion Mass

The Higgs mechanism can also give mass to the fermions. Reworking Eq. 2.21 in the U-gauge results in:

$$\begin{aligned}\mathcal{L}_{\mathrm{Yukawa}} &= \frac{\eta}{\sqrt{2}}\left[f_{jk}^\ell\bar{e}_{\mathrm{L},j}e_{\mathrm{R},k} + f_{jk}^u\bar{u}_{\mathrm{L},j}u_{\mathrm{R},k} + f_{jk}^d\bar{d}_{\mathrm{L},j}d_{\mathrm{R},k}\right] +\\ &\quad \frac{v}{\sqrt{2}}\left[f_{jk}^\ell\bar{e}_{\mathrm{L},j}e_{\mathrm{R},k} + f_{jk}^u\bar{u}_{\mathrm{L},j}u_{\mathrm{R},k} + f_{jk}^d\bar{d}_{\mathrm{L},j}d_{\mathrm{R},k}\right] + h.c. \end{aligned} \tag{2.36}$$

The first line shows the interactions between the Higgs field and the fermions, and the second line shows the masses of the fermions. However, there is one complication: the $M^b = \left[-\frac{v}{\sqrt{2}} f^b_{jk}\right]$ matrices are not required to be diagonal. One can define diagonal matrices, $M^b_d = S^{b\dagger} M^b T^b$, where S and T are unitary matrices, and then have

$$e_{\mathrm{L},i} = S^\ell_{ij} e^m_{\mathrm{L},j} \qquad u_{\mathrm{L},i} = S^u_{ij} u^m_{\mathrm{L},j} \qquad d_{\mathrm{L},i} = S^d_{ij} d^m_{\mathrm{L},j} \tag{2.37}$$

$$e_{\mathrm{R},i} = T^\ell_{ij} e^m_{\mathrm{R},j} \qquad u_{\mathrm{R},i} = T^u_{ij} u^m_{\mathrm{R},j} \qquad d_{\mathrm{L},i} = T^d_{ij} d^m_{\mathrm{R},j} \tag{2.38}$$

where the fields with the m superscript are the mass eigenstates. The neutral current interaction, Eq. 2.34, is invariant to whether the ψ or ψ^m fields are used, but the charged current interaction, Eq. 2.33, is not. However, any mixing can be localized to only the neutrinos in the lepton case (though in the standard GWS theory, where they are massless, there is no need for such mixing) and the down-type quarks in the quark case. Using this convention, the Yukawa couplings can be written as

$$\begin{aligned} \mathcal{L}_{\mathrm{Yukawa}} &= \frac{\eta}{\sqrt{2}} \left[f^\ell_j \bar{e}^m_{\mathrm{L},j} e^m_{\mathrm{R},j} + f^u_j \bar{u}^m_{\mathrm{L},j} u^m_{\mathrm{R},j} + f^d_j \bar{d}^m_{\mathrm{L},j} d^m_{\mathrm{R},j} \right] + \\ & \quad \frac{v}{\sqrt{2}} \left[f^\ell_j \bar{e}^m_{\mathrm{L},j} e^m_{\mathrm{R},j} + f^u_j \bar{u}^m_{\mathrm{L},j} u^m_{\mathrm{R},j} + f^d_j \bar{d}^m_{\mathrm{L},j} d^m_{\mathrm{R},j} \right] + h.c. \end{aligned} \tag{2.39}$$

The mass of the top quark, with a March 2007 D0 and CDF combined value of $m_t = 170.9 \pm 1.1(\mathrm{stat}) \pm 1.5(\mathrm{syst})\,\mathrm{GeV}$ [13], roughly that of a gold atom, is 10^4 times greater than the mass of an up quark. From Eq. 2.39, the top mass is given by $m_t = (v/\sqrt{2})|f^u_3|$. The numerical value of $v = (\sqrt{2}G_{\mathrm{F}})^{-\frac{1}{2}} \approx 246\,\mathrm{GeV}$ results in $|f^u_3| \approx 1$. The near unity Yukawa coupling to the Higgs boson, leads one to wonder if the top quark plays any special role in electroweak symmetry breaking (EWSB). The idea that a $t\bar{t}$ condensate plays the role of a Higgs boson, breaking the symmetry has been proposed [42, 43, 44], for example. Topcolor models add

a new strong interactions coupling to the third generation in order to form the $t\bar{t}$ condensate [45, 46, 47, 48, 49]. Topflavor places the third generation in a separate SU(2) group from the light quarks [50, 51, 52]. For a more complete review see [53].

2.3.2 The CKM Matrix

The charged current interaction term, $(1/2)\bar{u}_j\gamma^\mu(1-\gamma_5)d_jW^+$, in Eq. 2.33, can be written in terms of mass eigenstates as follows:

$$\begin{aligned}\bar{u}_j\gamma^\mu(1-\gamma_5)d_jW^+ &= 2\bar{u}_{\mathrm{L},j}\gamma^\mu d_{\mathrm{L},j}W^+ \\ &= 2\bar{u}^m_{\mathrm{L},j}\gamma^\mu[S^{u\dagger}S^d]_{jk}d^m_{\mathrm{L},k}W^+ \\ &= \bar{u}^m_j\gamma^\mu(1-\gamma_5)[S^{u\dagger}S^d]_{jk}d^m_kW^+. \end{aligned} \tag{2.40}$$

The matrix,

$$V_{\mathrm{CKM}} \equiv S^{u\dagger}S^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.41}$$

is known as the Cabibbo-Kobayashi-Mashawa matrix [54, 55]. Using the convention mentioned previously, where all mixing in the quark sector is localized in down-type quarks, one can view the CKM matrix as relating the mass eigenstates to the weak eigenstates:

$$\begin{pmatrix} d \\ c \\ b \end{pmatrix}_w = V_{\mathrm{CKM}} \begin{pmatrix} d \\ c \\ b \end{pmatrix}_m . \tag{2.42}$$

The CKM matrix is unitary, with a dimension of $n_{\mathrm{gen}} \times n_{\mathrm{gen}}$, or 3×3 given three generations. The CKM matrix does not have to be real, and interference caused by

imaginary phases can lead to the CP violation observed in, for example, K-mesons. In fact, CP violation was the original motivation of Koboyashi and Mashawa for expanding the original 2×2 Cabbibo matrix even before there was any evidence of a third generation.

Assuming three generations and unitarity, the magnitudes for the CKM matrix elements are [9]:

$$|V_{ij}| = \begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272 \pm 0.0010 & (3.96 \pm 0.09) \times 10^{-3} \\ 0.2271 \pm 0.0010 & 0.97296 \pm 0.00024 & \left(42.21^{+0.10}_{-0.80}\right) \times 10^{-3} \\ \left(8.14^{+0.32}_{-0.64}\right) \times 10^{-3} & \left(41.61^{+0.12}_{-0.78}\right) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix}. \quad (2.43)$$

The entries along the diagonal are the largest, meaning that the strongest interactions are between mass eigenstates of the same generation.

The CKM matrix element V_{tb} is of particular interest to single top searches, because the single top cross section is proportional to $|V_{tb}|^2$. Thus, measuring the single top cross section provides a direct way to measure $|V_{tb}|$, without assuming unitarity or three generations [56]. In contrast, the value of $|V_{tb}| = 0.999100^{+0.000034}_{-0.000004}$, given in Eq. 2.43, is derived assuming unitarity and three generations. The related measurement of $\mathrm{B}(t \to Wb)/\mathrm{B}(t \to Wq)$ measures

$$R = \frac{|V_{tb}|}{\sum_i |V_{ti}|}, \quad (2.44)$$

which D0 and CDF measured to be $1.03^{+0.19}_{-0.17}$ and $1.12^{+0.21}_{-0.19}(\mathrm{stat})^{+0.17}_{-0.13}(\mathrm{syst})$ respectively [57, 58]. If three generations are assumed, $\sum_i |V_{ti}| = 1$, so $|V_{tb}|$ can be extracted from the measurement. Both analyses set the same limit of $|V_{tb}| > 0.78$ at 95% confidence level.

An indirect measurement of $|V_{tb}|$ has also been done based on precision electroweak data results from LEP, SLC, the Tevatron, and neutrino experiments [59]. The result mostly comes from loop corrections to the $Z \to b\bar{b}$ interaction, examples of which are shown in Fig. 2.2. A result is $|V_{tb}| = 0.77^{+0.18}_{-0.24}$.

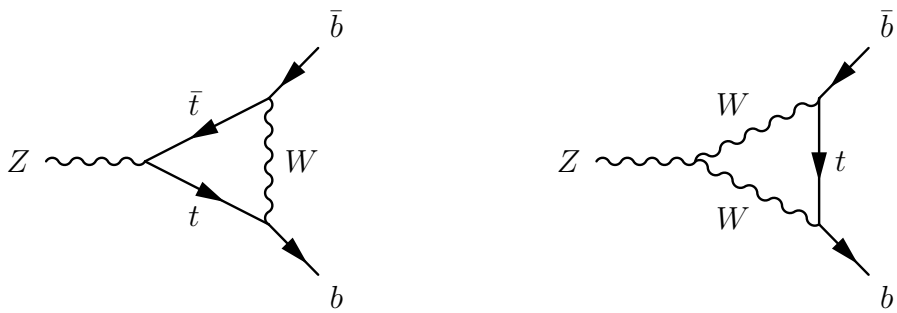


Figure 2.2: $Z \to b\bar{b}$ vertex corections containing a top quark.

Chapter 3

Single Top Production

Single top production refers to the production of one top quark by the charged-current electroweak interaction, making it a powerful probe of charged-current interactions involving the top quark. In $t\bar{t}$ events, on the other hand, the top quark's charged-current interactions can only be studied in the decays. Single top productions has therefore been extensively studied [60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71]. A useful review is provided in Ref. [56].

Single top production can be classified into three types:

- **s-channel:** Single top production via an s-channel process involves a the exchange of a time-like W boson with $Q_W^2 > (m_t + m_b)^2$, where Q_W is the four-vector squared of the W boson. A representative leading order (LO) Feynman diagram is given in Fig. 3.1(a).

- **t-channel:** Single top production via a t-channel process involves the exchange of a space-like W boson with $Q_W^2 < 0$. A representative LO Feynman diagram is given in Fig. 3.1(b).

- tW **associated production:** A single top quark can be created in association with an an on-shell W boson. Representative LO Feynman diagrams are given in Fig. 3.2.

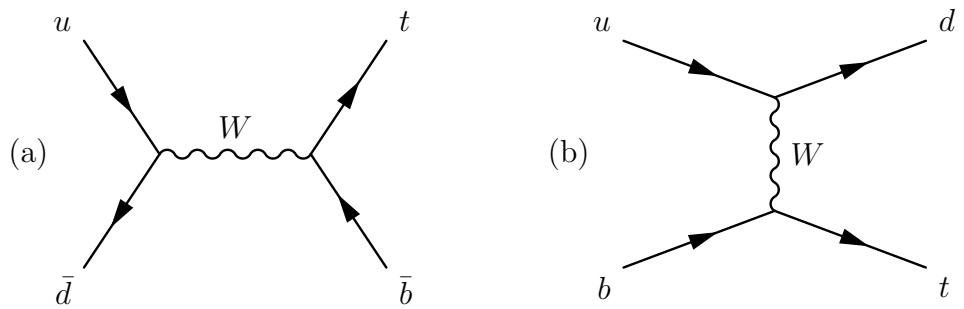


Figure 3.1: LO Feynman diagrams for single top production via the (a) s-channel and (b) t-channel process.

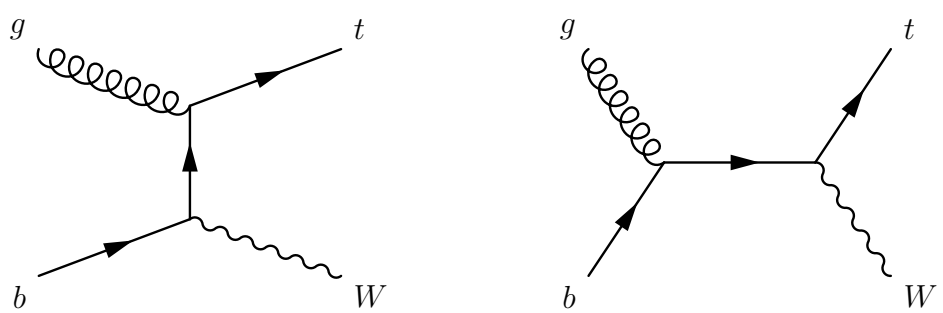


Figure 3.2: LO Feynman diagrams for tW associated production.

The cross sections have been calculated at the tree level [66] and at NLO [72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82], with many of the later ones being differential, providing theoretically-predicted event kinematics. The cross sections given in Table. 3.1 [83, 84] are for the three processes at Run II of the Fermilab Tevatron Collider and at LHC for $m_t = 171.4 \pm 2.1\,\text{GeV}$ at NLO plus NNLO and NNNLO threshold soft-gluon corrections. At the Tevatron the top and antitop rates are the same, so the values

Single Top Process	Cross Section (pb)		
	Run II Tevatron $t+\bar{t}$	LHC t	LHC $\bar{t}$
s-channel	1.08 ± 0.08	$7.80^{+0.70}_{-0.60}$	4.35 ± 0.26
t-channel	2.30 ± 0.14	150 ± 6	92 ± 4
tW associated	0.28 ± 0.06	43.5 ± 4.8	43.5 ± 4.8

Table 3.1: The cross sections for single top and antitop production. At the Tevatron, the antitop cross section is the same as the top cross section. The values are for $m_t = 171.4 \pm 2.1\,\text{GeV}$ at NLO plus NNLO and NNNLO threshold soft-gluon corrections, from [83, 84].

in the second column are twice the single top rates given in the paper. Associated production is small at the Tevatron because of its massive final state. Therefore, this analysis only looks for s-channel and t-channel production. At the LHC, however, tW production will be comparatively large because it involves an initial state gluon, and as can be seen from Fig 2.1, the gluon PDF grows much faster than the quark PDFs at low x. Because of the higher center of mass energy at the LHC, lower x will be needed for single top production. On the other hand, s-channel production will be relatively small at the LHC because the diagrams involve an initial antiquark, and because the LHC is a pp collider, antiquarks will only come from the sea. The Tevatron, being a $p\bar{p}$ collider, has valance antiquarks.

3.1 Single Top Event Signature

According to the SM, the top quark is expected to decay almost exclusively into a b-quark and a W boson. Other decays are suppressed by the small values of $|V_{ts}|$ and $|V_{td}|$ [9]. In the case where the W boson decays hadronically, the signal is swamped by QCD multijet background. Therefore, this analysis searches for events in which

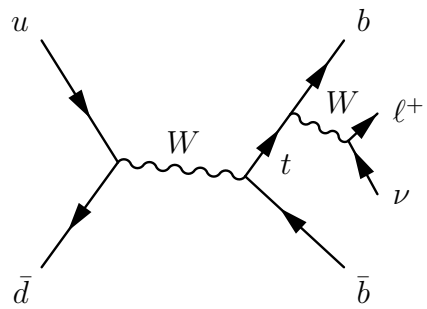


Figure 3.3: LO Feynman diagram for single top production via the s-channel, followed by a leptonic decay of the top quark.

the W boson decays into either an electron or a muon plus a neutrino[1].

This section uses the differential NLO calculations from Refs. [79, 80, 81] to show the theoretical distributions for single top production at Run II of the Fermilab Tevatron Collider.

3.1.1 s-channel

Figure 3.3 shows the representative LO s-channel diagram with the decay explicitly shown. The main components of the final state are the $\bar{b}$ that was created along with the top quark, the b-quark from the top quark decay, the charged lepton that comes from the W boson the top quark decays into, and the neutrino from the same W boson. The neutrino is evident only as $\not{E}_{\mathrm{T}}$ in the detector—it is not directly observed. Examples of NLO corrections to single top production via the s-channel

[1] Cases where the W boson decays into a tau and a neutrino are included as a signal if the tau subsequently decays into a muon or electron, though the efficiency of the multivariate techniques tends to be poorer in such cases. There are a number of reasons for the lower efficiency. First, the resulting electron or muon p_{T} is lower so triggering is less efficient. Second, the extra neutrinos make the tau's energy less well known than, for example, an electron's energy is. Third, at least for the matrix element analysis, the discriminant is built assuming the electron or muon comes directly from the W boson, not from a tau, so the discriminating power is smaller for tau events.

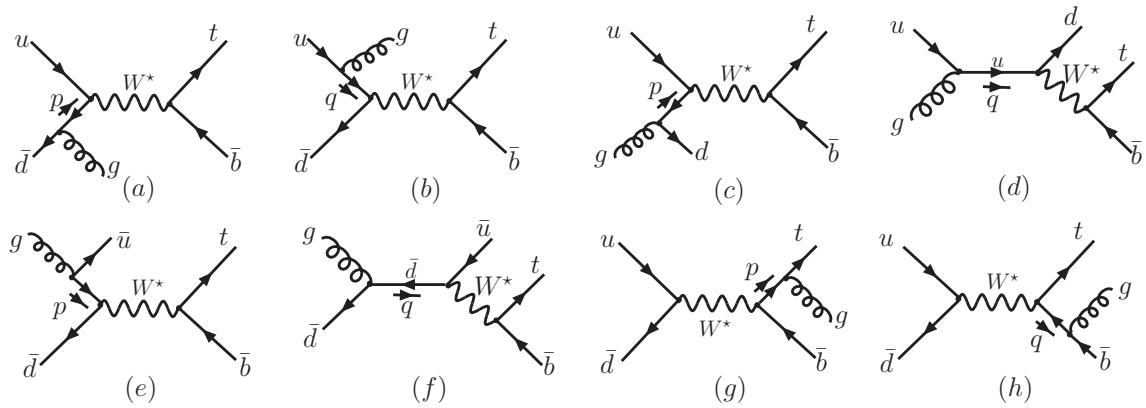


Figure 3.4: Representative diagrams of real emission correction for single top production via the s-channel process. (Diagrams from Ref. [79].)

that include real emission are shown in Fig. 3.4. These diagrams can produce a "third jet" in the event.

Figure 3.5 gives the Born level and NLO p_{T} and η distributions for the jets in s-channel single top production. The term, $\mathrm{O}(\alpha_s)$, refers to the NLO corrections applied to the Born level to get the NLO distributions. These plots are only for t (and not $\bar{t}$) production, which is why there is some asymmetry in the η plots, though not very large. The jets tend to be central, with the third jet least so. The p_{T} distribution of the b-jet that comes from the top quark decay has a higher peak than that of the $\bar{b}$ jet, but the p_{T} distribution of the $\bar{b}$ jet has a longer tail at the high end. The third jet p_{T} tends to be the lowest. Given the $p_{\mathrm{T}_{\mathrm{jet}}} > 15\,\mathrm{GeV}$ and $|\eta| < 3.0$ cuts used in this study, roughly 33% of the s-channel events have a third jet. The legend in Figure 3.5 refers to whether the third jet came from the $Wu\bar{d}$ vertex part of the graph (INIT), the $Wt\bar{b}$ vertex part of the graph (FINAL), or the top quark decay (SDEC).

Figure 3.6 gives the NLO p_{T} and η distributions for the charged lepton, and

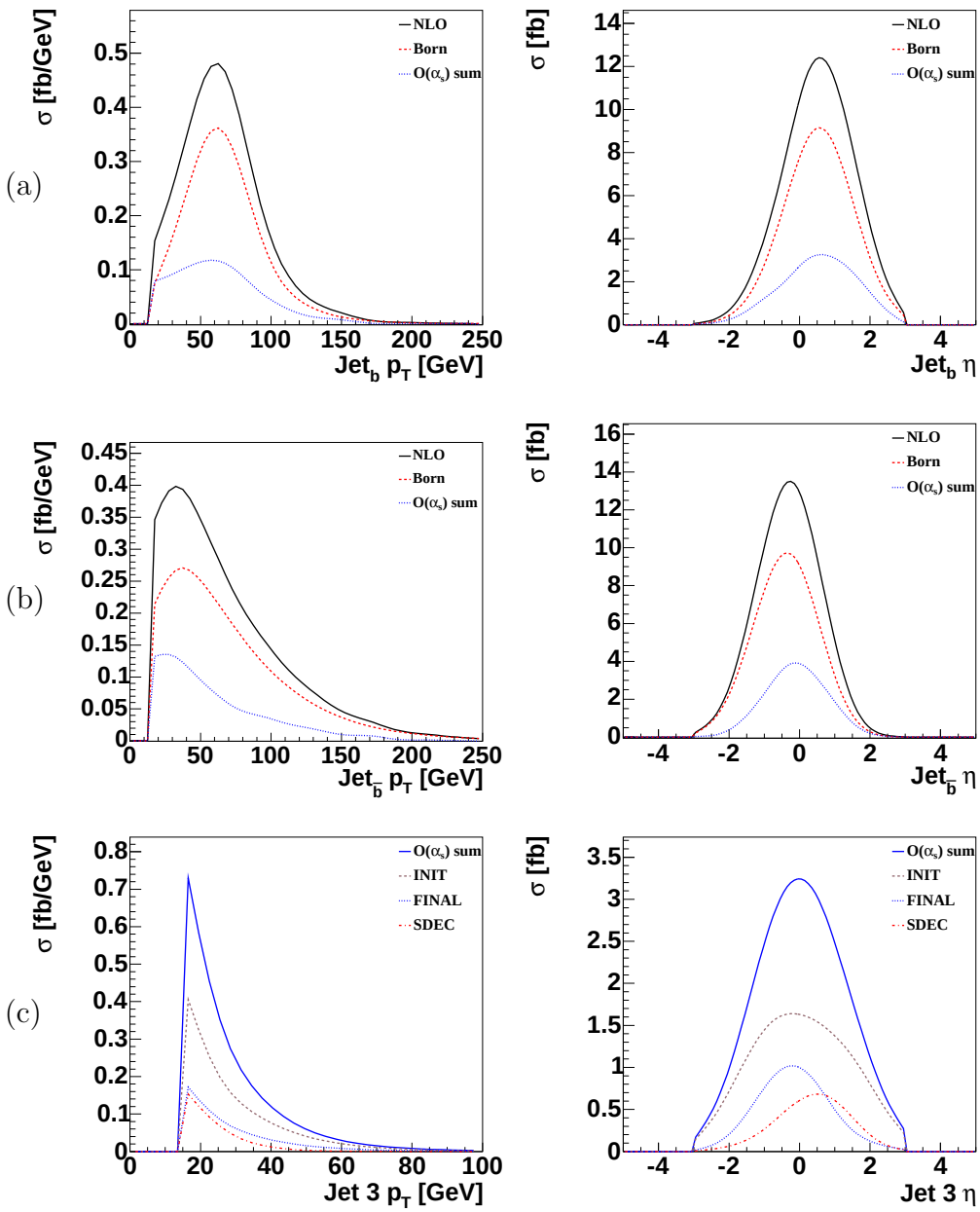


Figure 3.5: The NLO p_{T} and η distributions for the jets in s-channel single top production: (a) the b-jet from the top decay, (b) the $\bar{b}$-jet that gets created along with the top quark, and (c) the "third jet." (Plots from Ref. [80].)

also the p_{T} of the neutrino, which is observed as the $\not{E}_{\mathrm{T}}$, for s-channel single top production. The charged lepton is very central. The neutrino has slightly higher p_{T} on average compared to the charged lepton, which can be explained by the V−A structure of the charged-current interaction, as is explained Sec. 3.2.

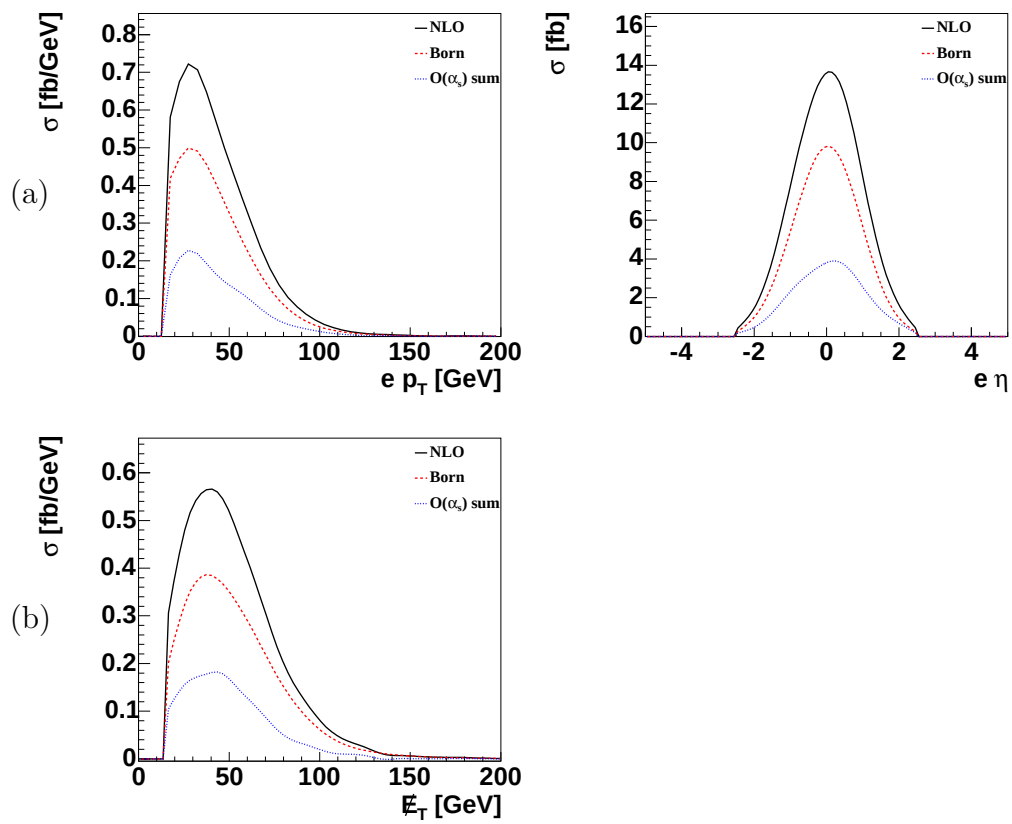


Figure 3.6: (a) The NLO p_{T} and η distributions for the charged lepton in s-channel single top production. (b) The neutrino p_{T}, observed as the event $\not{E}_{\mathrm{T}}$. (Plots from Ref. [80].)

3.1.2 t-channel

Figure 3.7 shows the representative LO t-channel diagram with the decay explicitly

shown. The main components of the final state are the b-quark from the top quark decay, the charged lepton that comes from the W boson the top quark decays into, the neutrino from the same W boson, and the d "spectator" quark. Sample NLO corrections to single top production via the t-channel that include real emission are shown in Fig. 3.8, including the well known W-gluon fusion diagrams, which are graphs (k)-(n) in the figure. Thus, a "third jet" produced by the NLO diagrams can come from the $\bar{b}$-quark from W-gluon fusion diagrams, in addition to gluon and light quark jets.

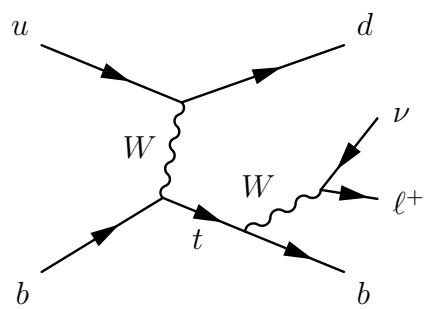


Figure 3.7: LO Feynman diagram for single top production via the t-channel, followed by a leptonic decay of the top quark.

Figure 3.9 gives the NLO p_{T} and η distributions for the jets in t-channel single top production. The b-jet from the top decay tends to have the highest p_{T} and be central. The spectator jet has a lower p_{T} distribution, but its most distinguishing feature is its forward and asymmetric η distribution, which can be used to help discriminate signal from background. These plots are only for t (and not $\bar{t}$) production; adding $\bar{t}$ production to the η plots would make them symmetric. In an experimental setup, the asymmetric plot can be retrieved by plotting "$Q \times \eta$," where Q is the charge of the lepton and η is the pseudorapidity of the leading jet without a b-tag. Given the $p_{\mathrm{T}_{\mathrm{jet}}} > 15\,\mathrm{GeV}$ and $|\eta| < 3.0$ cuts used in this study, roughly 40% of t-channel

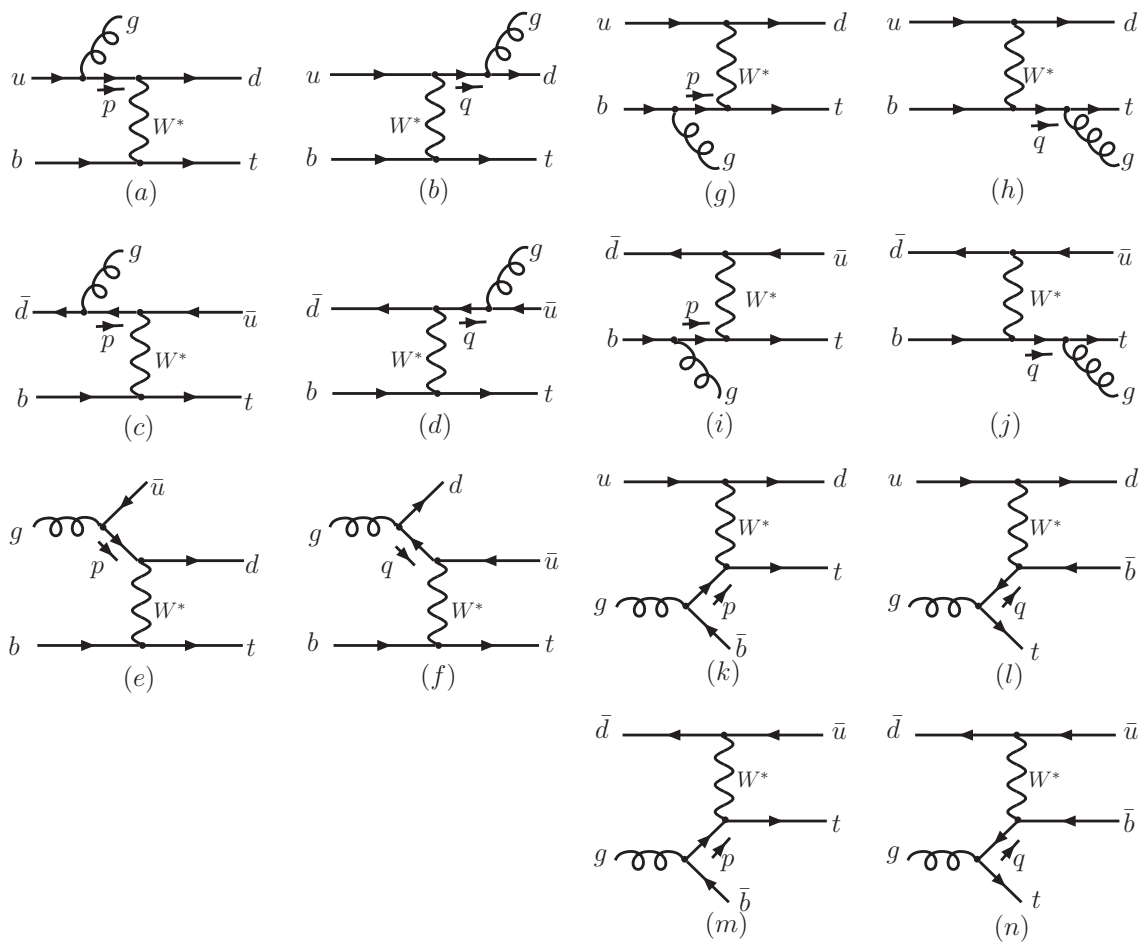


Figure 3.8: Representative diagrams of real emission correction for single top production via the t-channel process. (Diagrams from Ref. [79].)

events have a third jet. The legend in Figure 3.9 refers to whether the third jet came from the heavy quark line (HEAVY), the light quark line (LIGHT), or the top quark decay (TDEC). Note that the HEAVY correction is greater than the LIGHT or TDEC correction, and 80% of the HEAVY corrections include a $\bar{b}$ quark from W-gluon fusion processes. The p_{T} distribution of the third jet is the lowest, and the η distribution, though central, is fairly wide.

Figure 3.10 gives the NLO p_{T} and η distributions for the charged lepton, and also the p_{T} of the neutrino, which is observed as the $\not{E}_{\mathrm{T}}$, for t-channel single top production. As for s-channel production, the charged lepton is very central, and the neutrino has a slightly higher p_{T} on average compared to the charged lepton.

3.2 Polarization Effects

It was concluded before the top quark was discovered that its large mass would result in a decay time too short for hadrons to form [85]. Thus, unlike other quarks, the top quark does not form mesons and baryons, according to the SM. This fact has the advantage that it exposes raw quark properties, undiluted by hadronization.

The V−A weak force produces highly polarized top quarks [66, 86, 87, 88, 89, 80, 81]. The polarization of the top quark becomes evident in the angular correlations between the observed daughter particles. It had been suggested to use these correlations to improve the discovery potential [71, 56], and variables including those correlations have been directly used to set limits on single top production [90], and recently, to find evidence for single top production [91, 92]. This analysis uses those variables indirectly since the matrix elements include a full treatment of spins.

With more statistics, the polarizations and the resulting angular correlations can

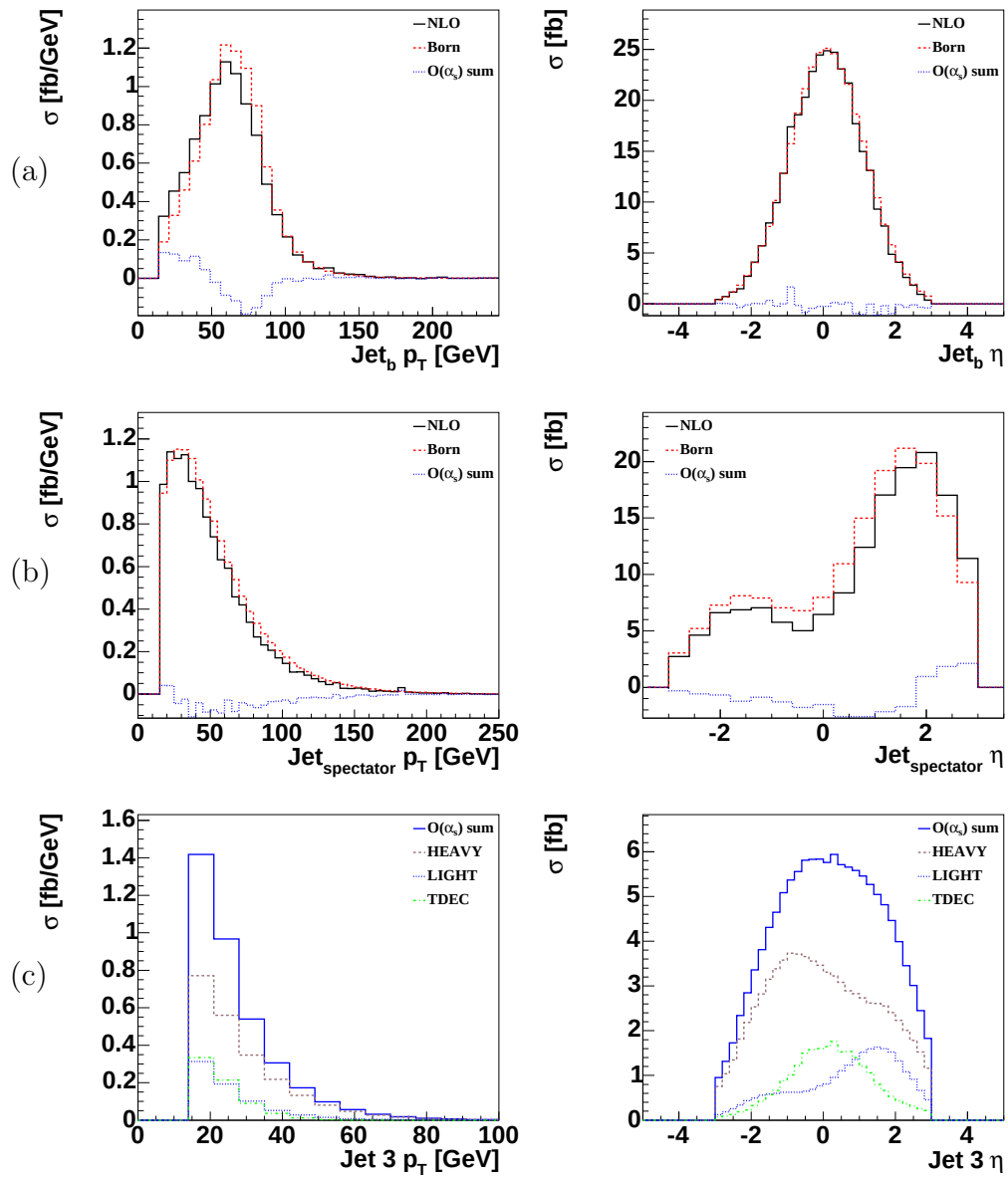


Figure 3.9: The NLO p_{T} and η distributions for the jets in t-channel single top production: (a) the b-jet from the top decay, (b) the "spectator jet," and (c) the "third jet." (Plots from Ref. [81].)

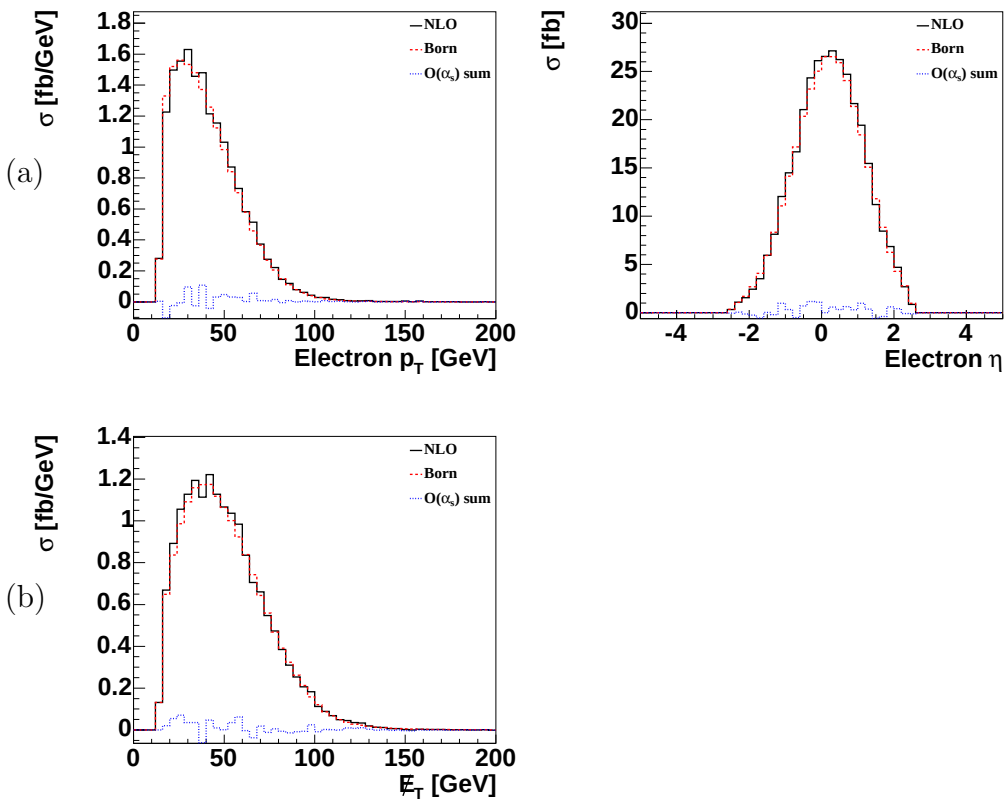


Figure 3.10: (a) The NLO p_{T} and η distributions for the charged lepton in s-channel single top production. (b) The neutrino p_{T}, observed as the event $\not{E}_{\mathrm{T}}$. (Plots from Ref. [81].)

be used to study the V−A form of the weak interaction. Beyond the SM scenarios that include right-handed couplings or scalar particles, such as charged technipions or Higgs particles, can produce different signatures [93].

To understand why the top quark is polarized in single top production, it is helpful to look at the decays of a polarized top quark in its rest frame. Being a heavy spin-1 particle, the W boson has three helicity states: left-handed, longitudinal, and right-handed. However, as shown in Ref. [94], in the top quark rest frame, W bosons with right handed polarization are heavily suppressed by the small mass of the b-quark relative to both the mass of the W boson and the mass of the top quark, so they can be ignored. In that approximation, the b-quark from a t-quark decay is always left-handed and the $\bar{b}$-quark from a $\bar{t}$-quark decay is always right-handed because of the V−A nature of the weak interaction.

Figure 3.11(a) shows the decay of a top quark into a longitudinally polarized W^+ boson and a b-quark. To conserve angular momentum, the b-quark, being left-handed, needs to go in the direction opposite to the spin of the top quark, so the W boson goes in the direction of the top quark spin. Note that the positron from the W boson's decay has spin aligned with the top quark and moves in the same direction as the top quark's spin. Figure 3.11(b) shows the decay of a top quark into a left-handed W^+ boson and a b-quark. To conserve angular momentum, the b-quark, being left-handed, needs to go in the direction of the spin of the top quark, so the W boson goes in the direction opposite of the top quark spin. Note that the positron, to conserve angular momentum in the W boson decay, again has its spin aligned with the top quark and moves in the same direction as the top quark's spin. Thus the top quark's spin is most strongly correlated with that of the positron, not its direct decay products. Figure 3.11(b) also explains why the $\not{E}_{\mathrm{T}}$ values on average tend to

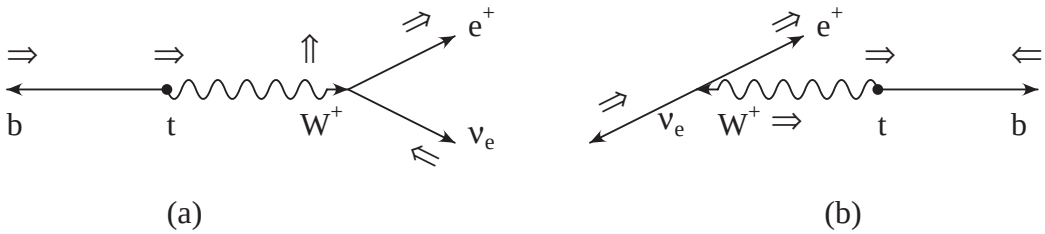


Figure 3.11: Correlation between the spin of the top quark and the direction of the charged lepton in the top quark rest frame for (a) a t-quark decaying into a W boson with longitudinal polarization and (b) a t-quark decaying into a W boson with left-handed polarization. (Figure from Ref. [93].)

be higher than the charged lepton p_{T} values: the neutrino in that diagram needs to have enough p_{T} to balance both the positron and the b quark.

There is an analogous correlation between $\bar{t}$ quarks and electrons from their decays. The spin of the charged lepton, therefore, is always aligned with that of the t or $\bar{t}$-quark, and the direction of the positron (electron) tends to be along (against) the spin of the t- ($\bar{t}$)-quark. As pointed out in Ref. [89], single top production can be viewed as a top quark decay going backwards in time. Relating the down quark to the electron because of its analogous role in the SU(2) electroweak doublet, one infers that the polarization with the top quark is most correlated with the polarization of the down quark. If one takes the spin basis to be the direction of the down quark, the correlation is 100% [86].

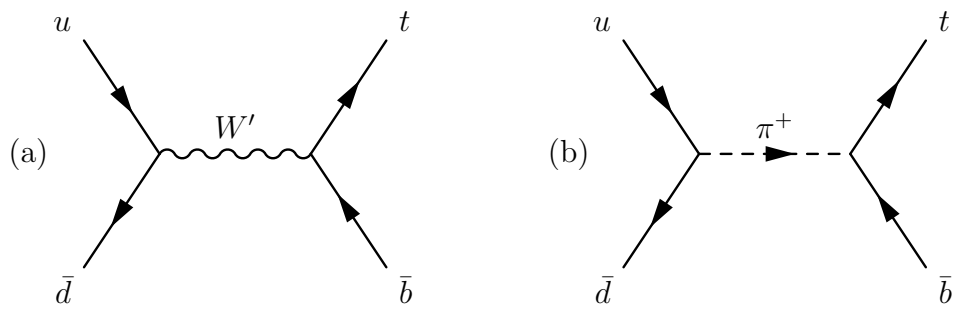


Figure 3.12: Feynman diagrams for anomalous single top production via s-channel processes with a (a) W' vector boson or a (b) top-pion as predicted by topcolor.

3.3 Single Top and the Search for Physics Beyond the Standard Model

Beyond the standard model (BSM) interactions can affect the cross sections, and interestingly, the three processes, s-channel, t-channel, and tW are differently sensitive to new effects [93]. The s-channel is most sensitive to new charged bosons, which can contribute to the production cross section in diagrams analogous to the SM one but with the W boson replaced by new particles. Representative figures are shown in Fig 3.12. Topflavor theories predict new gauge bosons [50, 51, 52], including a heavy W' that can contribute to s-channel single top production. Topcolor theories predict new scalar charged bosons, such as a top pion that couples to top and bottom quarks [45, 46, 47, 48, 49]. A charged Higgs particle from MSSM theories could play a similar role. Because there are additional diagrams, the s-channel cross section is typically enhanced, especially so if the boson can be produced on-shell. Figure 3.13 shows examples of how those theories would affect the s-channel cross section. In that figure, the top-flavor model is barely distinguishable from the SM because of the large mass of the gauge boson, but at the LHC, the large mass would not present a

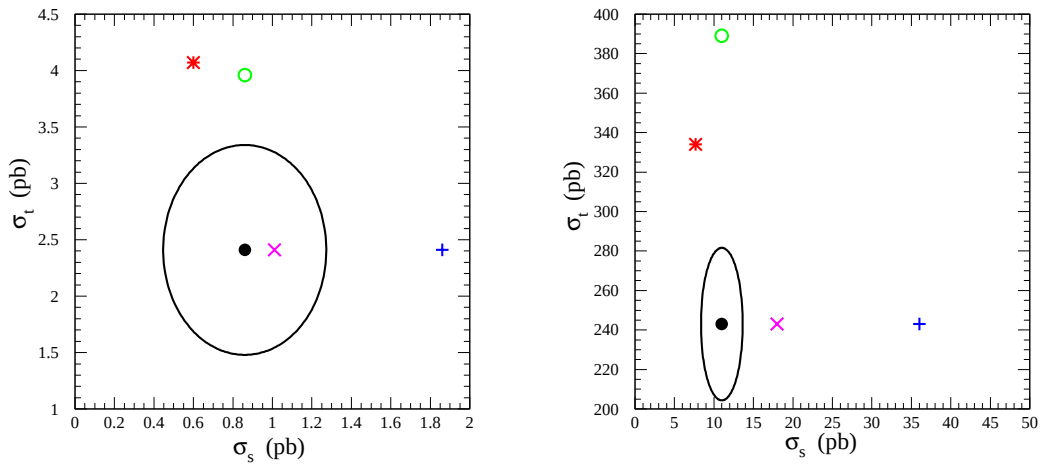


Figure 3.13: Examples of BSM effects to the single top cross section at the Tevatron Collider (left) and at LHC (right). The SM single top cross sections are represented by the black solid circle, with the oval line showing the 3σ theoretical uncertainty. The purple X shows how a top-flavor model ($M_{Z'} = 1\,\text{TeV}$, $\sin^2\phi = 0.05$) could affect the cross section, and the blue cross is how a charged top-pion ($m_\pi = 250\,\text{GeV}$ left, $450\,\text{GeV}$ right, and t_R-c_R mixing of $\sim 20\%$) could effect the cross section. Note that both are extra particles and thus affect the s-channel cross section. The green open circle shows how a FCNC Ztc vertex could affect the cross sections, and the red star shows the effect of fourth quark generation. The figure is from Ref. [93].

problem. On the other hand, because the additional bosons are heavy, their contribution would be suppressed in the t-channel by $1/M_{W'}^2$, where $M_{W'}$ is the mass of the boson.

The t-channel cross section is very sensitive to extra interactions. For example, flavor changing neutral current (FCNC) interactions, shown in Fig. 3.14 would increase the t-channel cross section because then additional diagrams of the type $qc \to qt$ would contribute, and the c-quark parton density in the proton is larger than the b-quark parton density, compensating for the presumably weaker interaction. The

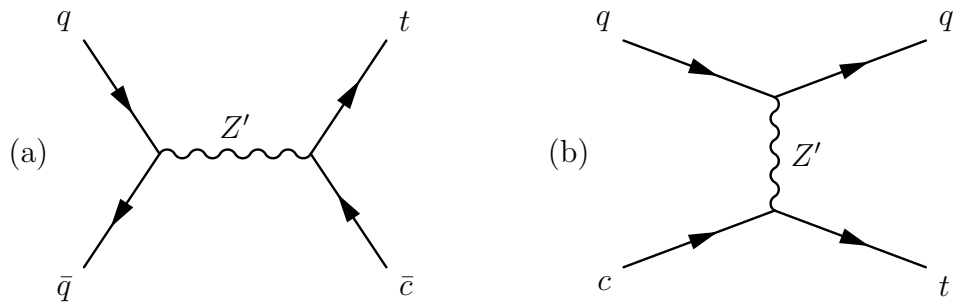


Figure 3.14: FCNC Feynman diagrams for single top production via the (a) s-channel and (b) t-channel process.

s-channel interaction would also experience an enhancement, but it would probably be negligible due to the FCNC interaction's weakness, and additionally there would be one fewer b-quark to tag in the final state, likely hurting the event selection efficiency. The t-channel cross section would also experience a similar enhancement from the PDFs if $|V_{ts}|$ is larger than predicted by the three-generational unitarity in the SM. The s-channel cross section would not have any enhancement, and could suffer because of lower b-tagging efficiencies. Figure 3.13 also shows examples how those theories would affect the t-channel cross section.

In summary, the s-channel and the t-channel cross sections have different sensitivities, and hence it would be useful to separately measure their cross sections, Associated tW production, because of the real W boson in the final state, is generally not sensitive to extra bosons, so it can further help untangle the results at the LHC.

BSM effects can also be observed in the polarizations of the top quark [93]. If the interactions have a V+A component, or if the interactions are mediated by a spin-0 boson, the angular correlations would change. For example, in the helicity basis, top quarks produced by top-pions would show close to 100% right-handed polarization,

while the SM predicts 70% left-handed top quarks in the s-channel [81].

Chapter 4

Experimental Apparatus

4.1 Accelerator Chain

The Fermilab Tevatron Collider is the highest energy collider currently in operation, with a center of mass energy of 1.96 TeV [95, 96, 97]. It was the world's first large scale superconducting synchrotron when it was commissioned in 1983, in proton-only fixed target mode. Currently the Tevatron Collider is used as a proton-antiproton ($p\bar{p}$) collider, with beams of protons and antiprotons traveling in opposite directions around a circular ring of radius 1 km. The current beam structure is "36 on 36," meaning that there are 36 bunches of protons colliding with 36 bunches of antiprotons. The bunches are arranged into three "superbunches" of 12 bunches each. Within a superbunch, the bunches are spaced 396 ns apart. Collisions can occur at two points on the ring: where the CDF detector is located and where the D0 detector is located. The beam spot at the D0 interaction point is 30 μm across, and the length of the interaction area along the beam is distributed according to a Gaussian distribution with $\sigma \approx 25$ cm. Current initial instantaneous luminosities are over $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$,

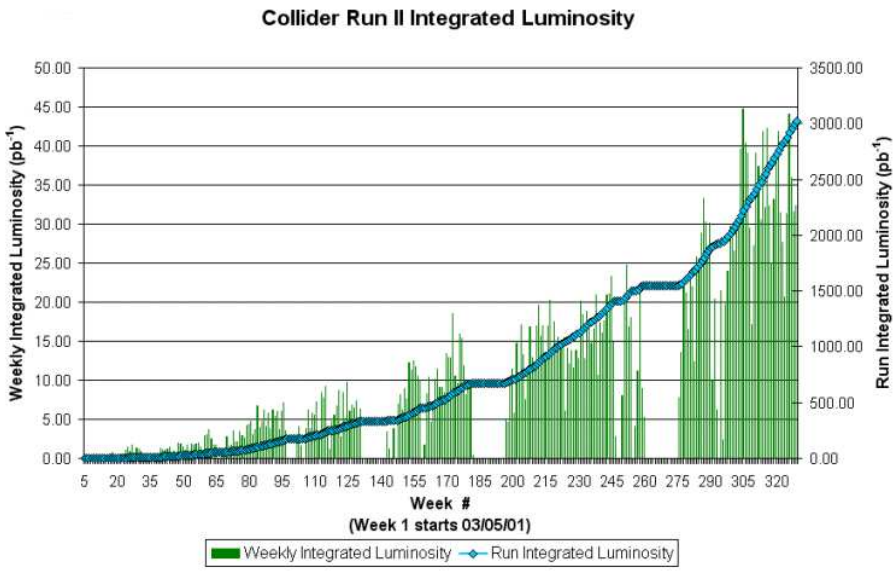


Figure 4.1: Integrated luminosity delivered by the Tevatron Collider.

and Fig. 4.1 shows the weekly and total integrated luminosities delivered from March 2001: the Run II running period. At the current luminosities, multiple collisions often occur per bunch crossing. In the time period used for this analysis, August 2002 to December 2005, most often just one primary vertex was reconstructed in an event, but the average number of primary vertices reconstructed was 1.9[1].

There are many steps involved to get protons and antiprotons circling in the Tevatron Collider with an energy of 0.98 TeV. Figure 4.2 shows the chain of accelerators involved in the process. The beam starts at the negative hydrogen ion (H^-) source, a magnetron surface-plasma source [95, 99]. The magnetron source consists of an oval cathode surrounded by an anode inside a magnetic field. Hydrogen gas (H_2) is sent into the chamber, and as a result of the electric and magnetic field, a

[1]This was measured in $Z \to e^+e^-$ events.

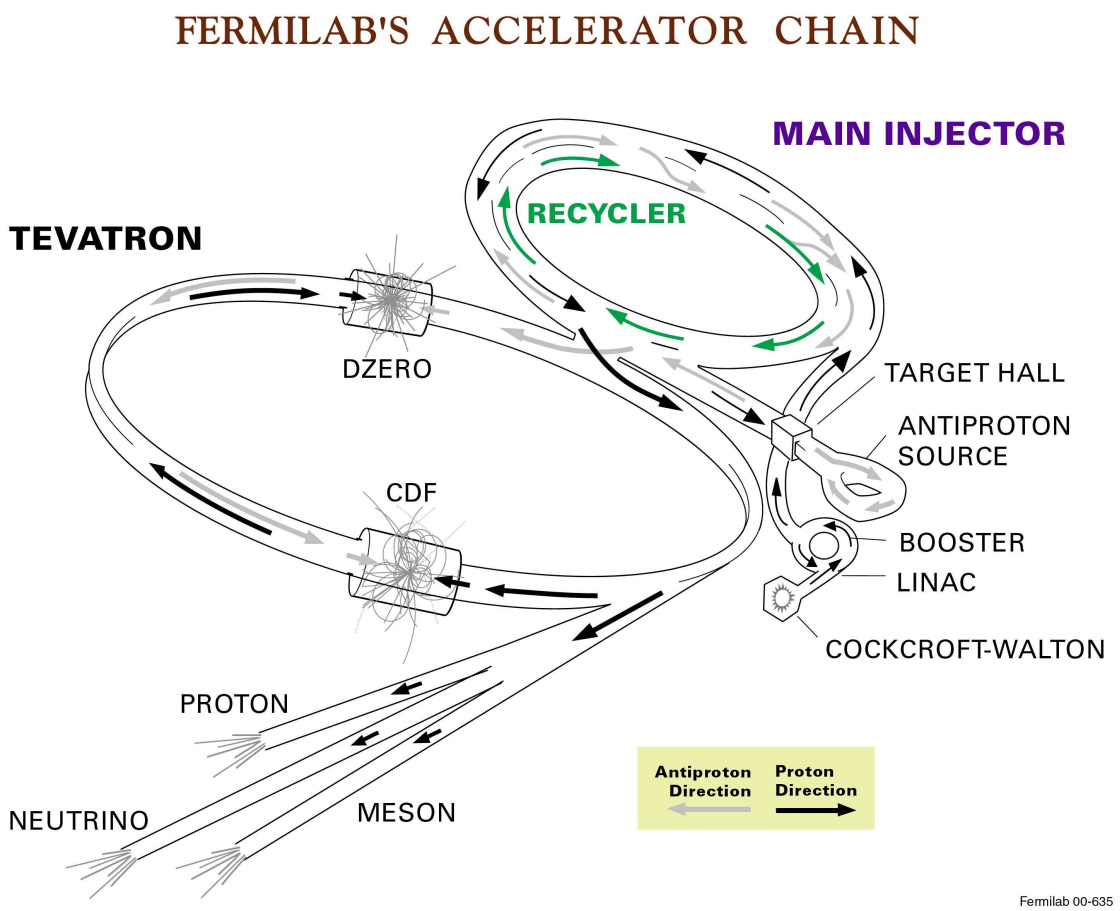


Figure 4.2: Fermilab's accelerator chain [98].

plasma is formed. When H^+ ions strike the cathode, they reflect or desorb hydrogen atoms that have been absorbed on the surface. Some of these are H^- ions. In order to increase the rate at which H^- ions are created, the cathode is coated with Cesium, which reduces the work function for removing an extra electron. The H^- ions are accelerated by the potential difference between the anode and cathode and leave through an aperture in the cathode with 18 keV of energy. The process is pulsed at 15 Hz.

The magnetron source is located within a dome kept at -750 kV by a Cockcroft-Walton generator [95, 100]. The Cockcroft-Walton generator uses a 75 keV AC input voltage (from a transformer) to charge capacitors in parallel. With the proper placement of diodes, the voltage from the capacitors is applied in series, thus multiplying its value. In this way, the H^- ions are accelerated to 750 keV, at which point they enter the Linac [100, 101, 102] for further acceleration.

The Linac is a two stage linear accelerator, whose purpose is to accelerate the 750 keV H^- ions up to 400 MeV. The first stage, which was a part of the original 200 MeV Linac built in 1971, is an Alvarez drift-tube accelerator. It accelerates the ions to 116 MeV. The second stage, a more modern, side-coupled accelerator that replaced the high energy portion of the original accelerator in 1993, accelerates the ions up to 400 MeV. The reason for the increased Linac energy was to decrease beam losses in the Booster [95, 103], the next step in the acceleration chain.

The purpose of the Booster, a fast cycling proton synchrotron with a 151 m radius, is to accept the 400 MeV H^- ions and output 8 GeV protons. Injection is done via a multi-turn charge-exchange injection [95, 104], which is the reason why H^- ions are used in the beginning of the accelerator chain. After debunching the H^- ions from the Linac to minimize their momentum spread, the H^- beam is merged with

the proton beam already in the Booster over multiple turns, and the electrons are stripped by a carbon foil.

After the Booster in the accelerator chain is the Main Injector [105]. It is also a synchrotron, but much larger, with a circumference of 3320 m. It was commissioned in 1998 as a replacement for the Main Ring. With regard to the running of the Tevatron Collider, the Main Injector has two duties: accelerating protons up to 120 GeV to send to the Antiproton Source [95, 96, 106], and accelerating protons or antiprotons to 150 GeV to inject into the Tevatron.

To make antiprotons, the Main Injector takes 8 GeV protons from the Booster, accelerates them to 120 GeV, and sends them to a nickel target. The target is followed by a lithium lens to focus the secondary particles, and then a dipole magnet to select 8 GeV antiprotons from the beam. These are then sent to the Debuncher, a triangular 8 GeV synchrotron, to reduce the momentum spread and to cool them. From there, the antiprotons are "stacked" in the Accumulator, another 8 GeV synchrotron.

The limiting factor to the luminosity of the Tevatron has been the number of antiprotons available. Therefore, a recent addition has been the Recycler [107], which is an 8 GeV storage ring with permanent magnets that shares the same tunnel as the Main Injector. When the Accumulator becomes full, stacking efficiency decreases, so the antiprotons are transferred to the Recycler, and stacking efficiency is improved. Additionally, electron cooling was recently accomplished to further improve the antiproton beam and potentially significantly improve the luminosity [108].

The Main Injector can take protons from the Booster accelerate them to 150 GeV, and inject them into the Tevatron. It can also take antiprotons from the Accumulator or the Recycler and do the same. The Tevatron then accelerates beams to 0.98 TeV, initializes collisions, and allows the beams to collide for the many hours of the "store."

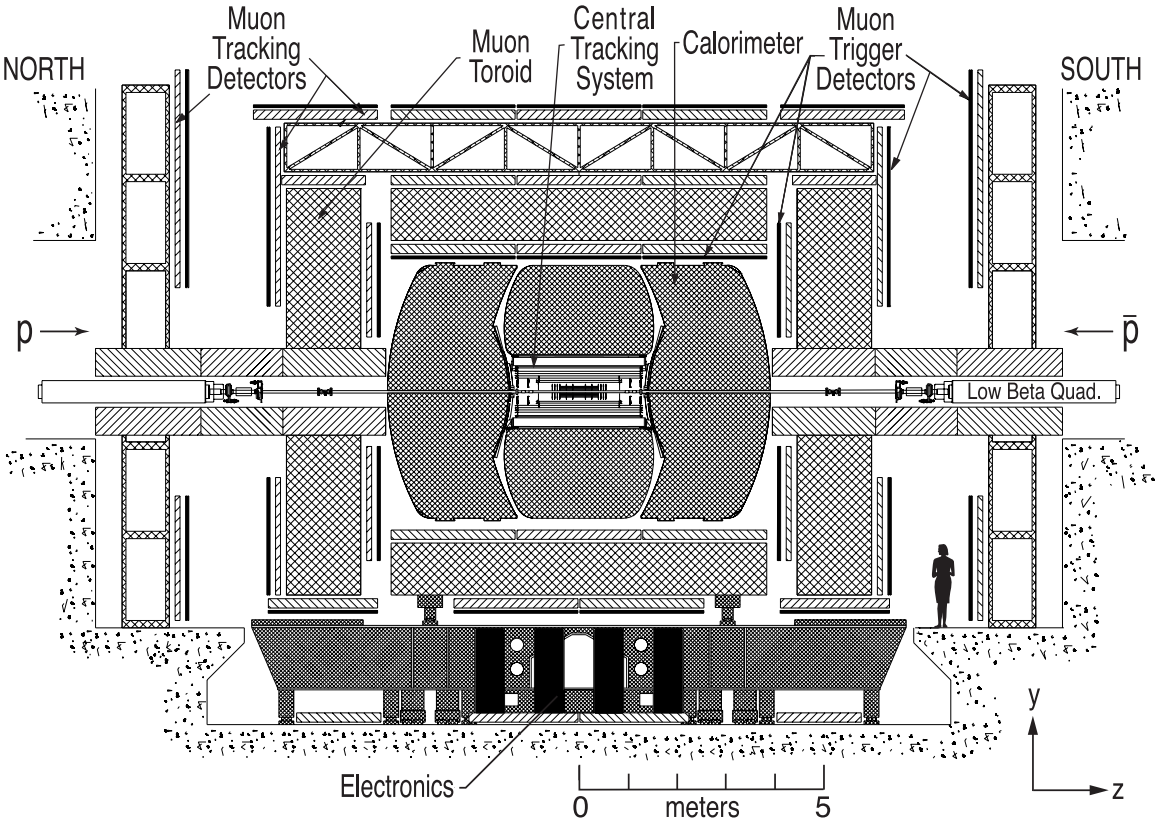


Figure 4.3: Schematic of the D0 detector [110].

