

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Measurement of The Single Top Quark Production Cross Section at $\sqrt{s}=1.96$ TeV

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Mark Anthony Padilla

December 2011

Dissertation Committee:

Dr. Stephen Wimpenny, Chairperson

Dr. Robert Clare

Dr. Jose Wudka

Copyright by
Mark Anthony Padilla
2011

The Dissertation of Mark Anthony Padilla is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would first and foremost like to thank my mother and father Cecilia and Rumaldo Padilla for their love and support throughout the years. None of this would have been possible without them. Secondly, I would like to thank my wife Mehgan for all her support throughout the past six years, thankyou honey, I love you. I would especially like to thank my advisor Stepen Wimpenny for all his guidance and support and in helping me understand many of the topics related to particle physics, thanks a million Steve.

I would like to give special thanks to the Single Top Group at Fermilab for all their help and support throughout the years. Liang, thanks for always being willing to help with root, C++ ect., you helped me more than you know. Victor, thanks for always pushing me and always having the time to help with code and many other things related to the single top group. Reinhard, thankyou for your help and always giving me clear and proper direction relative to the task at hand. Ann, thankyou for always being more than willing to help and for some hard lessons learned. Cecilia, thanks for keeping us on track with the big picture and helping us realize what needs to be done. Nathan, thanks for helping me trouble shoot those annoying pieces of code and collaborating with me on the BDTs. Jyoti, Wiegang and Yun-Tse thanks for always being more than willing to help and the hard work all three of you have put into the group, good luck in the future.

As for my peers and teachers that have had a direct influence on helping me gain a greater understanding of physics I would like to thank Medina, Gordon, Bob, Ryan, George, Ben, Tim, Kathy, Dennis, Gil, Santino, Vince and Bruce, thanks for the inspiration!

To Rumaldo Padilla 1940 - 2006,
the best father ever.

ABSTRACT OF THE DISSERTATION

Measurement of The Single Top Quark Production Cross Section at $\sqrt{s}=1.96$ TeV

by

Mark Anthony Padilla

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, December 2011
Dr. Stephen Wimpenny, Chairperson

Within the standard model top quarks are predicted to be most often produced in pairs via the strong interaction. However they can also be produced singly through the weak interaction. This is a rarer process with many experimental challenges. It is interesting because it provides a new window to search for evidence of physics beyond the standard model picture, such as a fourth generation of quarks or to search for insight into the Higgs Mechanism. Single top production also provides a direct way to calculate the CKM matrix element V_{tb} .

This thesis presents new measurements for single top quark production in the $s+t$, s and t channels using 5.4 fb^{-1} of data collected at the $D\bar{O}$ detector at Fermilab in Batavia, IL. The analysis was performed using Boosted decision trees to separate signal from background and Bayesian statistics to calculate all the cross sections. The results obtained are:

$$s \text{ channel : } \sigma(p\bar{p} \rightarrow tb + X) = 0.68^{+0.41}_{-0.39} \text{ pb,}$$

$$t \text{ channel : } \sigma(p\bar{p} \rightarrow tqb + X) = 3.03^{+0.78}_{-0.66} \text{ pb,}$$

$$\text{Combined } s+t \text{ channels : } \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.01^{+0.80}_{-0.75} \text{ pb.}$$

The s channel has a Gaussian significance of 1.6σ , the t channel 5.5σ and the $s + t$ 4.3σ . The results are consistent with the standard model predictions within one standard deviation. By combining these results with the results for two other analyses (using different MVA techniques) improved results of:

$$\text{Combined } s + t \text{ channels : } \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.43^{+0.73}_{-0.74} \text{ pb},$$

$$t \text{ channel : } \sigma(p\bar{p} \rightarrow tqb + X) = 2.90^{+0.59}_{-0.59} \text{ pb},$$

were obtained. These give a 5.6σ significance for the combined $s + t$ channel and 5.5σ for the t channel.

Contents

Acknowledgments	iv
Dedication	vi
Abstract	vii
List of Tables	ix
List of Figures	x
1 Introduction	1
2 Theory	3
2.1 The Standard Model	3
2.1.1 Particles	3
2.1.2 Forces	4
2.1.3 Gauge Theories	4
2.2 The Top Quark	6
2.2.1 Top Quark Production	7
2.2.2 Pair Production	7
2.3 Single Top Quark	11

2.3.1	Introduction	11
2.3.2	Production	11
2.3.3	Signal and Background	12
2.3.4	CKM Matrix Element $ V_{tb} $	14
2.3.5	Polarization	16
2.3.6	Beyond The Standard Model	16
3	Accelerators and Detectors	18
3.1	Introduction	18
3.2	Cockroft-Walton Pre-accelerator	19
3.3	Linear Accelerator (Linac)	21
3.4	Proton Booster Ring	21
3.5	Fermilab Main Injector	21
3.6	Antiproton Production	22
3.7	Tevatron	23
3.8	$D\bar{O}$ Detector	24
3.8.1	$D\bar{O}$ Coordinate System	24
3.8.2	Central Tracking System	26
3.8.3	$D\bar{O}$ Calorimeters	30
3.8.4	Muon Detectors	33
3.8.5	Luminosity Monitor	34
3.9	Trigger Framework	36
4	Event Reconstruction	39
4.1	Charged Track Reconstruction	39
4.2	Primary Vertex	41
4.3	Electron Identification	42

4.4	Muon Reconstruction	45
4.5	Jet Reconstruction	46
4.5.1	Calibration of the Jet Energy Scale	48
4.5.2	b -Jet Identification	51
4.6	Reconstruction of Missing Transverse Energy, \cancel{E}_T	52
5	Data Samples and Event Selection	55
5.1	The RunII Data Sample	55
5.2	Simulated Event Samples	56
5.2.1	Simulation of Single Top Signal	57
5.2.2	Background Event Simulation	57
5.3	Modeling the QCD Multijet Background	61
5.4	The Effects of the Top Mass Uncertainty	61
5.5	Determination of the W +Jets and Multijet Background Normalizations	62
5.6	Event Selection	64
5.6.1	General Preselection Requirements	64
5.6.2	Electron+Jets Channel Preselection	66
5.6.3	Muon+Jets Channel Preselection	67
5.7	Analysis Subsamples and Preliminary Data-Simulation Comparisons	69
6	Monte Carlo Model Corrections and Statistical Uncertainties	74
6.1	Detector Response Corrections	74
6.2	Theoretical Uncertainties	78
6.3	Residual Systematic Uncertainties	80
6.3.1	Normalization Effects	82
6.3.2	Shape-changing Effects	85
6.3.3	Summary of Residual Uncertainties	87

7	Boosted Decision Trees	89
7.1	Introduction	89
7.2	Sample Preparation	92
7.3	Training	92
7.3.1	Splitting a Node	93
7.3.2	Boosting	94
7.3.3	Parameters	95
7.3.4	Implementation	97
7.3.5	Selection of Input Variables	97
7.4	Transformation of the BDT Discriminant	107
7.5	Background Cross Checks	109
8	Analysis	112
8.1	Cross Section Measurement	112
8.1.1	Bayesian Approach	113
8.1.2	Systematic Uncertainties	115
8.2	Results	116
8.2.1	Expected Results	116
8.2.2	Observed Results	118
8.3	Ensemble Tests	122
8.3.1	Significance	126
8.3.2	Asymptotic approximation of the log-likelihood ratio	126
8.3.3	Measured Significance	127
8.4	$ V_{tb} $ Measurement	130
9	Comparison of Results	132
9.1	$D\bar{O}$ and CDF Measurements	132

9.1.1 Comparisons	133
10 Conclusions	139
A Decision Tree Outputs	141
B Residual Systematic Errors	145
B.1 Normalization Uncertainties	145
B.2 Shape and Normalization Uncertainties	149
Bibliography	156

List of Tables

2.1	Elementary Particles and their properties.	4
2.2	Top quark branching fractions [7]	7
2.3	On mass shell W boson branching ratios [8]	8
2.4	QCD prediction for the NLO cross section for single top quark production at $\sqrt{s}=1.96\text{TeV}$	12
4.1	Definitions for three classes of electron candidates.	45
5.1	Luminosity breakdown of the Run II dataset as a function of trigger version.	56
5.2	Sizes of the signal and background simulated event samples. The labels Run IIa and Run IIb, denote that these events were reconstructed with the software used for the data from these subsets of the data.	58
5.3	Fitted W+Jets and QCD Multijet Scale Factors for each of the 24 fits.	63
5.4	Comparison of observed and predicted yields for the single tagged event samples.	70
5.5	Comparison of observed and predicted yields for the double tagged event samples.	71

6.1	Residual systematic uncertainties after improvements to detector and process simulations. Where ranges are given, these correspond to 2, 3 and 4 jet events.	88
7.1	List of individual object kinematics.	100
7.2	List of event variables.	100
7.3	List of variables using the sign of lepton charge.	100
7.4	Calculated mass, p_T and the H_T variables.	101
7.5	Reconstructed top quark variables.	101
7.6	Reconstructed W boson variables.	102
7.7	Angular Correlations.	102
8.1	Expected cross sections for $s+t$, s and t channels in pb.	118
8.2	Expected posterior peak over half-width ratios for $s+t$, s and t channels.	118
8.3	Expected cross sections in individual analysis channels (e,μ) combined.	119
8.4	Observed cross sections for $s+t$, s and t channels in pb.	119
8.5	Observed posterior peak over half-width ratios for $s+t$, s and t channels.	119
8.6	Observed cross sections in individual analysis channels (e,μ) combined.	120
8.7	Cross section inputs for each individual channel.	122
8.8	Parameters of the likelihood which are used in the calculation of the significances using the AALR approach.	128
8.9	Expected and observed results for the single top cross sections.	130
9.1	$D\emptyset$ significance results for $s+t$, s and t channels.	135
9.2	$D\emptyset$ and CDF significance results for $s+t$, s and t channels.	137
B.1	Residual normalization systematic uncertainties: two jets with 1- b tag. These values are used as input parameters to the BDT analysis. . . .	146

B.2	Residual normalization systematic uncertainties: two jets with 2- <i>b</i> tags.	
	These values are used as input parameters to the BDT analysis. . . .	146
B.3	Residual normalization systematic uncertainties: three jets with 1- <i>b</i> tag.	
	These values are used as input parameters to the BDT analysis. . . .	147
B.4	Residual normalization systematic uncertainties: three jets with 2- <i>b</i> tags. These values are used as input parameters to the BDT analysis.	147
B.5	Residual normalization systematic uncertainties: four jets with 1- <i>b</i> tag.	
	These values are used as input parameters to the BDT analysis. . . .	148
B.6	Residual normalization systematic uncertainties: four jets with 2- <i>b</i> tags.	
	These values are used as input parameters to the BDT analysis. . . .	148

List of Figures

2.1	The fundamental particles and their interactions [2]. The lines connecting the particles indicate which interaction can take place.	5
2.2	Feynman diagrams for $t\bar{t}$ production. The top diagram shows quark-antiquark annihilation, the bottom three show gluon fusion [9].	9
2.3	Pie chart representing all branching fractions for for $t\bar{t}$ pair events [9].	10
2.4	Feynman diagram at LO for t -channel production.	12
2.5	Feynman diagram at LO for s -channel production.	13
2.6	Feynman diagrams for background processes related to single top quark production. W+jets (top-left), Z+jets(top-right), $t\bar{t}$ (middle-left), Diboson(middle-right) and QCD multijets (bottom).	15
2.7	Feynman diagrams showing BSM processes through the s -channel (left) and t -channel (right) single top quark production.	17
3.1	Diagram of Fermilab's chain of accelerators.	18
3.2	(a) The Cockroft-Walton Accelerator. (b) Voltage multiplier circuit diagram.	20
3.3	H^- ion production in the magnetron.	20
3.4	Schematic diagram of the Target and Lens system. [19].	23

3.5	Schematic of the experimental hall and the $D\bar{O}$ detector. The location of each of the major subsystems is shown, along with the position of the on-detector readout electronics.	25
3.6	$D\bar{O}$ coordinate system.	25
3.7	Schematic of the central tracking showing the location of SMT and CFT components with respect to the beam axis.	27
3.8	The elements of the SMT detector, showing the barrel and disc segmentation.	28
3.9	Schematic representation of the $D\bar{O}$ magnets and the magnetic field within the detector volume.	29
3.10	Schematic diagram shown the $D\bar{O}$ central endcap calorimeters. The cut away shows the details of the calorimeter modules [26].	30
3.11	Depiction of a photon shower where X_0 represents one radiation length.	31
3.12	Structure of a unit cell in the liquid Ar calorimeter modules.	32
3.13	Schematic of one octant of the $D\bar{O}$ central and endcap calorimeters. The shaded regions show the projective structure of the calorimeter towers and their map onto pseudorapidity.	33
3.14	Exploded view of the $D\bar{O}$ muon detection system. The top diagram shows the configuration of the A, B and C layers in the barrel and forward regions. The bottom diagram shows the location of the barrel and forward scintillation counters.	35
3.15	Diagram of the Luminosity monitors at $D\bar{O}$ [26].	36
3.16	Block diagram of the $D\bar{O}$ framework, showing the data-flow through the three levels of the trigger.	37

3.17	Block diagram of L1 and L2 showing what systems they are connected to [26]. All detectors are listed on the left where the calorimeter (CAL) is initially connected to the level 1 trigger, the central tracking trigger (CTT) is defined as both the preshower detectors (CPS and FPS) and the fiber tracker (CFT) and is also connected to the L1 trigger. The SMT goes straight to L2 where the muon system begins at L1. The forward proton detector (FPD) and the Luminosity monitor are also connected to L1 initially.	38
4.1	Diagram of track reconstruction using the AA algorithm.	41
4.2	Schematic of the parton fragmentation process and the particle showering used to detect the fragmentation products in the calorimeters. .	48
4.3	F_η measurements as a function of η for $ \eta < 3.5$	50
4.4	Diagram of the MPF method for a $\gamma + jet$ event.	51
4.5	Schematic of a 3-jet event in which one of the jets originates from a b -quark and has a displaced secondary vertex. The origin of the non-tagged jets is consistent with the primary vertex.	53
5.1	Lowest order Feynman diagrams for the s and t channel signal processes and the principle backgrounds.	60
5.2	Triangle cut in the $ \Delta\phi(\text{leading jet}, \cancel{E}_T) $ vs. \cancel{E}_T plane.	65
5.3	Triangle cut in the second leading jet p_T vs. $H_T(\text{all jets})$ plane.	66
5.4	Triangle cuts in the \cancel{E}_T vs. $ \Delta\phi(e, \cancel{E}_T) $ plane.	67
5.5	Triangle cuts in the \cancel{E}_T vs. $ \Delta\phi(\mu, \cancel{E}_T) $ plane.	69
5.6	Colour scheme plot key used to label Data, signal and background events for plots.	72

5.7	Comparison of the observed and predicted yields for 15 of the kinematic variables used as inputs to the BDT.	73
6.1	The two plots on the left are $\varepsilon_{\text{fake}-e}$ and $\text{QCDWeight} = \varepsilon_{\text{fake}-e}/(1 - \varepsilon_{\text{fake}-e})$. The two on the right are the same thing with the real electron contamination subtracted out.	81
6.2	The plot on the left has no correction whereas the plot on the right has been corrected.	81
7.1	Schematic of a decision tree showing the effects of branching and the terminal leaves. The blue circles represent nodes which are split into daughter nodes. Each daughter node is subjected to an independent test to see whether it will continue to be split or end as a terminal leaf (green). Here F and P stand for fail and pass the nodal test. . . .	91
7.2	Example plots showing separation power and cross section significance versus the number of boosts.	96
7.3	Data-Simulation comparison for selected variables I.	103
7.4	Data-Simulation comparison for selected variables II.	104
7.5	Data-Simulation comparison for selected variables III.	105
7.6	Data-Simulation comparison for selected variables IV.	106
7.7	Discriminant Transformation Histogram a.) is the raw distribution, b.) is the transformed distribution and c.) is the transformation function. . . .	108
7.8	BDT cross check samples for s+t, s and t where the histograms dominated by red correspond to $t\bar{t}$ events and the green dominated ones indicate W +Jet events.	110

7.9	BDT cross check samples for $s+t$, s and t where the histograms dominated by red correspond to $t\bar{t}$ events and the green dominated ones indicate W +Jet events.	111
8.1	Representation of a 1D posterior density distribution where the peak of the curve represents a cross section measurement and $\Delta\sigma$ its uncertainty.	115
8.2	Representation of shape changing uncertainties.	116
8.3	Expected posterior density distributions for $s+t$ (top left), s (top right) and t (bottom). The results include both statistical and systematic uncertainties.	117
8.4	BDT output distributions for $s+t$, s and t channel analyses with all six channels combined. The top row corresponds to the linear scale, the middle the log scale and the bottom are zoomed plots showing signal excess over background.	120
8.5	Observed posterior density distributions for the single top $s+t$ (top left), s (top right) and t (bottom) analyses.	121
8.6	Results of ensemble tests for the combined $s+t$ channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.	123
8.7	Results of ensemble tests for the combined s channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.	124
8.8	Results of ensemble tests for the combined t channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.	125

8.9	Probability density vs. log-likelihood significance for the $s + t$ channel analysis.	128
8.10	Probability density vs. log-likelihood significance for the s channel analysis.	129
8.11	Probability density vs. log-likelihood significance for the t channel analysis.	130
8.12	Posterior densities for both $ V_{tb}f_1^L ^2$ (left) and $ V_{tb} ^2$ (right).	131
9.1	$D\bar{O}$ cross section results for 2.3 fb^{-1} , 5.4 fb^{-1}	135
9.2	Cross section results for 2.3 fb^{-1} ($D\bar{O}$), 5.4 fb^{-1} ($D\bar{O}$) and 3.2 fb^{-1} (CDF) data sets.	136
9.3	2D posterior density for BNN Combination t . Below are the 1D posterior densities for both s and t obtained by integrating over both axes.	138
A.1	Boosted decision tree outputs for the $s + t$ channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.	142
A.2	Boosted decision tree outputs for the s channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.	143

A.3	Boosted decision tree outputs for the t channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.	144
B.1	Combined $s + t$ channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	150
B.2	Combined $s + t$ channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	151
B.3	Combined s channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	152
B.4	Combined s channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	153

B.5	Combined t channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	154
B.6	Combined t channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.	155

Chapter 1

Introduction

It all started when I was eight years old with firecrackers and toy cars. I always wondered what toy cars were made of so I decided to use a firecracker to blow up the minature car to find out. After many attempts I finally accomplished my task with a suitable sized firecracker (M-80), however I was not impressed. My experiment yielded burnt plastic and charred metal. Years later I find that I'm still doing the same thing, well, not demolishing toy cars but trying to figure out the fundamental constituent's of nature and their processes. The model that best describes the fundamental particles of nature and how those particles interact with one another is known as the Standard Model (SM). The interactions between fundamental particles occur via the four known forces of nature: electromagnetic, strong, weak and gravitaional. The SM only incorporates the first three of these forces and it has been very successful in describing nature and its interactions. However, like many theories it is still a work in progress. One way to probe the limits of a theory is through the use of a high energy particle accelerator.

The second largest particle accelerator in the world is the Tevatron which is located at Fermilab in Batavia, IL. This is the machine that was used for the discovery of

the Top Quark in 1995 by the $D\bar{O}$ and CDF collaborations. The discovery was huge, because not only did it further validate the SM, but it also resulted in the observation of the most massive fundamental particle ever observed. The top quark weighs in at approximately 172 GeV, which is over 170 times the mass of a proton. Because of its large mass, the top quark makes a good starting point to search for theories beyond the current standard model. In 1995 the discovery of the top quark was based on the production of $t\bar{t}$ pairs through the strong force. More recently (2009) single top quark production was observed for the first time. This proceeds through the weak force in three production channels, called the s , t , and tW channels. This analysis makes refined measurements of the combined $s+t$ channel cross section, and separate measurements of the s and t channel cross sections. The tW channel is not currently measurable because of the large background coming from $t\bar{t}$ pair production, which is kinematically almost indistinguishable.

This thesis presents the following chapters: Chapter 2 gives an overview of the standard model and current theories relating to particle physics. Chapter 3 describes the accelerator used to produce top quarks and the $D\bar{O}$ detector which was used to observe them. Chapter 4 discusses how objects such as electrons, muons and other physics objects are reconstructed from a proton antiproton collision. Chapter 5 describes the data collection, signal and background modeling, and the basic event selection. The corrections made to the Monte Carlo simulation and the systematic uncertainties are discussed in Chapter 6. Chapters 7 and 8 discuss the use of Boosted Decisions Trees to identify the signal and the measurement of the cross sections. Lastly, Chapters 9 and 10 compare the results to previous measurements and the predictions of theory.

Chapter 2

Theory

2.1 The Standard Model

2.1.1 Particles

In the Standard model the fundamental particles are divided into three categories which are, Leptons, Quarks and Bosons. The leptons and quarks are fermions and they both carry half integer spin. The bosons are the force carriers and they have integer spin. The lepton and quark families are separated into three generations as shown in Table 2.1. Each generation has two doublets: one of quarks and the second of leptons associated with it. Each quark doublet consists of a particle with $+2/3e$ charge and a second particle with $-1/3e$ charge whereas the lepton doublets have one particle with $+1e$ charge and a second which is neutral. The particle masses vary with the first generation being the lightest and the third the heaviest. For each particle in all generations there is an associated anti-particle with the same mass but opposite charge. The particles in the first generation are the basic building blocks for the known universe.

Generation	Quark			Lepton		
	Flavor	Charge	Mass(MeV/ c^2)	Flavor	Charge	Mass(MeV/ c^2)
I	Up(u)	$+2/3e$	7.5	Electron(e)	$-e$	0.511
	Down(d)	$-1/3e$	4.2	Neutrino(ν_e)	0	$<2.0 \times 10^{-4}$
II	Charm(c)	$+2/3e$	1.1×10^3	Muon(μ)	$-e$	105
	Strange(s)	$-1/3e$	150	Neutrino(ν_μ)	0	<0.19
III	Top(t)	$+2/3e$	173×10^3	Tau(τ)	$-e$	1784
	Bottom(b)	$-1/3e$	4.2×10^3	Neutrino(ν_τ)	0	<18.2

Table 2.1: Elementary Particles and their properties.

2.1.2 Forces

Whenever we look out into nature it is safe to say that we always observe things interacting. This takes place via the four fundamental forces: gravity, electromagnetism, weak and strong. Each force is transmitted through a mediating boson. In the case of gravity, the boson is called the graviton. However theorists have had serious issues reconciling general relativity with quantum mechanics, therefore gravity is not included within our current standard model. The photon is the mediator particle for the electromagnetic force and the mediators for the weak force are the W^+ , W^- and Z bosons. Gluons mediate the strong force. Last but not least is the Higgs Boson which is predicted by the SM and is supposed to give rise to the particle mass. The Higgs boson has not yet been observed, and there are ongoing searches taking place at CERN (Large Hadron Collider) and Fermilab (Tevatron). These searches have excluded some of the mass regions where the Higgs could exist, however no signal has been observed so far. A diagram of the particles and their associated interactions is shown in Figure 2.1.

2.1.3 Gauge Theories

Gauge Theory is a mathematical model created by physicists in order to explain the interactions of fermions. The field theory which combines the effects of special

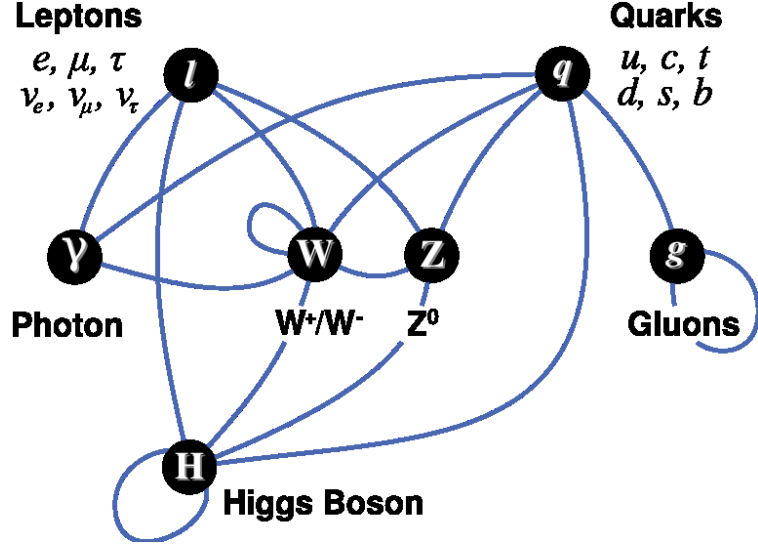


Figure 2.1: The fundamental particles and their interactions [2]. The lines connecting the particles indicate which interaction can take place.

relativity and quantum mechanics is known as quantum field theory (QFT). The main concept behind a gauge theory is that the Lagrangian remains invariant for local and global symmetry transformations where the global is just a subset of the local. The theory began in the 1920's when physicists were trying to create a quantum theory for the electromagnetic interaction. This is now called quantum electrodynamics (QED). If we extrapolate from the 1920's to now, one of the most successful gauge theories is the standard model of particle physics.

The gauge group $SU(2)_L \times U(1)_Y$ represents the electroweak interaction which is the unifying group for the electromagnetic ($U(1)$) and weak interactions ($SU(2)$). The subscript L points out that the weak force is associated with only left-handed particles and Y signifies the weak hypercharge. The gauge group $SU(3)_C$ is associated with the strong interaction where the C indicates color charge. Together these groups form the standard model gauge group, $SU(3)_C \times SU(2)_L \times U(1)_Y$.

One final note should be presented, that of “spontaneous symmetry breaking”. If the symmetry of the gauge group $SU(2)_L \times U(1)_Y$ is to remain unbroken then this

requires that all its mediator particles be massless. Through experiment we have seen that the W and Z bosons have masses of 80.4 GeV and 91.2 GeV [1]. To deal with this inconsistency theorists created the “Higgs Mechanism”. This allows for the W and Z bosons to have mass while preserving the theory. However the introduction of this produces another particle called the Higgs Boson, which has not yet been observed.

2.2 The Top Quark

The roots of the top quark lie in a postulate made by Makoto Kobayashi and Toshihide Maskawa in 1973. They were trying to understand CP violation in Kaon decays [3] by introducing a new generation of quarks which is now known to be the third generation. However, it was not until 1977 that one of the quarks in the third generation doublet, the b quark, was discovered [4]. Following the generational pattern of the SM the b quark was predicted to have an isospin partner named the t or top quark. Some 18 years later the top quark was finally observed by the $DØ$ and CDF collaborations [5].

The top quark is by far the heaviest particle in the SM model, weighing in at a massive 172 GeV ($\sim 170m_{proton}$). It is similar to the u and c quarks in that it is a spin-1/2 particle and carries a charge of $+2/3e$. However, because the top quark is so massive it decays very rapidly with a lifetime of $\sim 5 \times 10^{-26}$ s, which means it has no time to hadronize before it decays. This special property allows physicists to study the polarization and the angular correlations related to its decay products. The top quark interacts through both the strong and weak forces where the strong interaction produces $t\bar{t}$ pairs and the weak interaction produces top quarks singly. Both these production modes are discussed in the following sections. This thesis focuses on the analysis of the single top quark production mode.

2.2.1 Top Quark Production

Currently, the Tevatron at Fermilab and the LHC at CERN are the only two particle accelerators in the world capable of creating top quarks. Here, we will only focus on top quarks created at Fermilab since this thesis is based on data taken from there. The Tevatron creates top quarks by colliding a proton with an antiproton at a center of mass energy of 1.96 TeV. There are two types of production modes: $t\bar{t}$ pairs and single top quark production. Between these two, pair production dominates.

2.2.2 Pair Production

As mentioned earlier, the top quark was first observed in 1995 by the $D\bar{O}$ and CDF collaborations. It was observed in pairs, one top and one anti-top. The different sub-processes in which $t\bar{t}$ production occurs is shown by the Feynman diagrams in Figure 2.2. Of these, the top diagram ($q\bar{q}$ annihilation) is the dominant production mode at the Tevatron. Currently, the $t\bar{t}$ production cross section is measured to be $7.78^{+0.77}_{-0.64}$ [6]. The $q\bar{q}$ annihilation accounts for around 85% of the cross section, with the remaining 15% coming from the gluon fusion processes (bottom three diagrams). After the top quarks are created they decay very quickly to a W boson and a b quark almost 100% of the time. The branching fractions for the top quark decay can be seen in Table 2.2.

Decay Mode	Branching Fraction
$t \rightarrow Wd$	$\sim 0.006\%$
$t \rightarrow Ws$	$\sim 0.17\%$
$t \rightarrow Wb$	$\sim 99.8\%$

Table 2.2: Top quark branching fractions [7]

Clearly the $t \rightarrow Wb$ decay dominates. In this mode the b quark fragments into

a jet of particles while the W boson decays into a lepton and a neutrino or $q\bar{q}$ pair. The branching ratios for W decay are shown in Table 2.3. All $t\bar{t}$ pair branching

Decay Mode	Branching Fraction
$W^+ \rightarrow e\nu$	(10.75 ± 0.13)
$W^+ \rightarrow \mu\nu$	(10.57 ± 0.15)
$W^+ \rightarrow \tau\nu$	(11.25 ± 0.20)
$W^+ \rightarrow \textit{hadrons}$	(67.60 ± 0.27)

Table 2.3: On mass shell W boson branching ratios [8]

fractions are shown in Figure 2.3. The dilepton process accounts for 9% of all $t\bar{t}$ pair production, 45% for lepton+jets and 46% for alljets.

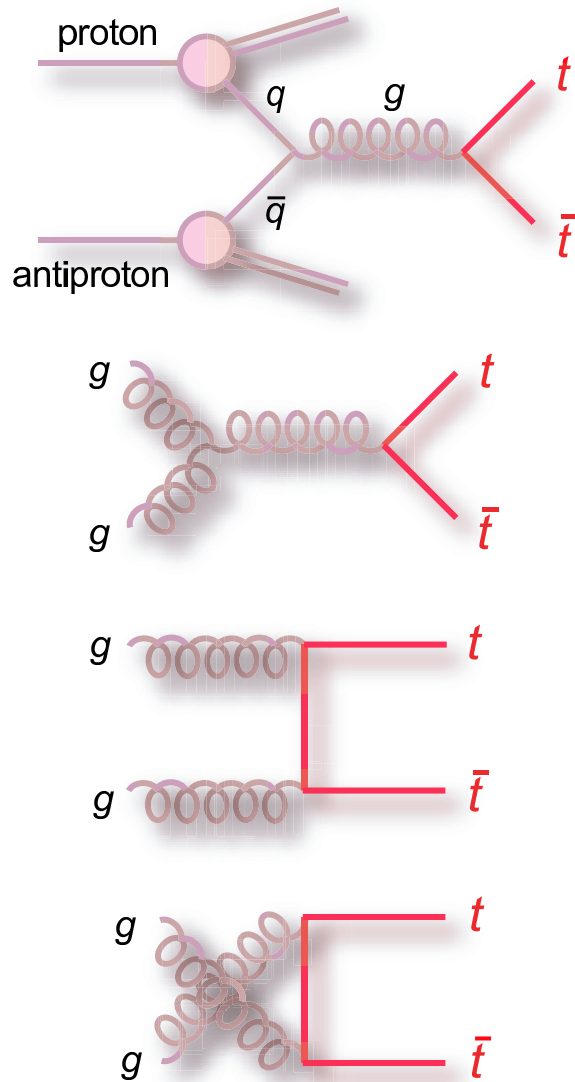


Figure 2.2: Feynman diagrams for $t\bar{t}$ production. The top diagram shows quark-antiquark annihilation, the bottom three show gluon fusion [9].

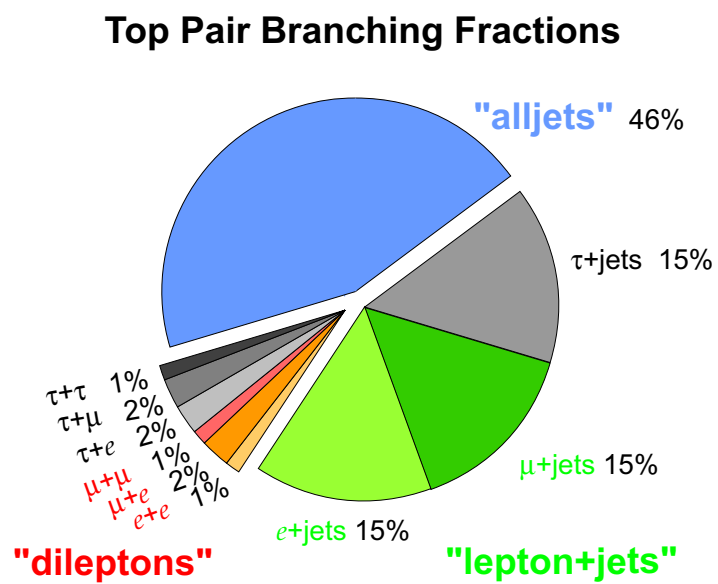


Figure 2.3: Pie chart representing all branching fractions for for $t\bar{t}$ pair events [9].

2.3 Single Top Quark

2.3.1 Introduction

Single top quark production is unlike $t\bar{t}$ pair production in that it does not occur through the strong interaction but the weak, also we only get one top quark instead of two. At the Tevatron this process occurs about once in every 10 billion events making it a very difficult particle to detect. In 2009, the single top quark process was first observed by the $D\bar{O}$ and CDF collaborations [10], 14 years after the top quark was first observed. This was quite a triumph in the physics community, not only because it took a long time to find but because this opens up a new area for searches for physics that may lie outside the SM. It also has allowed us to study different properties associated with the top quark. The last three sections of this chapter is devoted to a brief description of the properties of the top quark and Beyond the standard model (BSM) searches.

2.3.2 Production

The single top quark is produced in three different channels, that is the t channel, the s channel and the Wt channel. The latter channel has the smallest cross section at the Tevatron and is neglected here because of this and its similarity to the $t\bar{t}$ state. The t channel is the dominant channel followed by the s channel. The next-to-leading order (NLO) quantum chromodynamics (QCD) predictions for the cross sections for the three channels is shown in Table 2.4 [11].

The t channel process occurs when a quark from the proton (antiproton) exchanges an off mass shell W boson with a b quark producing a single top quark. The b quark comes from a gluon which originates from the antiproton (proton) sea then

Process	Cross Section(pb)
t -channel	2.26 ± 0.12
s -channel	1.04 ± 0.04
tW	0.28 ± 0.15

Table 2.4: QCD prediction for the NLO cross section for single top quark production at $\sqrt{s}=1.96\text{TeV}$

pair produces into a $b\bar{b}$ pair. The leading order Feynman diagram for this process is shown in Figure 2.4. We label this process “ tqb ” which means that it can either be a $tq\bar{b}$ or $\bar{t}qb$ (charge conjugate) state.

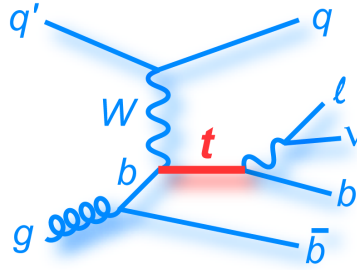


Figure 2.4: Feynman diagram at LO for t -channel production.

The s channel process occurs by $q\bar{q}$ annihilation where both originate from the proton and antiproton producing an off mass shell W boson which then decays into a t quark and a b quark. The leading order feynman diagram for this process is shown in Figure 2.5. We name this process “ tb ” which means that it can either be in a $t\bar{b}$ or $\bar{t}b$ state.