4.2 The D0 Detector

At one of the collision points of the Tevatron Collider sits the D0 detector [109, 110], a general purpose detector built to study $p\bar{p}$ collisions, with an emphasis on high-mass states and large p_{T} phenomena. The detector was commissioned in 1992 as part of Run I of the Tevatron, and it was heavily modified for the start of Run II in 2001. What is described here is the current, Run II version of the D0 detector.

Figure 4.3 shows a diagram of the D0 detector. Protons enter from the north side, antiprotons from the south, colliding in the center inside a beryllium beam pipe, which has a wall thickness of 0.508 mm and an outer diameter of 38.1 mm. Closest to the col-

lision point is the central tracking system, consisting of the silicon microstrip tracker (SMT) and the central fiber tracker (CFT), all inside a superconducting solenoidal magnet. Outside of the magnet are the preshower detectors, to compensate for the material in the tracking system, followed by the calorimeters. Finally, outside of the calorimeters is the muon system, with its toroidal magnets. These systems are explained in more detail below.

A right-handed coordinate system is used for the detector and data analysis. As shown in Fig. 4.3, the positive z direction is chosen to be the direction that protons travel, and the y direction is chosen to be up. The r coordinate is defined in the cylindrical coordinate system sense: $r = \sqrt{x^2 + y^2}$. The polar and azimuthal angles are written as θ and ϕ, with the standard definitions. The origin can be chosen to be in the middle of the detector or at the interaction vertex. When the origin is set in the middle of the detector, then the values are called *detector* values; when the origin is set at the interaction vertex, then the values are called *physics* values. If not specified, the values are physics values when discussing physics, detector values when discussing the detectors. For θ and derived values the distinction can be large because, due to the length of the bunches, the interaction vertices are distributed in the z coordinate with $\sigma_z \approx 25\,\mathrm{cm}$. Because $\phi^{\mathrm{det}} \approx \phi^{\mathrm{phys}}$, the distinction is often not made for ϕ.

The pseudorapidity, η, is defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{4.1}$$

which approximates the true rapidity, y (not to be confused with the coordinate y):

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{4.2}$$

for relativistic particles. Differences of rapidity are invariant to boosts parallel to the beam line. Distinction is often made between η^{det} and η^{phys}, as discussed above. ΔR, unless otherwise specified, is defined as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. (In jet algorithms ΔR is sometimes defined with rapidity instead of pseudorapidity.) "Transverse" values, such as p_{T}, the transverse momentum, unless otherwise stated, are transverse relative to the beam line. Transverse values are often used because the proton is a composite particle, so the longitudinal energy of the collision is not known.

4.2.1 Triggering and Data Acquisition

The collision rate at the D0 detector is 1.7 MHz, but only around 50 Hz of events can be written to tape. Most of the collisions, however, have only soft interactions, which we also call minimum bias interactions. We are interested, on the other hand, in the less frequent hard interactions, and the purpose of the trigger is to select events that have interesting, hard interactions.

The D0 detector uses three layers of triggering. Figure 4.4 shows an overview of the trigger and data acquisition (DAQ) system, and Fig. 4.5 shows the individual components of the first two trigger levels. The Level 1 (L1) trigger system must reduce the incoming rate of 1.7 MHz down to 2 kHz, with at most 3.3 μs of processing time per event. It is a hardware-based system because of the high input data rate it must handle and the tight timing required. The L1 accept rate is low compared to similar detectors, such as CDF. The reason is that reading out the trackers, as would happen on an L1 accept, introduces dead time, so the accept rate needs to be kept low to minimize the dead time.

The main trigger systems used for this analysis are the L1CAL, the calorimeter trigger, and the L1MUO, the muon trigger. They are described in more detail in the

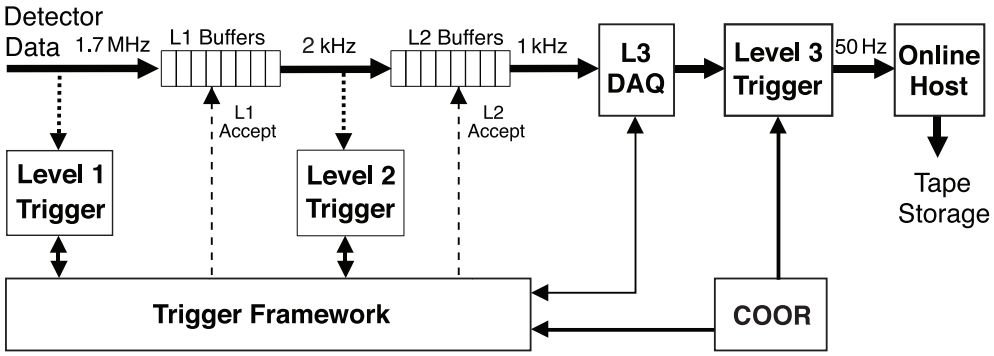


Figure 4.4: Overview of the trigger and DAQ systems [110].

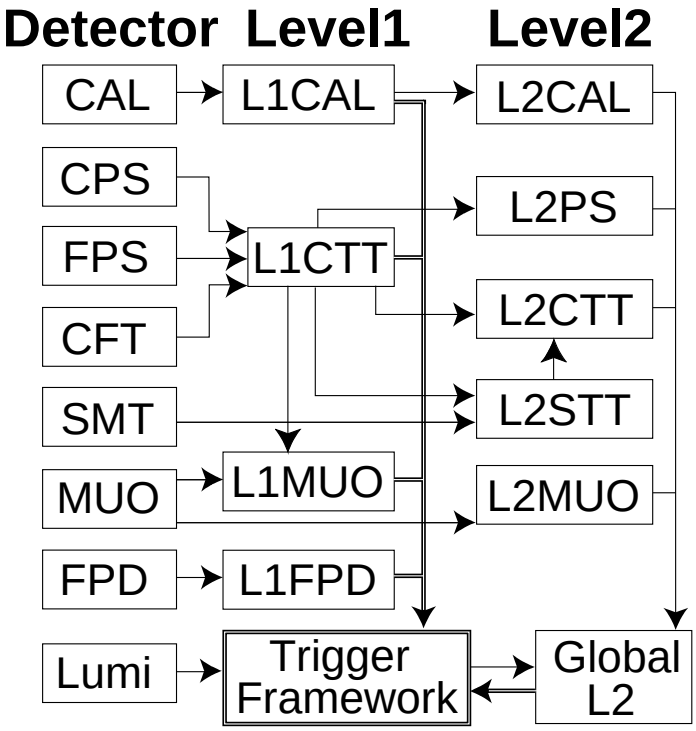


Figure 4.5: Overview of the L1 and L2 trigger systems [110].

calorimeter and muon system sections that follow. L1MUO also takes tracks as input from L1CTT, the track trigger.

The Level 2 (L2) trigger system reduces the rate by a factor of two. It is based on both special hardware and embedded microcontrollers. Since the L2 trigger system has more time to spend than the L1 trigger on the events, it can implement more sophisticated algorithms and make more correlations between the detectors; nevertheless, its time is still short, and it still makes do with the more limited trigger versions of some data inputs, such as the $\Delta\eta \times \Delta\phi \approx 0.2 \times 0.2$ trigger towers from the calorimeter or axial-only fiber hits from the CFT.

For this analysis, the only L2 triggers used are again the calorimeter trigger, L2CAL, and the muon trigger, L2MUO. They are described in their respective detector system sections.

The Level 3 (L3) trigger system reduces the accept rate to 50 Hz, which is written to tape. Currently extra capacity has been added and the accept rate can increase to 100 Hz during high instantaneous luminosity times at the beginning of a store. The L3 trigger is software-based, running on a computer farm, with access to the full information of the events and does partial reconstruction of the events with algorithms similar to those used offline.

4.2.2 Central Tracking

The central tracking system, consisting of the silicon microstrip tracker (SMT) [111] and the central fiber tracker (CFT) [112], all inside a 2 T superconducting solenoidal magnet [113], was newly designed for Run II. Figure 4.6 shows the location of the components.

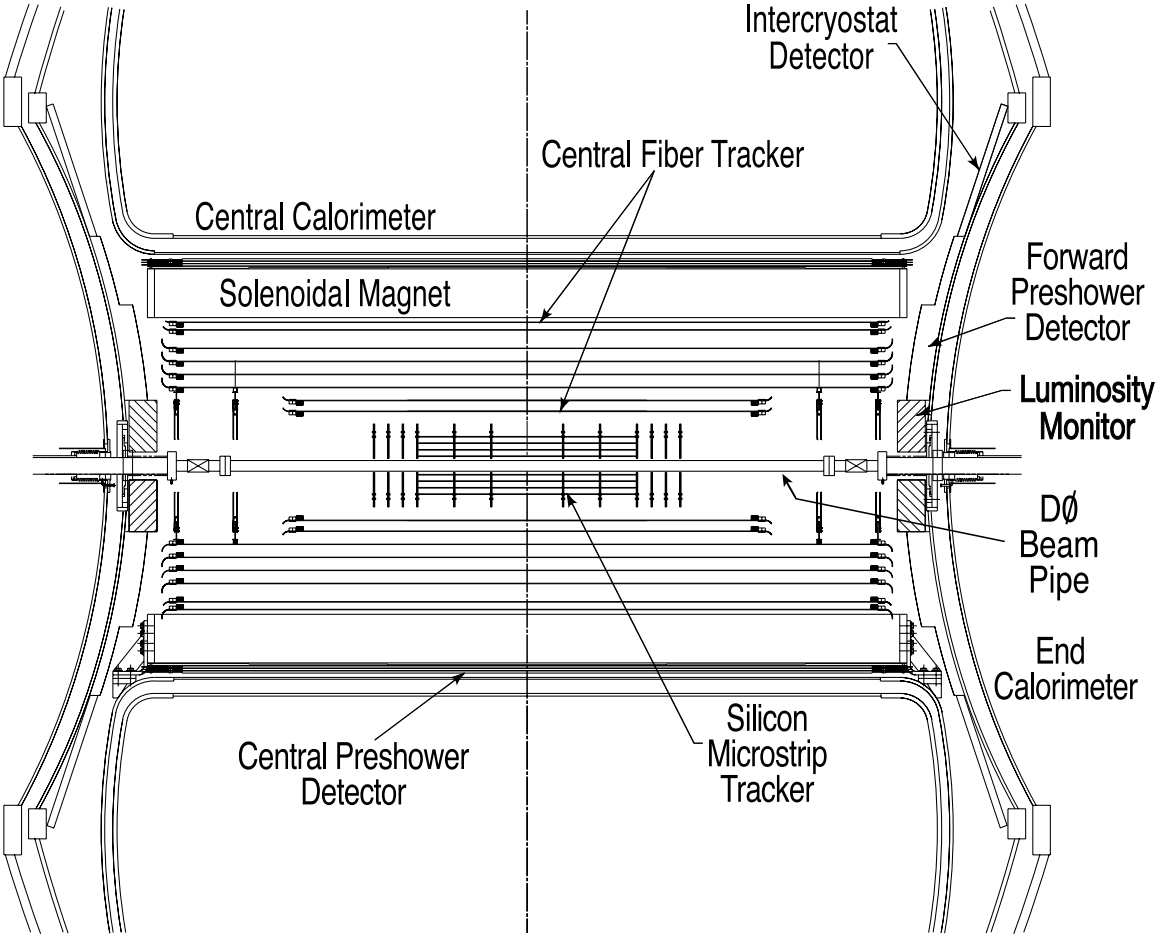


Figure 4.6: Schematic of the inner tracker [110].

The purpose of the central tracking system is to find tracks of charged particles. These tracks can then be used to measure the p_{T} of the particle that made the track, to find the interaction vertex, and for b-tagging. Furthermore, the tracks and vertices can be combined with information from other detectors for particle identification. Primary interaction vertices can be found with a resolution of about $35\,\mu$m along the z direction. For tracks with $\eta = 0$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$, the impact parameter resolution in the r-ϕ plane is better than $15\,\mu$m.

Silicon Microstrip Tracker

The SMT uses silicon microstrip sensors to provide the precise tracking close to the interaction point necessary for vertex finding and b-tagging. The main idea behind how silicon sensors work is that charged particles going through silicon ionize the material, creating electron-hole pairs. By appropriately biasing the sensor, which in effect is reverse-biasing a pn-junction, a large depletion zone is created, so called because it is an area depleted of free charge carriers such as electrons and holes. Furthermore, the depletion zone has an electric field across it, resulting in any electron-hole pairs created in the depletion zone to be quickly being swept out in opposite directions. Thus, if an ionizing particle goes through the depletion zone, the electron-hole pairs are quickly collected, providing a signal that can be read out from the electrodes. The probes collecting the charges can be made to be narrow strips, microstrips, in order to be able to measure spatially where the ionizing particle passed.

The SMT uses various types of sensors, some single sided (SS), with microstrips on one side only, and some double sided (DS), with microstrips on both sides. On DS detectors, the strips on each side can be at an angle relative to each other in order to be able to localize the path of the ionizing particle in more than one direction. However,

Module	Type	Layer	Pitch (μm)	Angle (°)	Length (cm)
F-disk	DS	–	p: 50, n: 62.5	30	7.93
H-disk	SS	–	40, 80 readout	15[a]	7.63 inner
		–			6.33 outer
Central Barrels (4)	DSDM[b]	1,3	p: 50, n: 153.5	90	12.0
	DS	2,4	p: 50, n: 62.5	2	6.0×2[c]
Outer Barrels (2)	SS	1,3	50	–	6.0×2
	DS	2,4	p: 50, n: 62.5	2	6.0×2

Table 4.1: SMT sensor specifications.

[a]two SS detectors mounted back to back
[b]Double Sided, Double Metal
[c]two 6 cm sensors bonded together

that does lead to more complicated fabrication and lower yield, and possibly higher sensitivity to radiation damage, which is why the new "Layer 0" upgrade to the SMT [114], which was installed after we finished taking data for this analysis, uses only SS sensors. Specifications of the sensors used are given in Table 4.1, including the pitch between the strips and the angle between the strips.

The fact that the interaction points are distributed over a relatively large distance along the beam line ($\sigma_z \approx 25$ cm) complicates the design of the SMT, which depends on tracks going across the sensors relatively perpendicularly. A design was chosen with sensors on barrel modules interspersed with sensors on disks in the central area and disks in the forward regions. Figure 4.7 shows a three-dimensional perspective of the SMT, and Fig. 4.6 is useful for seeing the spatial relations of the SMT modules to each other and to the CFT. There are six barrel detectors with centers at $|z| = 6.2$, 19.0, and 31.8 cm; twelve F-disks at $|z| = 12.5$, 25.3, 38.2, 43.1, 48.1, and 53.1 cm; and four H-disks at $|z| = 100.4$ and 121.0 cm. The barrel detectors extend in the radial direction from 2.7 cm to 7.6 cm, the F-disks from 2.6 cm to 10.0 cm, and the

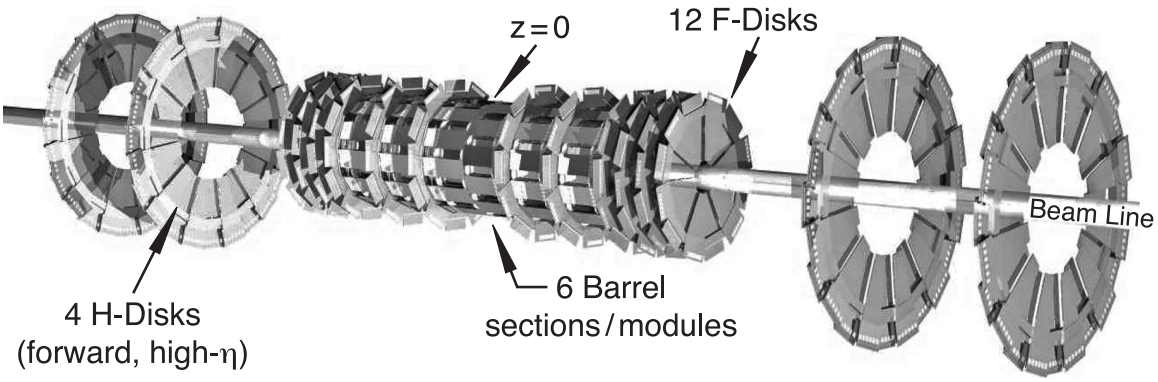


Figure 4.7: The silicon microstrip tracker [110].

H-disks from 9.5 cm to 26 cm. The barrel detectors measure primarily r-ϕ. The strips in the modules run in the axial direction, and if DS, at $\pm 2°$ or $90°$ relative to the axial direction. The disks primarily measure r-z, and also r-ϕ. The detector covers $|\eta| < 3$. For the barrel modules, the axial hit resolution is around $10\,\mu$m, and the z hit resolution is approximately $35\,\mu$m for the $90°$ stereo, $450\,\mu$m for the $2°$ stereo.

In total there are 793k channels in the SMT, which are read out using 128-channel SVXIIe chips. These chips include a preamplifier, a 32-cell analog pipeline, and an analog to digital converter (ADC) with a sparse readout.

The SMT is not used for Level 1 triggering because it would take too long to read out the channels. It is used starting with Level 2, in the silicon track trigger (STT), which can measure the impact parameter of tracks and thus determine which ones are likely to have come from b-quarks.

Central Fiber Tracker

Surrounding the SMT is the CFT. It is made up of scintillating fibers on eight concentric cylinders from 20 to 52 cm from the beam pipe. Each cylinder carries two

doublet layers of fiber, one layer pointing in the radial direction, and the other layer pointing at a stereo angle of $\pm 3°$, with the signs alternating between the cylinders. The doublets on the innermost cylinder have 2560 fibers each, for a total of 5120 fibers, and the doublets on the outermost cylinder have 6500 fibers each, for a total of 13 000 fibers. The inner two cylinders are 1.67 m long, and the other six are 2.52 m long. The coverage of the outer layer is $|\eta| < 1.7$. The fiber is only 835 μm in diameter, resulting in an inherent doublet resolution of about 100 μm.

The fibers are made of polystyrene (PS), doped 1% by weight with the organic fluorescent dye paraterphenyl (pT). Ionizing particles generally excite the PS, which transfers the excitations to the pT, which in a few nanoseconds has a fluorescent decay, giving off light with a wavelength of 340 nm. Light at that wavelength, however, has a mean free path of only a few hundred microns, so additionally, 3-hydroxyflavone (3HF) is added at 1500 ppm as a wavelength shifter. The final radiation has a wavelength of 530 nm. The fibers have two claddings to increase the attenuation length, and are connected on one end to clear fiber optical waveguides between 7.8 m and 11.9 m long, which take the signal to visible light photon counters (VLPCs). The other end is mirrored. The VLPCs are impurity-band silicon avalanche photodetectors, operating at 9 K, with a quantum efficiency greater than or equal to 75%, a gain of 22 000–65 000. They are capable of detecting single photons.

One important feature of the CFT is its fast speed. Signals from the axial doublets are used for Level 1 triggering in the central track trigger (L1CTT). The L1CTT also sends the tracks it finds to the L1MUO system for matching to muon candidate and the L2STT for finding tracks using the SMT. The CFT is used in all stages of triggering.

Solenoidal Magnet

The SMT and the CFT are both inside a 2 T magnetic field provided by the superconducting solenoidal magnet. The magnet's size, 1.42 m in diameter, 2.73 m in length, was set by the size of the tracking cavity determined by the calorimeter cryostats, which are from Run I. The magnet has a thickness of $0.9X_0$ at $\eta = 0$, where X_0 is known as a radiation length, or the length needed for an electron's energy to be reduced to $1/e$ of its original energy by radiation losses.

The solenoid is wound in two layers of superconducting conductor made of Cu:NbTi strands stabilized with aluminum. There are two types of conductors in order to make the current density greater at the ends of the solenoid, which results in better field uniformity. The coils are kept at a superconducting temperature by liquid helium.

The nominal current that flows through the solenoid is 4749 A. After the shutdown that took place in 2005, however, the solenoid could no longer hold such a high current due to resistive heating at a solder joint, so the current had to be reduced to 4550 A. Therefore, after the shutdown, the nominal magnetic field was reduced to 1.92 T. The direction of the current can be reversed so that the magnet can operate in both polarities.

4.2.3 Preshower Detectors

To compensate for the material in front of the calorimeters, especially in the solenoidal magnet, preshower detectors were installed outside of the central tracking system. Their purpose is to aid electron and photon identification, by both providing extra tracking to match tracks with calorimeter showers, as well as providing an energy measurement. There are two preshower detectors: the central preshower (CPS),

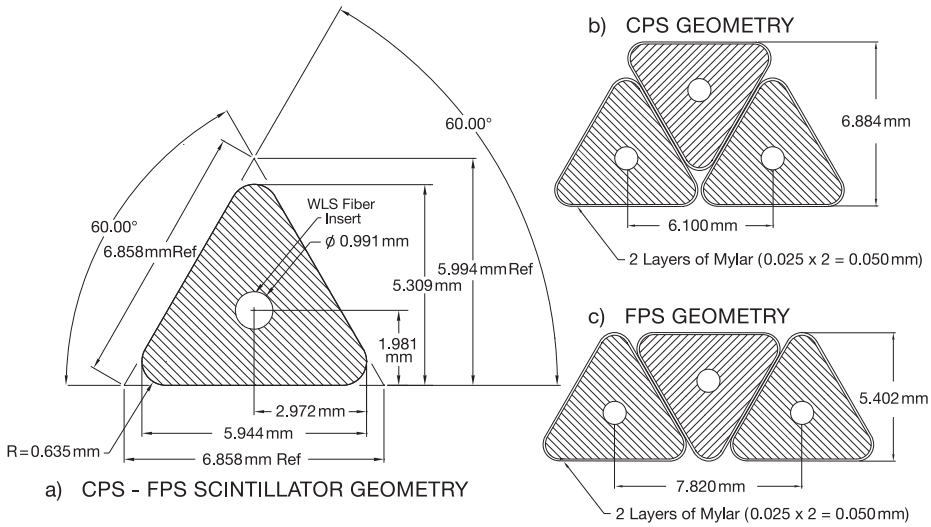


Figure 4.8: Cross-section of the scintillator strips used in the preshower detectors, with layout geometry shown for the CPS and FPS [110].

located between the solenoid and the central calorimeter and covering $|\eta| < 1.3$, and the forward preshower (FPS), located in front of the end calorimeter and covering $1.5 < |\eta| < 2.4$. Their position is shown on the diagram of the inner detector, Fig. 4.6.

The active layers of the preshower detectors are made from scintillating strips of triangular cross section, as shown in Fig. 4.8. The strips are made of PS plastic doped with 1% pT and 150 ppm diphenyl stilbene, and they are wrapped in aluminized mylar for optical isolation. In the middle of each strip there is a wave length shifting (WLS) fiber to collect and transmit the signal. These fibers are polished and silvered at one end, and read out using the same hardware as the CFT on the other end.

The CPS consists of three concentric cylindrical layers of scintillator placed behind a $1X_0$ lead radiator. One layer is in the axial direction, and two stereo layers are arranged at $\pm 24°$. Together, the solenoid and the lead radiator account for $2X_0$. At that point electrons should have started showering but heavier particles should not

have, thereby providing for some discrimination. Additionally, photons should have also started showering and thus become visible in the preshower.

The FPS consists of four layers of scintillator: two layers at angle of 22.5° relative to each other, then a $2X_0$ radiator made of lead and stainless steel, followed by two more layers of scintillator. The group of scintillators in front of the radiator are called the minimum ionizing particle (MIP) layers. Charged particles coming from the interaction point should appear as MIP particles in this region, and due to the more limited forward tracking, this is used to locate the particle in the FPS. The radiator causes electrons and photons to start showering, but heavier particles should not. The two layers behind the radiator are called the shower layers, and their purpose is to detect electron and photon showers, in order to discriminate them from hadrons.

The axial scintillator layer is read out for the Level 1 trigger similarly to the way axial doublets of the CFT are read out. However, the Level 2 trigger also uses the stereo layers, while for the case of the CFT, the stereo doublets are not used for triggering before Level 3.

4.2.4 Calorimeters

The goal of a calorimeter is to measure the energy of a particle by total absorption. Electrons and photons are made to produce an electromagnetic shower, where due to interactions with matter, electrons bremsstrahlung, i.e., emit photons, and photons, in turn, pair-produce electrons. The shower continues until a critical energy is reached, at which point electrons start losing energy more by collision loss and not by bremsstrahlung, and photons no longer have enough energy to pair-produce electrons. As already mentioned, the radiation length, X_0, is the length for an electron to have its energy reduced by a factor of $1/e$ by radiation losses. Additionally, it is found

that the mean path photons travel before pair-producing electrons is approximately $\frac{9}{7}X_0$ [115]. Thus, it is natural to express thicknesses in terms of radiation lengths.

Hadronic particles and jets also shower when interacting with matter, though the shower progression is more complex [116]. The interaction length (λ_A), or the mean path between interactions that are not elastic or diffractive, provides a scale for the shower development. Hadron showers develop over a longer distance than electromagnetic showers, especially for high atomic number (Z) materials.

The only SM particles that do not shower in the calorimeters are muons and neutrinos. There is a special detector for muons, covered in Sec. 4.2.5, but neutrinos are only implicitly observed by conservation of momentum.

One of the strengths of the D0 detector in Run I was the quality of its calorimeters. Therefore, though the readout was changed in order to handle the shorter time between bunch crossings, the calorimeters themselves are largely unchanged. The calorimeters are sampling calorimeters, mainly of uranium plates and liquid argon (LAr) active medium. Figure 4.9 shows a schematic of a calorimeter cell. In the uranium of the absorber plates, the radiation length and the interaction length is short, so the particles can be stopped in a relatively short distance. However, they are "dead material" in the sense that no signal is read from them. Instead, there are the gaps between the absorber plates filled with LAr where the shower is sampled, hence the name. The LAr is ionized by charged particles passing through it, and the charges are collected by the signal boards. In this way the total energy of the incident particle or jet can be measured.

Figure 4.10 shows the three calorimeters: the central calorimeter (CC) and the two end calorimeters (ECS, ECN, or collectively EC). Closest to the interaction point are the electromagnetic (EM) sections of the calorimeters, whose purpose is to accu-

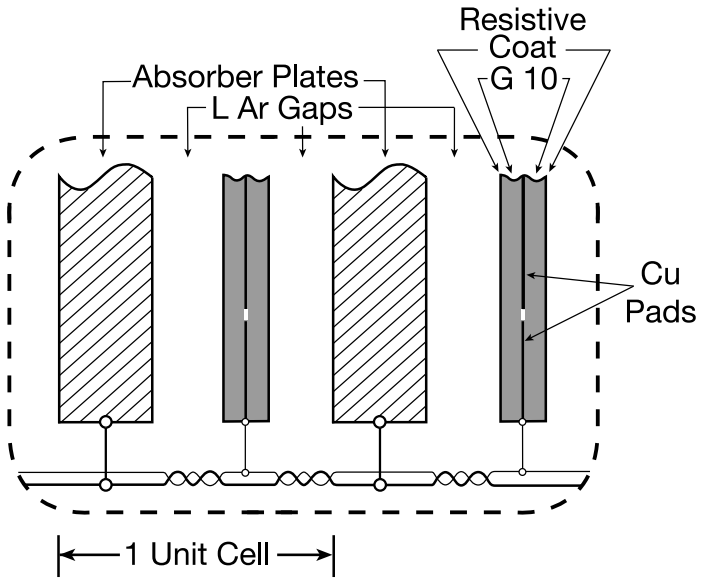


Figure 4.9: A schematic of a calorimeter cell [110].

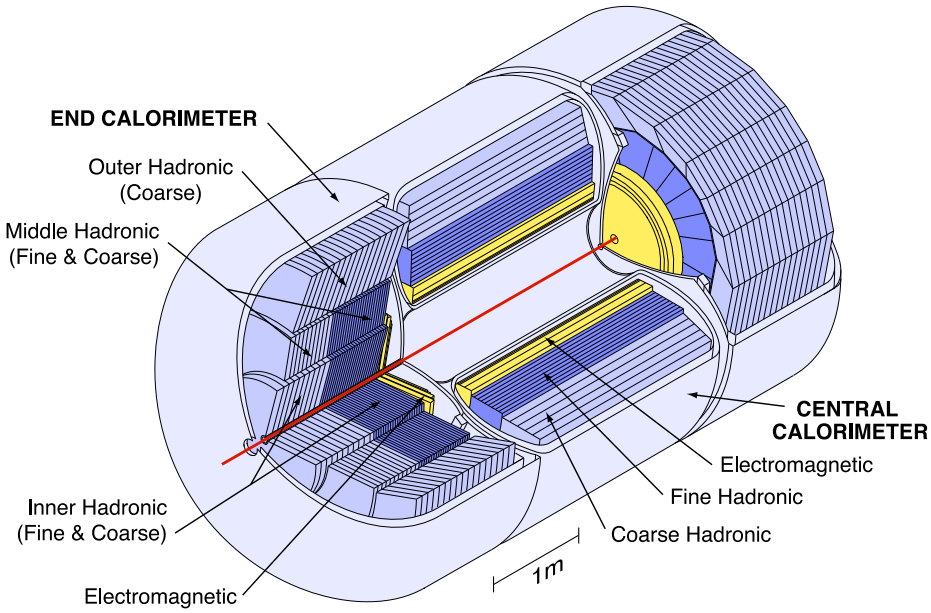


Figure 4.10: The calorimeters and their modules [110].

rate measure the energy of electrons and photons. They are built with 3 mm (CC) or 4 mm (EC) thick uranium plates from nearly pure depleted uranium[2]. The EM modules are divided into four depth layers, known as "floors," of approximately 1.4, 2.0, 6.8, and $9.8X_0$ thick in the CC, and 1.6, 2.6, 7.9, and $9.3X_0$ thick in the EC [110]. The number given for the first floor includes the material in the cryostat wall. The floors are called EM1, EM2, EM3, and EM4. Next in depth are the fine hadronic (FH) sections, which use 6 mm thick uranium-niobium (2%) alloy plates. The FH modules are divided into three floors of depth 1.3, 1.0, and $0.76\lambda_A$ in the CC, and four floors, each of either $1.1\lambda_A$ (inner hadronic) or $0.9\lambda_A$ (middle hadronic) in the EC. Those floors are called FH1, FH2, FH3, and FH4. Last in depth are the coarse hadronic (CH) sections, which uses 46.5 mm thick plates of copper in the CC or stainless steel in the EC. There is only one CH floor in the CC, approximately $3.2\lambda_A$ thick. In the inner hadronic and middle hadronic modules of the EC there is also one floor of thickness $4.1\lambda_A$ or $4.4\lambda_A$ respectively. The are three floors with a maximum thickness of $6.0\lambda_A$ in the outer hadronic module.

In addition to measuring energy, the calorimeters need to measure the location of the shower. For that, the floors are segmented into cells of $\Delta\eta = 0.1$ and $\Delta\phi = 2\pi/64 \approx 0.1$. In EM3, for better precision in determining the position of electromagnetic showers, the segmentation is $\Delta\eta = 0.05$ and $\Delta\phi = 2\pi/128 \approx 0.05$. In Run I the peak of electromagnetic showers occurred in EM3, which is why that floor was chosen to have finer resolution. In Run II, the shower peak tends to happen earlier because of the extra material in front of the calorimeter, the addition of the solenoidal magnet in particular. The cells are arranged into pseudo-projective towers, as shown in Fig. 4.11. They are called pseudo-projective and not just projective because the cell boundaries do not point towards the center of the detector. The η

[2]Depleted uranium has the uranium-235 isotope depleted.

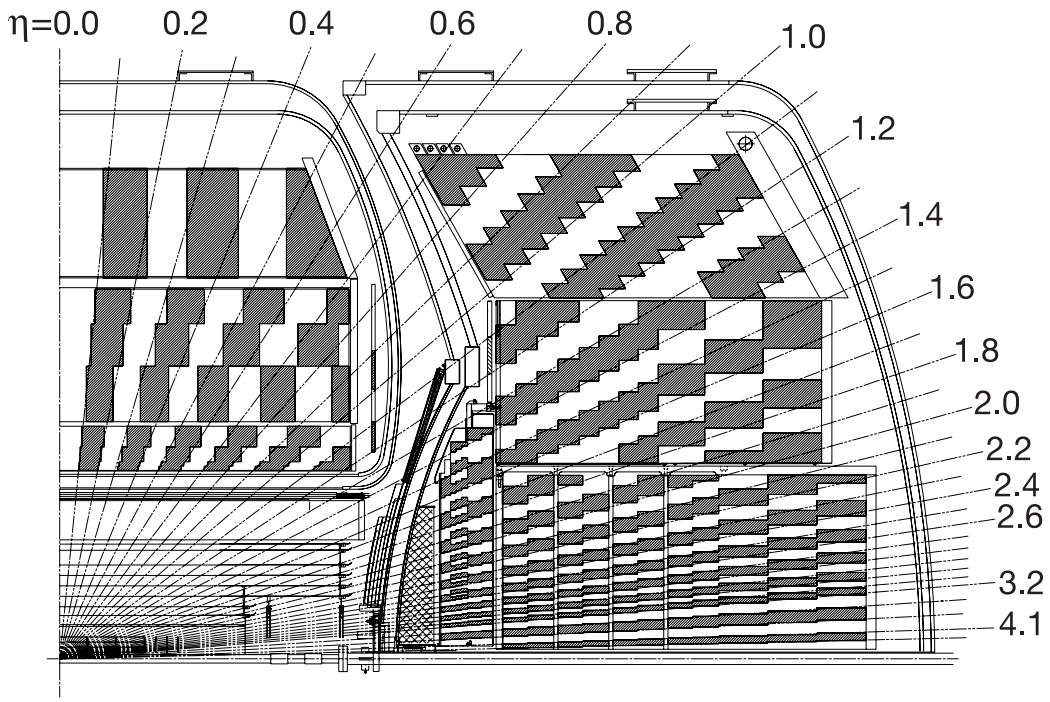


Figure 4.11: Schematic view of the calorimeter showing pseudo-projective towers and the calorimeter layout [110].

coverage of the calorimeter system extends out to about $|\eta| < 4$.

As can be seen in Fig. 4.11, in the area between the CC and the EC cryostats, there are certain trajectories that pass through much material before hitting the actual calorimeter structures. In order to add sampling to those trajectories, the massless gaps (MG) have been added within the cryostats, and the intercryostat detector (ICD) has been added between the cryostats. Those detectors can be seen in Fig. 4.11, though they have not been labeled in order to not over-complicate the figure. The ICD is also visible and labeled in Fig. 4.6. The massless gaps are standard calorimeter cells in front of the first layer of uranium in both the CC and the EC. The ICD which covers $1.1 < |\eta| < 1.4$ in Run II, is a series of 0.5 in thick Bicron BCF-400 scintillating tiles. The tiles are divided into subtiles of $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$ to match the calorimeter. The subtiles are read out by WLS fibers, which via clear fibers are sent to photo-multiplier tubes. The signal is shaped and made to be compatible to the standard calorimeter signals.

The relative momentum resolution for the calorimeter system is measured in data and found to be $\sigma(p_{\mathrm{T}})/ \approx 13\%$ for 50 GeV jets in the CC and $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} \approx 12\%$ for 50 GeV jets in the ECs. The energy resolution for electrons in the CC is $\sigma(E)/E \approx 16\%/\sqrt{E} \oplus 4\%$, where $\oplus$ means to add in quadrature.

The calorimeter is a key component at all levels of the trigger system. For Level 1 and 2, a faster readout is used, only distinguishing EM and FH layers in depth[3], and with lateral segmentation of $\Delta\eta \times \Delta\phi \approx 0.2 \times 0.2$. These 0.2×0.2 towers are called trigger towers. In Run IIa, the time period used for this analysis, the Level 1 triggers were only defined in terms of numbers of trigger towers with E_{T} above thresholds, measured either only in the EM layer or in the full trigger tower (EM + FH). At

[3]The CH is not used.

Level 2 there is some clustering of trigger towers done to define both EM objects and jets [117]. Electrons are made up of two towers, the seed tower, provided by L1, and its largest neighbor tower. Jet objects are 3×3 (before trigger version v9.31) or 5×5 trigger tower clusters, centered around a seed provided by L1.

4.2.5 Muon System

Outside of the calorimeter lies the muon system [118]. Muons generally pass through the calorimeter as minimum ionizing particles (losing around 1.6 GeV of energy), so it is the muon system's duty to identify them, measure their locations, and measure their transverse momenta. Because in Run II D0 has the solenoidal magnet for the central tracking systems, muon tracks also have their p_{T} measured there, usually with better accuracy. Nevertheless, the muon system is essential for muon identification. The muon system is also needed for the momentum measurement of muons with $|\eta| > 1.6$, for which the central tracking system has lower efficiency.

The muon system consists of a toroidal magnet, three layers of wire chambers, and scintillators. The purpose of the magnet is to measure the momentum of muon locally. The purpose of the wire chambers it to accurately measure the position of the muons (and for triggering). The purpose of the scintillators is to have a fast readout for accurate timing, rejecting cosmics, matching wire chamber hits with bunch crossings, and trigger. Figure 4.12 shows the exploded view of the arrangement of the wire chambers and scintillators. Note that the bottom of the detector is not fully instrumented. Figure 4.3 provides another view that shows the layout with regard to the toroidal magnet. The general arrangement is that there are three layers of sensors: the A layer first, then the magnet, after which are the B and C layers. There are two layers after the magnet so that the outgoing direction of the muons can be

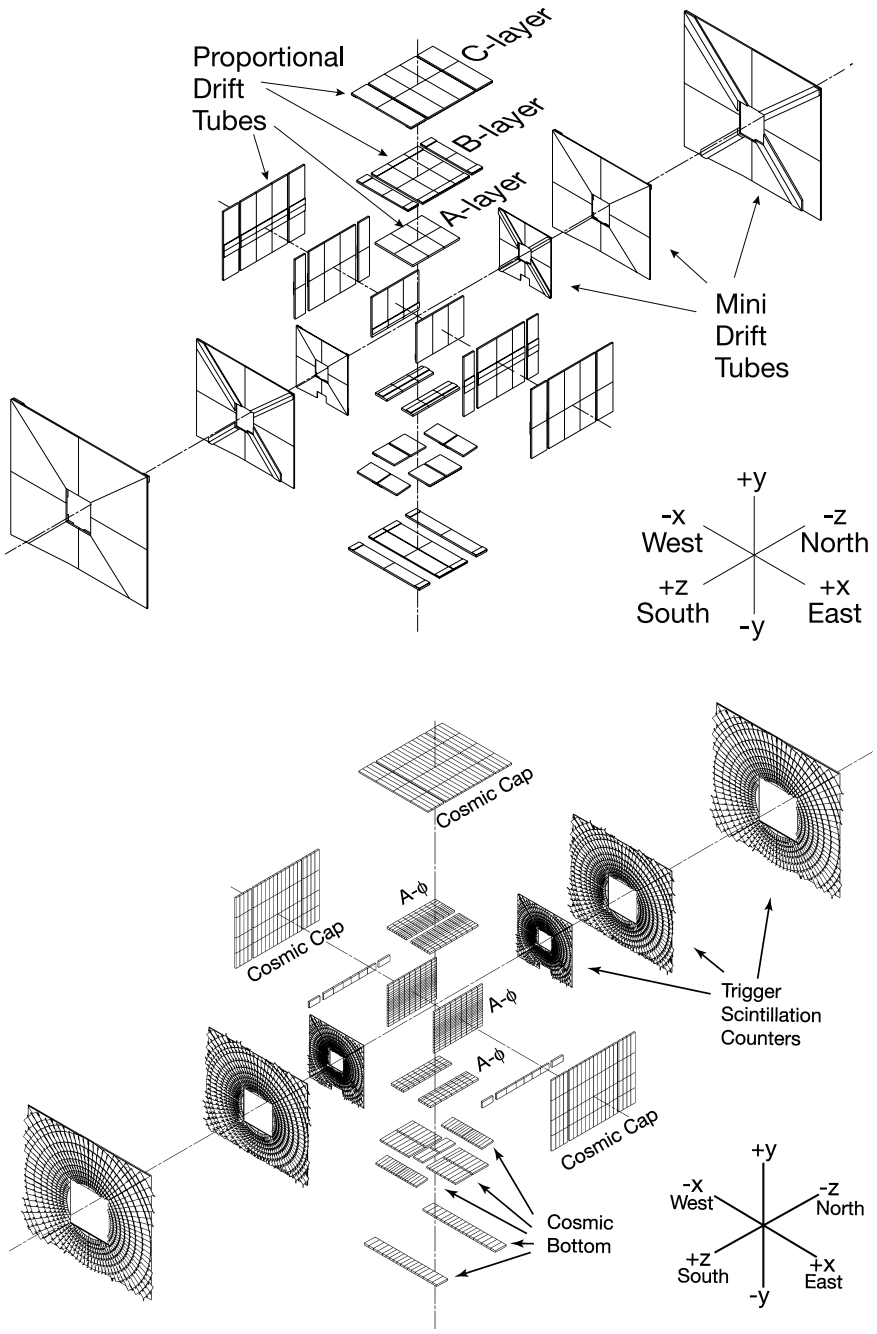


Figure 4.12: Exploded view of the muon system wire chambers (top) and scintillators (bottom) [110].

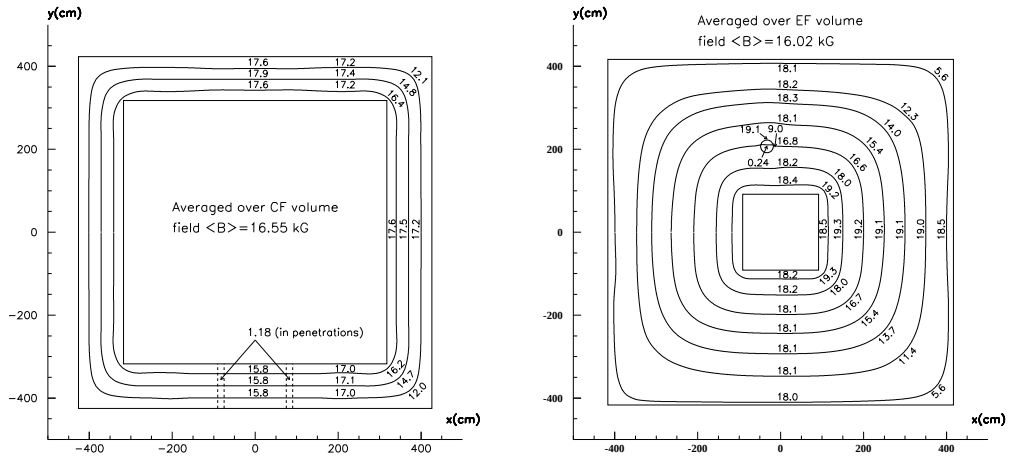


Figure 4.13: The magnetic field in the central (left) and end (right) toroidal magnets [118].

measured.

Figure 4.13 shows the magnetic field distribution in the central toroidal magnet (CF) and the end toroidal magnet (EF). The toroidal magnets runs at a lower current that they did in Run I as a cost-saving measure, because the momentum of muons is primarily measured in the central tracker now. The average energy loss across the toroidal magnet is 1.7 GeV.

The muon system can be divided into the central detector, covering roughly $|\eta| < 1$, and the forward muon system, extending the coverage to $|\eta| < 2$. The central detector uses the same proportional drift tubes (PDTs) from Run I, which have fairly large rectangular cells, 10.1 cm across. The tubes are filled with a faster gas mixture than in Run I, however: 84% argon, 8% methane, and 8% CF_4, for a maximum drift time of about 450 ns. The B and C layers have three rows, or "decks," of cells each, while the A layer has four decks, except along the bottom of the detector, where it has three decks. Approximately 55% of the central detector is covered by three layers

of PDTs, and 90% by at least two layers. The wires are arranged parallel to the magnetic field, which is in the y direction for the PDTs on the sides, and in the x direction for the top and bottom PDTs.

The anodes are read out in pairs, with the wires ganged together on one side. The time difference between the readouts from the two wires provides the position of the hit along the length of the wire. It has a resolution of between 10 cm and 50 cm depending on the location of the hit. Additionally, the cathodes use two separate electrodes that have thicknesses that vary with position, and the relative charge collected can be used to achieve a resolution of 5 mm along the direction of the wire. However, only the A layer and about 10% of the B and C layers are instrumented like this. The drift distance resolution is $\sigma \approx 1$ mm.

A problem found in the muon system during Run I was that the cathodes, made from copper-clad Glasteel strips, would outgas and over time coat the anodes. A method was developed to remove the coating, but for Run II, the electrodes were replaced in the A layer PDTs and the B layer PDTs under the detector, because access to those PDTs is difficult. The old cathodes remain in the other PDT's, however. The gas is recirculated independently for the A layer PDTs to keep them isolated for this problem, and in general there is molecular filtering to remove contaminants from the Glasteel.

The central detector has two layers of scintillators, the Aϕ system at the A layer, and the cosmic cap outside the toroid. The Aϕ scintillators are 4.5° in ϕ to match the CFT sectors for easier track matching, as are the new scintillators in the cosmic cap. However, the scintillators that were added in Run I have the long side oriented along ϕ. The time resolution of the Aϕ scintillators is $\sigma = 2.5$ ns for all counters combined without adjusting for differences in time-of-flight, light propagation in the counters,

or cable length within each group of fifteen counters.

The forward muon system has new for Run II mini drift tubes (MDTs), filled with 90% CF_4, 10% CH_4. The MDTs are much smaller than the wire chambers in the central system, of $0.94\,cm \times 0.94\,cm$ cross section. Their drift times are hence much shorter: less than 60 ns. The per plane muon hit detection efficiency is $(95 \pm 2)\%$. In test beam the coordinate resolution was measured to be around $350\,\mu$m, but due to digitization electronics, it becomes about 0.7 mm per hit. The MDTs do not measure the position along the length of the wire; better resolution is provided by the scintillators.

The forward detector has three layers of trapezoidal scintillators arranged in r-ϕ geometry. The ϕ segmentation is 4.5°, and there are about 12 concentric zones, representing η segmentation of 0.12 in the nine inner zones, 0.07 in the three outer zones. The time resolution for all counters combined for hits from reconstructed muon tracks is $\sigma = 2.2$ ns.

The muon system is used for triggering at all levels. At Level 1 there are two systems for triggering [119, 120], one based on the wire chambers and the scintillators, and one that attempt to match muon scintillator hits with tracks from the CTT. For the former, the scintillators are used to associated a signal in the wires with a bunch crossing, and the *loose* triggering requirement is defined if there is an A layer wire hit (or any layer in the bottom of the detector), while a *tight* trigger requires both an A layer and a B (in the forward detector) or C (in the central detector) layer wire hit. For the latter, there is a similar *loose* and *tight* requirement based on scintillator hits in various layers, and with the CTT track match, it is also possible to have a p_{T} cut using the CTT track. The Level 2 muon system [121] receives information from the L1 trigger and also information with fast calibration done straight from the

muon system. It can improve rejection by using better scintillator timing and finer tracking.

The upgraded trigger system for high-luminosity running during the later stages of Run II (Run IIb) uses the same hardware as the L1 Muon system for matching L1 calorimeter jet or electron to CTT tracks [122]. The new system is now in use.

4.2.6 Luminosity Monitor

The main purpose of the luminosity monitor (LM) is to determine the luminosity that D0 is observing, which it accomplishes by measuring the rate of inelastic $p\bar{p}$ collisions. The luminosity is measured as $\mathcal{L} = f\bar{N}_{\mathrm{LM}}/\sigma_{\mathrm{LM}}$, where $\bar{N}_{\mathrm{LM}}$ is the average number of inelastic collisions per beam crossings measured by the LM, f is the beam crossing frequency, and σ_{LM} is the effective cross section, taking into account the acceptance and efficiency of the LM [123]. The uncertainty on the luminosity is currently estimated to be 6.1% [124].

The LM consists of two arrays of 24 plastic scintillator counters with photomultiplier tube readouts covering $2.7 < |\eta| < 4.4$ placed in front of the End Calorimeters. The detectors can be seen on Fig. 4.6.

Chapter 5

Object Reconstruction

The data that is measured by the D0 detector needs to be processed for use in physics analyses. The hits in the SMT and CFT need to be combined into tracks. From the tracks the vertex needs to be determined. The calorimeter needs to be cleaned up from noise. The electrons, muons, and jets must be identified. Jets with b-quarks must be identified. Describing those steps is the purpose of this chapter.

5.1 Tracking

The data in the SMT and CFT detectors—the charge gathered by the microstrips or the light collected from the fibers—is first processed into "hits." Then the purpose of the tracking algorithms is to determine the trajectories of the charged particles that caused particular sequences of hits, called tracks.

Currently two algorithms are used to find tracks and their results are combined at the end. The first is called histogram track finding (HTF) [125]. The algorithm uses the notion that if there is a hit at point (x, y), then assuming a track coming

from the origin with curvature ρ and direction at the origin ϕ, the set of all parameter pairs, (ρ, ϕ), of tracks that go through (x, y) forms a line in parameter space. This mapping of points to lines in parameter space is called a Hough transform. All the hits of a given track would each map into a line, and these lines would intersect at the point, (ρ', ϕ'), that corresponds to the actual track. A similar Hough transform can be made between hits in (r, z) to lines in the parameter space (z_0, C), where z_0 is the position of the track origin along the z axis, and $C = dz/dr$ is the track inclination.

The HTF algorithm fills histograms in (ρ, ϕ) or (z_0, C) with the Hough transform lines. The intersection points, which correspond to real tracks, should correspond to peaks in the histogram; however, straightforward peak finding is not feasible. Instead, the histograms are cleaned up by removing bins with few entries, and then each bin in parameter space is considered to be a track template. Kalman filtering is used to select tracks from the templates [126, 127]. More details on how the algorithm is applied can be found in Ref. [125].

The other algorithm is called the alternate algorithm (AA) [128]. Working in 2D, the algorithm starts with three SMT hits to make a track hypothesis. The second hit must be on a following layer within $\Delta\phi$ of 0.08. The third, on a following layer, must be on a circle of radius greater than 30 cm and axial impact parameter with the beam spot of less than 2.5 cm. The overall fit must have $\chi^2 < 16$. Each track is extrapolated to the next layer of the SMT or CFT repeatedly, and hits are added to the track hypothesis if the increase of χ^2 is less than 16. If there are multiple hits in a given layer, they each become new hypotheses. A certain number of misses (i.e., no hits) in layers are allowed to improve the efficiency. The track hypotheses are ordered based on number of hits, and those that have equal number of hits by the fewest number of misses, and those that have the same number of hits and misses, by the

better the χ^2 of the fit. Tracks are accepted in that order if they pass a criteria that they do not share too many hits with already accepted tracks.

To further reduce the number of fake tracks, primary vertices are determined using the accepted tracks, and every track that comes close to a vertex is given two additional hits in the rankings. Using this new weight, the tracks are resorted, and a new pool of tracks is determined.

The description so far would preclude tracks with no SMT hits. Therefore, the same procedure is repeated starting with three CFT hits, but to control the huge combinatorics with stereo hit associations, the tracks must pass near a primary vertex determined from the SMT tracks.

At the end, the HTF tracks and the AA tracks are combined, refitted, and smoothed, as explained in Ref. [127], using the Kalman algorithm.

5.2 Primary Vertices

The location of the hard scatter, known as the primary vertex, is needed to correctly reconstruct the event. The main difficulty with determining the primary vertex is distinguishing which tracks come from it and which come from secondary vertices due to heavy quark decays or additional minimum bias interactions that happen to be close to the primary vertex. To accomplish this, we use an adaptive primary vertex algorithm [129].

The algorithm first clusters tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$, and two or more SMT hits if they are in an area with SMT acceptance, into different interaction areas 2 cm long along the z axis. Then, for each cluster, the location and width of the beam is determined using a Kalman filter with a "tear-down" approach, meaning that tracks

with the highest χ^2 contribution are removed, until the total vertex χ^2 per degree of freedom is smaller than 10. Subsequently only tracks with ($\mathrm{dca}/\sigma_{\mathrm{dca}} < 5$) are used, where dca stands for distance of closest approach to the beam position and dca is the standard deviation of the dca distribution.

Next comes the actual adaptive primary vertex algorithm, which assigns the following weighting function for each track i:

$$w_i = \frac{1}{1 + e^{\left(\chi_i^2 - \chi_{\mathrm{cutoff}}^2\right)/2T}} \tag{5.1}$$

where χ_i^2 is the χ^2 contribution of track i to the vertex, and $\chi_{\mathrm{cutoff}}^2 = 4$ and $T = 1$ tunable parameters, with their chosen values given. The algorithm consists of the following steps:

1. Start with all weights $w_i = 1$.
2. Weighing each track by its weight w_i, determine the primary vertex using a Kalman Filter.
3. Update the weights w_i using the new vertex. If $w_i < 10^{-6}$, set the weight to zero for that particular track.
4. If all the weights changed by less than 10^{-4}, then done. Otherwise, repeat from step 2. (There is also a maximum number of iterations exit in case convergence fails.)

This algorithm has been shown to be robust in cases with heavy flavor. For $Z \to q\bar{q}$ events the vertex resolution is $9.3\,\mu$m with a pull of 0.94, and more importantly, for $Z \to b\bar{b}$, the resolution is $12.8\,\mu$m with a pull of 1.02. The previous PV algorithm had a resolution of $14.1\,\mu$m with a pull of 1.40 for $Z \to b\bar{b}$ events.

After the list of primary vertices have been created, one must be chosen as the primary vertex of the hard scatter, while others are expected to come from minimum bias events. Minimum bias events have low p_{T} tracks, so a probability can be assigned to each track, based on its p_{T}, as to whether it comes from a minimum bias interaction [130]. From the probabilities of the tracks associated with a vertex, the probability that a vertex came from a minimum bias interaction can be determined. The vertex chosen as the primary vertex is the one with the lowest such probability.

5.3 Calorimeter Preprocessing

Before using the calorimeter data to reconstruct objects, the data is processed to remove noise. During data taking, individual cells can become "hot," that is, show high energy due to hardware problems. Such cells are usually suppressed so as to not affect data taking rates. At the Level 3 trigger and in the early stages of offline processing, the New Anomalous Deposits Algorithm (NADA) is used to further suppress hot cells and transient spikes in energy [131, 132, 133]. The general idea behind that algorithm is that if a cell has a high energy deposit ($E_{\mathrm{cell}} > 1\,\mathrm{GeV}$ in the original algorithm) while the sum of the energy of the 26 cells that surround it in a $3 \times 3 \times 3$ cube is low ($E_{\mathrm{cube}} < 100\,\mathrm{MeV}$ in the original algorithm, but only summing the energy of cells in the cube with $E > 100\,\mathrm{MeV}$), then the energy deposit is unphysical, and the cell's energy is set to $1\,\mathrm{MeV}$. There are special cases for certain layers (EM3, for example) and calorimeter boundaries in depth or η, and some of the thresholds can be dynamic, scaled as a fraction of the cell energy for energetic cells, but the basic idea is the same.

Another similar algorithm is used to remove noise: the T42 algorithm [134, 135,

136]. This noise could be electronics noise, uranium decays, or pile-up from interactions in previous bunch crossings, and so on. Zero-suppression zeros all cells with absolute energies less than 2.5σ above the pedestal value to remove most of this noise. The T42 algorithm additionally removes all negative energies. Furthermore, the T42 algorithm zeros all cells with energy between 2.5σ and 4σ above the pedestal value unless the cell has a neighbor with an energy over 4σ above the pedestal. Thus, around energy deposits that are considered to be signal-like, the zero-suppression is 2.5σ, but where there is no signal-like deposit, it is raised to 4σ. Neighbors are chosen similarly as in the NADA algorithm.

5.4 Jets

Outgoing quarks and gluons from the hard scatter form jets, which are detected in the calorimeter. The purpose of a jet algorithm is to reconstruct these jets from the energy deposits in the calorimeter towers in such a way that the kinematic properties can be related to those of the outgoing quark or gluon [137]. An ideal algorithm needs to fulfill certain requirements [137, 138]. It needs to be infrared safe, as demonstrated in Fig. 5.1. The jet algorithm must not be sensitive in its behavior to soft radiation. The ideal jet algorithm also needs to be collinear safe, which means that it should not be sensitive to collinear radiation. For example, if a jet algorithm depends on the energy deposit in just one tower as a seed, collinear radiation going into another cell could cause the energy to not be enough to form a seed. The algorithm should be invariant to boost in the longitudinal direction, which is not controlled in a hadron collider. The algorithm should be insensitive to the details of the final state (such as extra radiation). And of course, the algorithm should be experimentally well behaved: straightforward and efficient to implement, stable with luminosity, efficient, precise,

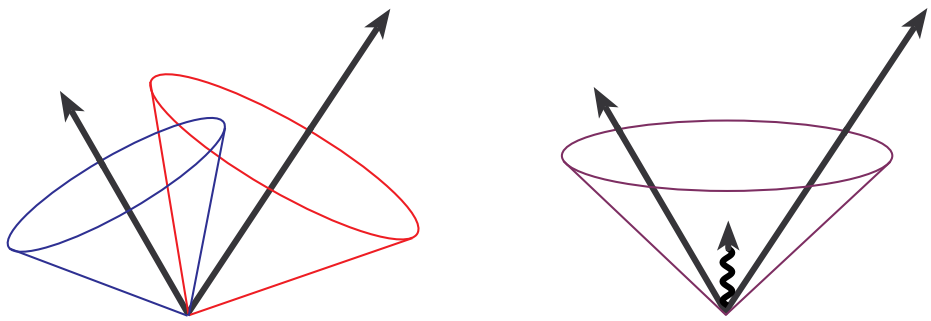


Figure 5.1: Diagram showing possible infrared sensitivity in jet algorithms. In this case the presence of soft radiation changes the algorithm's behavior from finding two jets to merging them into a single jet. (Figures from Ref. [137].

unbiased.

In an attempt to try to fulfill the above requirements, jets are reconstructed using the Run II Cone Algorithm [137, 138], which also goes by the name of Improved Legacy Cone Algorithm (ILCA). The algorithm is seed-based, in the sense that it starts looking for jets from seeds. Therefore, the first step is to find the seeds. For collinear safety the algorithm doesn't just use individual calorimeter towers as seeds but instead uses a *simple cone* algorithm to form preclusters for that role. The algorithm starts out with a p_{T}-ordered list of calorimeter towers. It takes the first tower, and if it has $p_{\mathrm{T}} > 500\,\mathrm{MeV}$, it makes it a precluster. (There are special conditions for towers with a lot of energy in the CH layers or in the EC massless gaps.) It then loops through the remaining towers, adding those within $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ and $p_{\mathrm{T}} > 1\,\mathrm{MeV}$ to the precluster. Whenever a tower is added to a precluster, it is removed from the list of towers. Then the algorithm continues to the next tower, and if it has $p_{\mathrm{T}} > 500\,\mathrm{MeV}$, it makes a new precluster, and repeats, until all the preclusters are made.

Next comes the clustering stage, which builds proto-jets starting from the preclusters. The preclusters are p_{T}-ordered, and the algorithm loops through all of them in turn. Using the definition, $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$, where y refers to the rapidity, the algorithm first checks if the precluster under consideration has $\Delta R < 0.25$ with any proto-jet. If so, it skips to the next precluster. Otherwise, the precluster is considered a seed for a proto-jet candidate PC. All towers within $\Delta R < 0.5$ of PC are added to it to make a new candidate PC'. If PC' has $p_{\mathrm{T}} < 3\,\mathrm{GeV}$, the algorithm skips to the next precluster. Otherwise, the algorithm checks to see if the proto-jet candidate was stable, that is if $\Delta R(PC, PC') < 10^{-3}$. If it was stable, and PC' is not a duplicate of an already found proto-jet, the algorithm makes PC' a proto-jet. Otherwise, the algorithm makes $PC' \rightarrow PC$ and iterates again adding towers within $\Delta R < 0.5$ until the proto-jet candidate is stable (or an iteration limit is reached).

To make the algorithm infrared safe, in addition to the proto-jets found above, there is an attempt to add additional proto-jets using seeds at the midpoints (sums) of any two proto-jets. The algorithm proceeds as when preclusters are used as seeds, except that it does not first check that $\Delta R < 0.25$ with any proto-jet or that a PC' is not a duplicate before making it into a proto-jet.

The last step is splitting or merging overlapping jets. Taking a p_{T}-ordered list of the proto-jets, the algorithm checks if the first proto-jet does not share any towers with any other proto-jets, and if it does not, it promotes it to a jet. Otherwise, taking the highest p_{T} neighbor proto-jet, the algorithm determines if the p_{T} of the shared towers is greater than 50% of the neighbor's p_{T}. If it is, the two jets are merged to create a new proto-jet and it is back into the list of proto-jets while the individual ones are removed. If shared p_{T} is less than 50%, then the shared towers are split by assigning them to the proto-jet that is closest in ΔR. In either case, the proto-jets

are resorted, and the algorithm iterates again.

After that step the list of jets is complete. Only those with $p_{\mathrm{T}} > 6\,\mathrm{GeV}$ are kept at this point. (The jet energy scale is not yet corrected.)

The following variables are used in the jet quality cuts:

- f_{CH}: the fraction of the jet's p_{T} contained in the CH layers of the calorimeter.
- f_{EM}: the fraction of the jet's p_{T} contained in the EM layers of the calorimeter.
- L1SET: the scalar sum of the trigger tower p_{T}s.
- $n90$: the minimum number of towers needed to contain 90% of the jet's p_{T}.
- $f90$: the fraction of the total towers needed to contain 90% of the jet's p_{T}.
- $\mathrm{L1}_{\mathrm{ratio}}$: L1SET divided by the uncorrected p_{T} of the jet, excluding the CH and MG layers.
- σ_{η}: the width of the jet in η.

Jets are considered to be of good quality if they fulfill the following criteria [139]:

- Pass coarse hadronic fraction cuts because the CH layers of the calorimeter tend to be noisy:
  - $f_{\mathrm{CH}} < 0.4$ or
  - $f_{\mathrm{CH}} < 0.44$ and $|\eta| < 0.8$ or
  - $f_{\mathrm{CH}} < 0.6$ and $0.85 < \left|\eta^{\mathrm{det}}\right| < 1.25$ and $n90 < 20$ (alternate cut for ICR jets where the detector has three CH layers) or
  - $f_{\mathrm{CH}} < 0.46$ and $1.5 < |\eta| < 2.5$.

- Pass EM fraction cuts to remove jets due to noise in the hadronic calorimeter:
  - $f_{\mathrm{EM}} > 0.05$ or
  - $\left|\left|\eta^{\mathrm{det}}\right| - 1.25\right| + \max[0, 40(\sigma_\eta - 0.1)] < 0.13$ (narrow jet in ICR where there is no EM layer) or
  - $f_{\mathrm{EM}} > 0.03$ and $1.1 < \left|\eta^{\mathrm{det}}\right| < 1.4$ (wide jet in ICR where there is no EM layer) or
  - $f_{\mathrm{EM}} > 0.04$ and $|\eta| > 2.5$.

  Jets that overlap with electrons or have $f_{\mathrm{EM}} > 0.95$ are also removed. We take *top_loose*, defined in Sec. 5.5, to be the electron criteria in that case.

- Pass L1 trigger tower confirmation to make sure that the jet is not just precision readout noise but is also visible in the trigger towers:
  - L1SET $> 55\,\mathrm{GeV}$ or
  - if outside of L1 coverage, $f90 < 0.5$ or $f_{\mathrm{CH}} < 0.15$
  - if inside L1 coverage,
    * $\mathrm{L1}_{\mathrm{ratio}} > 0.5$ or
    * $\mathrm{L1}_{\mathrm{ratio}} > 0.35$ and $p_{\mathrm{T}} < 15$ and $|\eta| > 1.4$ or
    * $\mathrm{L1}_{\mathrm{ratio}} > 0.1$ and $p_{\mathrm{T}} < 15$ and $|\eta| > 3.0$ or
    * $\mathrm{L1}_{\mathrm{ratio}} > 0.1$ and $p_{\mathrm{T}} \geq 15$ and $|\eta| > 3.0$.