2.3.3 Signal and Background

We define signal to be the processes we are looking for such as the t and s channels. However, there are several other processes that can mimic signal and we label these as background. In order to make a proper cross section measurement it is necessary

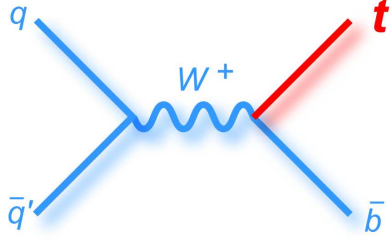


Figure 2.5: Feynman diagram at LO for s -channel production.

to be able to discriminate between the two. If we look back to Figures 2.4 and 2.5 we can see the decay products for the two signal channels. What we are looking for when these events occur are the following: One isolated lepton (e or μ) with high p_T , large \cancel{E}_T in the form of a neutrino, 1-2 b -tagged jets and 1-3 light quark jets both with high p_T . We also need to consider that some jets come from the remaining quarks in the proton and antiproton. This will be discussed in more detail in the selection chapter (Chapter 5).

The background processes that can mimic single top production are: W +jets, Z +jets, $t\bar{t}$, QCD and Diboson production. The feynman diagrams for each of these processes are shown in Figure 5.1. The following list summarizes how each of the background processes can mimic a single top signal:

- Two jets which are close together can be merged into a single jet leading to a lower jet multiplicity,
- A lepton can be misidentified as a jet and vice versa,
- Jets can be mistagged as b jets when they are not,
- A second lepton in the event (e or μ) may not be detected leading to a fake \cancel{E}_T .

Also some of the jet energies may be misreconstructed which can also lead to a

fake \cancel{E}_T .

As an example consider the QCD process in Figure 5.1 (bottom). This is an all jet final state. However if one of the jets fakes the lepton signature, another is misidentified by the b -tagging algorithm and the remaining jets energy is not fully measured, then the resulting signature is one lepton, \cancel{E}_T , and two jets. This would pass the event selection for a single top event.

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

CKM stands for Cabibbo-Kobayashi-Maskawa, which are the last names of the physicists that created the CKM matrix. Nicola Cabibbo initially introduced (1963) a matrix that described the probability that d and s quarks would decay to u quarks through the weak interaction. In 1974 the Charm quark was observed and this matrix became a 2×2 matrix. Kobayashi and Maskawa later came along and realized that another quark generation needed to be added to explain CP violation because the two generation of quarks did not do the job. The CKM matrix then became a 3×3 matrix which can be seen below.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.1)$$

Each of the CKM elements when squared ($|V_{xy}|^2$) denote the probability that quark x decay to quark y . If the CKM matrix assumes unitarity and only three quark generations then we get the following relation:

$$|V_{ub}|^2 + |V_{cb}|^2 + |V_{tb}|^2 = 1. \quad (2.2)$$

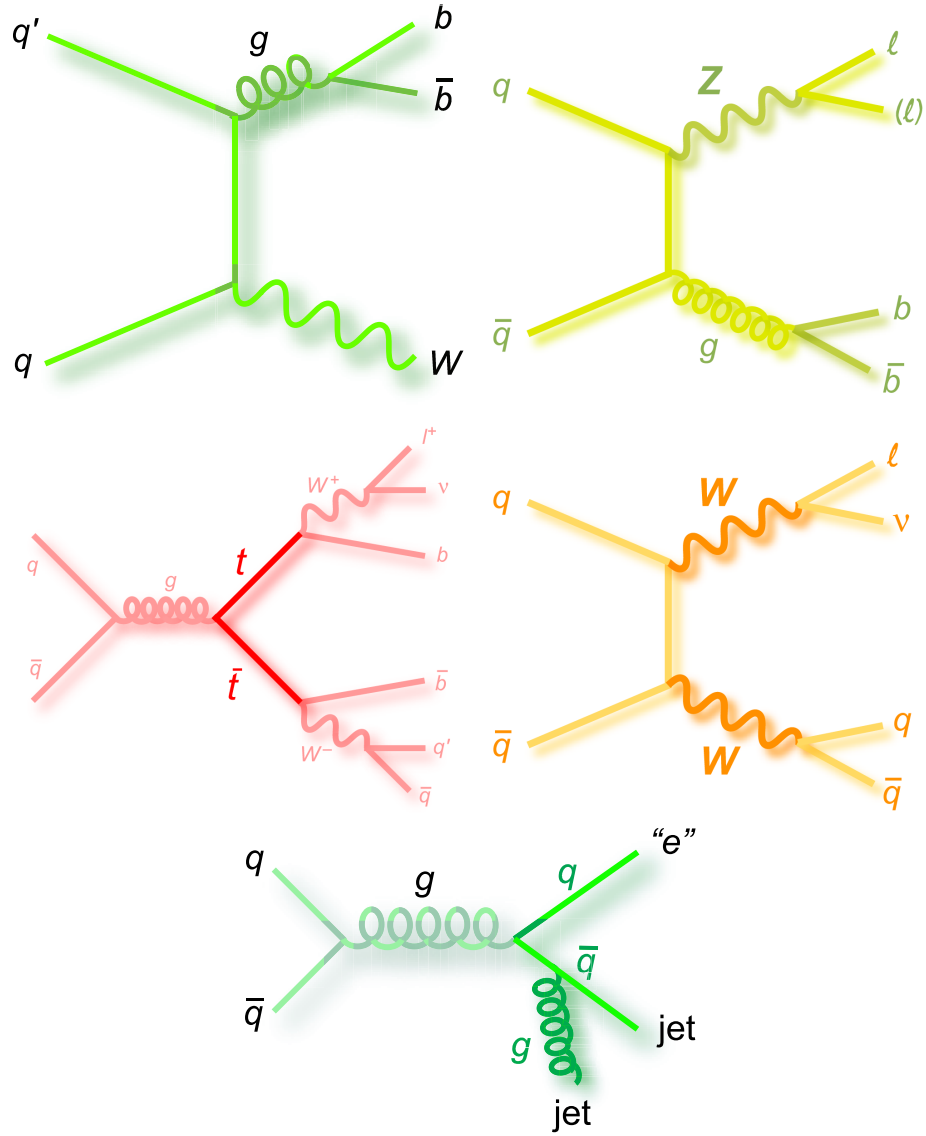


Figure 2.6: Feynman diagrams for background processes related to single top quark production. W+jets (top-left), Z+jets (top-right), $t\bar{t}$ (middle-left), Diboson (middle-right) and QCD multijets (bottom).

Here, the first two components ($|V_{ub}|$ and $|V_{cb}|$) have been measured very accurately [1] which leads to a value for $|V_{tb}|$. However, $|V_{tb}|^2$ is also directly proportional to the single top quark cross section. Thus a measurement of the singletop cross section allows us to measure $|V_{tb}|$ without imposing the three generations and unitarity constraint. The relationship between the Wtb vertex and the matrix element $|V_{tb}|$ can be seen through the following factor:

$$-\frac{ig_w}{2\sqrt{2}}V_{tb}\gamma^\mu(1-\gamma^5). \quad (2.3)$$

This thesis presents a measurement of $|V_{tb}|$ which can be seen in Section 8.4.

2.3.5 Polarization

As stated earlier, the lifetime of the top quark is $\sim 5 \times 10^{-26}$ s and because of this the top quark does not fragment. A consequence of this is that all of its spin information is transmitted to its decay particles (W and b). In the SM the weak interactions are left handed which means that single top quarks come in a polarized state. The W helicity (spin + polarization) is then transferred directly to the top decay products which can be measured. The angles with which the decay products form relative to the top quark can also be used to identify single top quark production. Studies on polarization and angular correlations of the top quark can be seen in the following references [12, 13, 14]. These measured properties also add to the validation of the SM.

2.3.6 Beyond The Standard Model

We briefly talked about one BSM search in Section 2.3.4 (the existence of more than three generations of quarks). Let's discuss a few more that have a direct connection to

the single top quark. In the left Feynman diagram of Figure 2.7 we see an s channel process where the decay products are $\bar{t}b$ coming from the charged Higgs boson. This leading order (LO) diagram has not been observed. However, we can use the s channel single top process to search for this decay. The diagram on the right of Figure 2.7 presents a t channel process where either a u or c quark exchanges a gluon with an antiquark and then changes flavor to a t quark. These types of decays are called flavor changing neutral currents (FCNC). Currently FCNC are not allowed in the SM, but we can search for such a process through the single top's t channel process. If either or both of these processes existed, then it would lead to a larger cross section for both the s and t channels cross sections. For further reading on other BSM predictions please refer to [15]

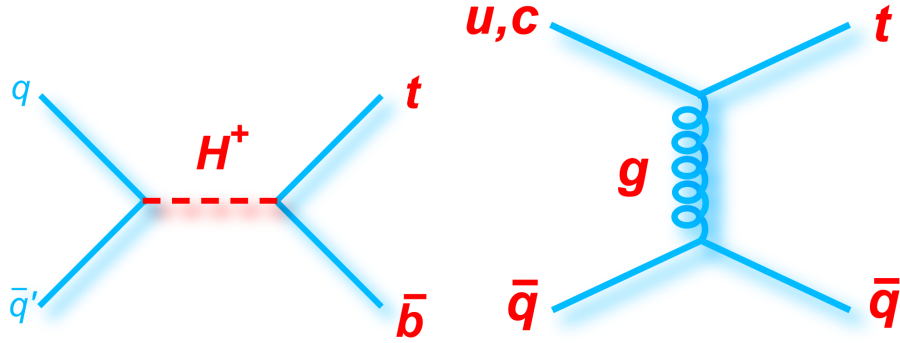


Figure 2.7: Feynman diagrams showing BSM processes through the s -channel (left) and t -channel (right) single top quark production.

Chapter 3

Accelerators and Detectors

3.1 Introduction

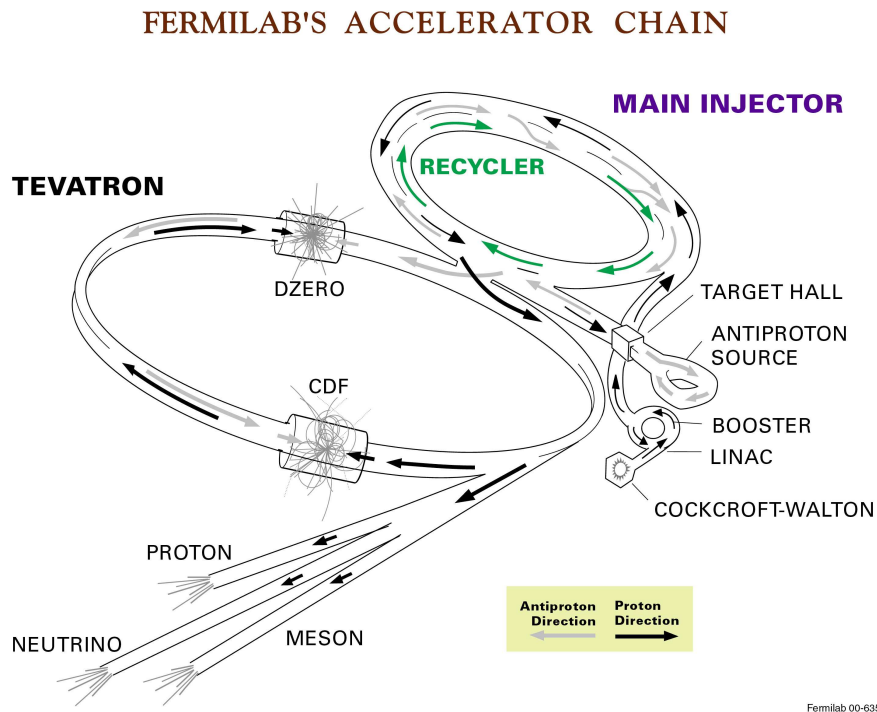


Figure 3.1: Diagram of Fermilab's chain of accelerators.

The Tevatron at Fermilab is the second largest synchrotron in the world. It accelerates protons and antiprotons up to ~ 1 TeV each and collides them at two detector sites ($D\bar{O}$ and CDF) located on opposite sides of the Tevatron ring. The resulting collisions occur at a center of mass energy of 1.96 TeV. The collision produces all SM particles excluding the Higgs boson. This chapter discusses the acceleration sequence used to reach collisions at this energy. This thesis is based on data taken from the detector site $D\bar{O}$.

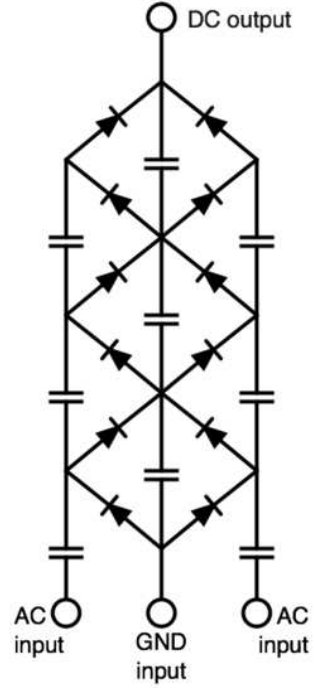
In order to get a collision to occur at the center of either the $D\bar{O}$ or CDF detectors we must first understand how the whole process unfolds. The following sections describe the chain of accelerators and storage rings used to produce $p\bar{p}$ collisions.

3.2 Cockroft-Walton Pre-accelerator

The first stage of acceleration begins with the Cockroft-Walton pre-accelerator. This consists of a dome which is mounted on three or four legs as there are two, however only one is used during acceleration. The legs act as a voltage multiplier (Figure 3.2). The multiplier takes low voltage AC power input (75 kV) and through the use of capacitors and diodes steps up the voltage to provide a DC output voltage of 750 kV. Located within the dome is a bottle of hydrogen gas (H_2) and a magnetron. The Hydrogen molecules get pumped into the magnetron which converts the gas to H^- ions (Figure 3.3). The hydrogen ions (H^-) are extracted by an extractor electrode and accelerated to 750 KeV through a tunnel which is connected to the dome and grounded at the wall. The H^- ions are then transferred to the “750 KeV line” which has an electrostatic chopper that chops the beam into bunches ($40\mu s$ each) before they enter the Linac.



(a)



(b)

Figure 3.2: (a) The Cockcroft-Walton Accelerator. (b) Voltage multiplier circuit diagram.

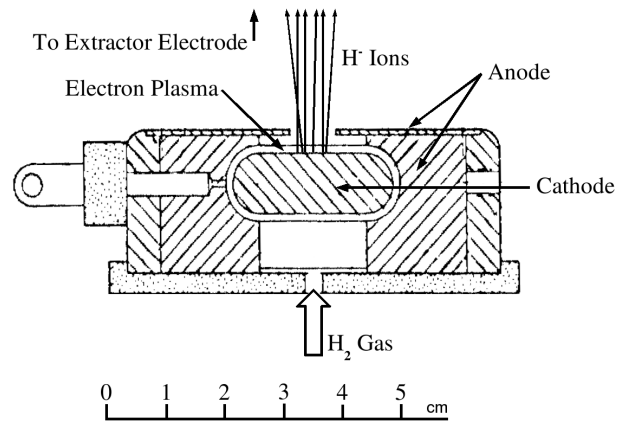


Figure 3.3: H^- ion production in the magnetron.

3.3 Linear Accelerator (Linac)

The Linac is divided into two parts, low energy drift tubes (DTL) and the high energy side coupled (SCL) linacs [16]. The DTL is 79 m long and consists of five tanks, each powered by a 5 MW amplifier which produces a 201 MHz RF signal. The DTL accelerates the H^- ion bunches from 750 KeV to 116 MeV. The SCL is 67 m long and consists of seven modules each powered by a 12 MW Klystron amplifier which creates a 805 MHz RF signal. This section takes the 116 MeV ion bunches and accelerates them to 400 MeV. The 400 MeV ion bunches from the Linac are sent to the Proton Booster ring for further acceleration.

3.4 Proton Booster Ring

This is the first synchrotron in the acceleration process [17]. It is a 75 m radius ring that uses 19 RF cavities to accelerate the beam from 400 MeV to 8 GeV. As the energy is increased the magnet current is synchronously increased to keep the beam on a circular orbit. At injection the 400 MeV ion bunches are passed through a carbon foil, which is used to strip off the electrons leaving only the protons to accelerate around the synchrotron. Once enough protons are accumulated at 8 GeV they are transferred to the Main Injector.

3.5 Fermilab Main Injector

The Main Injector is also a synchrotron and has a radius of ~ 528 m. The main injector is a very versatile accelerator that has many “Modes of Operation” [18]. One of the main modes performs the acceleration of both protons and antiprotons from 8 GeV to 150 GeV for injection into the Tevatron. For protons this is quite easy since

they are coming from the Booster. However, for antiprotons the process is involved as we first have to produce them. This defines a second mode of operation in which protons are accelerated from 8 GeV to 120 GeV and then sent to the antiproton source where they are used to produce antiprotons. The next section will cover the production of antiprotons.

3.6 Antiproton Production

Antiprotons are created by colliding 120 GeV protons from the Main Injector with a Nickel-Iron alloy target. This produces a jet of particles, some of which are antiprotons. A Lithium collection lens is then used to focus the negatively charged particles using a pulsed magnetic field. This is followed by a magnet which is used to momentum select the antiprotons that have an energy of about 8 GeV. Antiprotons are produced every 2.4 seconds through this process. A diagram of this process can be seen in Figure 3.4. The antiprotons are then transferred to the debuncher.

The Debuncher is a synchrotron that takes the shape of a rounded triangle with an average radius of 90 m. It is used to reduce the transverse momentum spread of the beam through stochastic cooling [19] and debunches the beam by using RF kicks to produce a continuous antiproton beam. The debuncher does no acceleration and it maintains the beam energy of 8 GeV.

From the debuncher the beam is transferred to the antiproton accumulator. This is very similar to the debuncher and resides in the same tunnel. It has an average radius of 75 m and is used to further cool the beam. Unlike the debuncher the beam stays in the accumulator for several hours and it is rebunched before being extracted to the last element of the antiproton production complex, the recycler.

The Recycler is a storage ring which resides just above the Main Injector. It is

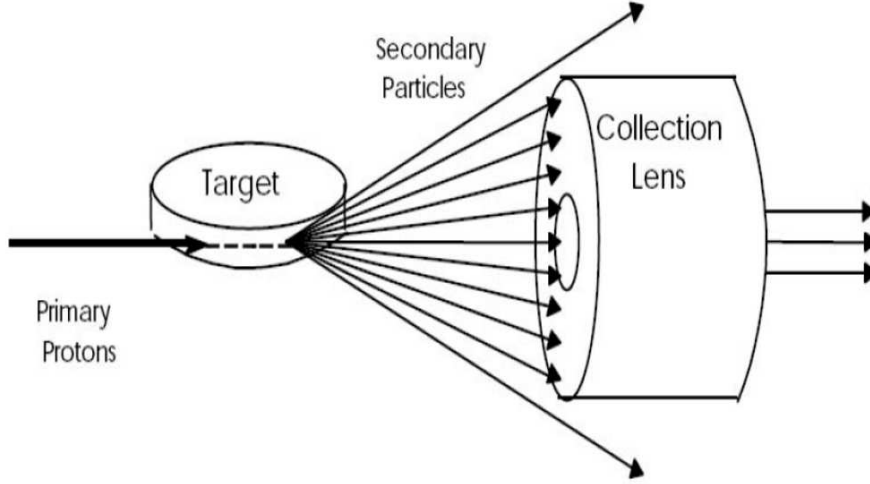


Figure 3.4: Schematic diagram of the Target and Lens system. [19].

used for the third phase of the beam cooling. It uses both stochastic and electron cooling [20]. Once the Recycler is sufficiently filled with antiprotons the 8 GeV beam is sent to the Main Injector for acceleration to 150 GeV.

3.7 Tevatron

Once there is a sufficient number of antiprotons the Main Injector injects the 150 GeV proton beam into the Tevatron in 36 bunches where each bunch is separated by 396 ns. In addition, after every 12th bunch there is an additional 2.64 μ s separation. Once the protons are loaded then the same thing is done with the antiprotons, the combined operation is referred to as a “store”. The Tevatron is the last synchrotron in the acceleration process. It is a ~ 4 mile ring which accelerates the beams from 150 GeV to 980 GeV each. It uses superconducting magnets and RF cavities to steer and accelerate the beams. The protons and antiprotons travel in opposite directions in

the same tunnel by rotating around one another in a helical path. At two locations there are two detectors ($D\bar{O}$ and CDF) where the beams are brought into collision. This gives $p\bar{p}$ collisions with a center of mass energy of 1.96 TeV.

3.8 $D\bar{O}$ Detector

The initial design for the $D\bar{O}$ detector was the merger of two somewhat different proposals, a non magnetic detector with precision calorimetry and a detector with extensive muon detection over a large range of rapidity. This is the basis of the Run I $D\bar{O}$ detector which was operated from 1992 until 1996 [22]. Between 1996 and the start of Run II in 2001, the detector was upgraded with the addition of a central magnetic field, a new central tracking system and a new forward muon spectrometer [23]. The present detector has four main detection systems: a central tracking system, a preshower system, central and forward calorimeters and central and forward muon spectrometers. Each of the sub-systems is described in the following sections.

3.8.1 $D\bar{O}$ Coordinate System

$D\bar{O}$ uses a right-handed coordinate system centered at the middle of the $D\bar{O}$ detector. The z -axis points along the path of the proton, the x -axis points outward along the radius of the Tevatron and the y -axis points upward (Figure 3.6). Thus the position of a particle originating from a collision at $x = y = z = 0$ can be defined in terms of its distance from the origin, $r = \sqrt{x^2 + y^2}$, and two angles θ and ϕ . θ is defined as the polar angle with respect to $+z$ -axis and ϕ is the azimuthal angle, defined with respect to the $+x$ -axis. Rather than using θ , it is often more convenient to use the

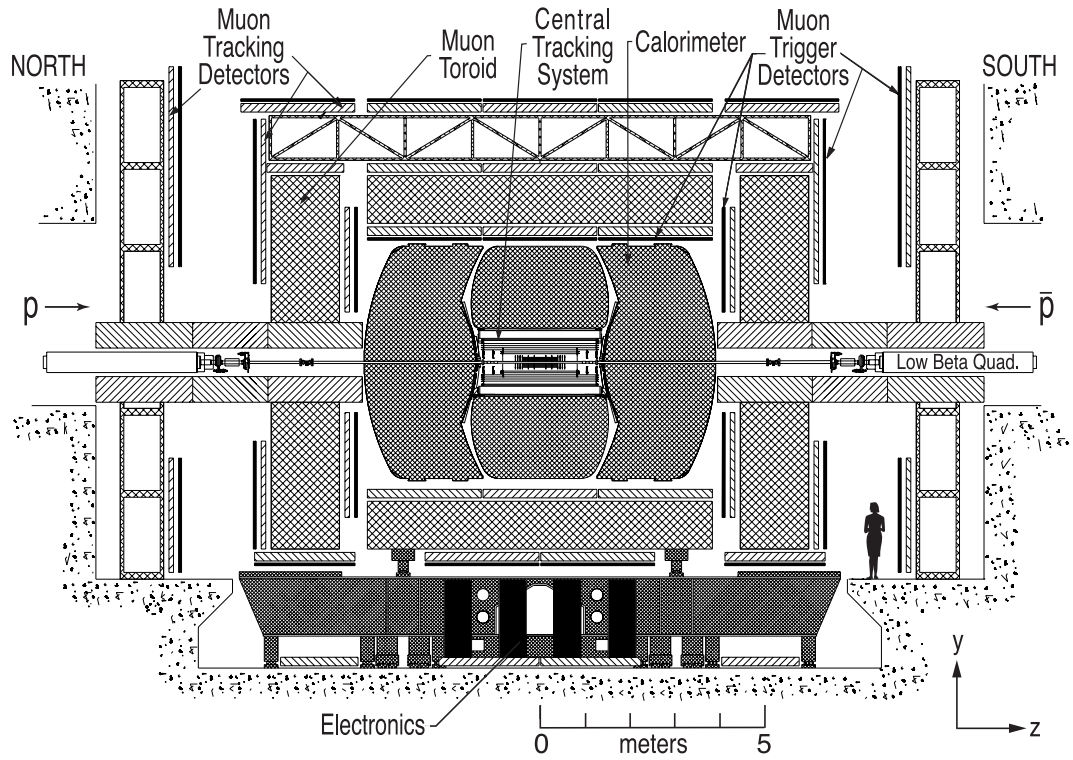


Figure 3.5: Schematic of the experimental hall and the $D\bar{O}$ detector. The location of each of the major subsystems is shown, along with the position of the on-detector readout electronics.

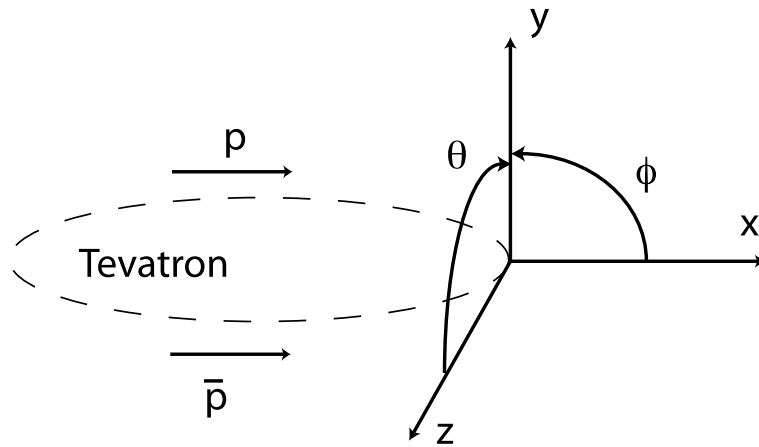


Figure 3.6: $D\bar{O}$ coordinate system.

pseudorapidity, η , which is defined by:

$$\eta \equiv -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.1)$$

This is an approximation of the true rapidity, y , in the relativistic limit $m \ll E$, which is valid in almost all of the cases considered here. The rapidity of a particle is Lorentz invariant and is given by:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (3.2)$$

The pseudorapidity is approximately Lorentz invariant and is more convenient variable than y , because the particle density per unit η is much more uniform than that of y .

In practice there is a finite spread in the z -position of the collision point, so that collisions often do not occur exactly at $x = y = z = 0$. In this case the pseudorapidity is defined with respect to the collision vertex.

3.8.2 Central Tracking System

Charged particle detection and measurement is performed by the central tracking system. For $D\bar{O}$ this has two components, the silicon microstrip tracker (SMT) and the central fiber tracker (CFT). The trackers are located inside of a solenoidal magnet which produces a 2 Tesla magnetic field (Figure 3.7). The combined tracker-magnet system is used to locate and identify particles and measure their momentum. It also locates the position of the collision or primary interaction. The SMT and CFT trackers locate the position of particles as they pass through them and the bend of the track in the magnetic field provides a measurement of the particle momentum. Once this is done we can then use sophisticated algorithms to identify and trace back

the particles point of origin.

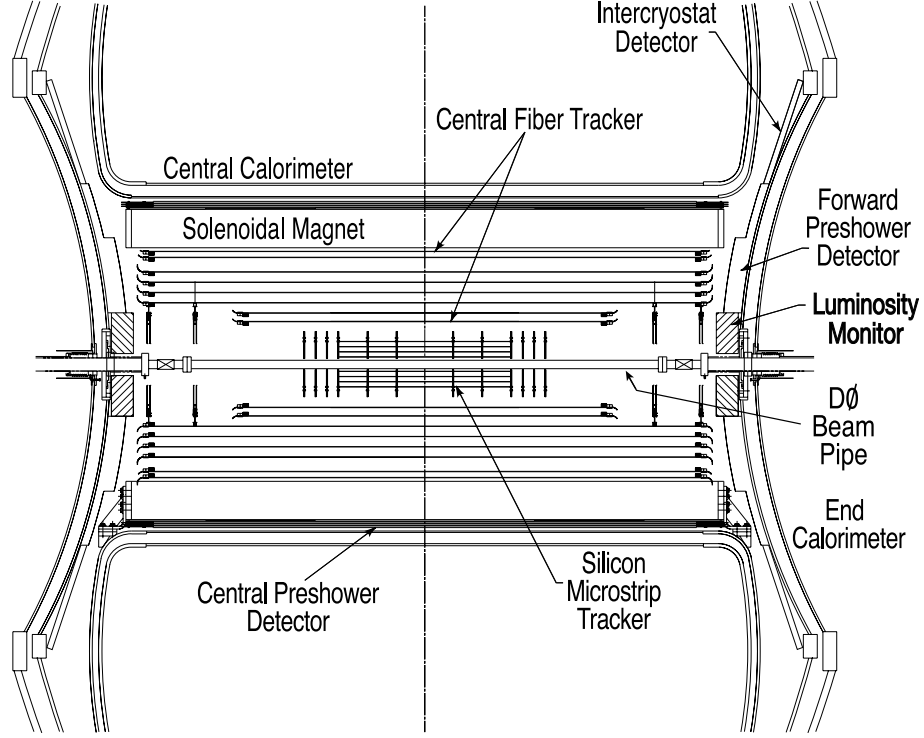


Figure 3.7: Schematic of the central tracking showing the location of SMT and CFT components with respect to the beam axis.

The SMT is used as the first line of detection and is therefore positioned at the center of the $DØ$ detector. As shown in Figure 3.8, it is comprised of six barrel sections, and sixteen transversely mounted disks. The barrel sections provide measurement of the coordinates r and ϕ and the disks measure the coordinates r , z and ϕ . The detector elements are formed from silicon wafers with readout strips etched onto either one or both sides of the wafer. Charged particles passing through the bulk of the silicon create electron-hole pairs which result in an induced pulse in the strips on the wafer surfaces. The strips on the top and bottom surfaces have a finite crossing angle, thereby providing a spatial measurement of the track position. The SMT covers the pseudorapidity range $|\eta| < 3.8$.

The CFT detector is composed of eight concentric cylinders which surround the

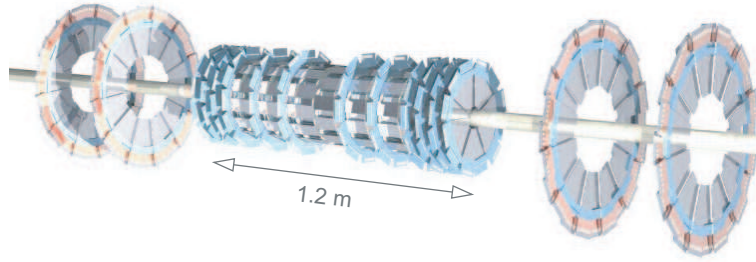


Figure 3.8: The elements of the SMT detector, showing the barrel and disc segmentation.

SMT. It has a regional coverage of $r = 20\text{-}52$ cm and $|\eta| = 1.7$. Each cylinder consists of two doublet layers of scintillating fibers. The first layer lies along the z -axis and the second has a stereo angle of $\phi = \pm 3^\circ$ relative to the z -axis. The scintillating fiber, when excited by a charged particle, gives off light which then travels down a wave guide to be readout by a visible light photon counter (VLPC). The CFT has a resolution of $100\text{ }\mu\text{m}$.

The $D\bar{O}$ solenoid is a superconducting magnet made from niobium-titanium wire which carries a current of 4.7 kA and produces a 2 T magnetic field that lies parallel to the z -axis. The magnet is used to bend charged particles so that their momentum can be measured from the track curvature. In the transverse plane, this can be expressed in the form:

$$p_T = 0.3qBr. \quad (3.3)$$

Where, q is the charge of the particle, B is the strength of the magnetic field and r is the radius of curvature. A schematic of the solenoid and the muon toroid magnets is shown in Figure 3.9, together with a representation of the the resulting field pattern inside of the detector.

The preshower detector is broken into two parts, the central preshower detector (CPS) and the forward preshower detector (FPS). These are located between the

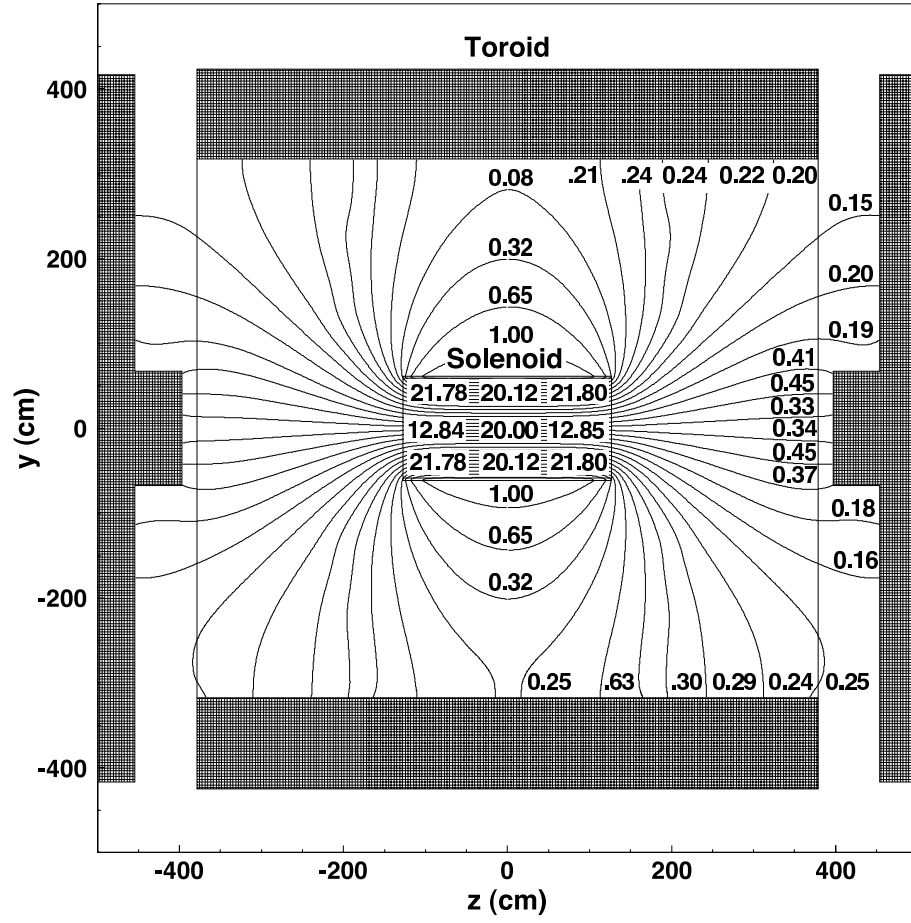


Figure 3.9: Schematic representation of the $D\bar{O}$ magnets and the magnetic field within the detector volume.

solenoid and the calorimeters. The CPS has a regional coverage of $|\eta| < 1.3$ at a radius of 72 cm and the FPS covers $1.5 < |\eta| < 2.5$. The CPS is composed of a lead radiator followed by three layers of scintillating fibers. The FPS consists of two layers of scintillating fibers, a lead radiator then another two layers of scintillating fibers. The location of the detectors is shown in Figure 3.7. The preshower detector helps in the identification of electrons and photons which begin to shower when hitting the radiator. It also aids in the identification of heavier particles which will leave only single tracks.

3.8.3 $D\bar{O}$ Calorimeters

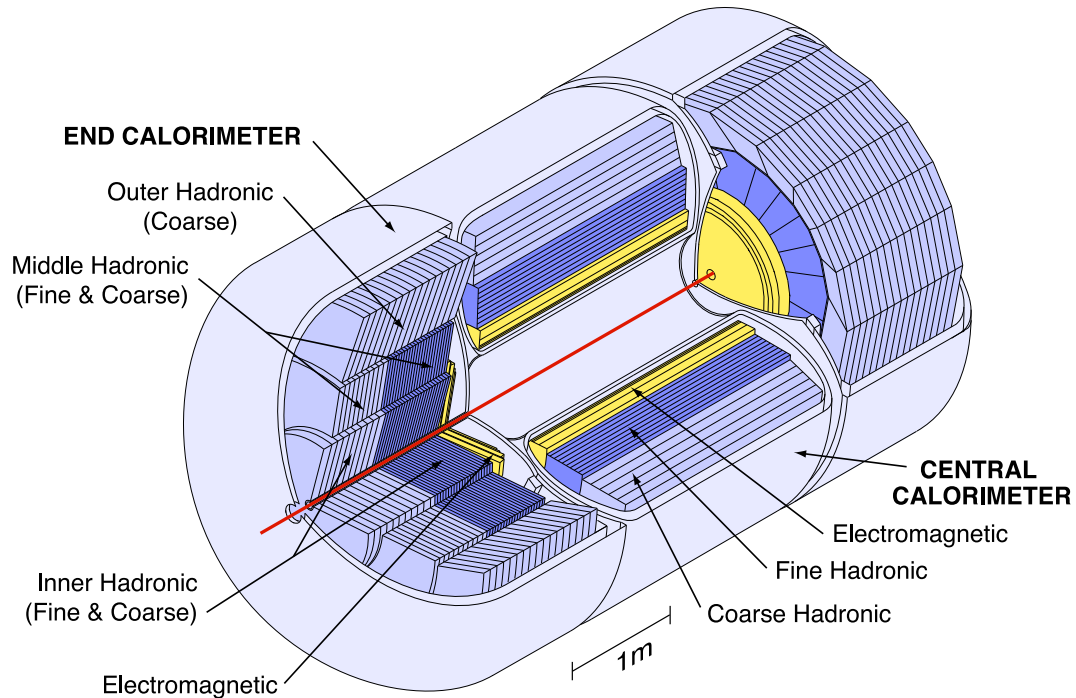


Figure 3.10: Schematic diagram shown the $D\bar{O}$ central endcap calorimeters. The cut away shows the details of the calorimeter modules [26].