The goal of the above criteria is to be 98–99% efficient on real jets without introducing significant unphysical jet background. The $f_{\mathrm{EM}} > 0.95$ cut is not included in that efficiency goal, however, and that can decrease the efficiency, as can be seen in Fig. 5.2. Real electrons and photon background might be included in in the dijet sample, but

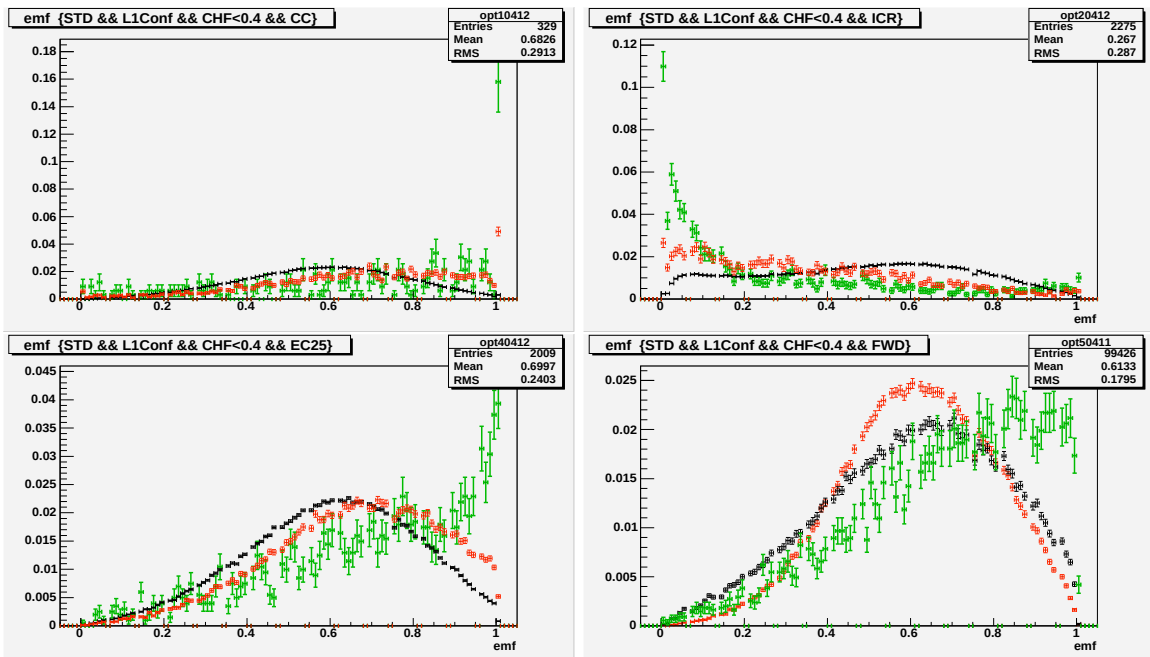


Figure 5.2: The jet f_{EM} distributions for various ranges of η. The distribution in black is for dijet events that is expected to be mostly good jets, while the red and the blue distributions are meant to model fakes. The figures are for the CC, ICR, EC ($1.5 < |\eta| < 2.5$), and EC–FWD ($|\eta| > 2.5$) (Figures from Ref [139].)

considering the shape of the f_{EM} for real electrons, Fig. 5.6, there is also significant jet content.

Finally, the energy of the jet must be corrected by the jet energy scale (JES) [140] so that it better matches the energy of the particle jet, that is, the jet of stable particles that comes out of the hard scatter after hadronization. The form of the correction is given in the following equation:

$$E_{\mathrm{particle}} = \frac{E_{\mathrm{raw}} - O}{F_{\eta} RS} \tag{5.2}$$

Energy not from the hard scatter (noise, additional minimum bias events, pile-up from previous bunch crossings) must be subtracted. This is the offset correction, given by O in Eq. 5.2. The correction is measured by determining the average per-tower energy in bunch crossings that consist only of minimum bias events (i.e., without a hard collision). The factors F_η and R are both response calibrations. F_η is η-intercalibration, which corrects the response to be uniform in η across the CC, ICR, and EC, and R is the absolute calibration. The corrections are measured by γ+jets events, because the EM energy is much more precisely measured, and also by di-jet events for intercalibration. Finally, S is the showering connection, which corrects for part of the shower leaking outside the jet cone, and also for outside energy leaking into the cone. It is measured with γ+jet data and Monte Carlo events. Figure 5.3 shows some of these corrections.

The total fractional uncertainties after making the above corrections is given in Fig. 5.4. However, one is often more concerned about differences in the energy scale between data and simulated (MC) events. For that, Fig 5.5 shows the difference between data and MC of the hemisphere observable (H), defined γ+jets events as the ratio of the p_{T} in the direction opposite the photon divided by the p_{T} in the direction of the photon.

Additionally, jets that contain a muon within $\Delta R(\mu, \mathrm{jetaxis}) < 0.5$ are corrected for the momentum carried away by the muon and its neutrino, which is assumed to have the same momentum as the muon. This correction is important for heavy flavor jets where the b-quark decays muonically.

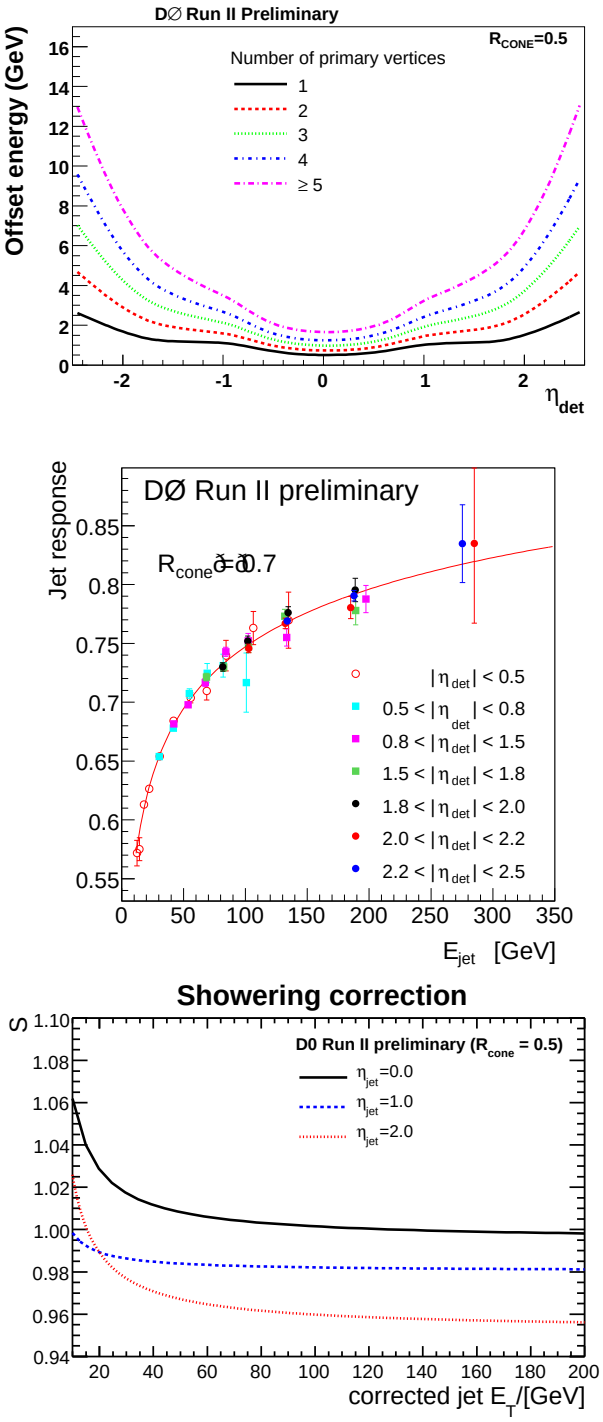


Figure 5.3: The offset term as a function of η and number of primary vertices (top), the uncorrected jet response as a fraction of the corrected jet response (middle), and the showering correction (bottom) [140].

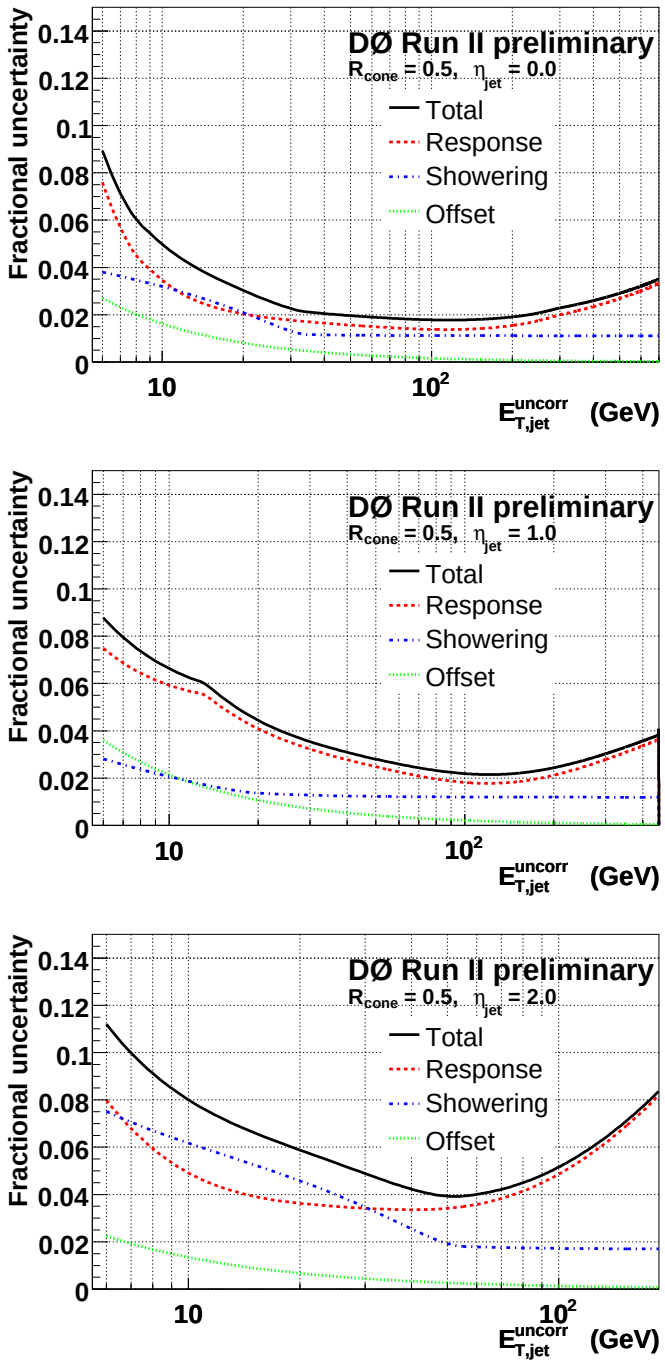


Figure 5.4: Jet energy scale uncertainties vs the uncorrected E_{T} at $\eta = 0$ (top), $\eta = 1$ (middle), and $\eta = 2$ (bottom) [140].

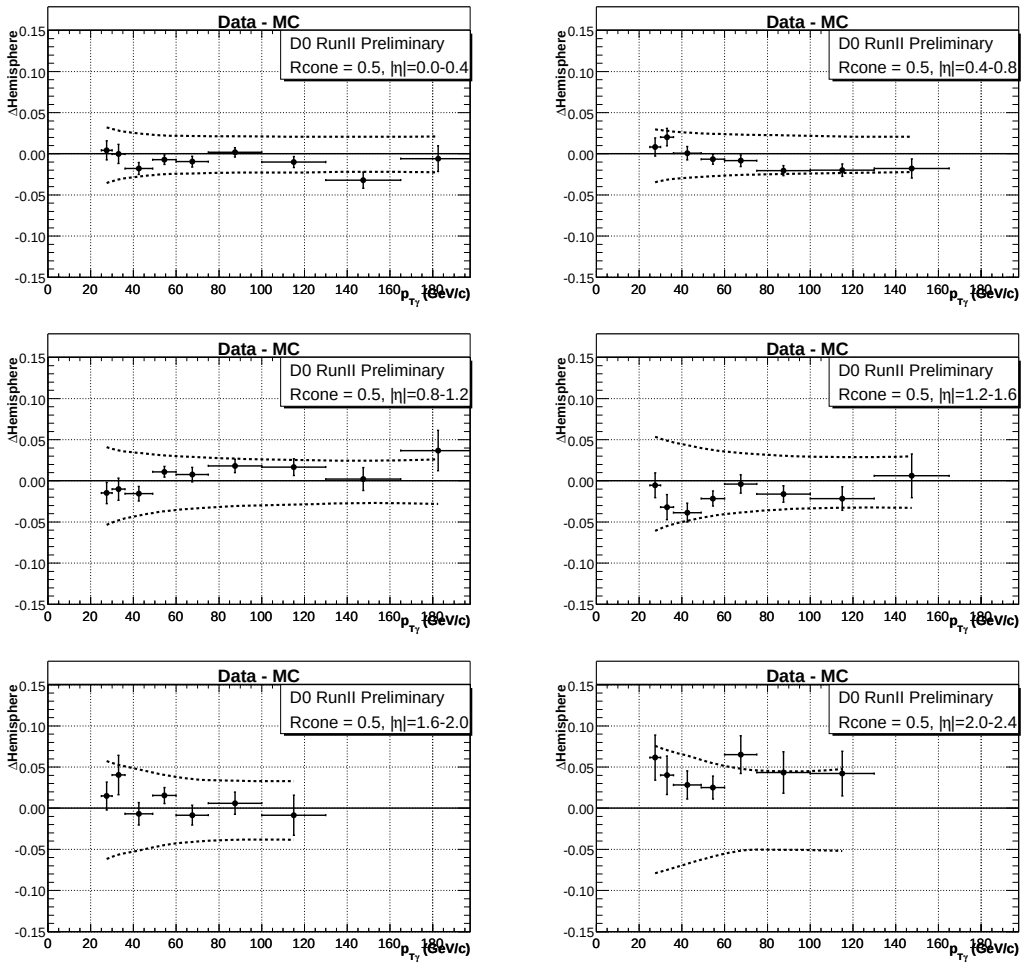


Figure 5.5: Difference between data and MC imbalances in γ+jet events for various $|\eta_{\text{jet}}|$ ranges. The dashed line is the quoted uncertainty [140].

5.5 Electrons

The electron identification algorithm should select electrons and reject background. Sources of background are π^0 showers which overlap with a nearby track, photons that convert to e^+e^- pairs, charged pions, and jet fluctuations. A study showed π^0 or η mesons to be the most common particles that fake electrons after requiring a track matched to the electron candidate [141].

Electrons, like jets, start out as calorimeter objects [142, 143]. The algorithm to build the EM clusters is the simple cone algorithm, similar to what is used to make preclusters for the jet algorithm. The algorithm uses a cone of $\Delta R < 0.4$ in (η, ϕ) around seed towers[1] with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$, and using the notation of E_{EM} being the energy in the EM layers and E_{tot} the total energy of the cluster, the cluster becomes an electron candidate if it has:

$$p_{\mathrm{T}} > 1.5\,\mathrm{GeV}$$

$$f_{\mathrm{EM}} \equiv \frac{E_{\mathrm{EM}}}{E_{\mathrm{tot}}} > 0.9$$

$$f_{\mathrm{iso}} \equiv \frac{E_{\mathrm{tot}}(R < 0.4) - E_{\mathrm{EM}}(R < 0.2)}{E_{\mathrm{tot}}(R < 0.4)} < 0.2$$

and at least 40% of its energy in the most energetic tower. In the last equation, $E(R < r)$ refers to the energy (EM or total) within a cone in (η, ϕ) of radius r. Thus, electrons are isolated calorimeter clusters with most of their energy in the EM layers of the calorimeter. We tighten the isolation requirement for the electrons to $f_{\mathrm{iso}} < 0.15$ to further reduce the jet background. Fig. 5.6 shows plots of f_{EM} and f_{iso} for CC electrons and fakes. For this plot and all others in this section, the real electrons are from dielectron samples dominated by $Z \to e^+e^-$ events, and the background is from

[1] For electrons, the towers only use the EM1, EM2, EM3, EM4, and FH1 floors.

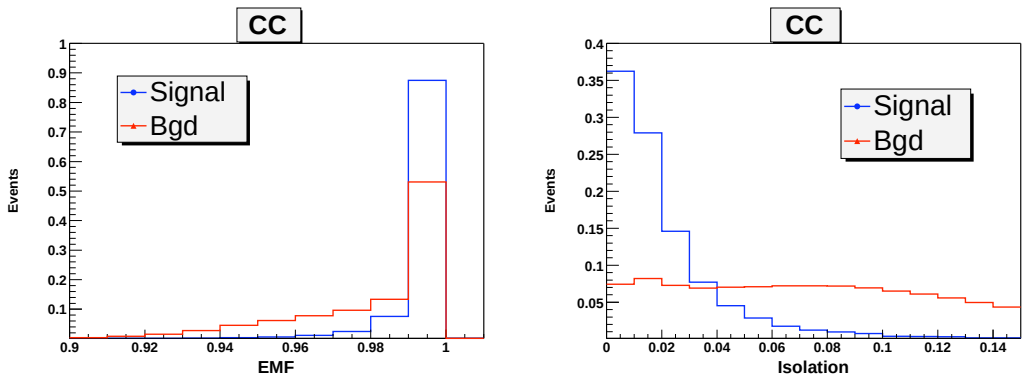


Figure 5.6: The f_{EM} (left) and f_{iso} (right) distributions for real and fake electrons (Figures from Ref [143].)

a "fake" electron back to back with a jet, with low $\not\!\!E_{\mathrm{T}}$.

The algorithm as described above is the loosest electron quality definition we use: *d0correct*, after the software package it was originally defined in. It is also called *Preselect_iso15* because it is the preselection step with $f_{\mathrm{iso}} < 0.15$ that many other electron quality definitions build upon. (There is a similar *Preselect_iso20* definition with only the $f_{\mathrm{iso}} < 0.2$ cut.)

Because the *d0correct*/*Preselect_iso15* algorithm does not have a track match, it also accepts photons with high efficiency, but the f_{EM} and f_{iso} requirements significantly suppress jets. The efficiency of this algorithm in the CC for real electrons with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ from $Z \to e^+e^-$ events is around $(96 \pm 2)\%$ [144], with the main culprit being intermodule cracks in the coverage in the CC, the "ϕ-cracks[2]." In the fiducial areas, the efficiency is $(99.2 \pm 0.8)\%$. The efficiency was measured was using a tag and probe technique on an electron+track data set dominated by Z events, as explained in more detail in Sec. 6.5.1. Electrons and photons of *d0correct* quality, $p_{\mathrm{T}} > 5\,\mathrm{GeV}$, and $|\eta| < 2.5$, are used for EM corrections to the $\not\!\!E_{\mathrm{T}}$ measurement, as explained in

[2]One of the main reasons that the ATLAS calorimeter has the accordion design is to avoid such inefficiencies.

Sec. 5.7.

Other electron quality definitions use the following variables:

- **H-matrix**:

  The purpose of the H-matrix is to measure how much the energy deposits in the calorimeter are compatible with an electron shower [145]. Given N electrons, the covariance matrix is calculated as:

$$M_{ij} = \frac{1}{N} \sum_{n=1}^{N} \left(x_i^{(n)} - \bar{x}_i \right) H_{ij} \left(x_j^{(n)} - \bar{x}_j \right) \tag{5.3}$$

  for a set of electron shower variables, x_i, and their means, $\bar{x}_i$. This is repeated in various η ranges to account for the different electron response. We then define $H = M^{-1}$. For a given electron candidate k, we can determine a χ^2 of how well its shower matches that of an electron:

$$\chi^2_{hm} = \sum_{i,j=1}^{7} \left(x_i^{(k)} - \bar{x}_i \right) H_{ij} \left(x_j^{(k)} - \bar{x}_j \right) \tag{5.4}$$

  The lower the value of χ^2_{hm}, the more electron-like the shower is. We use the H-matrix variant that uses seven variables:

  1. Energy fraction in the EM1 floor.
  2. Energy fraction in the EM2 floor.
  3. Energy fraction in the EM3 floor.
  4. Energy fraction in the EM4 floor.
  5. The $r\phi$ width of the shower in the EM3 floor.
  6. $\log_{10}(E)$

7. $z_{\mathrm{vtx}}/\sigma_{z_{\mathrm{vtx}}}$

- **Track match χ^2:**

  Electrons are expected to leave a track in the central tracking system, and this track should be well centered with the calorimeter cluster. Having a track thus provides a good way to distinguish electrons from neutral particles that do not leave a track. Furthermore, the quality of the track match can be used to distinguish between electrons and neutral particles that happen to overlap with a track by chance. There are two ways to define the track match χ^2:

$$\chi^2_{\mathrm{spatial}} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 \tag{5.5}$$

$$\chi^2_{\mathrm{with}\ E/p} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 + \left(\frac{E_{\mathrm{T}}/p_{\mathrm{T}} - 1}{\sigma_{E_{\mathrm{T}}/p_{\mathrm{T}}}}\right)^2 \tag{5.6}$$

$$\tag{5.7}$$

  where the variables are defined as

  - $\delta\phi = \phi_{\mathrm{track}} - \phi_{\mathrm{clus}}$, where the angles are measured at the EM3 floor. σ_ϕ is the expected width of $\delta\phi$ for an electron.
  - $\delta z = z_{\mathrm{track}} - z_{\mathrm{clus}}$, where the z values are measured at the EM3 floor. σ_z is the expected width of δz for an electron.
  - E_{T} is the transverse energy of the cluster.
  - p_{T} is the transverse momentum of the track.
  - $\sigma_{E_{\mathrm{T}}/p_{\mathrm{T}}}$ is the expected width of $E_{\mathrm{T}}/p_{\mathrm{T}}$ for an electron.

  For electrons $E_{\mathrm{T}}/p_{\mathrm{T}} \approx 1$, but if a track is randomly associated with the calorimeter cluster, this would not necessarily be the case. What is used to define the

quality of the track mach is the probability for a track to have a certain χ^2, $P(\chi^2)$.

- **Likelihood** [141]:

  A number of variables (noted as the vector x) are put together to define a likelihood that a track-mathed electron object is really an electron:

$$L = \frac{P_{\text{sig}}(x)}{P_{\text{sig}}(x) + P_{\text{bkg}}(x)} \tag{5.8}$$

  where $P_{\text{sig}}(x) = \prod_i P_{\text{sig},i}(x_i)$ and $P_{\text{bkg}}(x) = \prod_i P_{\text{bkg},i}(x_i)$, that is, the probabilities for signal and background are the product of the probabilities for the individual variables. The associated track is the one with the highest $P(\chi^2_{\text{spatial}})$ that has $\delta\phi < 0.05$ and $\delta\eta < 0.05$. The version of the likelihood that we use is based on seven variables:

  1. spatial track match probability, $P(\chi^2_{\text{spatial}})$
  2. (calorimeter cluster E_{T})/(track p_{T})
  3. The distance of closest approach (DCA) of the associated track to the primary vertex
  4. H-matrix (χ^2_{hm})
  5. EM fraction (f_{EM})
  6. The number of tracks in a cone of size $R = 0.05$ in (η, ϕ) around the track
  7. The sum of the transverse momenta of all the tracks other than the associated track in a cone of size $R = 0.4$

  The first two variables were discussed in the track match section; here the spatial part and $E_{\text{T}}/p_{\text{T}}$ are used as two separate variables. Isolated electrons should

come from the primary vertex, so the DCA should be small. The H-matrix and f_{EM} have also been discussed before. Electrons are expected to have low χ^2_{hm} values and $f_{\mathrm{EM}} \approx 1$. A loose cut of $\chi^2_{hm} < 50$ is applied before using the likelihood. Finally, the last two variables are track isolation variables. Electron tracks should be single, clean tracks, while electrons from photon conversion should have two tracks, and jets or parts of jets faking electrons should have extra tracks.

Plots of selected variables used in the likelihood are given in Fig 5.7. The real electrons are from a dielectron data sample dominated by $Z \to e^+e^-$, and the fake electrons are from a sample of a (fake) electron back to back with a jet with $\not{E}_{\mathrm{T}} < 15\,\mathrm{GeV}$.

The next tighter electron quality we use is called *top_loose*. It requires the following:

- Fulfill *Preselect_iso15* quality requirements
- $\chi^2_{hm} < 50$
- track match within $\delta\phi < 0.05$ and $\delta\eta < 0.05$ with $P(\chi^2_{\text{with } E/p}) > 0$
- track $p_{\mathrm{T}} > 5\,\mathrm{GeV}$

The H-Matrix cut is very loose, with the incremental efficiency given *Preselect_iso15* electrons (from Z events with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$) being $(98.7 \pm 0.6)\%$ in the CC. This is measured using a tag and probe method from a di-electron sample dominated by Z events, as explained in Sec. 6.5.1. For the track match, the $P(\chi^2_{\text{with } E/p}) > 0$ requirement is only a technical issue due to finite precision in floating point numbers on a computer. Tracks with zero probability are not returned, even if they are in the

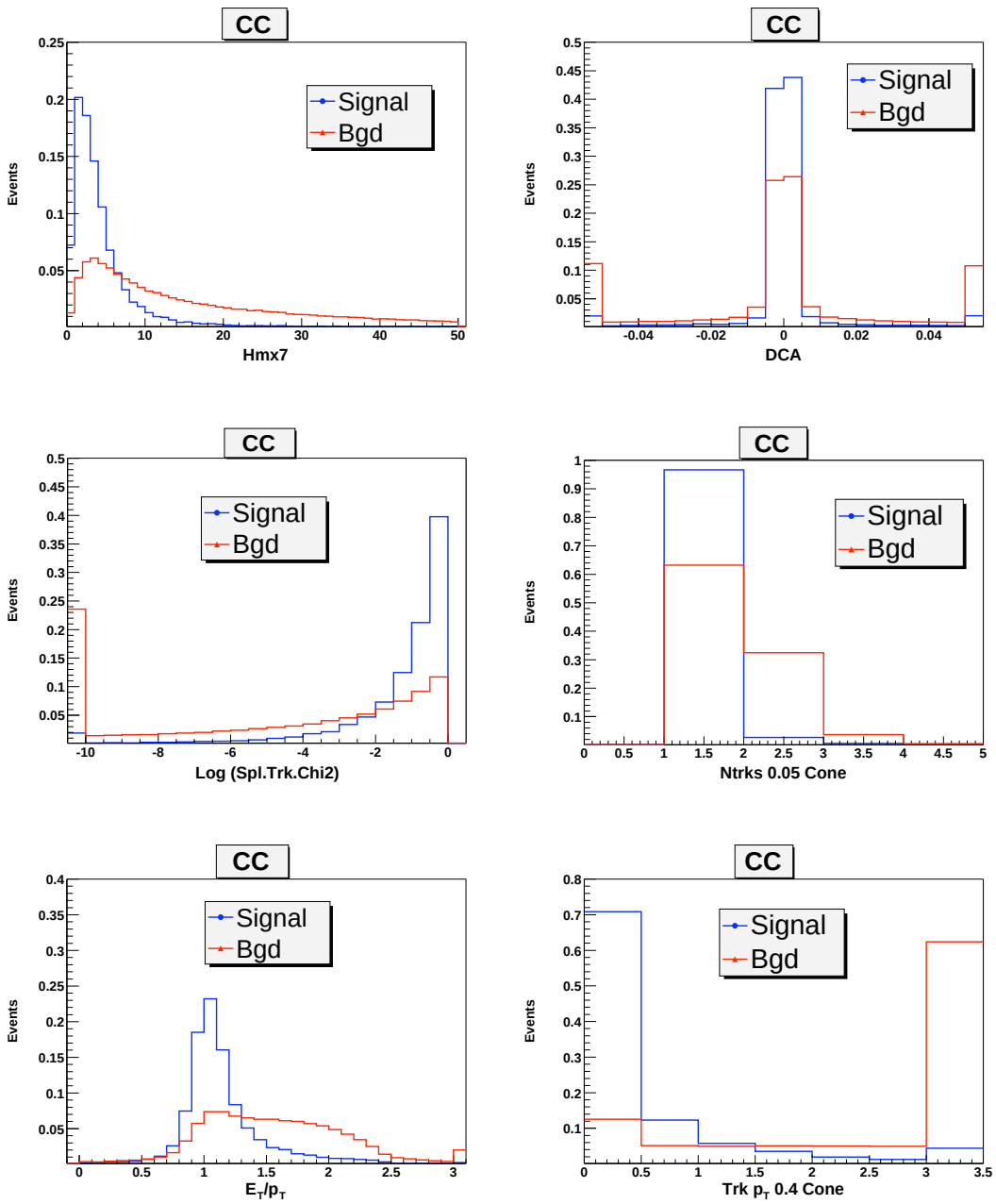


Figure 5.7: Selected electron discrimination variables distributions for real and fake electrons: H-matrix, $\log P(\chi^2_{\text{spatial}})$, and $E_{\text{T}}/p_{\text{T}}$ down the first column; DCA, the number of tracks in a cone of size $R = 0.05$, and he sum of the transverse momenta of all the tracks other than the associated track in a cone of size $R = 0.4$ down the second column. (Figures from Ref [143].)

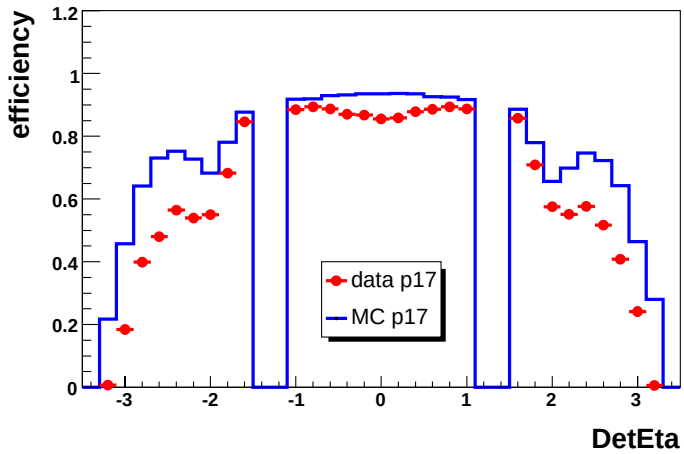


Figure 5.8: Efficiencies for data and MC electrons with quality *top_loose* as a function of η^{det}.

window. The track p_{T} cut is imposed by the data skims we use, so we added it to the electron quality definition. The incremental efficiency for the track match and track p_{T} cuts given *Preselect_iso15* electrons with $\chi^2_{hm} < 50$ and $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ in $Z \to e^+e^-$ events is $(89 \pm 2)\%$ in the CC. This efficiency drops quickly in the EC due to less efficient forward tracking, being around 86% at $\left|\eta^{\mathrm{det}}\right| = 1.6$ being around $(56 \pm 1)\%$ for $1.9 < \left|\eta^{\mathrm{det}}\right| < 2.5$, where it plateaus briefly, and falling to zero at $\left|\eta^{\mathrm{det}}\right| > 3.1$ Given the η distribution of electrons from $Z \to e^+e^-$ events, the EC efficiency is around 63%. A plot of the total efficiency for data and MC as a function of η^{det} is given in Fig. 5.8. The dip around $\eta^{\mathrm{det}} \approx 0$ is a result of incomplete modeling of tracking efficiencies in the MC.

The tightest requirement that we use is called *top_tight*. It adds a likelihood cut of $L > 0.85$ to the *top_loose* quality. The incremental efficiency for the likelihood cut given Z electrons with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ that pass *top_loose* is $(89 \pm 2)\%$ in the CC. In the

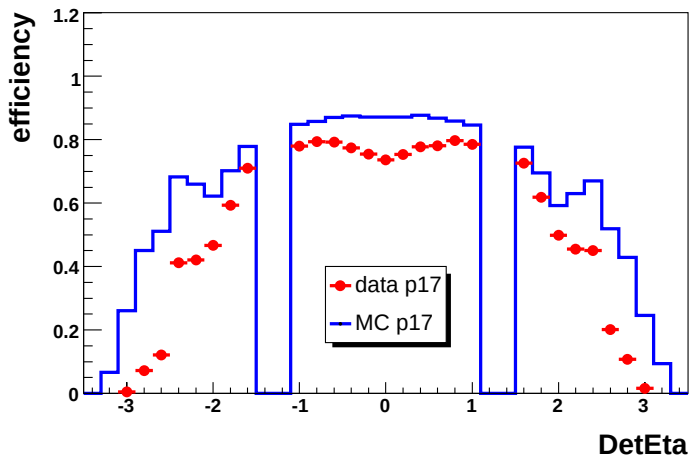


Figure 5.9: Efficiencies for data and MC electrons with quality *top_tight* as a function of η^{det}.

EC, it is around $(84 \pm 4)\%$ for $1.5 < \left|\eta^{\mathrm{det}}\right| < 2.5$, then immediately dropping to less than 40%, and approaching zero at $\left|\eta^{\mathrm{det}}\right| \approx 3$. Given the η distribution of electrons from $Z \to e^+e^-$ events, the efficiency in the EC is 78%. A plot of the total efficiency for data and MC as a function of η^{det} is given in Fig. 5.9.

Measurement of the incremental likelihood efficiency for both real electrons and fakes for events that pass our analysis's selection cuts is used to estimate the number of real electrons and fakes in our sample. Using the method explained in Sec. 6.7, the incremental efficiency for real electrons is estimated at $(88 \pm 2)\%$[3]. The events in our data sample have more jets, which tends to decrease the efficiency. The incremental efficiency for fakes was measured to be 13–22% depending on the trigger version. The selected events need to fire an electron trigger, which has a shower shape requirement and can be considered to be analogous to a tighter H-matrix cut, and since that cut

[3]We do not use that number directly but parameterize the efficiency in η^{det} and p_{T}. One-dimensional projections of that parametrization are shown in Sec. 6.7.

changed between trigger versions, the tighter that cut is, the more electron-like the fakes look. Note that the *top_loose* requirements remove most object that could fake electrons, so the incremental efficiency is for the most electron-like of the fakes.

5.6 Muons

Muon reconstruction involves several steps [146, 147]. First the hits need to be reconstructed at the three layers of the muon system. Then hits in each individual layer are reconstructed into straight-line segments [148]. Also, because there is no field between them, tracks between layers B and C are matched. Subsequently, the muon segments need to be reconstructed across the toroidal magnet to form what are called local muon tracks [149]. The algorithm works by grouping segments into pairs that pass certain cuts and propagating a track step by step from the BC segment, through the toroid, to the A segment, taking into account the energy loss in the magnet. The last step is to combine a local muon track with a central tracker track, which is done using error matrix propagation, taking into account the magnetic field in the toroid and solenoid and multiple scattering and energy loss in the calorimeter and toroid. The matching is done at the distance of closest approach to the beam axis.

The various certified muon identification requirements are explained in Ref. [150]. We use muons of quality $|\text{nseg}| = 3$ *medium*, which means that the muon has:

- at least two A layer wire hits
- at least one A layer scintillator hit
- at least two BC layer wire hits

- at least one BC scintillator hit (except for central muons with less than four BC wire hits)

We require that the muon be matched to a central track with quality *medium*, which means that the track reconstructed in the muon system must match a track reconstructed in the central tracker with $\chi^2/\text{dof} < 4$ and the central track must have a distance of closest approach with the primary vertex of $|\text{dca}| < 0.2\,\text{cm}$, tightened to $|\text{dca}| < 0.02\,\text{cm}$ if the track has an SMT hit. Furthermore, the muon must be away from a jet with $\Delta R(\mu, \text{jet}) > 0.5$ in (η, ϕ). Finally, we require a cosmic veto cut, which requires the scintillator hit times for both A and BC be within 10 ns of the expected time for a muon coming from a collision in that bunch crossing. A muon satisfying the above criteria we call a "loose" muon.

We also have a "tight" muon definition, which requires the same criteria as the "loose" muon with the addition of the *TopScaledLoose* isolation criteria: (a) the momenta of all tracks in a cone of radius $R < 0.5$ around the muon direction, except the track matched to the muon, must add up to less than 20% of the muon p_{T}; and (b) the energy deposited in an annular cone of radius $0.1 < R < 0.4$ around the muon direction must be less that 20% of the muon p_{T}.

Figure 5.10 shows the muon identification efficiencies, clearly showing the inefficiencies associated with the hole in coverage in the uninstrumented bottom of the detector. Figure 5.11 shows the isolation efficiencies for the *TopScaledLoose* and $\Delta R(\mu, \text{jet}) > 0.5$ criteria, among others.

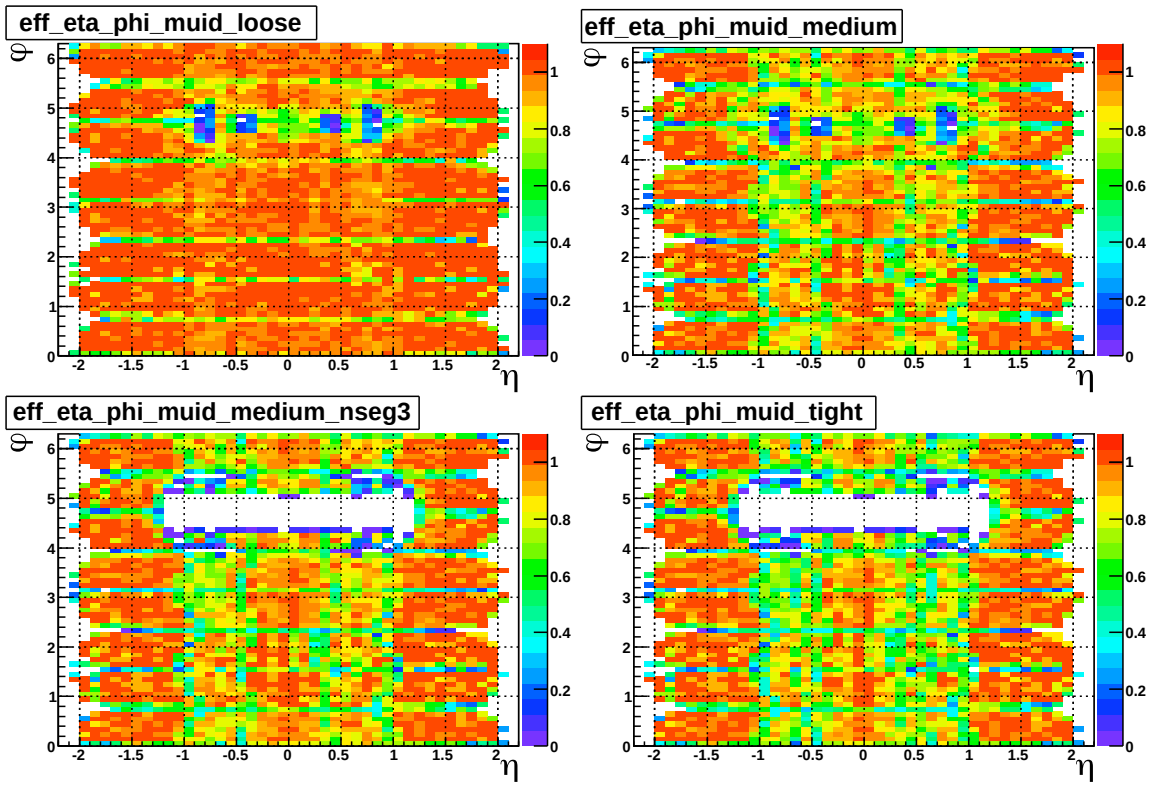


Figure 5.10: Muon reconstruction efficiencies as a function of η and ϕ [150].

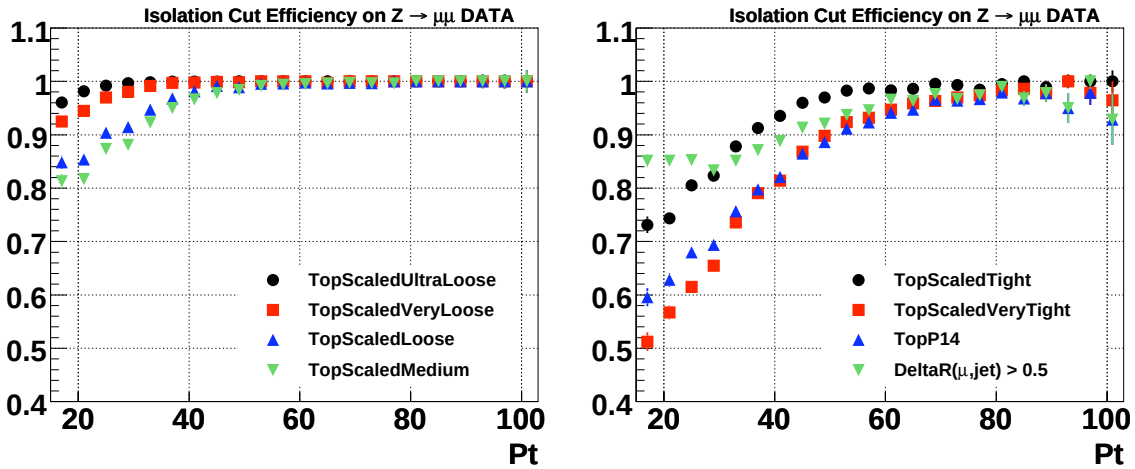


Figure 5.11: Muon isolation efficiencies for *TopScaledLoose* (left) and $\Delta R(\mu, \text{jet}) > 0.5$ (right) [150].

5.7 Missing Transverse Energy

Before the collision the proton and antiproton only have momentum in the z direction, so the momentum in the x-y plane, the transverse momentum, is zero. Conservation of momentum guarantees that to be the case after the collision as well[4]. This fact can be used to infer the presence of particles that escape detection, with neutrinos being a prime example. The missing transverse energy, $\not\!\!E_{\mathrm{T}}$, is the negative of the vectorial sum of the transverse momenta of particles observed in the detector. In practice we compute the $\not\!\!E_{\mathrm{T}}$ by adding up vectorially the transverse energies in all cells of the EM and FH layers of the calorimeter [151]. This value is corrected for the reconstructed particles with their associated energy scales. Jets are the prime example. Cells in the CH layers are too noisy to be added unless they are inside of jets, in which case they are included in the correction. The $\not\!\!E_{\mathrm{T}}$ is corrected for electrons and photons of *d0correct* quality, $p_{\mathrm{T}} > 5\,\mathrm{GeV}$, and $\eta < 2.5$, and for isolated

[4]Of course momentum is also conserved in the z direction, but we are not able to measure the momentum that goes down the beam pipe.

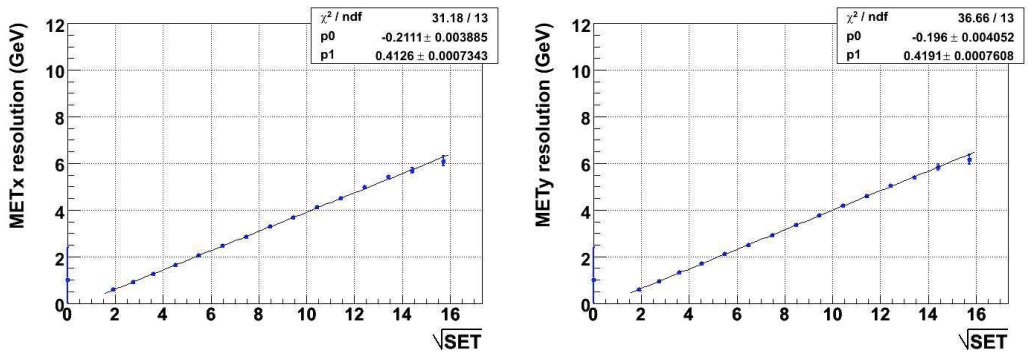


Figure 5.12: Missing transverse energy resolution along the x direction (left) and y direction (right) as a function of the square root of the scalar sum of the transverse energy in the calorimeter system [152].

muons with the "loose" quality we have defined. For isolated muons, the muons' energy deposit in the calorimeter is subtracted from the $\not\!\!E_{\mathrm{T}}$ measurement, and the reconstructed muon p_{T} is added instead. Note that muons inside of jets are treated through the JES correction, not through a muon object correction.

Figure 5.12 shows the $\not\!\!E_{\mathrm{T}}$ resolution during the time period used in this analysis for zero bias and minimum bias events as a function of the square root of the scalar sum of the transverse energy in the calorimeter system.

5.8 b-Tagging Jets

Many interesting physics processes, including single top production, result in b-quarks in the final state, so it is desirable to distinguish jets that come from b-quarks from other jets. There are two general categories of methods to distinguish b-jets from other jets. The first are methods that depend on being able to find a leptonic decay of the b-quark, such as the Soft Lepton Tag (SLT) algorithm [153]. They depend on the fact that about 20% of b-jets are associated with a muon because $B(b \to \mu X) =$

10.95%, b-hadrons decay with an almost 100% probability to charm hadrons, and $B(c \to \mu X) = 9.58\%$. The probability for light-quark jets to have muons is much smaller. Thus, a jet with a muon inside of it is considered to be tagged by the SLT algorithm.

The second category of methods to distinguish b-jets from other jets are those that depend on the lifetime of b-mesons, which usually travel some distance from the primary vertex before decaying. The distance traveled before decaying, the decay length, was determined to average around 3 mm for b-quarks in $t\bar{t}$ event [154]. The algorithms can either try to reconstruct a secondary vertex (SV) where the b-hadron decays, or they can rely on finding tracks offset from the primary vertex based on their impact parameter, defined as the minimum distance to the primary vertex.

These algorithms require first that the primary vertex be found, and then build jets out of the tracks [154, 155]. To build track-jets, first the tracks, ordered by p_{T}, are grouped one by one into clusters along the z axis, with a track being added to a cluster if $\Delta z < 2\,\mathrm{cm}$. Then for each cluster, using the closest primary vertex for the calculation of the distance of closest approach in the transverse plane ($|b_{\mathrm{transverse}}|$) and along z (b_z) measurements, tracks are selected with

- $N_{\mathrm{SMT\ hits}} \geq 1$
- $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$
- $|b_{\mathrm{transverse}}| < 0.15\,\mathrm{cm}$
- $|b_z| < 0.40\,\mathrm{cm}$.

The $|b_{\mathrm{transverse}}|$ and $|b_z|$ requirements are to reject long-lived particles such as K_s^0 and Λ hadrons. Then, using a seed track with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$, the simple cone algorithm,

the same one as used for preclustering in jet reconstruction, is used to build track jets with a cone size of $R = 0.5$ in (η, ϕ). Calorimeter jets are considered to have a matched track-jet if they lie within $\Delta R < 0.5$ of each other.

The Counting Signed Impact Parameter (CSIP) algorithm [156] considers a jet tagged if the impact parameter significance of two tracks (fulfilling certain quality requirements) is greater than three, or the signed impact parameter significance of three tracks is greater than two. For each jet, let b be the impact parameter, and σ_b its error. Then the impact parameter significance is $S = b/\sigma_b$. A sign is associated with it, positive if the impact parameter projection onto the jet axis is between the PV and jet, and negative if it is on the other side of the vertex. This signed impact parameter significance can be renormalized with a scale factor. The Jet LIfetime Probability (JLIP) algorithm [157] is similarly based on the impact parameter, but builds a probability based on all the impact parameters of the jets.

At D0 the algorithm to tag jets based on a secondary vertex is called the Secondary Vertex Tagger (SVT) [154, 155]. The algorithm selects tracks from the track-jet with an impact parameter significance $S > 3.0$ and builds all two-track seed vertices from them. The algorithm then tries to add other tracks pointing to the seeds according to the resulting χ^2 contribution, until no more tracks can be associated with the seed. The decay length is defined as $|L_{xy}| = |\mathbf{r}_{SV} - \mathbf{r}_{PV}|$, with a sign associated in the same way as for the signed impact parameter in the CSIP algorithm. The decay length significance is defined as $|L_{xy}/\sigma(L_{xy})|$. Operating points are defined based on the decay length significance, the minimum impact parameter significance, p_{T}, number of SMT hits, and χ^2 of the tracks used, the χ^2 of the SV, the vertex collinearity, and so on.

A neural network (NN) tagger was built upon primitives from the CSIP, JLIP,

Rank	Variable	Description
1	SVT_{SL} DLS	decay length significance of the Secondary Vertex (SV)
2	CSIP Comb	weighted combination of the tracks' IP significances
3	JLIP Prob	probability that the jet originates from the PV
4	SVT_{SL} χ^2_{dof}	chi square per degree of freedom of the SV
5	SVT_{SL} N_{Tracks}	number of tracks used to reconstruct the SV
6	SVT_{SL} Mass	mass of the SV
7	SVT_{SL} Num	number of SVs found in the jet

Table 5.1: The variables used by the NN tagger [159].

and SVT algorithms [158, 159]. The SLT algorithm is excluded because it is used to measure the efficiency of the tagger. The variables used to discriminate b-jets from other jets, ranked in order of separation power, are given in Table 5.1: The SVT_{SL} variables refer to the SVT algorithm run at the *super-loose* operating point. The CSIP combination is a weighted combination of the number of tracks with signed impact parameter significances greater than 2 or greater than 3, and also less than -2 or less than -3 (but with $\Delta\phi < 1.15$ with the jet). The SVT_{SL} Mass assumes the tracks are from pions, and corrects the mass for neutral particles.

For this analysis, we have chosen to use the NN tagger at the TIGHT operating point which requires the NN output to be greater than 0.775. The efficiencies for data are given in Fig 5.13, and the fake rates are given in Fig. 5.14[5]. The average fake rate for this operating point is 0.47% for jets in the CC, and its average b-tagging efficiency on data is 47% for jets up to $|\eta| = 2.5$[6] [159].

[5]How these are calculated is explained in Sec: 6.8.2
[6]Quoted are the preliminary versions used in the analysis.

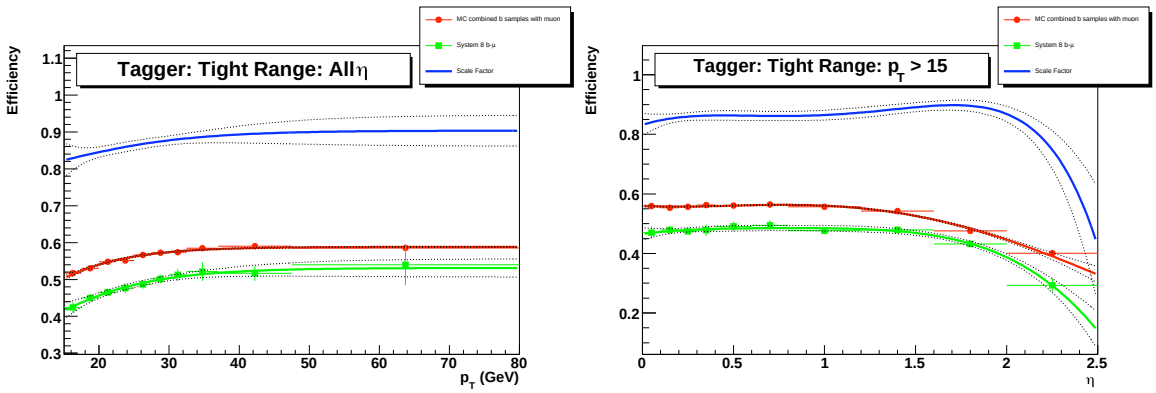


Figure 5.13: *b*-tagging efficiencies for data at the NN TIGHT operating point are given in green. (The red line shows the performance of running the tagger on MC and the blue line is the ratio between the data efficiency and the MC efficiency. In this analysis we do not run the tagger on MC.) [159].

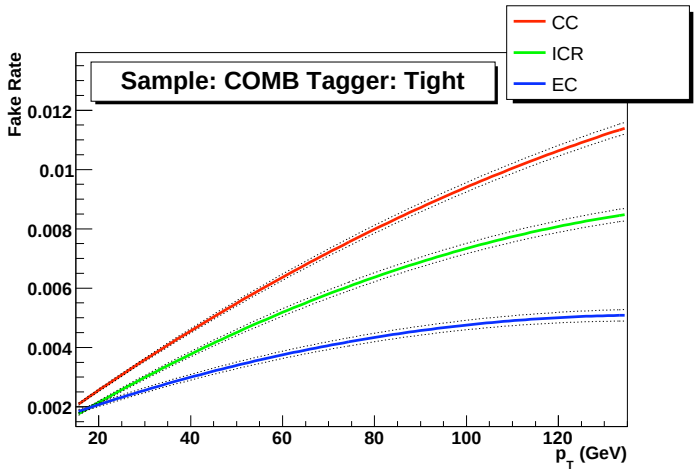


Figure 5.14: *b*-tagging fake rates (i.e., the probability to tag a light jet) at the NN TIGHT operating point [159].

Chapter 6

The Analysis: Event Selection

The analysis is divided into three steps: event selection, the development of a multivariate technique for discriminating between signal and background, and extracting a result. The selection step consists of loosely selecting the data to be analyzed, as well as constructing an appropriate background model. A common selection was shared by every analysis done in the single top subgroup for the $0.9\,\mathrm{fb}^{-1}$ data set. The second step, the multivariate analysis, is unique for each analysis technique. There were three primary techniques employed in the original published result [91]: Boosted Decision Trees (DT) [160], Bayesian Neural Networks (BNN) [161], and Matrix Elements (ME) [162]. A second iteration of the BNN has since been done [92, 163], and this thesis describes the second iteration of the ME technique. The third step, extracting a result, consists of applying systematic uncertainties and determining the cross section and significance.

This analysis only uses events with two or three jets, and one or two of those jets need to have a b-tag. The signal to background ratio is too low in zero-tagged events, and four-jet events are too computationally challenging for only a marginal benefit

at best. A graphical representation of how much signal and background falls in each combination of number of jets and number of b-tags can be seen in Fig. 6.1 for single top production via s-channel and t-channel processes. The different combinations of lepton type, number of jets, and number of b-tags are analyzed separately because of their different signal and background makeup, and combined only at the end. Thus, this analysis has eight different channels: 2 lepton types (e, μ) $\times$ 2 number of jet bins (2, 3) $\times$ 2 number of tag bins (1, 2).

The purpose of this chapter is to describe the first steps of the analysis, the selection. The two remaining steps are covered in the two subsequent chapters.

6.1 Background Processes

The single top event signature, described in Sec. 3.1, consists of one charged lepton, $\not{E}_{\mathrm{T}}$, and two, sometimes three jets, with at least one of them being a b-jet. Background processes that can mimic the signal event topology include:

- **W+jets**

  Events consisting of a W boson and extra jets form the largest background. The cross section for W+2jets is over 1000 pb [164, 165], including a 1% $Wb\bar{b}$ contribution, which has the same final state as s-channel single top. Requiring one or more b-tagged jets is the primary way to reduce the W+light jet background, though it still remains a major background. In addition to $Wb\bar{b}$, Wc and $Wc\bar{c}$ have higher b-tagging fake rates, so they become a higher percentage of the W+jets sample after b-tagging.

- **Z/γ^*+jets**

  Events with Z/γ^* have two isolated leptons (e^+e^-, $\mu^+\mu^-$, or $\tau^+\tau^-$) in the final

Percentage of s-channel *tb* selected events and S:B ratio (white squares = no plans to analyze)

Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	8% 1 : 11,000	19% 1 : 1,600	9% 1 : 1,200	3% 1 : 1,100	1% 1 :1,000
1 tag	6% 1 : 270	24% 1 : 55	12% 1 : 73	3% 1 : 130	1% 1 : 200
2 tags		9% 1 : 12	4% 1 : 27	1% 1 : 92	0% 1 :110

Percentage of t-channel *tqb* selected events and S:B ratio (white squares = no plans to analyze)

Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 4,400	27% 1 : 520	13% 1 : 400	4% 1 : 360	1% 1 : 300
1 tag	6% 1 : 150	20% 1 : 32	11% 1 : 37	4% 1 : 58	1% 1 : 72
2 tags		1% 1 : 100	2% 1 : 36	1% 1 : 65	0% 1 : 70

Figure 6.1: Percentages of the selected s-channel (upper table) and t-channel (lower table) signal events and signal:background ratios in each analysis channel.

state. For those with e^+e^- or $\mu^+\mu^-$ to mimic the single top event topology, one of the leptons needs to be lost, either by not being reconstructed or by overlapping with a jet. Some ways that an electron can fail to be reconstructed include if it falls in a ϕ-crack in the central calorimeter, or goes into the Intercryostat Region (ICR), or fails an electron identification cut. A muon can fail to be reconstructed if it falls in the uninstrumented area in the bottom of the detector or fails muon identification cuts. The main method for $Z/\gamma^*(\rightarrow \tau^+\tau^-)$+jets events to mimic single top is for one of the taus to decay leptonically and the other hadronically. Overall, after selection, the Z/γ^*+jets background is an order of magnitude smaller than that from W+jets.

- **WW, WZ, and ZZ**

  WW and WZ diboson events can mimic the signal topology if a W boson decays leptonically and the other boson decays into quarks. However, the cross sections for the diboson processes are a few picobarns each, so they are only a small background. ZZ diboson events are less likely to pass the selection because they require one Z boson to decay to e^+e^- or $\mu^+\mu^-$ and subsequently have one lepton to be lost, or to $\tau^+\tau^-$ with one of the taus decaying leptonically and the other hadronically, suffering from a low branching ratio.

- **QCD Multijets**

  Multijet events form a background in the electron channel when a jet is misidentified as an electron and mismeasurement produces a significant $\not\!\!E_{\mathrm{T}}$. The probability to misidentify a jet as an electron is rather small, about 10^{-4}, but the $\geq$3 jet cross section is so large that the overall contribution is significant. Multijet events can also form a background in the muon channel when a muon is created by the in-flight decay of a pion or kaon, or from punch-through, which

is when particles from a jet leak out of the calorimeter and enter the muon system. Tighter tracking and isolation requirements can suppress these backgrounds. Requiring b-tagging reduces the multijet backgrounds in both the electron and muon channels.

- **Heavy flavor production ($b\bar{b}$, $c\bar{c}$)**

  Heavy flavor production can contribute like the multijet background in the electron channel with a jet faking an electron. Though its cross section is much smaller than that for multijet events, it is not suppressed by b-tagging. In the muon channel, this process contributes to the background when one of the b-quarks decays muonically and the muon either travels wide of its jet or the jet is not reconstructed. (A similar background in the electron channel, of an electron traveling wide from a jet, is small.) This background will subsequently be included implicitly in the "multijet" background category.

- **$t\bar{t}$ production**

  The production of $t\bar{t}$ pairs by the strong force (NNLO cross section $= 6.77 \pm 0.42\,\text{pb}$, for $m_t = 175\,\text{GeV}$, $Q^2 = m_t^2$ [166]) is a significant background. The relevant processes have a larger multiplicity of final-state particles than single top production, but they contain top quarks and W bosons, so the kinematics are similar. The $t\bar{t}\to\ell\ell$ process can mimic the single top signature if one lepton is lost, in which case it usually falls in the two or three jet bin. A larger background is $t\bar{t}\to\ell+$jets, which can enter the three-jet bin, which has significant single-top content, if one jet is lost.

6.2 Signal and Background Modeling

To measure signal acceptance and extract expected results, we need a model of the signal. For that purpose, we have generated single top event samples with the CompHEP-SingleTop [167, 56] Monte Carlo event generator. SingleTop produces events whose kinematic distributions match those from NLO calculations. Ref. [77] showed that for s-channel single top production, the LO distributions times a K-factor match the NLO distributions. Therefore, SingleTop uses the LO diagrams times a K-factor. For the t-channel, however, the situation is more complicated because the LO diagrams, of the form $qb \to q't$, are found to not be adequate. The W-gluon fusion diagrams of the form $qg \to q't\bar{b}$ need to be applied as a correction. The diagrams cannot just be added, however, because the gluon splitting kernel is part of the b-quark PDF used in the LO diagram. The double counting is avoided by splitting the phase space of the diagrams, as explained below.

Within the proton, b-quarks are formed by gluon splitting, which can be modeled well by ISR in pythia [168] for the LO t-channel diagram. However, it is found that the hard "third jet" $\bar{b}$-jet distribution is not modeled well by the ISR jet. The W-gluon fusion diagrams, on the other hand, model the hard $\bar{b}$-jet distributions well, but not the soft because they do not model the large logarithmic corrections for soft $g \to b\bar{b}$. Those corrections are included in the b-quark PDFs used by the LO diagrams. Therefore, it is useful to split the phase space by the hardness of the $\bar{b}$-jet, and use the LO diagram for soft $\bar{b}$-jets and the W-gluon fusion diagrams for $\bar{b}$-jets. Figure 6.2 shows two attempts to merge the two processes, on the left with the split occurring at $p_{\mathrm{T}}^{\mathrm{cut}} = 20\,\mathrm{GeV}$, and on the right with the split occurring at $p_{\mathrm{T}}^{\mathrm{cut}} = 10\,\mathrm{GeV}$. The one on the right was chosen because it provides a smooth distribution. As can be seen in Refs. [167, 56], with this prescription, the distributions of the generated events match

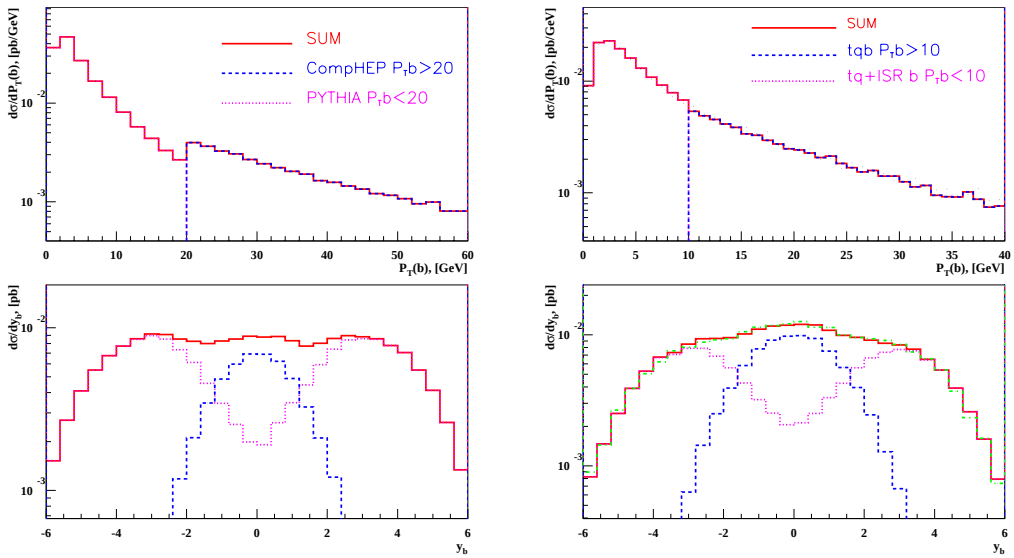


Figure 6.2: The p_{T} and rapidity (y) distributions of the "third jet" b-quark. The "soft" $p\bar{p} \to tq + b_{\mathrm{ISR}}$ process is combined with the "hard" $p\bar{p} \to tq + b_{\mathrm{LO}}$ with the split falling at $p_{\mathrm{T}}^{\mathrm{cut}} = 20\,\mathrm{GeV}$ (left) and $p_{\mathrm{T}}^{\mathrm{cut}} = 10\,\mathrm{GeV}$ (right). (Plots from [56].)

those of NLO calculations.

SingleTop decays the top quarks and W bosons to ensure the spins are properly treated. In our setup, tauola [169] decayed tau leptons and EvtGen [170] decayed b hadrons. pythia was used to add the underlying event and initial- and final-state radiation. A top quark mass of 175 GeV has been used for event generation throughout this analysis. The CTEQ6L1 [20] pdf set was used, with the scales set to m_t^2 for the s-channel and $(m_t/2)^2$ for the t-channel in order for there to be "maximum closeness" of the LO cross section to the NLO cross section [167, 171].

For background modeling, we use three categories: W+jets, multijet, and $t\bar{t}$. The Z/γ^*+jets and diboson background is small, so by normalizing the W+jets background yield to data, as explained in Sec. 6.7, we implicitly add the Z/γ^*+jets and diboson yield. It is expected that the discriminator output shape of this simplified

scheme matches, within error, the shape it would have if we had explicitly modeled the Z/γ^*+jets and diboson backgrounds.

The W+jets and $t\bar{t}$ backgrounds are modeled by events generated using the ALPGEN Monte Carlo event generator [164], interfaced to PYTHIA for the parton shower. This version of ALPGEN includes a matching scheme, the MLM prescription, to match parton showers with the matrix elements [172]. Since there is no factorization theorem to separate the contributions of the hard process and the showering, the matching scheme provides a way to not double-count any of the phase-space. The CTEQ6L1 pdf set was used, with the scales set to $m_W^2 + p_T^2(W)$ for the W+jets sample and $m_{\text{top}}^2 + \sum p_T^2(\text{jets})$ for the $t\bar{t}$ sample.

For the W+jets we separately generated $W + m$ light jets, $W + c\bar{c} + n$ light jets, and $W + b\bar{b} + n$ light jets using ALPGEN, where m goes from zero to five and n goes from zero to three. As required by the MLM matching routines, all the samples have additional PYTHIA jets removed, except for the $W + 5$ light jets, $W + c\bar{c} + 3$ light jets, and $W + b\bar{b} + 3$ light jets samples, to avoid double-counting the phase space. Also, heavy flavor added by PYTHIA was removed form the light jets sample. The Wcj subprocesses, with massless charm quarks, are included in the W+light jets sample because there was a bug in ALPGEN's Wc subprocess at the time of generation

Table 6.1 shows the cross sections, branching ratio, initial numbers of events, and corresponding integrated luminosities of the Monte Carlo samples. The $t\bar{t}$ cross sections are from Ref. [166], the single top cross sections are from Ref. [77], and the W+jets cross sections, since they are not used in the analysis, are the values calculated by ALPGEN. Wjj refers to W+light jets, and except for that sample, all other samples correspond to much higher integrated luminosities than the data. The single top and $t\bar{t}$ samples are weighted to the integrated luminosity of our dataset

using the cross sections and branching ratio listed in that table. The W+jets sample yield is normalized to data, as is explained in Sec. 6.7. The relative weight between W+heavy flavor and W+light flavor is not well known [173, 174], so we measure it when we do b-tagging. That procedure is also explained in Sec. 6.7.

Event Type	SM Cross Section [pb]	Branching Ratio	Number of Events	Int. Lum. [fb^{-1}]
Signals				
$tb \to e$+jets	0.88 ± 0.14	0.1111 ± 0.0022	92,620	947
$tb \to \mu$+jets	0.88 ± 0.14	0.1111 ± 0.0022	122,346	1,251
$tb \to \tau$+jets	0.88 ± 0.14	0.1111 ± 0.0022	76,433	782
$tqb \to e$+jets	1.98 ± 0.30	0.1111 ± 0.0022	130,068	591
$tqb \to \mu$+jets	1.98 ± 0.30	0.1111 ± 0.0022	137,824	626
$tqb \to \tau$+jets	1.98 ± 0.30	0.1111 ± 0.0022	117,079	532
Backgrounds				
$t\bar{t} \to \ell$+jets	6.8 ± 1.2	0.4444 ± 0.0089	474,405	157
$t\bar{t} \to \ell\ell$	6.8 ± 1.2	0.1111 ± 0.0022	468,126	620
$Wb\bar{b} \to \ell\nu bb$	142	0.3333 ± 0.0066	1,335,146	28
$Wc\bar{c} \to \ell\nu cc$	583	0.3333 ± 0.0066	1,522,767	8
$Wjj \to \ell\nu jj$	$18,734$	0.3333 ± 0.0066	8,201,446	1

Table 6.1: The cross sections, branching ratios, initial numbers of events, and corresponding integrated luminosities of the Monte Carlo event samples.

All the Monte Carlo samples have been run through a GEANT-based simulator of the detector called DØGSTAR [175, 110], which simulated the response of the D0 detector. Next the events were processed by the program, DØSIM [176], which:

- overlays minimum bias events to model multiple interactions per bunch crossing,
- models calorimeter pileup from previous bunch crossings,
- models calorimeter noise,
- models SMT, CFT, and Muon system noise and inefficiencies.

Finally, the events were processed through the same reconstruction algorithms as data, which were discussed in Chapter 5.

Because there is no good MC description of QCD multijet events, that background is modeled with data. An orthogonal sample created by reversing part of the lepton identification requirement is used and properly scaled, as is explained in Sec. 6.7.

6.3 Data and Triggers

The data sample was collected between August 2002 and December 2005, corresponding to approximately $0.9\,\mathrm{fb}^{-1}$ of good quality data. Given the signal event signature, lepton+jet triggers were used to select events. Over this time period, there were many versions of trigger conditions used. Tables 6.2 and 6.3 show the specific triggers used, as well as the integrated luminosity collected with each trigger. Tables 6.4 and 6.5 show specifically what the various triggers require.

6.3.1 Electron Channel

For the electron channel, the triggers are calorimeter-based. At Level 1, there are only two types of trigger terms used. The main one is CEM(1,x), where x is 10, 11, or 12, which fires if there is a trigger tower in the calorimeter with electromagnetic $E_{\mathrm{T}} \geq x\,\mathrm{GeV}$. Trigger versions v8–v11 also require CJT(2,5), which fires if there are two or more trigger towers with total (electromagnetic+hadronic) $E_{\mathrm{T}} \geq 5\,\mathrm{GeV}$. The trigger tower that satisfies CEM(1,10) automatically satisfies half of CJT(2,5), so only one additional trigger tower with $E_{\mathrm{T}} \geq 5\,\mathrm{GeV}$ is required.

Trigger Version	Trigger Name	Integrated Luminosity [pb^{-1}] Delivered	Recorded	Good Quality
v8	EM15_2JT15	6	5	5
v9	EM15_2JT15	48	42	25
v10	EM15_2JT15	20	18	10
v11	EM15_2JT15	79	72	63
v12	E1_SHT15_2J20	273	251	227
v13.00–v13.23	E1_SHT15_2J_J25	80	73	55
v13.30–v13.90	E1_SHT15_2J_J30	354	325	294
v14	E1_SHT15_2J_J25	290	271	234
Total Integrated Luminosity		1,150	1,056	**913**

Table 6.2: Triggers and integrated luminosities by trigger version for the electron channel.

At Level 2, trigger versions v8–v11 require one calorimeter electron object, with $E_{\mathrm{T}} \geq 10\,\mathrm{GeV}$ and $f_{\mathrm{EM}} > 0.85$. Those trigger versions also require two jet objects with $E_{\mathrm{T}} \geq 10\,\mathrm{GeV}$, though almost always the object that satisfies the electron part of the Level 2 trigger satisfies half of the jet part, meaning that only one additional jet is required. Trigger version v12 has no Level 2 requirement, and trigger version v13 and v14 require one calorimeter electron object with $E_{\mathrm{T}} \geq 15\,\mathrm{GeV}$.

At Level 3, there are two types of electron triggers used: ELE_LOOSE_SH_T(1,15) for trigger versions v8–v11, and ELE_NLV_SHT(1,15) for v12–v14. Both types require a Level 3 electron object with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $f_{\mathrm{EM}} > 0.85$, and both make a "tight" cut on the shower widths on the EM1, EM2, and EM3 floors of the calorimeter. However, ELE_NLV_SHT(1,15) uses rescaled widths when it cuts on the shower shape [177]. The rescaling is a function of the values of η^{det} and

$$\phi_{\mathrm{mod}} \equiv \frac{32\phi}{2\pi} \mod 1 \tag{6.1}$$

of the electron. The variable ϕ_{mod} is a measure of the location of the electron within a calorimeter module so that effects from cell boundaries and cracks in the sensitive region can be taken into account in the shower width. Furthermore, those filters started making nonlinear corrections and using L3 event vertex to improve the E_{T} resolution and hence sharpen the turn-on. The jet part of the triggers, JET(1,x) and JET(2,x), fire when there are one or more or two or more, respectively, Level 3 jets with $E_{\mathrm{T}} > x\,\mathrm{GeV}$. For example, JET(1,25)JET(2,20) requires one jet to have $E_{\mathrm{T}} > 25\,\mathrm{GeV}$ and another, $E_{\mathrm{T}} > 20\,\mathrm{GeV}$, because the first jet also satisfies half of the JET(2,20) requirement. Where it gets more complicated is that the electron, if it has a high enough E_{T}, may also be one of the "jets" mentioned above.

To simplify trigger modeling we explicitly make the requirement offline that the reconstructed electron in our data set has fired the electron part of the triggers by requiring $\Delta R(\mathrm{reco}, \mathrm{trigger}) < 0.4$ at all levels, where $\Delta R = \sqrt{(\Delta\eta^{\mathrm{det}})^2 + (\Delta\phi^{\mathrm{det}})^2}$.

6.3.2 Muon Channel

For the muon channel, at Level 1, the triggers all require a muon and a calorimeter trigger tower (jet). The calorimeter requirement is either CJT(1,3) or CJT(1,5), namely, one or more trigger towers with total (electromagnetic+hadronic) $E_{\mathrm{T}} \geq 3\,\mathrm{GeV}$ or $E_{\mathrm{T}} \geq 5\,\mathrm{GeV}$, respectively. The muon part of the triggers are mu1ptxatxx or mu1ptxatlx, and those refer to at least one muon (the "1" in the name) with:

- no L1CTT track match or p_{T} requirement ("ptx"),
- in any region ("a": all),
- a tight scintillator requirement ("t")
- no requirement ("x") or loose ("l") wire chamber requirement,

Trigger Version	Trigger Name	Integrated Luminosity [pb^{-1}] Delivered	Recorded	Good Quality
v8	MU_JT20_L2M0	7	6	6
v9	MU_JT20_L2M0	48	42	25
v10	MU_JT20_L2M0	21	19	11
v11	MU_JT20_L2M0	79	74	65
v12	MU_JT25_L2M0	277	255	231
v13.00–v13.11	MUJ2_JT25	56	39	31
v13.20–v13.23	MUJ2_JT25_LM3	26	22	16
v13.30–v13.90	MUJ2_JT30_LM3	382	277	252
v14.00–v14.10	MUJ1_JT25_LM3	0	0	0
v14.20–v14.21	MUJ1_JT25_ILM3	25	23	21
v14.30–v14.90	MUJ1_JT35_LM3	265	248	214
Total Integrated Luminosity		1,187	1,006	**871**

Table 6.3: Triggers and integrated luminosities by trigger version for the muon channel.

- and no optional requirement (the last "x") [120].

At Level 2, all the triggers require a medium quality muon with no p_{T} requirement. Additionally, from trigger version v12, there has been a requirement for one or more jets with $E_{\mathrm{T}} \geq 10\,\mathrm{GeV}$ (v12) or $E_{\mathrm{T}} \geq 8\,\mathrm{GeV}$ (v13, v14).

At Level 3, before trigger v13.30, there was simply a one or more jet requirement, with $E_{\mathrm{T}} > 20\,\mathrm{GeV}$ or $E_{\mathrm{T}} > 25\,\mathrm{GeV}$. Starting with trigger version v13.30, MUON(1,3,loose) was added, which requires one or more loose Level 3 muons with $p_{\mathrm{T}} > 3\,\mathrm{GeV}$. The jet thresholds have had to increase to lower the trigger rates. For trigger version v14.20 and v14.21, ISO_MUON(1,3,loose) was also used at Level 3. This filter requires one or more loose isolated muons with $p_{\mathrm{T}} > 3\,\mathrm{GeV}$. The isolation is determined with respect to the transverse energy in the calorimeter in a cone around the track, which has to be smaller that $8\,\mathrm{GeV}$.

Trigger Name	Level 1 Condition	Level 2 Condition	Level 3 Condition
EM15_2JT15	CEM(1,10)CJT(2,5)	EM(0.85,10)JET(2,10)	ELE_LOOSE_SH_T(1,15)_JET(2,15)
E1_SHT15_2J20	CEM(1,11)	None	ELE_NLV_SHT(1,15)_JET(2,20)
E1_SHT15_2J_J25	CEM(1,11)	L2CALEM(15,x)	ELE_NLV_SHT(1,15)_JET(1,25)JET(2,20)
E1_SHT15_2J_J30	CEM(1,11)	L2CALEM(15,x)	ELE_NLV_SHT(1,15)_JET(1,30)JET(2,20)
E1_SHT15_2J_J25	CEM(1,12)	L2CALEM(15,x)	ELE_NLV_SHT(1,15)_JET(1,25)JET(2,20)

Table 6.4: Definitions of triggers used in the electron channel.

Trigger Name	Level 1 Condition	Level 2 Condition	Level 3 Condition
MU_JT20_L2M0	mu1ptxatxx_CJT(1,5)	MUON(1,med)	JET(1,20)
MU_JT25_L2M0	mu1ptxatxx_CJT(1,3)	MUON(1,med)JET(1,10)	JET(1,25)
MUJ2_JT25	mu1ptxatxx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,25)
MUJ2_JT25_LM3	mu1ptxatlx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,25)MUON(1,3,loose)
MUJ2_JT30_LM3	mu1ptxatlx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,30)MUON(1,3,loose)
MUJ1_JT25_LM3	mu1ptxatlx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,25)MUON(1,3,loose)
MUJ1_JT25_ILM3	mu1ptxatlx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,25)MUON(1,3,loose)
			ISO_MUON(1,3,loose)
MUJ1_JT35_LM3	mu1ptxatlx_CJT(1,5)	MUON(1,med)JET(1,8)	JET(1,35)MUON(1,3.,loose)

Table 6.5: Definitions of triggers used in the muon channel.

To simplify trigger modeling we explicitly make the requirement offline that the reconstructed muon in our data set has fired the muon part of the triggers

6.3.3 Trigger Efficiencies

These lepton+jets triggers allows for a lower lepton threshold compared to single-lepton triggers while still maintaining straightforward modeling. The average trigger efficiency has been 87% for s-channel signal for both electron and muon channels, 86% for t-channel signal in the electron channel, and 82% for t-channel signal in the muon channel, The efficiencies were determined by modeling the triggers on MC events, to be explained in Sec. 6.6. With increasing luminosity, the threshold for the jet has had to be increased, particularly in the muon channel, thereby reducing the efficiency to trigger on the signal. This effect is shown in Tables 6.6 and 6.7. Therefore, subsequent versions of this analysis will be including more triggers in the future, in particular, an "OR" of single lepton triggers and lepton+jet triggers.

Trigger Version	s-channel	t-channel
v8	92%	92%
v9	91%	91%
v10	91%	91%
v11	91%	91%
v12	86%	85%
v13.00–v13.23	87%	86%
v13.30–v13.90	86%	85%
v14	88%	87%
Overall average	87%	86%

Table 6.6: Average electron-channel trigger efficiencies for single top events after selection.

Trigger Version	s-channel	t-channel
v8	96%	94%
v9	93%	90%
v10	91%	88%
v11	92%	88%
v12	91%	88%
v13.00–v13.11	91%	87%
v13.20–v13.23	91%	87%
v13.30–v13.90	86%	81%
v14.00–v14.10	92%	89%
v14.20–v14.21	92%	89%
v14.30–v14.90	80%	72%
Overall average	87%	82%

Table 6.7: Average muon-channel trigger efficiencies for single top events after selection.

6.4 Event Selection

The analysis strategy is to make the event selection cuts loose and then to depend on the multivariate analysis to discriminate the signal from background. The idea is that the multivariate techniques are more powerful than simple cuts, so they should be allowed to do the most work. In addition, however, it is important to filter out events that are difficult to model. The "triangle" cuts, explained below, provide an example, since they are designed to remove poorly measured events.

The event selection common for both the electron and muon channels makes the following requirements:

- Good quality (for data): the data quality must have been declared good for tracking, calorimetry, and muon system, and there must not be excessive noise in the calorimeter.

- Pass trigger (for data): the event must have fired one of the triggers described in Sec. 6.3. Additionally, offline electrons and muons in the data must be matched to the lepton object at all three trigger levels that fired in the appropriate trigger for that run period. This requirement allows better modeling of the trigger behavior in Monte Carlo.
- Good primary vertex: $|z_{\mathrm{PV}}| < 60\,\mathrm{cm}$ with at least three tracks attached, where z_{PV} is the z coordinate of the primary vertex.
- Missing transverse energy $15 < \not{E}_{\mathrm{T}} < 200\,\mathrm{GeV}$. The upper criteria serves to remove badly mismeasured (usually μ+jets) events.
- Two or three jets that pass standard jet ID requirements (see Sec. 5.4), each with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 3.4$. Also available for the other multivariate techniques and for cross checks are events with one jet and with four jets, though the ME analysis does not make use of these for extracting the single top measurement.
- The leading jet is required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.5$.
- The second leading jet is required to have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$.
- There must be fewer than three jets in the event that fail the standard jet ID requirements. This is an event quality cut to remove noisy events.
- Jet triangle cut: Events must have $(\Delta\phi(\text{leading jet}, \not{E}_{\mathrm{T}}), \not{E}_{\mathrm{T}})$ fall below the line that goes through the points $(1.5\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(\pi\,\mathrm{rad}, 35\,\mathrm{GeV})$; see Fig. 6.3. $\Delta\phi(\text{leading jet}, \not{E}_{\mathrm{T}})$ is defined to be the (positive) angle between the leading jet and $\not{E}_{\mathrm{T}}$. The regions being cut out are when the $\not{E}_{\mathrm{T}}$ is opposite a jet, which suggests that the $\not{E}_{\mathrm{T}}$ is poorly measured because, for example, the jet's energy is poorly measured.

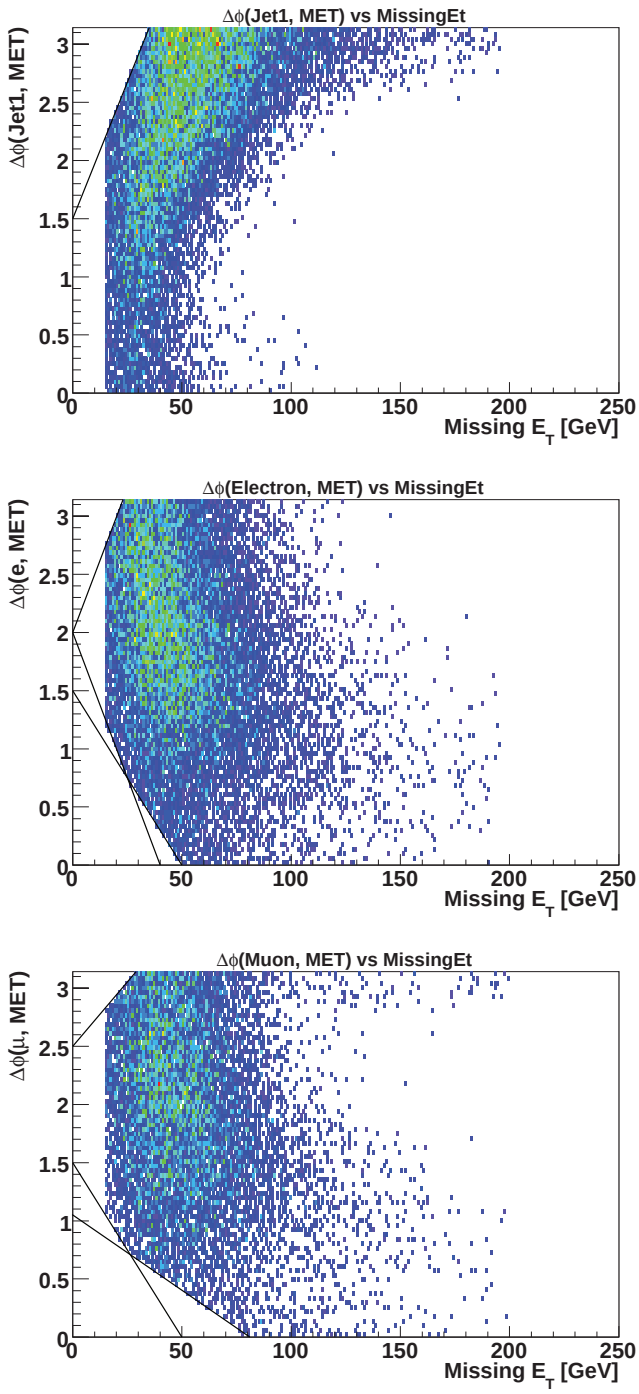


Figure 6.3: The triangle cuts for jets, electrons, and muons. The cuts are indicated by the diagonal lines in the left upper and lower corners. The blue-green areas are where s-channel single top falls after applying the event selection, including the $\not{E}_{\mathrm{T}}$ and triangle cuts.

The electron channel has, in addition to the common requirements, the following specific selection cuts:

- Exactly one electron of quality *top_tight* (see Sec. 5.5) with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 1.1$.
- Electron coming from the primary vertex: $|\Delta z(e, \mathrm{PV})| < 1\,\mathrm{cm}$.
- Second electron veto: No second electron of quality *top_loose* with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ with any η. The purpose of this cut is to remove most of the $Z/\gamma^*(\rightarrow e^+e^-)$+jets and $t\bar{t} \rightarrow e^+e^-$+jets backgrounds.
- No tight muon with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.0$. This requirement is mainly to keep the electron and muon channels orthogonal, but it also functions to remove the $t\bar{t} \rightarrow \mu e$+jets (and a bit of $Z/\gamma^*(\rightarrow \tau^+\tau^-)$+jets) backgrounds.
- Electron triangle cuts: Events must have $(\Delta\phi(e, \not{E}_{\mathrm{T}}), \not{E}_{\mathrm{T}})$ fall
  - **below** the line that goes through the points $(2\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(\pi\,\mathrm{rad}, 24\,\mathrm{GeV})$
  - **above** the line that goes through the points $(2\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(0\,\mathrm{rad}, 40\,\mathrm{GeV})$
  - **above** the line that goes through the points $(1.5\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(0\,\mathrm{rad}, 50\,\mathrm{GeV})$;

  see Fig. 6.3. The regions being cut out are when the $\not{E}_{\mathrm{T}}$ is opposite or along the direction of the electron (in ϕ), which suggests that the $\not{E}_{\mathrm{T}}$ is poorly measured.

The muon channel has, in addition to the common requirements, the following specific selection cuts:

- Exactly one tight muon (see Sec. 5.6) with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.0$. This cut also removes most of the $Z/\gamma^*(\rightarrow \mu^+\mu^-)$+jets and $t\bar{t} \rightarrow \mu^+\mu^-$+jets backgrounds.

- Muon coming from the primary vertex: $|\Delta z(\mu, \mathrm{PV})| < 1\,\mathrm{cm}$.
- No electron of quality *top_tight* with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.5$. This is mostly to keep the electron and muon channels orthogonal, but it also functions to remove $t\bar{t} \to \mu e$+jets (and a bit of $Z/\gamma^*(\to \tau^+\tau^-)$+jets) background.
- Muon triangle cuts: Events must have $(\Delta\phi(\mu, \not{E}_{\mathrm{T}}), \not{E}_{\mathrm{T}})$ fall
  - **below** the line that goes through the points $(2.5\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(\pi\,\mathrm{rad}, 30\,\mathrm{GeV})$
  - **above** the line that goes through the points $(1.5\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(0\,\mathrm{rad}, 50\,\mathrm{GeV})$
  - **above** the line that goes through the points $(1.1\,\mathrm{rad}, 0\,\mathrm{GeV})$ and $(0\,\mathrm{rad}, 80\,\mathrm{GeV})$;

  see Fig. 6.3. The regions being cut out are when the $\not{E}_{\mathrm{T}}$ is opposite or along the direction of the muon (in ϕ), which suggests that the $\not{E}_{\mathrm{T}}$ is poorly measured.

6.5 Correcting the Monte Carlo Model Efficiencies and Resolutions

The Monte Carlo samples for single top, W+jets, and $t\bar{t}$ have to pass the event selection described above. However, the efficiency to reconstruct, identify, and select electrons and muons in the MC is higher than in data, so correction factors are needed to correct for these differences. For the leptons and jet, the energy resolution is better for MC than for data, so the MC sample is further smeared. These corrections are further explained below.

6.5.1 Electron Corrections

We correct each MC event in the electron channel with a scale factor that accounts for the differences in electron cluster finding and identification efficiency between data and MC events. The correction factor is divided into two parts: preselection and post-preselection. Preselection refers to the basic electron criteria that is common among many electron quality definitions: ID, electromagnetic fraction, and isolation. The corresponding scale factor is parametrized as a function of η^{det}, and is shown in Fig 6.4(a). The post-preselection criteria, which is unique for our particular electron quality definition, consists of the H-matrix cut, the track matching requirements, and the likelihood cut. The corresponding scale factor is parametrized as a function of η^{det} and ϕ^{det}, and is shown in Fig 6.4(b). The overall correction factor thus becomes:

$$\text{cf}_{e-\text{ID}} = \frac{\varepsilon^{\text{Data}}_{\text{Presel}}}{\varepsilon^{\text{MC}}_{\text{Presel}}} \times \frac{\varepsilon^{\text{Data}}_{\text{PostPresel}}}{\varepsilon^{\text{MC}}_{\text{PostPresel}}}$$

These factors were derived from $Z \to e^+e^-$ data and simulated events using a tag and probe technique, as described in more detail in Ref [144]. To measure the preselection efficiency, events in data and $Z \to e^+e^-$ MC were selected to have one tight electron that has fired a single-electron trigger (if data), the *tag*, plus an isolated good quality track opposite of the electron, the *probe*. For the electron to be considered tight it has to pass a tight H-matrix cut and a tight track match ($\chi^2_{hm} < 12$, $P(\chi^2_{\text{spatial}}) > 0.01$), and the trigger requirement is to minimize trigger biases in the measured quantity—the event will have been triggered on regardless of the probe. The *probe* track had to be away from any muon, and the invariant mass of the electron and the track had to be greater than 60 GeV. The efficiency was calculated by determining the probability to construct an electron cluster where

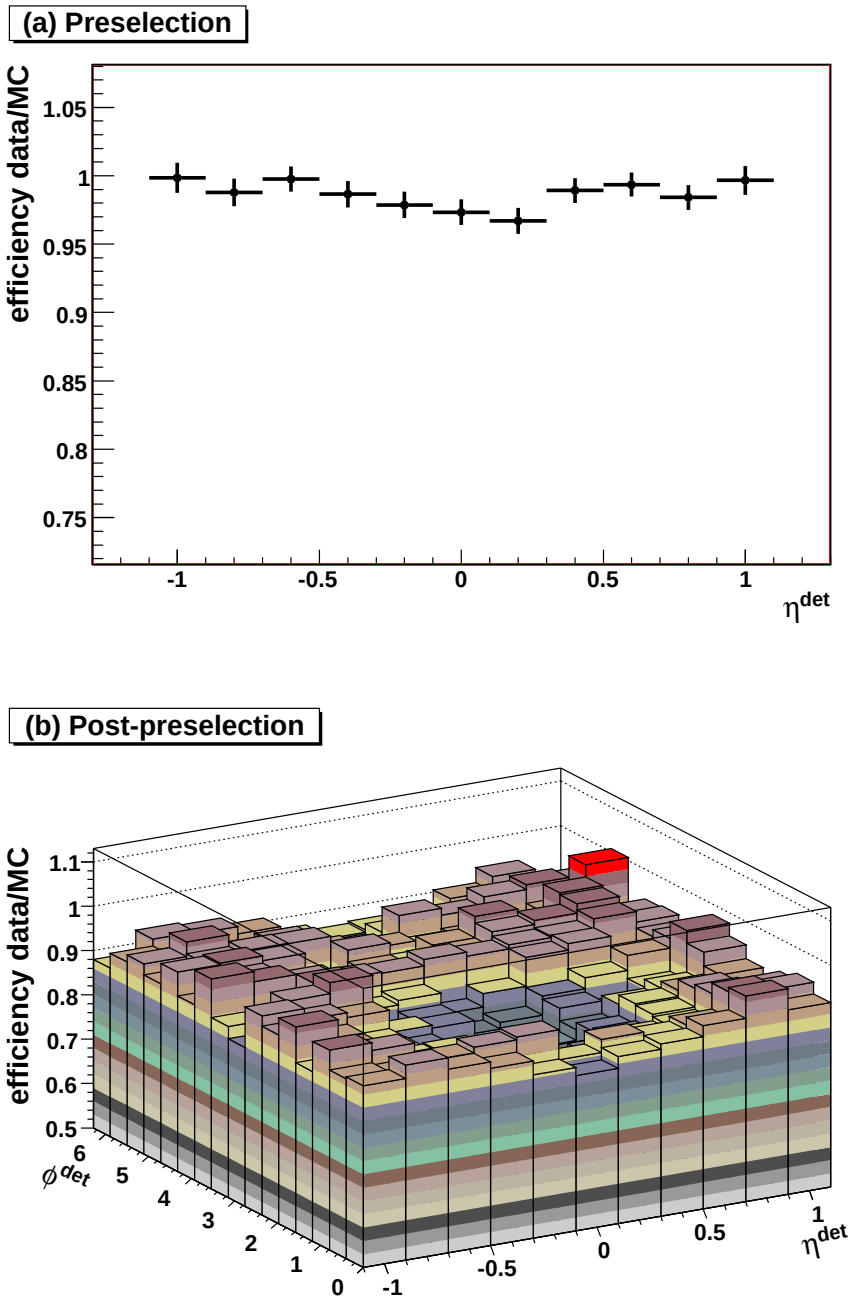


Figure 6.4: The (a) preselection and (b) post-preselection correction factors for electrons.

the *probe* track was pointing. The background subtraction was determined by using cases where the *tag* electron track and the *probe* track have the same sign.

To measure the post-preselection efficiency, events in data and $Z \rightarrow e^+e^-$ MC were selected to have two electrons, one being the *tag*, defined in this case to be of *top_tight* quality and to have fired a single-electron trigger (if data), and the other being the *probe*. The probability to pass the H-matrix requirement given a preselected electron was measured by taking a *probe* electron that passed preselection and seeing how often it passed the H-matrix requirement. The probability to pass the track requirement given a preselected electron that has passed the H-matrix was measured by taking a *probe* electron that passed preselection and H-matrix and seeing how often it passed the track matching requirement. The probability to pass the likelihood given that it has passed preselection, H-matrix, and track matching was calculated in an analogous manner. The post-preselection efficiency can then be determined by putting these partial probabilities together. Background subtraction was done by fitting a Voigt function plus exponential around the Z invariant mass peak. In all cases, if both electrons were able to satisfy the *tag* requirement, each was separately used as the *tag*.

The electron energies in the MC samples are smeared so that they match the resolution of data. The smearing parameters were determined by making the invariant mass of the two electrons from a $Z \rightarrow e^+e^-$ data and MC sample have the same distribution.

6.5.2 Muon Corrections

We correct each MC event in the muon channel with a scale factor that accounts for the differences in muon identification efficiency between data and MC events. The

correction factor is divided into three parts: the muon type and quality, the global track match, and the isolation. The muon type and quality refers to the $|\text{nseg}| = 3$ *medium* criteria, and the corresponding correction factor is parametrized as a function of η^{det} and ϕ. The correction factor for the *medium* quality global track match is parametrized as a function of track-z and η_{CFT}. The isolation correction factor is a function of η. The correction factors and their parametrizations have been derived from $Z \to \mu^+\mu^-$ data and MC, as described in Ref. [150]:

$$\text{cf}_{\mu-\text{ID}} = \frac{\varepsilon^{\text{Data}}_{\text{MediumID}}}{\varepsilon^{\text{MC}}_{\text{MediumID}}} \times \frac{\varepsilon^{\text{Data}}_{\text{TrackMatch}}}{\varepsilon^{\text{MC}}_{\text{TrackMatch}}} \times \frac{\varepsilon^{\text{Data}}_{\text{TightIsol}}}{\varepsilon^{\text{MC}}_{\text{TightIsol}}}$$

The muon energies in the MC samples are also smeared to match the resolution of the data. Because the solenoid current was reduced after the shutdown that took place in 2005, the muon resolution became 15% worse. Therefore, there are different smearing parameters for pre-shutdown and post-shutdown data. We smear the MC samples to have the same pre-/post-shutdown ratio as the data.

6.5.3 Jet Corrections

MC jets do not need to have a correction factor applied to them, but they do need to have their energy corrected and smeared. As for data jets, MC jets have a jet energy scale applied to them. The energy then needs to be smeared [178], and only then can the jet p_{T} cuts can be applied to them.

6.6 Modeling the Trigger for Monte Carlo Events

The trigger effects need to be modeled in MC events in order for them to accurately represent data events. We do that by assigning a weight to each event based on how likely it would be triggered [179, 180, 181]. The triggering efficiencies for the leptons and jets are measured in data and parametrized. The individual object parametrizations are then combined into a probability for each MC event that that particular event would have been triggered.

The probability for an electron to fire a trigger was measured by a tag and probe technique in the same way that the post-preslection efficiencies were measured. The trigger efficiency that is needed is the probability for an electron to fire the electron part of the trigger given that it has passed the selection requirement (i.e., be of quality *top_tight*). Data events were selected to have two *top_tight* electrons, and one, the *tag*, must have fired a single-electron trigger to minimize trigger biases. The probability for the *probe* electron to fire the electron part of the electron+jet trigger gives the desired efficiency. If both electrons satisfied the *tag* requirements, then both were separately treated as the *tag*. The efficiency was fitted and parametrized as a function of p_{T}. Additionally, the electron can fire part of the jet trigger. This is measured by seeing how often the probe electron fired the single-jet part of the jet trigger given that it had fired the electron part of the trigger. For v8–v11, this was found to be roughly 100%, so it is taken to be 100% for those trigger versions with no further parametrization. However, from v12 on, the jet part of the trigger had a higher E_{T} threshold than the electron part, so the efficiency was fitted and parametrized as a function of the electron p_{T}. Figure 6.5 shows the trigger turn-on curves for electrons to fire the electron part and the jet part of E1_SHT15_2J_J30, which has the highest integrated luminosity of all the e+jets triggers we used.

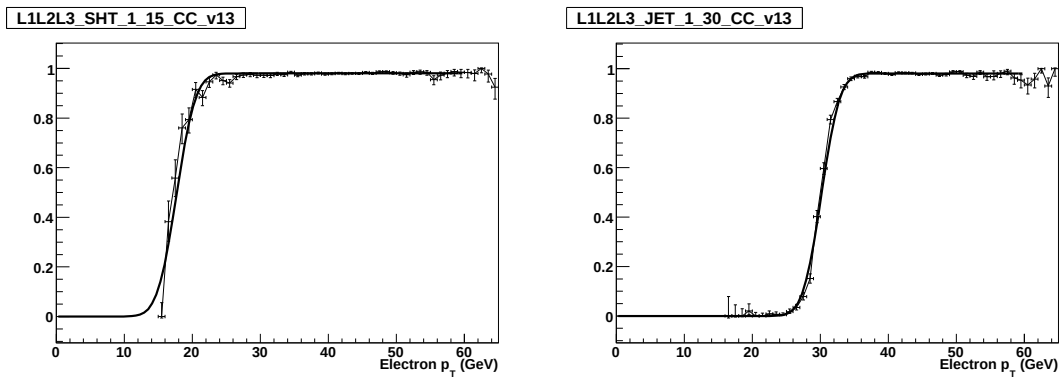


Figure 6.5: The trigger turn-on curves for electrons to fire the electron part (left) and the jet part (right) of E1_SHT15_2J_J30.

The muon trigger efficiency was similarly measured and parametrized using $Z \to \mu^+\mu^-$ events, this time as a function of η^{det} and ϕ. The jet trigger efficiencies were calculated using a sample of muon-triggered jet events. The trigger efficiencies were calculated by seeing how often there were matching trigger objects to jets that passed the offline reconstruction values. The efficiency was parametrized in p_{T} and η^{det}.

6.7 Normalizing the W+jets and Multijets Backgrounds to Data

The amount of fake lepton background in our selected data samples can be estimated using the matrix method (MM) [182], which proceeds as follows. Two data samples are defined, one with a loose lepton requirement, and one with a tight lepton requirement. The definition of the tight requirement must be such that all tight leptons are also loose leptons. We need to externally have an estimate of the incremental efficiency for loose leptons to pass the tight lepton requirement, which we label $\varepsilon^{\text{real}}$. This efficiency does not need to be a constant; it can be a function of the lepton

variables, such as p_{T} and η, and also event variables, like instantaneous luminosity. We also similarly need to externally have an estimate of the fake rate, that is, the efficiency for fake leptons that pass the loose quality to also pass the tight quality. This fake rate is labeled as $\varepsilon^{\mathrm{fake}}$.

Defining N_{loose} to be the number of events that fulfill the loose lepton criteria and N_{tight} the number that fulfill the tight criteria, the number of real and fake leptons in the loose sample, $N_{\mathrm{loose}}^{\mathrm{real}}$ and $N_{\mathrm{loose}}^{\mathrm{fake}}$, can be determined by solving this system of two equations and two unknowns:

$$\begin{aligned} N_{\mathrm{loose}} &= N_{\mathrm{loose}}^{\mathrm{real}} + N_{\mathrm{loose}}^{\mathrm{fake}} \\ N_{\mathrm{tight}} &= \varepsilon^{\mathrm{real}} N_{\mathrm{loose}}^{\mathrm{real}} + \varepsilon^{\mathrm{fake}} N_{\mathrm{loose}}^{\mathrm{fake}} \end{aligned} \tag{6.2}$$

Because the lepton efficiency and fake rate are a function of the event variables, logically we can think of solving the system of equations separately in different bins defined such that within a bin the ε values are constant. Then the number of real and fake electron events in the tight sample in bin i can be estimated as $N_{\mathrm{tight}\ i}^{\mathrm{real}} = \varepsilon_i^{\mathrm{real}} N_{\mathrm{loose}\ i}^{\mathrm{real}}$ and $N_{\mathrm{tight}\ i}^{\mathrm{fake}} = \varepsilon_i^{\mathrm{fake}} N_{\mathrm{loose}\ i}^{\mathrm{fake}}$, and the estimates of the total number of real and fake leptons in the tight sample can be estimated as $N_{\mathrm{tight}}^{\mathrm{real}} = \sum_i N_{\mathrm{tight}\ i}^{\mathrm{real}}$ and $N_{\mathrm{tight}}^{\mathrm{fake}} = \sum_i N_{\mathrm{tight}\ i}^{\mathrm{fake}}$.

In the electron channel the loose electron quality criteria is *top_loose*, and the tight is *top_tight*. The tight sample is the sample that passes all the criteria in Sec. 6.4, while the loose sample changes the "exactly one electron of quality *top_tight* with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 1.1$" criteria to "exactly one electron of quality *top_loose* with $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 1.1$." For muons the loose and tight criteria are defined in Sec. 5.6. Similarly to the electron case, the tight sample is is the sample that passes all the criteria in Sec. 6.4, while the loose sample changes the "exactly

one tight muon with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.0$" to "exactly one loose muon with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2.0$."

The real electron efficiency is calculated as

$$\varepsilon^{\mathrm{real}} = \varepsilon^{\mathrm{real}}_{\mathrm{MC}} \times \frac{\varepsilon^{\mathrm{real}}_{Z\to e^+e^-\ \mathrm{data}}}{\varepsilon^{\mathrm{real}}_{Z\to e^+e^-\ \mathrm{MC}}}. \tag{6.3}$$

The $\varepsilon^{\mathrm{real}}_{\mathrm{MC}}$ variable is the real electron efficiency measured in our MC samples scaled according to their approximate proportions in the loose sample (which is similar to the proportions in the tight sample, given in Table 6.10). MC truth is used to measure this efficiency. The term on the right is the incremental likelihood[1] data/MC correction factor, as explained in Sec. 6.5.1. The only difference is that the correction factors, since they are applied on MC before the trigger model, need to be as free from trigger biases as possible. On the other hand, we apply the MM to triggered data, so we want ε to model the performance of triggered data. Therefore, when doing tag and probe in data, the *probe* electron is required to match an electron in the trigger system. The difference is measurable but only around 1% for $\varepsilon^{\mathrm{real}}$, which is smaller than the systematic uncertainties associated with the method. The difference is more significant for $\varepsilon^{\mathrm{fake}}$, however. We parameterize $\varepsilon^{\mathrm{real}}$ in the electron channel as a function of p_{T} and η^{det}, and one-dimensional projections of the two-dimenisonal parametrization that we use are given in Fig. 6.6.

The muon efficiencies are measured using tag and probe with $Z\to\mu^+\mu^-$ events. We parameterize $\varepsilon^{\mathrm{real}}$ in the muon channel as a function of p_{T} and the number of jets in the event, and one-dimensional projections of the this parametrization are given in Fig. 6.7.

The lepton fake rates are calculated using our loose data sample but with $\not{E}_{\mathrm{T}} <$

[1]The difference between *top_loose* and *top_tight* is that *top_tight* has a $L > 0.85$ cut.

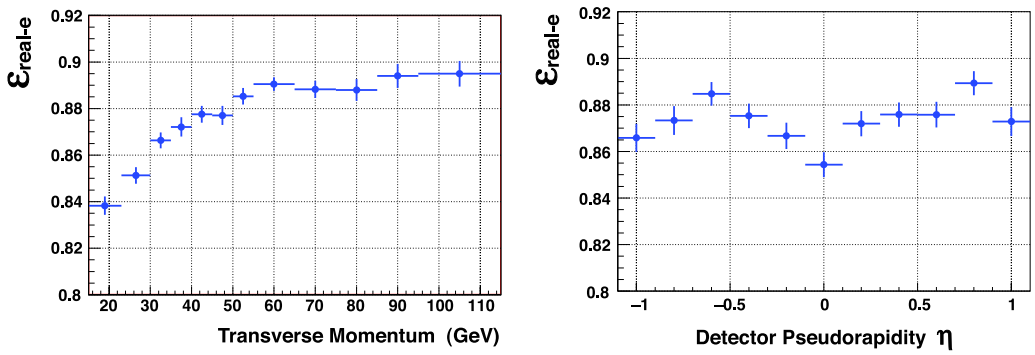


Figure 6.6: One-dimensional projections in p_{T} and η^{det} of the electron efficiencies used in the matrix method.

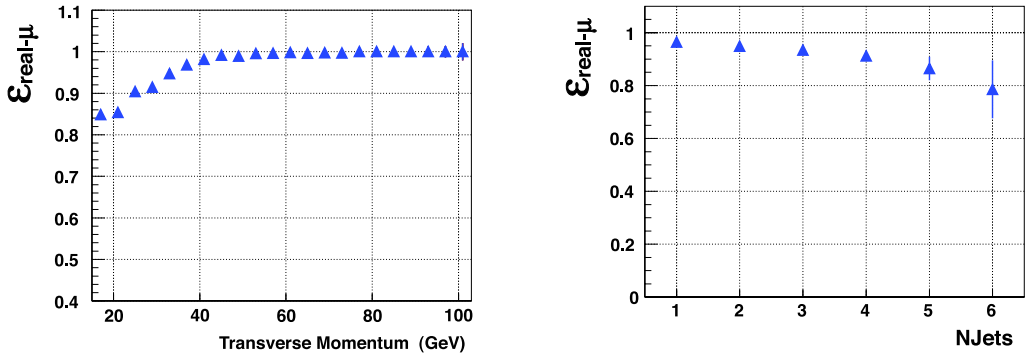


Figure 6.7: One-dimensional projections in p_{T} and number of jets of the muon efficiencies used in the matrix method.

No. jets	v8–v11	v12	v13.0–v13.11	v13.20–v13.90	v14
1	$(11.2 \pm 0.5)\%$	$(17.9 \pm 0.6)\%$	$(18.7 \pm 1.2)\%$	$(19.1 \pm 0.6)\%$	$(18.5 \pm 0.6)\%$
2	$(12.8 \pm 1.0)\%$	$(19.2 \pm 1.0)\%$	$(18.8 \pm 2.2)\%$	$(19.4 \pm 1.1)\%$	$(22.0 \pm 1.2)\%$
3	$(13.6 \pm 1.5)\%$	$(19.5 \pm 1.6)\%$	$(19.8 \pm 3.4)\%$	$(19.2 \pm 1.6)\%$	$(19.4 \pm 1.7)\%$
4	$(10.0 \pm 2.8)\%$	$(15.5 \pm 2.9)\%$	$(20.9 \pm 8.6)\%$	$(17.7 \pm 3.3)\%$	$(20.7 \pm 3.7)\%$

Table 6.8: The electron fake rates used in the matrix method as a function of trigger version and number of jets.

10 GeV instead of $\not\!\!E_{\mathrm{T}} > 15$ GeV. The low $\not\!\!E_{\mathrm{T}}$ region is dominated by multijet background, and it is assumed that the fake rate is independent of the $\not\!\!E_{\mathrm{T}}$ of the event. The lepton fake rates are determined by seeing what fraction of the loose leptons in the above samples fulfill the tight lepton criteria.

Our study of the electron fake rate found no dependence on the lepton p_{T} or lepton η, but dependence on the trigger version and jet multiplicity. As mentioned in Sec. 6.3, the Level 3 shower shape requirement became significantly tighter starting with trigger version v12, making the electrons that are triggered look much more like real electrons. Therefore, the incremental fake rate for the likelihood requirement became higher (though the total fake rate is smaller). A summary of the results is presented in Table 6.8.

For the muon fake rate, we performed a similar set of studies and found no need to parametrize as a function of the trigger version, but we did find a small dependence on the muon η^{det}. Although there is also a dependence on the muon p_{T}, parametrizing on that variable did not change the overall results. Therefore, the muon fake rate is parametrized as a function of η^{det} and the number of jets in the event, as shown in Fig. 6.8.

The number of events in the loose and tights samples, as well as the resulting estimates of the number of real and fake lepton events, are given in Table 6.9. The

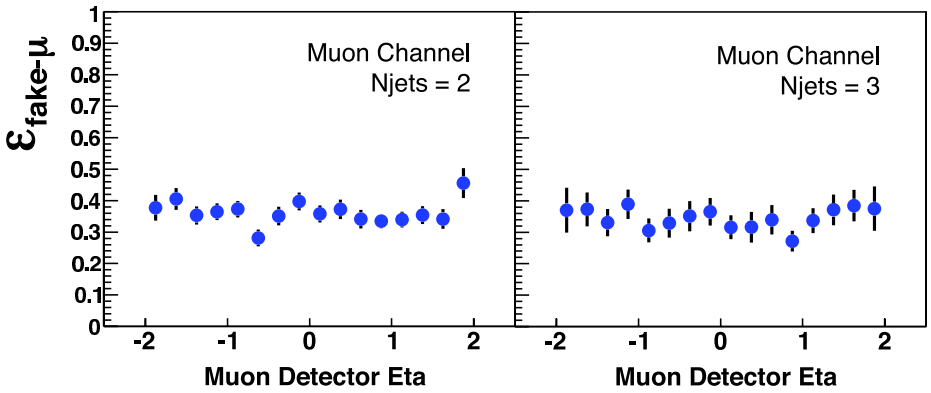


Figure 6.8: The muon fake rates used in the matrix method as a function of η^{det} and number of jets.

	Electron Channel		Muon Channel	
	2 jets	3 jets	2 jets	3 jets
N_{loose}	15,213	7,118	7,092	3,054
N_{tight}	8,220	3,075	6,432	2,590
$N_{\text{tight}}^{\text{real}}$	6,787	2,215	6,105	2,369
$N_{\text{tight}}^{\text{fake}}$	1,433	860	329	223

Table 6.9: The number of events in the loose and tight samples that are given as input to the matrix method, and the resulting estimates for the number of real and fake lepton events in the tight sample..

MM is run separately for each jet bin of the electron and muon channels. To model the multijet background, in each jet bin we take the loose sample and subtract out the tight sample, thus creating a sample orthogonal to the tight sample. This orthogonal sample, scaled to $N_{\text{tight}}^{\text{fake}}$ in each bin, becomes our multijet background model. This sample has a small $W+$jets contamination, but because the contamination is within the uncertainties we apply to the model, we do not correct for it.

We scale the W+jets sample in each channel so that

$$N_{W+\text{jets}} = N^{\text{real}}_{\text{tight}} - N_{t\bar{t}} \tag{6.4}$$

where $N_{t\bar{t}}$ is normalized using the cross section given in Table 6.1. The relative fraction of heavy flavor production W+heavy flavor to W+light jets ratio is determined as follows. Let

$$N_{W+\text{jets}} = w_{\text{MM}} \left[N^{\text{ALPGEN}}_{W+\text{light jets}} + \alpha \left(N^{\text{ALPGEN}}_{Wb\bar{b}} + N^{\text{ALPGEN}}_{Wc\bar{c}} \right) \right] \tag{6.5}$$

where the N^{ALPGEN} counts are the unscaled yields using the ALPGEN cross sections, w_{MM} is the matrix method weight needed to make $N_{W+\text{jets}}$ agree with the value set in Eq. 6.4, and α is an extra scaling factor given to the W+heavy flavor samples. $Wb\bar{b}$ and $Wc\bar{c}$ are scaled by the same factor.

Up to this point we have not applied b-tagging to the data, and the yields are made to exactly match the data by the MM. We can choose different values for α in Eq. 6.5 and if we compensate with w_{MM}, we can still satisfy Eq. 6.4. We can select the combination of α and w_{MM} only with more input points, and for that, we apply b-tagging to the data (directly) and MC (using TRFs), to be explained in Sec. 6.8. We use the zero-tagged data set yields, where the signal content is negligible, to determine the heavy flavor fraction. The zero-tagged data set is not used elsewhere within the analysis. We scale the heavy flavor fraction (and adjust w_{MM} to keep Eq. 6.4 satisfied) so that the yields match in the zero-tagged channel, and the measured α is shown in Fig. 6.9. We take $\alpha = 1.5$ and assign a 30% uncertainty.

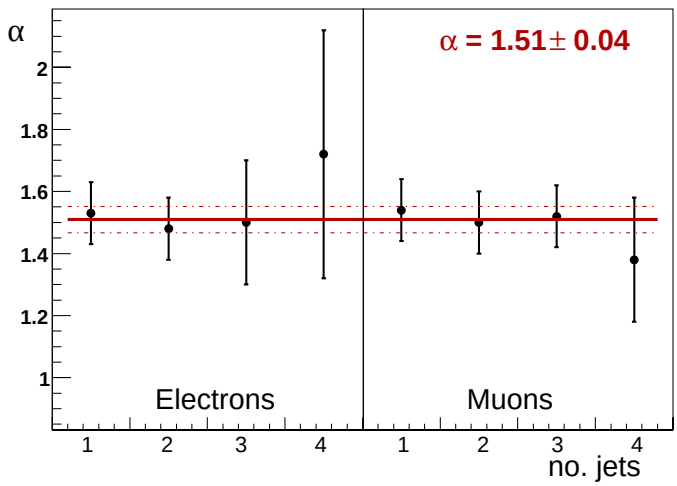


Figure 6.9: The values of α in the zero-tag sample for electron and muon channel in each jet multiplicity.

6.8 b-Tagging

Fig. 6.1 strongly makes the case for the benefits of b-tagging to the single-top analysis. In the two-jet bin for s-channel production, requiring one b-tag changes the signal to background ration from 1:1600 to 1:55, and two b-tags makes it even purer, 1:12, though with a smaller yield. Similarly in the two-jet bin for t-channel production, requiring one b-tag changes the signal to background ration from 1:520 to 1:32. We therefore require for events to enter our analysis that they have one or two b-tags.

We use the NN b-tagging algorithm at the TIGHT operating point [158, 159]. We apply the tagger to data, but we apply a tag-rate function (TRF) to the MC, which is a parametrization of the b-tagging performance, taking into account the jet flavor and jet variables. Applying the tagger directly to the MC results in too high of a b-tag efficiency prediction. The steps are explained in more detail below.

6.8.1 Taggability

For data events, before applying the b-tagging algorithm, we first determine whether the jet is taggable, which is a criteria independent of the particular b-tagging algorithm[2]. The specific criteria is that the (calorimeter) jet is required to be within $\Delta R < 0.5$ to a track-jet, as defined in Sec. 5.8. It is clearly a requirement for the tracking-based b-tagging algorithms to be able to associate a set of tracks with a jet.

For MC events, we do not apply the criteria directly but instead apply a taggability rate function, which is a parametrization of the efficiency for a jet to be taggable, parametrized as a function of jet p_{T}, jet η and the primary vertex z. The taggability rate functions have been derived on the loose data samples, as defined for the matrix method, to keep enough statistics for the measurement. The result of applying the functions to the tight samples are given in Fig. 6.10 for the electron channel and in Fig. 6.11 for the muon channel. The observed and predicted results agree within the uncertainty.

6.8.2 b-Tagging

After applying the taggability criteria, the next step is to apply the actual b-tagging. For data, we run the b-tagging algorithm. For MC, we apply a flavor-dependent parametrized function that gives the probability that an event is tagged.

The procedure to calculate the TRF is as follows. In a heavily b-enriched data sample, with one jet tagged with the JLIP requirement, and the probe jet containing a muon within it (for SLT tagging), we apply the NN tagger on the probe jet to determine the semileptonic data tagging efficiency, $\varepsilon_{b\to\mu}^{\mathrm{DATA}}$. In an admixture of $Z \to b\bar{b}$

[2]This considers only lifetime-based tagging, not SLT.

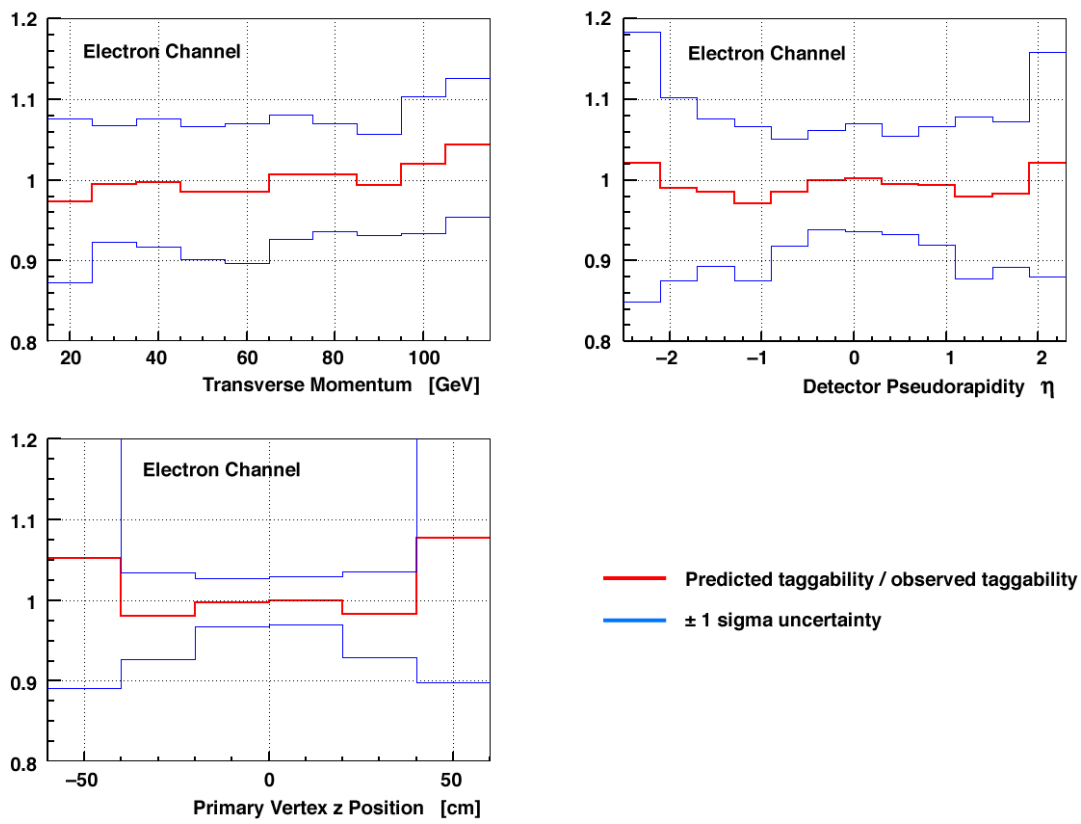


Figure 6.10: The ratio of the predicted taggability rate function over the observed taggability on the electron channel tight data sample.

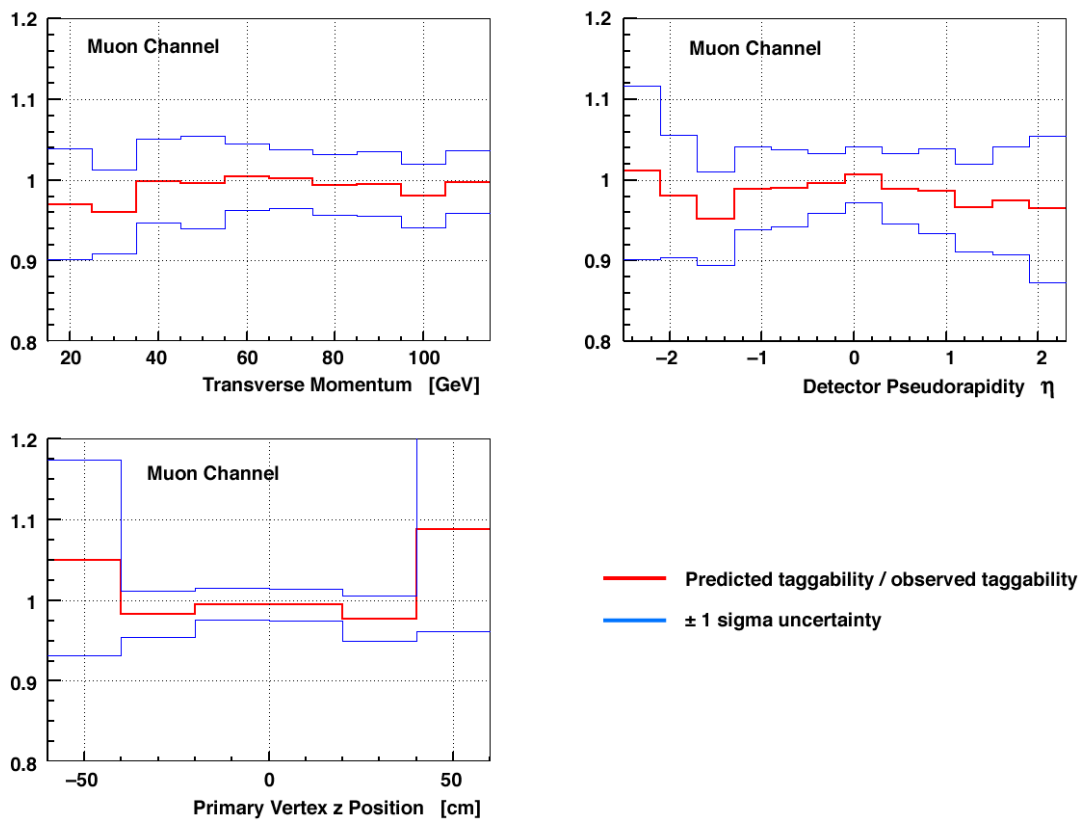


Figure 6.11: The ratio of the predicted taggability rate function over the observed taggability on the muon channel tight data sample.

and $t\bar{t}$ MC events we determine $\varepsilon^{MC}_{b\to\mu}$, in order to calculate the data/MC scale factor:

$$SF_b = \frac{\varepsilon^{\mathrm{DATA}}_{b\to\mu}}{\varepsilon^{MC}_{b\to\mu}}. \tag{6.6}$$

We assume that this scale factor is unchanged for b-jets without a muon and also for c-jets. Therefore,

$$\varepsilon_b = \varepsilon^{\mathrm{MC}}_b \times SF_b \tag{6.7}$$

$$\varepsilon_c = \varepsilon^{\mathrm{MC}}_c \times SF_b. \tag{6.8}$$

The MC samples used for the $\varepsilon^{\mathrm{MC}}_c$ calculation were $Z \to c\bar{c}$, $t\bar{t}$, and QCD samples. The measured ε_b and ε_c dependences are shown in Fig. 6.12 together with the direct tagger efficiency.

The probability to tag a light jet, ε_{light}, is calculated using the concept of a negative tag. In Sec. 5.8, it was mentioned how CSIP has a signed impact parameter significance and how SVT has a signed decay length significance. Real b-quarks are expected to have a positive significance, because, taking SVT as an example, the b-jet is on the side of the SV, not opposite it. When the SV is opposite of the jet, then the notion is that the SV is a result of mismeasurement. Since light quarks are not supposed to have a SV, the sign of the significance should average to zero. Therefore, ε_{light} can be determined by measuring the negative tag rate, $\varepsilon^{-}_{\mathrm{data}}$, in data and EM skims.

The probability so derived needs to be corrected, however, for the presence of

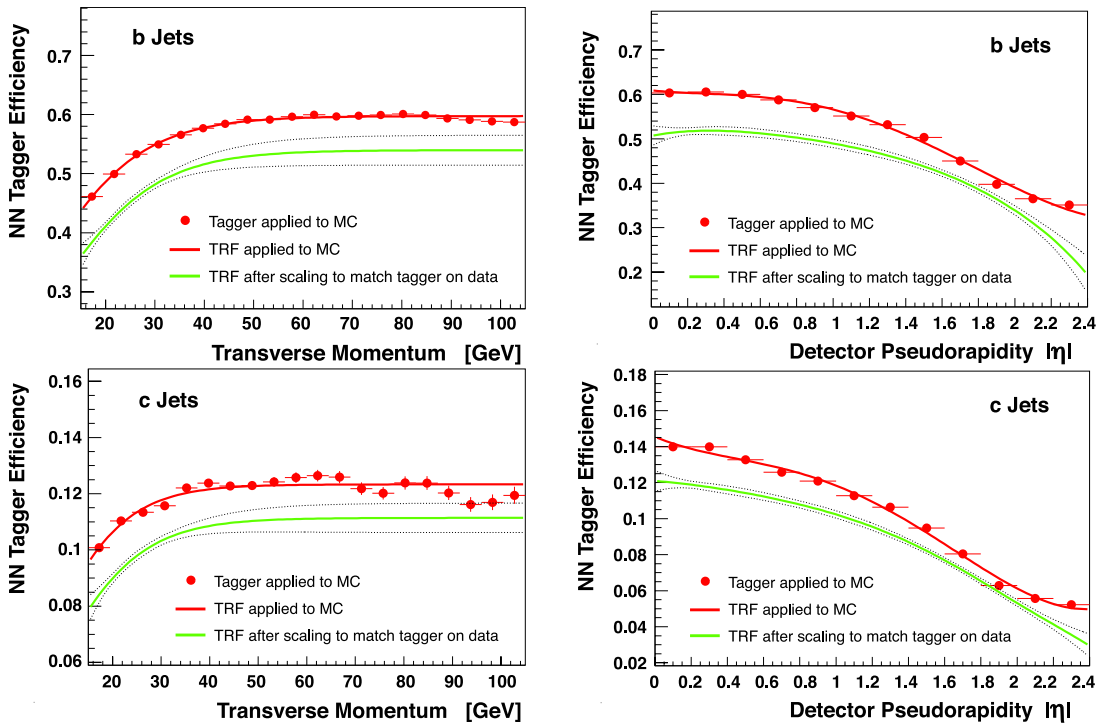


Figure 6.12: NN TIGHT tagger b-jet (upper row) and c-jet (lower row) efficiencies as a function of p_{T} (left column) and η (right column) in the inclusive b and c MC samples. The corresponding data tag-rate functions (ε_b and ε_c) are also displayed.

b-quarks and c-quarks in the samples, as follows:

$$\varepsilon_{\text{light}} = \varepsilon^{-}_{\text{data}} \times SF_{hf} \times SF_{ll} \tag{6.9}$$

$$SF_{hf} = \varepsilon^{-}_{\text{QCD light}} / \varepsilon^{-}_{\text{QCD all}} \tag{6.10}$$

$$SF_{ll} = \varepsilon^{+}_{\text{QCD light}} / \varepsilon^{-}_{\text{QCD light}}, \tag{6.11}$$

where the two scale factors are derived in MC. The scale factor, SF_{hf}, which is the number of negative tagged jets from light quarks divided by the total number of negative tagged jets in the QCD MC, serves to correct for heavy flavor contamination. The second scale factor, SF_{ll}, is to correct for long-lived hadron decays in light-quark jets, and it is defined as the number of positive tagged jets from light quarks divide by the number of negative tagged jets from light quarks in the QCD MC.