The $D\bar{O}$ calorimetry is divided into three sections, the central module (CC) and the two endcap modules (ECN, ECS) (Figure 3.10). It is designed to measure the en-

ergy of electrons, muons, photons, jets and can be used to infer any missing transvers energy \cancel{E}_T through the conservation of momentum. The CC covers a region of $|\eta| < 1.1$ where the two ECs cover $1.2 < |\eta| < 4.2$. Each module consists of layers which are designated the electromagnetic, the fine hadronic and coarse hadronic layers. The electromagnetic layers are used to measure showers coming from electrons and photons (Figure 3.11) and the two hadronic layers are used to measure the showers from pions, Kaons, protons and muons.

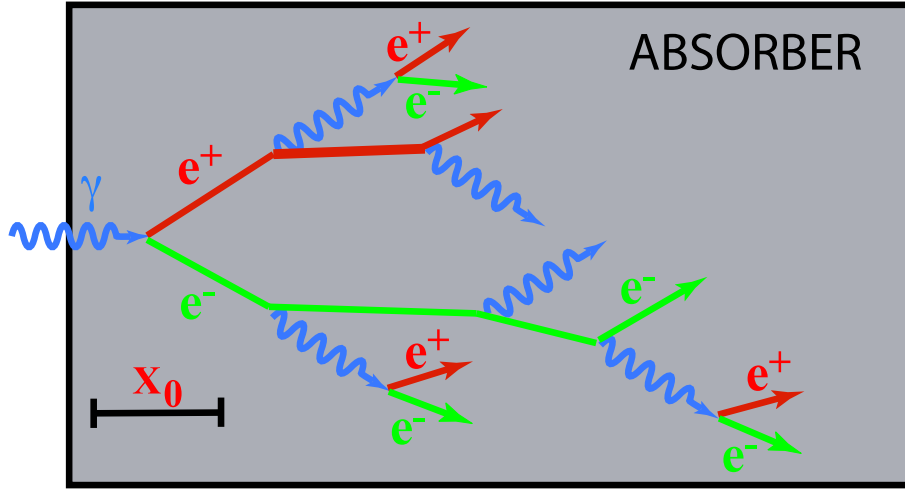


Figure 3.11: Depiction of a photon shower where X_0 represents one radiation length.

Energy measurements are made by summing the calorimeter energy depositions along the track directions. The calorimeter has a projective tower topology (Figure 3.13) which is readily adapted to this process. Each layer in the calorimeter has different absorber materials associated with them. The active medium in all of the calorimetry is liquid argon. The absorber varies between the module types and is depleted uranium for the electromagnetic and fine hadronic modules. In the outer (coarse hadronic) modules this is replaced by copper because of cost considerations.

Figure 3.12 shows the structure of a unit cell within a calorimeter module. This

works by applying a high 2.0-2.5 kV positive potential to the outer surfaces of the readout boards while grounding the absorber plates, thus creating an electric field between the two. Each readout board consists of two layers with copper pads etched into the middle of the board and its exterior surface is covered with a high resistive coating which forms the high voltage electrodes. Charged particles emanating from a hadronic or electromagnetic shower leave the absorber plates and ionize the liquid argon. The ionization drifts to the anode readout boards and is capacitively coupled to the Cu pads. The current measured is proportional to the energy deposited by the particle. The readout cells are combined into projective towers (Figure 3.13) using a cell size of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ apart from the third electromagnetic layer (shower max) for which this is reduced to 0.05×0.05 .

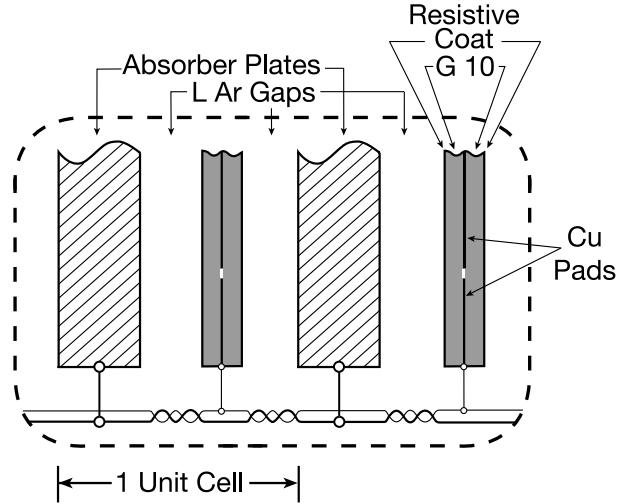


Figure 3.12: Structure of a unit cell in the liquid Ar calorimeter modules.

Calorimeter noise is the collective term used to describe anything that can produce a false signal in the readout. In the $D\bar{O}$ case, this is predominantly from low level signals from uranium decays and electronic noise. It is essential to eliminate this noise before reconstructing physics objects (Chapter 4) so as not to degrade the energy resolution. Cells which contain this noise are labelled as “hot cells”. $D\bar{O}$

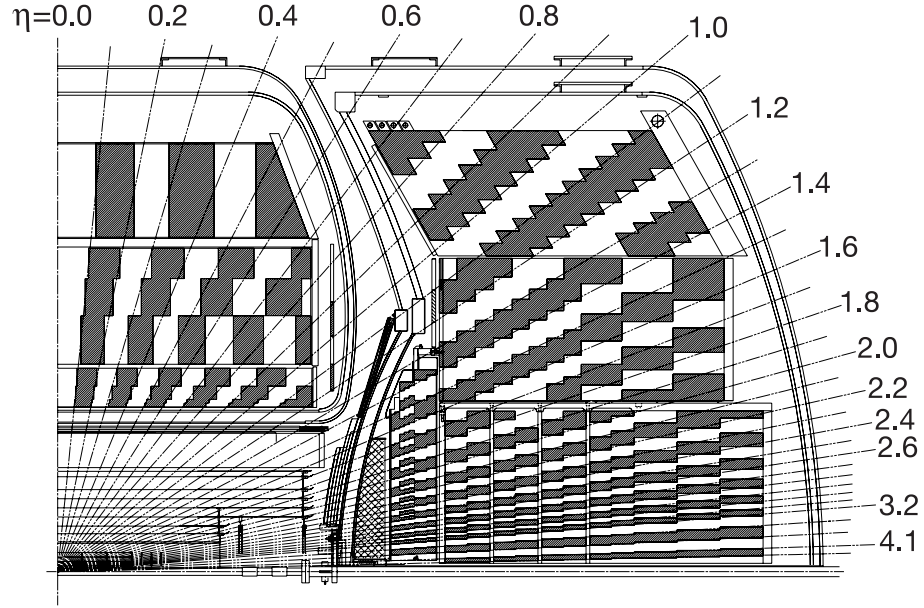


Figure 3.13: Schematic of one octant of the $DØ$ central and endcap calorimeters. The shaded regions show the projective structure of the calorimeter towers and their map onto pseudorapidity.

uses two algorithms to identify and remove them: the New Anomalous Deposits Algorithm (NADA [24]) and the T42 algorithm [25].

Each cell's energy is checked for noise (σ_{cell}) when the accelerator is not running. Both algorithms use this baseline measurement to identify potentially hot cells before reconstruction takes place. As an example the T42 algorithm will first identify cells that have $E_{cell} > 4\sigma$ as well neighbouring cell that have $E_{cell} > 2.5\sigma$, otherwise the rest of the cells are removed. The NADA algorithm works in a similar fashion.

3.8.4 Muon Detectors

Muons leave only a minimum ionizing deposition of energy in the tracking and calorimeter systems, hence the need for a muon detection system. The $DØ$ muon detector [27] consists of three elements: proportional drift tubes (PDT); mini drift

tubes (MDT); an array of scintillation counters and three 1.8 T toroidal magnets. The central part of the muon system covers the region $|\eta| < 1$ and the forward part covers $1 < |\eta| < 2$. The muon spectrometer is divided into the A, B and C layers where the toroidal magnets lie between layers A and B. An expanded view of the chambers and scintillation counters is shown in Figure 3.14.

The PDTs and MDTs measure the position and momentum of the muons passing through them, with the PDTs in the barrel (central) part of the detector, and the MDTs the forward regions. Both types of drift tubes have a similar construction. The two sides of each drift tube have a negative voltage applied to them, while the wire running down the middle of the tube that has a positive voltage. The drift tubes are filled with a inert gas mixture that gets ionized when muons pass through. The electrons that drift toward the anode wire and produce a signal pulse on the wire. The ionization takes a finite (drift) time to reach the anode. For the PDTs, this is 500 ns, and for the MDTs it is 100 ns. The scintillation counters work much the same way as the scintillating fibers and have a response time of 1.6 ns that allows for precise timing measurements. Because of this they are able to correlate muon hits with bunch crossings and can also determine if a muon came from a $p\bar{p}$ collision or from outside of the detector (Cosmic ray Muons). The toridal magnets are used to bend muons in the 1.8 T field which allows for more precise measurements of momentum.

3.8.5 Luminosity Monitor

The $D\bar{O}$ luminosity monitors (LM) consists of two circular detectors made of 24 scintillating tiles each. The detectors are located at ± 144 cm along the z -axis and has a regional coverage of $2.7 < |\eta| < 4.4$ (Figure 3.15). The luminosity is measured

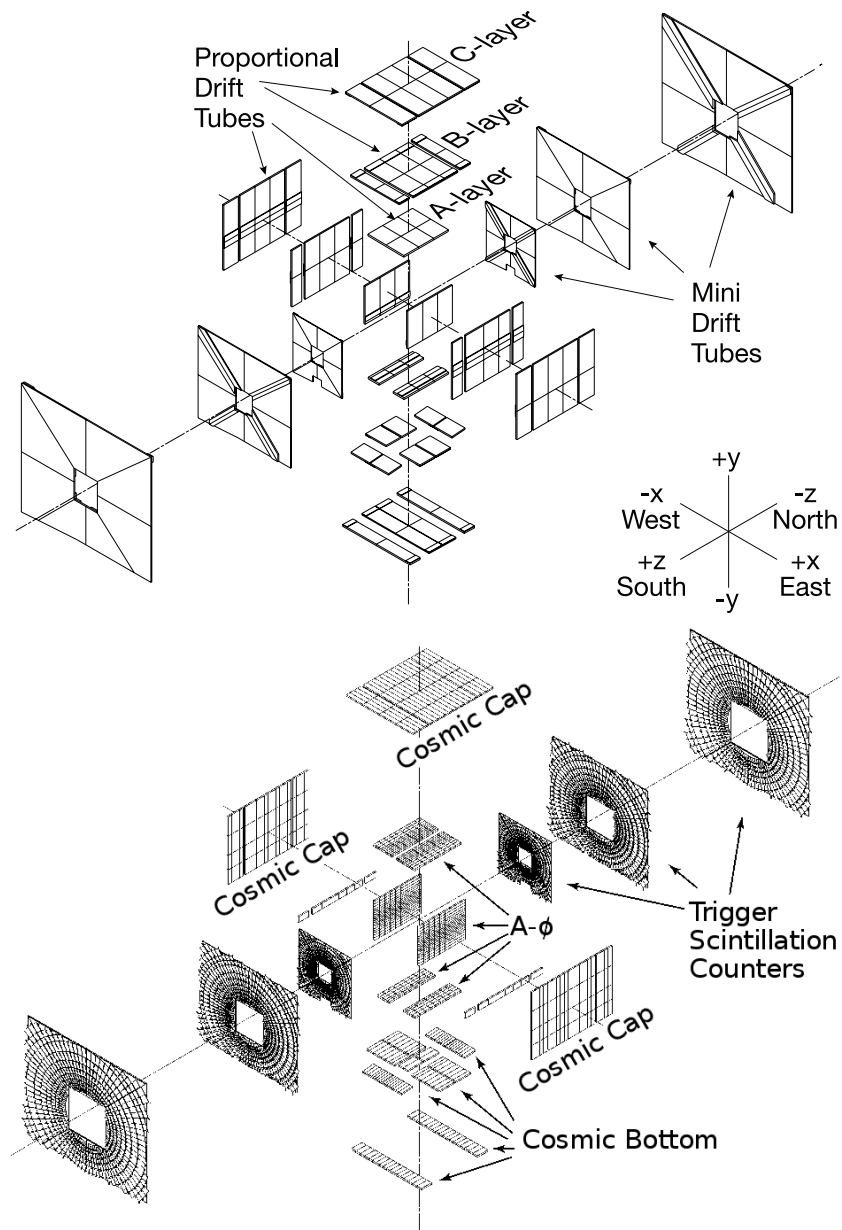


Figure 3.14: Exploded view of the $DØ$ muon detection system. The top diagram shows the configuration of the A, B and C layers in the barrel and forward regions. The bottom diagram shows the location of the barrel and forward scintillation counters.

by the following equation:

$$L = \frac{f \bar{N}_{LM}}{\sigma_{LM}}. \quad (3.4)$$

Where, f is the frequency of the crossing beam, \bar{N}_{LM} is the average number of inelastic collisions and σ_{LM} is the effective cross section [26]. The LM is also used to discriminate between $p\bar{p}$ interactions and the beam and it can also be used to make a quick measurement of the position of the interaction vertex along the z -axis.

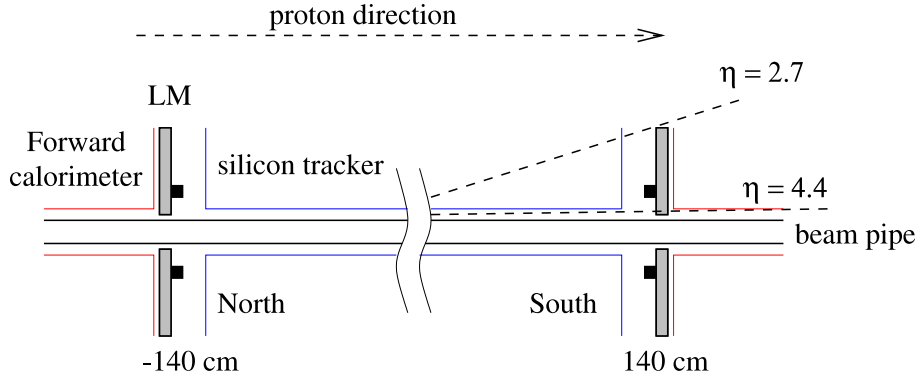


Figure 3.15: Diagram of the Luminosity monitors at $D\bar{O}$ [26].

3.9 Trigger Framework

The job of the trigger framework is set up to analyze the data coming from all the detection systems that comprise the $D\bar{O}$ detector and decide which of the events should be recorded for analysis. The raw event rate of 1.7 MHz would result in the need to record more than one petabyte of data per second, which is far too large to handle. The practical limit is a trigger rate of ~ 100 Hz for recording to tape, requiring a rate reduction of $\sim 2 \times 10^4$. This is done using the three trigger level $D\bar{O}$ framework (Figure 3.16).

The lowest level trigger (L1) consists of firmware and hardware and reduces the event rate of 1.7 MHz to 2 KHz. This needs to make decisions very quickly ($3.5 \mu s$)

and it only takes in signal from the subset of the detector systems with very fast readouts (Figure 3.17).

The second level trigger (L2) is more sophisticated and it uses a combination of firmware and microprocessors to make decisions. It has the capability of correlating the data between the detector systems and is used to reduce the trigger rate to 1 kHz in around $100 \mu\text{s}$.

The highest level trigger (L3) is comprised of over 100 Linux computers which take in the 1 kHz event rate from L2 and reduces it to around 100 Hz. The L3 trigger uses sophisticated algorithms to quickly reconstruct events to see if they are of interest then stores them for offline reconstruction.

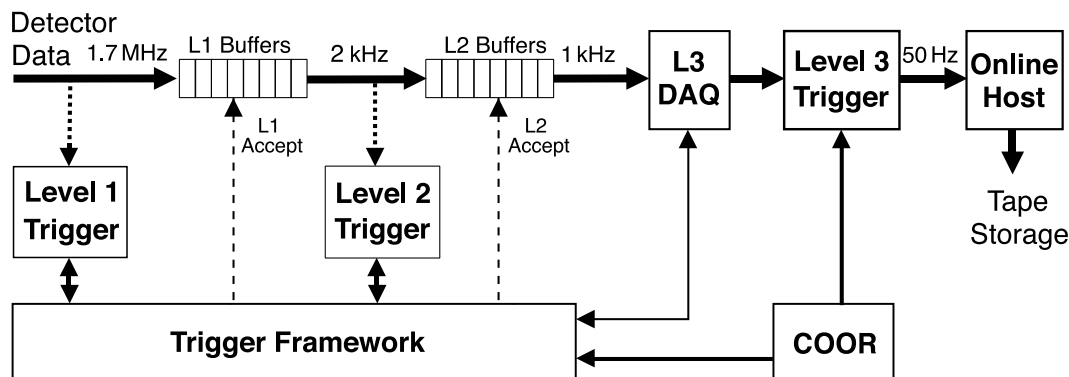


Figure 3.16: Block diagram of the $D\bar{O}$ framework, showing the data-flow through the three levels of the trigger.

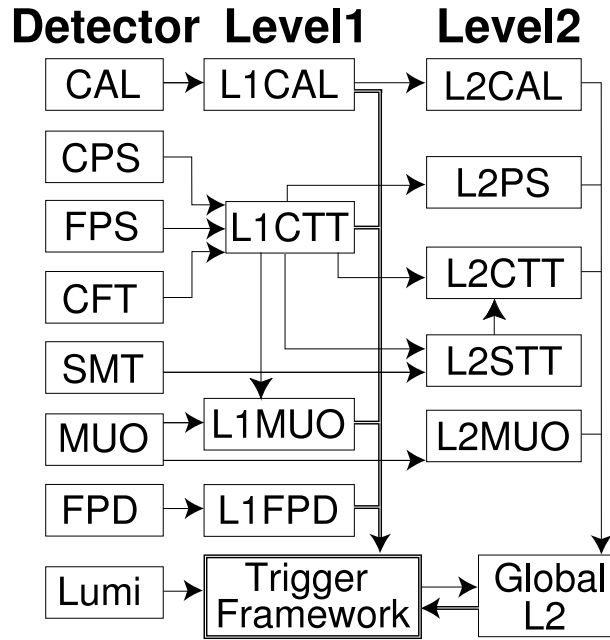


Figure 3.17: Block diagram of L1 and L2 showing what systems they are connected to [26]. All detectors are listed on the left where the calorimeter (CAL) is initially connected to the level 1 trigger, the central tracking trigger (CTT) is defined as both the preshower detectors (CPS and FPS) and the fiber tracker (CFT) and is also connected to the L1 trigger. The SMT goes straight to L2 where the muon system begins at L1. The forward proton detector (FPD) and the Luminosity monitor are also connected to L1 initially.

Chapter 4

Event Reconstruction

When a $p\bar{p}$ inelastic collision occurs at $D\bar{O}$ there are many particles produced that travel through the detector and are measured by the tracking system and the calorimeters. It is the job of the reconstruction code to convert the raw signals (pulse-heights and positions) into reconstructed charged and neutral particles. These in turn are grouped into jets and projected to their point of origin, to define production vertices. For this analysis, the principal objects of interest are electrons, muons, jets (light and heavy quark), the position of primary and secondary vertices, and the missing energy (\cancel{E}_T). To reconstruct and identify physics objects there are certain requirements that must be followed, which are discussed in the following sections.

4.1 Charged Track Reconstruction

When charged particles traverse the SMT and CFT they deposit only a small amount of energy via $\frac{dE}{dx}$ and their presence is identified by a pulse or hit in a specific part of the detector readout. The reconstruction code takes these hits and builds them into tracks emanating from a vertex or point of production for a particular event. Since

both the SMT and CFT are inside a strong magnetic field the extrapolated tracks are curved. For any particular hard interaction there will be many hits corresponding to many tracks making the process of track finding complicated. In order to solve this problem $DØ$ uses two different algorithms to define tracks. These are the histogram track finding (HTF) [28] and the alternative AA [29] algorithms. To define the final set of tracks in an event these algorithms are combined using global track reconstruction (GTR). GTR relies on a Kalman Filter algorithm [30] which cleans and smooths the tracks before they are stored for analysis.

The HTF method [28] uses all possible hits in (x, y) space and maps them to points in (ρ, ϕ) space while taking account of the point of interaction. Here, ρ is the curvature of the track and ϕ is the direction of the track. Using these two parameters 2D histograms are built allowing the identification of tracks. If there is a particle corresponding to a collection of hits in the tracking volume then those hits should all have the same curvature and direction. These appear as peaks in the 2D histograms thus allowing the identification of tracks. In contrast, the hits associated with other tracks will be spread uniformly throughout the histogram. For a detailed description of the algorithm please see Ref. [28].

The AA algorithm [29] begins with a hypothesis that a track must have at least three hits in the SMT in order to be reconstructed. All three hits must occur in succession where the first hit can be in either an SMT barrel or F-disk. The second hit must form an angle with the first hit and relative to the beam spot where $\Delta\phi < 0.08$ (Figure 4.1). The third hit is required to have a “radiusMin” > 30 cm and an “impactMax” < 2.5 cm (Figure 4.1), where radiusMin is the radius of curvature of the track and impactMax is the axial impact parameter with respect to $x = y = z = 0$. Also, all three hits are required to have a fit of $\chi^2 < 16$. If all these requirements are met the algorithm extrapolates outwards to the following tracking layers searching for

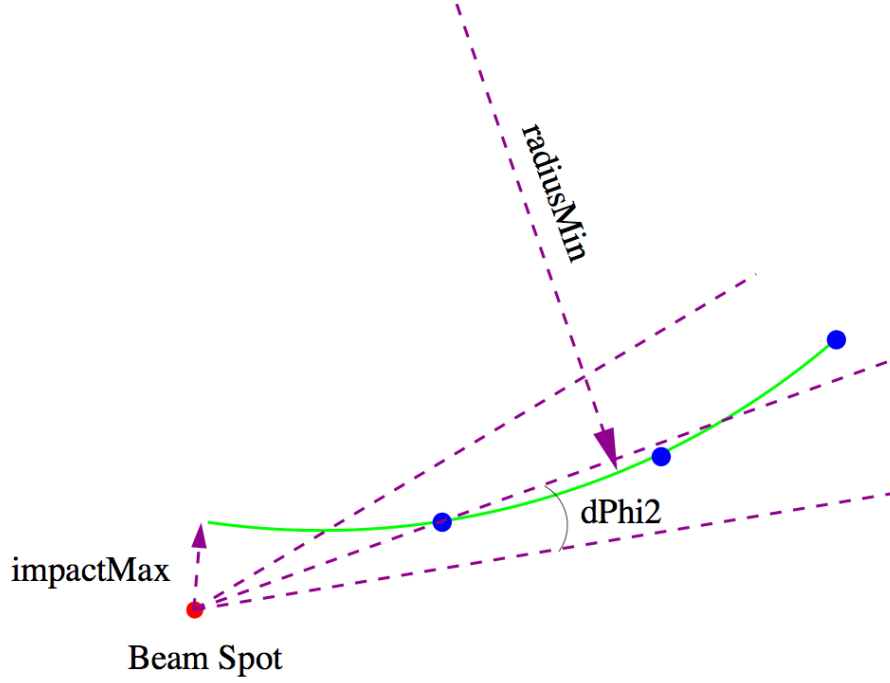


Figure 4.1: Diagram of track reconstruction using the AA algorithm.

more hits associated with that track. As the number of hits grows the track fit is still required to have a $\chi^2 < 16$. If there are many hits on one layer then the algorithm creates a new hypothesis. The whole process continues until the detector ends or three successive layers are missed. For a detailed discussion please see Ref. [29].

Tracks from both the HTF and AA are then analyzed using global track reconstruction (GTR). GTR uses a Kalman Filter to refit, clean and smooth the tracks resulting in a final set of tracks.

4.2 Primary Vertex

The primary vertex is defined as the location of the hard inelastic collision between the proton and antiproton. The process of locating the interaction point is complicated by the presence of secondary vertices, such as those from b quark decays, and the

presence of multiple interactions when running at high luminosity.

To deal with these complications $D\bar{O}$ uses an adaptive primary vertex algorithm [31] to identify the position of the primary vertex. This is done in three steps: track selection, vertex fitting, and vertex selection.

Tracks are selected by requiring at least two hits in the SMT and a $p_T > 0.5$. The selected tracks are then clustered relative to the z -axis and must be within 2 cm of each other in z . Vertex fitting then uses a 2-pass process, where the first pass fits the clustered tracks to a common vertex, and the second pass chooses the clusters that are consistent with the position of the beam spot. The clustering is done using a Kalman Filter which removes tracks with a high χ^2 until a $\chi^2/ndf < 10$ is achieved for the cluster. In the second pass only the clusters which lie within 5σ of beam spot are retained. Lastly, the primary vertex is chosen from the remaining candidates by selecting the vertex which has the minimum probability of being a minimum bias interaction [32].

4.3 Electron Identification

The signature of an electron in the $D\bar{O}$ detector is a track in the central tracker which matches to an electromagnetic shower in calorimeters. Electromagnetic cluster candidates are identified from the energy deposited in the calorimeter towers using a basic cone algorithm [33] with the requirement that almost all of the energy is found in the EM calorimeter. To be considered an electron candidate, a cluster must have a track from the central tracker which spatially matches to the cluster. The criteria used for both steps are discussed below.

Cluster candidates are identified using three criteria for the energy profile within the calorimeter. These are the cluster electromagnetic fraction, f_{EM} , its isolation,

f_{iso} , and its H-matrix χ^2 , χ_H^2 . These are defined below. The EM fraction is defined by:

$$f_{EM} = \frac{E_{EM}}{E_{total}}. \quad (4.1)$$

Where, E_{EM} is the amount of electromagnetic energy in the cluster and E_{total} is the total cluster energy. A cone size $\Delta R = 0.2$, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, is used for energy clustering and clusters are only retained if they have an electromagnetic fraction, $f_{EM} > 0.9$. The next step is the cluster isolation requirement, which is defined using the equation:

$$f_{iso} = \frac{E_{total}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)}. \quad (4.2)$$

Where, $E_{total}(R < 0.4)$ is the energy observed in a cone of size $\Delta R = 0.4$, centered on the electromagnetic cluster and $E_{EM}(R < 0.2)$ is the cluster energy. For an isolated electron, f_{iso} should peak close to zero. The requirements used in this analysis are listed in Table 4.1, below.

The last requirement for the cluster selection is a cut on transverse and depth profiles of the electromagnetic shower in the calorimeter. This is done using the H-matrix χ^2 , χ_H^2 [34], which compares the observed profiles to those expected for an electron on the same energy. This uses seven shower shape variables to define a 7×7 covariance matrix. The matrix is then used to calculate a χ_H^2 by comparing the observed and predicted energy depositions. The seven inputs to χ_H^2 are: the fractions of the energy in each of the four layers of the electromagnetic calorimeter, the transverse width of the shower in the fourth layer (position of shower max), $\log_{10}(E_{EM})$ and the z position of the primary vertex. To be retained, a cluster is required to have $\chi_H^2 < 50$.

At this point, the cluster list contains both electron and photon candidates. The last part of the selection is to require a match between the cluster and a track from the central tracker. This is done using a cut on the probability of the matching χ^2 , $P(\chi^2)$. This contains contribution from both the spatial matching and the energy-momentum matching between the cluster and the track.

An important factor for this analysis is the ability to simulate the background from QCD events in which the electron signature is faked due to jet fluctuations or charge/neutral pion overlaps. As this is very difficult to simulate reliably with a monte carlo, this is done using a subset of the data. For this we define one last measure of electron quality, the electron likelihood,

$$\mathcal{L}_e(\vec{x}) = \frac{P_S(\vec{x})}{P_S(\vec{x}) + P_B(\vec{x})}, \quad (4.3)$$

where the vector \vec{x} has seven dimensions and the subscripts S and B denote signal (genuine electrons) and background (fake electrons). The dimension of \vec{x} are: f_{EM} , $P(\chi^2_{spatialmatch})$, track-cluster $\frac{E}{p}$ ratio, χ_H^2 , track distance of closest approach to the primary vertex, $\Sigma_{\Delta R} = 0.4$ (track p_T), $\Sigma_{\Delta R} = 0.05$ (no tracks) [35]. The cones for the last two of these are defined with respect to the cluster axis. For the purpose of later clarity it is convenient at this point to define three classes of electron candidates, which are denoted ultraloose, loose and tight. The ultraloose category is used to define the misidentification background from QCD and the loose and tight categories are used in the selection of signal events (see later). These are defined in Table 4.1 below.

Quantity	Ultraloose	Loose	Tight
f_{EM}	> 0.9	> 0.9	> 0.9
χ_H^2	< 50	< 50	< 50
f_{iso}	< 0.2	< 0.2	< 0.2
p_T	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$
$P(\chi^2)$	NA	> 0	> 0
Track match p_T	NA	$> 5 \text{ GeV}$	$> 5 \text{ GeV}$
Tracks match with PV	NA	$\Delta z < 1 \text{ CM}$	$\Delta z < 1 \text{ CM}$
\mathcal{L}_e	NA	NA	> 0.85

Table 4.1: Definitions for three classes of electron candidates.

4.4 Muon Reconstruction

If a muon track has a momentum $\geq 3.5 \text{ GeV}$, it will be seen by the tracking system, the calorimeters and the muon detectors. In this analysis, muon candidates are required to have hits in all three layers of the muon system. The B and C layer chambers see no magnetic field so that the BC segment of the track is found as a straight line. To define a muon track candidate, the extrapolation of a segment found in the B and C layers is required to match to a segment in the A layer chambers and to a track in the central tracker. For this analysis two categories of muon candidates are used. These are based on the extent to which the track is isolated and both are used in the selection of signal events (see later).

Loosely isolated muons candidates are required to have at least two hits in both the A layer and combined B+C layer chambers. To reject the background from cosmic ray muons, candidates are also required to have hits in the A and B+C layer muon scintillators which are within 10 ns of the bunch crossing time. They must also have a match to a track in the central tracker with a $\chi^2 < 4$ if the track has a $dca < 0.2 \text{ cm}$ with respect to the primary vertex or a $dca < 0.02 \text{ cm}$ for hits in the SMT. To be retained for further analysis, muon candidates must have $p_T > 4 \text{ GeV}$ and have

$\Delta R(\mu, \text{nearest jet}) > 0.5$.

A second category, designated as tightly isolated muons, must satisfy the loose isolation requirements and cuts on two additional quantities which look at the p_T -normalized energy and momentum flow about the muon track. These are the scaled track halo [36] and the scaled calorimeter halo [36] cuts, which are defined as:

$$\text{ScaledTrackHalo} = \left| \sum_{\text{Tracks}} \frac{p_T}{p_T(\mu)} \right|, \text{ for } \Delta R(\text{track}, \text{muontrack}) < 0.5, \quad (4.4)$$

$$\text{ScaledCalorimeterHalo} = \left| \sum_{\text{cells}} \frac{E_T}{p_T(\mu)} \right|, \text{ for } 0.1 < \Delta R(\text{cells}, \text{muoncal-track}) < 0.4. \quad (4.5)$$

In equation 4.4 the requirement is that the total transverse momentum of tracks surrounding the muon track in a cone of $\Delta R < 0.5$ must be less than 20% of the muon p_T . Same idea is used in equation 4.5 however, here the sum of the energy excludes the region $\Delta R < 0.1$, which is occupied by the muon $\frac{dE}{dx}$ deposition.

4.5 Jet Reconstruction

In a hadron collider, jets are frequently produced and can be defined as an associated group of particles which come from the hadronization of a quark or gluon. Since a typical jet contains both charged and neutral particles, full jet reconstruction requires the information from both the hadronic and electromagnetic calorimeters. This is done by clustering the energy deposited within a cone in ΔR space. The resulting clustered energy is then associated with a hadronic jet. This process is discussed in the following paragraphs.

$D\mathcal{O}$ uses the Run II Improved Legacy Cone Algorithm (ILC) [38, 39] for jet reconstruction. This algorithm is implemented in three steps. The first step takes

a number of calorimeter towers with $p_T > 0.5 \text{ GeV}$ and forms pre-clusters. To accomplish this, the algorithm starts from a p_T -ordered list of towers and associates the towers with $p_T > 1 \text{ MeV}$ that lie within $\Delta R = 0.3$ to form pre-clusters. The associated towers are then removed from the list and the procedure is repeated until all of the towers have been pre-clustered.

The next step is to form proto-jets. This starts from a p_T -ordered list of the pre-clusters. Starting from the highest energy pre-cluster, the pre-clusters are associated within $\Delta R = 0.5$ to form a proto-jet. The jet axis is then re-determined and the process is iterated until either the jet energy converges or the p_T of the jet falls below 4 GeV. All proto-jets with $p_T < 4 \text{ GeV}$ are discarded, as are any pre-clusters that lie within $R = 0.25$ of a stable proto-jet. The algorithm iterates the association until all of the stable proto-jets have been identified.

The final step in the reconstruction process involves the merging and splitting of proto-jets. Proto-jets which share more than 50% of their energy with another jet are merged to form a single jet, whereas if the shared fraction is less than 50% the energy is split to form two jets.

The jets used in this analysis are required to have $p_T > 15 \text{ GeV}$ and to lie within $|\eta| < 3.4$. They are also required to satisfy the quality criteria defined by the $D\phi$ Jet-ID group [40]. These are designed to remove fake jets produced by noise in the detector modules (hot cells) or problems in the associated electronics. Jets are removed if more than 90% of their energy is coming from a single calorimeter cell or the ratio of the most energetic to second most energetic cell is higher than ten. They are also required to have a coarse hadronic energy fraction of below 44% and an electromagnetic energy fraction between 5 and 95%. Five of the main requirements are listed below.

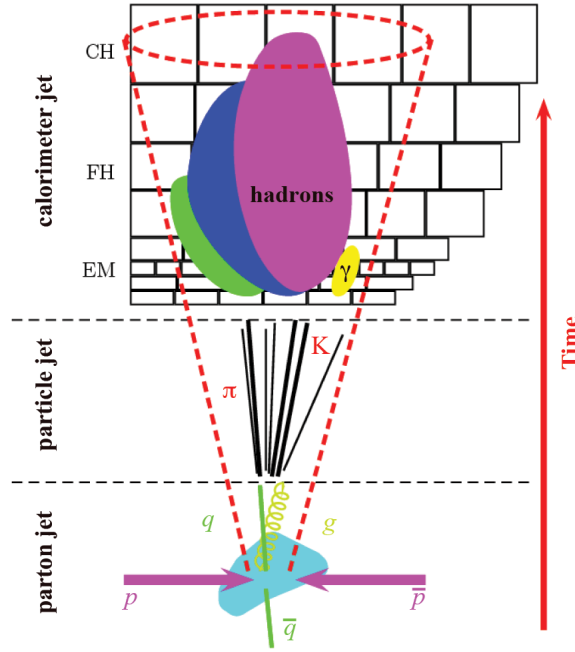


Figure 4.2: Schematic of the parton fragmentation process and the particle showering used to detect the fragmentation products in the calorimeters.