6.8.3 b-Tagging Event Weights and b-Jet Assignment Combinations

Given the three different efficiencies ε_α outlined above for $\alpha = b$, c and light jets, the probability to tag a jet of flavor α, can be expressed as the product of the taggability and the tagging efficiency:

$$\mathcal{P}_\alpha(p_{\text{T}}, \eta) = P^{\text{taggable}}(p_{\text{T}}, \eta) \times \varepsilon_\alpha(p_{\text{T}}, \eta) \tag{6.12}$$

Using the per jet probability, we can deduce the probability for an event to contain a given combination of tagged and untagged jets. When working with our MC samples, we permute over all these combinations, assigning the jets as tagged or untagged based on the particular combination, and give each particular permutation a b-tagging

weight based on the probability to have that combination. Thus, each MC event enters our analysis numerous times.

6.9 Event Yields

Tables 6.10, 6.11, 6.12, and 6.13 show the signal and background yields for before b-tagging, with zero b-tags, with one b-tagged jet, and with two b-tagged jets.

6.10 Cross Checks of the Background Model

Figure 6.13 show the distributions of W transverse mass, defined as

$$M_{W\mathrm{T}} = \sqrt{(p_{\mathrm{T}}^{\ell} + \not{E}_{\mathrm{T}})^2 - (p_x^{\ell} + \not{E}_x)^2 - (p_y^{\ell} + \not{E}_y)^2}, \tag{6.13}$$

separately for each lepton flavor and jet multiplicity, after the matrix method has been applied. More cross-check plots after selection before b-tagging, with one b-tag, and with two b-tags are given in Appendix B. Good agreement between data and the background model is seen everywhere.

	Electron Channel					Muon Channel				
	1 jet	2 jets	3 jets	4 jets	≥ 5 jets	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
Signals										
tb	4	14	7	2	0	3	10	5	1	0
tqb	9	27	14	5	1	6	20	11	3	1
Backgrounds										
$t\bar{t}\rightarrow ll$	9	35	28	10	4	5	27	22	8	3
$t\bar{t}\rightarrow l$+jets	2	26	103	128	67	1	14	71	99	43
$Wb\bar{b}$	659	358	149	42	5	431	312	161	47	10
$Wc\bar{c}$	1,592	931	389	93	10	1,405	1,028	523	131	21
Wjj	23,417	5,437	1,546	343	51	15,476	4,723	1,591	385	85
Multijets	1,691	1,433	860	256	86	498	329	223	58	10
Background Sum	27,370	8,220	3,075	874	223	17,816	6,434	2,592	727	172
Data	27,370	8,220	3,075	874	223	17,816	6,432	2,590	727	173

Table 6.10: Yields after selection and before b tagging.

	Electron Channel					Muon Channel				
	1 jet	2 jets	3 jets	4 jets	≥ 5 jets	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
Signals										
tb	3	5	2	1	0	1	4	2	1	0
tqb	6	16	7	2	1	4	11	6	2	0
Backgrounds										
$t\bar{t}\rightarrow ll$	5	14	11	4	1	3	10	8	3	1
$t\bar{t}\rightarrow l$+jets	2	13	43	47	24	1	7	28	35	15
$Wb\bar{b}$	471	222	92	27	3	300	187	97	28	6
$Wc\bar{c}$	1,511	856	352	84	9	1,341	953	475	117	19
Wjj	23,242	5,376	1,526	338	50	15,351	4,665	1,569	379	84
Multijets	1,655	1,365	808	236	78	481	302	198	49	7
Background Sum	26,886	7,845	2,832	735	165	17,476	6,124	2,375	610	131
Data	26,925	7,833	2,831	752	178	17,527	6,122	2,378	599	125

Table 6.11: Yields after selection for events with no b-tagged jets.

	Electron Channel					Muon Channel				
	1 jet	2 jets	3 jets	4 jets	≥ 5 jets	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
Signals										
tb	2	7	3	1	0	1	5	2	1	0
tqb	3	11	6	2	1	2	9	5	2	0
Backgrounds										
$t\bar{t}\rightarrow ll$	4	16	13	5	2	2	13	10	4	1
$t\bar{t}\rightarrow l$+jets	1	11	47	58	30	0	6	32	45	20
$Wb\bar{b}$	188	120	50	14	2	131	110	56	16	4
$Wc\bar{c}$	81	74	36	9	1	64	74	46	13	2
Wjj	175	61	20	5	1	125	58	23	6	2
Multijets	36	66	48	18	7	17	26	24	8	2
Background Sum	484	348	213	110	43	340	286	191	93	30
Data	445	357	207	97	35	289	287	179	100	38

Table 6.12: Yields after selection for events with exactly one b-tagged jet.

	Electron Channel					Muon Channel				
	1 jet	2 jets	3 jets	4 jets	≥ 5 jets	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
Signals										
tb	—	2.3	1.1	0.3	0.1	—	1.9	0.9	0.3	0.1
tqb	—	0.3	0.8	0.4	0.2	—	0.2	0.7	0.4	0.1
Backgrounds										
$t\bar{t}\rightarrow ll$	—	5.5	4.6	1.7	0.7	—	4.6	3.8	1.4	0.5
$t\bar{t}\rightarrow l$+jets	—	1.7	13.6	21.8	11.7	—	1.0	10.2	18.0	8.1
$Wb\bar{b}$	—	16.2	6.8	1.8	0.3	—	15.3	8.2	2.3	0.6
$Wc\bar{c}$	—	1.6	1.1	0.4	0.1	—	1.6	1.5	0.5	0.1
Wjj	—	0.1	0.1	0.0	0.0	—	0.1	0.1	0.0	0.0
Multijets	—	2.5	3.2	2.7	1.4	—	1.5	1.9	0.4	0.8
Background Sum	—	27.5	29.4	28.4	14.2	—	24.1	25.7	22.7	10.1
Data	—	30	37	22	10	—	23	32	27	10

Table 6.13: Yields after selection for events with exactly two b-tagged jets.

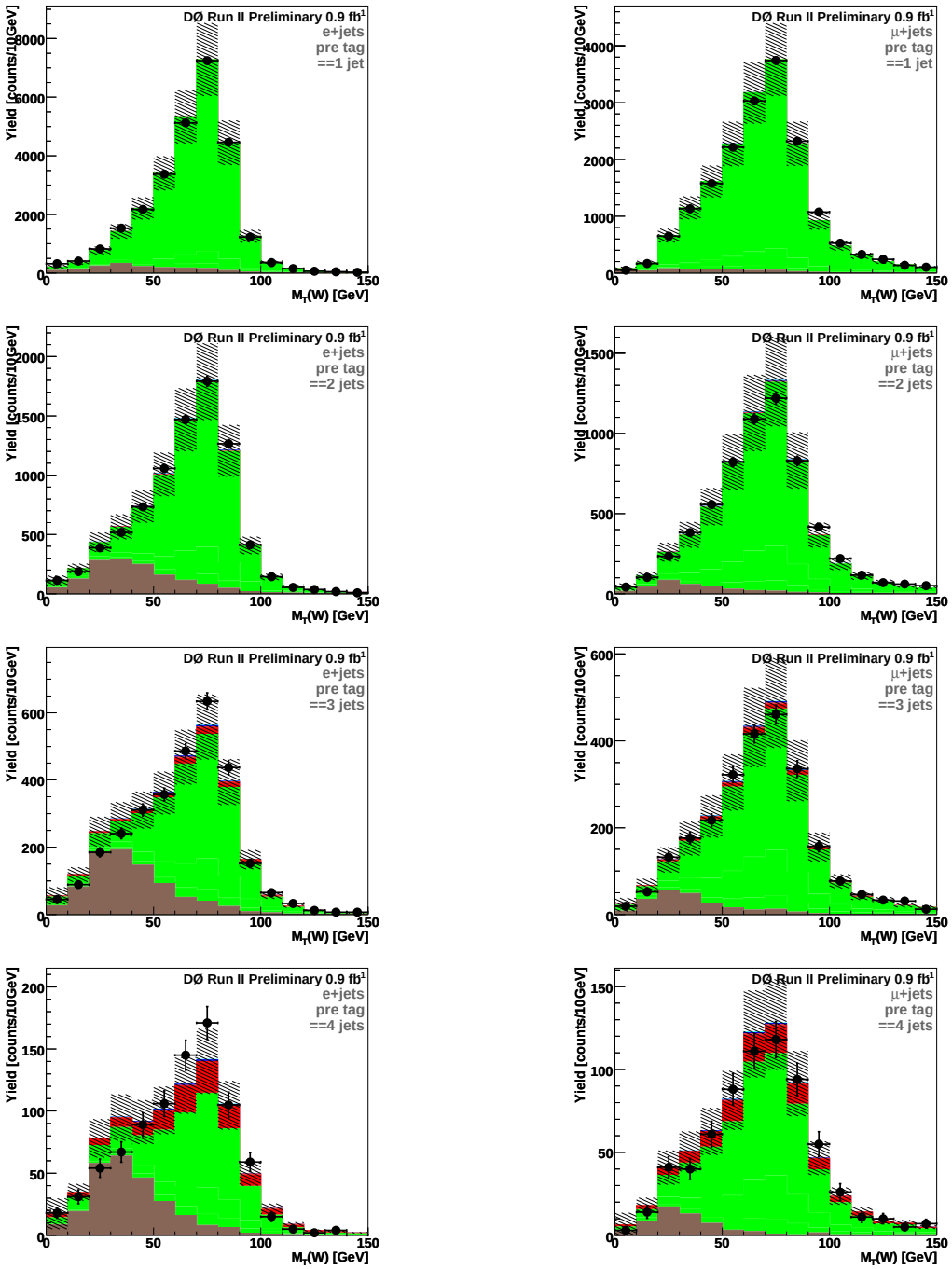


Figure 6.13: The W transverse mass distribution for one, two, three, and four jets events in the electron channel (left column) and in the muon channel (right column).

Chapter 7

The Analysis: Matrix Element Technique

The central idea behind the matrix element (ME) technique is that the physics of a collision is contained in the matrix element:

$$d\sigma = \frac{|\mathcal{M}|^2}{F} d\Phi \tag{7.1}$$

where $d\sigma$ is the differential cross section, F is the flux factor, $\mathcal{M}$ is the matrix element, and $d\Phi$ is the Lorentz invariant phase space factor. All the correlations between the observables are contained in this equation. Therefore, instead of using machine learning methods to infer the correlations from the data, the way a neural network (NN) or decision tree (DT) would, it would be preferable to build a discriminant that directly uses the information contained in Eq. 7.1. That is what the ME method does. The main advantage of the ME method is therefore that it maximally uses the information contained in the physics of the problem, without trying to infer it from

the distributions. Furthermore, since it uses the information from first principles, not from training, the ME method is not susceptible to the overtraining problem of the other methods. It will not be led astray by peculiarities of the MC sample.

Of course, reality diminishes the discriminant power somewhat. The ME analysis does not use the full matrix elements of the processes, but LO approximations. Nor does the analysis include matrix elements for all possible background processes. Nevertheless, the ME method remains arguably the most powerful and well motivated method for discriminating signal from background.

The ME method has been used at D0 and CDF in the past for parameter estimation such as the top quark mass [183, 184, 185, 186, 187, 188] or the longitudinal W boson helicity fraction in top quark decays [189]. However, along with an analogous analysis at CDF [190], this is the first use for a search for single top production.

7.1 Matrix Element Technique Overview

The matrix element method uses the matrix elements of a process to calculate the probability to observe a particular event assuming that it is the given process. The key equation is:

$$P(x|\text{process}_i) = \frac{1}{\sigma_i}\frac{d\sigma_i}{dx} \tag{7.2}$$

where x is the configuration of the event, and $P(x|\text{process}_i)$ is the probability density to observe x given that the physics process is process_i. More concretely, x is the set of reconstructed jet and lepton four-vectors, and possibly other information such as b-tagging state, of the event. For each event, we can calculate $P(x|\text{signal})$, which uses the matrix elements of the signal processes, and $P(x|\text{background})$, which uses the matrix elements of the background processes. Bayes's Theorem allows us to invert

the relation:

$$P(\text{signal}|x) = \frac{P(x|\text{signal})P(\text{signal})}{P(x|\text{signal})P(\text{signal}) + P(x|\text{background})P(\text{background})} \tag{7.3}$$

Providing what is needed to separate signal from background: the probability that the event is signal given the event's configuration. Actually, the analysis uses a discriminant $D(x)$ defined by a related equation that contains the same information:

$$D(x) = \frac{P(x|\text{signal})}{P(x|\text{signal}) + P(x|\text{background})}. \tag{7.4}$$

For each event in each channel, two discriminant values are calculated: a t-channel discriminant and an s-channel discriminant. These discriminant values are plotted in a two-dimensional histogram, which is provided as input to `top_statistics` [191], a program that applies systematic errors and extracts the result using a Bayesian approach, as discussed in the next chapter.

7.2 Calculation of the Event Probability Density Functions

The event configuration, x, which was discussed above, refers to the reconstructed event configuration. However, the matrix element, $\mathcal{M}$, depends on the parton-level configuration of the event, which we label y. The differential cross section, $d\sigma/dx$, can be related to the parton-level variant, $d\sigma/dy$, by integrating over all the possible parton values, using the parton distribution functions to relate the initial state partons to the proton and antiproton, and using a "transfer function" to relate the outgoing

partons to the reconstructed objects. The relevant expression is:

$$\frac{d\sigma}{dx} = \sum_j \int dy \left[f_{1,j}(q_1, Q^2)\, f_{2,j}(q_2, Q^2)\, \frac{d\sigma_{hs,j}}{dy}\, W_j(x, y)\, \Theta_{\text{parton}}(y) \right] \tag{7.5}$$

where

- $\sum_j$ is a sum of different configurations that contribute to the differential cross section: it is the discrete analogue to $\int dy$. Specifically, this sum includes summing over the initial parton flavors in the hard scatter collision and the different permutations of assigning jets to partons. With regards to the initial parton flavors, for example, an s-channel collision can occur via $u\bar{d}$, $\bar{d}u$, $c\bar{s}$, and $\bar{s}c$ annihilation, where the first element of the pair is associated with the proton and the second with the antiproton. The differential cross section is the sum over those configurations. The CKM matrix is taken to be diagonal when choosing which pairs to sum.

- $\int dy$ is an integration over the phase space:

$$\int dy = \int dq_1 dq_2 d^3p_\ell d^3p_\nu d^3p_{q_1} d^3p_{q_2} \ldots . \tag{7.6}$$

  Many of these integrations are reduced by delta functions. Energy and momentum conservation removes 4 degrees of freedom, and as an approximation we take the angular part of the per-object transfer functions to be delta functions, removing 6 degrees of freedom in two-jet events and 8 for three-jet events. More details are given in Section 7.5.

- $f_{n,j}(q, Q^2)$ is the parton distribution function in the proton or antiproton ($n =$ 1 or 2, respectively) for the initial state parton associated with configuration j,

Process Category	Factorization Scale
s-channel single top	m_{top}^2
t-channel single top	$(m_{\text{top}}/2)^2$
W + jets	$m_W^2 + p_{\text{T}\,W}^2$
$t\bar{t}$	$m_{\text{top}}^2 + \sum_{\text{jets}} p_{\text{T}\,i}^2$

Table 7.1: The factorization scales, Q^2, used by the various categories of processes. The factorization scales were chosen to be the same as those used for the Monte Carlo generation.

carrying momentum q, evaluated at the factorization scale Q^2. We use the same factorization scales as used for the Monte Carlo generation, which are given in Table 7.1. This analysis uses CTEQ6L1 [20] leading-order parton distribution functions via LHAPDF [192].

- $d\sigma_{hs}/dy$ is the differential cross section for the hard scatter collision. It is proportional to the square of the leading order matrix element as given by (cf. Eq: 7.1):

$$d\sigma_{hs,j} = \frac{(2\pi)^4}{4\sqrt{(q_1 \cdot q_2)^2 - m_1^2 m_2^2}} |\mathcal{M}|^2 d\Phi \tag{7.7}$$

  where q and m are the four-momenta and masses of the initial state partons. The matrix elements are discussed in more detail in Section 7.4.

- $W_j(x, y)$, which can also be written as $W(x \,|\, y, j)$, is called the transfer function, and represents the conditional probability to observe configuration x in the detector given the original parton configuration (y, j). The transfer function is divided into two parts:

$$W(x \,|\, y, j) = W_{\text{perm}}(x \,|\, y, j)\; W_{\text{reco}}(x \,|\, y, j) \tag{7.8}$$

where $W_{\text{perm}}(x \,|\, y, j)$ is the weight assigned to the given jet-to-parton permutation, and $W_{\text{reco}}(x \,|\, y, j)$ relates the reconstructed value to parton values for a given permutation. Permutation weights are discussed in Sec. 7.6. More information about W_{reco} is provided in Sec. 7.7.

- $\Theta_{\text{parton}}(y)$ represents the parton level cuts applied in order to avoid singularities in the matrix element evaluation. No parton-level cuts were applied when integrating over $t\bar{t}$ matrix elements. For all others, there was a pseudorapidity cut for each parton: $|\eta| < 8.0$. For the t-channel matrix elements, and for Wcg, Wgg, $Wcgg$, and $Wcgg$ matrix elements, an additional quark and gluon transverse momentum cut of $p_{\text{T}} > 2.0\,\text{GeV}$ was required.

7.3 Normalization of the Probabilities

The probability to observe a particular event given a process hypothesis, Eq. 7.2, also requires the total cross section ($\times$ branching ratio) as a normalization. Logically, the total cross section (σ) is just an integration of Eq. 7.5:

$$\sigma = \int dx \, \frac{d\sigma}{dx} \, \Theta_{\text{reco}}(x). \tag{7.9}$$

The term, $\Theta_{\text{reco}}(x)$, approximates the selection cuts. While conceptually simple, Eq. 7.9 represents a huge integral: 13 dimensions for two-jet events, 17 dimensions for three-jet events other than $t\bar{t}$, and 20 dimensions for $t\bar{t}$. However, this integral needs to be calculated only once, not once per event, so the actual integration time is insignificant. Also, the values are not very important because any error just becomes a multiplicative factor to the true probability, and the discriminating power of the discriminant remains unchanged, though perhaps the optimal binning or weights

would be affected. That same argument is used to say that $P(\text{signal}|x)$ and $D(x)$, Eq. 7.3 and Eq. 7.4, have the same discriminating power.

7.4 The Matrix Elements

The matrix elements used in this analysis are listed in Table 7.2. The code to calculated the matrix elements is taken from the Madgraph [193] LO matrix-element generator and uses the HELAS [194] routines to evaluate the diagrams. In the table, for the single top processes, the top quark is understood to decay leptonically: $t \to \ell^+\nu b$, and for the $W+$jets processes, the W boson is also understood to decay leptonically: $W^+ \to \ell^+\nu$. The charge conjugate processes are also used. The same matrix elements are used for both the electron and muon channels. Furthermore, we use the same matrix elements for heavier generations of incoming quarks, assuming a diagonal CKM matrix. In other words, for the $t\bar{b}$ process, we use the same matrix element for $u\bar{d}$ and $c\bar{s}$ initial state partons.

New to the analysis after the result published in Ref. [91] is an optimization of the third jet bin. In the third jet bin, a significant fraction of the background is $t\bar{t} \to \ell +$jets, as can be seen from the yield tables or more directly, the observed results in Ref. [162]. Adding a matrix element for that process was what motivated this iteration of the analysis. While no new processes were added to the two jet bin, in the three jet bin, tqg, $Wcgg$, $Wggg$, and lepjets are all new to this analysis.

Two Jets		Three Jets	
Name	Process	Name	Process
Signals		*Signals*	
tb	$u\bar{d} \to t\bar{b}$ (1)	tbg	$u\bar{d} \to t\bar{b}g$ (5)
tq	$ub \to td$ (1)	tqg	$ub \to tdg$ (5)
	$\bar{d}b \to t\bar{u}$ (1)		$\bar{d}b \to t\bar{u}g$ (5)
		tqb	$ug \to td\bar{b}$ (4)
			$\bar{d}g \to t\bar{u}\bar{b}$ (4)
Backgrounds		*Backgrounds*	
Wbb	$u\bar{d} \to Wb\bar{b}$ (2)	$Wbbg$	$u\bar{d} \to Wb\bar{b}g$ (12)
Wcg	$\bar{s}g \to W\bar{c}g$ (8)	$Wcgg$	$\bar{s}g \to W\bar{c}gg$ (54)
Wgg	$u\bar{d} \to Wgg$ (8)	$Wggg$	$u\bar{d} \to Wggg$ (54)
		lepjets	$q\bar{q} \to t\bar{t} \to \ell^+\nu b\bar{u}d\bar{b}$ (3)
			$gg \to t\bar{t} \to \ell^+\nu b\bar{u}d\bar{b}$ (3)

Table 7.2: The Matrix Elements used in the analysis. The number in parentheses specifies the number of Feynman diagrams included in the process. For simplicity, only the processes that contain a positively-charged lepton in the final state are shown. The charge conjugated processes are also used.

7.5 Integration Details

As shown in Sec. 7.2, the integration for the differential cross section, $d\sigma/dx$ consists of an integration over:

$$\int dy = \int dq_1 dq_2 d^3p_\ell d^3p_\nu d^3p_{q_1} d^3p_{q_2} \ldots \tag{7.10}$$

where the ellipsis represents the integration over any extra quarks or gluons. Energy and momentum conservation can be used to collapse the d^3p_ν integration and one of the two degrees of freedom associated with $dq_1 dq_2$, the initial state parton momenta, to get:

$$\int dy = \int dp_z d^3p_\ell d^3p_{q_1} d^3p_{q_2} \ldots, \tag{7.11}$$

where dp_z is an integration over the total p_z of the hard scatter. Changing to polar integration:

$$\int dy = \int dp_z d|p_\ell| d\Omega_\ell d|p_{q_1}| d\Omega_{q_1} d|p_{q_2}| d\Omega_{q_2} \ldots, \tag{7.12}$$

where the Jacobian, $|p_\ell|^2 |p_{q_1}|^2 |p_{q_2}|^2 \ldots$, has been omitted. This description omits all Jacobians, though they were of course included in the calculation.

The analysis program uses the VEGAS Monte Carlo integration algorithm [195], as implemented in the GNU Scientific Library [196]. That algorithm converges more quickly if it can find the areas that contribute significantly to the integral. This fact motivates the changes of variables for the various matrix elements, discussed in the following subsections.

7.5.1 Single Top, W+jets

For the single top and $W+$jets matrix elements, we assume one-to-one matching of parton to reconstructed object. As an approximation the angular part of the transfer functions is assumed to be a delta function (see Eqs. 7.26, 7.28 and 7.32), collapsing the $d\Omega$ integrations:

$$\int dy = \int dp_z d|p_\ell| d|p_{q_1}| d|p_{q_2}| \ldots . \tag{7.13}$$

Thus, in the end, the integrations are over four dimensions for two-jet events and over five dimensions for three-jet events.

To speed up the integration we change the variables to integrate over the resonances, and we use a Briet-Wigner random number distribution to generate the samples inside the peaks. Specifically, instead of integrating over $d|p_\ell|$, the program integrates over $dm^2_{\ell\nu}$, and the random numbers used to evaluate that integral have a Breit-Wigner distribution. For the single top processes, there is a similar top quark resonance that is treated in the same way. In summary, the integration becomes:

$$\int dy = \int dp_z dm^2_{\ell\nu} d|p_{q_1}| d|p_{q_2}| \ldots \ (W+\text{jets}) \tag{7.14}$$

$$\int dy = \int dp_z dm^2_{\ell\nu} dm^2_{\ell\nu b} d|p_{q_2}| \ldots \ (\text{singletop}). \tag{7.15}$$

7.5.2 $t\bar{t}\rightarrow\ell$+jets

For the $t\bar{t}\rightarrow\ell$+jets integration, we can no longer assume a one-to-one matching of parton to reconstructed object. The final state has four quarks, so one-to-one matching would lead to a four-jet event. In two-jet events, $t\bar{t}$ is not a significant problem, so this analysis does not use any $t\bar{t}$ matrix elements in the two-jet bin. We are interested, however, in using the lepjets matrix element in the three-jet bin because that bin

contains significant $t\bar{t}$ background. The $t\bar{t}$ events therefore have to "lose" one jet to enter this bin. One way that this jet could be lost is because the reconstructed p_{T} of the jet is below our cutoff. The p_{T} cuts for jets, as explained in Sec. 6.4, are 25 GeV for the leading jet, 20 GeV for the second leading jet, and 15 GeV for subsequent jets. A $t\bar{t}\to\ell$+jets event that has a fourth jet with $p_{\mathrm{T}} = 13\,\mathrm{GeV}$, for example, would consider the fourth jet lost. Another way to lose a jet is to merge it with another nearby object. Overlapping a jet with an electron would probably kill the electron, but two jets could be merged to bring the jet count down to three. The jet could also be outside the η acceptance of the analysis, which is $\left|\eta^{\mathrm{det}}\right| < 3.4$. There is also general reconstruction inefficiency that could cause a jet to be lost, but the probability for that is small: a few percent [139].

We studied $t\bar{t}\to e$+jets Monte Carlo events before tagging and found that for 80 % of the time when a jet was lost, there was no jet that passed the selection cuts within $\Delta R < 0.5$ of the corresponding parton. Thus, the jet was not merged with another jet but was lost some other way. Fig. 7.1 shows the p_{T} and η distribution of quarks not matched to a jet passing the selection cuts. The low p_{T} of the quark suggests that the jet is often lost because it falls below the jet p_{T} threshold. Looking at the quark flavor of lost jets, the study showed that the light quark jets were 1.7 times as likely to be lost (not by merging) as the heavy quarks. This observation can be understood from the p_{T} distributions: the b-quarks tend to be harder, as seen in Fig. 7.2. This observation motivated us to make the simplifying assumption that it is a light quark coming from the hadronically-decaying W that is lost.

Specifically, the following algorithm was used to model losing a jet. The algorithm involves determining two things for each parton-level configuration (y, j): which light quark jet is lost and how likely would it be lost. The steps are done in sequence.

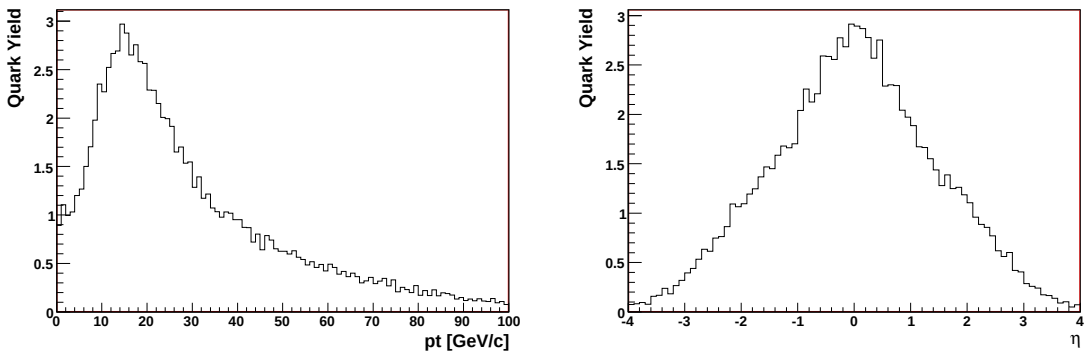


Figure 7.1: The p_{T} and η distribution of quarks that are not matched to a jet and not merged with another jet for a $t\bar{t}\rightarrow e$+jets MC sample before tagging.

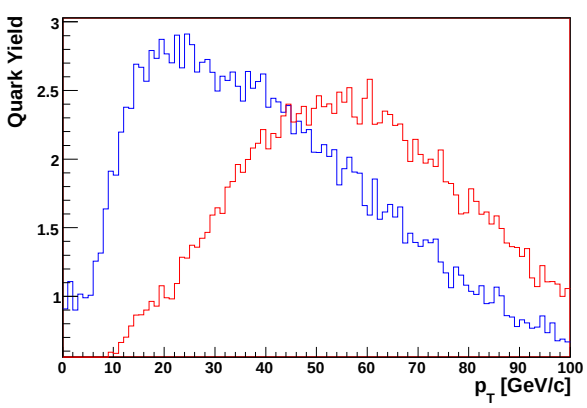


Figure 7.2: The p_{T} distribution of b-quarks (red) and light quarks (blue) for a $t\bar{t}\rightarrow e$+jets sample before tagging.

1. If the two light quarks are within $\Delta R < 0.6$, it is assumed they would be merged into a single jet, and the likelihood to lose the jet, w, is set to 1. The rationale for the value of 0.6 is that the analysis uses jets with a cone of $r = 0.5$, and two jets are merged when 50 % of the energy of the lower energy jet is in the overlap area. If one assumes the distribution of energy is uniform within a 0.5 cone around each initiating parton, the point where merging occurs is when the distance between the two jets is 0.6 [197]. The jet is treated for transfer function purposes as if it were a single light jet associated with the summed parton four-vector: $p_{\text{not lost}} \equiv p_q + p_{\bar{q}'}$. This case does not happen very often.

2. Otherwise, randomly choose which light quark to lose. Associate it with the "lost jet" four-vector p_{lost} that is integrated over, and associate the jet that is not lost with the reconstructed jet four-vector $p_{\text{not lost}}$.

3. If the lost parton has $|\eta| > 3.4$ it is assumed that an associated jet would not found because of the η cut in the selection. The likelihood to lose the jet, w, is set to 1.

4. Otherwise, return a weight based on the transfer function as a function of the E_{T} of the parton that is to be lost:

$$w(E_{\text{T,parton}}) = \max\left\{\int_0^{15} dE_{\text{T,reco}} W_{jet}(E_{\text{T,reco}}|E_{\text{T,parton}}), 0.05\right\}. \tag{7.16}$$

   Integrating over the transfer function for reconstructed jet values below the jet p_{T} threshold models cases where the reconstructed jet p_{T} is below the jet threshold, and the minimum returned value of 0.05 models other jet identification efficiencies.

Continuing on from Eq. 7.12, the approximation that the angular part of the

transfer functions is a delta function collapses the lepton and b-quark angular integrations, just as it does for single top and W+jets. It also collapses the angular integral of the "found" light quark jet. The other light jet four-vector remains, and is renamed p_{lost}. Mapping the resonances in the same method as for single-top, the final form of the integration becomes:

$$\int dy = \int w\, dp_z dm^2_{\ell\nu} dm^2_{\ell\nu b} dm^2_{q\bar{q}'} dm^2_{q\bar{q}'b} d|p_{\text{lost}}| d\Omega_{q_{\text{lost}}}. \tag{7.17}$$

The weight for losing the jet is included explicitly in this equation as a reminder, even though other weights, like the transfer function weight, have been omitted. The integration is over eight dimensions: there are the same five degrees of freedom as in other three-jet events plus three degrees for the momentum of the lost jet.

7.6 Assignment Permutations

The (discrete) summation over different configurations incorporated in Eq. 7.5 includes the summation over the different ways to assign the partons to the jets. A weight for each permutation is included as the W_{perm} part of the transfer function. This analysis uses two pieces of information to determine the weight, namely b-tagging and muon charge:

$$W_{\text{perm}} = W_{b\text{tag}} W_{\mu\text{charge}}. \tag{7.18}$$

7.6.1 b-Tagging

The b-tagging weight is assumed to factor by jet:

$$W_{b\text{tag}} = \prod_i w_{b\text{tag}}(\text{tag}_i \,|\, \alpha_i, p_{\text{T}i}, \eta_i), \tag{7.19}$$

where α_i is the flavor of jet i and tag_i is true or false depending on whether or not the jet is tagged. Since this analysis uses the neural network b tagger [159] for the event selection, it is natural to use the tag-rate functions in $w_{b\text{tag}}$. In particular, the weights assigned to cases with and without a b-tag are:

$$w_{b\text{tag}}(\text{tag} = \text{true} \,|\, \alpha, p_{\text{T}}, \eta) = P^{\text{taggable}}(p_{\text{T}}, \eta)\varepsilon_\alpha(p_{\text{T}}, \eta) \tag{7.20}$$

$$w_{b\text{tag}}(\text{tag} = \text{false} \,|\, \alpha, p_{\text{T}}, \eta) = 1 - P^{\text{taggable}}(p_{\text{T}}, \eta)\varepsilon_\alpha(p_{\text{T}}, \eta), \tag{7.21}$$

where ε_α is the tag-rate function for the particular flavor and P^{taggable} is the taggability-rate function, which is explained in Sec 6.8.1. The taggability is really also a function of the z position of the vertex, using six z bins, though this analysis always uses the $z = 0^-$ bin, that is, just 0, approaching from the negative side since 0 is a bin boundary. The logic behind that is that the normalization integral does not include an integration over the vertex position, so assuming $z = 0$ makes everything consistent. It is possible in future iterations to include the vertex position as an integration variable.

7.6.2 Muon Charge

For the s-channel and the lepjets matrix elements, there is both a b-quark and a $\bar{b}$-quark in the final state. Furthermore the matrix elements are not symmetric with

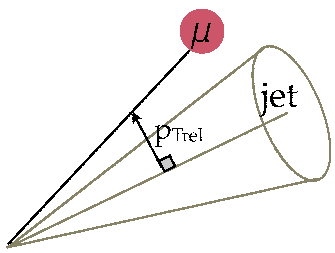


Figure 7.3: The definition of $p_{\mathrm{T_{rel}}}$.

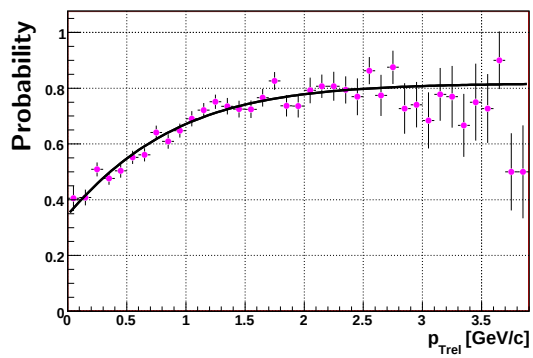


Figure 7.4: The probability for b quark jet to have a μ^- or a $\bar{b}$ quark jet to have a μ^+, given that there is a muon in the jet.

respect to the interchange of the b and $\bar{b}$ quarks, unlike for the Wbb matrix elements. Because b-tagging depends on the kinematic variables, it can still help to correctly assign jets to partons for single-tagged events; however, it would be useful to be able to distinguish between b-jets and $\bar{b}$-jets. Two possible ways to do this are by the jet charge, since b and $\bar{b}$ quarks have opposite charges, and in the case of muonic decays of the b quark, by the charge of the decay muon. The former is a possible future improvement, while the latter was implemented because it is relatively straightforward to do so, though it suffers from the low muonic branching ratios.

The sought after decays are:

$$b \to c\mu^-\nu \qquad\qquad \bar{b} \to \bar{c}\mu^+\bar{\nu} \tag{7.22}$$

but there are also the following:

$$b \to cX\bar{X}' \to s\mu^+\bar{\nu}X\bar{X}' \qquad\qquad \bar{b} \to \bar{c}X\bar{X}' \to \bar{s}\mu^-\nu X\bar{X}'. \tag{7.23}$$

The muon charge can still be used, however, because the distribution of $p_{\mathrm{T}_{\mathrm{rel}}}$, the muon p_{T} relative to the jet axis (see Fig. 7.3), is different in the two cases. Fig. 7.4 shows the probability to have a direct decay, that is, the muon charge to quark type association in Eq. 7.22, given that the jet has a muon, along with the fit that was used:

$$P^{\,\mathrm{direct}}(p_{\mathrm{T}_{\mathrm{rel}}}) = -0.182 + \tanh\left(0.68 p_{\mathrm{T}_{\mathrm{rel}}} + 0.587\right). \tag{7.24}$$

The probability ranges from around 40% for low $p_{\mathrm{T}_{\mathrm{rel}}}$ and goes up to 80% for high $p_{\mathrm{T}_{\mathrm{rel}}}$.

7.7 Object Transfer Functions

The matrix elements are a function of the parton four-vectors, but what we experimentally have access to are reconstructed objects, not partons. As was shown in Eq. 7.5, what we need is the "transfer function," the function relating the reconstructed configuration given the parton configuration: $W(x \,|\, y, j)$, or as we usually write it, $W_j(x, y)$. As shown in Eq. 7.8, W is the product of the probability associated with the assignment of jets to partons, W_{perm}, and the probability to measure the reconstructed values for the given parton values for a given parton to jet assignment, W_{reco}. In this section we are only concerned with W_{reco} and subsequently omit the subscript.

An assumption we make is that the transfer function can be factorized into individual per-object transfer functions:

$$W(x, y) = \prod_i W_i(x_i, y_i) \tag{7.25}$$

where $W_i(x_i, y_i)$ is a transfer function for one object—a jet, a muon, an electron—and x_i and y_i are reconstructed and parton-level information, respectively, for that object. Technically, the transfer functions can also depend on the vertex though that still allows them to factor.

7.7.1 Jets

The transfer function for jets [198] is assumed to be a function only of the relative energy difference between the parton and the jet, and all angles are assumed to be well measured:

$$W_{jet}(x_{jet}, y_{parton}) = W(E_{jet} - E_{parton}) \times \delta(\Omega_{jet} - \Omega_{parton}) \tag{7.26}$$

where $W(E_{jet} - E_{parton})$ is parametrized using a double Gaussian:

$$W(E_{jet} - E_{parton}) = \frac{\exp\left[-\frac{(E_{jet}-E_{parton}-p_1)^2}{2p_2^2}\right] + p_3 \exp\left[-\frac{(E_{jet}-E_{parton}-p_4)^2}{2p_2^5}\right]}{\sqrt{2\pi}(p_2 + p_3 p_5)} \tag{7.27}$$

with $p_i = \alpha_i + \beta_i E_{parton}$. The five α and five β parameters are determined by minimizing a likelihood formed by measuring the parton energy and the matched jet energy in a $t\bar{t} \to \ell$+jets MC sample. The parameters used for this analysis were determined in several pseudorapidity regions of the calorimeter to account for resolution differences in the detector. The parameters are determined separately for three types of jets: those that have a b-tag and a muon within the jet, those that have a b-tag but not a muon, and those that do not have a b-tag.

Figures 7.5 and 7.6 show the derived transfer functions for partons with parton energies of 15, 25, 50, and 90 GeV in the different η^{det} bins. These transfer functions

are the ones derived for and used in the top mass analysis [185], which we also used.

7.7.2 Electrons

The transfer function for electrons is based on the electron response modeling in the Parametrized MC Simulation (PMCS) package for fast simulation [199, 200]. The transfer function is assumed to be a function of the reconstructed energy of the electron, E_{reco}, the parton energy of the electron, E_{parton}, and θ, the production angle with respect to the beam axis:

$$W_{electron}(x_{reco}, y_{parton}) = W(E_{reco}, E_{parton}, \theta) \times \delta(\Omega_{reco} - \Omega_{parton}) \tag{7.28}$$

where $W(E_{reco}, E_{parton}, \theta)$ is parametrized using the following Gaussian form:

$$W(E_{reco}, E_{parton}, \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(E_{reco} - E_{center})^2}{2\sigma^2}\right] \tag{7.29}$$

with

$$\begin{aligned}
E_{center} &= 1.0002 E_{parton} + 0.324\,\mathrm{GeV} \\
\sigma &= 0.028 E_{center} \oplus \mathrm{Sampling}(E_{center}, \theta) E_{center} \oplus 0.4\,\mathrm{GeV} \\
\mathrm{Sampling}(E, \theta) &= \left[\frac{0.164}{\sqrt{E}} + \frac{0.122}{E}\right] \exp\left[\frac{\mathrm{p1}(E)}{\sin\theta} - \mathrm{p1}(E)\right] \\
\mathrm{p1}(E) &= 1.35193 - \frac{2.09564}{E} - \frac{6.98578}{E^2}.
\end{aligned} \tag{7.30}$$

The symbol, $\oplus$, means to add in quadrature.

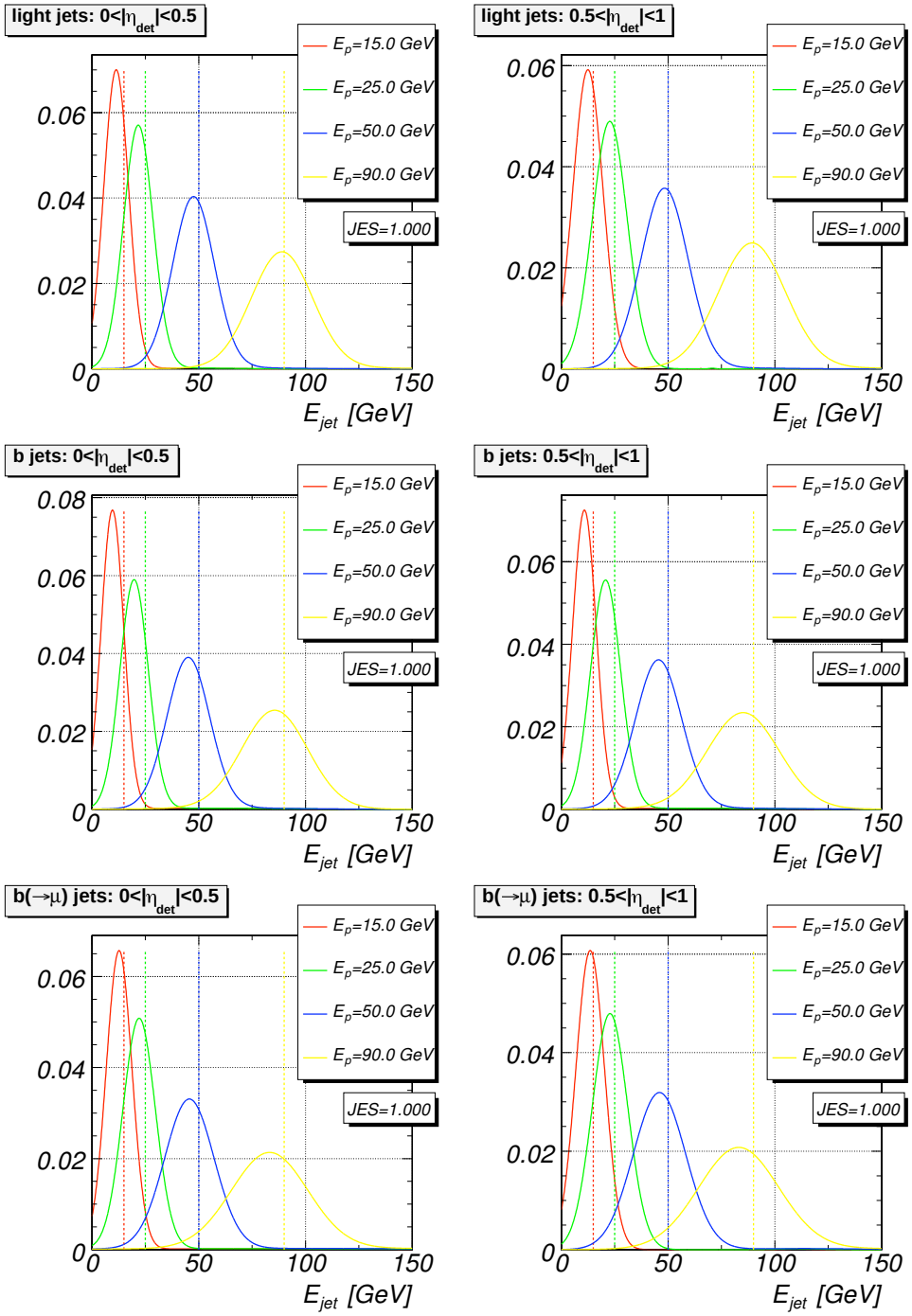


Figure 7.5: The jet transfer function parametrizations for $\left|\eta^{\mathrm{det}}\right| < 0.5$ and $0.5 < \left|\eta^{\mathrm{det}}\right| < 1.0$. (Figures from [198].)

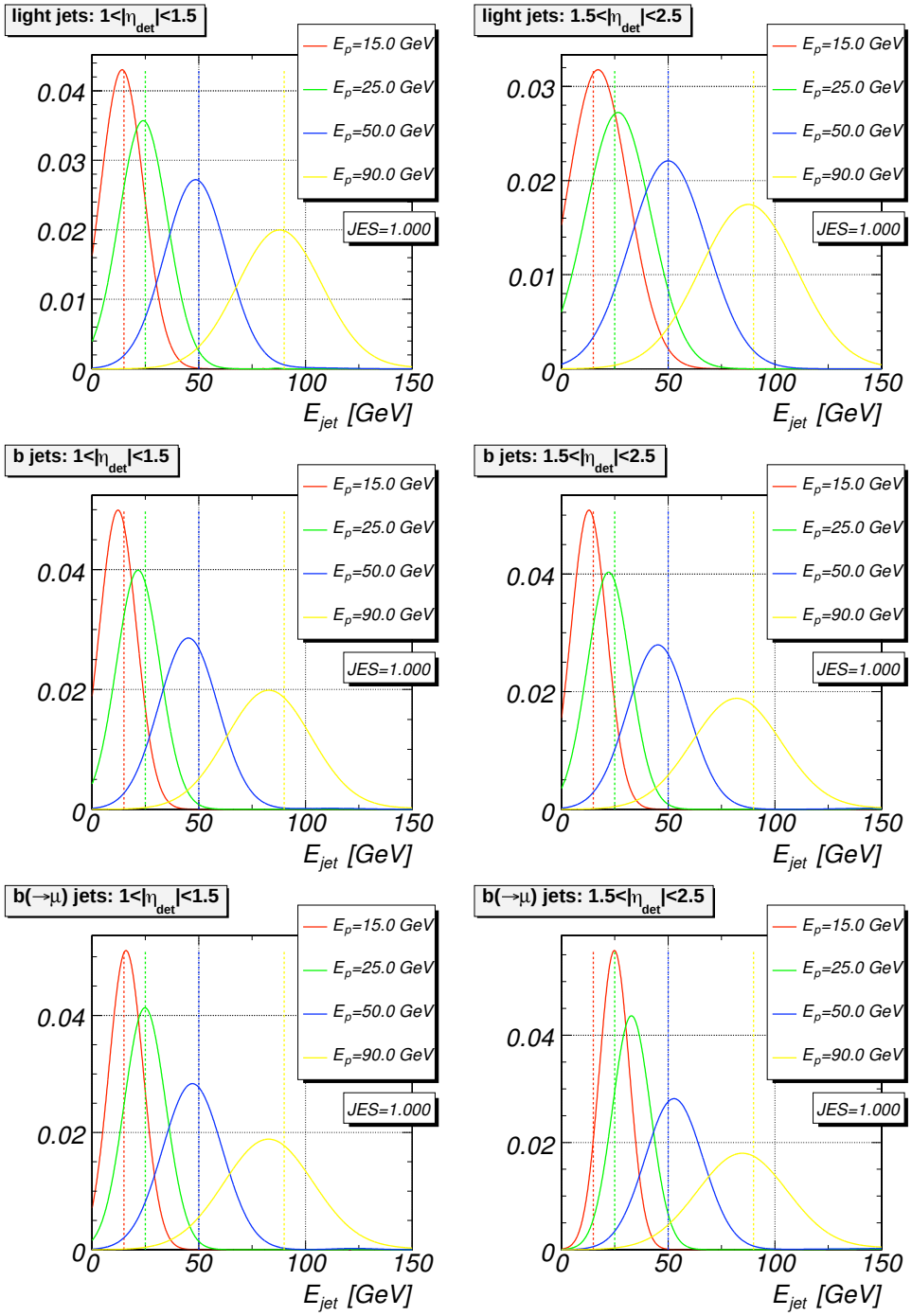


Figure 7.6: The jet transfer function parametrizations for $1.0 < \left|\eta^{\mathrm{det}}\right| < 1.5$ and $1.5 < \left|\eta^{\mathrm{det}}\right| < 2.5$. (Figures from [198].)

7.7.3 Muons

The muon transfer functions were derived on PYTHIA $t\bar{t}$ samples, smeared first to model the data [201, 202]. The transfer function is assumed to be a function of

$$\Delta\left(\frac{q}{p_t}\right) = \left(\frac{q}{p_t}\right)_{reco} - \left(\frac{q}{p_t}\right)_{parton} \tag{7.31}$$

and of η_{CFT}. Therefore

$$W_{muon}(x_{reco}, y_{parton}) = W\left(\Delta\left(\frac{q}{p_t}\right), \eta_{\mathrm{CFT}}\right) \times \delta(\Omega_{reco} - \Omega_{parton}) \tag{7.32}$$

where $W\left(\Delta\left(q/p_t\right), \eta_{\mathrm{CFT}}\right)$ is parametrized using a single Gaussian:

$$W\left(\Delta\left(\frac{q}{p_t}\right), \eta_{\mathrm{CFT}}\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{\left[\Delta\left(\frac{q}{p_t}\right)\right]^2}{2\sigma^2}\right\} \tag{7.33}$$

$$\sigma = \begin{cases} \sigma_o & : \quad |\eta_{\mathrm{CFT}}| \leq \eta_o \\ \sqrt{\sigma_o^2 + [c(|\eta_{\mathrm{CFT}}| - \eta_o)]^2} & : \quad |\eta_{\mathrm{CFT}}| > \eta_o \end{cases} \tag{7.34}$$

There are three fitted parameters in the above equations: σ_o, c, and η_o, each of which is fitted by two sub-parameters:

$$par = par(0) + par(1) \cdot 1/p_t. \tag{7.35}$$

Furthermore, these parameters are derived for four classes of events: those that were from before or after the 2004 shutdown, when the magnetic field strength changed, and in each run range, those that have an SMT hit on the track and those that do not.

These transfer functions are the same ones derived and used in the top mass analysis. This analysis uses only the post-2004 parametrization. Charge misidentification is modeled by adding over both the correct and incorrect charge assignment in Eq. 7.31, though the effect should be small. A study measuring the $W \to \mu\nu$ charge asymmetry found the muon charge misidentification to be approximately 2% when no track quality requirements were made, 0.2% when more than eight CFT hits and at least one SMT hit was required, and $(0.1 \pm 0.1)\%$ for the tight track requirements that analysis used [203].

7.8 Single Top Discriminant

The form of the discriminant is given in Eq. 7.4 in Section 7.1, and it is repeated below:

$$D(x) = \frac{P(x|\text{signal})}{P(x|\text{signal}) + P(x|\text{background})} \tag{7.36}$$

where

$$P(x|\text{process}_i) = \frac{1}{\sigma_i}\frac{d\sigma_i}{dx}. \tag{7.37}$$

We build separate s-channel and t-channel discriminants, D_s and D_t. The signal probabilities for the various channels are:

$$P(x\,|\,2\text{jet}, s\text{channel}) = \frac{1}{\sigma_{tb}}\frac{d\sigma_{tb}}{dx} \tag{7.38}$$

$$P(x\,|\,2\text{jet}, t\text{channel}) = \frac{1}{\sigma_{tq}}\frac{d\sigma_{tq}}{dx} \tag{7.39}$$

$$P(x\,|\,3\text{jet}, s\text{channel}) = \frac{1}{\sigma_{tbg}}\frac{d\sigma_{tbg}}{dx} \tag{7.40}$$

$$P(x\,|\,3\text{jet}, t\text{channel}) = \frac{1}{(\sigma_{tqb} + \sigma_{tqg})}\frac{d(\sigma_{tqb} + \sigma_{tqg})}{dx}. \tag{7.41}$$

Equation 7.41 can also be written as:

$$P(x \,|\, 3\text{jet}, t\text{channel}) = \frac{\sigma_{tqb}}{(\sigma_{tqb}+\sigma_{tqg})}\frac{1}{\sigma_{tqb}}\frac{d\sigma_{tqb}}{dx} + \frac{\sigma_{tqg}}{(\sigma_{tqb}+\sigma_{tqg})}\frac{1}{\sigma_{tqg}}\frac{d\sigma_{tqg}}{dx} \quad (7.42)$$

$$= w_{tqb}P(x|tqb) + w_{tqg}P(x|tqg), \quad (7.43)$$

where w_{tqb} and w_{tqg} are the relative yields of the two signal processes. For single-tagged events, we use $w_{tqb} = 0.6$ and $w_{tqg} = 0.4$, while for double-tagged events, we use $w_{tqb} = 1$ and $w_{tqg} = 0$, based on the yield fractions given in Eq. 7.42 using the cross sections as calculated by the normalization integral, Eq. 7.9.

We keep the same methodology of using weights based on yield fraction for the $P(x|\text{background})$ calculations. We do not use a matrix element for every background that exists, however, so the yield fractions cannot be determined as for the signal probabilities. Some, such as $u\bar{d} \to Wc\bar{c}$, are not included because they have similar characteristics as ones that are included, $u\bar{d} \to Wb\bar{b}$ in this particular case. Therefore, we use the yields as determined from the official MC samples and consider what background the matrix elements were meant to discriminate against. Defining f_{sample} as the yield fraction for each official MC sample, two alternate sets of weights were evaluated:

$$w_{wbb} = f_{Wbb} + 0.5 f_{Wcc}$$

$$w_{wcg} = 0.5 f_{Wcc}$$

$$w_{wgg} = f_{Wlp} + f_{QCD}$$

	Weight	1 tag Electron	1 tag Muon	2 tags Electron	2 tags Muon
2-jet	w_{wbb}	0.55	0.60	0.83	0.87
	w_{wcg}	0.15	0.15	0.04	0.04
	w_{wgg}	0.35	0.30	0.13	0.09
3-jet	w_{wbbg}	0.35	0.45	0.30	0.40
	w_{wcgg}	0.10	0.10	0.02	0.03
	w_{wggg}	0.30	0.25	0.13	0.10
	w_{lepjets}	0.25	0.20	0.55	0.47

Table 7.3: Background weights chosen for each analysis channel in two-jet and three-jet events.

and because the $\bar{s}g \to W\bar{c}g$ process is in the Wlp ALPGEN sample:

$$\begin{aligned} w_{wbb} &= f_{Wbb} + f_{Wcc} \\ w_{wcg} &= 0.5 f_{Wlp} \\ w_{wgg} &= 0.5 f_{Wlp} + f_{QCD}. \end{aligned}$$

In the three-jet bin, the weights were calculated the same way, with the addition that $w_{\text{lepjets}} = f_{\text{lepjets}}$. These two alternate sets of weights proved to be close to each other, and because the discriminating power of the discriminants proved to not be very sensitive to the weights provided they are reasonable, the two sets of weights functioned as a guide. The actual values used are summarized in Table 7.3.

The performance of the s-channel and t-channel discriminants can be seen in Figs. 7.7–7.10. These are plots of the fraction of s-channel (t-channel) signal that passes versus the fraction of background that passes, varying the cut value c in $D_s > c$ ($D_t > c$), for various backgrounds. The color of the lines indicates the background considered, using the standard colors as specified by the legend in the preface to

this thesis. The three-jet discriminants include the lepjets matrix element, while the two-jet discriminants do not. A comparison between the two with regards to $t\bar{t}$ performance hints at the benefit of explicitly including the $t\bar{t}$ matrix element.

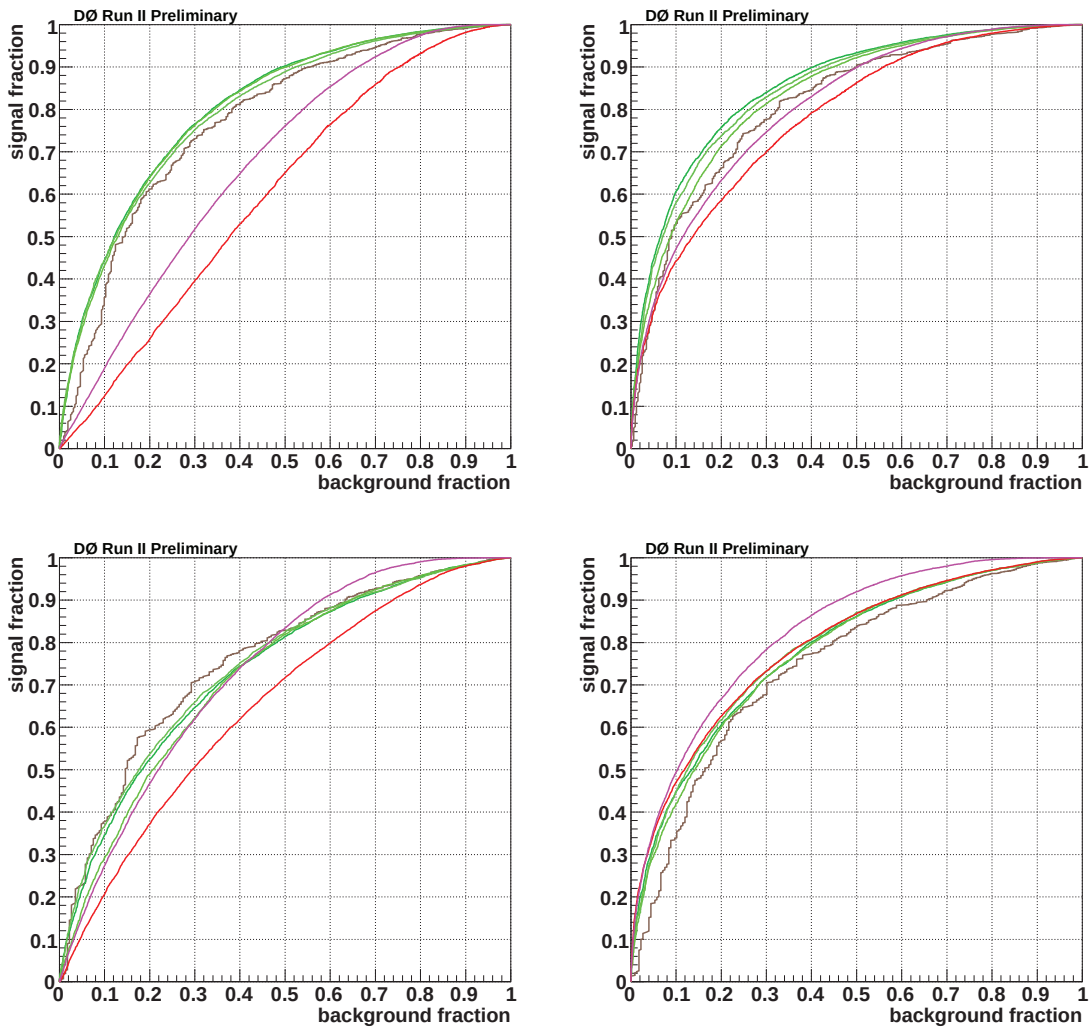


Figure 7.7: The discriminant performance for e+jets events with one b-tag. The upper row is for two-jet events, while the lower for three-jet events. On the left is the performance of the s-channel discriminant, on the right that of the t-channel discriminant.

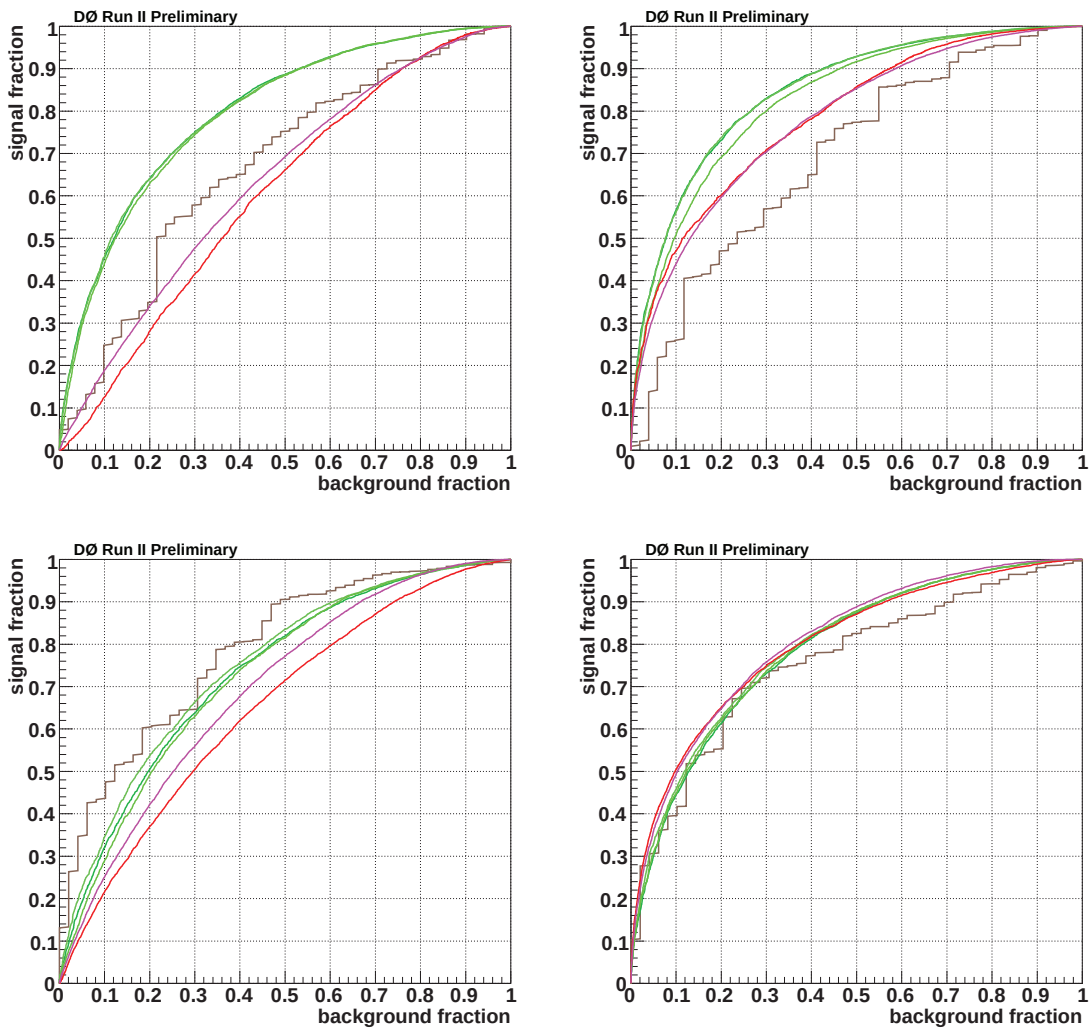


Figure 7.8: The discriminant performance for μ+jets events with one b-tag. The upper row is for two-jet events, while the lower for three-jet events. On the left is the performance of the s-channel discriminant, on the right that of the t-channel discriminant.