- The EM fraction is required to be $0.05 < f_{EM} < 0.95$.
- Jets occurring in the coarse hadronic layers must have a $f_{CH} < 0.44$ for the CC region and $f_{CH} < 0.46$ for the EC.
- If a cell with the most energy has ten times more energy than the second most energetic cell it is removed where $f_{HOT} = \frac{E_{cell1}}{E_{cell2}} < 10$.
- If 90% of a jets energy is contained in one calorimeter tower it is removed.

4.5.1 Calibration of the Jet Energy Scale

Figure 4.2 shows a schematic of the evolution of a jet from production through to detection by the calorimeters. Each parton from the hard scattering process undergoes fragmentation and is detected by the calorimeters as the collection of particle showers,

which we have identified as a jet. The jet energy measured by the detector is different from that of the parent parton because of the complications of the fragmentation process folded with the effects of finite jet resolution and non-associated backgrounds. To account for this difference we make corrections to the reconstructed jet in order to obtain a better estimate of the energy of the parent parton. This is done using the equation:

$$E_{jet}^{corrected} = \frac{E_{jet}^{uncorrected} - E_o}{F_\eta \times R \times S}, \quad (4.6)$$

where, $E_{jet}^{uncorrected}$ is the jet energy measured after reconstruction and E_o is an offset correction which allows for non-associated background. F_η represents the detector relative response, R is the response, and S is the showering correction. The role of these corrections is discussed below.

E_o: The offset correction, E_o , is associated with energy reconstructed as part of the jet that is not coming from the parent parton. This energy comes from electronic noise, uranium decays, interactions occurring in the same bunch crossing and pile-up (energy due to other collisions which are not part of the same bunch crossing).

R: The response correction, R , is determined using the Missing E_T Projection Fraction method (MPF). This calculates how much of the jet energy is not measured by the detector. This missing energy comes from detector regions that are not fully instrumented, and the regions of the detector which have a non-uniform hadronic/electromagnetic energy ratio. This correction is the largest, and accounts for 30% of the overall correction.

F_η: The relative response correction, F_η , is also determined using the MPF method. The CC and EC calorimeters are reasonably uniform whereas the regions between the calorimeters ($0.5 < |\eta| < 1.8$) are not. F_η is used to correct these non-uniformities and poorly instrumented regions. It is also used as a function of η to do

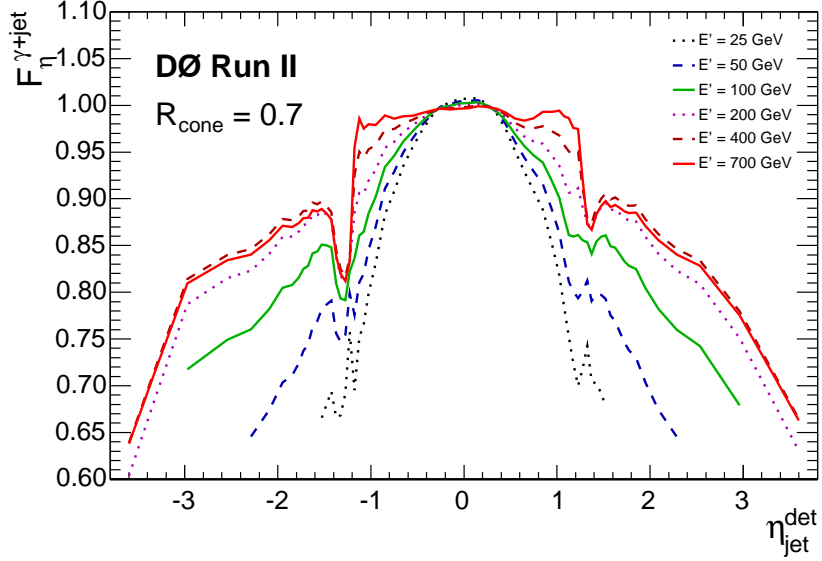


Figure 4.3: F_η measurements as a function of η for $|\eta| < 3.5$.

a relative calibration of the jets in the forward region with respect to the jets in the central region. Figure 4.3 shows the values of F_η obtained from a sample of $\gamma + jet$ events at six different energies. It is clear from this plot that the poorly instrumented regions need correction as the response is low.

S: The showering correction, S , accounts for particles inside the jet cone that leak out of the jet cone and for the particles outside the jet cone that leak energy in.

MPFMethod: (Figure 4.4) gives a schematic representation of how the MPF method works for a $\gamma + jet$ event. At the detector level the energy of a photon is a very well defined, unlike that of the recoiling hadronic jet. Using conservation of momentum we can construct an equation at the detector level to measure a response for the recoiling jet. This is basically a measure of the fraction of energy missing in the jet.

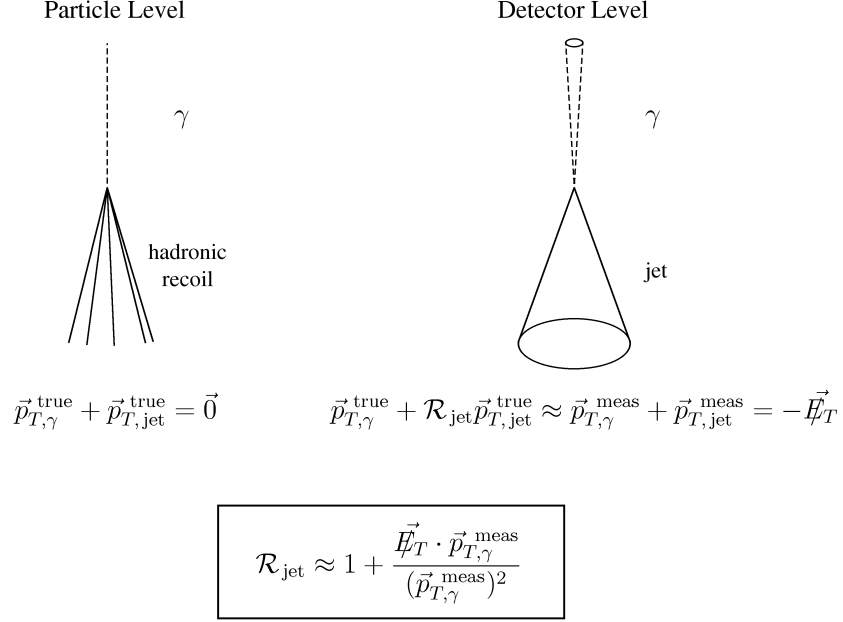


Figure 4.4: Diagram of the MPF method for a $\gamma + \text{jet}$ event.

4.5.2 b -Jet Identification

The association of a jet with a b quark hadronization and decay is an essential part of this analysis because events from both the s and t channels decays have b quarks as part of their signal. A b jet is formed when a b quark hadronizes to form a B hadron, which then decays. The B hadrons have significant lifetimes ($\sim 10^{-12}s$) so that they typically decay a few mm away from the position of the primary vertex. Figure 4.5 shows a schematic of the production of such a displaced or secondary vertex. In order for the reconstruction software to be able to search for a secondary vertex it must be taggable. This means there must be a minimum of two good tracks emanating from a common vertex with at least one SMT hit per track. If these requirements are fulfilled then a jet is taggable and a neural network (NN) b -tagging algorithm [41] is used to differentiate b jets and non b jets. The neural network b -tagger has seven inputs which are used to distinguish between b and non- b quark jets. These are:

- the significance of the decay length,
- the weighted sum of the track impact parameters,
- the probability that the jet originates from the primary vertex,
- the χ^2/ndf of the secondary vertex fit,
- the number of tracks in the secondary vertex,
- the mass of the secondary vertex,
- the number of secondary vertices reconstructed within the jet.

The algorithm outputs a value between 0 and 1 that is associated to the probability that the jet comes from a b -quark. The choice of a cut value in the network discriminant defines the operating point for the tagger. For this analysis we use the tight and loose operating points, for which the discriminant cuts are set to > 0.775 and 0.5 , respectively. The loose operating point is used for the case where an event is required to have two b -tagged jets and the tight operating point is used if only one b -tagged jet is required.

4.6 Reconstruction of Missing Transverse Energy,

$$\cancel{E}_T$$

In this analysis neutrinos are part of the signal we are looking for, however they are not detected at the $D\bar{O}$ detector. We can however, infer their transverse momentum by using the conservation of momentum in the transverse plane. This is a two step process in which we first form the sum of the calorimeter cell depositions and then we apply corrections to account for the unmeasured energy due to any muons in the

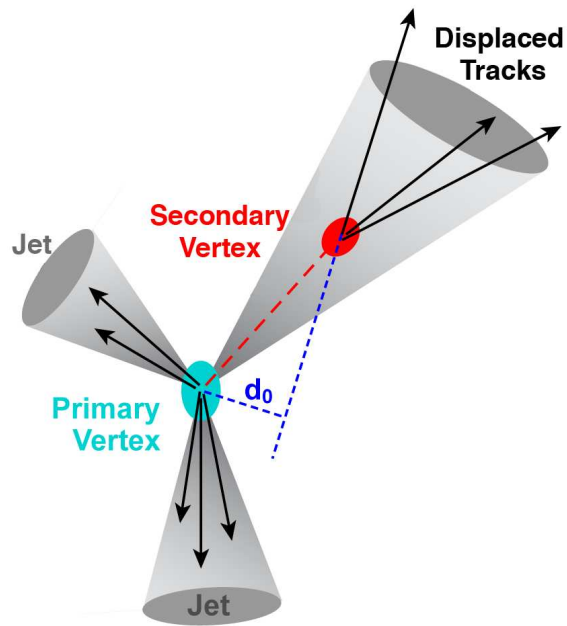


Figure 4.5: Schematic of a 3-jet event in which one of the jets originates from a b -quark and has a displaced secondary vertex. The origin of the non-tagged jets is consistent with the primary vertex.

event. The calorimeter cell sum runs over the cell energies, assuming that they have a common origin at the primary vertex. This uses the cell 3-momentum, \vec{p}_i^{cell} for each of the i electromagnetic and hadronic cells which should vectorially sum to \sim zero if no neutrinos or muons are present and the detector had a perfect resolution.

$$\vec{E}_T^{\text{uncorr}} = - \sum_i \vec{p}_i^{\text{cell}}, \quad (4.7)$$

The presence of one or more muons in the event will modify the imbalance and only a fraction of their momentum (that due to $\frac{dE}{dx}$ in the calorimetry) is included in the sum. In the second step, this is corrected for by modifying the summation to include the muon 3-momentum, taken from the tracker, with a small adjustment to the cell sum for the cells traversed by the muon track. The resolution can be further improved if the event contains identifiable electrons or photons as their energies are typically better measured than the simple cell sum would give. This is also covered in the second level of correction. The best estimate of the transverse energy due to a neutrino is,

$$\vec{E}_T^{\text{corr}} = \vec{E}_T^{\text{uncorr}} - \sum_{\text{objects}} \left(\vec{p}_{\text{object}} - \sum_{i \in \text{object}} \vec{p}_i^{\text{cell}} \right), \quad (4.8)$$

where the object sum runs over all of the identified electrons, photons and muons, \vec{p}_{object} is the particle momentum and cell sum runs over all of the cells traversed by the particle.

Chapter 5

Data Samples and Event Selection

After the Data and simulated event samples have been reconstructed we can begin to use these samples to optimize the event selection. Initially the quality of the data is checked in order to see if there are any needed corrections. Selection cuts are then implemented to maximize the sensitivity for events that may contain signal. After the selection of basic objects we then look at the signal:background (S:B) ratio which is initially 1:185. This ratio improves to 1:20 when b tagging requirements are applied. The following sections discuss the preliminary treatment of the data and monte carlo events.

5.1 The RunII Data Sample

This analysis is based on the data taken during Tevatron run II, which was collected during two data-taking runs between August 2002 and February 2006 (Run IIa) and between June 2006 and June 2009 (RunIIb), which corresponds to an integrated luminosity of 5.4 fb^{-1} (Table 5.1). To be accepted for analysis, an event is required to satisfy at least one trigger from the list of several hundred active triggers from these

data periods. This requirement is an OR of all active triggers used in RunII data. These are the $\gamma/e/\mu$ and e/μ +jets triggers. The resulting selection has been shown to have a trigger efficiency 100% with $\pm 5\%$ uncertainty [42].

Data sample	Trigger Version	Integrated Luminosity [pb^{-1}]
Run IIa	v8 – v14	1079
Run IIb	v15a	534
	v15b	688
	v15c	397
	v16	2662
Total	v8-16	5361

Table 5.1: Luminosity breakdown of the Run II dataset as a function of trigger version.

5.2 Simulated Event Samples

To simulate the response of the $D\bar{O}$ detector a model based on GEANT [43] is used. This simulates the detector response to particles as they pass through it, including the effects of the magnetic fields and any interactions with the material of the detector. In order to simulate real data the output of GEANT is processed by three $D\bar{O}$ packages, $D\bar{O}GSTAR$, $D\bar{O}SIM$ and $D\bar{O}TRIGSIM$. $D\bar{O}GSTAR$ formats the raw data output from GEANT such that it can be used by the other two packages [44]. $D\bar{O}SIM$ and $D\bar{O}TRIGSIM$ are then used to include the effects of electronic noise and a model of the trigger framework, respectively.

Due to the detector being upgraded after RunIIa both RunIIa and IIb are simulated separately. The following sections discuss the comparison of the data and simulated events. This is followed in Chapter 6 by a discussion of corrections that are needed to correct for additional resolution and modeling insufficiencies.

5.2.1 Simulation of Single Top Signal

We have generated samples of the s and t channel single top processes using the CompHEP-SINGLETOP [45] event generator assuming a top quark mass of 172.5 GeV. These have been processed through GEANT and reconstructed using the same software used for data. We have compared these samples with samples generated using NLO QCD calculations and we find good agreement [45]. We have chosen to use the CompHEP-SINGLETOP generator, rather than the NLO QCD generator because it preserves the correct spin information in the particle decays. PYTHIA [47] was used to process the generated samples which add the underlying event and initial and final state radiation corrections. This occurs before GEANT processing and reconstruction. Within PYTHIA Tau and B hadron decays are modeled using , TAUOLA [48] and EVTGEN [49], respectively.

5.2.2 Background Event Simulation

The dominant backgrounds to single top production are the events from W +Jets, Z +Jets, QCD multijets, $t\bar{t}$ and diboson production. Each of these are modeled using monte carlo simulations except the QCD multijet background, which is obtained from data.

The ALPGEN v-2.11 [50] event generator was used to simulate the backgrounds from W +Jets, Z +Jets and $t\bar{t}$ production, and the the diboson backgrounds (WW , WZ and ZZ) were simulated using PYTHIA. For this analysis a detailed discussion of these calculations is given in reference [52]. For this analysis a zero-bias overlay of events was used to improve the modeling of higher instantaneous luminosity. These are events that are recorded without any trigger requirement.

Summarized in Table 5.2 is a list of all the cross sections and branching fractions

used in these simulations along with the amount of events generated in each sample.

Event Type	Cross Section [pb]	Branching Fraction	RunIIa Events	RunIIb Events
$tb \rightarrow \ell + \text{jets}$	$1.04^{+0.04}_{-0.04}$	0.3240 ± 0.0032	0.6M	0.5M
$tqb \rightarrow \ell + \text{jets}$	$2.26^{+0.12}_{-0.12}$	0.3240 ± 0.0032	0.6M	0.5M
Total	$3.30^{+0.16}_{-0.16}$	0.3240 ± 0.0032	1.2M	1.0M
$t\bar{t} \rightarrow \ell + \text{jets}$	$7.46^{+0.48}_{-0.67}$	0.4380 ± 0.0044	1.5M	1.3M
$t\bar{t} \rightarrow \ell\ell$	$7.46^{+0.48}_{-0.67}$	0.1050 ± 0.0010	1.5M	1.3M
Top pairs total	$7.46^{+0.48}_{-0.67}$	0.5430 ± 0.0054	3.0M	1.6M
$Wbb \rightarrow \ell\nu bb$	90.5	0.3240 ± 0.0032	2.7M	3.0M
$Wc\bar{c} \rightarrow \ell\nu cc$	260	0.3240 ± 0.0032	2.7M	3.0M
$Wjj \rightarrow \ell\nu jj$	23,831	0.3240 ± 0.0032	55M	97M
W+jets total	24,182	0.3240 ± 0.0032	60.4M	103M
$Zbb \rightarrow \ell\ell bb$	38.7	0.03366 ± 0.00002	0.7M	0.7M
$Zc\bar{c} \rightarrow \ell\ell cc$	106	0.03366 ± 0.00002	0.7M	0.7M
$Zjj \rightarrow \ell\ell jj$	7,032	0.03366 ± 0.00002	14M	4.0M
Z+jets total	7,177	0.03366 ± 0.00002	15.4M	5.4M
$WW \rightarrow \text{anything}$	11.6 ± 0.4	1.0 ± 0.0	2.0M	0.7M
$WZ \rightarrow \text{anything}$	3.25 ± 0.11	1.0 ± 0.0	1.0M	0.6M
$ZZ \rightarrow \text{anything}$	1.33 ± 0.04	1.0 ± 0.0	1.0M	0.5M
Diboson total	16.2 ± 0.6	1.0 ± 0.0	4.0M	1.8M

Table 5.2: Sizes of the signal and background simulated event samples. The labels Run IIa and Run IIb, denote that these events were reconstructed with the software used for the data from these subsets of the data.

The experimental signatures for the signal processes are events which contain a isolated lepton with high p_T , large \cancel{E}_T (neutrino), 1-2 b jets and 1-3 light quark jets all having a high p_T . The backgrounds have slightly different characteristics which can be used in their rejection. Each of these is discussed below, ranked in importance, from greatest to least.

Figure 5.1 shows the Feynman diagrams for two $W + b$ jet processes which can mimic the single top signal. The one on the left clearly mimics a single top signal with two b jets, a lepton and missing energy and the one on the right can also mimic the signal if the jet on the top is misidentified as a b jet. These can be suppressed by making use of the differences between the angular correlations for the signal and backgrounds.

$t\bar{t}$ events can easily mimic the single top signatures as each event naturally contains two b -quark jets. These are lepton+jets and dilepton events. The lepton+jets

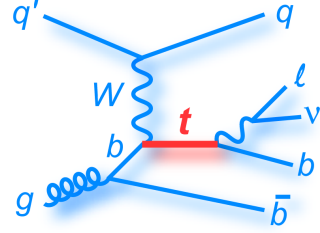
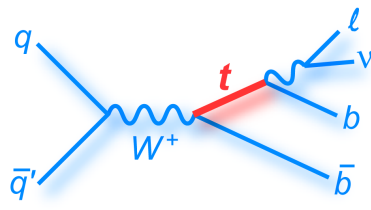
events (Figure 5.1) produce events with additional high p_T jets. Dilepton events can also mimic single top if one of the leptons is not reconstructed. In this case the reconstructed W mass and the amount of the missing energy can be used as discriminators.

QCD multijet events start out as all jet events without significant \cancel{E}_T or high p_T leptons. To mimic the single top signal requires that the lepton comes from a b or c quark decay within a jet, which gets misidentified as an isolated lepton, or the jet may fluctuate and fake the signal of an electron. Most of the \cancel{E}_T part of the signature generally arises due to mismeasurement of one or more of the jets energy in the event. These events are unlikely, however the cross section for QCD multijet production is so large that they occur at a detectable rate. These events are handled by knowing that there are no W bosons or top quarks in the event.

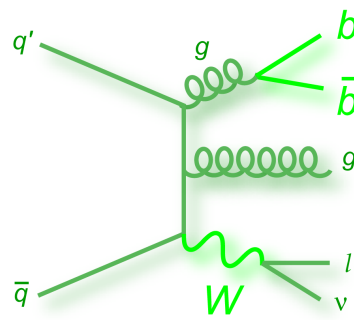
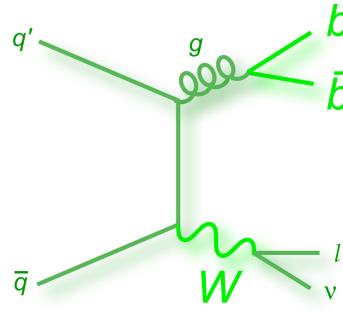
Z +Jet production is also shown in Figure 5.1. For this process to fake a signal, one of the leptons from the Z decay is not reconstructed, which provides the \cancel{E}_T signature. Similar to the W case, the b jets can be real. Suppression of this background is done by using the size of the \cancel{E}_T and by imposing a minimum p_T cut on the b jets. Overall Z +Jets make up a small portion of the total background.

Diboson events (WW , WZ , ZZ) can also fake the single top signature. In the WW and WZ case one W decays leptonically and the other hadronically, giving the high p_T lepton + \cancel{E}_T and the b jets arise via bremsstrahlung, as in the $W(Z)$ +jets cases. In the ZZ case, one Z must decay hadronically and the other leptonically where the lepton gets misreconstructed. As for other cases, the b jets are due to bremsstrahlung. The production cross sections for these three processes are so small that they can safely be neglected.

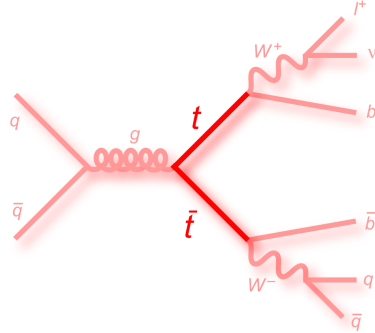
Single Top Production



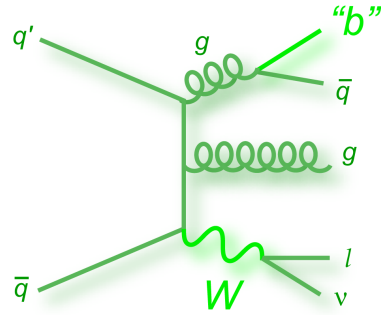
W+Jets



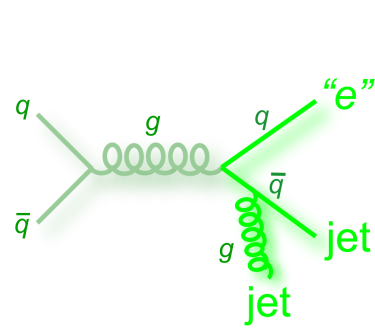
$t\bar{t}$ production



Z+Jets



Mulijets



Diboson

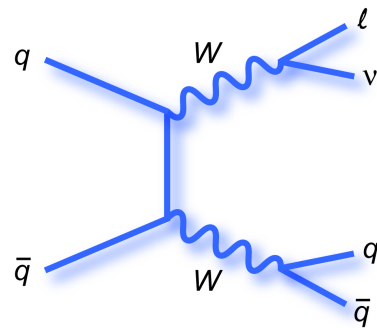


Figure 5.1: Lowest order Feynman diagrams for the s and t channel signal processes and the principle backgrounds.

5.3 Modeling the QCD Multijet Background

As stated previously, the QCD multijet background is modeled using data. Two samples are produced separately, one for the electron channel and one for the muon channel. These event samples are derived from a subset of the data which passes almost all of the selection criteria. The samples are made orthogonal to the analysis samples by inverting and removing some of the lepton identification cuts. For the electron+jets channel the requirement for track matching is removed and the electron likelihood cut is reversed to $\mathcal{L}_e < 0.85$. This allows the selection of fake electrons and significantly increases the number of events for the QCD multijet sample. This also distorts the shape of the lepton p_T distribution. The method used to correct this is discussed in Chapter 6. For muon+jets channel, the QCD sample is selected by removing the muon and $\Delta R(\mu, jet)$ requirements.

5.4 The Effects of the Top Mass Uncertainty

The top quark mass has a uncertainty of around 2 GeV [51]. This analysis is sensitive to such a change in mass. To study this we have simulated single top and $t\bar{t}$ production with three different top quark masses (170.0 GeV, 172.5 GeV and 175.0 GeV) [53]. The added benefit of this triples the statistics allowing a lower dependence on the top quark mass. This new merged sample is only used to train Boosted Decision Trees discussed in Chapter 7.

5.5 Determination of the W +Jets and Multijet Background Normalizations

The normalization of the single top and background samples other than W +Jets and QCD multijets is done using the integrated luminosity and the cross sections listed in Table 5.2. The W +Jets and QCD multijet are normalized to data using the equation:

$$N = \lambda_{\text{wjets}} N_{\text{wjets}} + \lambda_{\text{multijets}} N_{\text{multijets}}, \quad (5.1)$$

where, N is the total number of W +Jet and multijet events ($N = N_{\text{Data}} - N_{\text{non-wjets-MC}}$), λ_{wjets} and $\lambda_{\text{multijets}}$ are the normalization scale factors which are anticorrelated. All samples used here do not have any b tagging requirements. A simultaneous fit is performed using three kinematic distributions which are sensitive to both λ_{wjets} and $\lambda_{\text{multijets}}$. The distributions included in the fit are the lepton p_T , \cancel{E}_T and the transverse mass, $M_T(l, \cancel{E}_T)$ [42]. In order to obtain values for both λ s an iterative procedure is used by first setting $\lambda_{\text{wjets}} = 0$ and solving for $\lambda_{\text{multijets}}$ using Eq. 5.1. A Kolmogorov-Smirnov (KS) test is then used to measure the quality of the fit then repeated by increasing $\lambda_{\text{multijets}}$ in steps of 0.001 to 4.0. The final normalization is chosen from the fit result with the best combined KS fit for all three distributions. Table 5.3 shows the final results used for RunIIa and RunIIb 2, 3 and 4 jet data samples and for the electron and muon channels.

	λ_{wjets}				$\lambda_{\text{multijets}}$			
	Run IIa		Run IIb		Run IIa		Run IIb	
	e	μ	e	μ	e	μ	e	μ
2 jets	1.071	1.042	0.971	0.954	0.444	0.014	0.388	0.034
3 jets	1.326	1.288	1.082	1.105	0.351	0.025	0.343	0.045
4 jets	1.371	1.553	1.070	1.294	0.309	0.033	0.383	0.024

Table 5.3: Fitted W+Jets and QCD Multijet Scale Factors for each of the 24 fits.

5.6 Event Selection

The main point of the event selection cuts is to identify events that are W -like with one isolated lepton, missing energy, and between two and four high p_T jets. After a loose preselection, the data is partitioned into samples with two, three and four jets. These are collectively categorized as pertagged events, since no b tagging requirement has been applied. The majority of the events are W +Jets with a small number of $t\bar{t}$ events which dominate for higher jet multiplicities. Discussed below are the requirements used for event preselection which are separated into three groups: general (applied to all events), and electron (muon) channel which are additional cuts made specifically for that channel.

5.6.1 General Preselection Requirements

Below are the general selection requirements used to pre-select data and simulated events for the electron+jets and muon+jets channels.

For good quality data, runs that are identified as having detector or readout problems are rejected before any kinematic selections. Also, any runs having a negative instantaneous luminosity is thrown out and the total integrated luminosity is corrected for both effects.

Events are required to have at least one trigger firing in the OR list and have a good primary vertex where $|z_{PV}| < 60$ cm with at least three tracks connected to it. This rejects events which are poorly defined trigger objects and that are at the edges of the beamspot region. To be consistent with single top jet topology, it is required that two to four jets have a $p_T > 15$ GeV and exist in $|\eta_{det}| < 3.4$. To remove some of the backgrounds the leading jet must have a $p_T > 25$ GeV. RunIIb has the highest instantaneous luminosity, therefore additional cuts are made to remove fake jets and

soft p_T backgrounds. This consists of jets that have matching tracks to be consistent with a primary vertex and for the three (four) jet sample to have the second leading jet $p_T > 20 \text{ GeV}$ ($> 25 \text{ GeV}$), respectively.

As with many $D\bar{O}$ analyses triangle cuts are implemented to suppress background contributions where the missing energy is a result of the mismeasurement of a jet. For this analysis, triangle cuts reject much of the QCD multijet background in which \cancel{E}_T is produced by jet mismeasurement. We apply two cuts where the first considers the correlation between \cancel{E}_T and the opening angle $|\Delta\phi|$. This angle lies between the leading jet direction and the \cancel{E}_T vector. In order for events to be kept, they must lie outside of the triangular shaded region of Figure 5.6.1. The shaded region is defined such that $|\Delta\phi(\text{leading jet}, \cancel{E}_T)| = \frac{\pi}{2}$ for $\cancel{E}_T = 0$ to $|\Delta\phi(\text{leading jet}, \cancel{E}_T)| = \pi$ with $\cancel{E}_T = 35 \text{ GeV}$. The second triangular cut considers the second leading jet p_T in events

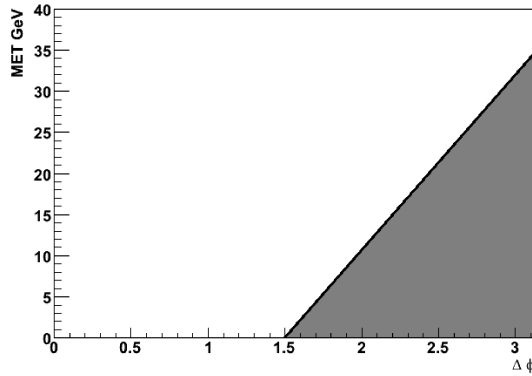


Figure 5.2: Triangle cut in the $|\Delta\phi(\text{leading jet}, \cancel{E}_T)|$ vs. \cancel{E}_T plane.

only having exactly two jets and is defined as the p_T of this jet versus the H_T of the event ($H_T = \sum p_T^{jets}$). Figure 5.6.1 shows the implementation of this cut where events that lie within the triangle are rejected. The line drawn is defined from $p_T = 27.5 \text{ GeV}$ at $H_T = 0 \text{ GeV}$ to $p_T = 0 \text{ GeV}$ at $H_T = 165 \text{ GeV}$. In order to provide further cleaning we apply boundary values to \cancel{E}_T to reject poorly mismeasured events. Two jet events

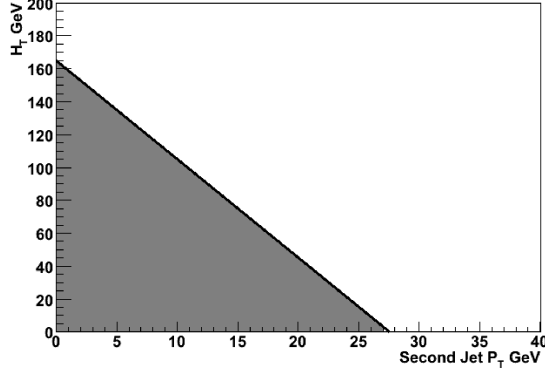


Figure 5.3: Triangle cut in the second leading jet p_T vs. $H_T(\text{alljets})$ plane.

are retained if their \cancel{E}_T lie between 20 GeV to 200 GeV and events with three to four jets are kept if they lie between 25 GeV to 200 GeV.

Lastly, threshold cuts are required for event H_T (alljets) and $H_T(\text{alljets, lepton}, \cancel{E}_T)$. Also, these are implemented to reduce the background for soft QCD multijet events. For H_T (alljets) minimum thresholds of 50, 75 and 100 GeV are required for the two, three and four jet events (RunIIa only). For RunIIb these are extended to 55, 80 and 110 GeV due to higher luminosity running conditions. For $H_T(\text{alljets, lepton}, \cancel{E}_T)$ minimum thresholds of 120, 140 and 160 GeV are required for two, three and four jet events for both RunIIa and RunIIb.

5.6.2 Electron+Jets Channel Preselection

For a candidate electron+jet event to be kept it is required to have one tight electron with $|\eta^{\text{det}}| < 1.1$ and $p_T > 15$ that passes the tight electron ID requirements and is consistent with primary vertex $|\Delta z(e, \text{PV})| < 1$ cm. For events with ≥ 2 good jets the p_T is raised to 20 GeV. Events with an additional loose electron with $p_T > 15$ GeV and passes the loose electron definition are vetoed. Likewise, events with a tight isolated muon with $p_T > 15$ GeV and within $|\eta_{\text{det}}| < 2.0$ are rejected as well. This

gets rid of potential background coming from $t\bar{t}$ dilepton and Z^0 decays.

Additionally, three triangle cuts are implemented for the electron+jets channel in order to suppress the QCD multijet background. These are defined in the 2D plane of \cancel{E}_T vs. $|\Delta\phi(e, \cancel{E}_T)|$. Events are removed if they exist in any of the triangular shaded regions shown in Figure 5.6.2. The three lines are drawn as follows:

- I.) $|\Delta\phi| = 2$ rad at $\cancel{E}_T = 0$ GeV to $|\Delta\phi| = \pi$ at $\cancel{E}_T = 24$ GeV;
- II.) $|\Delta\phi| = 0$ rad at $\cancel{E}_T = 50$ GeV to $|\Delta\phi| = 1.5$ rad at $\cancel{E}_T = 0$ GeV;
- III.) $|\Delta\phi| = 0$ rad at $\cancel{E}_T = 40$ GeV to $|\Delta\phi| = 2$ rad at $\cancel{E}_T = 0$ GeV.

The initial cut is a back-to-back restriction and the second two define a nearside restriction.

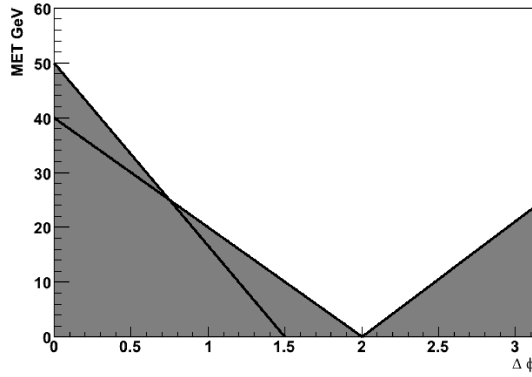


Figure 5.4: Triangle cuts in the \cancel{E}_T vs. $|\Delta\phi(e, \cancel{E}_T)|$ plane.

5.6.3 Muon+Jets Channel Preselection

For a candidate muon+jet event to be retained it is required to have only one tight muon with $p_T > 15$ GeV and $|\eta_{det}| < 2.0$ that passes the muon ID requirements and is consistent with primary vertex $|\Delta z(\mu, PV)| < 1$ cm. Events with additional

loose muon with $p_T > 4$ GeV and pass the loose muon requirements are vetoed. Likewise, events with a loose electron that has $p_T > 15$ GeV and within $|\eta_{det}| < 2.5$ gets rejected. This gets rid of potential background coming from $t\bar{t}$ dilepton and Z^0 decays.

The momentum of high p_T muons can be severely miss measured, because of this an additional track quality cut is required to reject events with this type of signature. This cut is defined by imposing a restriction on the significance of the track curvature. The track curvature is defined as:

$$|TrackCurvSig| = \left| \frac{q/p_T}{\sigma(1/p_T)} \right|, \quad (5.2)$$

where, q and p_T are the charge and transverse momentum of the charged track associated with the muon. RunIIa and RunIIb both have different values due to the change in the amount of background considered in the tracker. There are two cuts which are defined in the $|TrackCurvSig|$ vs. $|\Delta\phi(\mu, \cancel{E}_T)|$ plane. They are:

- I.) $|\Delta\phi|$ from 0.875π to π rad when $|TrackCurvSig| = 0$, and $|TrackCurvSig|$ from 0 to 4 (6) when $|\Delta\phi| = \pi$ rad for Run IIa (Run IIb) period;
- II.) $|\Delta\phi|$ from 2 to π rad when $|TrackCurvSig| = 0$, and $|TrackCurvSig|$ from 0 to 2 (3) when $|\Delta\phi| = \pi$ rad for Run IIa (Run IIb) period.