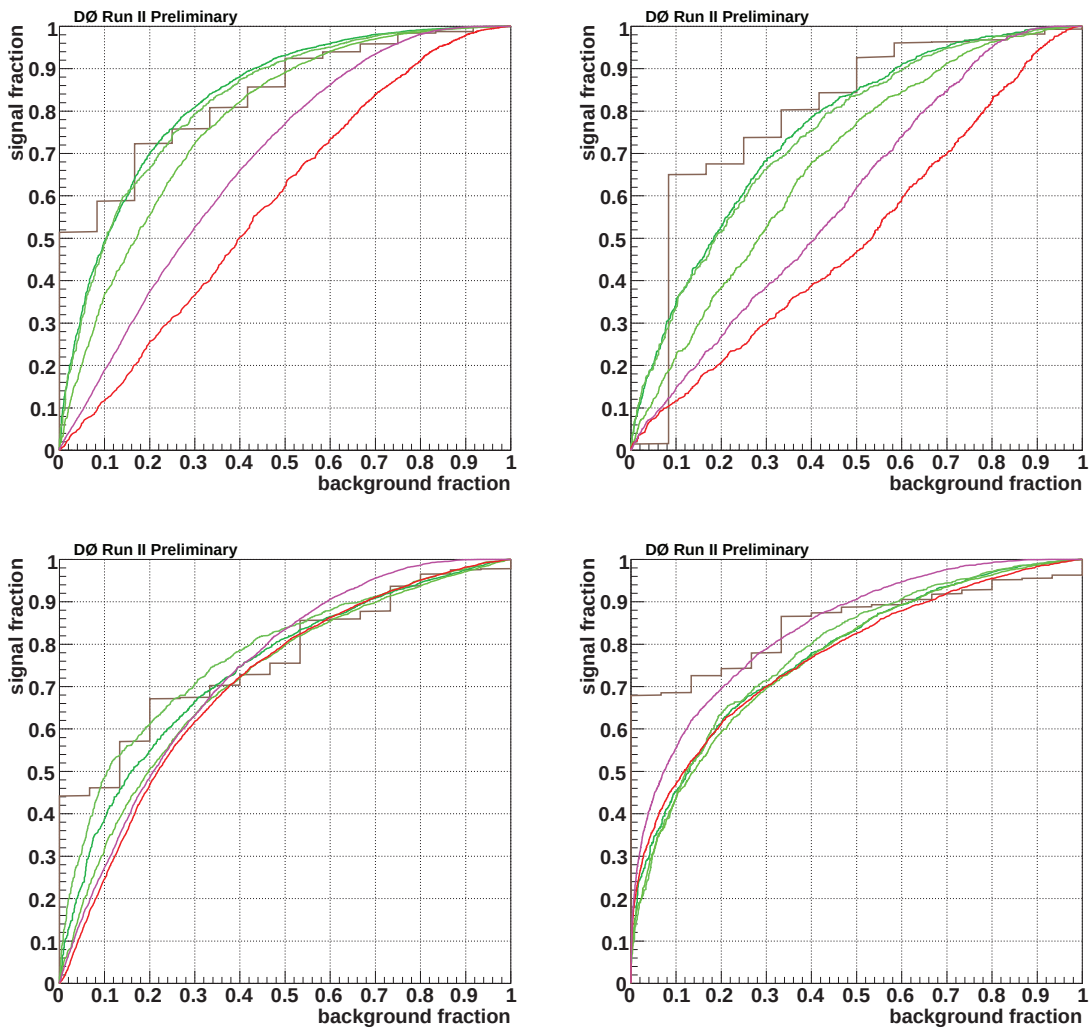


Figure 7.9: The discriminant performance for e+jets events with two b-tags. The upper row is for two-jet events, while the lower for three-jet events. On the left is the performance of the s-channel discriminant, on the right that of the t-channel discriminant.

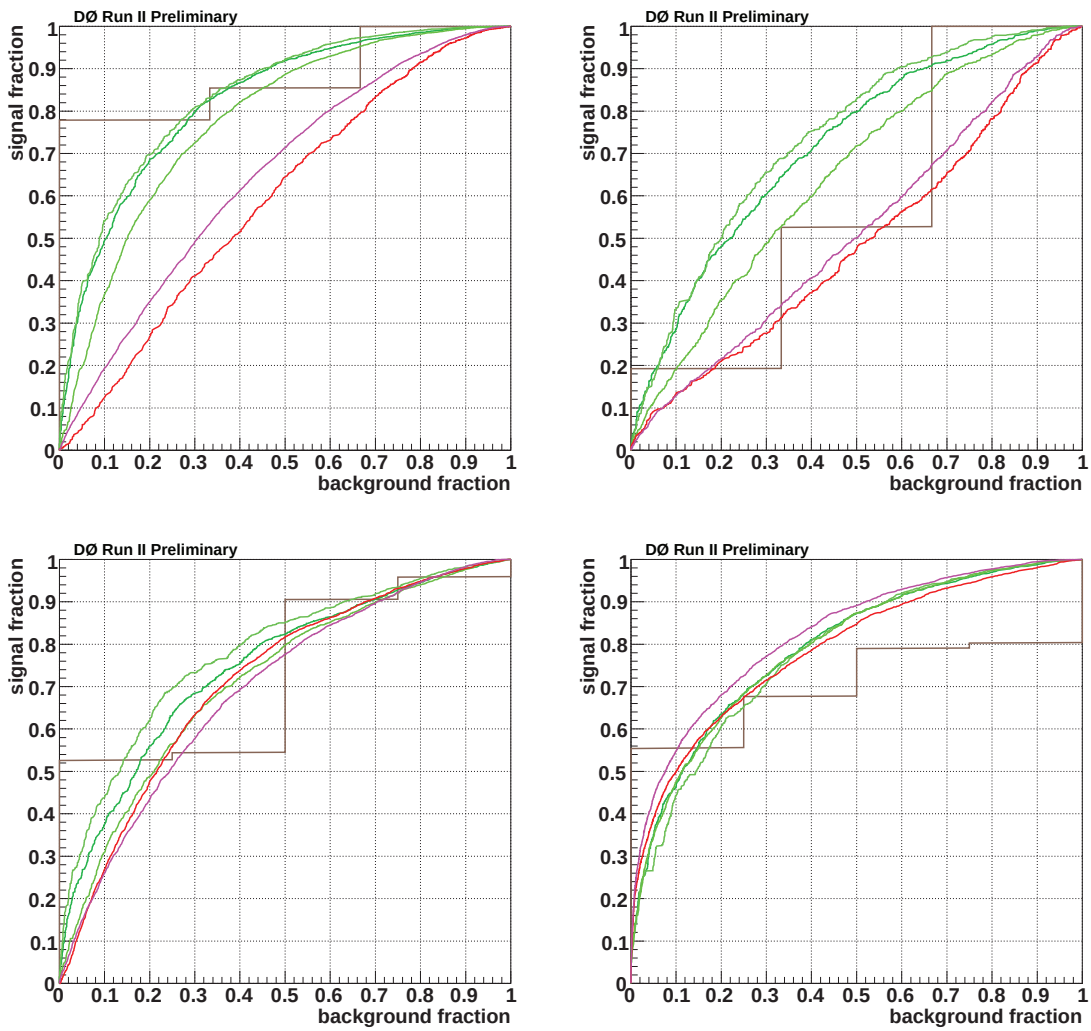


Figure 7.10: The discriminant performance for μ+jets events with two b-tags. The upper row is for two-jet events, while the lower for three-jet events. On the left is the performance of the s-channel discriminant, on the right that of the t-channel discriminant.

7.9 Cross-Check Samples

An important step in the single top search is to establish that the background model is appropriate while minimizing examining the search region. For this purpose, two background-dominated control samples are defined, and the 1D discriminants in data and in the Monte Carlo background model are compared.

The two control samples are defined to consist of events that pass the standard event selection and in addition $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ or $H_{\mathrm{T}} > 300\,\mathrm{GeV}$, respectively, where H_{T} is the scalar sum of the charged lepton p_{T}, the $\not{E}_{\mathrm{T}}$, and the jet p_{T}s. Because two-jet events are largely dominated by W+jets background, these cross-checks mainly test the modeling of that background. For three-jet events, the $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ sample also contains a significant fraction of $t\bar{t}$, so it is also a test of how well the $t\bar{t}$ background is modeled.

Most of the signal, and thus the W+jets background that is of concern, has H_{T} between these cuts, as can be seen in Fig. 7.11. Therefore, by confirming that the observed discriminant distribution is well reproduced by the background model for the softest and hardest W+jets events, we gain confidence that the W+jets background in the signal region is also well modeled.

Figures 7.12 and 7.13 compare the s-channel and t-channel discriminants between data and the background model for events with two and three jets respectively for the $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ sample. Figures 7.14 and 7.15 compare the s-channel and t-channel discriminants between data and the background model for events with two and three jets respectively for the $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ sample. The plots have the electron and muon channels, as well as the one and two b-tag channels combined for increased statistics, and the single top quark content is scaled to the measured cross section. Good agreement is seen for both control samples

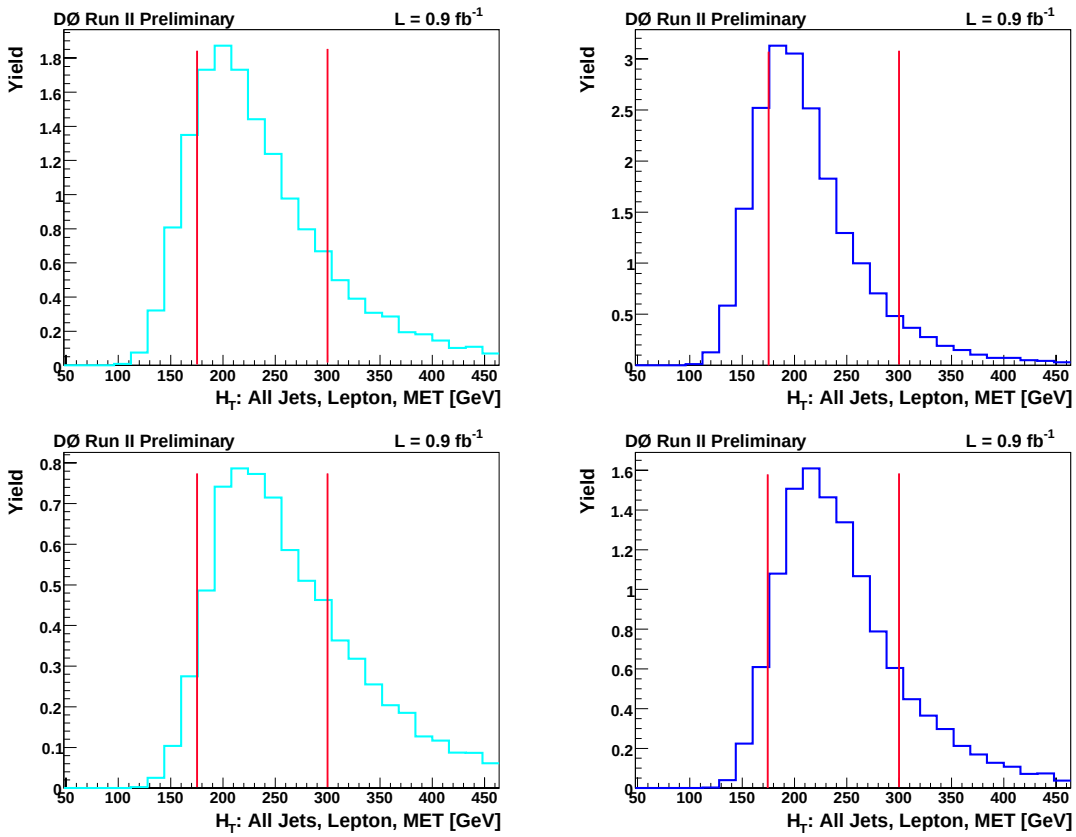


Figure 7.11: The H_{T} distributions for single-top MC events. The upper row is for two-jet events, while the lower row is for three-jet events. The left column is for s-channel, and the right for t-channel. Lines are drawn at $H_{\mathrm{T}} = 175\,\mathrm{GeV}$ and $H_{\mathrm{T}} = 300\,\mathrm{GeV}$.

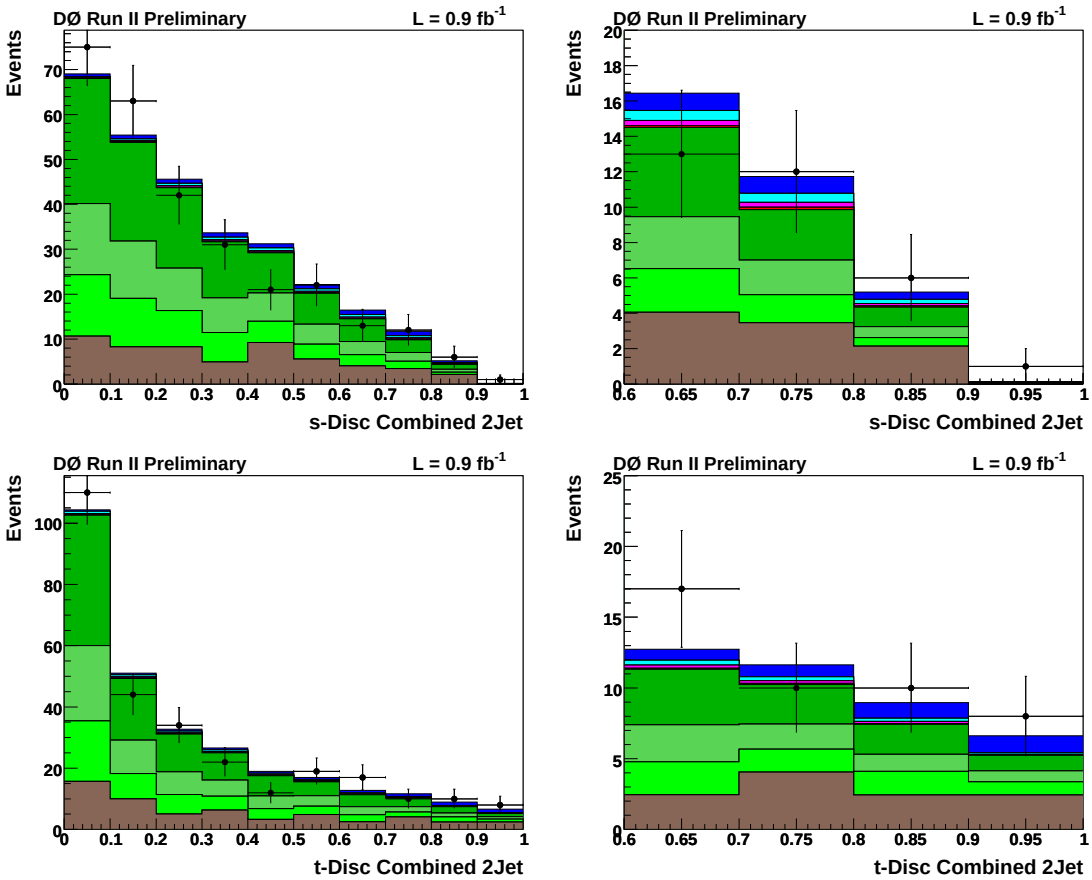


Figure 7.12: $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ cross-check plots in two-jet events for the s-channel discriminant (upper row) and the t-channel discriminant (lower row). The left column shows the full discriminant region while the right column shows the high discriminant region.

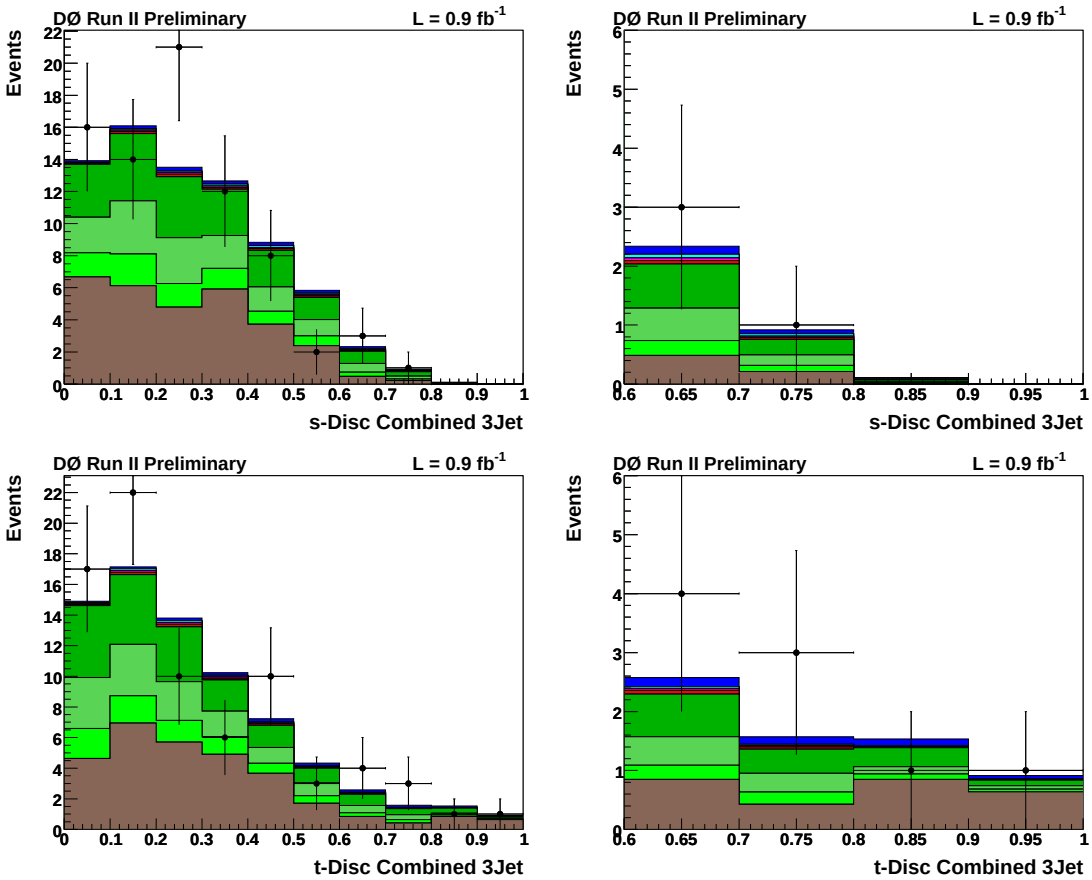


Figure 7.13: $H_{\mathrm{T}} < 175\,\mathrm{GeV}$ cross-check plots in three-jet events for the s-channel discriminant (upper row) and the t-channel discriminant (lower row). The left column shows the full discriminant region while the right column shows the high discriminant region.

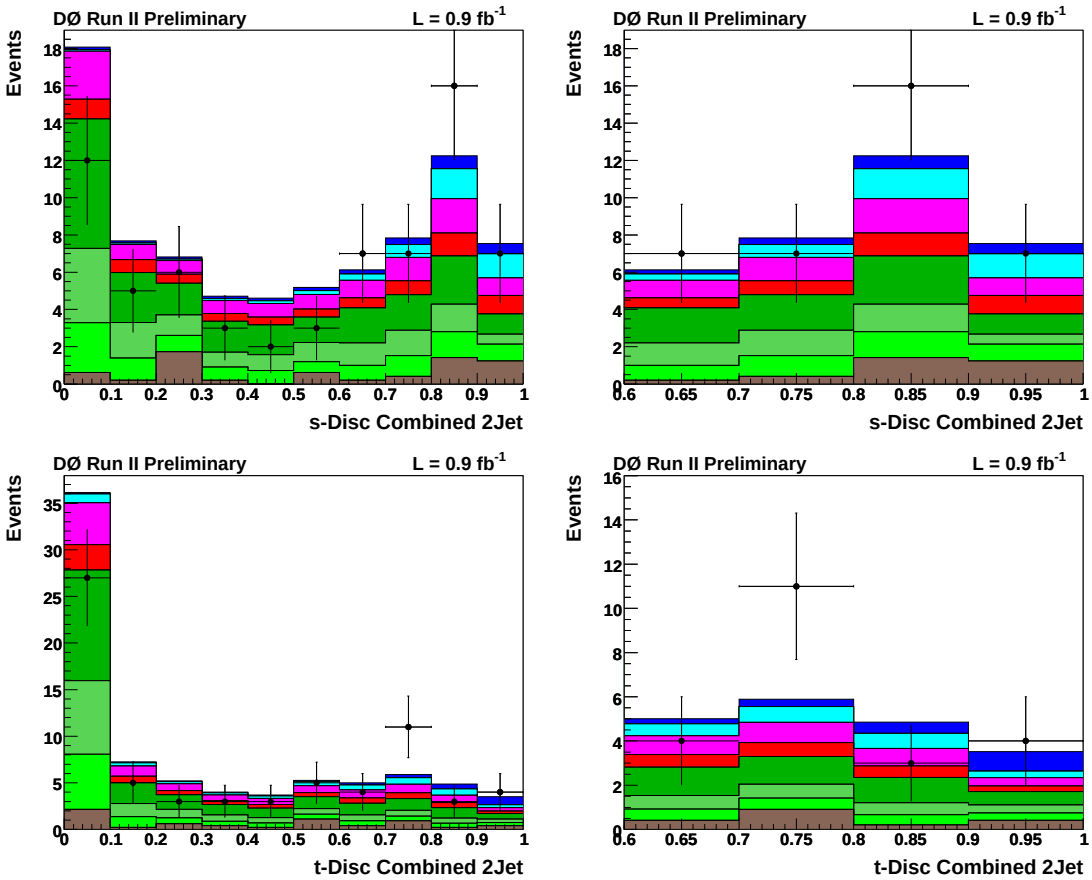


Figure 7.14: $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ cross-check plots in two-jet events for the s-channel discriminant (upper row) and the t-channel discriminant (lower row). The left column shows the full discriminant region while the right column shows the high discriminant region.

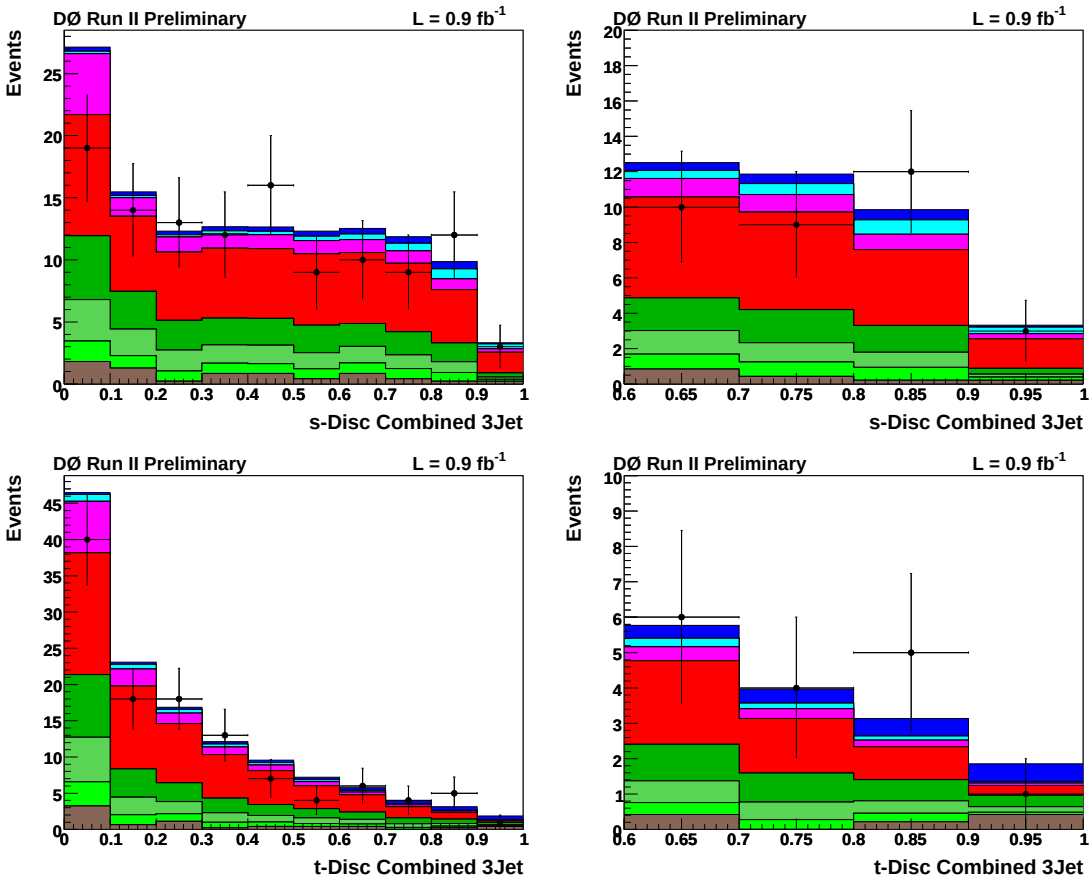


Figure 7.15: $H_{\mathrm{T}} > 300\,\mathrm{GeV}$ cross-check plots in three-jet events for the s-channel discriminant (upper row) and the t-channel discriminant (lower row). The left column shows the full discriminant region while the right column shows the high discriminant region.

Chapter 8

The Analysis: Extracting a Result

For each event two discriminant values are calculated: the s-channel single-top discriminant and the t-channel single-top discriminant. These are plotted in eight two-dimensional histograms, one for each channel (electron or muon, one or two b-tagged jets, two or three jets), and provided as input to `top_statistics` [191], which applies systematics and extracts a measurement using a Bayesian approach, as explained in Sec. 8.2. As a compromise between having too few statistics in the bins of the histograms, which would increase the uncertainty, and too few bins, which would decrease the discriminating power, 9×9 variable-sized binning was chosen. The size of the bins is such that in one-dimensional projections of the histogram, the amount of "signal" is uniform in each bin. For binning purposes, in the projection where the t-channel discriminant is collapsed, leaving only a one-dimensional histogram binned by the s-channel discriminant value, s-channel is considered signal, and similarly in the projection where the s-channel discriminant is collapsed, t-channel is considered the signal. However, we extract a combined s-channel+t-channel cross section assuming the SM cross section ratio of $\sigma_s/\sigma_t = 0.44$, so s-channel+t-channel is considered

signal subsequently.

8.1 Systematic Uncertainties

We treat systematic uncertainties in two ways: as an uncertainty on the normalization of a sample (a flat uncertainty) and as an uncertainty that change the shapes of the discriminant distributions (a shape-changing uncertainty). We find that only two uncertainties noticeably affect the shapes of the discriminant: the uncertainty in the jet energy scale, and the uncertainty in the TRFs used to parametrize b-tagging. Everything else we treat as a flat uncertainty.

For the JES uncertainties, the JES is shifted up by one standard deviation, and the analysis is repeated to produce a discriminant output. Then the JES is shifted down one standard deviation, and the discriminant is calculated again. In this way, separate shifts are calculated for each bin of the discriminant distributions. For MC events, the jet energy resolution uncertainty is not taken into account in the JES uncertainty, so to account for this, the jet energy smearing is varied by the size of the jet energy resolution. The TRF uncertainty is treated analogously.

We have considered the following systematic uncertainties in the analysis. The particular values for the electron channel, two-jets, one b-tag are given in Fig. 8.1, while the uncertainties for the other channels are in Appendix A.

- **Integrated luminosity**

  The uncertainty on the integrated luminosity is 6.1%. It affects the signal and $t\bar{t}$ yields.

- **Theoretical cross sections and branching ratios**

  The uncertainty on the cross section for signal and $t\bar{t}$ includes the theoretical

error and the uncertainty from the top quark mass uncertainty. Because W+jets is normalized to data, there is no cross section uncertainty associated with it.

- **Trigger efficiency**

  The uncertainty in the trigger efficiency is calculated by shifting the trigger turn-on curves up and down by one standard deviation. This shift in the yield is taken as an uncertainty in the yield.

- **Primary vertex selection efficiency**

  The primary vertex selection efficiency in data and MC are not the same. We assign a systematic uncertainty for the difference in the z position of the primary vertex taking into account the beam profile along the longitudinal direction [204].

- **Jet reconstruction and identification**

  A 1.5% uncertainty in the reconstruction efficiency for jets is assigned to take into account differences in the efficiency with regard to the number of jets and the η distribution of the jets.

- **Jet energy scale and jet energy resolution**

  This is a shape-changing uncertainty on the scale and resolution of jet energy.

- **Jet fragmentation**

  This systematic is modeled by the difference in the jet fragmentation models of PYTHIA and HERWIG as well as the uncertainty in the modeling of initial-state and final-state radiation.

- **Electron preselection efficiency**

  The electron preselection correction factors are parametrized as a function of

η^{det}. The uncertainty in the efficiency accounts for its dependence on variables other than η^{det}, and as a result of limited data statistics in determining the correction factor. The uncertainty is determined to be 2.2%.

- **Electron post-preselection efficiency**

  The electron post-preselection correction factors are parametrized as a function of η^{det} and ϕ^{det}. The uncertainty in the efficiency accounts for its dependence on other variables, such as the number of jets and the instantaneous luminosity, and as a result of limited data statistics in determining the correction factor. The uncertainty is determined to be 5%.

- **Muon reconstruction and identification efficiency**

  An uncertainty of 7% is assigned to the muon ID reconstruction efficiency, mostly due to limited statistics in determining the correction factor, but also including uncertainties in the tag and probe and background subtraction methods.

- **Muon central track matching uncertainty**

  The muon tracking uncertainty is assigned to be 1.5%. It includes uncertainties from the tag and probe and background subtraction method, biases due to luminosity and timing bias, and averaging over ϕ, and the limited statistics in each bin of the scale factor.

- **Muon isolation efficiency**

  The muon isolation efficiency uncertainty was estimated to be 2% to cover the dependences not taken into account in the muon isolation correction factor.

- **Matrix method normalization** [182]

  The determination of the number of real-lepton events in data is affected by the

uncertainties associated with the determination of the lepton efficiencies and fake rates. It is also affected by the limited statistics of the data sample.

- **Heavy flavor ratio**

  The error on the additional scale factor (α) we apply to set the $Wb\bar{b}$ and $Wc\bar{c}$ contributions in the W+jets sample is estimated to cover several effects: dependence on the b-quark p_{T}, the difference between the zero tag samples where it is estimated and the signal samples where it is used, and the intrinsic uncertainty on the value of the LO cross section.

- **MC tag-rate functions**

  This is a shape-changing uncertainty on the TRFs originating from statistical errors due to limited MC statistics, the assumed fraction of heavy flavor in the MC QCD used for the mistag rate determination, and the dependence that b-tagging has on variables other than the ones the TRFs are parametrized in.

8.2 Extracting a Measurement Using a Bayesian Approach

We model systematics and extract a result using a Bayesian approach [191]. We assume the probability to observe a count D if the mean is d, $P(D|d)$, is given by a Poisson distribution. In each individual bin of the 9×9 histograms, the expected count is $d = s + \sum_i b_i$, where s stands for signal and b_i are the various backgrounds. Furthermore, $s = \sigma a$, where σ is the cross section we are trying to extract, and a is the luminosity times the acceptance, which we measure with the single top MC.

	Single-Tagged Two-Jets Electron Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	18.2	18.2	18.2	18.2
Primary vertex	2.4	2.4	2.4	2.4	—	—	—	—
Electron ID	5.5	5.5	5.5	5.5	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	3.0	3.0	3.0	3.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	1.4	0.3	9.9	1.7	—	—	—	—
Flavor-dependent TRFs	2.1	5.9	4.6	2.4	4.4	6.3	7.4	—
Statistics	0.7	0.7	1.3	0.8	0.9	0.9	0.4	5.6
Combined								
Acceptance uncertainty	10.8	12.1	—	—	—	—	—	—
Yield uncertainty	19.3	19.3	24.1	21.1	18.8	19.3	19.7	19.1

Table 8.1: Electron channel uncertainties for one b-tag and two jets.

Using $L(D|d) \propto P(D|d)$, we can obtain the posterior probability density,

$$P(\sigma|D) = \frac{1}{\mathcal{N}} \int L(D|\sigma, a, \mathbf{b})\pi(\sigma, a, \mathbf{b}) da\, d\mathbf{b}, \tag{8.1}$$

where $\mathbf{b}$ is a vector of all the backgrounds, and $\pi(\sigma, a, \mathbf{b})$ is the prior. The measured cross section is taken to be peak of the posterior.

The prior, $\pi(\sigma, a, \mathbf{b})$, is assumed to factor: $\pi(\sigma, a, \mathbf{b}) = \pi(\sigma)\pi(a, \mathbf{b})$. We use a flat prior for the cross section:

$$\pi(\sigma) = \begin{cases} 1/\sigma_{\max} & : \quad 0 < \sigma < \sigma_{\max} \\ 0 & : \quad \text{otherwise} \end{cases} \tag{8.2}$$

The prior, $\pi(a, \mathbf{b})$, represents our prior knowledge of those variables. Thus systematic uncertainties are modeled by integrating over Gaussian priors. For the shape-changing uncertainties, the prior in each bin has a different positive and negative width, taken from how big a shift each bin experienced for the one standard deviation positive and negative shift studies. Correlations in the shifts can be taken into account by using the same random number for all the bins.

8.3 Generating Ensembles

In order to calibrate the method and determine significances, we make use of ensembles, which were generated as follows. Working separately for the electron channel and muon channel, but merged in number of jets and number of b-tags, we can determine nominal yields for $t\bar{t}$, W+jets, multijet background, and the two single-top signals for the ensemble. Then, to model systematics, for each member of an ensemble, the

nominal yields are shifted by a random number with a Gaussian distribution with a width 20% of the nominal yields, since the systematics as calculated more carefully in `top_statistics` are near that value. The same random number is used for the W+jets and multijet yields to model their anti-correlation. The number of events in this member of the ensemble is determined by sampling a poisson distribution with a mean equal to the sum of the shifted yields.

Separately for the electron channel and muon channel, we have a pool of signal+background events made up of the MC models and the orthogonal sample used for modeling the multijet background. The events have a weight associated with them so that the sum of all the events reproduces the nominal yields. For the backgrounds these are just the weights that are used in the background model, but the signal yields are scaled to produce the appropriate cross sections. Because we determined how many events a given sample in an ensemble has, we choose that many events from the pool at random, with the probability to select a given event being proportional to its weight. These events are all given a weight of one and treated as data events.

8.4 Calibration of the Method

Several ensembles of around 2000 simulated data sets were produced from the background model to test the calibration of the method. The ensembles were generated with a non-SM σ_{s+t} cross section but with the cross section ratio fixed at the SM value of $\sigma_s/\sigma_t = 0.44$. The results are shown in Fig. 8.1. In each case, the extracted cross section was in good agreement with the input value, although in the lower results, some shift in the mean can be attributed to not allowing a negative cross section to be measured.

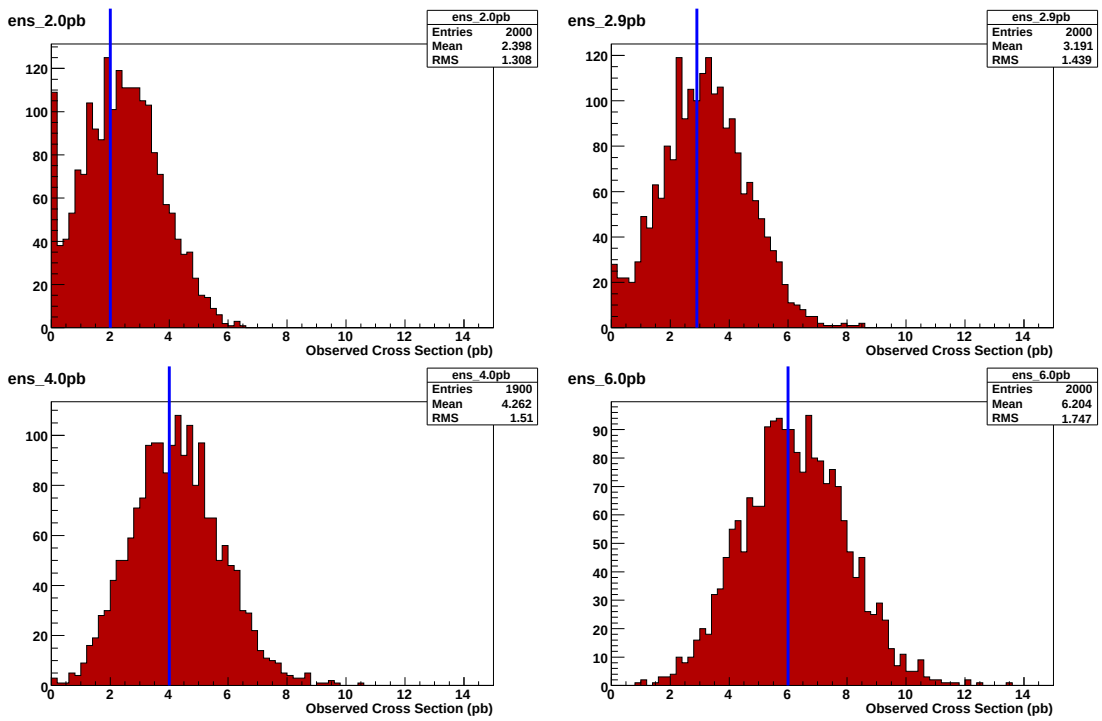


Figure 8.1: Results using the ensembles with a non-SM cross section but a SM σ_s:σ_t ratio. The blue lines indicates the input s-channel+t-channel cross section values, which are $2.0\,pb, 2.9\,pb, 4.0\,pb$, and $6.0\,pb$, respectively.

Figure 8.2 shows the measured cross section values versus the input cross sections, with a straight line fit. The fitted function is $\sigma_{\text{meas}} = 0.95\sigma_{\text{in}} + 0.46\,\text{pb}$. Given the relatively low precision of our final measurement and the fact that at least part of the reason for the small positive shift in the average measured values is because the measured cross section was not allowed to be negative, we do not apply any correction.

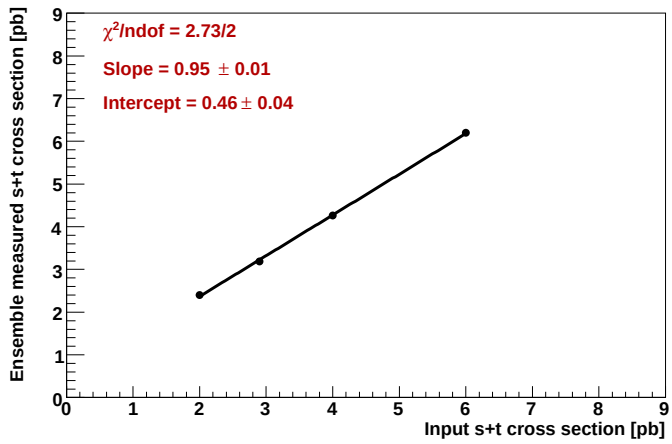


Figure 8.2: Measured signal cross section versus input cross section for the calibration ensembles.

8.5 Expected Results

This section presents the expected performance of the analysis. For these results the number of observed events has been set equal to the expected signal, according to the SM prediction, plus the expected background.

Figures 8.3 and 8.4 show the resulting s+t-channel posterior for the combined e+μ $\geq 1b$-tag channel in two-jet and three-jet events. Figure 8.5 shows the $s+t$-channel posterior for the combination of all channels. Table 8.2 shows the expected cross

sections for various combinations of analysis channels. Fig. 8.6 shows the expected cross sections for the various channels, along with a measure of the significance, $\sigma_{s+t}/\Delta\sigma_{s+t}$, where $\Delta\sigma_{s+t}$ is the half-width of the posterior. In all cases the expected result for each combination is consistent with the standard model cross section of 2.9 pb. We can see that the greatest sensitivity comes from the electron channel with two jets, one b-tagged, with the similar configuration in the muon channel being next. Nevertheless, the 3-jet channels and the two-tagged channels do contribute to the sensitivity. Systematic uncertainties make the sensitivity lower, but this analysis is dominated by low statistics.

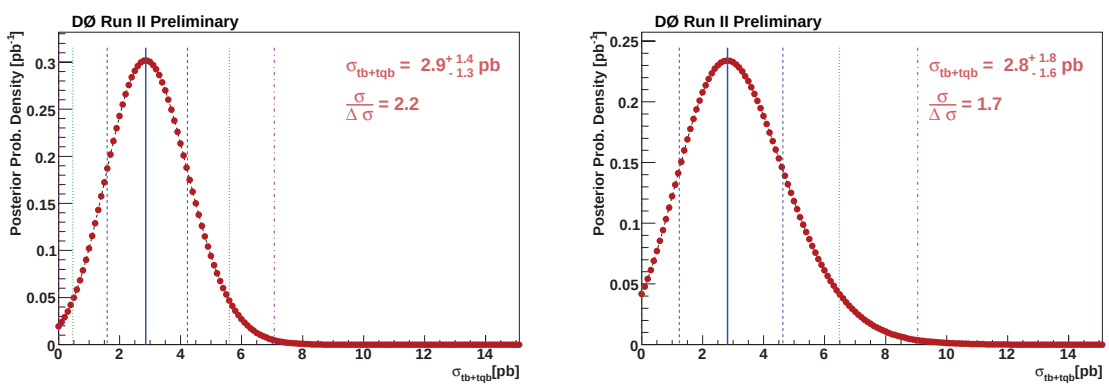


Figure 8.3: Expected 1D posterior plots for the combined $e+\mu \geq 1$ b-tag channel in two-jet events, with statistical uncertainties only (left plot) and with statistical and systematic uncertainties (right plot).

Another measure of the significance of our measurement is the p-value, which is the probability that the background alone could fluctuate up to or above the measured cross section. Figure 8.7 shows the distribution of measured cross sections for an ensemble made up of simulated zero-signal data sets. The data sets had both statistical and systematic variations applied when they were created. From this distribution, we calculate a p-value of 3.1% assuming there is SM single-top with $\sigma = 2.9$ pb, meaning

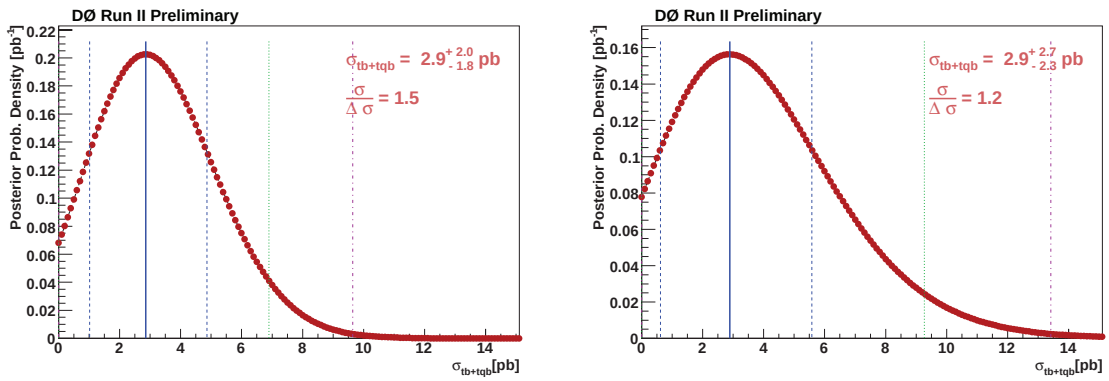


Figure 8.4: Expected 1D posterior plots for the combined $e+\mu \geq 1$ b-tag channel in three-jet events, with statistical uncertainties only (left plot) and with statistical and systematic uncertainties (right plot).

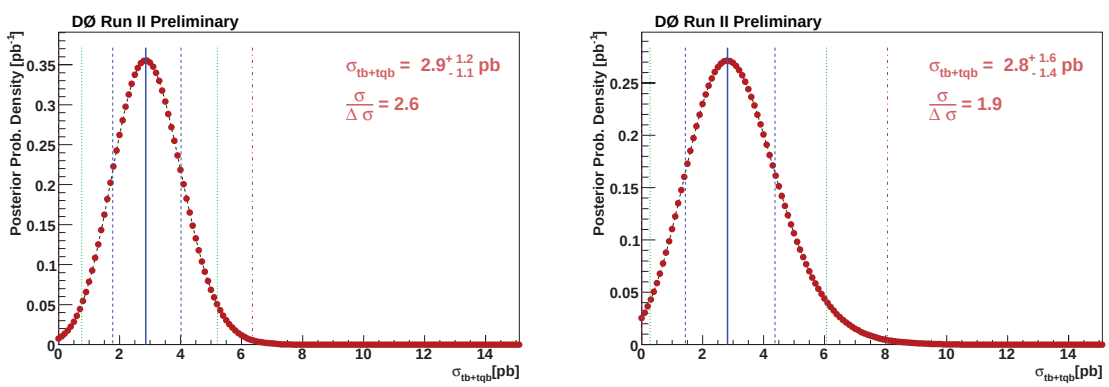


Figure 8.5: Expected 1D posterior plots for all channels combined, with statistical uncertainties only (left plot) and with statistical and systematic uncertainties (right plot).

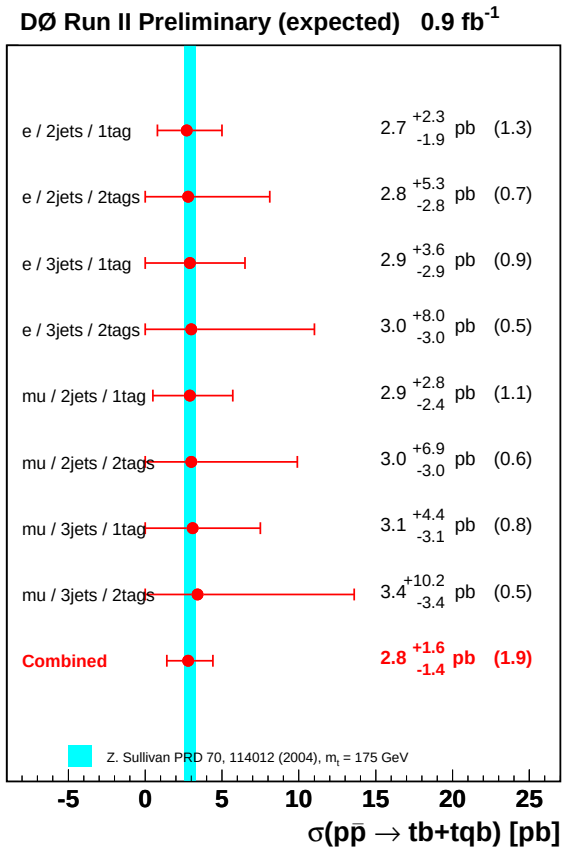


Figure 8.6: Summary plot of the expected single top quark cross sections showing the individual measurements and their combination. The number in parenthesis is a measure of the significance of the measurement, $\sigma_{s+t}/\Delta\sigma_{s+t}$.

	1,2tags + 2,3jets		e,μ + 2,3jets		e,μ + 1,2tags		All
	e-chan	μ-chan	1 tag	2 tags	2 jets	3 jets	channels
Statistics only	$2.9^{+1.5}_{-1.4}$	$2.9^{+1.7}_{-1.6}$	$2.9^{+1.3}_{-1.2}$	$2.9^{+2.4}_{-2.1}$	$2.9^{+1.4}_{-1.3}$	$2.9^{+2.0}_{-1.8}$	$2.9^{+1.2}_{-1.1}$
With systematics	$2.7^{+2.0}_{-1.7}$	$2.9^{+2.3}_{-2.0}$	$2.8^{+1.7}_{-1.5}$	$2.7^{+3.3}_{-2.6}$	$2.8^{+1.8}_{-1.6}$	$2.9^{+2.7}_{-2.3}$	$\mathbf{2.8^{+1.6}_{-1.4}}$

Table 8.2: Expected $s+t$-channel cross sections, without and with systematic uncertainties, for many combinations of the analysis channels. The final expected result of this analysis is shown in the lower right hand corner in bold type.

that only 3.1% of the time zero signal could cause the ME method to measure 2.9 pb or higher. This p-value corresponds to a 1.9σ Gaussian-equivalent significance.

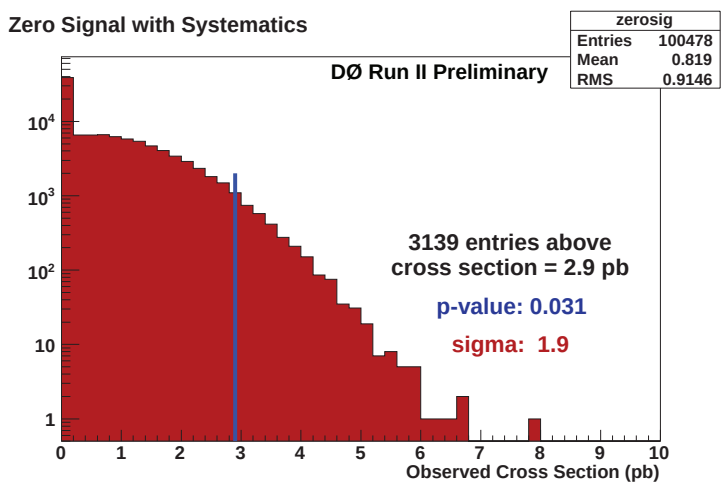


Figure 8.7: Distribution of cross sections from a zero-signal ensemble, with full systematics included, and the expected p-value using $\sigma = 2.9$ pb.

8.6 Observed Results

8.6.1 Discriminant Output

Figures 8.8 and 8.9 shows the s-channel and t-channel discriminant histograms for two-jet and three-jet events respectively including data. The different lepton and number of b-tags channels are combined. The single top quark content is scaled to the measured cross section. It is evident that there is better agreement between data and background+signal versus background-only.

Note that this combination does not take into account the binning or the significance of the measurement of the individual channels, so it is provided only as a check. The individual channel 1D projections of the discriminants that are provided as input to `top_statistics` are shown in Appendix C.

8.6.2 Measured Cross Section

Figure 8.10 shows the observed $s + t$-channel posterior without and with systematic uncertainties for all channels combined. The measured cross section with full systematics is

$$\sigma\left(p\bar{p} \to tb + X,\ tqb + X\right) = 4.8^{+1.6}_{-1.4}\ \text{pb}.$$

This result assumes a SM cross section ratio of $\sigma_s/\sigma_t = 0.44$. Both Table 8.3 and Figure 8.11 shows the measured cross sections from various combinations of analysis channels. The results are consistent with each other, though somewhat above the SM prediction.

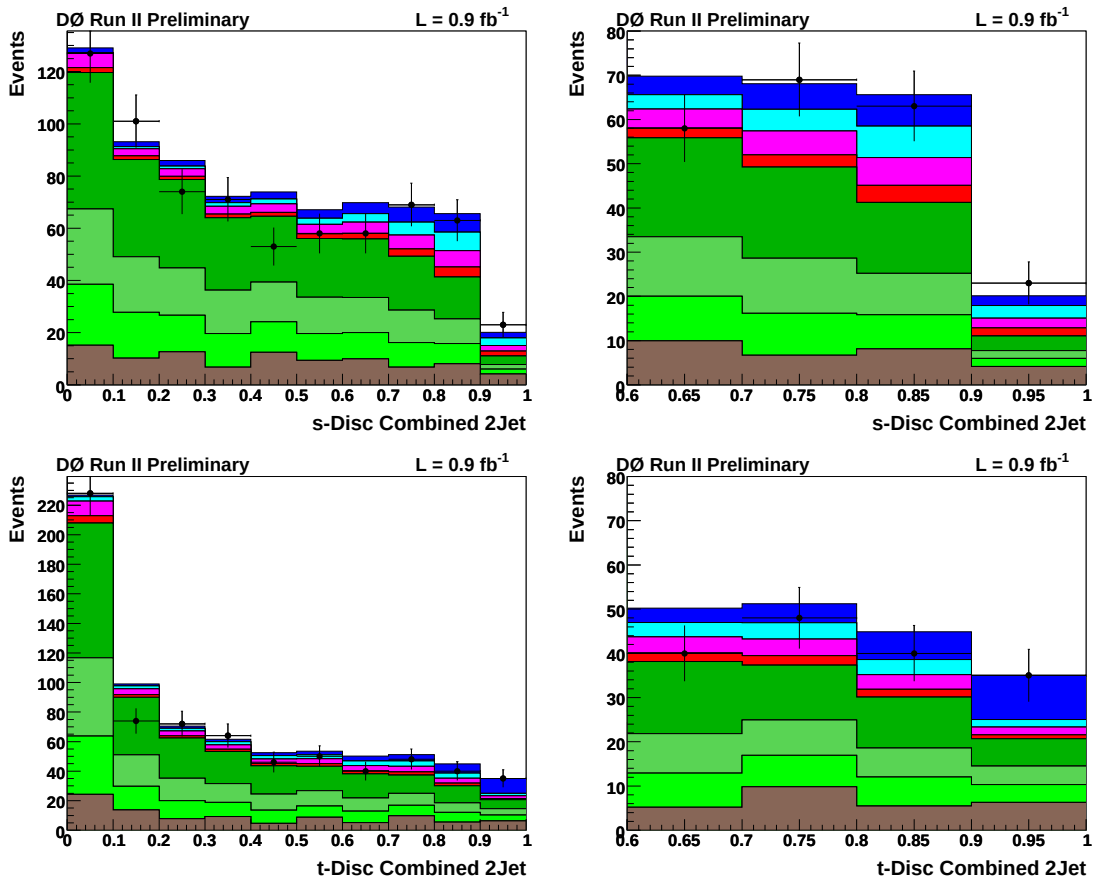


Figure 8.8: Discriminant results for the e+μ channel with two jets and ≥ 1 b tag. Upper row: s-channel discriminant; lower row: t-channel discriminant. Left column: full output range; right column: close-up of the high end of the distributions.

	1,2tags + 2,3jets		e,μ + 2,3jets		e,μ + 1,2tags		All
	e-chan	μ-chan	1 tag	2 tags	2 jets	3 jets	channels
Statistics only	$3.6^{+1.5}_{-1.4}$	$3.7^{+1.7}_{-1.6}$	$2.9^{+1.2}_{-1.1}$	$7.1^{+3.1}_{-2.7}$	$4.0^{+1.4}_{-1.3}$	$3.0^{+1.9}_{-1.7}$	$3.7^{+1.1}_{-1.1}$
With systematics	$4.3^{+2.0}_{-1.7}$	$5.9^{+2.6}_{-2.2}$	$4.6^{+1.8}_{-1.5}$	$7.5^{+4.3}_{-3.5}$	$5.1^{+2.0}_{-1.7}$	$4.4^{+2.7}_{-2.2}$	$\mathbf{4.8^{+1.6}_{-1.4}}$

Table 8.3: Measured s+t-channel cross sections, without and with systematic uncertainties, for many combinations of the analysis channels. The final result of this analysis is shown in the lower right hand corner in bold type.

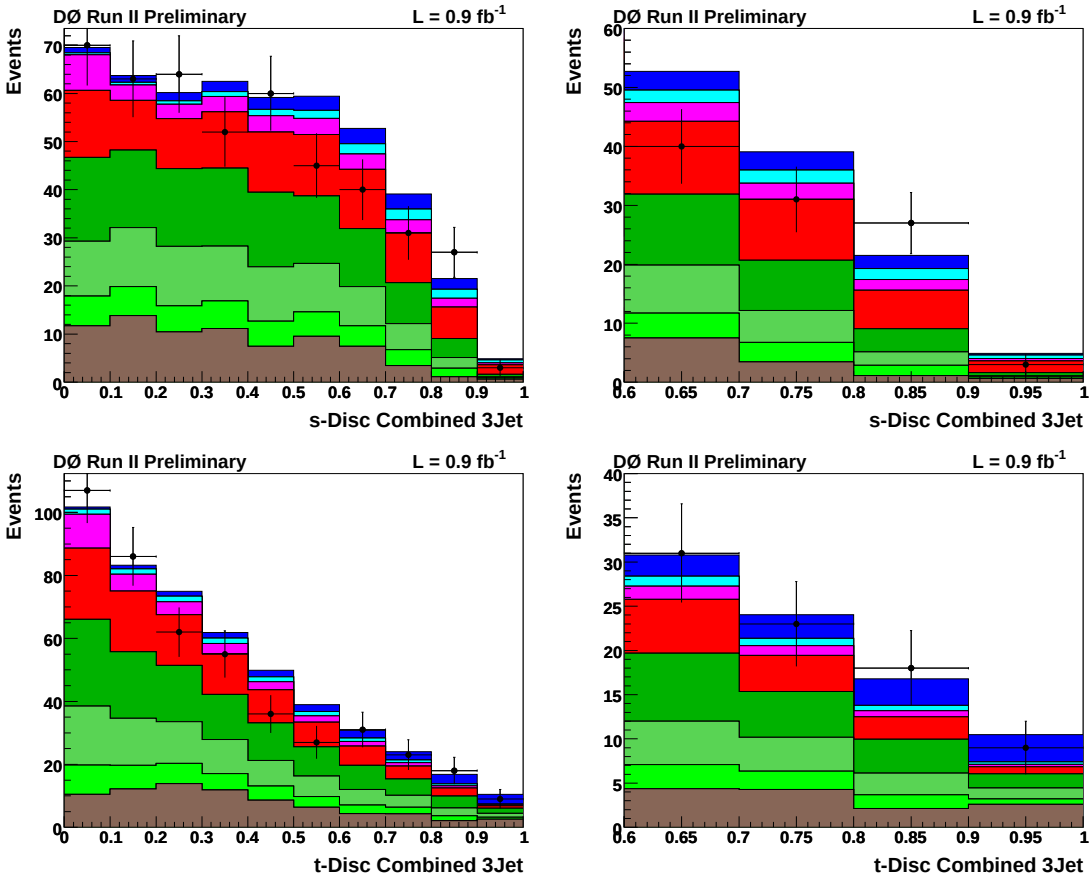


Figure 8.9: Discriminant results for the e+μ channel with three jets and ≥ 1 b tag. Upper row: s-channel discriminant; lower row: t-channel discriminant. Left column: full output range; right column: close-up of the high end of the distributions.

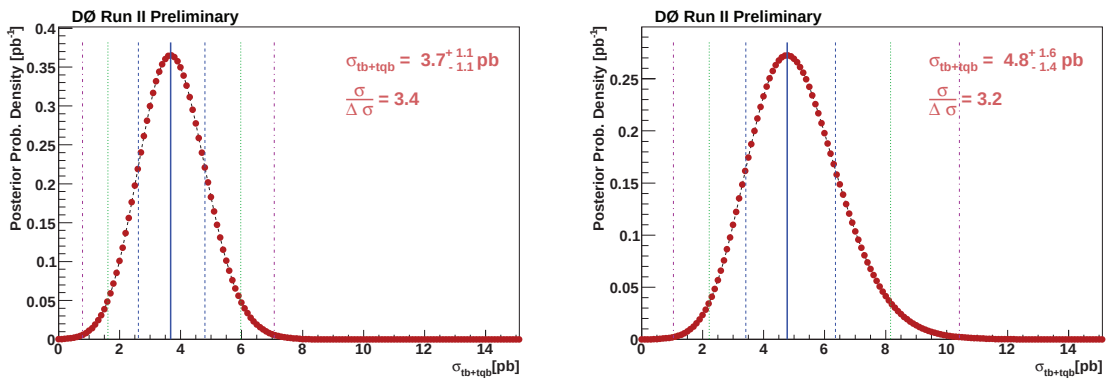


Figure 8.10: Measured 1D posterior plots for all channels combined, with statistical uncertainties only (left plot) and with statistical and systematic uncertainties (right plot).

8.6.3 Signal Significance

Figure 8.12 shows the previously seen distribution of measured cross sections for an ensemble made up of simulated zero-signal data sets. Using the measured cross section of 4.8 pb, from this distribution, we calculate a p-value of 0.08%, meaning that only 0.08% of the time zero signal could cause the ME method to measure $\sigma \geq 4.8$ pb. That p-value corresponds to a Gaussian-equivalent significance of 3.2 standard deviations.

Figure 8.13 shows the distribution of measured cross sections for an ensemble made up of simulated SM signal ($\sigma = 2.9$ pb) data sets. From this distribution, we calculate a p-value of 13%, meaning that 13% of the time a Standard Model signal could cause the ME method to measure $\sigma \geq 4.8$ pb, corresponding to 1.1σ.

8.7 Event Characteristics

As a check that the ME discriminant selects events which resemble the characteristics expected for single-top events, we plot the distributions of variables after applying a

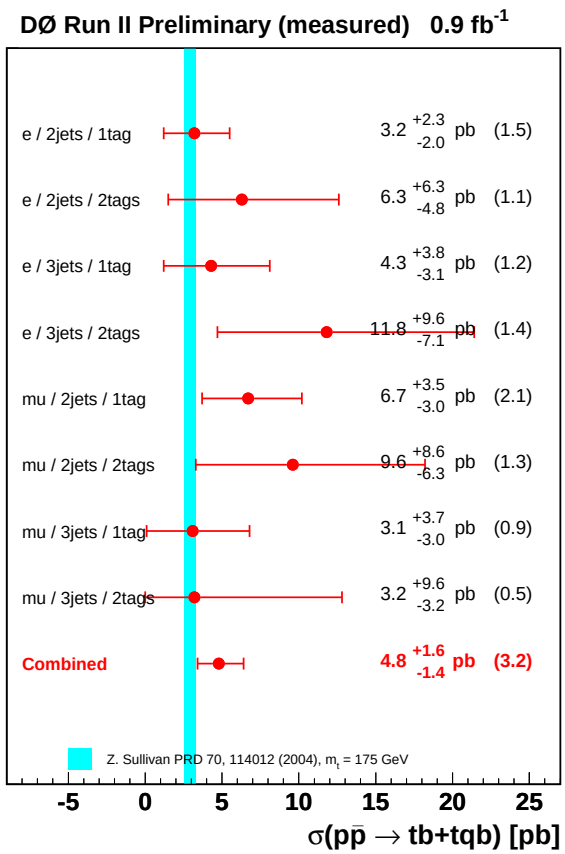


Figure 8.11: Summary plot of the measured single top quark cross sections showing the individual measurements and their combination. The number in parenthesis is a measure of the significance of the measurement, $\sigma_{s+t}/\Delta\sigma_{s+t}$.

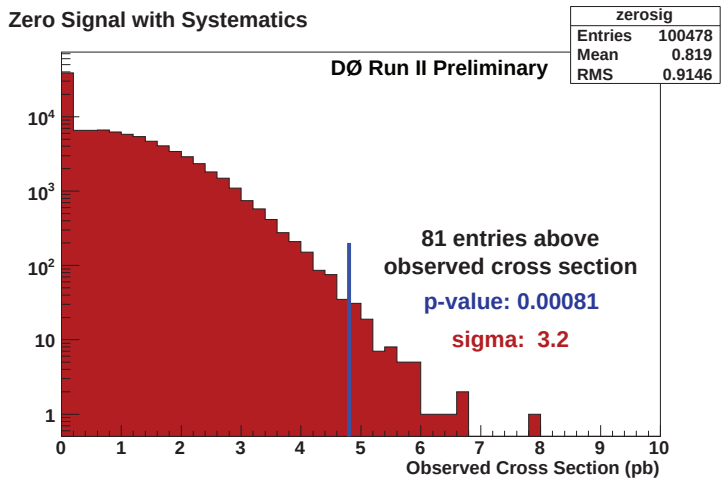


Figure 8.12: Distribution of cross sections from a zero-signal ensemble, with full systematics included, and the observed zero-signal p-value.