Additionally, two triangle cuts are implemented to suppress the QCD multijet background. Similar to the electron+jets channel, these are defined in the \cancel{E}_T vs. $|\Delta\phi(\mu, \cancel{E}_T)|$ plane. Events are removed if they exist in any of the triangular shaded regions shown in Figure 5.6.3. The two bounding lines are drawn as follows:

- I.) $|\Delta\phi| = 2.5$ rad at $\cancel{E}_T = 0$ GeV to $|\Delta\phi| = \pi$ at $\cancel{E}_T = 30$ GeV;
- II.) $|\Delta\phi| = 0$ rad at $\cancel{E}_T = 85$ GeV to $|\Delta\phi| = 1.2$ rad at $\cancel{E}_T = 0$ GeV;

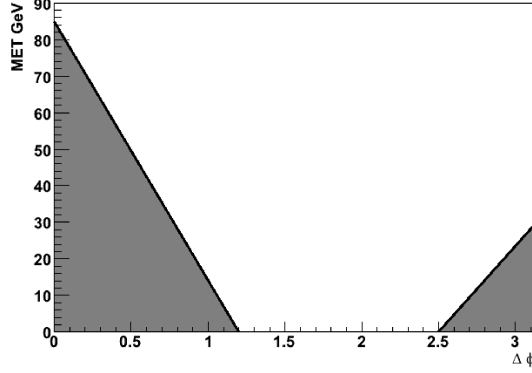


Figure 5.5: Triangle cuts in the E_T vs. $|\Delta\phi(\mu, E_T)|$ plane.

The last two cuts which are applied to the muon+jets channel are where either of the leading or second leading jets lie within the ICD region which is located between the central and endcap calorimeters ($1.0 < |\eta^{\text{det}}| < 1.5$). Both RunIIa and RunIIb are treated separately relative to the two cuts where RunIIa only applies the first cut and RunIIb both. These cuts are defined as:

- I.) Leading jet cut: $p_T > 30$ GeV for leading jets in the ICD region;
- II.) Second leading jet cut: $p_T > 30$ GeV for next to leading jets in the ICD region for events with exactly three good jets.

5.7 Analysis Subsamples and Preliminary Data-Simulation Comparisons

To optimize the tuning of the Boosted Decision Trees (Chapter 7) the data is split into six subsamples relative to the event topology. Both RunIIa and RunIIb are combined for this purpose. It has been shown that by splitting the samples in this way maximizes the discrimination between signal and background [53]. The six subsamples

are divided by two, three or four jets and either one or two b tags. Tables 5.4 and 5.5 present a comparison between the observed and expected yields for all six subsamples. All simulated samples are broken down to show the amount of each background process. For all cases the predicted and observed yields have good agreement.

Samples	Electron+Muon, Run IIa+Run IIb			All Channels
	2 jets	3 jets	4 jets	
tb	66 ± 0.46	27 ± 0.30	7.7 ± 0.17	100 ± 0.57
tqb	133 ± 0.78	58 ± 0.55	18 ± 0.32	209 ± 1.0
$tb+tqb$	198 ± 0.91	85 ± 0.63	26 ± 0.36	309 ± 1.2
$t\bar{t} \rightarrow ll$	168 ± 0.84	121 ± 0.75	37 ± 0.38	326 ± 1.2
$t\bar{t} \rightarrow l + \text{jets}$	144 ± 1.6	424 ± 2.7	462 ± 2.8	$1,030 \pm 4.2$
$Wb\bar{b}$	989 ± 9.4	315 ± 4.6	84 ± 2.2	$1,388 \pm 11$
$Wc\bar{c}$	554 ± 9.5	208 ± 4.9	56 ± 2.3	818 ± 11
Wcj	612 ± 15	137 ± 6.9	24 ± 2.2	774 ± 16
Wjj	$1,066 \pm 4.1$	301 ± 2.2	74 ± 1.1	$1,441 \pm 4.8$
$Zb\bar{b}$	78 ± 1.6	28 ± 1.2	7.3 ± 0.53	113 ± 2.1
$Zc\bar{c}$	39 ± 1.4	17 ± 1.2	4.5 ± 0.51	60 ± 1.9
Zjj	82 ± 5.7	31 ± 3.6	5.3 ± 1.0	118 ± 6.8
Dibosons	158 ± 3.1	48 ± 1.7	12 ± 1.1	218 ± 3.7
Multijets	258 ± 2.2	117 ± 1.9	37 ± 0.83	412 ± 3.1
Total Background	$4,150 \pm 22$	$1,745 \pm 11$	804 ± 5.2	$6,699 \pm 25$
Background + Signal	$4,348 \pm 22$	$1,830 \pm 11$	830 ± 5.3	$7,008 \pm 25$
Data	$4,284 \pm 65$	$1,772 \pm 42$	851 ± 29	$6,907 \pm 83$

Table 5.4: Comparison of observed and predicted yields for the single tagged event samples.

Samples	Electron+Muon, Run IIa+Run IIb			All Channels
	2 jets	3 jets	4 jets	
$t\bar{t}b$	38 ± 0.31	17 ± 0.21	5.2 ± 0.12	60 ± 0.39
$tq\bar{b}$	7.8 ± 0.21	14 ± 0.27	7.3 ± 0.19	29 ± 0.39
$t\bar{t}b+tq\bar{b}$	46 ± 0.37	31 ± 0.34	13 ± 0.22	90 ± 0.55
$t\bar{t} \rightarrow l\bar{l}$	90 ± 0.56	83 ± 0.52	27 ± 0.28	199 ± 0.82
$t\bar{t} \rightarrow l+jets$	31 ± 0.80	203 ± 1.7	334 ± 2.1	568 ± 2.8
$Wb\bar{b}$	252 ± 4.5	90 ± 2.6	26 ± 1.2	367 ± 5.3
$Wc\bar{c}$	41 ± 2.3	25 ± 1.8	11 ± 1.2	77 ± 3.2
Wcj	16 ± 0.65	7.3 ± 0.55	1.6 ± 0.16	24 ± 0.87
Wjj	30 ± 0.35	17 ± 0.28	6.0 ± 0.15	53 ± 0.47
$Zb\bar{b}$	16 ± 0.64	8.5 ± 0.49	2.6 ± 0.33	27 ± 0.87
$Zc\bar{c}$	3.2 ± 0.39	1.4 ± 0.27	0.56 ± 0.15	5.1 ± 0.50
Zjj	2.9 ± 1.3	1.8 ± 0.81	0.64 ± 0.28	5.4 ± 1.6
Dibosons	20 ± 0.81	6.6 ± 0.49	2.0 ± 0.27	29 ± 0.98
Multijets	19 ± 0.62	13 ± 0.64	6.3 ± 0.33	39 ± 0.95
Background Sum	520 ± 5.5	456 ± 3.9	418 ± 2.8	$1,394 \pm 7.3$
Background + Signal	566 ± 5.5	487 ± 3.9	431 ± 2.8	$1,484 \pm 7.3$
Data	597 ± 24	535 ± 23	432 ± 21	$1,564 \pm 40$

Table 5.5: Comparison of observed and predicted yields for the the double tagged event samples.

Figure 5.7 shows the comparison of data with all the simulations for fifteen of the variables used in the BDT analysis. Each comparison plot is a combination of all six subsamples and show good agreement. At the split subsample level comparisons similar to these also exhibit good agreement [42]. Figure 5.6 below presents the colour scheme used throughout this analysis to label data, signal and background events occurring in all plots.

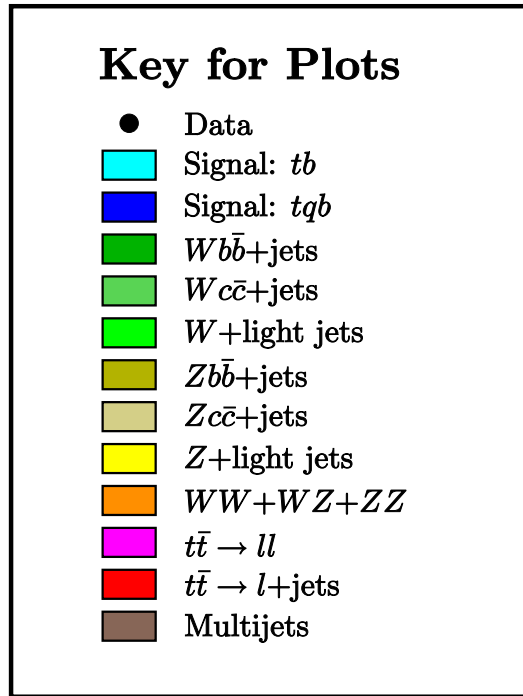


Figure 5.6: Colour scheme plot key used to label Data, signal and background events for plots.

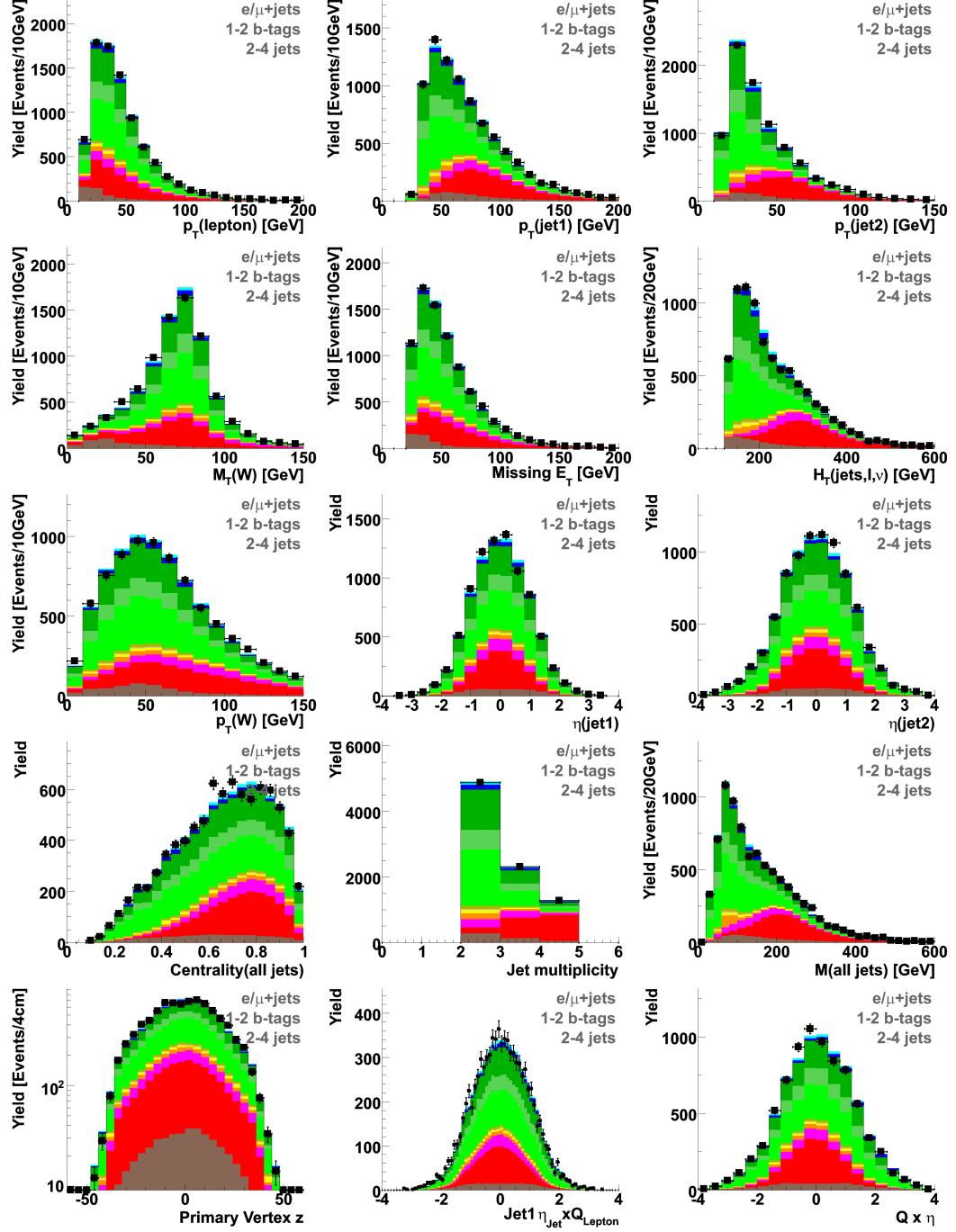


Figure 5.7: Comparison of the observed and predicted yields for 15 of the kinematic variables used as inputs to the BDT.

Chapter 6

Monte Carlo Model Corrections and Statistical Uncertainties

The $D\bar{O}$ detector has been around for quite awhile where the effect of its aging is present in the efficiencies of some of its components and resolution. The GEANT model does take most of these effects into account, however it is not entirely correct and additional corrections are needed to compensate for the residual differences. Corrections to the underlying physics simulations are also needed to agree with current values. These are discussed in the following sections.

6.1 Detector Response Corrections

Both RunIIa and RunIIb have different instantaneous luminosities. This is modeled by adding together zero-bias events with corresponding simulated events. Because zero-bias events are limited in size the MC instantaneous luminosity does not reflect that of the instantaneous luminosity for data. Therefore, the MC instantaneous luminosity is reweighted to reflect the same value as data. Zero-bias events are recorded

data events which have no trigger requirements.

The position and size of the beam spot is also different for both run periods. Specifically, when comparing these to simulated events we find that the primary vertices z distributions are narrower than what is seen in data. Since this also affects the acceptance calculations, a correction factor is given to simulated events so as to be consistent with data. The correction factor is given as a function of instantaneous luminosity and run period [54].

The electron identification efficiencies for monte carlo and data do not match, with the former having higher efficiencies than that of data. Therefore, the monte carlo events are reweighted using the equation

$$\varepsilon_{e-ID} = \frac{\varepsilon_{Presel}^{Data}}{\varepsilon_{Presel}^{MC}} \times \frac{\varepsilon_{PostPresel}^{Data}}{\varepsilon_{PostPresel}^{MC}} \quad (6.1)$$

where the terms in the correction are expressed in two parts, labeled preselection and post preselection. The pre-selection ratio focuses on the electromagnetic fraction and isolation efficiency contribution and is parametrized as a function of η_{det} . The latter ratio is parameterized in (η_{det}, ϕ) and is determined by using a tag-and-probe analysis of $Z \rightarrow ee$ events and takes account of the differences in the H-matrix term.

A similar correction is needed to obtain the muon identification efficiency values. This is reweighted using the equation

$$\varepsilon_{\mu-ID} = \frac{\varepsilon_{Reco}^{Data}}{\varepsilon_{Reco}^{MC}} \times \frac{\varepsilon_{Track|Reco}^{Data}}{\varepsilon_{Track|Reco}^{MC}} \times \frac{\varepsilon_{Isolation|Track}^{Data}}{\varepsilon_{Isolation|Track}^{MC}} \quad (6.2)$$

where each of the three correction terms relate to the uncertainties in the simulation of the ID efficiency (Reco), the muon-central tracking matching efficiency (Track—Reco) and the isolation efficiency (Isolation—Track), respectively. A tag and probe method

of $Z \rightarrow \mu\mu$ events is used to determine each of these. The muon ID term is parametrized as a function of (η_{det}, ϕ) , the track matching term in (z, η_{CFT}) and the isolation term in η [58, 59].

The simulation of the jet resolution also needs some adjustment to make it match the data. For this analysis the $D\phi$ JSSR (Jet Smearing Shifting and Removal) is used [60]. The re-weighting corrections are obtained from the comparison of simulated and actual γ +jet events. These corrections are applied to each of the W +Jet, Z +Jet, and diboson samples used in this analysis.

The last set of detector response corrections are used to match the simulated b tagging efficiencies in monte carlo to those in data. The corrections are divided into two parts for the b tagging process. Initially a jet must have enough information for the b tagging algorithm to designate a jet as taggable, called the taggability efficiency. Second is the efficiency for the algorithm to tag b jets, called the tagging efficiency. Both these efficiencies are determined on a per jet basis from data and monte carlo simulations. Ratios of the data and monte carlo efficiencies are then used to provide the necessary correction factors. Since this procedure is quite involved, it is discussed in detail below.

The taggability scale factor SF_T can be defined as the number of taggable jets divided by the total number of good jets. It is expressed in the form

$$SF_T(p_T, \eta, PV_z) = \frac{T^{\text{data}}(p_T, \eta)}{T^{\text{mc}}(p_T, \eta)}, \quad (6.3)$$

where $T(p_T, \eta)$ is the taggability efficiency as a function of the jet p_T and η . SF_T is parametrized as a function of the jet p_T , η , and the z position of the PV_z . A scale facotor can then be formed to preserve normalization for untaggable jets and is given

as:

$$SF_{\bar{T}}(p_T, \eta, PV_z) = \frac{1 - T^{\text{data}}(p_T, \eta)}{1 - T^{\text{mc}}(p_T, \eta)}. \quad (6.4)$$

The scale factor per event can then be derived as:

$$\prod_{i=1}^{N_T} SF_T(p_{T,i}, \eta_{det,i}, PV_{z,i}) \times \prod_{j=1}^{N_{\bar{T}}} SF_{\bar{T}}(p_{T,j}, \eta_{det,j}, PV_{z,i}), \quad (6.5)$$

where N_T and $N_{\bar{T}}$ stand for the number of taggable and untaggable jets in each event.

To account for some of the correlations between the jets an additional correction,

$SF(N_T, N_{\bar{T}})$ is required, where:

$$SF(N_T, N_{\bar{T}}) = \frac{E^{\text{data}}(N_T, N_{\bar{T}})}{E^{\text{MC}}(N_T, N_{\bar{T}})}. \quad (6.6)$$

Here E represents the number of events in each sample. Thus the final taggability scale factor is:

$$SF(N_T, N_{\bar{T}}) \times \prod_{i=1}^{N_T} SF_T(p_{T,i}, \eta_{det,i}, PV_{z,i}) \times \prod_{j=1}^{N_{\bar{T}}} SF_{\bar{T}}(p_{T,j}, \eta_{det,j}, PV_{z,i}). \quad (6.7)$$

This correction is applied to all of the simulated event samples.

The $D\bar{O}$ b -ID provides the b tagging efficiencies for both data and monte carlo events. These efficiencies are a function of jet p_T , η and the operating point of the neutral network tagger, NN_{out} . The following scale factors are then used to match the monte carlo simulations to data:

$$SF_{tag}(p_T, \eta, NN_{out}) = \frac{\epsilon^{\text{data}}(p_T, \eta, NN_{out})}{\epsilon^{\text{mc}}(p_T, \eta, NN_{out})}. \quad (6.8)$$

Thus the per event scale factor is given by:

$$\prod_{i=1}^{N_T} SF_{tag}(p_{T,i}, \eta_i, NN_{out}). \quad (6.9)$$

N_T is the number of taggable jets.

The regular b tagging algorithm is not used for the simulated W +light jets background due to the tagging algorithm accepting fake b tagged jets from this sample. Using the regular b tagging algorithm would result in low statistical samples due to the small fake probability. To get around this problem we use a random tagging technique [61]. This is implemented by selecting random untagged jets in the simulated W +light jets samples and assigning them a 50% tag probability. All of the selected jets which are taggable are then weighted using the appropriate scale factor.

6.2 Theoretical Uncertainties

This section provides a discussion for corrections made for various theoretical uncertainties that have an affect on the background simulations. Specifically, the most significant of these are the treatment of the W +heavy flavor background, the corrections to the shape of the Zp_T spectrum and the bias in the QCD multijet sample that is used for the analysis in the electron channel.

The W +heavy flavor production cross sections are not accurately known. This must be considered because the $W + b\bar{b}$ background is sensitive to this analysis due to the presence of real b quarks in events. In the previous $D\bar{O}$ analysis [52], a heavy flavor correction (λ_{HF}) was measured and applied to the $Wb\bar{b}$ and $Wc\bar{c}$ samples to provide better data-background agreement [52]. For this analysis studies were also performed, resulting in a measured factor $\lambda_{HF}=1 \pm 12\%$ [52]. Based on these measurements, a

correction was not needed, however its uncertainties are included.

ALPGEN [55] was used to simulate the Z +jets background. There is a difference in shape of the Zp_T spectrum between ALPGEN and current NLO predictions. Because of this the simulated events are corrected so as to follow the NLO distribution.

The W +Jets and Z +Jets simulated backgrounds show good agreement with data, however there are a few discrepancies in the the jet rapidity and angular correlations. It is assumed that these discrepancies come about by using the leading log approximation in ALPGEN associated with the zero bias overlay being limited in the ICR region. These discrepancies are corrected by applying a reweighting factor R_i to simulated events before the b tag requirements. Each distribution is reweighted in the i^{th} bin and is defined by

$$R_i = \frac{N_i}{N_{V+jets,i}}, \quad (6.10)$$

where, N_i is the number of observed events minus the predicted yields from all processes other than W and Z +Jets ($N_i = N_{Data,i} - N_{QCD+non-vjets,i}$). $N_{vjets,i}$ is the sum of the W and Z +Jets predicted event yields in bin i . These ratios are then plotted for each of the angular distributions to be corrected and fitted to determine a reweighting function. Five distributions are considered: $\eta(jet_1)$, $\eta(jet_2)$, $\Delta R(jet_1, jet_2)$, $\eta(jet_3)$ and $\eta(jet_4)$. The reweighting is done sequentially, starting with $\eta(jet_1)$ and ending with $\eta(jet_4)$, where $\eta(jet_3)$ and $\eta(jet_4)$ are used in the three jet and four jet samples, respectively. For a more detailed discussion of the reweighting process please see reference [42].

The QCD multijet sample used for the electron channel analysis was derived from data by applying all standard electron requirements except for the track matching requirement and the electron likelihood cut, which was inverted ($\mathcal{L}_e < 0.85$). This introduces a kinematic bias coming from fake electrons which needs to be corrected.

This was corrected using the QCDWeight, defined below, which was determined using the matrix method [62].

$$QCDWeight = \varepsilon_{\text{fake-}e} / (1 - \varepsilon_{\text{fake-}e}), \quad (6.11)$$

where, $\varepsilon_{\text{fake-}e}$ is the efficiency of a ultraloose electron passing the tight electron requirements. The fake electron efficiency, $\varepsilon_{\text{fake-}e}$, was measured using an orthogonal sample of events that was selected by imposing the requirement of $\cancel{E}_T < 10$ GeV, thereby creating a sample that has a significant amount of multijet events. $\varepsilon_{\text{fake-}e}$ is given by the ratio of the number of tight electron events over the number of ultraloose electron events. However, this is assuming that there are no real electrons in this sample when indeed there are. These real electrons come mainly from Z +Jet events and some from W +Jet events. These events are measured from monte carlo and subtracted out before measuring $\varepsilon_{\text{fake-}e}$. It was found that $\varepsilon_{\text{fake-}e}$ has a dependence on lepton p_T [62]. Figure 6.1 shows this dependence where the figure on the left is $\varepsilon_{\text{fake-}e}$ and the one to the right of that is the correction factor called QCD weight. The other two plots are just a repeat of the first two, however the real electrons are subtracted out. Figure 6.2 shows the affect once the correction factor has been applied to each event in lepton p_T .

6.3 Residual Systematic Uncertainties

After making all of the above corrections, the residual systematic uncertainties can be divided into two categories: those which affect the normalization of the signal and background, and those that change the shapes of the distribution of those samples. These are discussed in the next two sections.

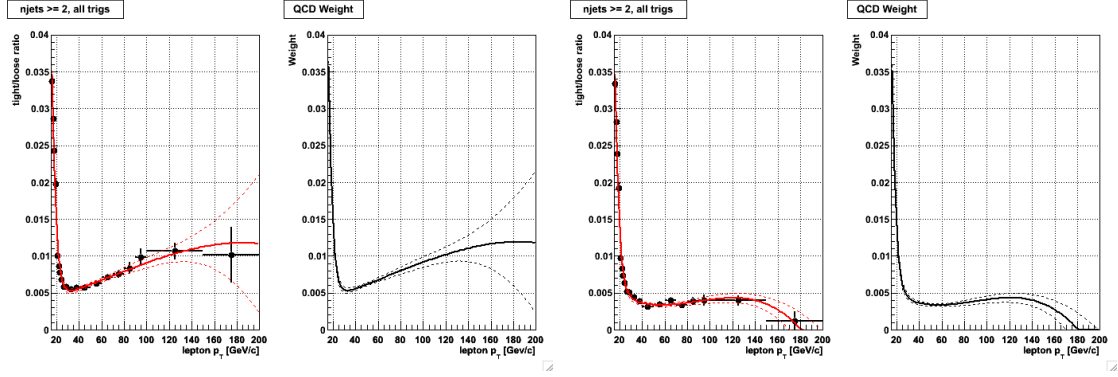


Figure 6.1: The two plots on the left are $\varepsilon_{\text{fake}-e}$ and $\text{QCDWeight} = \varepsilon_{\text{fake}-e}/(1 - \varepsilon_{\text{fake}-e})$. The two on the right are the same thing with the real electron contamination subtracted out.

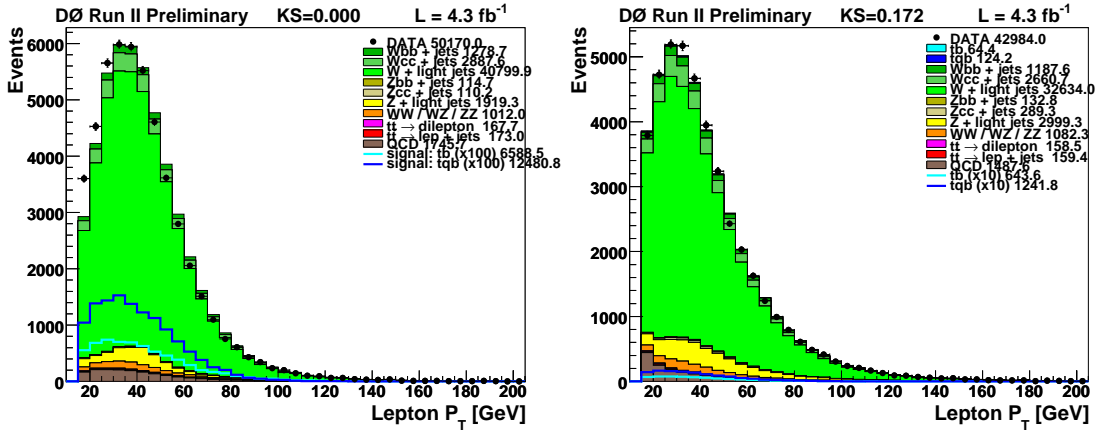


Figure 6.2: The plot on the left has no correction whereas the plot on the right has been corrected.

6.3.1 Normalization Effects

There are two luminosity uncertainties that affect this analysis. The calculation of the integrated luminosity has an uncertainty of $\pm 6.1\%$. This uncertainty was calculated using RunIIa data where the main contribution is due to the $\pm 5.4\%$ uncertainty in the knowledge of the inelastic $p\bar{p}$ cross section [63, 64]. The $\pm 6.1\%$ uncertainty is then obtained by adding in quadrature the $\pm 5.4\%$ uncertainty with the uncertainties which come from the modeling of the radiation damage to the luminosity monitor. Additionally there is an uncertainty of $\pm 1\%$, which originates in the reweighting process used to model the differences in the instantaneous luminosity within the Run IIa and IIb simulations.

The theoretical cross sections pertaining to this analysis also have an uncertainty. For signal and $t\bar{t}$ the uncertainties are calculated in Refs. [65, 66]. These are calculated by summing in quadrature the uncertainties due to the factorization scales, kinematics, PDF and the knowledge of the top quark mass. For a nominal top quark mass of 172.5 GeV the uncertainties are $\pm 3.8\%$ for the s channel, $\pm 5.3\%$ for the t channel, $\pm 4.8\%$ for the combined s and t channels and $^{+6.4\%}_{-9.0\%}$ for $t\bar{t}$ pair production. The $\pm 7.0\%$ uncertainties in the diboson (WW , WZ , ZZ) cross sections are taken from the values used by the $D\bar{O}$ Higgs group [42]. The Z +Jets cross section uncertainty is set to $\pm 3.3\%$ using with the results from reference [42].

The signal acceptance has an uncertainty of $\pm 2.0\%$ coming from the PDFs. This was calculated by reweighting monte carlo events using 40 different CTEQ PDFs and then calculating the acceptance with each PDF. The uncertainty was conservatively estimated by taking the maximum change in the acceptance between the 40 PDF results. The PDF uncertainties for $t\bar{t}$, diboson and Z +Jets are included in the theoretical cross section uncertainties for each.

The uncertainty on branching fractions for a W boson decaying to any given lepton is taken from the PDG results [1]. This gives an average uncertainty of $\pm 1.5\%$ which is added into the uncertainties for monte carlo normalization.

The trigger efficiency for this analysis is approximately $\sim 100\%$ with an uncertainty of $\pm 5.0\%$. The uncertainty was obtained by looking at the efficiencies of several variables which are sensitive to the triggers. Nearly all efficiencies lie between 95% and 105% and $\pm 5.0\%$ is taken as the representative uncertainty. The trigger efficiency plots used in this estimate can be seen in Ref. [42].

There are two different uncertainties attributed to the primary vertex, one of modeling and the other, selection. The primary vertex distribution in monte carlo is reweighted to match that of data [54]. This reweighting has a $\pm 0.05\%$ uncertainty. The selection efficiency for the primary vertex differs between monte carlo and data. The uncertainty in this difference gives an additional uncertainty of $\pm 1.4\%$ [54].

The uncertainties for the electron and muon scale factors have been determined by the $D\bar{O}$ electron and muon ID groups using tag and probe methods. For electrons the scale factor uncertainty is $\pm 4.8\%$ [67], and for muons the corresponding uncertainty is $\pm 2.1\%$ [42, 68].

The effects of modeling color reconnection has an uncertainty. This is assessed by comparing the results from the tune A Pro and tune AC Pro versions of the PYTHIA generator. Tune A Pro does not account for reconnection effects, whereas Tune AC Pro does try to model them. The difference in the two values for a top mass analysis is $\pm 1.0\%$, which is taken as the estimated uncertainty [69].

The uncertainties due to jet fragmentation and higher order affects were estimated by comparing the acceptances for $t\bar{t}$ events using different monte carlo generators. For this study, events from ALPGEN+PYTHIA were compared to those from ALPGEN+HERWIG. This results in uncertainties of $\pm 0.7\%$, $\pm 3.7\%$ and $\pm 4.7\%$ for the

two, three and four jet bins for the signal, Z +Jets and diboson samples. By including the effects of higher order QCD contributions an estimation of those effects can be obtained by comparing events from ALPGEN+HERWIG and MC@NLO+HERWIG. This gives additional uncertainties of $\pm 1.6\%$, $\pm 1.7\%$ and $\pm 7.0\%$ for the two, three and four jet bins, respectively. These are applied to the $t\bar{t}$ samples used in this analysis.

$t\bar{t}$ samples were also used to estimate the uncertainties in the modeling of the initial and final state radiation. This is done by generating a number of $t\bar{t}$ acceptances while changing the ISR/FSR settings in PYTHIA. These are then compared to the values generated with the nominal settings to measure the uncertainty. Taking into account jet multiplicity the uncertainty ranges from $\pm 0.8\%$ to $\pm 10.9\%$ for the two, three and four jet bins, respectively.

By taking into account what is known about b jet fragmentation, uncertainties can be estimated using $t\bar{t}$ simulations. This is done by changing the tune settings of the Bowler fragmentation function used by PYTHIA [70]. By comparing the results of the SLD and LEP tunes, the b fragmentation uncertainty was estimated to be $\pm 2.0\%$. This uncertainty is applied to the signal, $t\bar{t}$ and $Zb\bar{b}$ samples.

The uncertainty for the relative b /light jet response can be measured by first taking the double ratio of the b and light jets p_{Ts} (measured from response) from a simulated $t\bar{t}$ sample, given by:

$$\frac{\left(\frac{p_T^{Data}}{p_T^{MC}}\right)_{b-jet}}{\left(\frac{p_T^{Data}}{p_T^{MC}}\right)_{light-jet}}. \quad (6.12)$$

This results in a difference of $\pm 1.8\%$ [71]. The b jets response is then shifted down by 1.8% to correct for this. The corresponding uncertainties are taken from the $D\bar{O}$ $t\bar{t}$ cross section analysis [72]. This gives uncertainties of $\pm 0.4\%$ (2 or 3 jet events) and $\pm 0.3\%$ (4 jet events). For single top, the corresponding uncertainties were measured to be $\pm 0.5\%$, $\pm 0.8\%$ and $\pm 0.9\%$ for two, three and four jet events, respectively.

Lastly, the $Zb\bar{b}$ sample has uncertainties of $\pm 1.0\%$, $\pm 0.6\%$ and $\pm 0.8\%$ for two, three and four jet topologies.

The uncertainty on the taggability efficiencies are measured by fluctuating the scale factor up and down by one sigma for each monte carlo sample. The whole analysis is repeated for both up and down samples and uncertainties range from $\pm 3.1\%$ to $\pm 21.5\%$ [42], where the range corresponds to one tag 2, 3 and 4 jet samples and two tag 2, 3 and 4 jet samples.

The W +Jets heavy scale factor (λ_{HF}) is used as an additional correction to the $Wb\bar{b}$ and $Wc\bar{c}$ monte carlo samples. For this analysis it is measured with an uncertainty of $\pm 12.0\%$ [42]. For a detailed discussion please see Ref. [42]. The uncertainty on the Z +Jets heavy scale factor is measured in a similar fashion to that of W +Jets and is also found to be $\pm 12.0\%$.

The uncertainty coming from both W +Jet and QCD multijet normalizations is calculated by comparing the final scale factors to ones that vary the most. This is done separatley for each variable in both event samples. The resulting uncertainty in the W +jets normalization is $\pm 1.8\%$ and for the QCD multijet normalization is between 30 and 40%.

Sample statistics uncertainty is due to the finite size of both the data and monte carlo samples and is given by $\frac{1}{\sqrt{N}}$. This uncertainty is considered in the final discriminant output for each background sample.

6.3.2 Shape-changing Effects

The vertex confirmation (VC), jet energy scale (JES), jet energy resolution (JER), b tagging (BTag) and angular corrections (AC) result in simulation changes which not only change the sample yield (normaliztion), but also change the shape of some of

the kinematic distributions. The uncertainties due to these are estimated by running the whole analysis for each correction, producing two samples for each, one that fluctuates the corrections up one sigma and another down one sigma. The uncertainty is taken as the difference in fractional yield between each of these and the nominal simulated sample. The jet reconstruction and identification (JetID) shape uncertainty is estimated by first not applying the jet ID correction factor. This produces a systematic sample which is compared to the nominal sample in an MVA output. The difference between the two is then cut in half where $JetID_{up}$ is the nominal plus one half and $JetID_{down}$ is the nominal minus one half.

The b tagging uncertainty is estimated from the uncertainty in the taggability and tagging scale factors added in quadrature. The origin of these uncertainties is the limited statistics of the event samples used to calculate the scale factors. The uncertainty for single tagged events ranges from $\pm 4.3\%$ to $\pm 14.0\%$ for 2, 3 and 4 jets, and for double tagged events $\pm 5.8\%$ to $\pm 11.2\%$ for 2, 3 and 4 jets.

The V +Jets angular correction procedure preserves the normalization of the W and Z +Jets pretag samples. By b tagging these samples a normalization uncertainty is introduced through the b tagging scale factors, giving a normalization uncertainty of $\pm 0.3\%$.

Uncertainties for jet reconstruction and identification originate from the dependence on the primary vertex, luminosity and the calculated jet efficiencies. These combine to give a normalization uncertainty of between $\pm 0.04\%$ and $\pm 3.7\%$, where the range corresponds to the 2, 3 and 4 jet samples. The uncertainty coming from the JER corrections comes mainly from fits used to measure the resolution and dependence on the luminosity [73]. The normalization uncertainty here ranges from $\pm 0.2\%$ to $\pm 11.6\%$, where the range corresponds to the 2, 3 and 4 jet samples.

The JES corrections introduce an uncertainty of between $\pm 0.3\%$ and $\pm 14.6\%$,

where the uncertainty comes from the uncertainties in all of its correction factors [74]

The final uncertainty considered here is that introduced by the vertex confirmation. The precision of this correction is limited by the size of the samples used in its calculation. For instance, very forward jets are complicated because there is almost no tracker and only a few jets can be confirmed, which makes it harder to derive a correction, and consequently an increased uncertainty. These uncertainties lie between $\pm 0.1\%$ and $\pm 9.6\%$ which corresponds to the position of the jet in the detector.