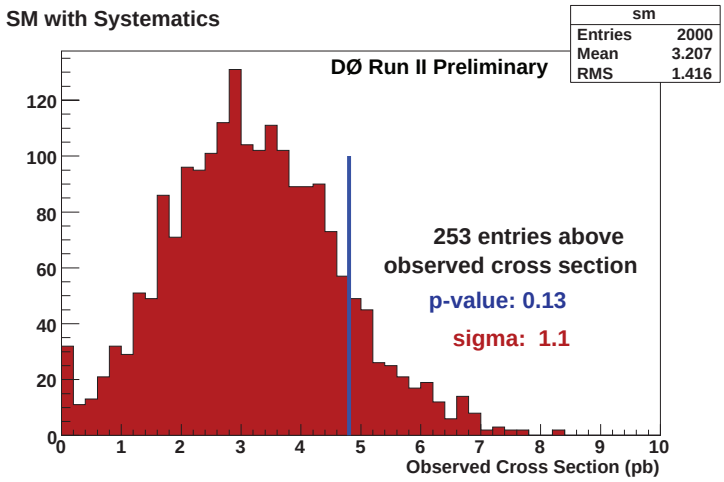


Figure 8.13: Distribution of cross sections from a SM ensemble, with full systemetics included, and the observed SM p-value.

cut on the discriminant value. One variable that has a distinct shape for t-channel single-top events is "$Q \times \eta$," or to be more specific, the charge of the lepton times the pseudorapidity of the untagged jet. It is also expected that the invariant mass of the lepton, $\not{E}_{\mathrm{T}}$, and b-tagged jet be consistent with the top quark mass. Figure 8.14 shows these two variables for different t-channel discriminant cuts. The top row is for all events, the middle is signal-suppressed, and the bottom row is signal-enhanced. The $Q \times \eta$ variable in the single-enhanced row resembles the theoretical prediction, as given in Fig. 3.9(b), suggesting that the discriminant performs as expected. Note that background that mimics this behavior is also enhanced. The invariant mass is consistent with a top quark.

Fig. 8.15 shows the same variables for different s-channel discriminant cuts. As expected, the "$Q \times \eta$" variable does not have a distinct shape for s-channel, but the invariant mass is again consistent with the top quark mass. In all plots, the single top quark content is scaled to the measured cross section.

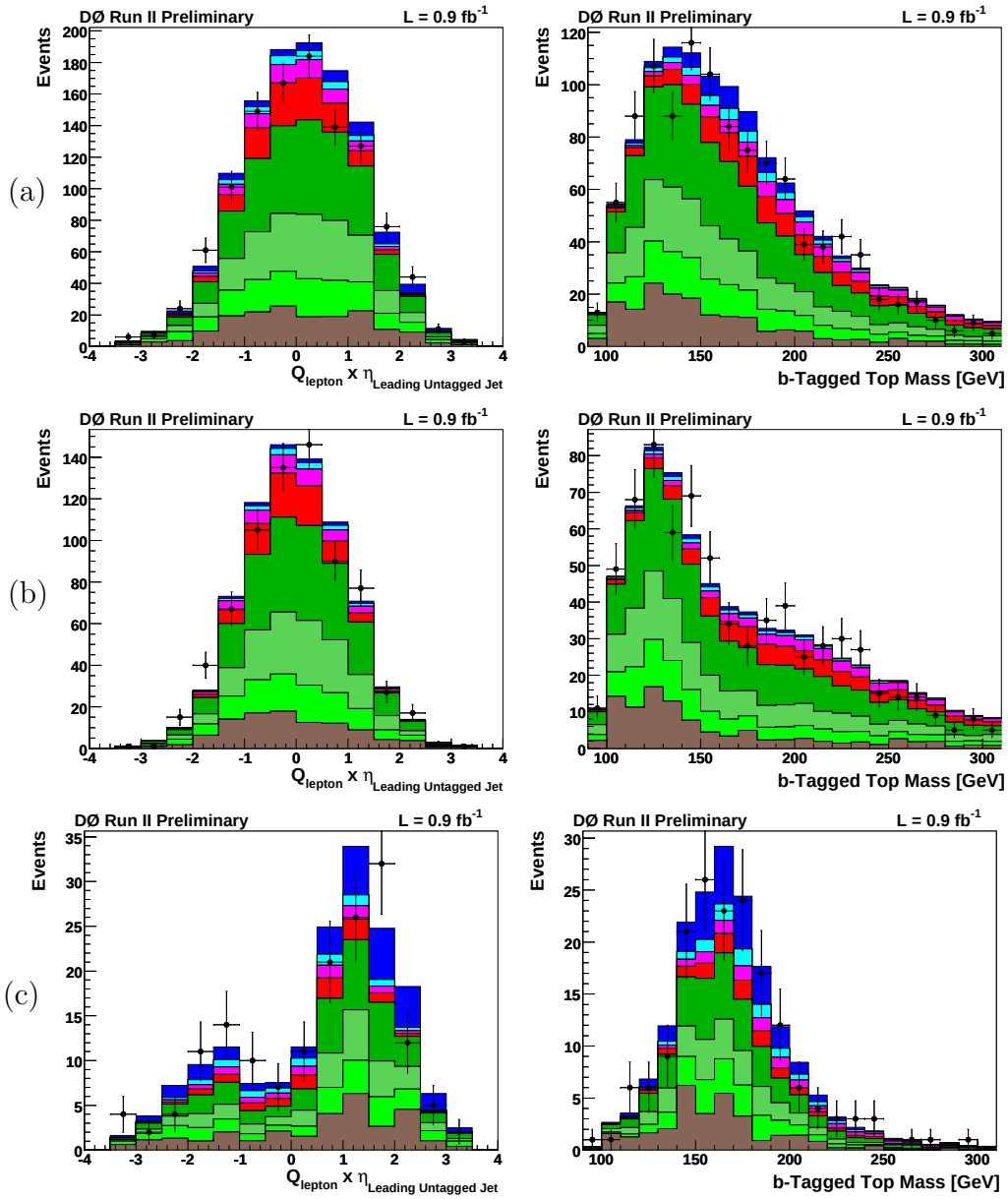


Figure 8.14: The lepton charge times the pseudorapidity of the untagged jet, $Q \times \eta$, (left column) and the invariant mass of the lepton, $\not\!\!E_{\mathrm{T}}$, and tagged jet (right column) for (a) all events, (b) only events that pass a $D_t < 0.4$ t-channel discriminant cut, and (c) only events that pass a $D_t > 0.7$ t-channel discriminant cut.

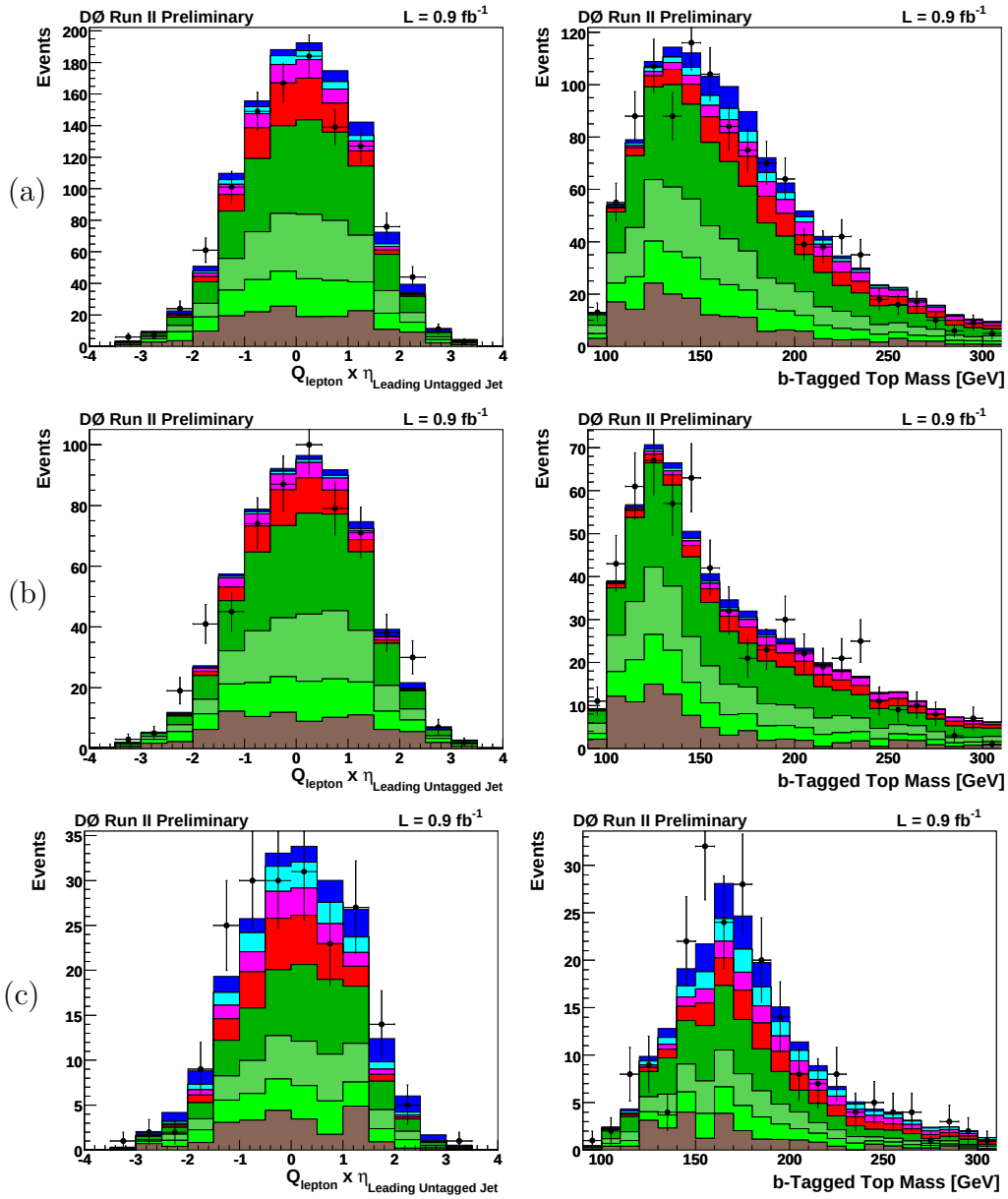


Figure 8.15: The lepton charge times the pseudorapidity of the untagged jet, $Q \times \eta$, (left column) and the invariant mass of the lepton, $\not{E}_{\mathrm{T}}$, and tagged jet (right column) for (a) all events, (b) only events that pass a $D_s < 0.4$ s-channel discriminant cut, and (c) only events that pass a $D_s > 0.7$ s-channel discriminant cut.

Chapter 9

Summary and Conclusions

We have used the matrix elements method to discriminate single top quark processes from background processes in nearly $1\,\mathrm{fb}^{-1}$ of Run II data. It is a second, optimized iteration, the main difference being the addition of a $t\bar{t}$ discriminant for events with three jets. Assuming a SM cross section ratio of $\sigma_s/\sigma_t = 0.44$, we measure:

$$\sigma\left(p\bar{p} \to tb + X,\ tqb + X\right) = 4.8^{+1.6}_{-1.4}\ \mathrm{pb}.$$

This result has a p-value of 0.08%, corresponding to a 3.2 standard deviation Gaussian-equivalent significance, an improvement over the p-value of 0.22% (2.9σ) that was measured in the first iteration.

Analysis Technique	Measured σ_{s+t} (pb)	Observed p-value
BNN	5.0 ± 1.9	0.89% (2.4σ)
ME	$4.6^{+1.8}_{-1.5}$	0.21% (2.9σ)
DT	4.9 ± 1.4	0.04% (3.4σ)

Table 9.1: The single top results published in Ref. [91].

Analysis Technique	Measured σ_{s+t} (pb)	Observed p-value
BNN	$4.4^{+1.6}_{-1.4}$	0.08% (3.1σ)
ME	$4.8^{+1.6}_{-1.4}$	0.08% (3.2σ)

Table 9.2: The improved single top results.

9.1 The Current D0 Single Top Results

9.1.1 Cross Section Measurements

There were three primary techniques employed in the original published result [91]: Bayesian neural networks, matrix elements, and decision trees. The results that were obtained are summarized in in Table 9.1. The analysis using the BNN technique has since been improved, and this thesis has presented the analysis using an improved ME technique. The new results are given in Table 9.2.

The DT result and the improved versions of the BNN and ME results have been combined using the Best Linear Unbiased Estimate (BLUE) method [205], as follows. The combined cross section is calculated with the linear estimate as:

$$\sigma_{\text{comb}} = \sum_j w_j \sigma_j \tag{9.1}$$

where σ_j are the cross sections as calculated by the individual analyses, and w_j is a

weight, determined by:

$$w_i = \frac{\sum_j \mathrm{Cov}^{-1}(\sigma_i, \sigma_j)}{\sum_k \sum_l \mathrm{Cov}^{-1}(\sigma_k, \sigma_l)}. \tag{9.2}$$

The covariances were calculated by running over the SM ensembles, and the following weights were calculated: $w_{\mathrm{DT}} = 0.127$, $w_{\mathrm{BNN}} = 0.386$, and $w_{\mathrm{ME}} = 0.488$. The correlation matrix was found to be

$$\rho = \begin{pmatrix} \mathrm{DT} & \mathrm{ME} & \mathrm{BNN} & \\ 1 & 0.64 & 0.66 & DT \\ 0.64 & 1 & 0.59 & ME \\ 0.66 & 0.59 & 1 & BNN \end{pmatrix}. \tag{9.3}$$

These analyses are highly correlated. Finally, the uncertainty to the combined result is taken to be:

$$\Delta\sigma_{\mathrm{comb}} = \sqrt{\sum_i \sum_j w_i w_j \rho_{ij} \Delta\sigma_i \Delta\sigma_j}. \tag{9.4}$$

Assuming a SM cross section ratio of $\sigma_s/\sigma_t = 0.44$, the combined result for the single top cross section is

$$\sigma\left(p\bar{p} \to tb + X,\ tqb + X\right) = 4.7 \pm 1.3 \text{ pb}.$$

These results are summarized in Fig. 9.1.

Figure 9.2 shows the measured cross sections for the ME, BNN, and DT analyses for an ensemble made up of simulated zero-signal data sets, along with the measurement using the combination. Using the combined measured cross section of 4.7 pb, from this distribution, we calculate a p-value of 0.014%, meaning that only 0.014% of the time zero signal could cause the combination to measure $\sigma \geq 4.7$ pb. That

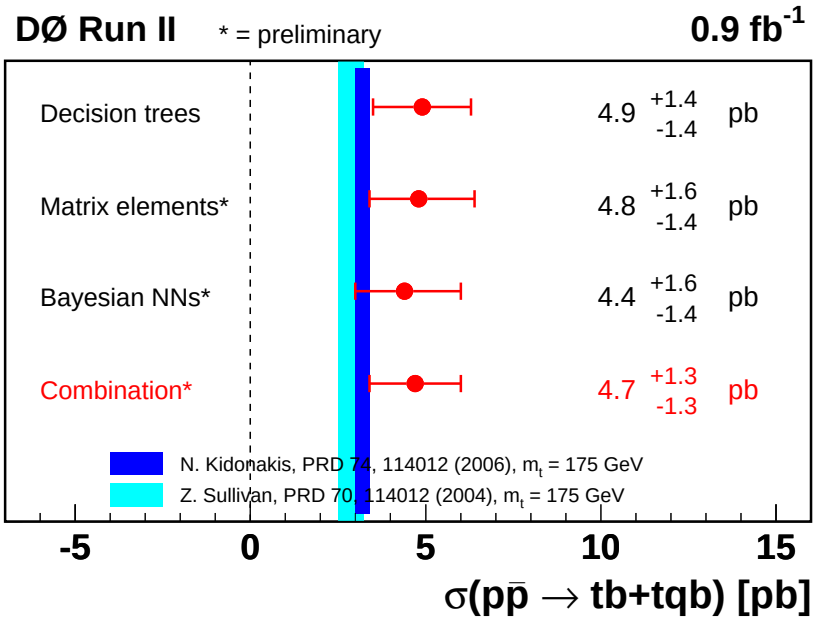


Figure 9.1: The single top cross section measurement of the individual analyses and their combination [205].

p-value corresponds to a Gaussian-equivalent significance of 3.6 standard deviations.

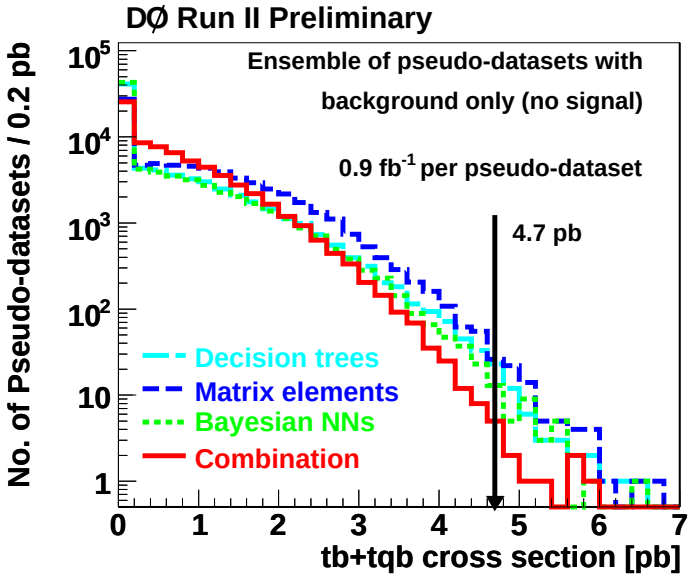


Figure 9.2: Distribution of cross sections from a zero-signal ensemble, with full systematics included, and the observed zero-signal p-value for the combined cross section [205].

9.1.2 The First Direct Measurement of $|V_{tb}|$

One of the primary reasons to study single top production is to measure $|V_{tb}|$. This was done for the DT analysis and was included in published result [91], and the method is explained with more detail in Ref. [206]. This result makes no assumptions on the unitarity of CKM or the number of families, but it does require a few assumptions. The first assumption is that the observed single top results were produced by an interaction with a W boson, not from one of the various BSM scenarios that include extra scalar and vector bosons or FCNC interactions. The second assumption is that

Uncertainty Component	s-channel	t-channel
Top quark mass	8.5	13.0
Factorization scale	4.0	5.5
PDF	4.5	10.0
α_s	1.4	0.01

Table 9.3: Systematic uncertainties in percent on the cross section factor required to extract $|V_{tb}|$.

$|V_{tb}|^2 \gg |V_{td}|^2 + |V_{ts}|^2$, which is experimentally supported by the $\mathrm{B}(t \to Wb)/\mathrm{B}(t \to Wq)$ measurements done on $t\bar{t}$ events [58, 57]. Lastly, it is assumed that the Wtb vertex is CP-conserving and of the V−A form, though possibly of anomalous strength:

$$\Gamma^{\mu}_{Wtb} = -\frac{g}{\sqrt{2}} V_{tb} \bar{u}(p_b) \gamma^{\mu} f_1^L P_L u(p_t). \tag{9.5}$$

Additional theoretical uncertainties needed to be added to the measurement of $|V_{tb}|$ [77], and they are given in Table 9.3.

Two measurements were performed, one for $|V_{tb} f_1^L|$, with no requirement that it be less than one, and one assuming $f_1^L = 1$, resulting in $|V_{tb}|$ being restricted to between zero and one. The limits were calculated using a Baysian approach, and the resulting posteriors for $|V_{tb} f_1^L|^2$ and $|V_{tb}|^2$ are given in Fig. 9.3. The final results for the two sets of assumptions are:

- $|V_{tb} f_1^L|^2 = 1.72^{+0.64}_{-0.54}$, corresponding to $|V_{tb} f_1^L| = 1.31^{+0.25}_{-0.21}$
- $|V_{tb}|^2 = 1.0^{+0.0}_{-0.24}$, corresponding to $|V_{tb}| > 0.68$ at 95% C.L.

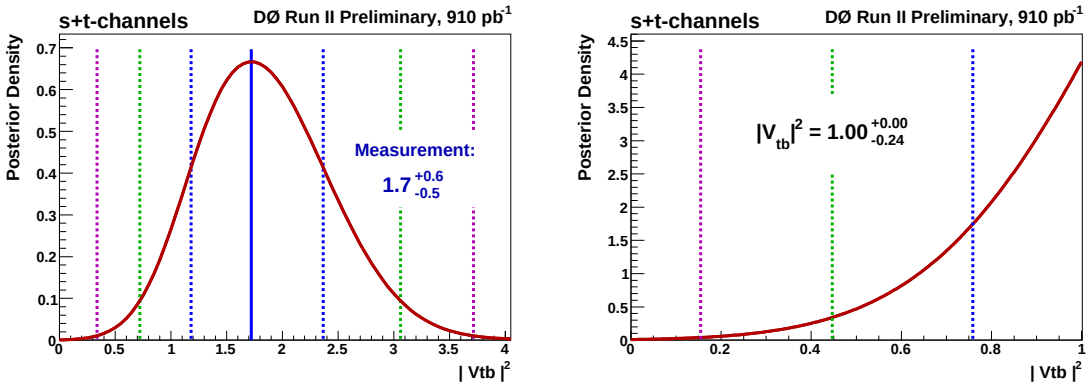


Figure 9.3: Posterior probability density for $|V_{tb}f_1^L|^2$ (left)and for $|V_{tb}|^2$ (right).

9.2 Outlook

Figure 9.4 shows the projected significance of the s-channel and t-channel signals as a function of the integrated luminosity at D0, assuming a SM cross section. The current measurements are somewhat higher than the SM, so it will be interesting to see if that feature stays. At the end of Run II at the Tevatron Collider, single top production will have been discovered both $s+t$-channel combined, and t-channel by itself. It is not known if s-channel by itself will have a significance over five standard deviations, but it should have at least three sigma evidence.

The LHC presents a transition from having a search for single top to having precision measurements with single top [56]. Though s-channel production will remain difficult, the order of magnitude increase in the t-channel and tW associated production channels will allow for clean single top samples. There can be an emphasis on trying to extract cleaner samples. Also, the analyses could become more cuts-based, which has the benefit of distorting the background less, allowing for lower systematics. As Fig. 8.14 shows, multivariate techniques tend to shape the background to

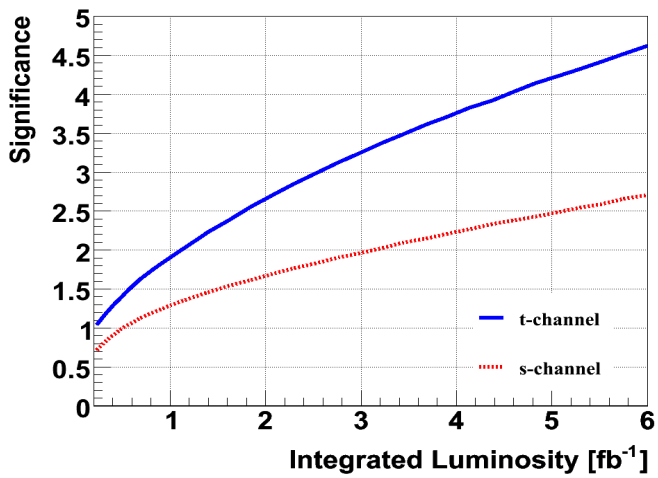


Figure 9.4: The single top measurement projections from D0.

look like the signal.

After $30\,\mathrm{fb}^{-1}$ of integrated luminosity, the statistical uncertainty on $|V_{tb}|^2$ will be less that 2% [207], though it will be a challenge to reduce other uncertainties to that level. If all other uncertainties are kept at 5% or less, the uncertainty to $|V_{tb}|^2$ should be 10% or less. Therefore, the goal is to measure V_{tb} with an accuracy of 5% or less. A measurement of the polarization of single top quarks produced via the t-channel process should have a statistical accuracy of 1.6% with $10\,\mathrm{fb}^{-1}$ of integrated luminosity.

Single top physics at the LHC will present plenty of opportunities to probe the EW interactions of the top quark. From the initial searches that set limits to single top production, to the current evidence for single top, to the discovery of single top production in the later phases of Run II at the Tevatron Collider, to precision physics at the LHC, one could argue that the physics is getting richer and richer, and potentially more exciting. There are many reasons to believe that BSM effects might

become first visible in electroweak interactions of the top quark, so the precision tests of such interactions have the potential of great rewards.

Bibliography

[1] LEP Electroweak Working Group. LEP EW WG plots for the winter 2007. URL `http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2007/`.

[2] Abachi, S. *et al.* Observation of the top quark. *Phys. Rev. Lett.* **74**, 2632–2637 (1995).

[3] Abe, F. *et al.* Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.* **74**, 2626–2631 (1995).

[4] Quigg, C. *Gauge Theories of the Strong, Weak and Electromagnetic Interactions*, vol. 56 (Westview Press, 1983).

[5] Cheng, T. P. & Li, L. F. *Gauge Theory of Elementary Particle Phyiscs* (Oxford, UK: Claredon, 1984).

[6] Greiner, W. & Muller, B. *Gauge Theory of Weak Interactions* (Berlin, Germany: Springer, 1993), third edn.

[7] Griffiths, D. J. *Introduction to Elementary Particles* (New York, USA: Wiley, 1987).

[8] Peskin, M. E. & Schroeder, D. V. *An Introduction to Quantum Field Theory* (Reading, USA: Addison-Wesley, 1995).

[9] Yao, W. M. *et al.* Review of particle physics. *J. Phys.* **G33**, 1–1232 (2006). URL `http://pdg.lbl.gov`.

[10] Fukuda, Y. *et al.* Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.* **81**, 1562–1567 (1998).

[11] Ahmad, Q. R. *et al.* Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.* **89**, 011301 (2002).

[12] Arnold, R. *et al.* First results of the search of neutrinoless double beta decay with the NEMO 3 detector. *Phys. Rev. Lett.* **95**, 182302 (2005).

[13] The Tevatron Electroweak Workging Group. A combination of CDF and D0 results on the mass of the top quark (2007). `hep-ex/0703034`.

[14] Noether, E. Invariante variationsprobleme. *Nachr. Kgl. Ges. Wiss. Göttingen* 235–257 (1918).

[15] Weyl, H. A new extension of relativity theory. *Annalen Phys.* **59**, 101–133 (1919).

[16] Fock, V. On the invariant form of the wave equations and the equations of motion for a charged point mass. *Z. Phys.* **39**, 226–232 (1926).

[17] London, F. Quantum mechanical interpretation of the Weyl's theory. *Z. Phys.* **42**, 375–389 (1927).

[18] Weyl, H. Electron and gravitation. *Z. Phys.* **56**, 330–352 (1929).

[19] Yang, C.-N. & Mills, R. L. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.* **96**, 191–195 (1954).

[20] Pumplin, J. *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *JHEP* **07**, 012 (2002).

[21] Fermi, E. An attempt of a theory of beta radiation. 1. *Z. Phys.* **88**, 161–177 (1934).

[22] Fermi, E. Trends to a theory of beta radiation. *Nuovo Cim.* **11**, 1–19 (1934).

[23] Lee, T. D. & Yang, C.-N. Question of parity conservation in weak interactions. *Phys. Rev.* **104**, 254–258 (1956).

[24] Wu, C. S. *et al.* Experimental test of parity conservation in beta decay. *Phys. Rev.* **105**, 1413–1414 (1957).

[25] Christenson, J. H. *et al.* Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.* **13**, 138–140 (1964).

[26] Schwinger, J. S. A theory of the fundamental interactions. *Annals Phys.* **2**, 407–434 (1957).

[27] Nambu, Y. Quasi-particles and gauge invariance in the theory of superconductivity. *Phys. Rev.* **117**, 648–663 (1960).

[28] Nambu, Y. & Jona-Lasinio, G. Dynamical model of elementary particles based on an analogy with superconductivity. I. *Phys. Rev.* **122**, 345–358 (1961).

[29] Nambu, Y. & Jona-Lasinio, G. Dynamical model of elementary particles based on an analogy with superconductivity. II. *Phys. Rev.* **124**, 246–254 (1961).

[30] Goldstone, J. Field theories with superconductor solutions. *Nuovo Cim.* **19**, 154–164 (1961).

[31] Goldstone, J. *et al.* Broken symmetries. *Phys. Rev.* **127**, 965–970 (1962).

[32] Higgs, P. W. Broken symmetries, massless particles and gauge fields. *Phys. Lett.* **12**, 132–133 (1964).

[33] Englert, F. & Brout, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13**, 321–322 (1964).

[34] Guralnik, G. S. *et al.* Global conservation laws and massless particles. *Phys. Rev. Lett.* **13**, 585–587 (1964).

[35] Higgs, P. W. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.* **145**, 1156–1163 (1966).

[36] Glashow, S. L. Partial symmetries of weak interactions. *Nucl. Phys.* **22**, 579–588 (1961).

[37] Salam, A. & Ward, J. C. Electromagnetic and weak interactions. *Phys. Lett.* **13**, 168–171 (1964).

[38] Weinberg, S. A model of leptons. *Phys. Rev. Lett.* **19**, 1264–1266 (1967).

[39] 't Hooft, G. Renormalization of massless Yang-Mills fields. *Nucl. Phys.* **B33**, 173–199 (1971).

[40] 't Hooft, G. Renormalizable Lagrangians for massive Yang-Mills fields. *Nucl. Phys.* **B35**, 167–188 (1971).

[41] Hasert, F. J. *et al.* Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment. *Phys. Lett.* **B46**, 138–140 (1973).

[42] Nambu, Y. Quasisupersymmetry, bootstrap symmetry breaking and fermion masses. In *Proc. of 1988 Int. Workshop New Trends in Strong Coupling Gauge Theories* (Nagoya,, 1988).

[43] Miransky, V. A. *et al.* Dynamical electroweak symmetry breaking with large anomalous dimension and t quark condensate. *Phys. Lett.* **B221**, 177 (1989).

[44] Miransky, V. A. *et al.* Is the t quark responsible for the mass of W and Z bosons? *Mod. Phys. Lett.* **A4**, 1043 (1989).

[45] Hill, C. T. Topcolor: Top quark condensation in a gauge extension of the standard model. *Phys. Lett.* **B266**, 419–424 (1991).

[46] Hill, C. T. Topcolor assisted technicolor. *Phys. Lett.* **B345**, 483–489 (1995).

[47] Lane, K. D. & Eichten, E. Natural topcolor assisted technicolor. *Phys. Lett.* **B352**, 382–387 (1995).

[48] Dobrescu, B. A. & Hill, C. T. Electroweak symmetry breaking via top condensation seesaw. *Phys. Rev. Lett.* **81**, 2634–2637 (1998).

[49] Chivukula, R. S. *et al.* Top quark seesaw theory of electroweak symmetry breaking. *Phys. Rev.* **D59**, 075003 (1999).

[50] Muller, D. J. & Nandi, S. Topflavor: A separate SU(2) for the third family. *Phys. Lett.* **B383**, 345–350 (1996).

[51] Malkawi, E. *et al.* A model of strong flavor dynamics for the top quark. *Phys. Lett.* **B385**, 304–310 (1996).

[52] He, H.-J. *et al.* New topflavor models with seesaw mechanism. *Phys. Rev.* **D62**, 011702 (2000).

[53] Hill, C. T. & Simmons, E. H. Strong dynamics and electroweak symmetry breaking. *Phys. Rept.* **381**, 235–402 (2003).

[54] Cabibbo, N. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.* **10**, 531–532 (1963).

[55] Kobayashi, M. & Maskawa, T. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.* **49**, 652–657 (1973).

[56] Gerber, C. E. *et al.* Tevatron-for-LHC report: Top and electroweak physics (2007). `arXiv:0705.3251 [hep-ph]`.

[57] Abazov, V. M. *et al.* Measurement of $b(t \to wb)/b(t \to wq)$ at $\sqrt{s} = 1.96\,\mathrm{TeV}$. *Phys. Lett.* **B639**, 616–622 (2006).

[58] Acosta, D. *et al.* Measurement of $b(t \to wb)/b(t \to wq)$ at the Collider Detector at Fermilab. *Phys. Rev. Lett.* **95**, 102002 (2005).

[59] Swain, J. & Taylor, L. First determination of the quark mixing matrix element V_{tb} independent of assumptions of unitarity. *Phys. Rev.* **D58**, 093006 (1998).

[60] Willenbrock, S. S. D. & Dicus, D. A. Production of heavy quarks from w gluon fusion. *Phys. Rev.* **D34**, 155 (1986).

[61] Yuan, C. P. A new method to detect a heavy top quark at the tevatron. *Phys. Rev.* **D41**, 42 (1990).

[62] Cortese, S. & Petronzio, R. The single top production channel at tevatron energies. *Phys. Lett.* **B253**, 494–498 (1991).

[63] Ellis, R. K. & Parke, S. J. Top quark production by W-gluon fusion. *Phys. Rev.* **D46**, 3785–3788 (1992).

[64] Carlson, D. O. & Yuan, C. P. Studying the top quark via the W-gluon fusion process. *Phys. Lett.* **B306**, 386–390 (1993).

[65] Stelzer, T. & Willenbrock, S. Single top quark production via $q\bar{q} \to t\bar{b}$. *Phys. Lett.* **B357**, 125–130 (1995).

[66] Heinson, A. P. *et al.* Single top quarks at the Fermilab Tevatron. *Phys. Rev.* **D56**, 3114–3128 (1997).

[67] Belyaev, A. S. *et al.* Single top quark at future hadron colliders: Complete signal and background study. *Phys. Rev.* **D59**, 075001 (1999).

[68] Stelzer, T. *et al.* Single top quark production at hadron colliders. *Phys. Rev.* **D58**, 094021 (1998).

[69] Tait, T. M. P. The tW^- mode of single top production. *Phys. Rev.* **D61**, 034001 (2000).

[70] Bowen, M. T. *et al.* In search of lonely top quarks at the Tevatron. *Phys. Rev.* **D72**, 074016 (2005).

[71] Sullivan, Z. Angular correlations in single-top-quark and Wjj production at next-to-leading order. *Phys. Rev.* **D72**, 094034 (2005).

[72] Bordes, G. & van Eijk, B. Calculating QCD corrections to single top production in hadronic interactions. *Nucl. Phys.* **B435**, 23–58 (1995).

[73] Smith, M. C. & Willenbrock, S. QCD and Yukawa corrections to single-top-quark production via $q\bar{q} \to t\bar{b}$. *Phys. Rev.* **D54**, 6696–6702 (1996).

[74] Stelzer, T. *et al.* Single-top-quark production via W-gluon fusion at next-to-leading order. *Phys. Rev.* **D56**, 5919–5927 (1997).

[75] Belyaev, A. & Boos, E. Single top quark $tW + X$ production at the LHC: A closer look. *Phys. Rev.* **D63**, 034012 (2001).

[76] Harris, B. W. *et al.* The fully differential single top quark cross section in next-to-leading order QCD. *Phys. Rev.* **D66**, 054024 (2002).

[77] Sullivan, Z. Understanding single-top-quark production and jets at hadron colliders. *Phys. Rev.* **D70**, 114012 (2004).

[78] Campbell, J. *et al.* Single top production and decay at next-to-leading order. *Phys. Rev.* **D70**, 094012 (2004).

[79] Cao, Q.-H. & Yuan, C. P. Single top quark production and decay at next-to-leading order in hadron collision. *Phys. Rev.* **D71**, 054022 (2005).

[80] Cao, Q.-H. *et al.* Next-to-leading order corrections to single top quark production and decay at the Fermilab Tevatron. I: s-channel process. *Phys. Rev.* **D71**, 054023 (2005).

[81] Cao, Q.-H. *et al.* Next-to-leading order corrections to single top quark production and decay at the Fermilab Tevatron. II: t-channel process. *Phys. Rev.* **D72**, 094027 (2005).

[82] Campbell, J. & Tramontano, F. Next-to-leading order corrections to Wt production and decay. *Nucl. Phys.* **B726**, 109–130 (2005).

[83] Kidonakis, N. Single top production at the Tevatron: Threshold resummation and finite-order soft gluon corrections. *Phys. Rev.* **D74**, 114012 (2006).

[84] Kidonakis, N. Higher-order soft gluon corrections in single top quark production at the LHC. *Phys. Rev.* **D75**, 071501 (2007).

[85] Bigi, I. I. Y. *et al.* Production and decay properties of ultraheavy quarks. *Phys. Lett.* **B181**, 157 (1986).

[86] Mahlon, G. & Parke, S. J. Improved spin basis for angular correlation studies in single top quark production at the Tevatron. *Phys. Rev.* **D55**, 7249–7254 (1997).

[87] Mahlon, G. & Parke, S. J. Single top quark production at the LHC: Understanding spin. *Phys. Lett.* **B476**, 323–330 (2000).

[88] Mahlon, G. Observing spin correlations in single top production and decay (2000). `hep-ph/0011349`.

[89] Boos, E. E. & Sherstnev, A. V. Spin effects in processes of single top quark production at hadron colliders. *Phys. Lett.* **B534**, 97–105 (2002).

[90] Abazov, V. M. *et al.* Multivariate searches for single top quark production with the D0 detector. *Phys. Rev.* **D75**, 092007 (2007).

[91] Abazov, V. M. *et al.* Evidence for production of single top quarks and first direct measurement of $|V_{tb}|$. *Phys. Rev. Lett.* **98**, 98181802 (2007).

[92] The D0 Collaboration. An improved search for single top quarks using Bayesian neural networks (2007). D0 Note 5397–CONF.

[93] Tait, T. & Yuan, C. P. Single top quark production as a window to physics beyond the standard model. *Phys. Rev.* **D63**, 014018 (2001).

[94] Dalitz, R. H. & Goldstein, G. R. The decay and polarization properties of the top quark. *Phys. Rev.* **D45**, 1531–1543 (1992).

[95] Thompson, J. Introduction to colliding beams at Fermilab (1994). FERMILAB-TM-1909.

[96] Design report Tevatron 1 project (1984). FERMILAB-DESIGN-1984-01.

[97] Run II handbook. URL `http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html`.

[98] Provided by Fermilab Visual Media Services.

[99] Schmidt, C. W. & Curtis, C. D. A 50-milliampere negative hydrogen ion source. *IEEE Trans. Nucl. Sci.* **26**, 4120–4122 (1979).

[100] Patterson, D. The FNAL 200-MeV Linac (1986). FERMILAB-MISC-1986-01.

[101] Curtis, C. D. *et al.* Linac H^- beam operation and uses at Fermilab. *IEEE Trans. Nucl. Sci.* **26**, 3760–3762 (1979).

[102] Fermilab Linac upgrade conceptual design revision 4A (1989). FERMILAB-LU-CONCEPTUAL-DESIGN.

[103] Hubbard, E. L. Booster synchrotron (1973). FERMILAB-TM-0405.

[104] Hojvat, C. *et al.* The multiturn charge exchange injection system for the Fermilab Booster accelerator. *IEEE Trans. Nucl. Sci.* **26**, 3149–3151 (1979).

[105] The Fermilab Main Injector technical design handbook. URL `http://www-fmi.fnal.gov/fmiinternal/MI_Technical_Design/index.html`.

[106] The antiproton source rookie book, version 1.1 (1999). URL `http://www-bdnew.fnal.gov/pbar/documents/PBAR_Rookie_Book.pdf`.

[107] Jackson, G. The Fermilab recycler ring technical design report. rev. 1.2 (1996). FERMILAB-TM-1991.

[108] Nagaitsev, S. *et al.* Experimental demonstration of relativistic electron cooling. *Phys. Rev. Lett.* **96**, 044801 (2006).

[109] Abachi, S. *et al.* The D0 detector. *Nucl. Instrum. Meth.* **A338**, 185–253 (1994).

[110] Abazov, V. M. *et al.* The upgraded D0 detector. *Nucl. Instrum. Meth.* **A565**, 463–537 (2006).

[111] D0 Upgrade Collaboration. D0 silicon tracker technical design report (1994). URL `http://www-d0.fnal.gov/~lipton/tdr_final.ps`. D0 Note 2169.

[112] Adams, D. *et al.* The D0 upgrade: Central fiber tracker, technical design report (1999). D0 Note 4164.

[113] Brzezniak, J. *et al.* Conceptual design of a 2 Tesla superconducting solenoid for the Fermilab D0 detector upgrade (1994). FERMILAB-TM-1886.

[114] D0 Layer 0 Silicon Group. Conceptual design report of the layer 0 silicon detector. D0 Note 4415.

[115] Leo, W. R. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach* (Berlin, Germany: Springer, 1987).

[116] Fabjan, C. W. & Gianotti, F. Calorimetry for particle physics. *Rev. Mod. Phys.* **75**, 1243–1286 (2003).

[117] Adams, M. *et al.* Level-2 calorimeter preprocessor technical design report (1999). D0 Note 3615.

[118] Abazov, V. M. *et al.* The muon system of the Run II D0 detector. *Nucl. Instrum. Meth.* **A552**, 372–398 (2005).

[119] Davis, K. *et al.* Level 1 muon trigger technical design report (1999). URL `http://atlas.physics.arizona.edu/~johns/l1muo/l1mu/Summary.PDF`.

[120] Temple, J. Advanced CalMuo shifter tasks for Level 1 Muon. URL `http://www-d0online.fnal.gov/www/groups/l1muo/guides/AdvShifter.pdf`.

[121] Maciel, A. *et al.* The l2-muon trigger methods and algorithms (2005). D0 Note 4756.

[122] Abolins, M. *et al.* The Run IIb trigger upgrade for the D0 experiment. *IEEE Trans. Nucl. Sci.* **51**, 340–344 (2004).

[123] Edwards, T. L. *et al.* Determination of the effective inelastic $p\bar{p}$ cross-section for the D0 Run II luminosity measurement. FERMILAB-TM-2278-E.

[124] Andeen, T. *et al.* The D0 experiment's integrated luminosity for Tevatron Run IIa. FERMILAB-TM-2365.

[125] Khanov, A. HTF: histogramming method for finding tracks. the algorithm descpription. (2000). D0 Note 3778.

[126] Frühwirth, R. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.* **A262**, 444–450 (1987).

[127] Greenlee, H. The D0 Kalman track fit (2004). D0 Note 4303.

[128] Borissov, G. Ordering a chaos or... techical details of AA tracking (2003). URL `http://www-d0.fnal.gov/atwork/adm/d0_private/2003-02-28/adm_talk.ps`.

[129] Schwartzman, A. & Tully, C. Primary vertex reconstruction by means of adaptive vertex fitting (2005). D0 Note 4918.

[130] Schwartzman, A. & Narain, M. Probabilistic primary vertex selection (2002). URL `http://www-clued0.fnal.gov/~aran/m_4042.ps`. D0 Note 4042.

[131] Bernardi, G. *et al.* NADA: a new event by event hot cell killer (2000). D0 Note 3687.

[132] Bernardi, G. & Kado, M. Hot cell suppression at level 3 (2003). D0 Note 4039.

[133] Bernardi, G. & Trincaz-Duvoid, S. Improvement of the NADA algorithm: Hot cell killing in D0 Run II data (2002). D0 Note 4057.

[134] Bassler, U. & Bernardi, G. Towards a coherent treatment of calorimetric energies: Missing transverse energy, jets, E.M. objects and the T42 algorithm (2002). D0 Note 4124.

[135] Vlimant, J.-R. *et al.* Technical description of the T42 algorithm for the calorimeter noise suppression (2003). D0 Note 4146.

[136] Bernardi, G. *et al.* Improvements from the T42 algorithm on calorimeter objects reconstruction (2004). D0 Note 4335.

[137] Blazey, G. C. *et al.* Run II jet physics. In *Proceedings of the Physics at Run II: QCD and Weak Boson Physics Workshop* (Batavia, 1999). `hep-ex/0005012`.

[138] Busato, E. & Andrieu, B. Jet algorithms in the D0 Run II software: Description and user's guide (2004). D0 Note 4457.

[139] Harel, A. Jet ID optimization (2005). D0 Note 4919.

[140] D0 preliminary jet energy scale. URL http://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/.

[141] Kozminski, J. *et al.* Electron likelihood in p14 (2004). D0 Note 4449.

[142] Getting started with electron/photon ID. URL http://www-d0.fnal.gov/phys_id/emid/d0_private/emid_intro.html.

[143] Kumar, A. *et al.* Electron likelihood study (2005). D0 Note 4769.

[144] Hays, J. *et al.* Single electron efficiencies in p17 data and Monte-Carlo using p18.05.00 d0correct (2006). D0 Note 5105.

[145] Narain, M. Electron identification in the D0 detector. In *The Fermilab Meeting: DPF 92* (Batavia, 1992). FERMILAB-CONF-93-054-E.

[146] Run II D0 muon identification. URL http://www-d0.fnal.gov/phys_id/muon_id/d0_private/muonid_dh.html.

[147] Luo, C. *Muon Identification and B Physics Studies at the Tevatron Collider Experiment D0.* Ph.D. thesis, Indiana University (2003).

[148] Peters, O. Muon segment reconstruction: Linked list algorithm (2001). D0 Note 3901.

[149] Déliot, F. The fit algorithm in muo_trackreco (2000). URL http://www-d0.fnal.gov/~deliot/fitalg.ps.

[150] Calfayan, P. *et al.* Muon identification certification for p17 data. D0 Note 5157.

[151] Trincaz-Duvoid, S. & Verdier, P. Missing ET reconstruction in p17 (2004). D0 Note 4474.

[152] Verdier, P. Status of missing Et in Run IIb data. Presented at the Joint Algorithms/Physics Conveners Meeting, September 1, 2006.

[153] Hanagaki, K. & Kasper, J. Identification of b-jet by soft muon (2005). D0 Note 4867.

[154] Schwartzman, A. & Narain, M. b-quark jet identification via secondary vertex reconstruction (2003). D0 Note 4080.

[155] b-id secondary vertex tagger p14-pass2 certification results (2005). URL `http://www-d0.fnal.gov/phys_id/bid/d0_private/certification/p14Pass2/SVT/secvertex_v4.html`.

[156] Demina, R. *et al.* Measurement of b-tagging efficiency and mis-tagging rates with CSIP method (2004). D0 Note 4432.

[157] Bloch, D. & Clément, B. Update of the JLIP b-tagger performance in p14/pass2 with JES 5.3 (2005). D0 Note 4824.

[158] Scanlon, T. A neural network b-tagging tool (2005). D0 Note 4889.

[159] Anastasoaie, M. *et al.* Performance of the NN b-tagging tool on p17 data (2007). D0 Note 5213.

[160] Aguiló, E. *et al.* Using boosted decision trees to search for single top quarks in $1\,\mathrm{fb}^{-1}$ of data (2006). D0 Note 5286.

[161] Aguiló, E. *et al.* Using Bayesian neural networks to search for single top quarks in $1\,\mathrm{fb}^{-1}$ of data (2006). D0 Note 5286.

[162] Aguiló, E. *et al.* The search for single top quark production using the matrix element analysis technique in $1\,\mathrm{fb}^{-1}$ of data (2006). D0 Note 5287.

[163] Aguiló, E. *et al.* A second look at Bayesian neural networks in the search for single top quarks in $1\,\mathrm{fb}^{-1}$ of data (2007). D0 Note 5361.

[164] Mangano, M. L. *et al.* ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP* **07**, 001 (2003).

[165] Mrenna, S. & Richardson, P. Matching matrix elements and parton showers with HERWIG and PYTHIA. *JHEP* **05**, 040 (2004).

[166] Kidonakis, N. & Vogt, R. Next-to-next-to-leading order soft-gluon corrections in top quark hadroproduction. *Phys. Rev.* **D68**, 114014 (2003).

[167] Boos, E. E. *et al.* Method for simulating electroweak top-quark production events in the NLO approximation: SingleTop event generator. *Phys. Atom. Nucl.* **69**, 1317–1329 (2006).

[168] Sjöstrand, T. *et al.* PYTHIA 6.4 physics and manual. *JHEP* **05**, 026 (2006).

[169] Jadach, S. *et al.* The tau decay library TAUOLA: Version 2.4. *Comput. Phys. Commun.* **76**, 361–380 (1993).

[170] Lange, D. J. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.* **A462**, 152–155 (2001).

[171] Boos, E. *et al.* Complete calculations of $Wb\bar{b}$ and $Wb\bar{b}$ + jet production at Tevatron and LHC: Probing anomalous Wtb couplings in single top production. *Eur. Phys. J.* **C11**, 473–484 (1999).

[172] Höche, S. *et al.* Matching parton showers and matrix elements (2006). `hep-ph/0602031`.

[173] Ellis, R. K. & Veseli, S. Strong radiative corrections to $Wb\bar{b}$ production in $p\bar{p}$ collisions. *Phys. Rev.* **D60**, 011501 (1999).

[174] Febres Cordero, F. *et al.* Nlo qcd corrections to w boson production with a massive b- quark jet pair at the tevatron p anti-p collider. *Phys. Rev.* **D74**, 034007 (2006).

[175] Brun, R. & Carminati, F. GEANT detector description and simulation tool (1993). CERN Program Library Long Writeup W5013.

[176] D0sim. URL `http://www-d0.fnal.gov/computing/MonteCarlo/simulation/d0sim.html`.

[177] Private communication with V. Buescher.

[178] Makovec, N. & Grivaz, J.-F. Shifting, smearing and removing simulated jets (2005). D0 Note 4914.

[179] Christofek, L. *et al.* Probability calculation for multi-object trigger configurations (2005). D0 Note 4882.

[180] Agelou, M. *et al.* Top trigger efficiency measurements and the `top_trigger` package (2004). D0 Note 4512.

[181] Using `caf_trigger`. URL `http://www.phys.ualberta.ca/~kwchan/d0/caf_trigger/readme.html`.

[182] Barberis, E. *et al.* The matrix method and its error calculation (2004). D0 Note 4564.

[183] Abazov, V. M. *et al.* A precision measurement of the mass of the top quark. *Nature* **429**, 638–642 (2004).

[184] Abazov, V. M. *et al.* Measurement of the top quark mass in the lepton+jets final state with the matrix element method. *Phys. Rev.* **D74**, 092005 (2006).

[185] The D0 Collaboration. Measurement of the top quark mass with a matrix element method using the lepton+jets $1\,\mathrm{fb}^{-1}$ data set (2007). D0 Note 5362–CONF.

[186] The CDF Collaboration. Measurement of the top quark mass using the matrix element analysis technique in the lepton + jets channel with in-situ $W \to jj$ calibration (2006). CDF Note 8375.

[187] Abulencia, A. *et al.* Precision measurement of the top quark mass from dilepton events at CDF II. *Phys. Rev.* **D75**, 031105 (2007).

[188] The CDF Collaboration. Top mass measurement in the lepton + jets channel using a multivariate method and *in situ* jet calibration (2007). CDF Note 8780.

[189] Abazov, V. M. *et al.* Helicity of the W boson in lepton+jets $t\bar{t}$ events. *Phys. Lett.* **B617**, 1–10 (2005).

[190] The CDF Collaboration. Search for single top quark production in $955\,\mathrm{pb}^{-1}$ using the matrix element technique. CDF Note 8588.

[191] Jain, S. *et al.* Computing limits using a Bayesian approach in package `top_statistics` (2006). D0 Note 5123.

[192] The Les Houches Accord PDF interface. URL `http://hepforge.cedar.ac.uk/lhapdf/`.

[193] Maltoni, F. & Stelzer, T. MadEvent: Automatic event generation with MadGraph. *JHEP* **02**, 027 (2003).

[194] Murayama, H. *et al.* HELAS: HELicity Amplitude Subroutines for Feynman diagram evaluations (1992). KEK-91-11.

[195] Lepage, G. P. Vegas: An adaptive multidimensional integration program (1980). CLNS-80/447.

[196] The GNU scientific library. URL `http://www.gnu.org/software/gsl/`.

[197] Private communication with A. Harel.

[198] Shiefferdecker, P. & Wang, M. Jet transfer functions derived from p17 $t\bar{t}$ Monte Carlo (2006). D0 Note 5136.

[199] Private communication with L. Wang. Parametrization found in the file `/work/cole-clued0/leiwang/wz_epmcs/p170303_sampling/wz_epmcs/src/pmcsana.cpp`.

[200] PMCS home page. URL `http://www-d0.fnal.gov/computing/MonteCarlo/pmcs/pmcs_doc/pmcs.html`.

[201] Haefner, P. & Fiedler, F. Determination of the muon transfer function for top mass measurements (2005). D0 Note 4818.

[202] Haefner, P. Derivation of muon transfer function parameters for p17 (2006). Presented at the top production/properties meeting, July 13, 2006.

[203] The D0 Collaboration. A measurement of the $W \to \mu\nu$ charge asymmetry with the D0 detector at $\sqrt{s} = 1.96\,\text{TeV}$ (2006). D0 Note 5061–CONF.

[204] Schellman, H. The longitudinal shape of the luminous region at D0 (2006). D0 Note 5142.

[205] The D0 Collaboration. Updated combination results from three single top quark cross section measurements using the BLUE method (2007). D0 Note 5396–CONF.

[206] Aguiló, E. *et al.* First direct measurement of $|V_{tb}|$ (2007). D0 Note 5440.

[207] Beneke, M. *et al.* Top quark physics (2000). `hep-ph/0003033`.

Appendix A

Systematic Uncertainties

	Single-Tagged Three-Jets Electron Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	16.8	16.8	16.8	16.8
Primary vertex	2.4	2.4	2.4	2.4	—	—	—	—
Electron ID	5.5	5.5	5.5	5.5	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	3.0	3.0	3.0	3.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	5.3	5.8	4.1	3.2	—	—	—	—
Flavor-dependent TRFs	2.1	4.5	2.9	2.1	4.4	6.2	7.6	—
Statistics	1.0	1.0	0.5	0.5	1.0	1.0	0.5	6.7
Combined								
Acceptance uncertainty	12.0	12.9	—	—	—	—	—	—
Yield uncertainty	20.0	19.8	22.0	21.2	17.4	18.0	18.5	18.1

Table A.1: Electron channel uncertainties for one b-tag and three jets.

	Single-Tagged Two-Jets Muon Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	20.7	20.7	20.7	20.7
Primary vertex	3.0	3.0	3.0	3.0	—	—	—	—
Muon ID	7.4	7.4	7.4	7.4	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	6.0	6.0	6.0	6.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	5.3	6.1	20.1	6.8	—	—	—	—
Flavor-dependent TRFs	1.8	5.9	4.5	2.0	4.4	6.3	7.5	—
Statistics	9.0	0.7	1.0	1.0	0.8	0.8	0.4	14.0
Combined								
Acceptance uncertainty	16.7	15.4	—	—	—	—	—	—
Yield uncertainty	23.1	21.5	30.7	23.2	21.2	21.7	22.0	25.0

Table A.2: Muon channel uncertainties for one b-tag and two jets.

	Single-Tagged Three-Jets Muon Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	20.8	20.8	20.8	20.8
Primary vertex	3.0	3.0	3.0	3.0	—	—	—	—
Muon ID	7.4	7.4	7.4	7.4	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	6.0	6.0	6.0	6.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	9.3	9.0	10.8	7.6	—	—	—	—
Flavor-dependent TRFs	1.8	4.4	2.6	1.9	4.3	6.2	7.6	—
Statistics	2.0	2.0	0.8	0.7	1.0	1.0	0.7	14.3
Combined								
Acceptance uncertainty	16.1	16.5	—	—	—	—	—	—
Yield uncertainty	22.7	22.3	25.2	23.5	21.2	21.7	22.1	25.2

Table A.3: Muon channel uncertainties for one b-tag and three jets.

	Double-Tagged Two-Jets Electron Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	26.5	26.5	26.5	26.5
Primary vertex	2.4	2.4	2.4	2.4	—	—	—	—
Electron ID	5.5	5.5	5.5	5.5	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	3.0	3.0	3.0	3.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	0.8	4.1	8.0	1.8	—	—	—	—
Flavor-dependent TRFs	12.9	12.9	13.5	13.0	12.2	13.6	16.1	—
Statistics	0.7	0.7	1.3	0.8	0.9	0.9	0.4	28.9
Combined								
Acceptance uncertainty	16.7	17.2	—	—	—	—	—	—
Yield uncertainty	23.1	22.8	26.6	24.6	29.1	29.8	31.0	39.2

Table A.4: Electron channel uncertainties for two b-tags and two jets.

	Double-Tagged Three-Jets Electron Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	22.1	22.1	22.1	22.1
Primary vertex	2.4	2.4	2.4	2.4	—	—	—	—
Electron ID	5.5	5.5	5.5	5.5	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	3.0	3.0	3.0	3.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	4.8	4.0	3.5	2.9	—	—	—	—
Flavor-dependent TRFs	12.7	12.4	12.6	12.8	12.0	13.3	16.4	—
Statistics	1.0	1.0	0.7	0.7	1.0	1.0	0.5	25.8
Combined								
Acceptance uncertainty	17.2	16.8	—	—	—	—	—	—
Yield uncertainty	23.5	22.5	25.1	24.6	25.2	25.8	27.5	34.0

Table A.5: Electron channel uncertainties for two b-tags and three jets.

	Double-Tagged Two-Jets Muon Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	27.6	27.6	27.6	27.6
Primary vertex	3.0	3.0	3.0	3.0	—	—	—	—
Muon ID	7.4	7.4	7.4	7.4	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	6.0	6.0	6.0	6.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	5.2	9.1	19.7	6.9	—	—	—	—
Flavor-dependent TRFs	12.9	12.8	13.4	12.9	12.2	13.5	16.1	—
Statistics	1.3	0.9	0.7	0.7	1.0	1.0	0.5	57.7
Combined								
Acceptance uncertainty	19.0	20.3	—	—	—	—	—	—
Yield uncertainty	24.8	25.2	32.9	26.5	30.2	30.7	31.9	64.0

Table A.6: Muon channel uncertainties for two b-tags and two jets.

	Double-Tagged Three-Jets Muon Channel Percentage Errors							
	tb	tqb	$t\bar{t}lj$	$t\bar{t}ll$	Wbb	Wcc	Wjj	Mis-ID e
Components for Normalization								
Luminosity	(6.1)	(6.1)	6.1	6.1	—	—	—	—
Cross section	(16.0)	(15.0)	18.0	18.0	—	—	—	—
Branching fraction	(1.0)	(1.0)	1.0	1.0	—	—	—	—
Matrix method	—	—	—	—	25.0	25.0	25.0	25.0
Primary vertex	3.0	3.0	3.0	3.0	—	—	—	—
Muon ID	7.4	7.4	7.4	7.4	—	—	—	—
Jet ID	1.5	1.5	1.5	1.5	—	—	—	—
Jet fragmentation	5.0	5.0	7.0	5.0	—	—	—	—
Trigger	6.0	6.0	6.0	6.0	—	—	—	—
Components for Normalization and Shape								
Jet energy scale	10.2	7.6	10.1	7.8	—	—	—	—
Flavor-dependent TRFs	12.6	12.3	12.4	12.7	12.0	13.1	16.4	—
Statistics	2.0	2.0	0.8	0.6	1.0	1.0	0.6	50.0
Combined								
Acceptance uncertainty	20.8	19.4	—	—	—	—	—	—
Yield uncertainty	26.3	24.5	27.7	26.7	27.7	28.2	29.9	55.9

Table A.7: Muon channel uncertainties for two b-tags and three jets.