6.3.3 Summary of Residual Uncertainties

Table 6.1 presents an overall summary of the residual uncertainties. Detailed lists of the source and its uncertainties as a function of tag and jet multiplicity can be seen in the tables of Appendix B. The second part of Appendix B presents several plots showing the shape changing effects of these uncertainties that are present in the BDT discriminants used for the s , t and the combined $s + t$ channel analyses.

Source	Uncertainty %
Normalization Uncertainties	
Integrated luminosity	6.1
$t\bar{t}$ cross section	9.0
Z +jets cross section	3.3
WW , WZ , ZZ , cross sections	7.0
Branching fractions	1.5
Parton distribution functions	2.0
(signal acceptances only)	
Triggers	5.0
Instantaneous luminosity reweighting	1.0
Primary vertex selection	1.4
Color reconnection	1.0
b /light jet response	(0.3-1.0)
Electron identification	(2.8-3.8)
Muon identification	2.1
Jet fragmentation and higher order effects	(0.7-7.0)
ISR and FSR corrections	(0.8-10.9)
b -jet fragmentation	2.0
Taggability	(3.1-21.5)
W +jets heavy-flavor correction	12.0
Z +jets heavy-flavor correction	12.0
W +jets normalization	1.8
QCD multijet normalization	(30-40)
MC and multijets statistics	(0.2-16)
Normalization and Shape Uncertainties	
Jet reconstruction and identification	(0.04-3.7)
JER corrections	(0.2-11.6)
JES corrections	(0.3-14.6)
Vertex confirmation	(0.1-9.6)
b tagging (single tag)	(4.3-14.0)
b tagging (double tag)	(5.8-11.2)
Angular correction	0.3

Table 6.1: Residual systematic uncertainties after improvements to detector and process simulations. Where ranges are given, these correspond to 2, 3 and 4 jet events.

Chapter 7

Boosted Decision Trees

7.1 Introduction

Once the selection and b tagging requirements have been applied (Chapter 5) the signal to background ratio is around 1:20. In a traditional analysis further cuts are applied to obtain a suitable S:B ratio (cut based analysis). However in this case this proves ineffective as the excess of signal is smaller than the predicted background uncertainty. Also, signal is lost when additional cuts are applied making the cut based approach impractical for a singletop cross section measurement. In this analysis a multivariate analysis (MVA) technique called Boosted Decision Trees (BDT) is used for the analyses optimization. BDTs take as input selected variables for each event and provide an output discriminant between 0 and 1, where events closer to 0 tend to be background whereas events closer to 1 tend to be signal.

Historically decision trees first came to light within the statistics community in 1963 [75] and later became more popular in the 1980s as data mining and machine learning algorithms [76]. It was not until the 1990s when more sophisticated decision tree algorithms were developed such as boosted decision trees [77]. Boosted deci-

sion trees were first used in the high energy physics community by the MiniBooNE experiment [78, 79].

Figure 7.1 shows an example of how a decision tree is created. It starts off with an initial node which applies a cut criteria where events pass if their $H_T < 242$ GeV and fail if $H_T \geq 242$ GeV. The cut is chosen such that it maximizes the separation between signal and background events in the initial data sample. This cut creates two daughter nodes where new criteria can be applied to separate signal from background for these nodes. The process is continued until certain criteria are not met which turns the node into a terminal leaf. These criteria will be discussed later on in this chapter. The leaf then represents the discriminant output value $D(\vec{x}_i)$ where \vec{x}_i corresponds to the variables used to obtain that value. This is also referred to as the purity which is defined as:

$$p = \frac{\sum_i s_i}{\sum_i s_i + b_i} = \frac{s_{leaf}}{s_{leaf} + b_{leaf}}, \quad (7.1)$$

where s_i and b_i are the weighted sums of signal and background events for the leaf. This whole process is continued until all nodes become terminal leafs, thereby creating a tree. This process is also called decision tree training.

The training of a single tree does not provide enough separation power between signal and background for this analysis, however this can be enhanced by using the technique called boosting. Boosting basically creates a forest of trees by reweighting the misclassified events within the tree. All of the trees created by boosting are then averaged for the final output. This recovers misclassified events and gives a smoother approximation to the final discriminant output. It is discussed in more detail in Section 7.3.2.

Some of the advantages for using decision trees in an analysis are that events which would be rejected in a cut based analysis are still retained. BDTs make it easy to

identify why an event gets labeled as signal or background and their training is much faster than a neural network. Another advantage is that discrete variables can be used directly and the performance of the decision tree is not really affected by the addition or removal of variables. Some disadvantages are the discrete values of the discriminant outputs and that trees can become unstable relative to the training sample used if the input variables are poorly chosen. Both disadvantages are minimized by using the boosting algorithm which takes an average of all of the trees, thereby smoothing the discriminant output.

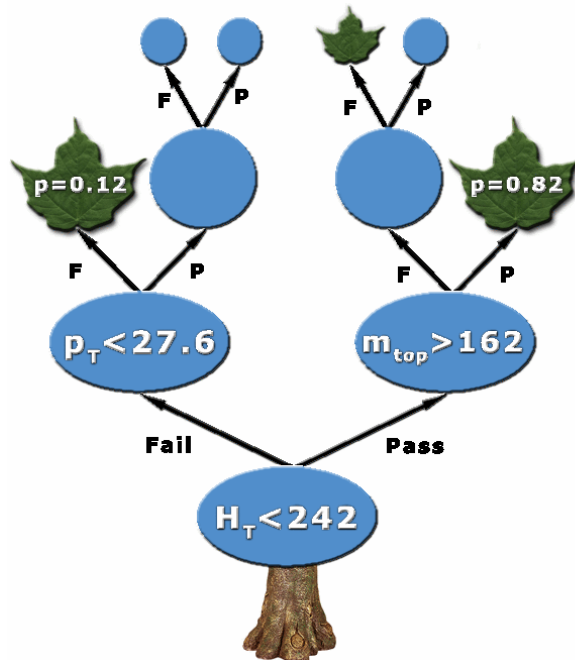


Figure 7.1: Schematic of a decision tree showing the effects of branching and the terminal leaves. The blue circles represent nodes which are split into daughter nodes. Each daughter node is subjected to an independent test to see whether it will continue to be split or end as a terminal leaf (green). Here F and P stand for fail and pass the nodal test.

7.2 Sample Preparation

Three independent monte carlo subsamples are used for this analysis. These are created by dividing the MC signal and background samples into three separate data sets. The first of these is called the “training sample” and is used to train the BDTs. The second is the “testing sample” which is used to train the MVAs used to combine different BDTs. The third is called the “yield sample” and is used to calculate all final cross section measurements. The data sample is unchanged and the simulated samples are the only ones used in this way. The monte carlo datasets are split such that the first event goes to the testing sample, the second the training sample and the third and fourth are taken by the yield sample. This process is iterated until all of the available events have been assigned. In total one fourth of all the MC samples go to the training, one fourth to the testing and one half to the yield samples. Each sample is then normalized back to the nominal luminosities set by multiplying the training and testing samples by four and the yield sample by two.

7.3 Training

By using the training sample we can define an algorithm which describes how a decision tree is built. Initially signal and background samples are normalized relative to one another taking account of the normalization scales and correction factors discussed in Chapter 6. A node is then produced with both signal and background events. A list of variables is then read into the BDT code where the variables have been chosen to provide the best discrimination between signal and background. If no splitting values are found that provide sufficient signal:background separation then the node is defined to be a terminal leaf. If the data at a node can be split, then one

daughter node contains the events that pass the cut and one that fail the cut. If the number of events in either daughter node falls below the minimum required, then the node becomes a terminal leaf. This process is repeated many times until all of the remaining daughter nodes have become terminal leaves. Each leaf is then assigned a purity value using equation 7.1. This is shown schematically in Figure 7.1.

7.3.1 Splitting a Node

The splitting of a node into two daughters is by far the most significant and powerful part in the training of a decision tree. This is done by first measuring the sample impurity at the node. For this analysis the Gini Index [80] was used for this where:

$$G = \frac{sb}{s+b}. \quad (7.2)$$

Similar to the purity measure, s and b are the weighted sum of both signal and background events for each node. Other impurity indexes are available, however it has been shown that the Gini Index provides the best performance for this type of analysis [80]. The best separation corresponds to the largest change of impurity ΔG which is defined by:

$$\Delta G = G - G_P - G_F = \frac{sb}{s+b} - \frac{s_P b_P}{s_P + b_P} - \frac{s_F b_F}{s_F + b_F}, \quad (7.3)$$

where the subscripts P and F represent pass and fail for the daughter nodes created by the split. A cut value is chosen such that it maximizes ΔG , thereby providing the best separation between signal and background. This is done initially by reading in the list of variables and sorting the events in the node relative to the variable being used. ΔG is then calculated by going through each event in order and the highest

ΔG is recorded. The algorithm moves to the next variable in the input list and the process is repeated until the variable list is exhausted. The best ΔG is used to define the nodal split.

7.3.2 Boosting

Boosting is used in this analysis to overcome some of the limitations previously stated. It has also been shown to increase the overall performance by 20% [80]. The basic idea behind boosting is to create an initial decision tree and then identify the misclassified events within the tree. Misclassified events are then reweighted and a new decision tree is created. This process is repeated until 30 trees are produced. Thirty represents the number of boosting cycles and was chosen because it gives the overall best performance. The trees are then averaged to provide the final discriminant output. Several different types of boosting algorithms are available in the literature, however for this analysis the AdaBoost algorithm was chosen [77].

The AdaBoost algorithm works by first creating an initial decision tree. The discriminant output of that tree is then used to find misclassified events by using the relation $|D(\vec{x}_i) - z_i| > 0.5$ where z_i is 1 (0) for signal (background). A misclassification error (ϵ_n) is then calculated for the discriminant output as a weighted fraction of the misclassified events. Using this error a tree weight is defined by:

$$\alpha_n = \beta \times \ln \frac{1 - \epsilon_n}{\epsilon_n}, \quad (7.4)$$

where β is known as the boosting parameter. Each of the misclassified events are then reweighted by e^{α_n} (which is always greater than one). A new decision tree is then trained focusing on the misclassified events. This process is repeated for 30 cycles, producing thirty decision trees. The decision trees are then averaged to obtain the

final discriminant $D(x_i)$:

$$D(x_i) = \frac{1}{\sum_{n=0}^{30} \alpha_n} \sum_{n=0}^{30} \alpha_n D_n(x_i). \quad (7.5)$$

Two figures of merit are then used to evaluate the performance of the BDT. These are the separation performance and the cross section expected significance which are defined as:

$$\sum_{i=0}^N \frac{(s_i - b_i)^2}{s_i + b_i}, \quad (7.6)$$

$$\sum_{i=0}^N \frac{s_i^2}{s_i + b_i}. \quad (7.7)$$

Figure 7.2 shows these for one of the analysis channels as a function of the number of cycles (trees). From this it is clear that the results have converged for ~ 30 cycles.

7.3.3 Parameters

There are several BDT parameters that can be tuned and which influence the decision tree in the training process. The principal parameters used in this analysis are the : number of boosts, minimal leaf size, impurity measure and the AdaBoost parameter. The first corresponds to the number of trees used in the boosting portion of the training and is set to 30. The second represents the minimal number of events which turns into a terminal leaf. This is set to 100 events. The Gini Index is used for the impurity measure and the final parameter is the weight given to events that are misclassified. This is set to $\beta = 0.2$

In the previous $D\bar{O}$ analysis an optimization study was done to find the optimal settings for each of the parameters [80]. From this it was determined that the decision tree performance changed very little as the parameters were varied. The only excep-

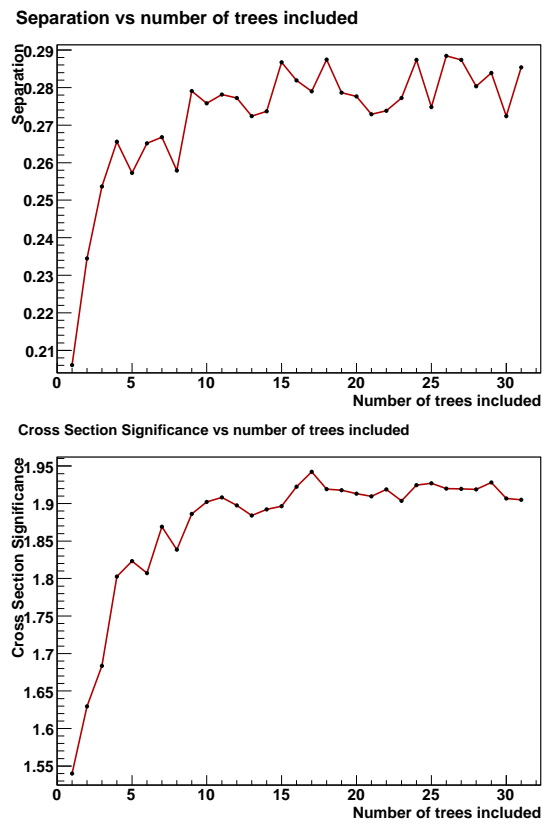


Figure 7.2: Example plots showing separation power and cross section significance versus the number of boosts.

tion was the effect of changing the number of boosting cycles. Because of this a new study was performed for this analysis and it was found that 30 boosting cycles showed optimal performance. The other parameters values were set to the same values as were used in the previous $D\bar{O}$ analysis [80].

7.3.4 Implementation

Decision trees are created and evaluated using the code package classifier located in the $D\bar{O}$ CVS code repository. Previous $D\bar{O}$ Single Top Groups have also used this code resulting in evidence and observation publications [80, 81]. The differences with the code used for this comes mainly from the improvements and options which were added to cope with large output files and the production of extensive plot lists for each channel.

7.3.5 Selection of Input Variables

Selecting the input variables is a very important step in the definition of a BDT. The variables must be chosen so as to effectively separate signal and background. If the list of variables is too small, then the outcome is reflected in poor separation power. On the other hand the addition of extra variables may have a negligible effect, unless they have significant discrimination power. The addition of variables also makes the training process more complex and slows down the construction of decision trees. Because of this it is important to optimize both the number and the choice of variables in order to minimize the complexity of the decision tree process. Fifty input variables were chosen and used in the creation of decision trees for this analysis. To reduce the added complexity, these same fifty variables are used for each of the six 1 tag \times 2,3 and 4 jet and 2 tag \times 2,3 and 4 jet channels.

The variable selection is divided into two steps. The first uses the ROOT implementation of the Kolmogorov-Smirnov (KS). This compares the data and simulated distributions for any variable chosen by applying a quality test to determine if they come from the same underlying distribution. The test gives a value between 0 and 1 where values closer to 0 suggest that the MC and data do not come from the same underlying function and 1 suggests that they do and are typically well modeled. The second step uses the BDT ranking method where each variable is assigned a ΔG value. This is used to select the variables which provide the best separation between signal and background. The combination of both these steps result in the list of fifty input variables which were chosen for the analysis.

The first step is implemented by considering the KS value for all variables in each of the six analysis channels. The KS values for each variable are added and then divided by six to give the average for each variable. Only variables having an average KS value ≥ 0.30 are kept for further analysis and individual KS values < 0.01 are rejected. This step provides a list of 136 variables which passes to the final BDT ranking.

For the second step the list of 136 variables was used as inputs to train six BDTs, one for each channel. Each BDT yields 30 trees, resulting in a ΔG sum over 180 calculations, per variable. The ΔG s were then added and averaged and the fifty variables with the highest ΔG were retained for the analysis. The data/MC distributions for the fifty selected variables are shown in Figs 7.3-7.6 and they are listed in Tables 7.1- 7.7 where they have been grouped into topological categories. A brief discussion on the nomenclature of the variables is given below.

Jets are indexed by the order of their p_T . For instance, jet1 has the highest p_T , jet2 the second highest. The label “best” jet means the invariant mass $M(W, \text{jet})$ is the jet that gives a top quark mass closest to 172.5 GeV. Each event has one best

jet where the others are labeled “not best”. For once tagged samples, labels such as “leading tagged jet” and “most b like jet” correspond to jets passing the b tagging requirements, also the most b like jet is a jet with the highest BNN output. For twice tagged samples, labels such as “leading tagged jet” and “second tagged jet” refer to the order of their p_T as well as passing the b tagging requirements. Labels such as “leading light quark jet” and “second leading light quark jet” refer to jets which are not b tagged and are also ordered by p_T . Both the top quark and W boson are reconstructed for many of the variables. The top quark is reconstructed using the combined four vectors of the lepton, jet and neutrino. The W boson is reconstructed by using the lepton four momentum and the reconstructed neutrino p_z . The neutrino p_z is unknown in the detector, however it can be estimated using the lepton four momentum and the mass of the W . The p_z solution found is quadratic yielding two solutions. The first solution is labeled S1 and is chosen as the lower of the two values. The second solution is labeled S2. Typically S1 is used.

Kinematic Variables
Leading Light Quark Jet E: Leading light quark jet energy
η_{det} (jet1): η_{det} fo the most b like jet
η width ($jet_{most\ b\ like}$): Width of the most b-like jet in η
η width ($jet_{not\ best\ 1}$): Not the best leading jet width in η
p_T(lepton): p_T of the lepton
$p_T(jet_{best})$: p_T of the best jet
B tag $NN_{Best\ Jet}$: NN output of the best jet
Lepton ϕ_{det} direction: ϕ_{det} of the lepton
Neutrino p_z: first solution for p_z of the neutrino (S1)
Not Best 1 ϕ width: Not the best leading jet width in ϕ
Not Best 2 ϕ width: Not the best 2 nd leading jet width in ϕ

Table 7.1: List of individual object kinematics.

Event Variables
Run Period: RunIIa vs. RunIIb
Centrality (all jets): $\frac{\sum H_T}{\sum E}$
Missing E_T: p_T of the \cancel{E}_T
Missing E_T phi direction: ϕ direction of the \cancel{E}_T
Primary vertex z: Primary vertex position of z

Table 7.2: List of event variables.

Charge of the lepton times η
$Q \times \eta$: of the leading light jet
$Q(lepton) \times \eta(lepton)$: of the lepton
$Q(lepton) \times \eta(best1)$: of the most b like jet
Jet1 $\eta_{jet} \times Q_{lepton}$: of the leading jet
Not Best 1 $\eta_{jet} \times Q_{lepton}$: of not the best leading jet

Table 7.3: List of variables using the sign of lepton charge.

Reconstructed Objects
m (jet1,jet2): Invariant mass of the object assembled from the first two leading jets
m (jet1_{lq1},jet2_{lq2}): Invariant mass of the object assembled from the first two leading light quark jets
m (alljets – jet_{tag1}): Invariant mass of the object assembled from all jets minus the leading <i>b</i> tagged jet
m (alljets – jet_{best}): Invariant mass of the object assembled from all jets minus the best jet
M(all jets): Invariant mass of the object assembled from all jets
p_T(jet1,jet2): <i>p_T</i> of the object assembled from the first two leading jets
H_T(jets,l,ν): <i>H_T</i> of all jets, lepton and \cancel{E}_T
H_T All Jets-Best Jet: <i>H_T</i> of all jets minus the best jet
H_T MET, Lepton: <i>H_T</i> of the lepton and \cancel{E}_T

Table 7.4: Calculated mass, *p_T* and the *H_T* variables.

Reconstructed Top
χ²(<i>m_{top}</i>): (Reconstructed top mass - the MC top mass) ² which is the smallest for all reconstructed top quarks
χ² Top Mass: Reconstructed top mass that has the nearest mass to theory
significance_{min}(<i>m_{top}</i>): Smallest significance of the top quark mass out of all the reconstructed top quarks
Jet1 Top Mass Solution 2: Invariant mass of the reconstructed top quark. The top quark was reconstructed from the leading jet and the <i>W</i> from the second neutrino solution
m(W,jet_{blike}): Invariant mass of the reconstructed top quark. The top quark was reconstructed from the most <i>b</i> like jet
m(W,jet_{blike},S2): Invariant mass of the reconstructed top quark. The top quark was reconstructed from the most <i>b</i> like jet and the <i>W</i> from S2
M(l,ν,b): Invariant mass of the reconstructed top quark. The top quark was reconstructed from the leading <i>b</i> tagged jet

Table 7.5: Reconstructed top quark variables.

Reconstructed W
Shericity (W, all jets): Shericity of all jets and the W
Aplanarity (all jets, W): Aplanarity of all jets and the W
$p_T(W)$: p_T of the reconstructed W
$M_T(W)$: Transverse mass of the reconstructed W were $m_t = m^2 + p_x^2 + p_y^2$

Table 7.6: Reconstructed W boson variables.

Angular Correlations
Delta R all jets: minimum ΔR between any two jets
Delta ϕ Lepton MET: $\Delta\phi$ between the lepton and \cancel{E}_T
Delta ϕ: $\Delta\phi$ between jet2 and \cancel{E}_T
Most B like jet Delta ϕ: $\Delta\phi$ between lepton and the most b like jet
cos Best Lepton in CMF B-Tag: Cosine between the lepton in the b tagged top frame and the b tagged top in the center of mass frame. Here the b tagged top is assembled from the leading b tagged jet
cos Best Lepton in CMF: Cosine between the lepton in the best top frame and the best top in the center of mass frame. Here the b tagged top is assembled from the leading b tagged jet
cos Light quark Jet B-Tag: Cosine between the leading light quark jet and the lepton in b tagged top frame
cos(tag1,lepton)_{btaggedtop}: Cosine between the leading tagged jet and the lepton in the b tagged top frame
cos Lepton, $Q \times Z$ Best top: Cosine between the lepton and the z axis $\times Q(lepton)$ in the best top frame

Table 7.7: Angular Correlations.

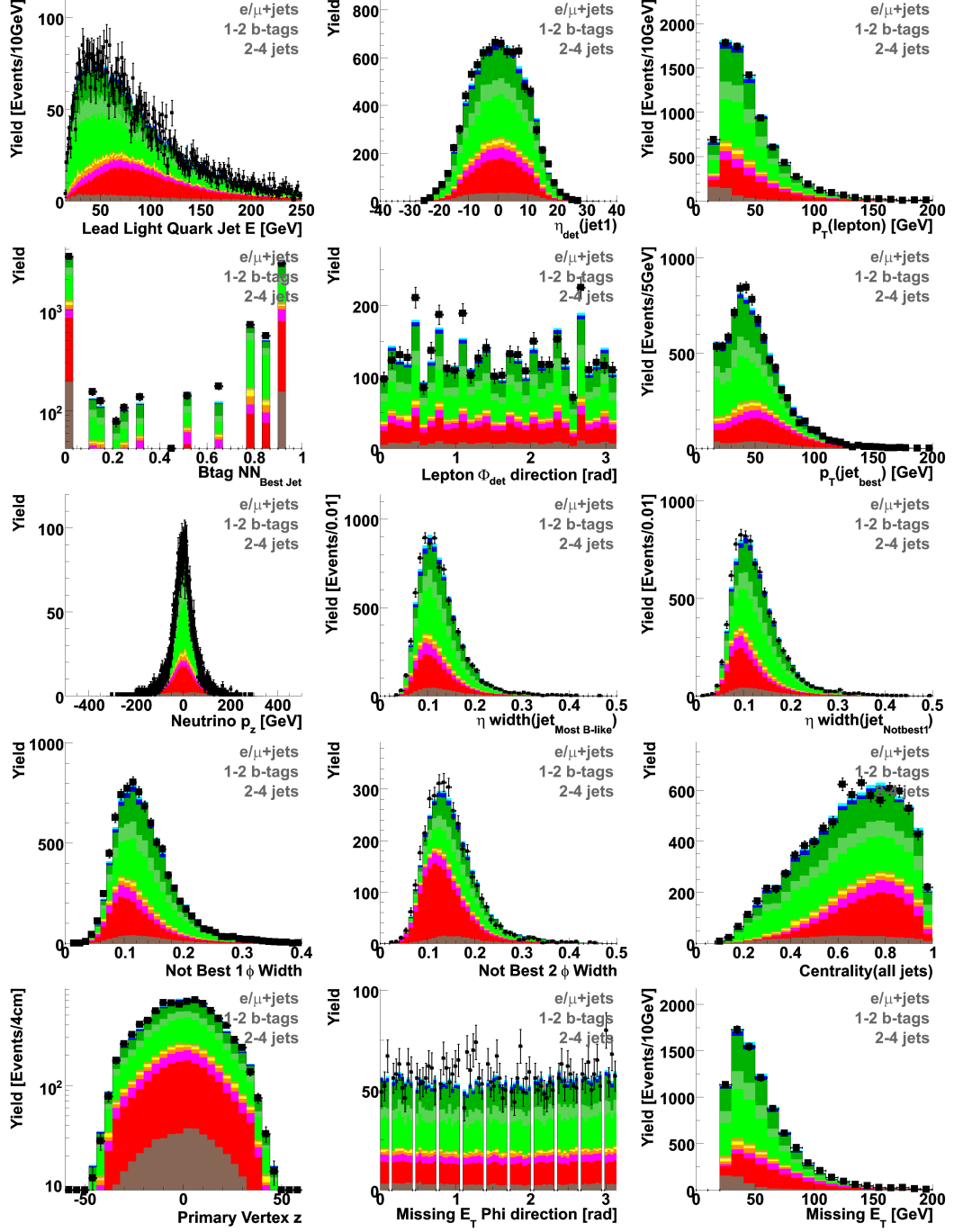


Figure 7.3: Data-Simulation comparison for selected variables I.

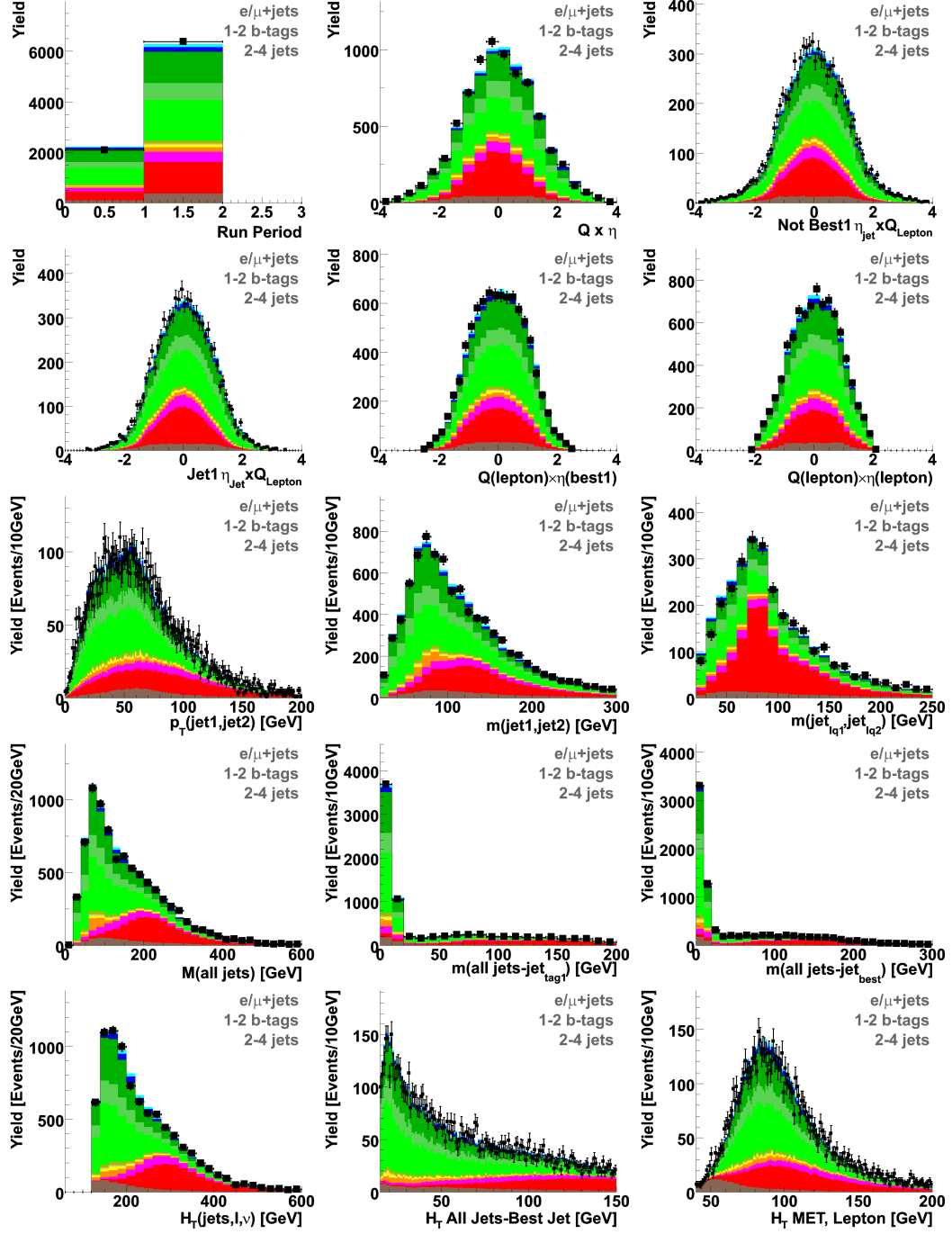


Figure 7.4: Data-Simulation comparison for selected variables II.

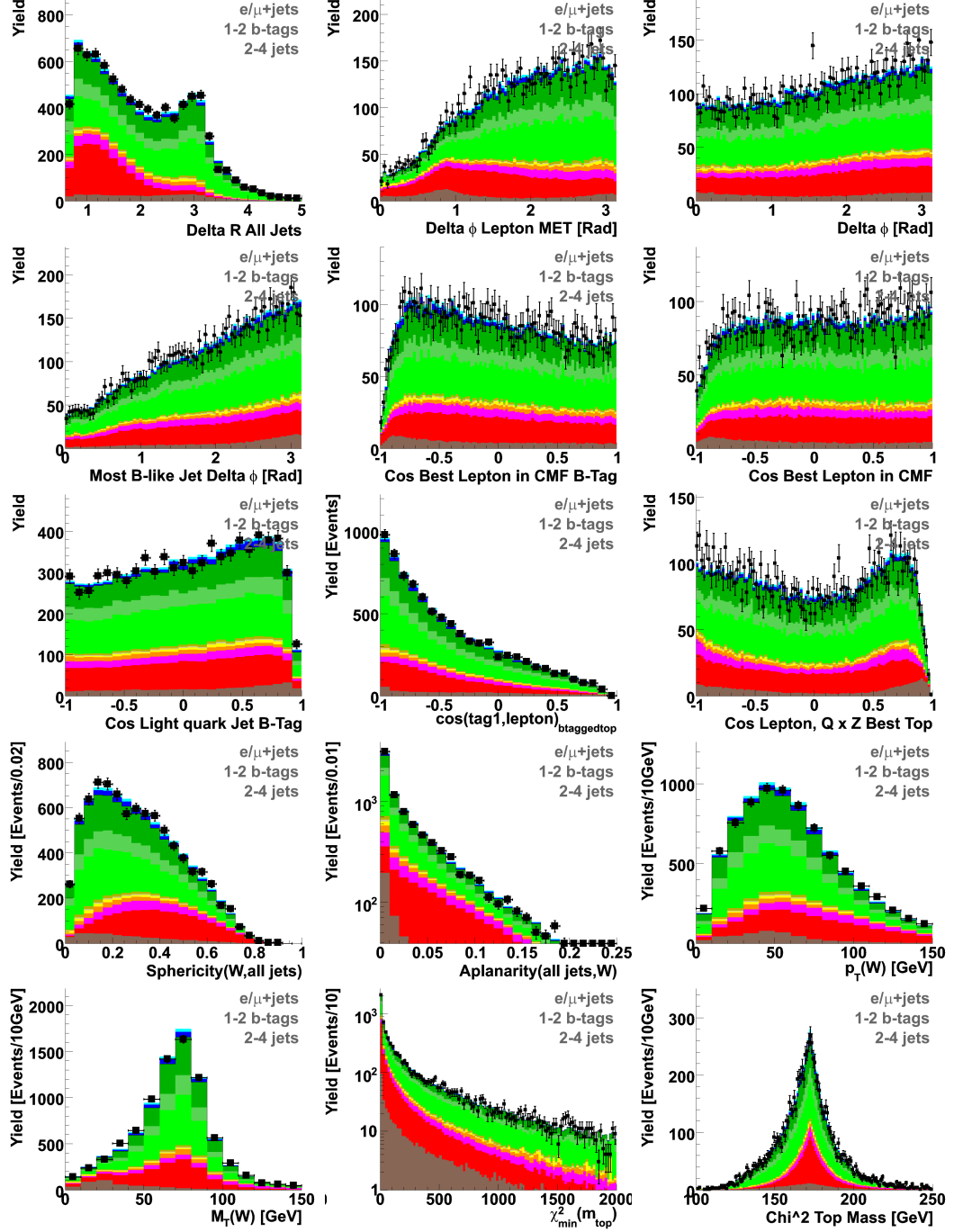


Figure 7.5: Data-Simulation comparison for selected variables III.

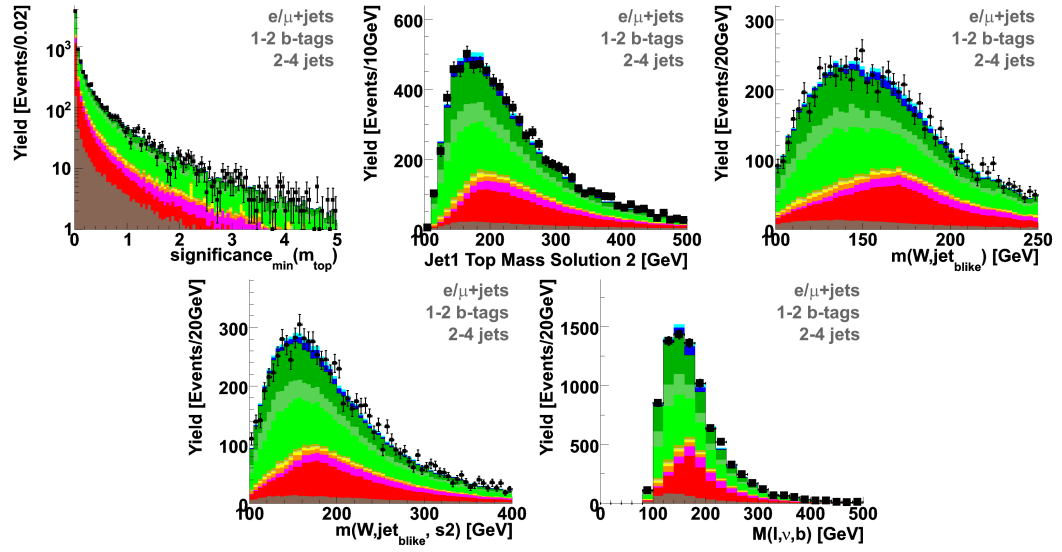


Figure 7.6: Data-Simulation comparison for selected variables IV.

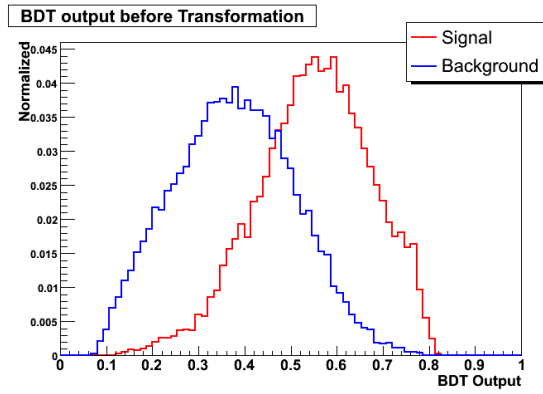
7.4 Transformation of the BDT Discriminant

The plot on the top left of Figure 7.7 shows a BDT output histogram before the transformation. It is clear by looking at this plot that there are very few events if any at all which populate the regions close to 0 and 1. This plot also shows that the signal and background distributions are centralized around 0.5. This is the result of averaging over the 30 BDTs and it causes instabilities in cross section measurements because the large statistical uncertainties in weakly populated bins. To deal with this the BDT discriminants are transformed by shaping the background as $1/x$. The transformation process is constrained such that a minimum of 64 effective background events will be in each of the 25 bins for the final BDT discriminant.

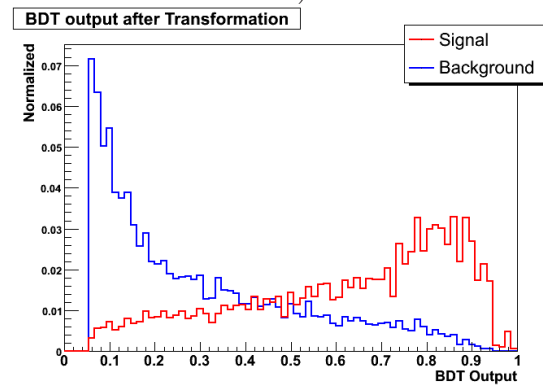
The function used for the transformation is shown in the bottom plot of Figure 7.7 and is defined as follows:

$$y = \begin{cases} \frac{k}{x}, & x \leq 0.8 \\ M - Kx, & 0.8 < x < 0.95 \\ 1, & x \geq 0.95, \end{cases} \quad (7.8)$$

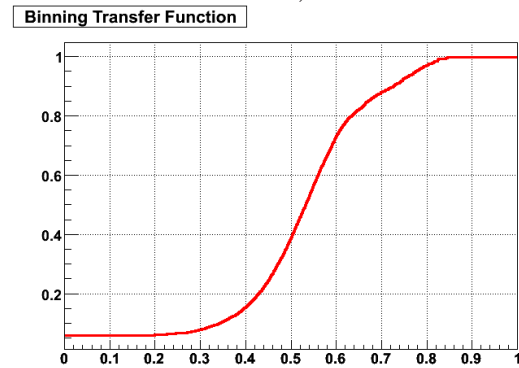
where k , K and M are 0.346, 2.88 and 0.274. In the region from 0 to 0.8 the function is defined by k/x . In the 0.8 to 0.95 region the function is linear, and is 1 for the region ≥ 0.95 . The constants are assigned a value by making sure the transformation function is continuous and by preserving normalization. The top right plot of Figure 7.7 shows the BDT discriminant output after the transformation function is applied.



a.)



b.)



c.)

Figure 7.7: Discriminant Transformation Histogram a.) is the raw distribution, b.) is the transformed distribution and c.) is the transformation function.

7.5 Background Cross Checks

Two cross check samples are created for this analysis in order to check the modelling of the background. This is done for the W +Jets and $t\bar{t}$ backgrounds. To ensure that the discriminant outputs are dominated by these contributions additional selection cuts are added. For the W +Jets check, $H_T < 175$ GeV cut is applied and only events with two jets are considered. For the $t\bar{t}$ check, a $H_T > 300$ GeV cut is applied and only events with four jets are considered. Figures 7.8 and 7.9 shows cross checks for the $s+t$, s and t channel analyses. The first two rows are the cross checks for $s+t$, the second two for s and the last two for t . Red indicates the $t\bar{t}$ contribution and green indicates the W +Jets contribution. The last column for all three channels represent the first two columns combined. These plots show good agreement between data and simulation ensuring that the two main background for this analysis are well modeled.

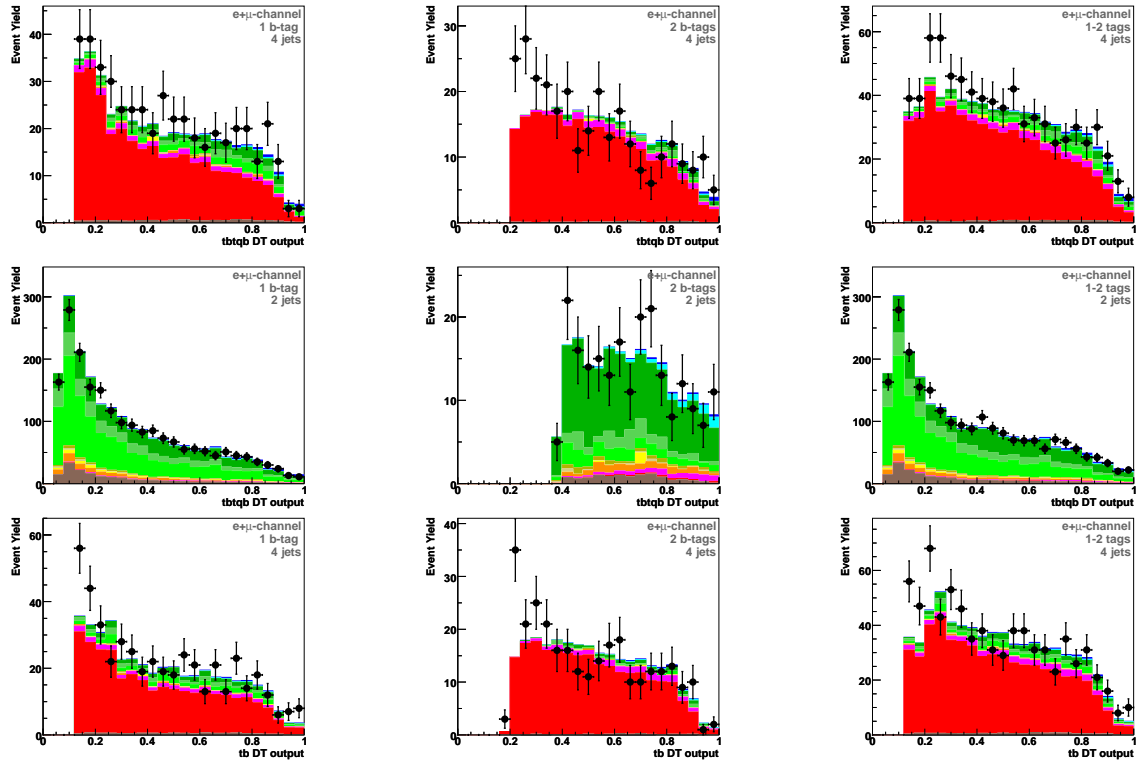


Figure 7.8: BDT cross check samples for $s+t$, s and t where the histograms dominated by red correspond to $t\bar{t}$ events and the green dominated ones indicate W +Jet events.

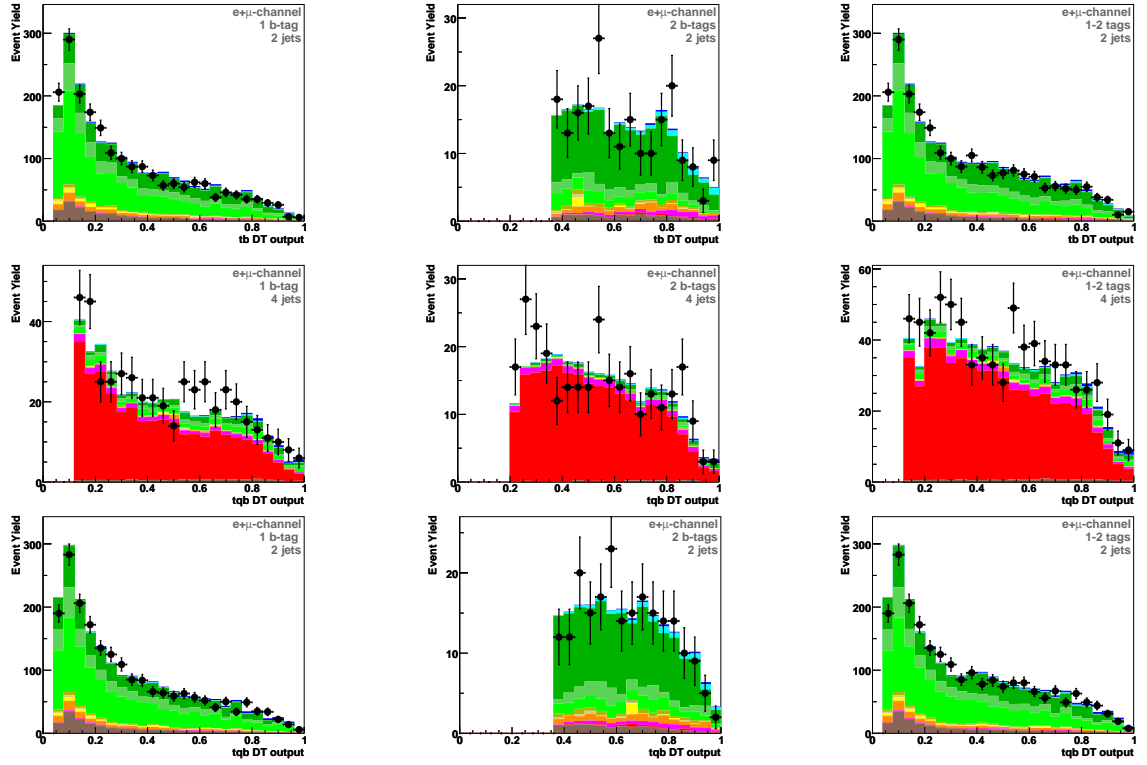


Figure 7.9: BDT cross check samples for $s+t$, s and t where the histograms dominated by red correspond to $t\bar{t}$ events and the green dominated ones indicate W +Jet events.

Chapter 8

Analysis

This chapter gives an overview of this analysis starting with a discussion of how the cross section measurements are made. This is followed by a discussion of systematic uncertainties, results, significance and ends with measurement of the CKM matrix element $|V_{tb}|$.

8.1 Cross Section Measurement

The measurement of the single top production cross sections is calculated using the sets of six BDT discriminant outputs for each of the s channel, the t channel (see Appendix A). For example, for the $s+t$ channel analysis we use six BDT outputs, each of which contains a 25 bin histogram giving a total of 150 independent $s + t$ channel measurements. All of the cross section measurements and uncertainties are calculated using the `top_statistics` package [82, 83]. This employs both the frequentist and bayesian approach, however for this analysis the latter is used. The package is also used to validate the results which will be discussed in Section 8.3.

8.1.1 Bayesian Approach

In order to make a cross section measurement using the bayesian approach we initially calculate the mean count d of both signal and background together. This is done for each individual bin and defined by:

$$d = s + b = \alpha \mathcal{L} \sigma + \sum_{i=1}^N b_i \equiv a\sigma + \sum_{i=1}^N b_i, \quad (8.1)$$

where signal $s = \alpha \mathcal{L} \sigma$. Here α is the signal acceptance, \mathcal{L} is the integrated luminosity, σ is the cross section for the signal and a is defined as the effective luminosity $\alpha \mathcal{L}$. The next step is to calculate the likelihood of observing a count D given the mean count d . This is done using a Poisson distribution where the likelihood in each bin is defined as:

$$L(D|d) = \frac{e^{-d} d^D}{\Gamma(D+1)}, \quad (8.2)$$

where Γ is a Gamma function and is equal to $D!$ when dealing with integer counts of D as is the case for data. Since we are dealing with 150 bins we can build a total likelihood by considering all bins as a product of each of the binned likelihoods:

$$L(\vec{D}|\vec{d}) \equiv L(\vec{D}|\sigma, \vec{a}, \vec{b}) = \prod_{j=1}^{150} L(D_j|d_j), \quad (8.3)$$

where \vec{D} represents the vector for observed counts, \vec{d} the vector for mean counts, and \vec{a} and \vec{b} the vectors for the effective luminosity and background.

The next step in the measurement is to use Baye's theorem to calculate the posterior probability density $p(\sigma, \vec{a}, \vec{b}|\vec{D})$. This is done by integrating over vectors \vec{a} and

\vec{b} to get the cross section posterior density given \vec{D} which can be expressed as:

$$p(\sigma|\vec{D}) = \frac{1}{\mathcal{N}} \int \int L(\vec{D}|\sigma, \vec{a}, \vec{b}) \pi(\sigma, \vec{a}, \vec{b}) d\vec{a} d\vec{b}. \quad (8.4)$$

where the normalization constant \mathcal{N} is calculated from $\int_0^{\sigma_{max}} p(\sigma|\vec{D}) d\sigma = 1$. The prior probability density function is given by $\pi(\sigma, \vec{a}, \vec{b})$ and contains within it all the information pertaining to σ , \vec{a} and \vec{b} . This information consists of the cross section, yields for \vec{a} and \vec{b} an all correlated systematic uncertainties. Since it is assumed that \vec{a} and \vec{b} are independent of σ we can then factorize the prior probability density function as:

$$\pi(\sigma, \vec{a}, \vec{b}) = \pi(\vec{a}, \vec{b}) \pi(\sigma). \quad (8.5)$$

By assuming $\pi(\sigma)$ to be flat in σ , where:

$$\begin{aligned} \pi(\sigma) &= \frac{1}{\sigma_{max}}, 0 < \sigma < \sigma_{max} \\ &= 0, \text{ otherwise,} \end{aligned} \quad (8.6)$$

we avoid any potential bias towards a preferred cross section. Equation 8.4 can then be rewritten as:

$$p(\sigma|\vec{D}) = \frac{1}{\mathcal{N} \sigma_{max}} \int \int L(\vec{D}|\sigma, \vec{a}, \vec{b}) \pi(\vec{a}, \vec{b}) d\vec{a} d\vec{b}. \quad (8.7)$$

In order to extract a cross section from $p(\sigma|\vec{D})$ we plot $p(\sigma|\vec{D})$ versus the cross section and take the cross section as the mode (peak) of the curve shown in Figure 8.1

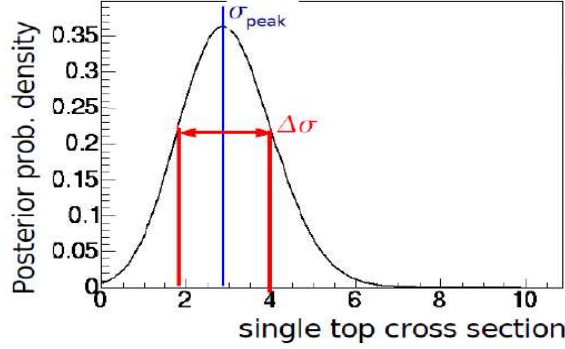


Figure 8.1: Representation of a 1D posterior density distribution where the peak of the curve represents a cross section measurement and $\Delta\sigma$ its uncertainty.

8.1.2 Systematic Uncertainties

As discussed in Chapter 6, the systematic uncertainties are accounted for as normalization and shape uncertainties. Uncertainties which only change the normalization are dealt with by modeling them as widths of a gaussian while setting the yields as the mean [82]. Uncertainties which affect the shape of the discriminant outputs are modeled using a gaussian where the mean is set to 0 and width set to 1 for individual systematics. Three distributions are then used; the nominal and the $\pm\sigma$ shifted distributions mentioned in Chapter 6. These are used to calculate the shift in the yield, Δy_{ibin} using:

$$\Delta y_{ibin} = s_{tot}^{\pm} \times g(0, 1) \times \delta_{ibin}^{\pm}, \quad (8.8)$$

where $s_{tot}^{\pm} = \frac{\Sigma y}{\Sigma y_{ibin}^{\pm}}$ and δ_{ibin}^{\pm} represents the $\pm\sigma$ fluctuations for that bin. Figure 8.2 shows an example of this fluctuation. This is done for individual bins for each source for a given channel. The yield for each bin is then calculated by:

$$y' = y + \Sigma_{isys} \Delta y_{ibin}, \quad (8.9)$$

where y is the nominal yield.

These uncertainties are then encoded into the prior density $\pi(\vec{a}, \vec{b})$ which is then used to numerically evaluate $p(\sigma|\vec{D})$. This is done by “MC importance sampling” [82]. For this analysis 1 million points (\vec{a}_k, \vec{b}_k) are sampled from $\pi(\vec{a}, \vec{b})$ and σ_{max} is set to 12 pb. The posterior density is then:

$$p(\sigma|\vec{D}) = \frac{1}{\mathcal{N}_{\sigma_{max}} N_{samples}} \sum_{k=1}^{N_{samples}} L(\vec{D}|\sigma, \vec{a}_k, \vec{b}_k). \quad (8.10)$$

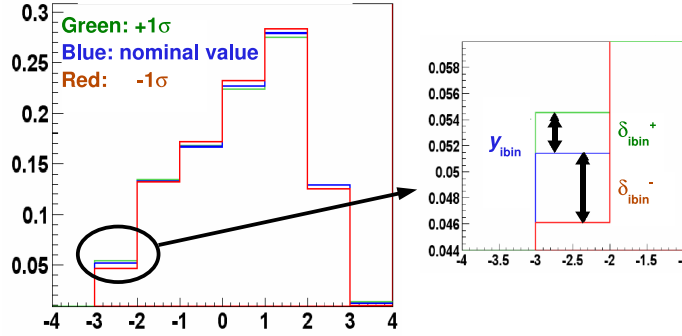


Figure 8.2: Representation of shape changing uncertainties.

8.2 Results

8.2.1 Expected Results

The boosted decision tree output histograms used to measure the three single top cross sections $(s+t, s, t)$ are presented in Appendix A. The results for the $s+t$, s and t cross sections are combinations of the six BDT discriminants obtained for each

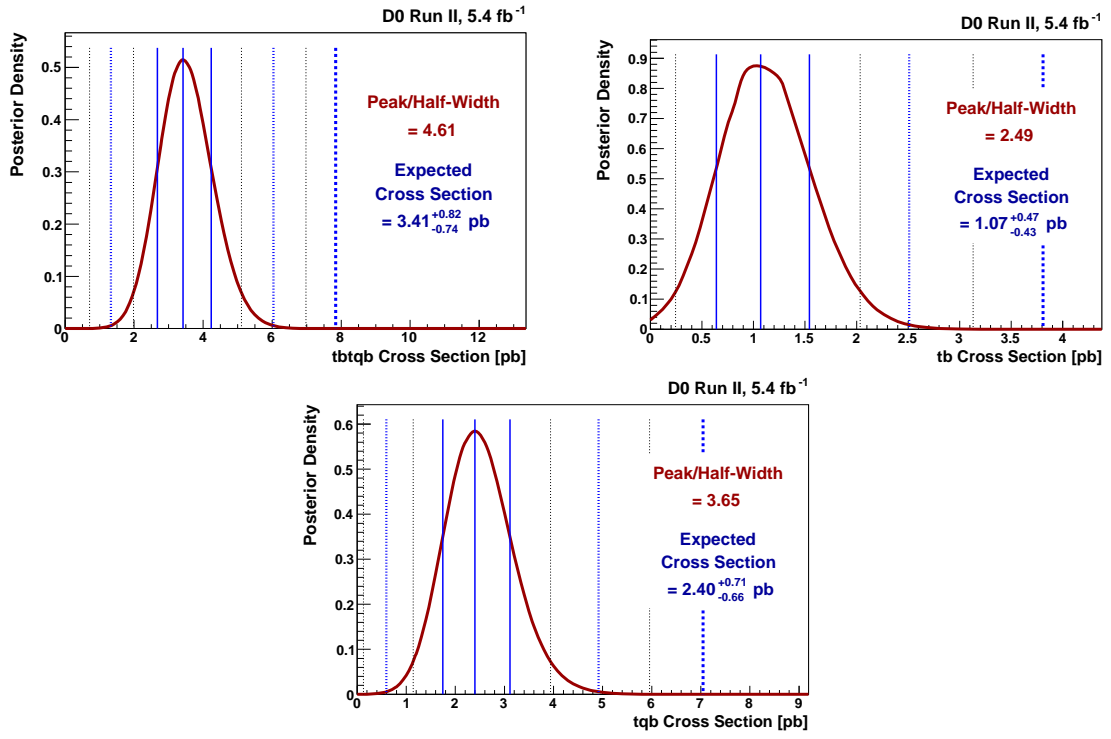


Figure 8.3: Expected posterior density distributions for $s+t$ (top left), s (top right) and t (bottom). The results include both statistical and systematic uncertainties.

cross section. The term expected refers to using only the MC signal and background samples to make a cross section measurement. These measurements are done initially so as not to bias subsequent measurements using the data. This also allows the trouble shooting of unstable results and the optimization of the sensitivity in the expected results. If the expected and observed results are consistent, this provides some confidence in the validity of the observed measurements.

The expected cross section results are shown in Table 8.1 and they are found to be in good agreement with the SM values of 3.30 pb ($s+t$ -channel), 1.04 pb (s -channel) and 2.26 pb (t -channel). The corresponding bayesian posteriors are shown in Fig 8.3 for each cross section and the “peak over half-widths” are shown in Table 8.2. The aim of these values is to provide an estimate of the sensitivity of the measurement. Table 8.3 shows all the cross section measurements for the six individual channels. In comparing the uncertainties of Tables 8.1 to 8.3 it is clear that by combining the channels the uncertainty is reduced. The measurements for all individual channels are presented in Table 8.3.

s+t-channel (pb)	s-channel (pb)	t-channel (pb)
$3.41^{+0.82}_{-0.74}$	$1.07^{+0.47}_{-0.43}$	$2.40^{+0.71}_{-0.66}$

Table 8.1: Expected cross sections for $s+t$, s and t channels in pb.

s+t-channel	s-channel	t-channel
4.61	2.49	3.65

Table 8.2: Expected posterior peak over half-width ratios for $s+t$, s and t channels.

8.2.2 Observed Results

This section presents the final cross section measurements obtained from the 5.4 fb⁻¹ dataset. All observed BDT discriminant outputs are shown in Appendix A.

Channel	s+t-channel (pb)	s-channel (pb)	t-channel (pb)
RunIIa+b, (e,μ) , 1tag, 2jets	$3.26^{+1.18}_{-1.14}$	$0.98^{+1.08}_{-0.89}$	$2.42^{+0.94}_{-0.85}$
RunIIa+b, (e,μ) , 1tag, 3jets	$3.79^{+1.84}_{-1.61}$	$1.52^{+1.86}_{-1.55}$	$2.46^{+1.55}_{-1.31}$
RunIIa+b, (e,μ) , 1tag, 4jets	$3.30^{+3.80}_{-3.15}$	$1.36^{+5.35}_{-1.36}$	$2.07^{+3.00}_{-2.07}$
RunIIa+b, (e,μ) , 2tag, 2jets	$3.37^{+1.97}_{-1.75}$	$1.00^{+0.67}_{-0.60}$	$1.07^{+7.95}_{-1.05}$
RunIIa+b, (e,μ) , 2tag, 3jets	$3.97^{+2.28}_{-2.01}$	$1.25^{+1.22}_{-1.04}$	$2.66^{+2.36}_{-2.00}$
RunIIa+b, (e,μ) , 2tag, 4jets	$3.70^{+3.91}_{-3.29}$	$0.81^{+4.07}_{-0.81}$	$3.88^{+4.05}_{-3.25}$

Table 8.3: Expected cross sections in individual analysis channels (e,μ) combined.

Figure 8.4 shows the discriminant distributions for the combined $s+t$, s and t analysis channels. The first column represents $s+t$, the second the s and the third the t channel analyses. The top row presents the results on the linear scale, the second row on a log scale and third row, zoomed in plots showing the excess of signal over background at large discriminant values.

The corresponding single top cross sections are shown in Table 8.4 and the peak over half widths in Table 8.5. Figure 8.5 shows the observed posterior density distributions for each of the cross section measurements for the $s+t$, s and t channels, respectively. For completeness, Table 8.6 shows each of the 18 individual cross section measurements.

s+t-channel (pb)	s-channel (pb)	t-channel (pb)
$3.01^{+0.80}_{-0.75}$	$0.68^{+0.41}_{-0.39}$	$3.03^{+0.78}_{-0.66}$

Table 8.4: Observed cross sections for $s+t$, s and t channels in pb.

s+t-channel	s-channel	t-channel
4.01	1.76	4.59

Table 8.5: Observed posterior peak over half-width ratios for $s+t$, s and t channels.

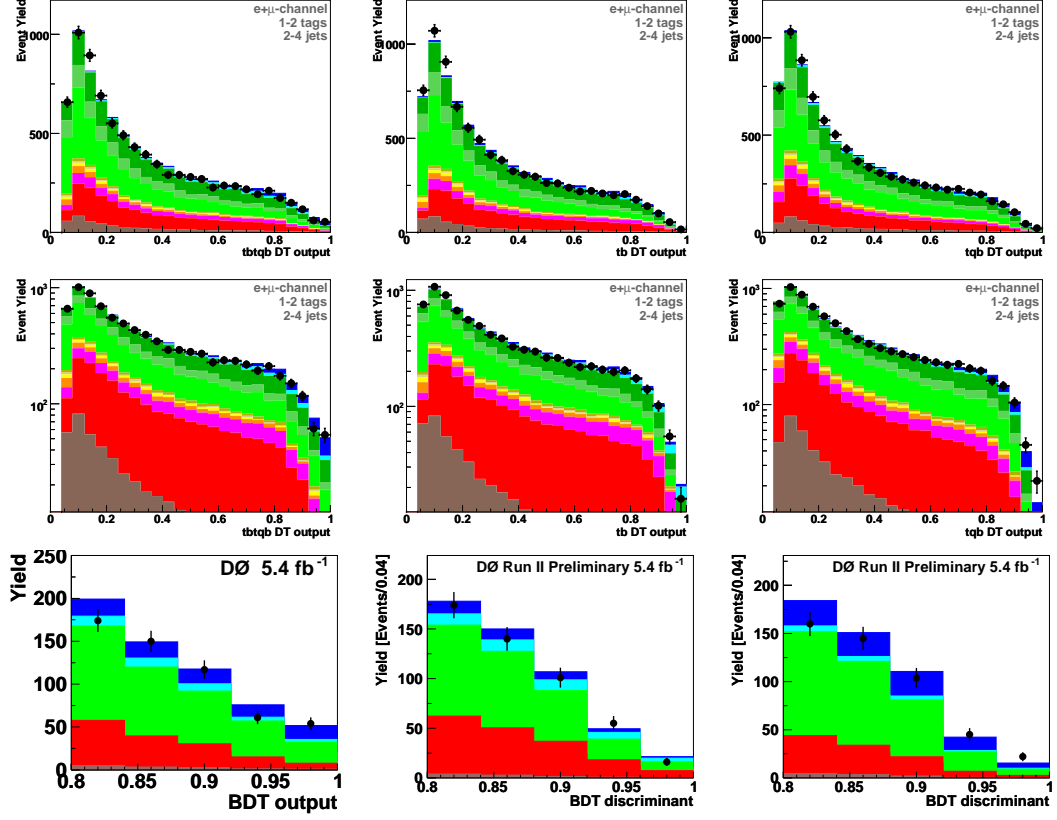


Figure 8.4: BDT output distributions for $s+t$, s and t channel analyses with all six channels combined. The top row corresponds to the linear scale, the middle the log scale and the bottom are zoomed plots showing signal excess over background.

Channel	s+t-channel (pb)	s-channel (pb)	t-channel (pb)
RunIIa+b, (e,μ) , 1tag, 2jets	$3.24^{+1.25}_{-1.22}$	$0.47^{+1.05}_{-0.47}$	$2.96^{+1.05}_{-0.95}$
RunIIa+b, (e,μ) , 1tag, 3jets	$2.57^{+1.52}_{-1.37}$	$0.00^{+2.00}_{-0.00}$	$3.66^{+1.53}_{-1.27}$
RunIIa+b, (e,μ) , 1tag, 4jets	$0.00^{+3.48}_{-0.00}$	$0.00^{+4.82}_{-0.00}$	$0.67^{+2.67}_{-0.67}$
RunIIa+b, (e,μ) , 2tag, 2jets	$2.36^{+1.64}_{-1.50}$	$0.42^{+0.47}_{-0.42}$	$2.67^{+6.50}_{-2.67}$
RunIIa+b, (e,μ) , 2tag, 3jets	$1.76^{+1.76}_{-1.62}$	$1.17^{+1.20}_{-1.06}$	$4.04^{+2.36}_{-2.02}$
RunIIa+b, (e,μ) , 2tag, 4jets	$3.63^{+3.57}_{-3.02}$	$0.00^{+3.85}_{-0.00}$	$2.58^{+3.38}_{-2.58}$

Table 8.6: Observed cross sections in individual analysis channels (e,μ) combined.

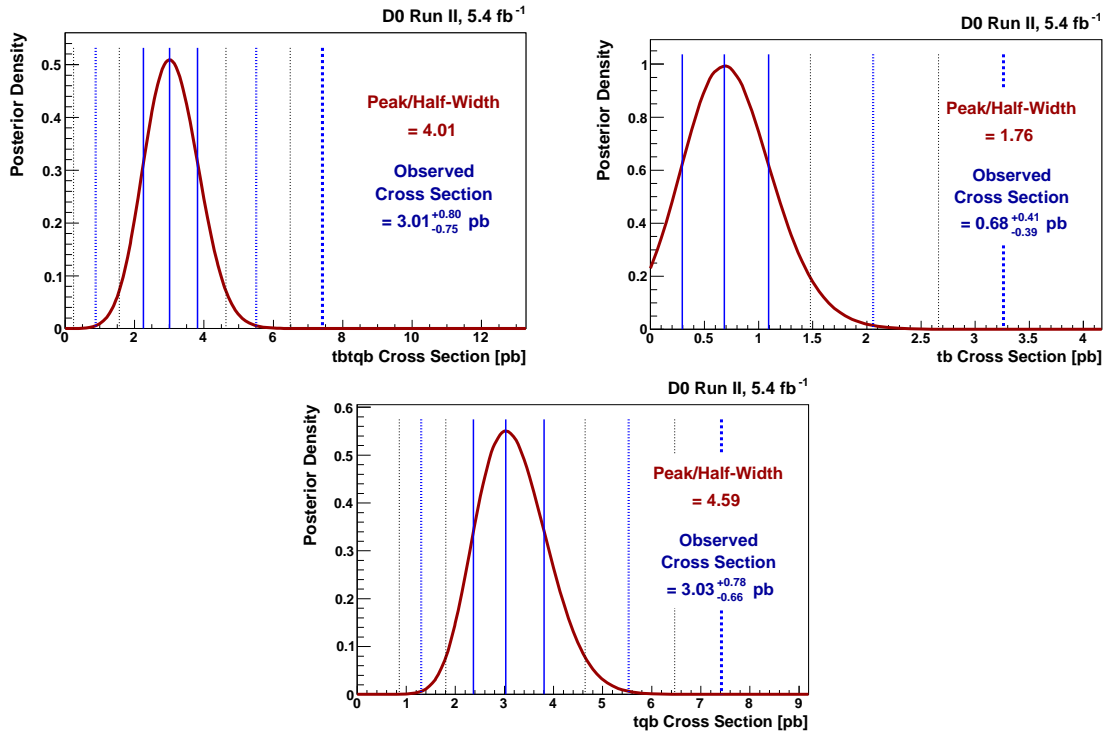


Figure 8.5: Observed posterior density distributions for the single top $s+t$ (top left), s (top right) and t (bottom) analyses.

8.3 Ensemble Tests

Ensemble testing is used to test for any bias in the cross section measurements. This is done by creating ensembles of pseudo-datasets, given an input cross section, and using the signal and background events contained in the discriminant outputs. Pseudo-datasets are created by first randomly selecting a systematic sample and then a count is chosen randomly from a Poisson distribution for each bin. The mean of the Poisson distribution comes from both the background and signal yields, however the signal yields are adjusted to match the input cross section. All systematic uncertainties are taken into account when performing the analysis. For a more detailed discussion please see reference [82].

Once the ensembles are created they are then run through the full analysis to produce new measurements of each cross section. These are summarized in Table 8.7 and corresponding distributions are shown in Figures 8.6-8.8. Each of the distributions

Input Cross sections		
$s + t$ channel (pb)	s channel (pb)	t channel (pb)
2	0.5	1.5
3.30 (SM)	1.04 (SM)	2.26 (SM)
5	1.5	3.5
7	2.5	5
10	3.5	7

Table 8.7: Cross section inputs for each individual channel.

are fitted by a gaussian where the mean provides the measurement of the cross section. The results are then plotted against the input cross section used to test the linearity of the results. In all cases a smooth linear correlation is observed indicating that the fits are free from significant bias and that no additional corrections are needed.

LINEARITY TEST FOR BDT $s+t$ CHANNEL

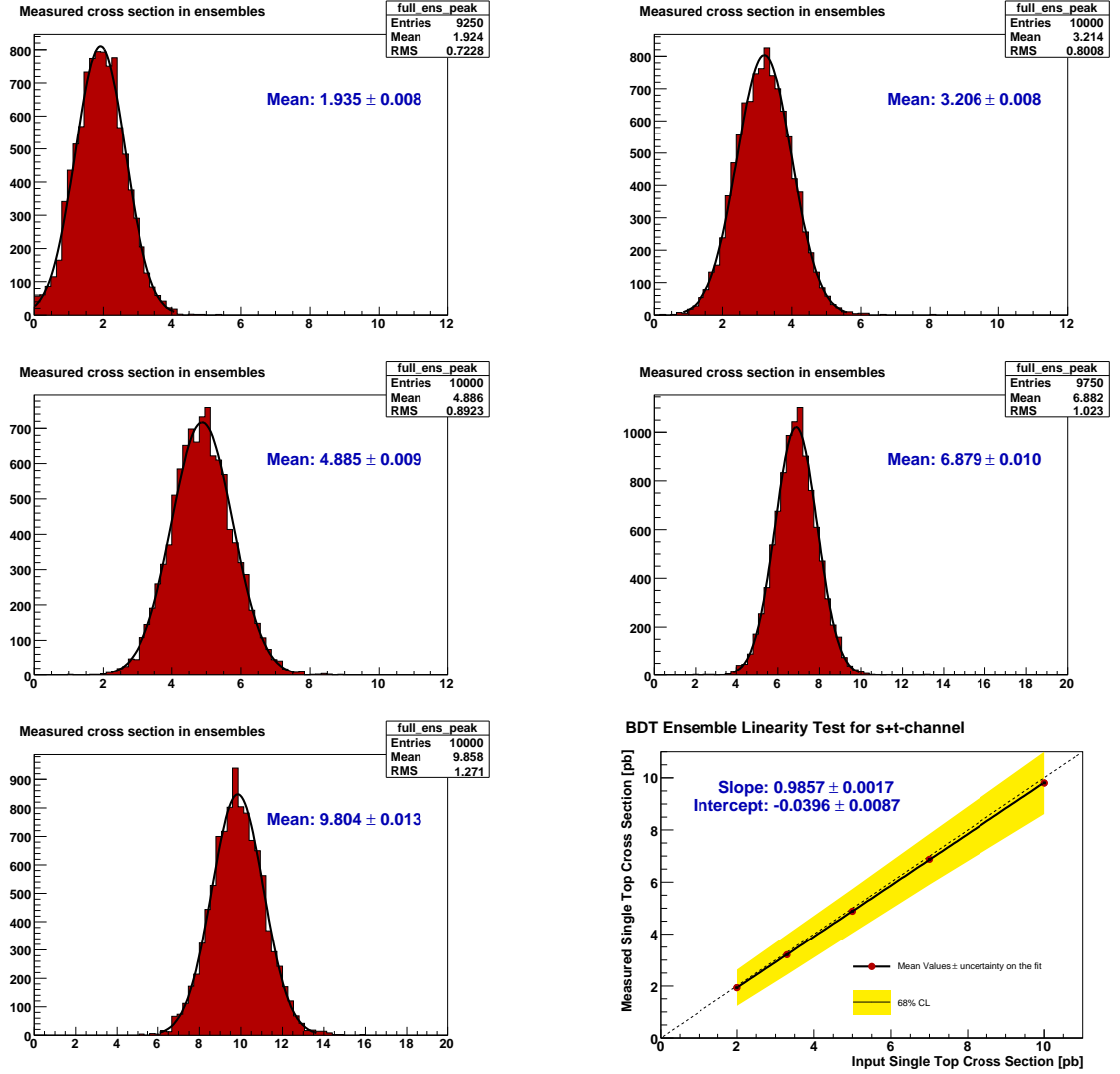


Figure 8.6: Results of ensemble tests for the combined $s + t$ channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.