Appendix B

Plots After Selection

LEADING JET TRANSVERSE MOMENTUM, ELECTRON

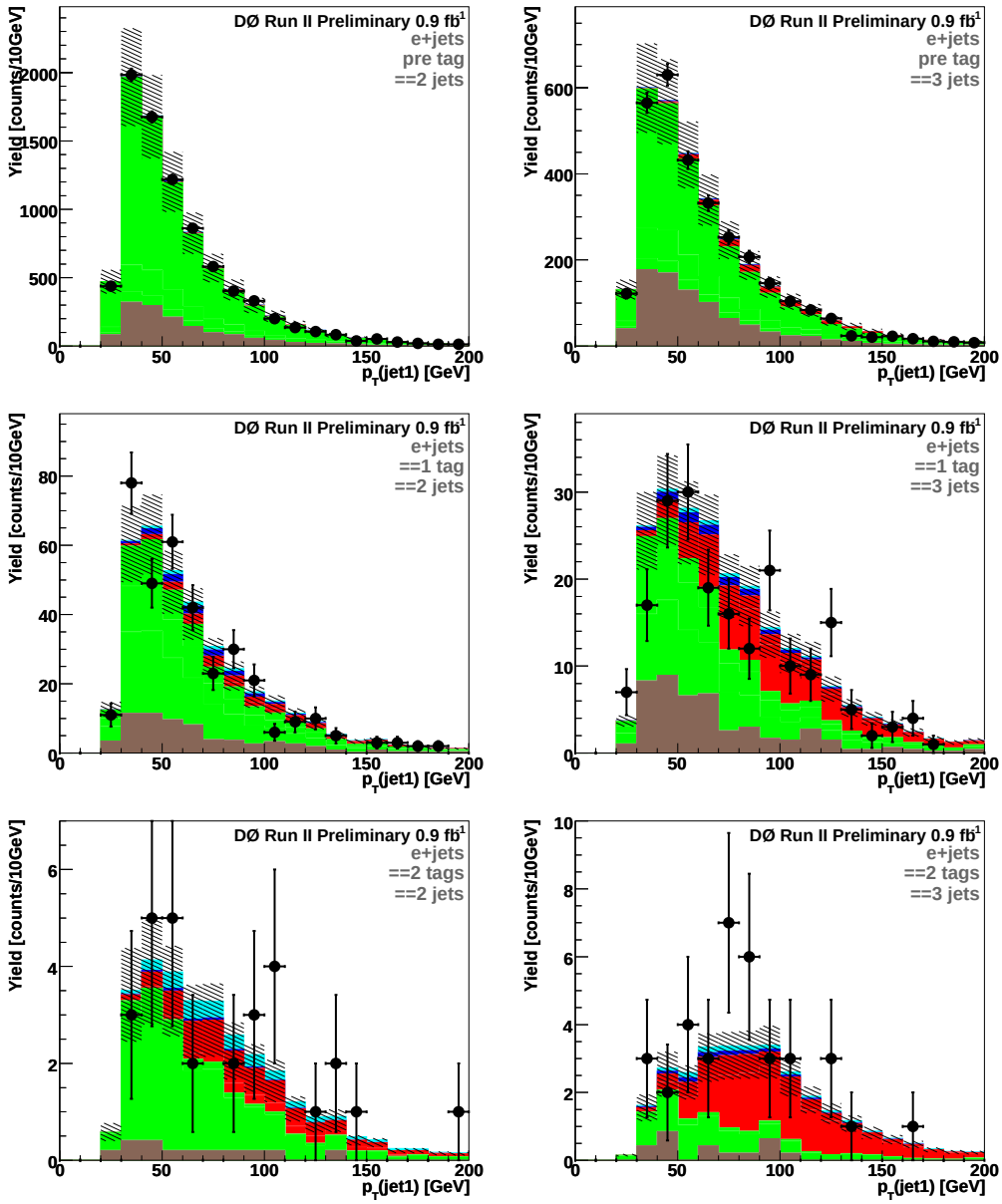


Figure B.1: The transverse momentum of the leading jet in the electron channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

LEADING JET TRANSVERSE MOMENTUM, MUON

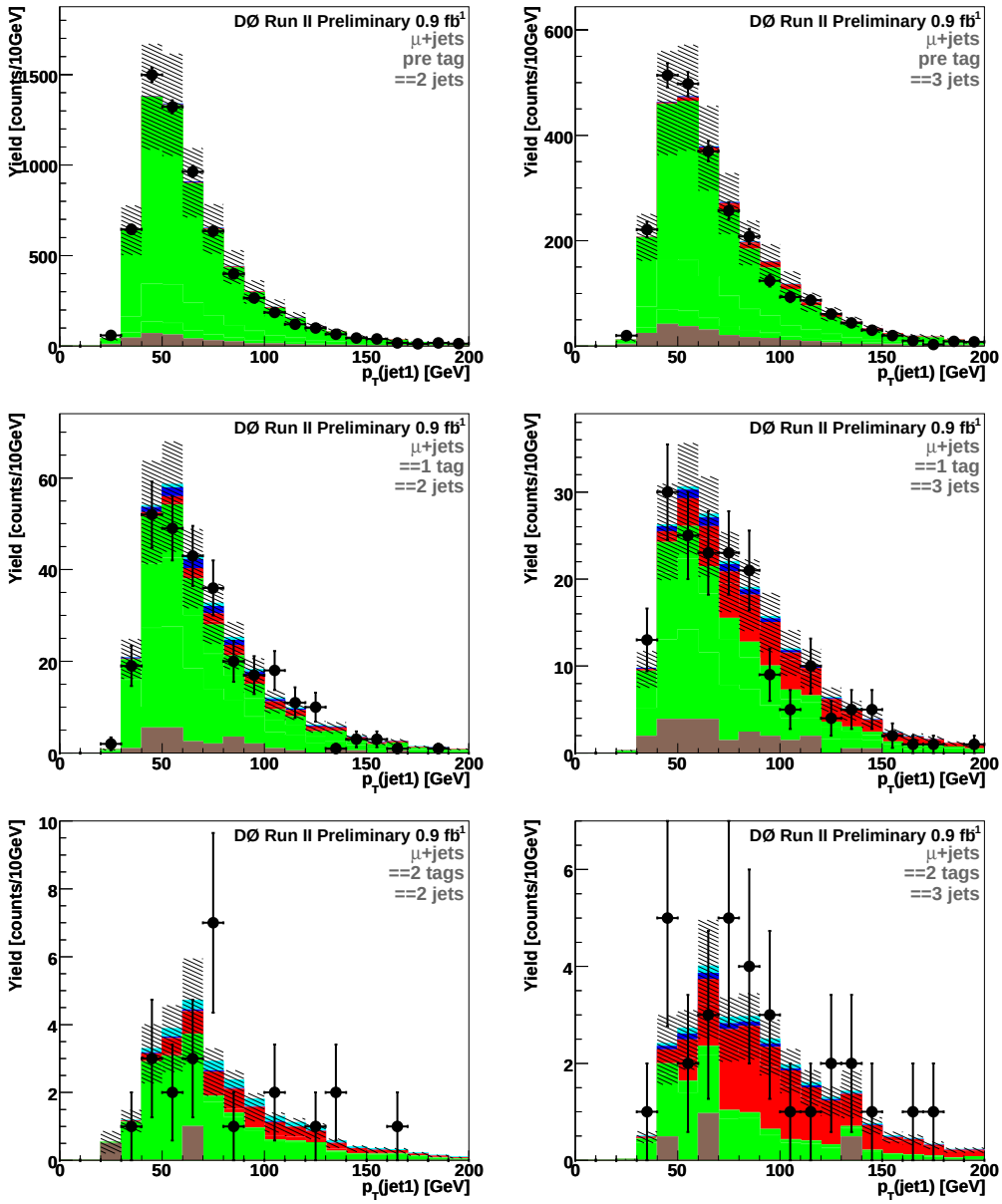


Figure B.2: The transverse momentum of the leading jet in the muon channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

SECOND JET TRANSVERSE MOMENTUM, ELECTRON

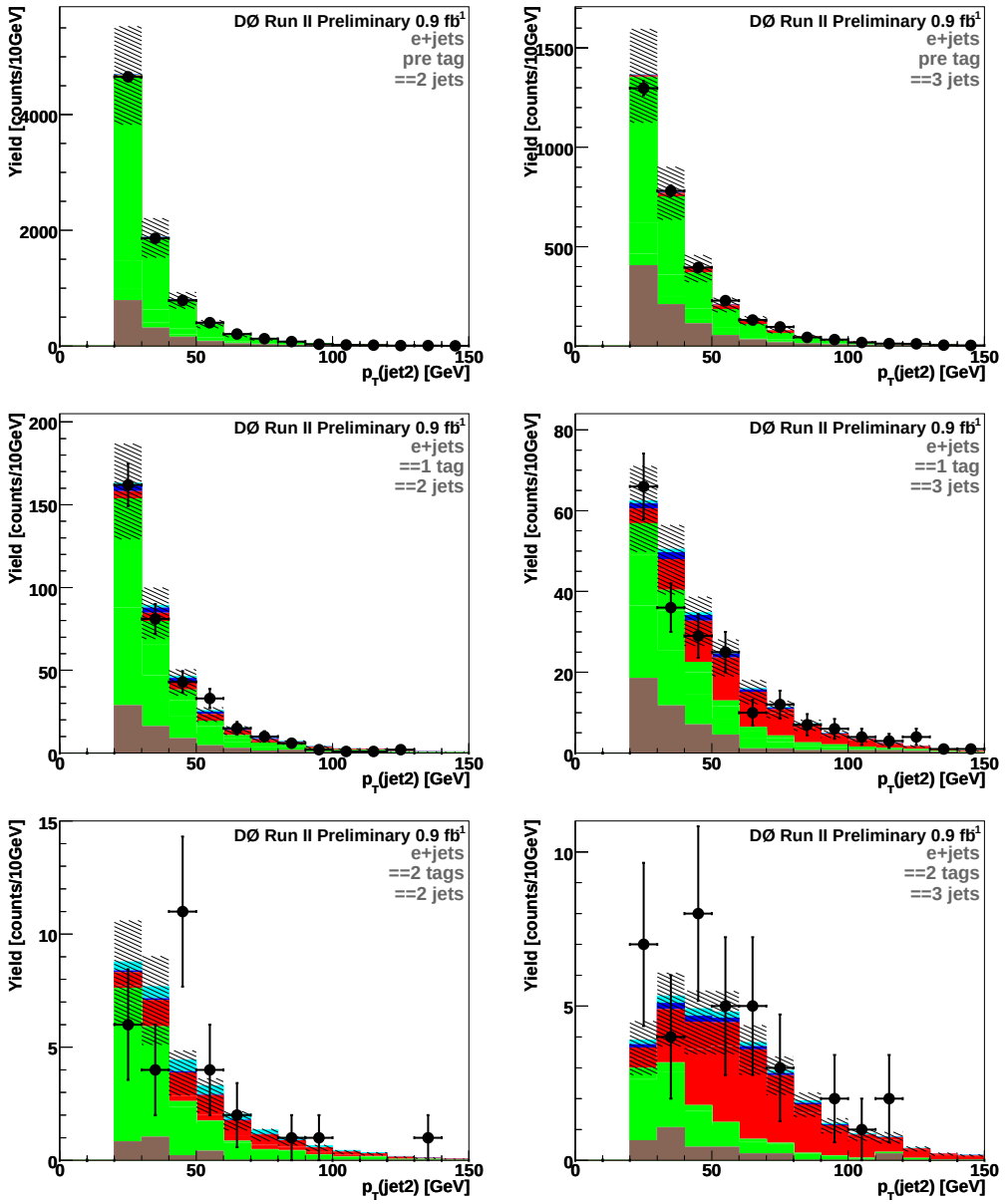


Figure B.3: The transverse momentum of the second jet in the electron channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

SECOND JET TRANSVERSE MOMENTUM, MUON

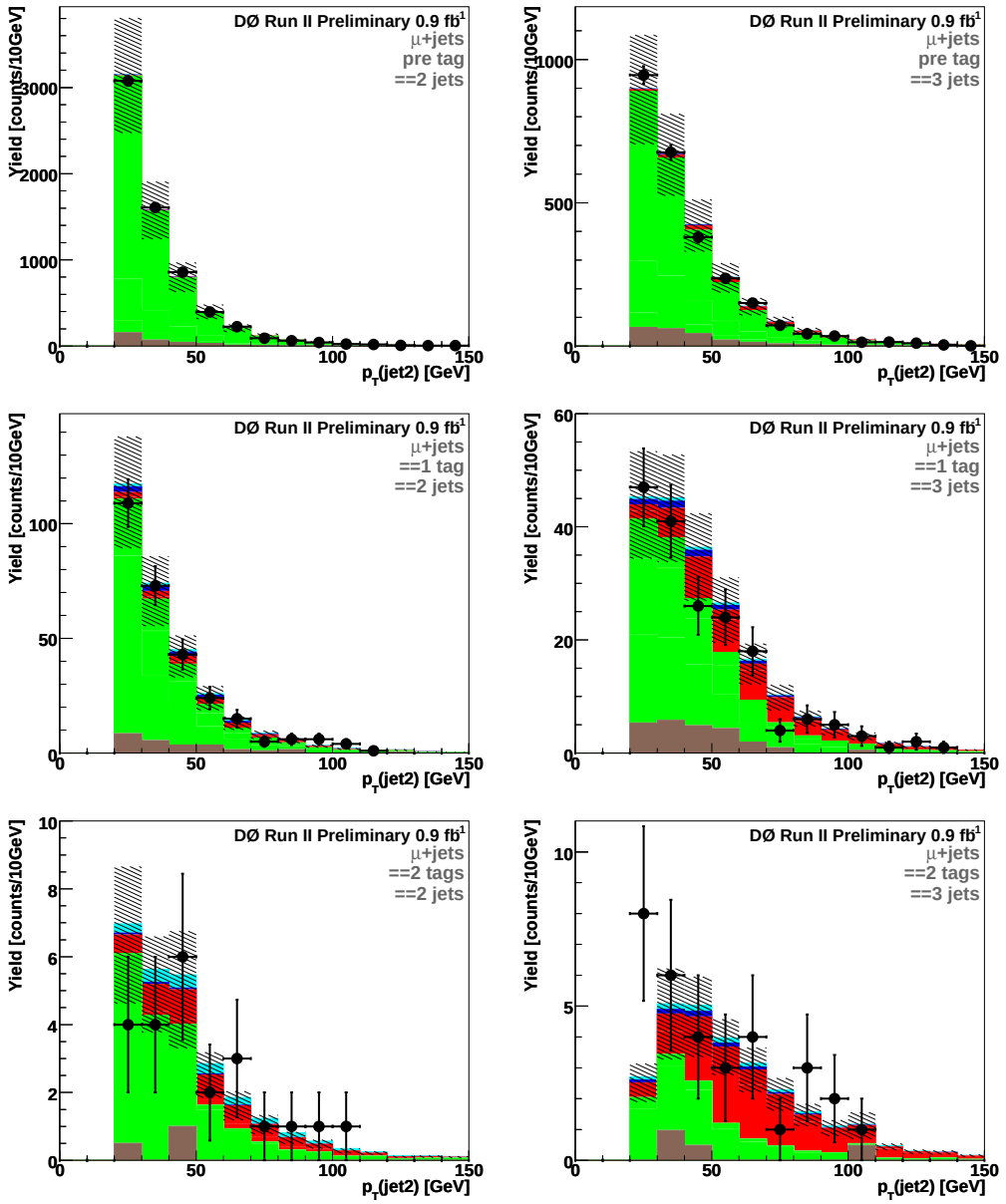


Figure B.4: The transverse momentum of the second jet in the muon channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

ELECTRON TRANSVERSE MOMENTUM

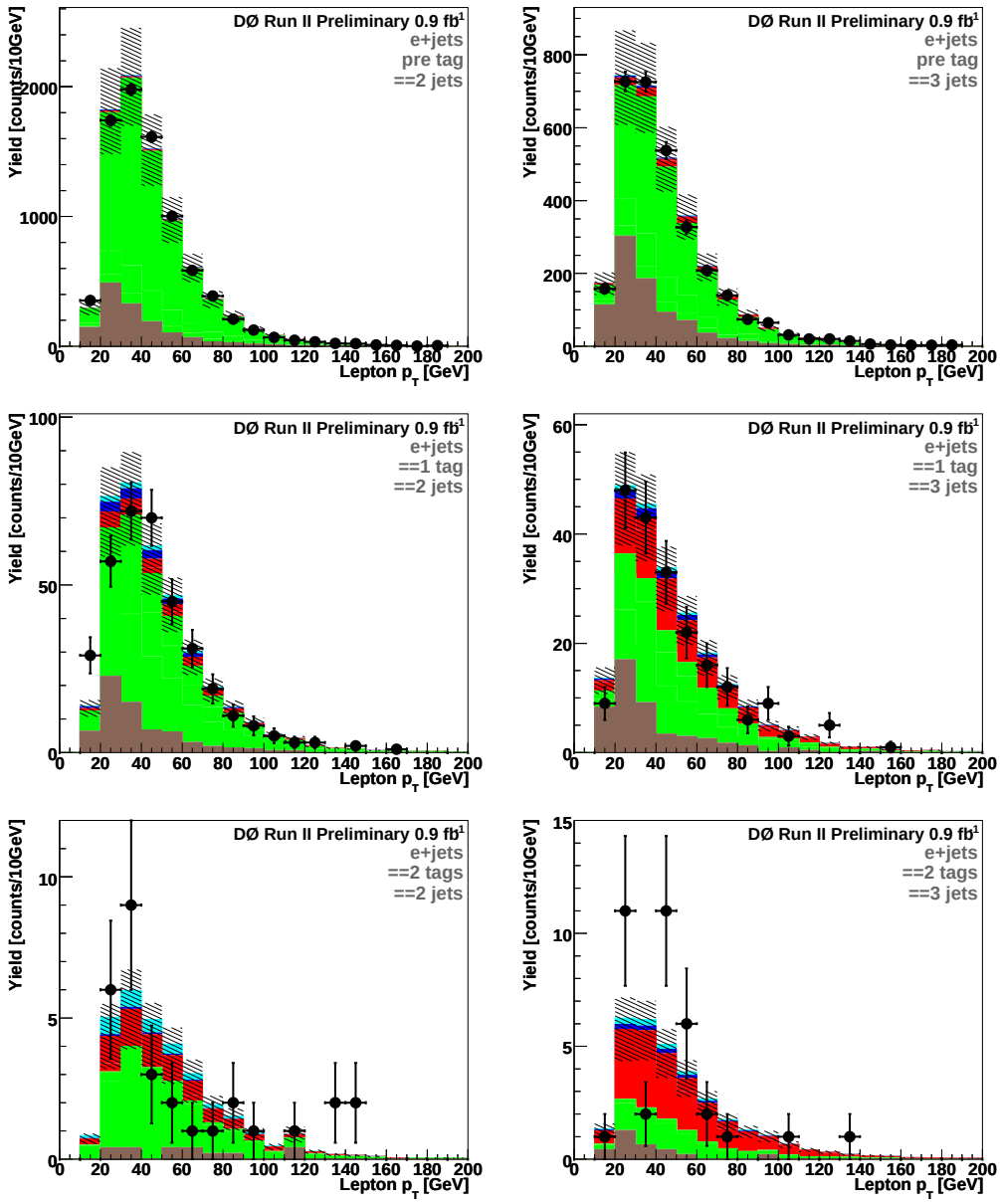


Figure B.5: The transverse momentum of the electron for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

MUON TRANSVERSE MOMENTUM

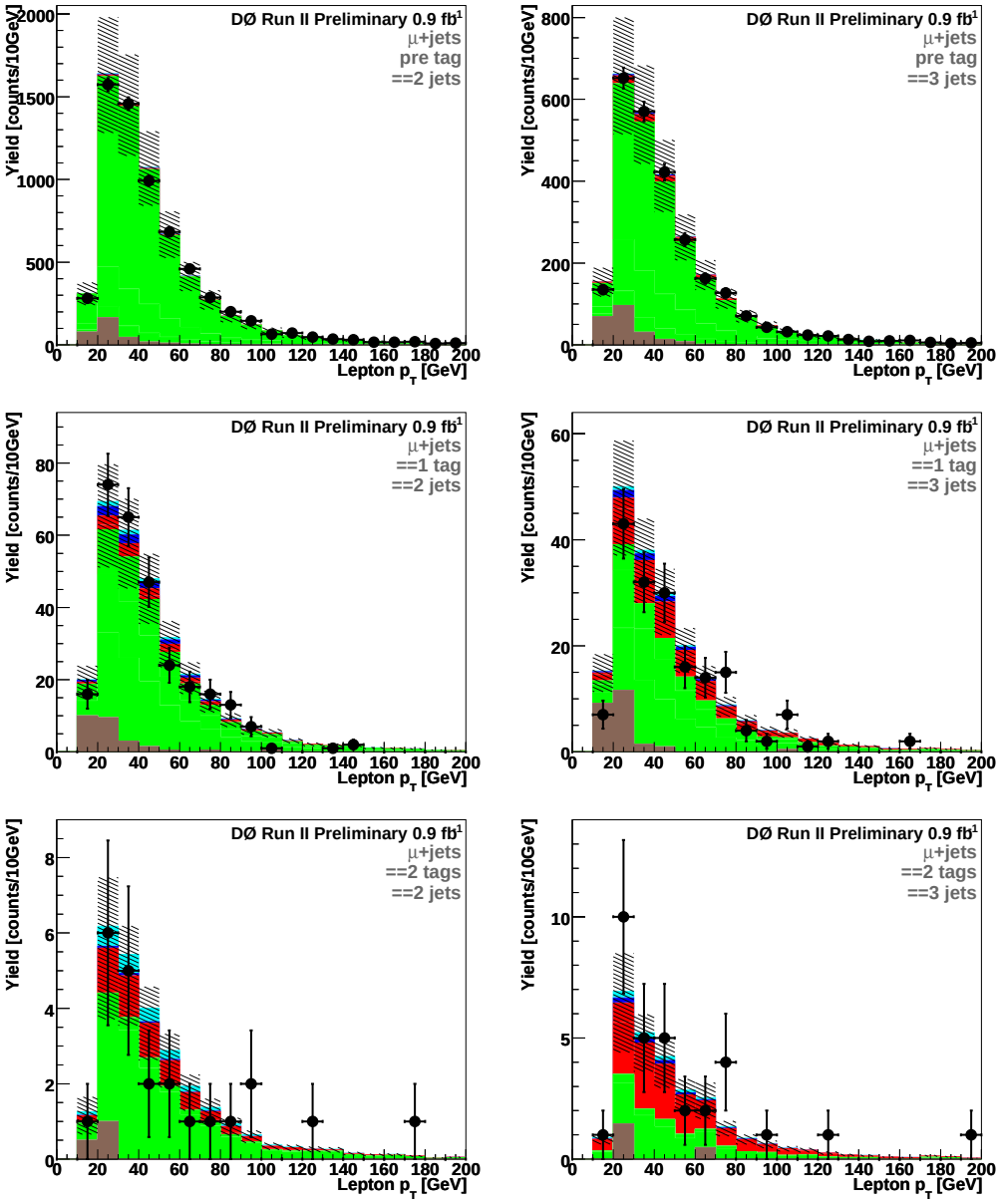


Figure B.6: The transverse momentum of the muon for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

MISSING TRANSVERSE ENERGY, ELECTRON

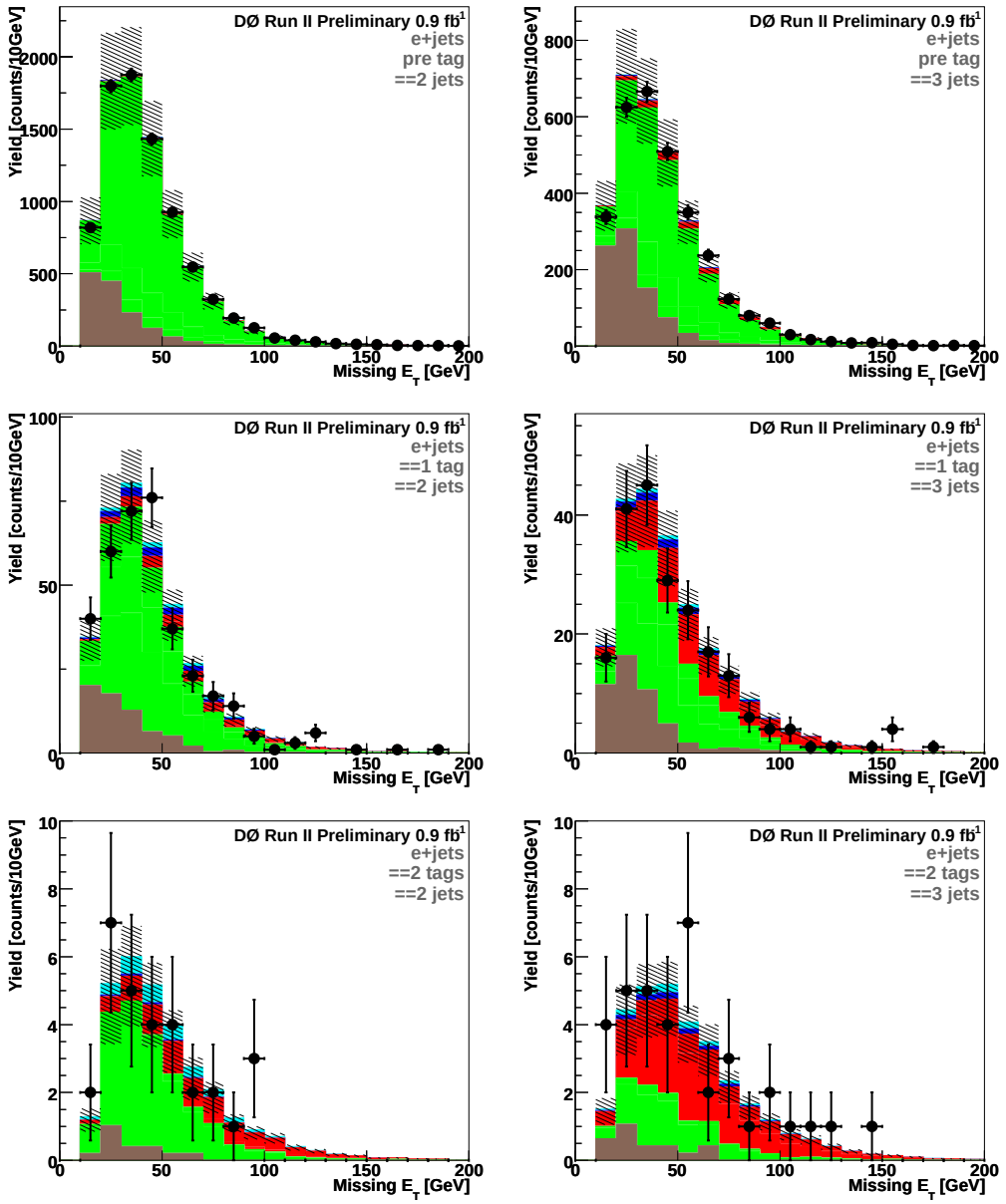


Figure B.7: The $\not\!\!E_T$ in the electron channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

MISSING TRANSVERSE ENERGY, MUON

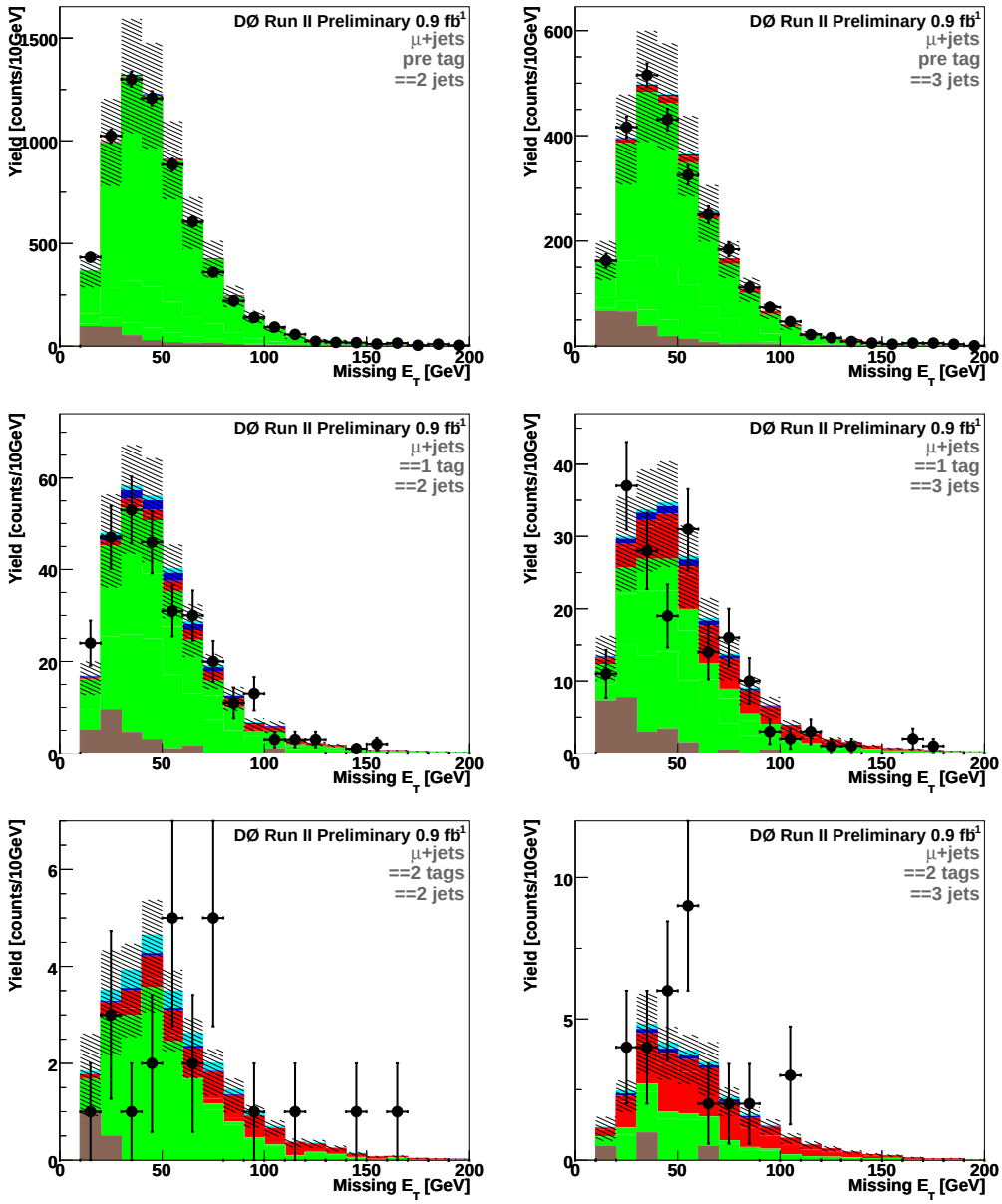


Figure B.8: The $\not\!\!E_{\mathrm{T}}$ in the muon channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

OPENING ANGLE BETWEEN JET 1 AND JET 2, ELECTRON

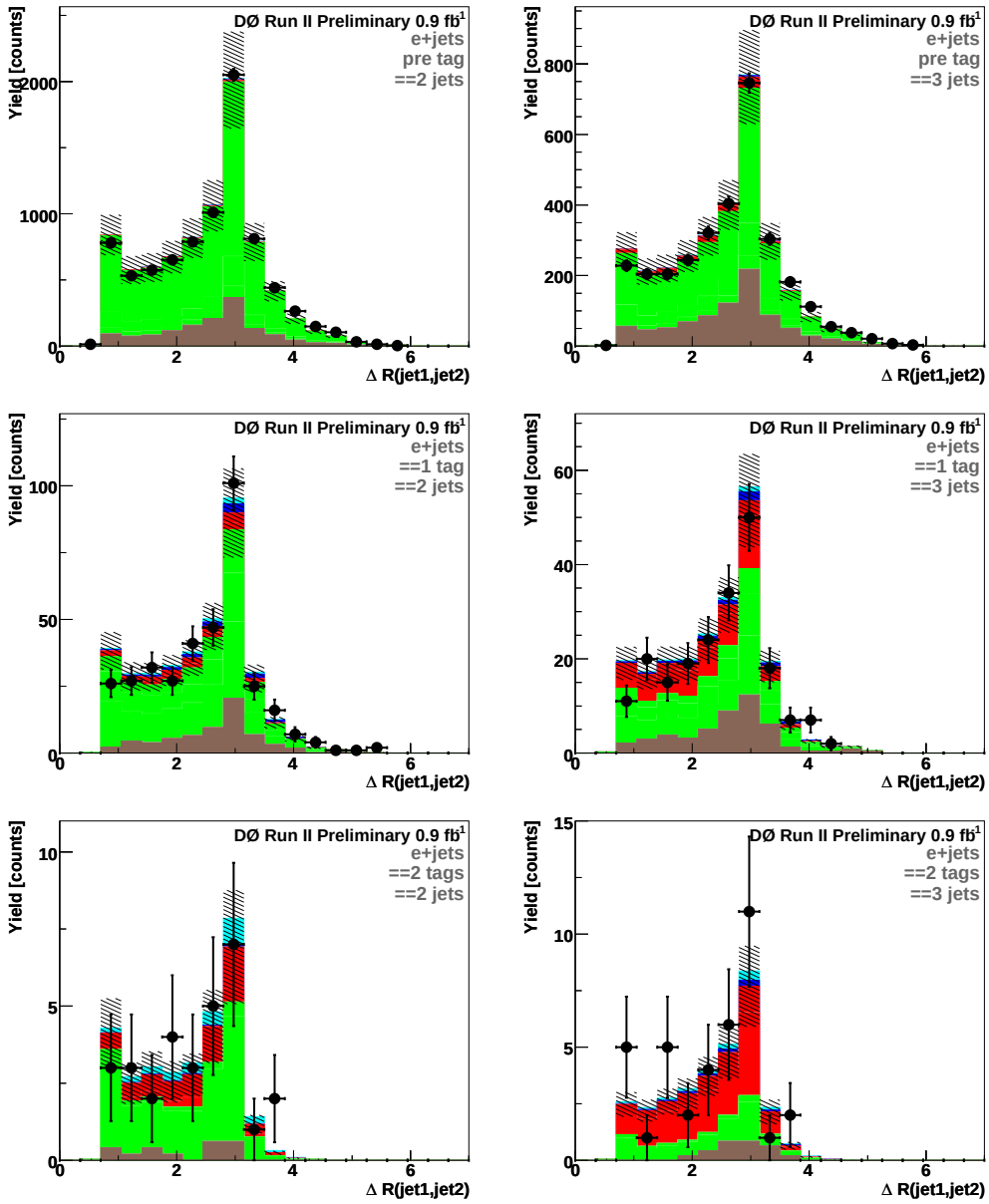


Figure B.9: The opening angle $\Delta R(\text{jet1}, \text{jet2})$ in the electron channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

OPENING ANGLE BETWEEN JET 1 AND JET 2, MUON

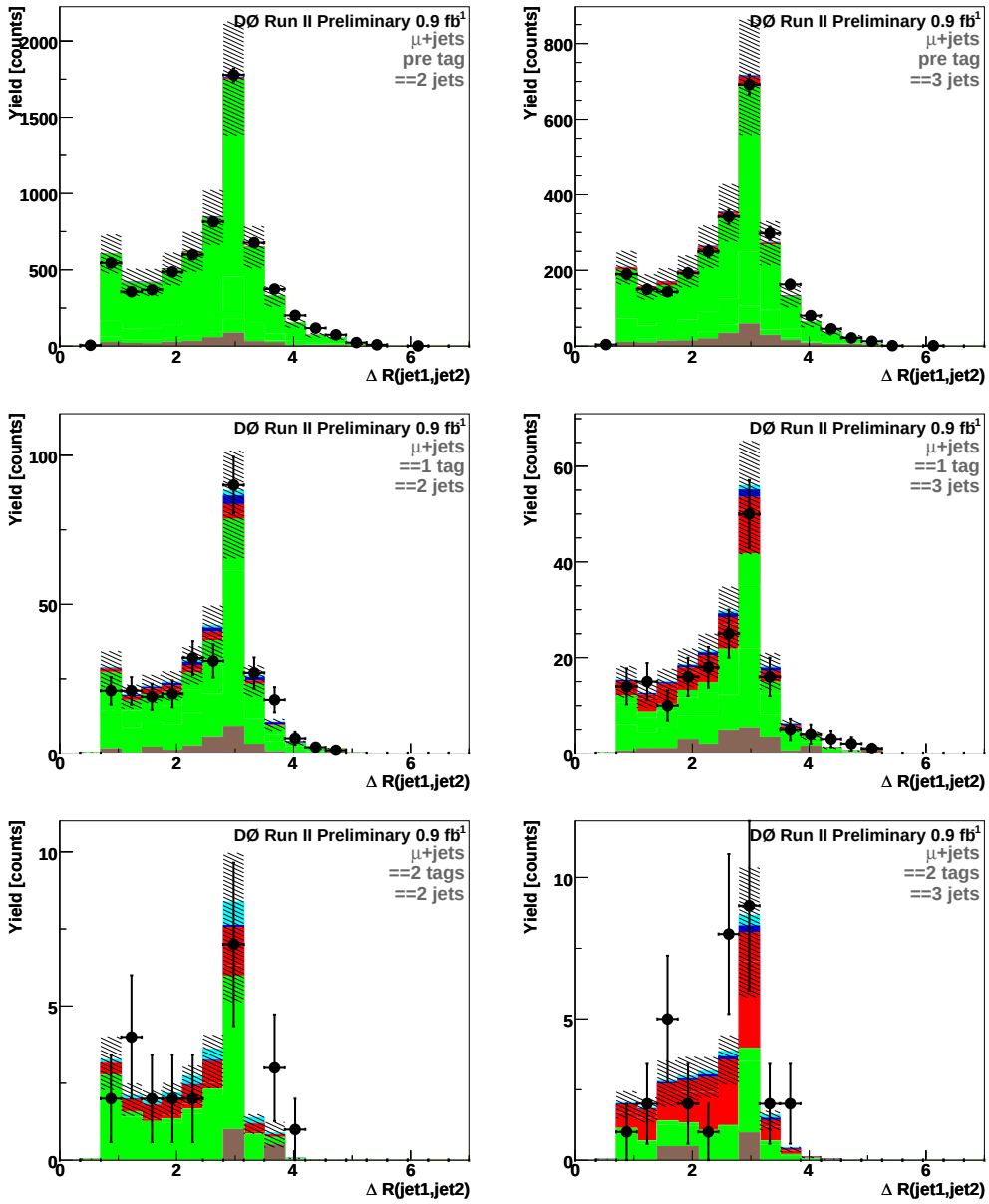


Figure B.10: The opening angle $\Delta R(\text{jet1}, \text{jet2})$ in the muon channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

W TRANSVERSE MASS, ELECTRON

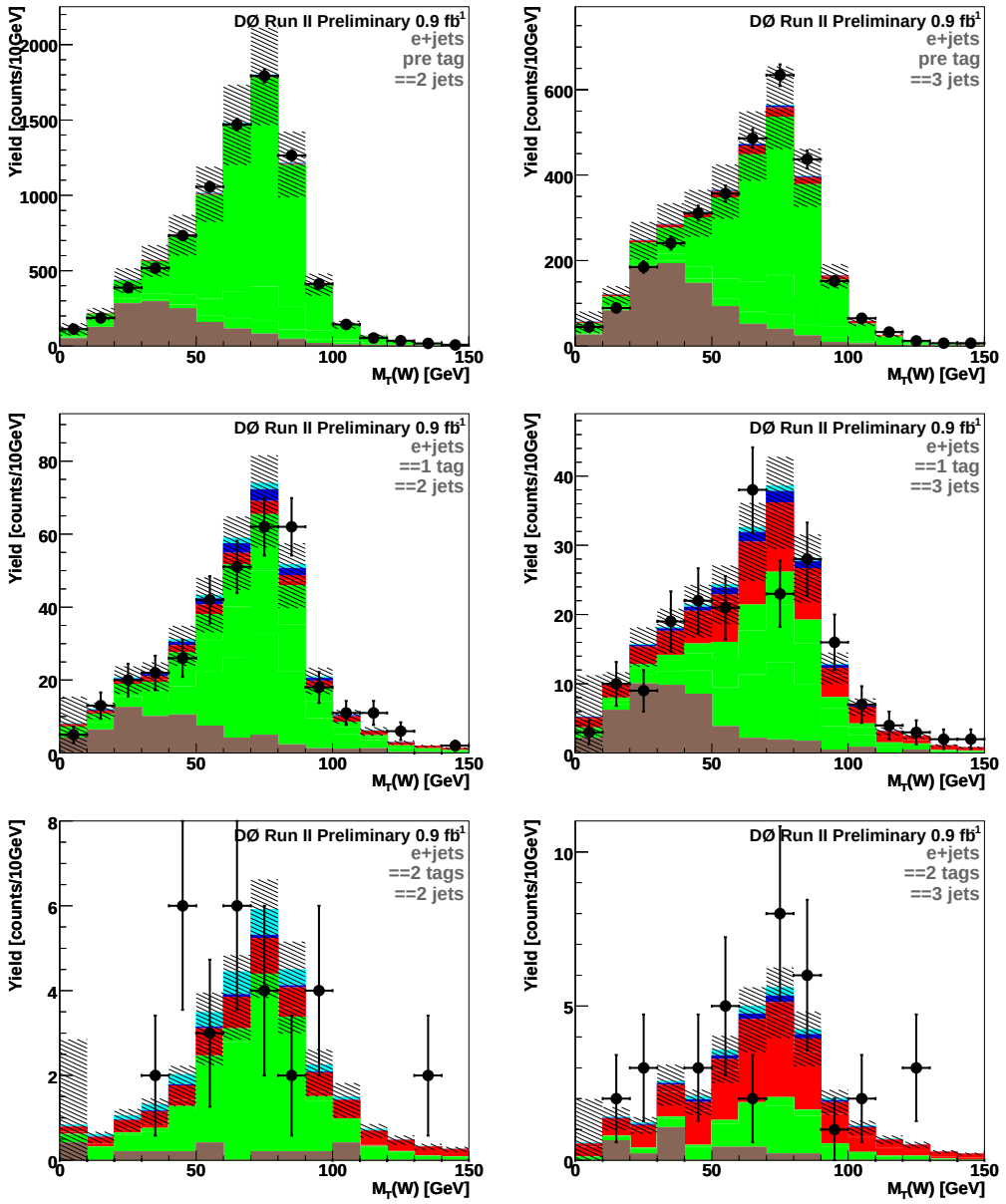


Figure B.11: The W transverse mass in the electron channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

W TRANSVERSE MASS, MUON

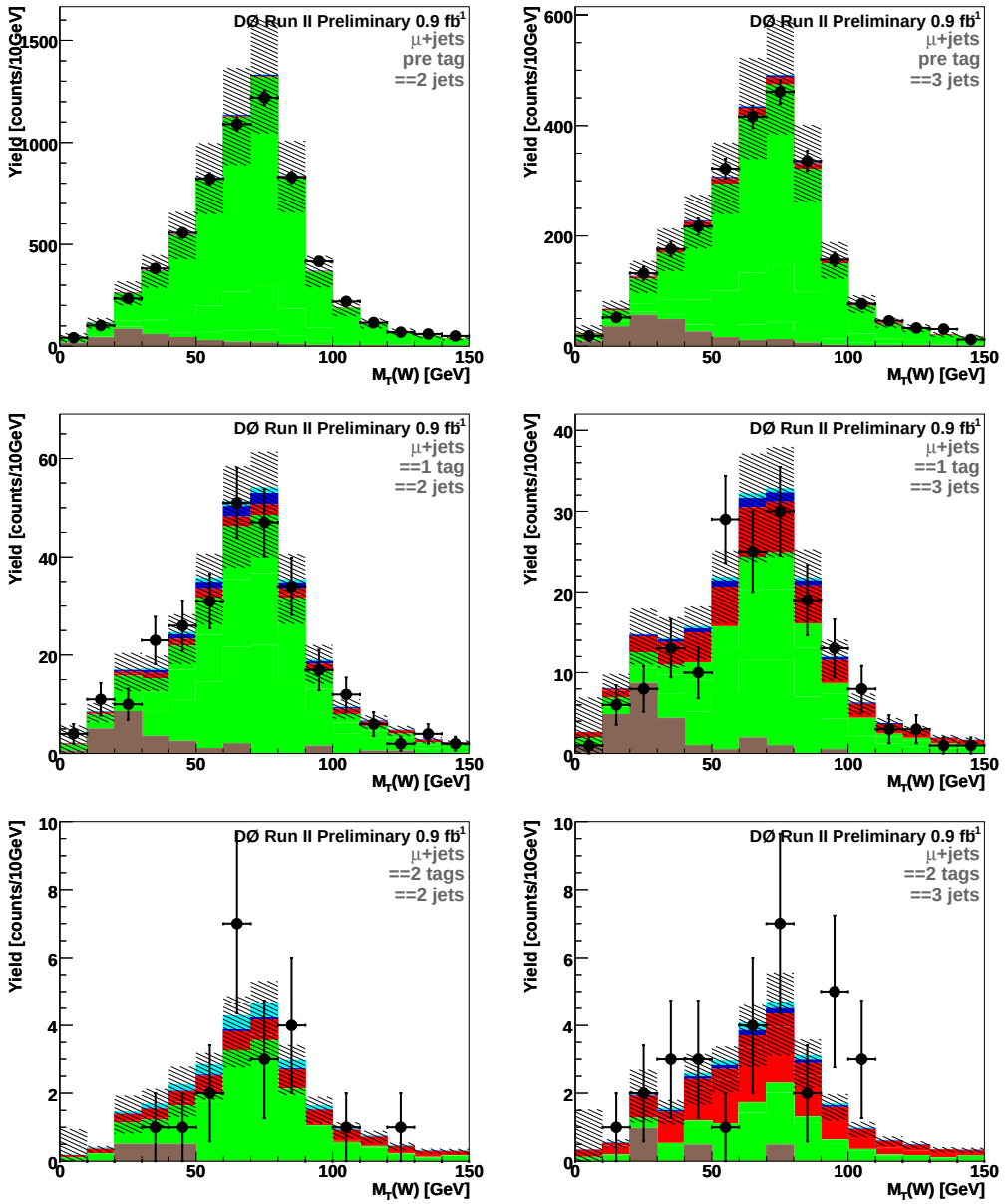


Figure B.12: The W transverse mass in the muon channel for events with two jets (left column) and three jets (left column), before tagging (first row), one b-tag (second row), and two b-tags (third row).

Appendix C

Discriminant Output Plots

MATRIX ELEMENT OUTPUTS FOR THE ELECTRON CHANNEL WITH TWO JETS

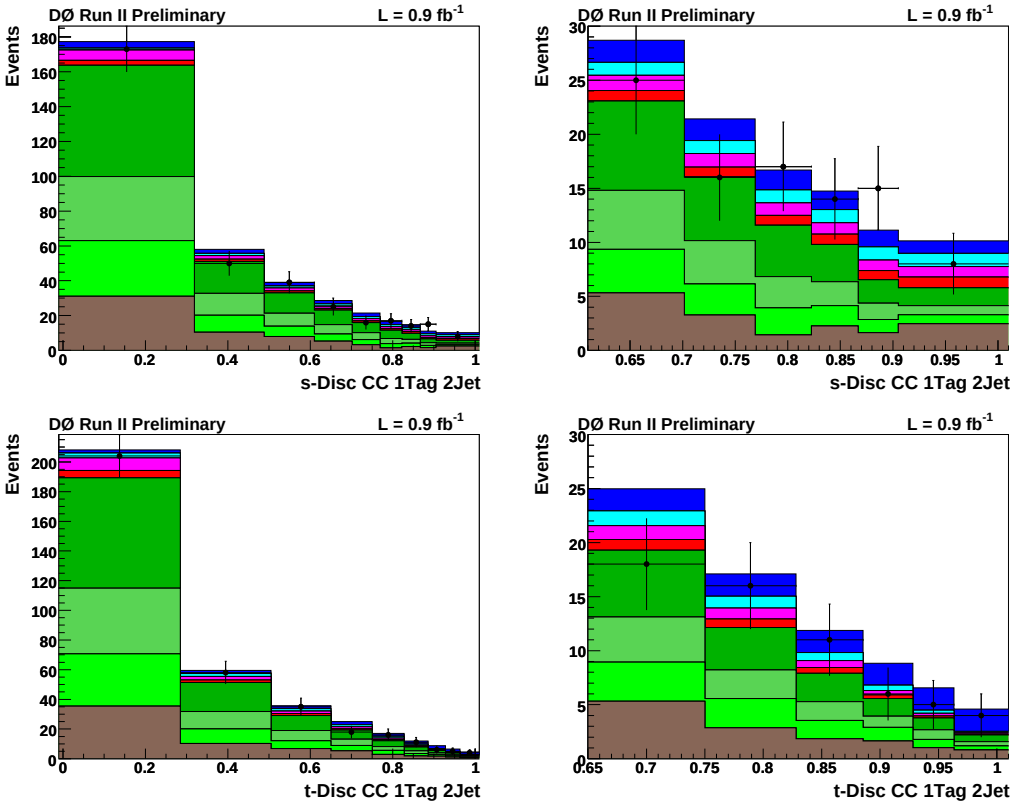


Figure C.1: 1D discriminant projections for the electron channel with one b tag. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE ELECTRON CHANNEL WITH TWO JETS

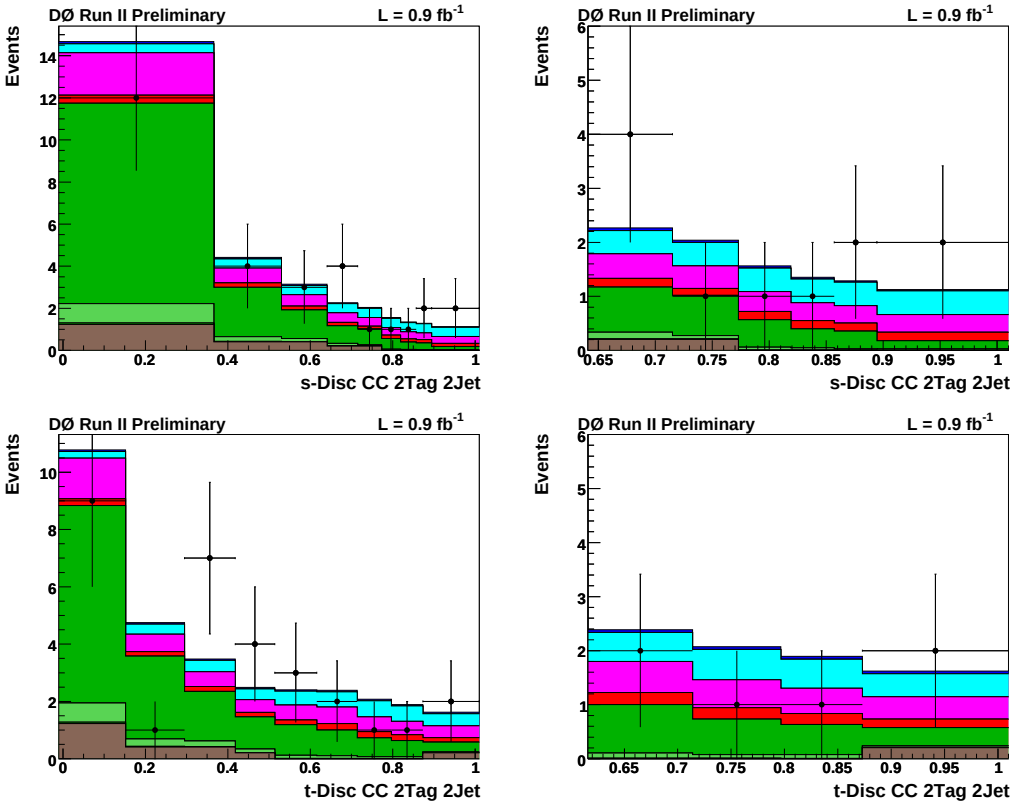


Figure C.2: 1D discriminant projections for the electron channel with two b tags. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE MUON CHANNEL WITH TWO JETS

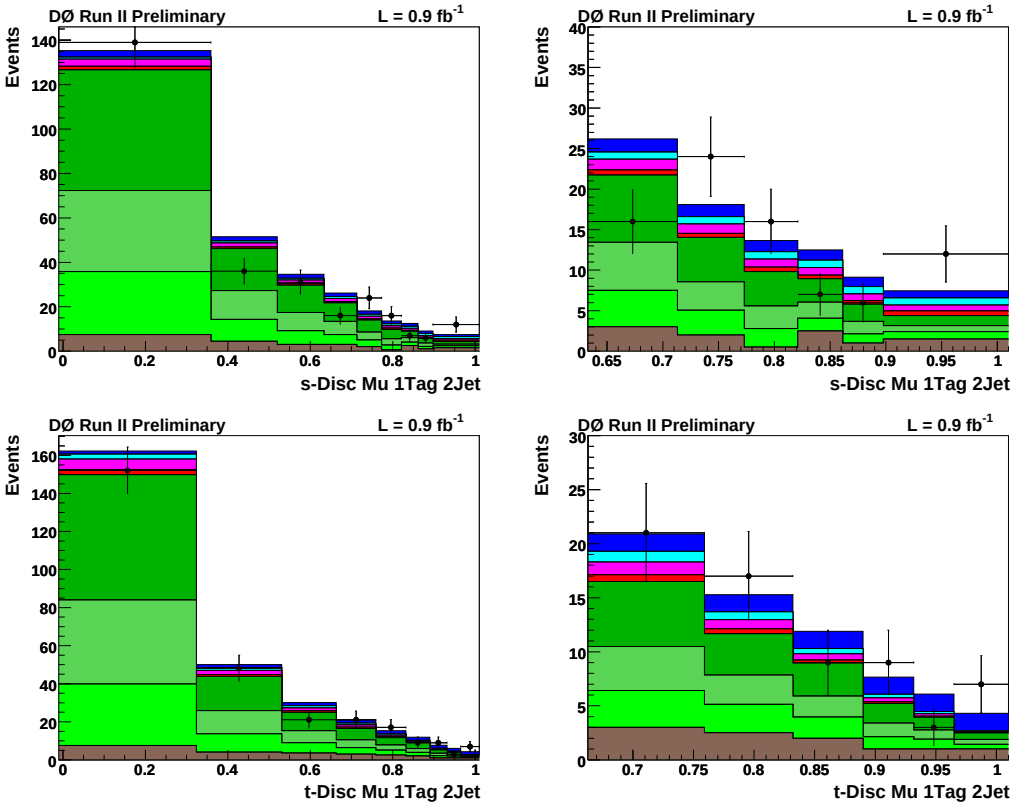


Figure C.3: 1D discriminant projections for the muon channel with one b tag. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE MUON CHANNEL WITH TWO JETS

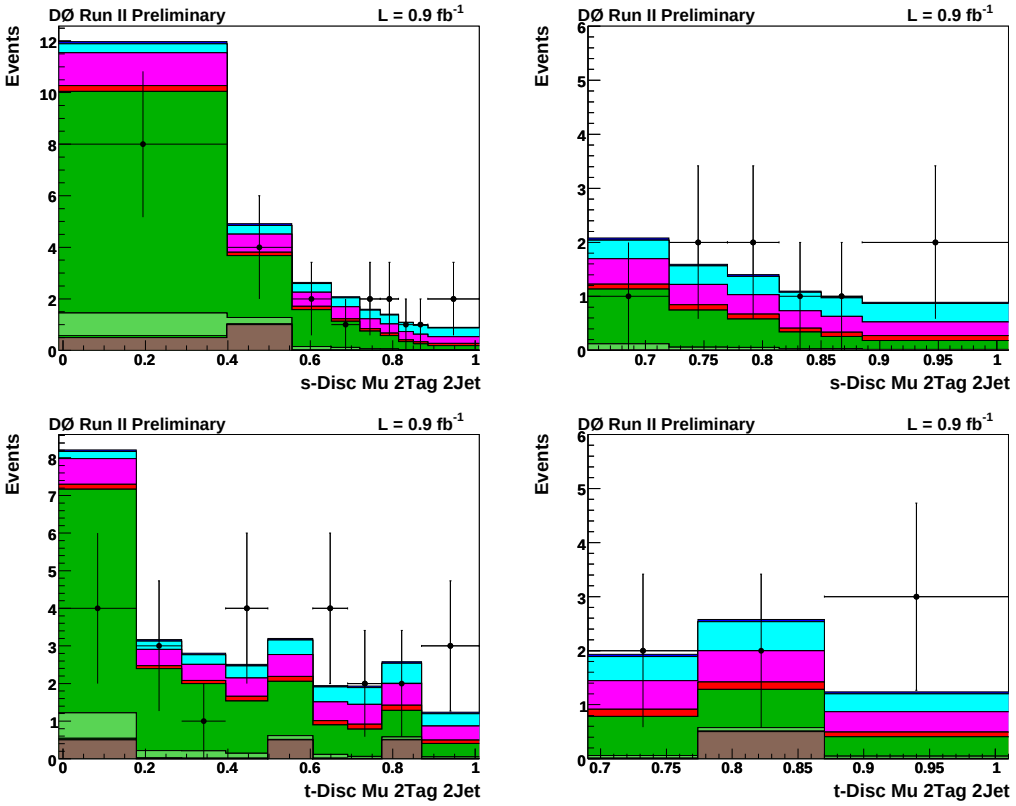


Figure C.4: 1D discriminant projections for the muon channel with two b tags. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE ELECTRON CHANNEL WITH THREE JETS

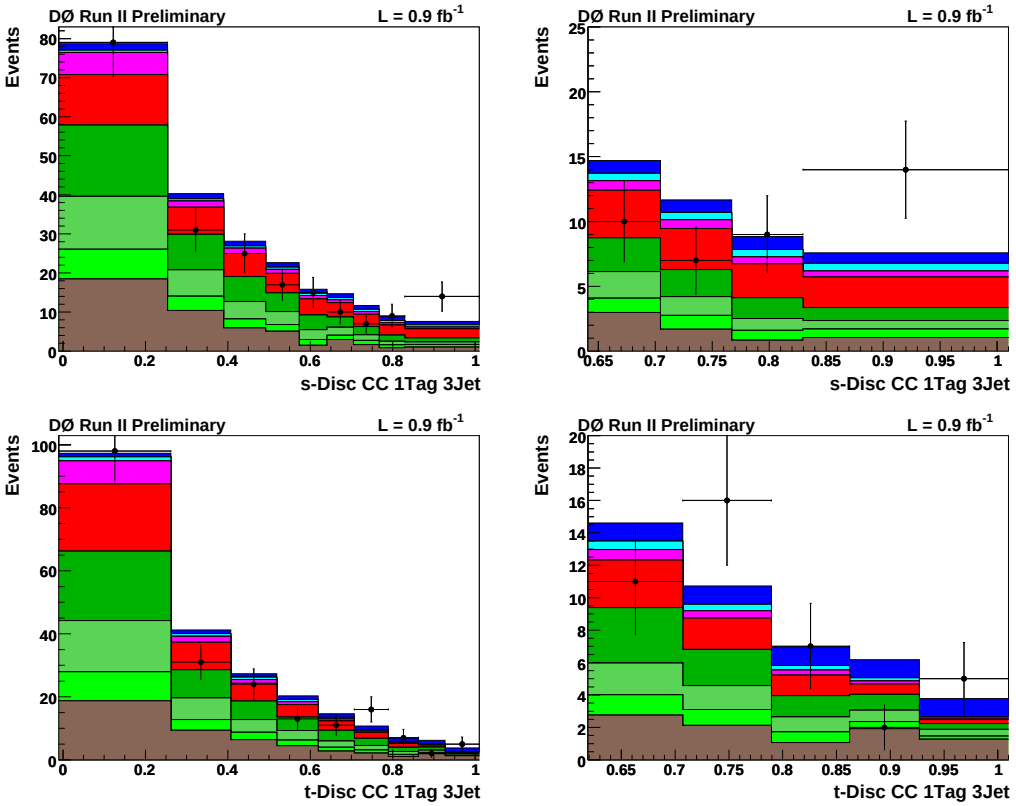


Figure C.5: 1D discriminant projections for the electron channel with one b tag. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE ELECTRON CHANNEL WITH THREE JETS

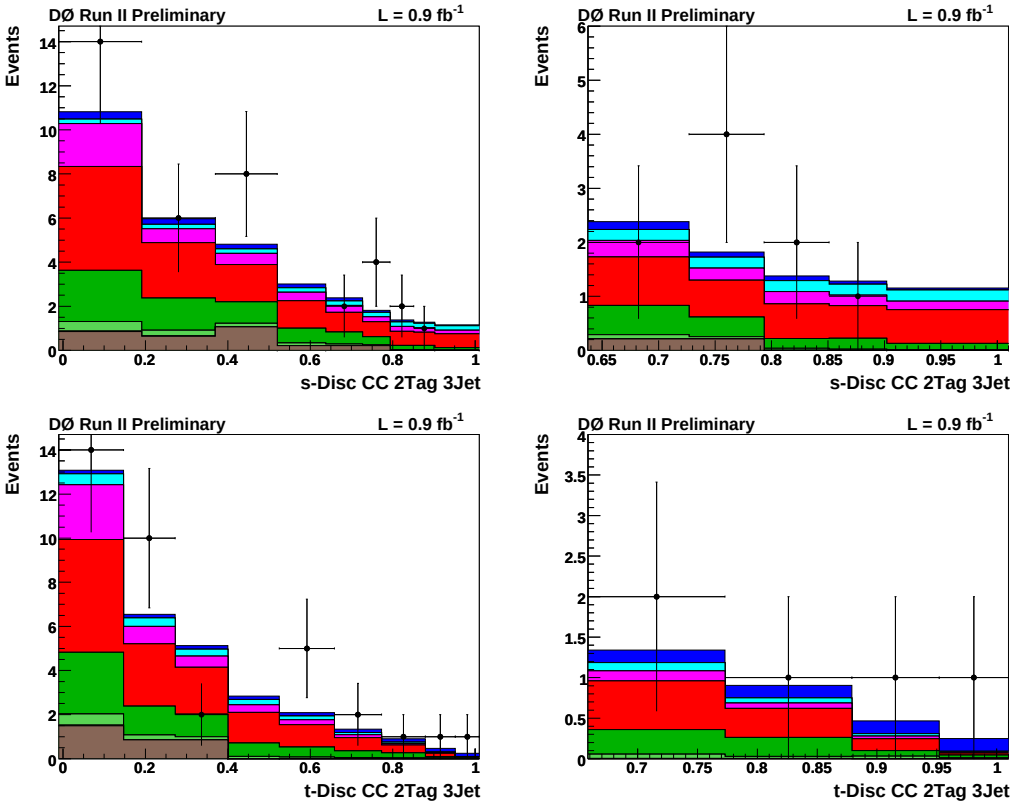


Figure C.6: 1D Discriminant projection for the electron channel with two b tags. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE MUON CHANNEL WITH THREE JETS

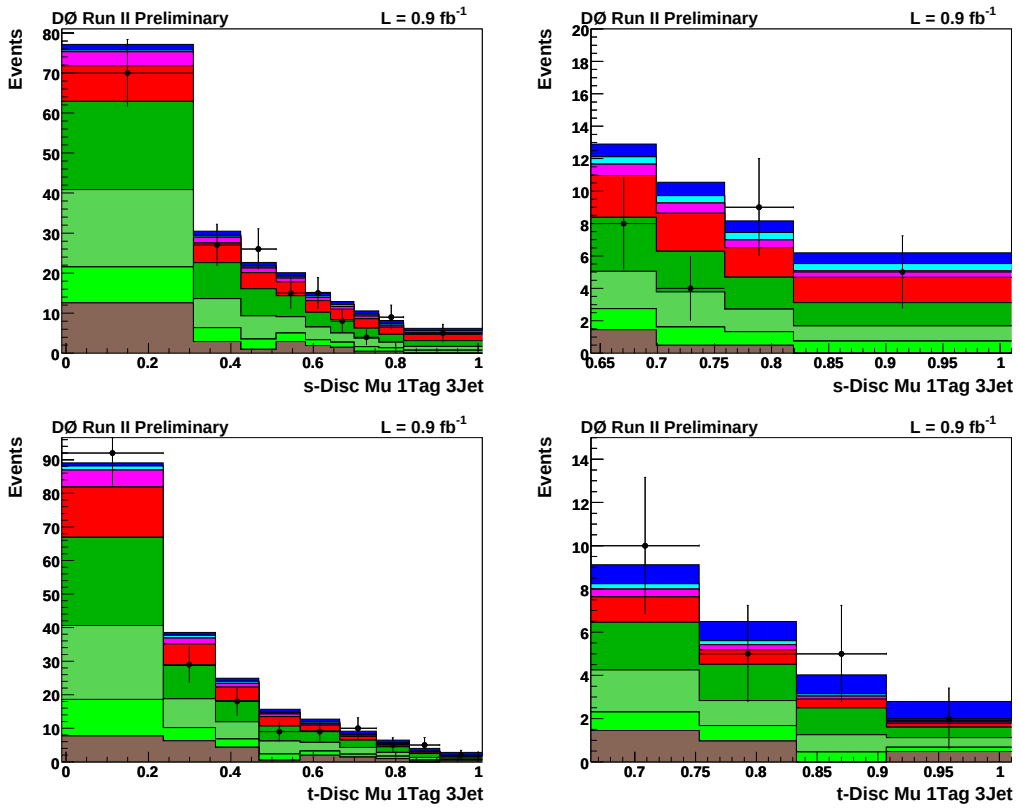


Figure C.7: 1D Discriminant projection for the muon channel with one b tag. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.

MATRIX ELEMENT OUTPUTS FOR THE MUON CHANNEL WITH THREE JETS

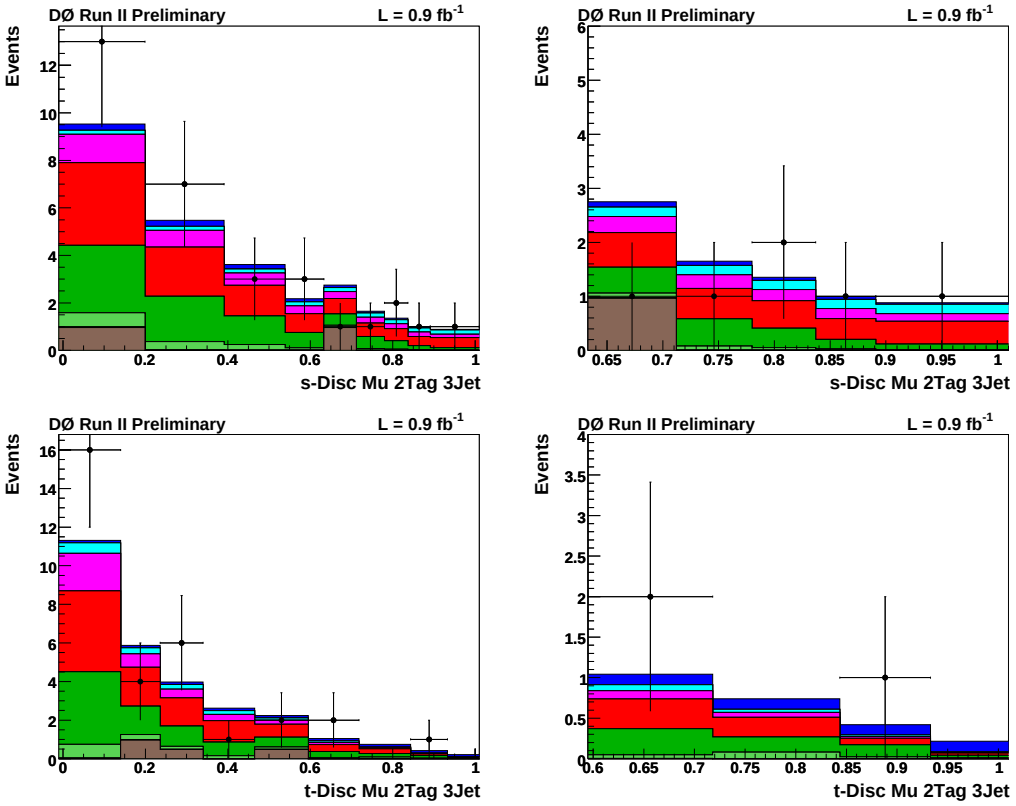


Figure C.8: 1D Discriminant projection for the muon channel with two b tags. Upper row: s-channel discriminant, lower row: t-channel discriminant. Left column, full discriminant range; right column, the high end of the distribution.