LINEARITY TEST FOR BDT s CHANNEL

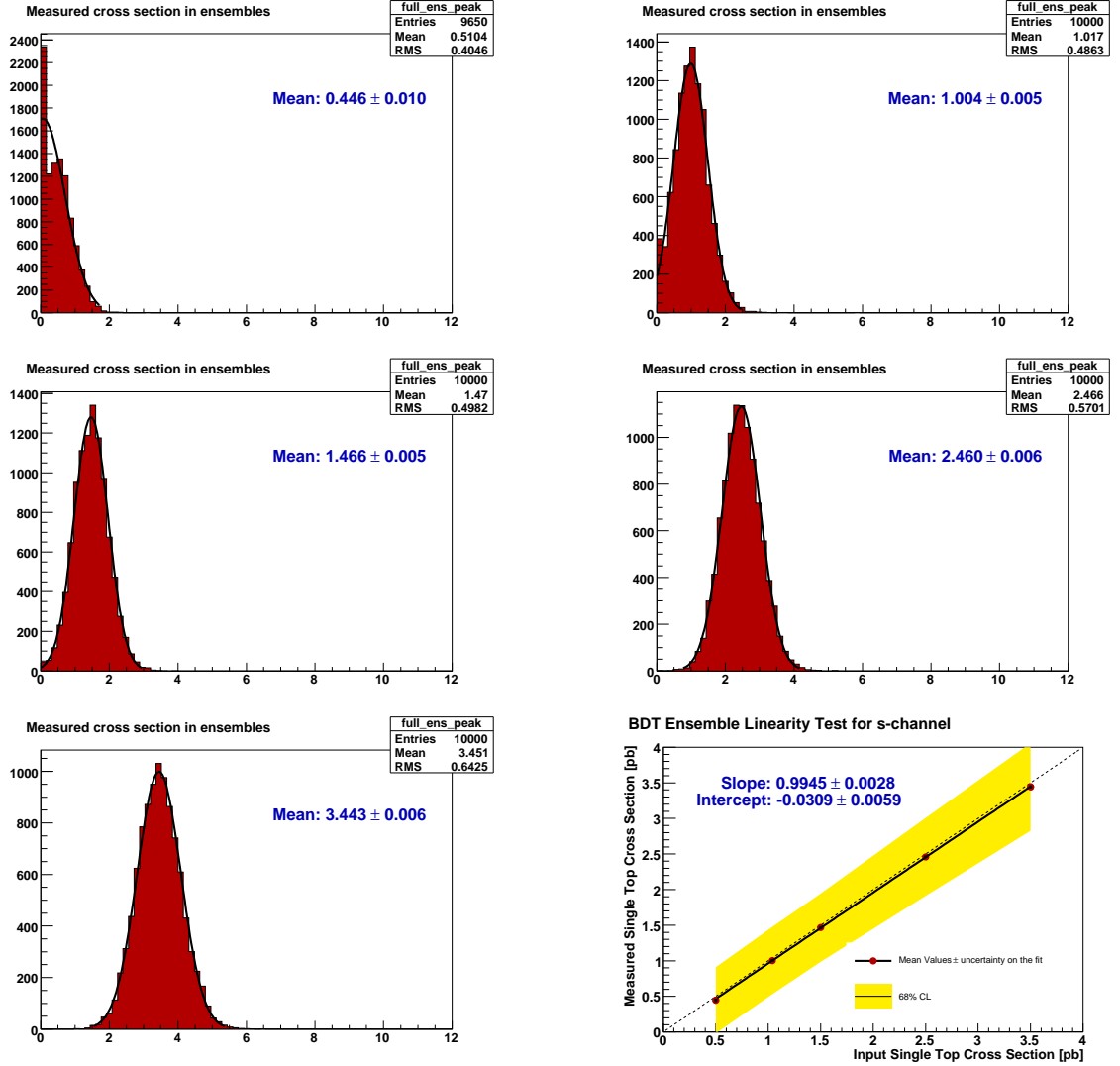


Figure 8.7: Results of ensemble tests for the combined s channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.

LINEARITY TEST FOR BDT t CHANNEL

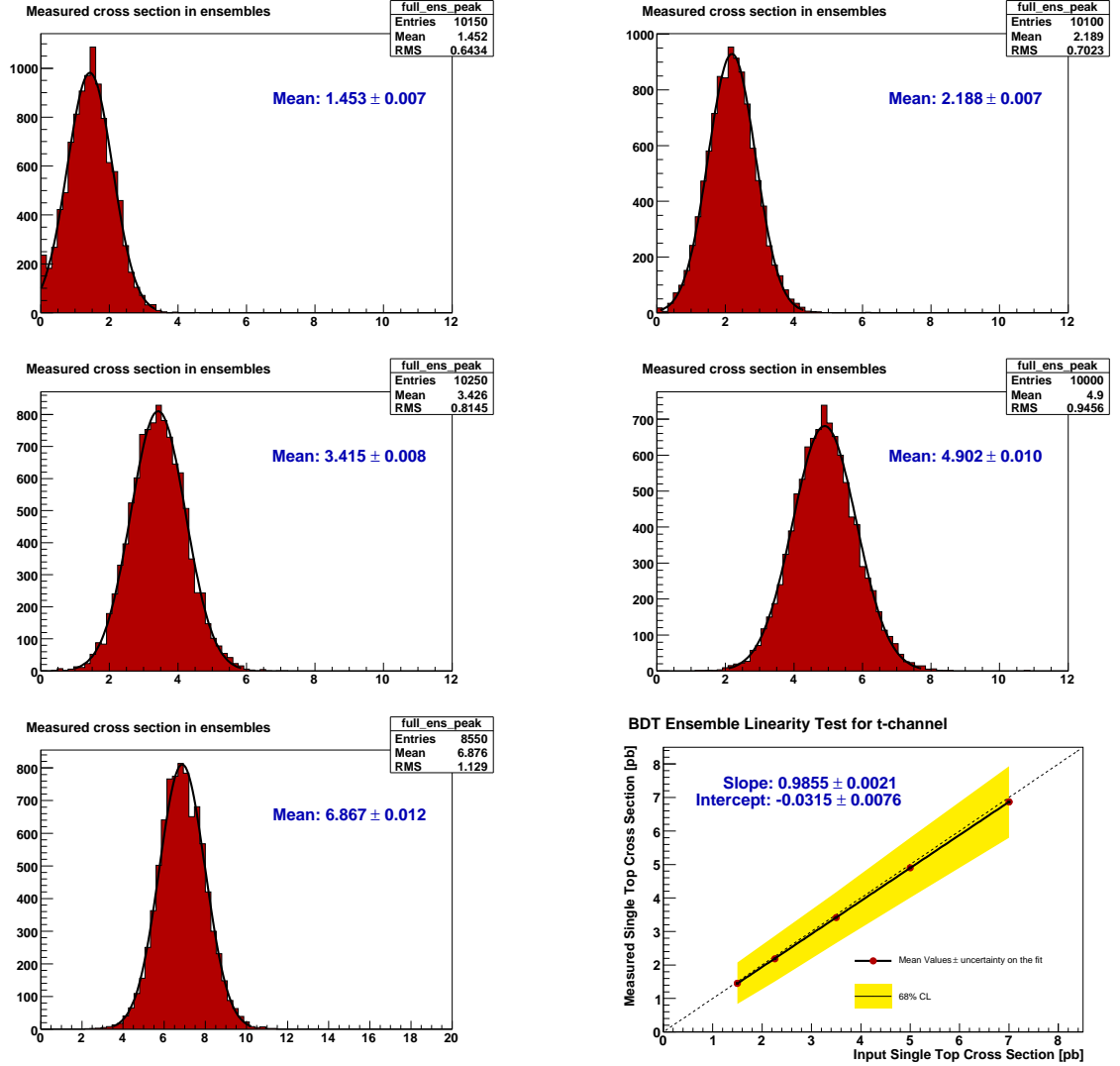


Figure 8.8: Results of ensemble tests for the combined t channel analysis. The observed linear behavior indicates the absence of bias from the analysis procedure.

8.3.1 Significance

The statistical significance of a measurement is generally calculated by finding the “p-value” of that measurement. The p-values from this analysis can be interpreted as the probability of measuring a given cross section in data if the dataset contained no single top events. Typically this is expressed in standard deviations where α represents the p-value. The number of standard deviations is then given by

$$N_\sigma = \sqrt{2} \operatorname{erf}^{-1}(1 - 2\alpha). \quad (8.11)$$

8.3.2 Asymptotic approximation of the log-likelihood ratio

In the previous $D\bar{O}$ analysis, the cross section significance was calculated using ensembles of pseudo-datasets that contained no signal [80]. This approach takes a large amount of cpu time and translates to about a month of computing time. A more effective approach called the log likelihood method channel analysis was used in the t channel analysis [84, 86]. However, this too is very cpu intensive but less so than ensemble testing. For this analysis the asymptotic approximation of the log-likelihood method (AALR) has been used [86]. The results have been compared to the previous analyses and are found to be consistent [53].

As stated earlier, the goal of measuring a significance is to obtain a p-value. In the AALR method this is done by first defining the log-likelihood ratio

$$Q = -2 \ln \frac{L(\mu = 1)}{L(\mu = 0)}, \quad (8.12)$$

where the likelihood is defined by $L(\mu)$. The likelihood includes all systematic for $\sigma = \mu\sigma_o$, where μ is the strength parameter and σ_o the theoretical cross section for

the single top signal. The 1 represents the likelihood for a theoretical signal cross section compared to the likelihood of no theoretical cross section (background only).

If a number of events are to be assessed by AALR given as signal strength of μ' then the output distribution is a Gaussian whose mean and variance is

$$E[Q] = \alpha^2(1 - 2\mu') \quad (8.13)$$

$$V[Q] = 4\alpha^2, \quad (8.14)$$

where α is defined as

$$\alpha = \frac{\sqrt{-2 \ln \frac{L_A(\mu)}{L_A(\hat{\mu})}}}{|\mu - \mu'|}, \quad (8.15)$$

and $\hat{\mu}$ corresponds to the maximum value for the likelihood. Assuming that α is independent of μ then we can then set $\mu' = 0$ and μ then becomes $\mu_{exp} = \sigma_{exp}/\sigma_o$ where σ_{exp} is the expected cross section. This is the background only scenario where the probability density $\rho(Q)$ is given by a Gaussian with a mean and variance defined as α_B^2 and $4\alpha_B^2$. For the signal plus background case the mean and variance become $-\alpha_{S+B}^2$ and $4\alpha_{S+B}^2$. In order to calculate the p-value we use the cumulative distribution for a Gaussian Φ with mean 0 and variance 1 for which the p-value is defined as

$$p = \int_{-\infty}^{Q_{obs}} \rho(Q) dQ = \Phi \left(\frac{Q_{obs} - \alpha^2}{2\alpha} \right). \quad (8.16)$$

8.3.3 Measured Significance

Table 8.9 summarizes the parameters and significance calculations for $s+t$, s and t channels analyses. Figs 8.9-8.11 show the expected and observed significances for the three analyses. In each figure the green curve represents the background only scenario and blue curve the signal plus background case. The position of the black

Channels	α_B	α_{S+B}	Q_{obs}	p_{exp}	p_{obs}
s+t	4.40	4.82	-18.83	4.9σ	4.3σ
s	2.27	2.54	-2.06	2.6σ	1.6σ
t	3.59	4.03	-25.68	4.1σ	5.5σ

Table 8.8: Parameters of the likelihood which are used in the calculation of the significances using the AALR approach.

arrow indicates the observed log-likelihood ratio in each case. The expected p-value is then obtained by integrating from the mean of the blue curve moving left till the end of the green curve. The observed p-value is done in a similar way, where the integration starts from the observed log-likelihood ratio.

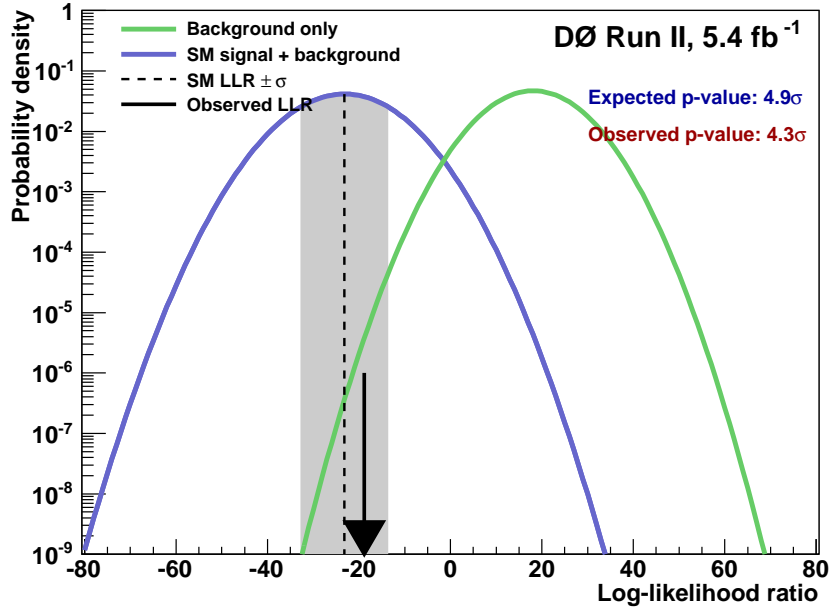


Figure 8.9: Probability density vs. log-likelihood significance for the $s + t$ channel analysis.

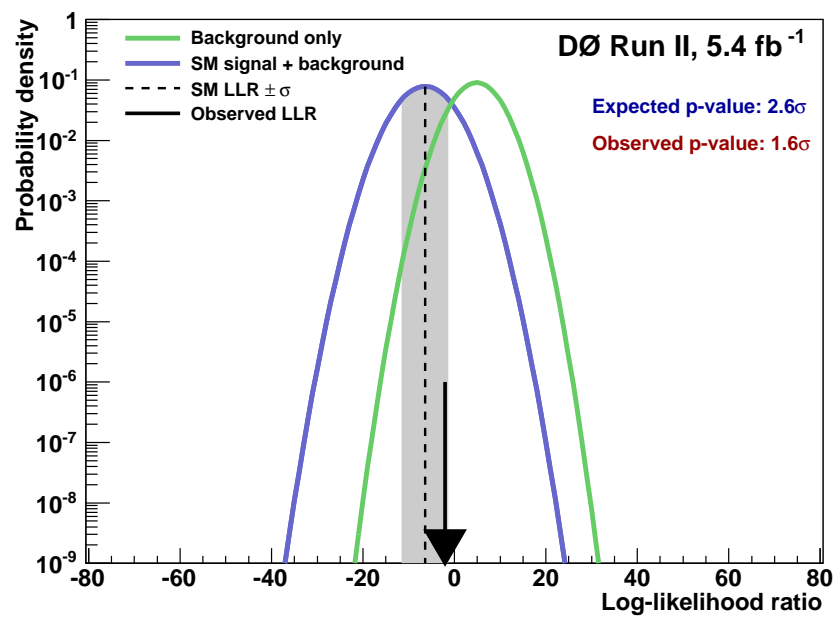


Figure 8.10: Probability density vs. log-likelihood significance for the s channel analysis.

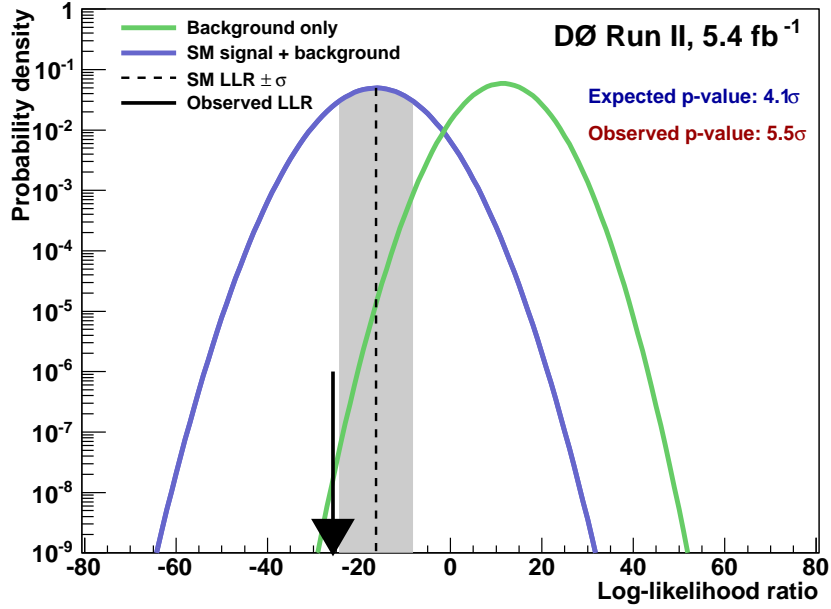


Figure 8.11: Probability density vs. log-likelihood significance for the t channel analysis.

Channels	Expected Cross Section (pb)	Observed Cross Section (pb)	p_{obs}
s+t	$3.41^{+0.82}_{-0.74}$	$3.01^{+0.80}_{-0.75}$	4.3σ
s	$1.07^{+0.47}_{-0.43}$	$0.68^{+0.41}_{-0.39}$	1.6σ
t	$2.40^{+0.71}_{-0.66}$	$3.03^{+0.78}_{-0.66}$	5.5σ

Table 8.9: Expected and observed results for the single top cross sections.

8.4 $|V_{tb}|$ Measurement

The single top production cross section is directly proportional to the square of the CKM matrix element $|V_{tb}|$. This measurement of the cross section can also be used to measure $|V_{tb}|$ without the requirement of assumptions for the number of generations or unitarity. However it is still necessary to assume that $t \rightarrow Wb \sim 100\%$ of the time and that it is a CP-conserving V-A interaction.

The measurement of $|V_{tb}|$ is done using the same bayesian technique as was used in measuring the cross section. The difference is that the results from three different MVA analyses are combined in order to measure $|V_{tb}|$. This is done using a bayesian neural network (BNN) which takes the results from this ananalysis and combines them with the results from the $D\bar{O}$ BNN [87] and NEAT [88] analyses [53]. All systematic uncertainties are considered plus additional theoretical uncertainties from of the Top quark mass, Factorization scale, PDF and α_s uncertainties. Assuming a mass of 172.5 GeV the resulting uncertainties are $\pm 5.3\%$ for the t channel and $\pm 3.8\%$ for the s channel.

Specifically we are measuring the V-A coupling strength represented as $|V_{tb}f_1^L|$, where f_1^L is the Wtb coupling. The resulting Bayesian probability denstiy for $|V_{tb}f_1^L|^2$ can then be used to extract $|V_{tb}|$. This is shown in the plot on the left of Figure 8.12 where $|V_{tb}f_1^L|^2$ is given as the peak of the posterior density. This gives a result of 1.03 ± 0.22 for $|V_{tb}f_1^L|^2$ and $1.02^{+0.10}_{-0.11}$ for $|V_{tb}f_1^L|$. When constraining $|V_{tb}f_1^L|^2$ between 0 and 1 we set $f_1^L = 1$ and obtain a limit of $|V_{tb}|^2 > 0.62$ which corresponds to $|V_{tb}| > 0.79$ at the 95% CL. This is the best limit attained thus far.

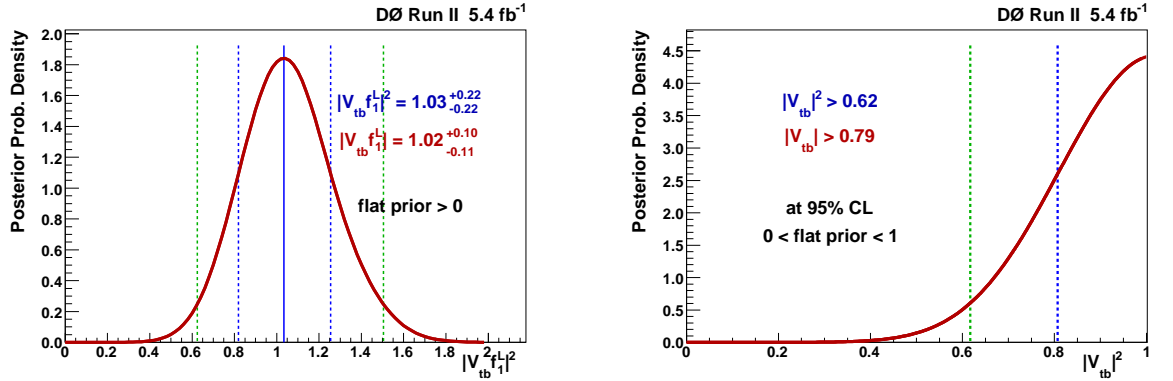


Figure 8.12: Posterior densities for both $|V_{tb}f_1^L|^2$ (left) and $|V_{tb}|^2$ (right).

Chapter 9

Comparison of Results

9.1 $D\bar{O}$ and CDF Measurements

In this chapter we compare results from $D\bar{O}$ and CDF single top analyses and the predictions of the NLO QCD. We also compare the result of this analysis with those obtained from the previous (2009) $D\bar{O}$ analyses [80]. The integrated luminosities for these results are, 2.3fb^{-1} for ($D\bar{O}$ 2009) [80], 5.4fb^{-1} for (this analysis) and 3.2fb^{-1} (CDF) [85].

The $D\bar{O}$ single top group uses two different approaches to calculate cross sections. The first is the one dimensional approach which was discussed in Section 8.1.1. This approach was used for the three MVA analyses: Boosted Decision Trees (this analysis), Bayesian Neural Networks (BNN) and, Neural Evolution of Augmenting Topologies (NEAT). Each MVA provides a cross section measurement for $s+t$, s and t channels. The results were combined using two different BNN methods. The first used as input the results from the three sets of MVA results and is labelled BNN Combination. The second, labelled BNN Combination 6, takes as input the s and t measurements from the three MVA analyses, hence the number six. In both cases

the combined $s + t$, s and t channel cross sections are measured.

The second method is a two dimensional approach. This is used specifically to measure the separated s and t channel cross sections. It uses a two dimensional posterior density function for the combined s and t channels. This is done without making any assumption on the predicted cross sections. Using the 2D posterior density we can then integrate over either s or t to obtain a 1D posterior density. This is shown in Figure 9.3 for the BNN, which is labelled BNN Combination t method. In this case the training was implemented so as to focus on the t channel. The BNN Combination 6 method also uses the 2D approach and these results are compared to the other methods in Fig 9.1. The significance values presented in Figures 9.1 and 9.2 are labelled 2D if the cross sections were obtained from the 2D method, otherwise they are 1D measurements.

9.1.1 Comparisons

Fig 9.1 a) shows the comparison of the $D\bar{O}$ results from the three different MVA analyses and for the combinations for the combined $s + t$ -channel cross section. The measurements are consistent within one sigma as well as being in agreement with the predicted cross section of 3.30 ± 0.13 pb. Fig 9.1 b) shows the corresponding results for the s channel and Fig 9.1 c), for the t channel. The results are all consistent within one standard deviation. The s channel results are also consistent with the predicted cross section ($1.04 \pm .04$ pb) at the one standard deviation level, apart from the NEAT result, which is within 2 standard deviations. The t -channel results are all consistent with theory (2.26 ± 0.12 pb) and with eachother at the one standard deviation level. Table 9.1 shows a list of significances for the three channels. The results of note are the observed p-values for the t channel using BDTs (5.5σ), BNN

Combination t (5.5σ (2D)) and the $s + t$ channel using BNN Combination 6 (5.6σ). These are the best measurements for each of these channels.

Figure 9.2 compares these results with those from the previous $D\bar{O}$ analysis and with those from CDF. Good agreement is observed between the $D\bar{O}$ results for the $s + t$, s and t channels. The CDF combination is consistent within a 1-2 sigma uncertainty and they tend to be less consistent with the theoretical predictions. The significances of these results can be seen in Table 9.2.

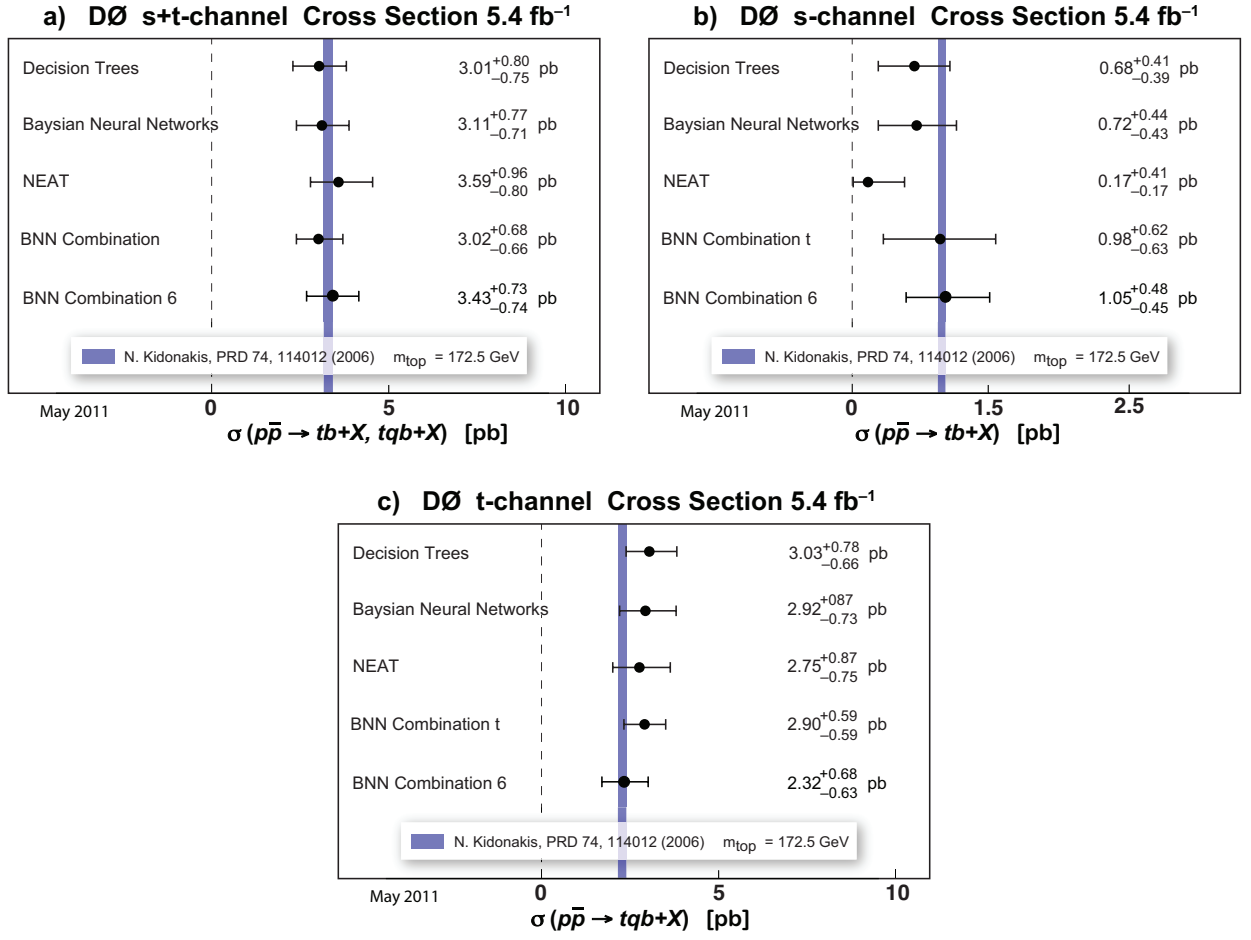
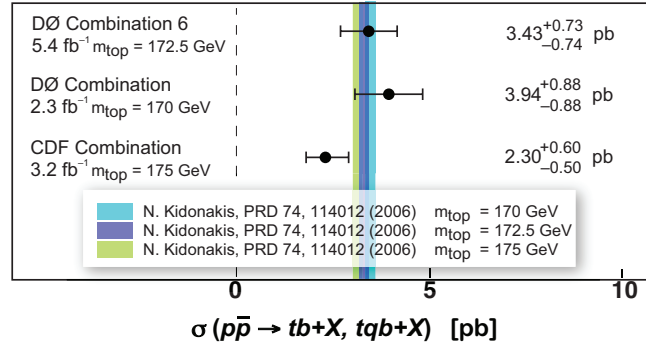


Figure 9.1: $D\bar{O}$ cross section results for 2.3 fb^{-1} , 5.4 fb^{-1} .

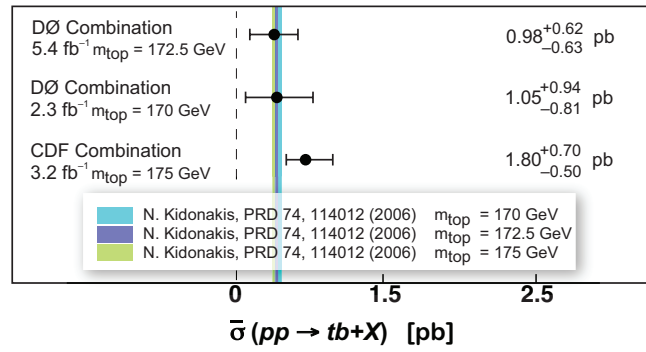
$D\bar{O}$ 5.4 fb^{-1} Single Top Results						
Analysis Method	Significance					
	s+t		s		t	
	expected	observed	expected	observed	expected	observed
Boosted Decision Trees	4.9σ	4.3σ	2.6σ	1.6σ	4.1σ	5.5σ
BNN Combination	NA	NA	NA	NA	4.6σ (2D)	5.5σ (2D)
BNN Combination 6	5.7σ	5.6σ	2.1σ (2D)	2.4σ (2D)	4.4σ (2D)	4.2σ (2D)

Table 9.1: $D\bar{O}$ significance results for $s + t$, s and t channels.

CDF and DØ s+t-channel Cross Section



CDF and DØ s-channel Cross Section



CDF and DØ t-channel Cross Section

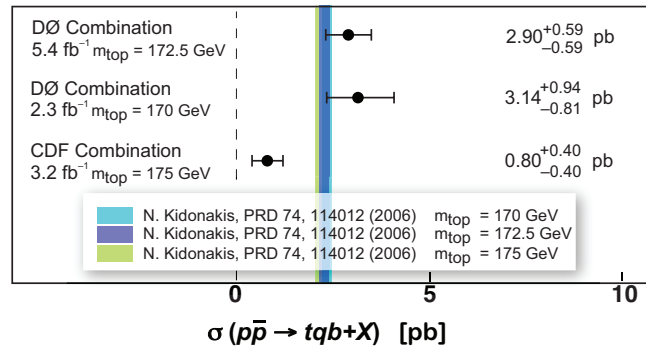


Figure 9.2: Cross section results for 2.3 fb^{-1} (DØ), 5.4 fb^{-1} (DØ) and 3.2 fb^{-1} (CDF) data sets.

DØ and CDF Combination Single Top Results						
Analysis Method	Significance					
	s+t		s		t	
	expected	observed	expected	observed	expected	observed
DØ 5.4 fb ⁻¹	NA	NA	2.7σ	1.8σ	4.6σ (2D)	5.5σ (2D)
DØ Comb6 5.4 fb ⁻¹	5.7σ	5.6σ	2.1σ (2D)	2.4σ (2D)	4.4σ (2D)	4.2σ (2D)
DØ 2.3 fb ⁻¹	4.5σ	5.0σ	NA	NA	3.7σ	4.8σ
CDF 3.2 fb ⁻¹	5.9σ	5.0σ	NA	NA	NA	NA

Table 9.2: $DØ$ and CDF significance results for $s + t$, s and t channels.

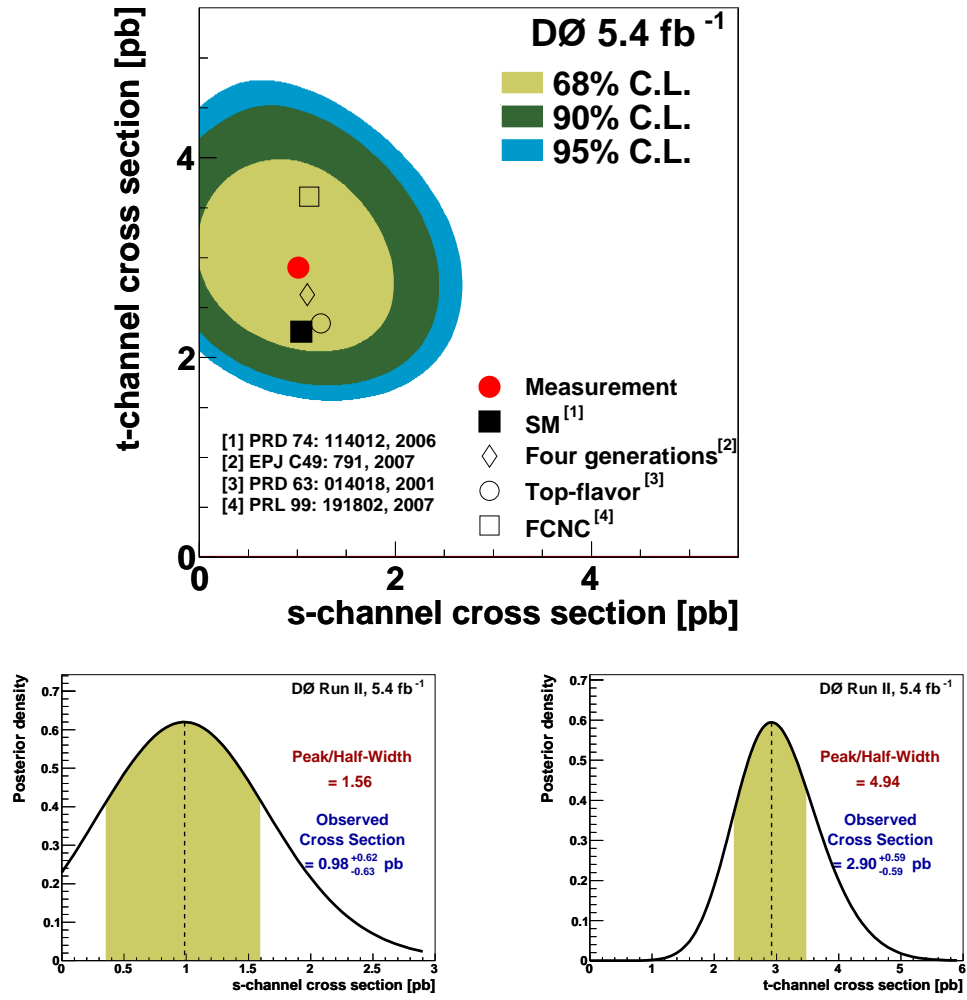


Figure 9.3: 2D posterior density for BNN Combination t . Below are the 1D posterior densities for both s and t obtained by integrating over both axes.

Chapter 10

Conclusions

This dissertaion presents measurements of the $s + t$, s and t channel single top quark cross sections using 5.4fb^{-1} of data collected at the $D\bar{O}$ detector at Fermilab in Batavai, IL. The data was classified into six subsamples according to the number of jets and b tags for each of the channels. These samples were then used to train six BDTs for each channel using Bayesian statistics and used to calculate the individual cross sections. The final results are:

$$s \text{ channel : } \sigma(p\bar{p} \rightarrow tb + X) = 0.68^{+0.41}_{-0.39} \text{ pb,}$$

$$t \text{ channel : } \sigma(p\bar{p} \rightarrow tqb + X) = 3.03^{+0.78}_{-0.66} \text{ pb,}$$

$$\text{Combined } s + t \text{ channels : } \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.01^{+0.80}_{-0.75} \text{ pb.}$$

The s channel has a significance of 1.6σ , the $s + t$ channel 4.3σ and the t channel 5.5σ . The corresponding theoretical predictions are 1.04 ± 0.04 pb, 3.30 ± 0.13 pb and 2.26 ± 0.12 pb. The observed results are consistent with the predictions to within one standard deviation for each of the three measurements. These results

were combined with two parallel analyses using different MVA methods to provide the final $D\bar{O}$ results for 5.4fb^{-1} . They give:

$$\text{Combined } s + t \text{ channels : } \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.43_{-0.74}^{+0.73} \text{ pb},$$

$$t \text{ channel : } \sigma(p\bar{p} \rightarrow tqb + X) = 2.90_{-0.59}^{+0.59} \text{ pb}.$$

The $s + t$ channel result has a significance of 5.6σ , and the t channel a significance of 5.5σ . Both results are in good agreement with theoretical predictions and with the results from this analysis.

Appendix A

Decision Tree Outputs

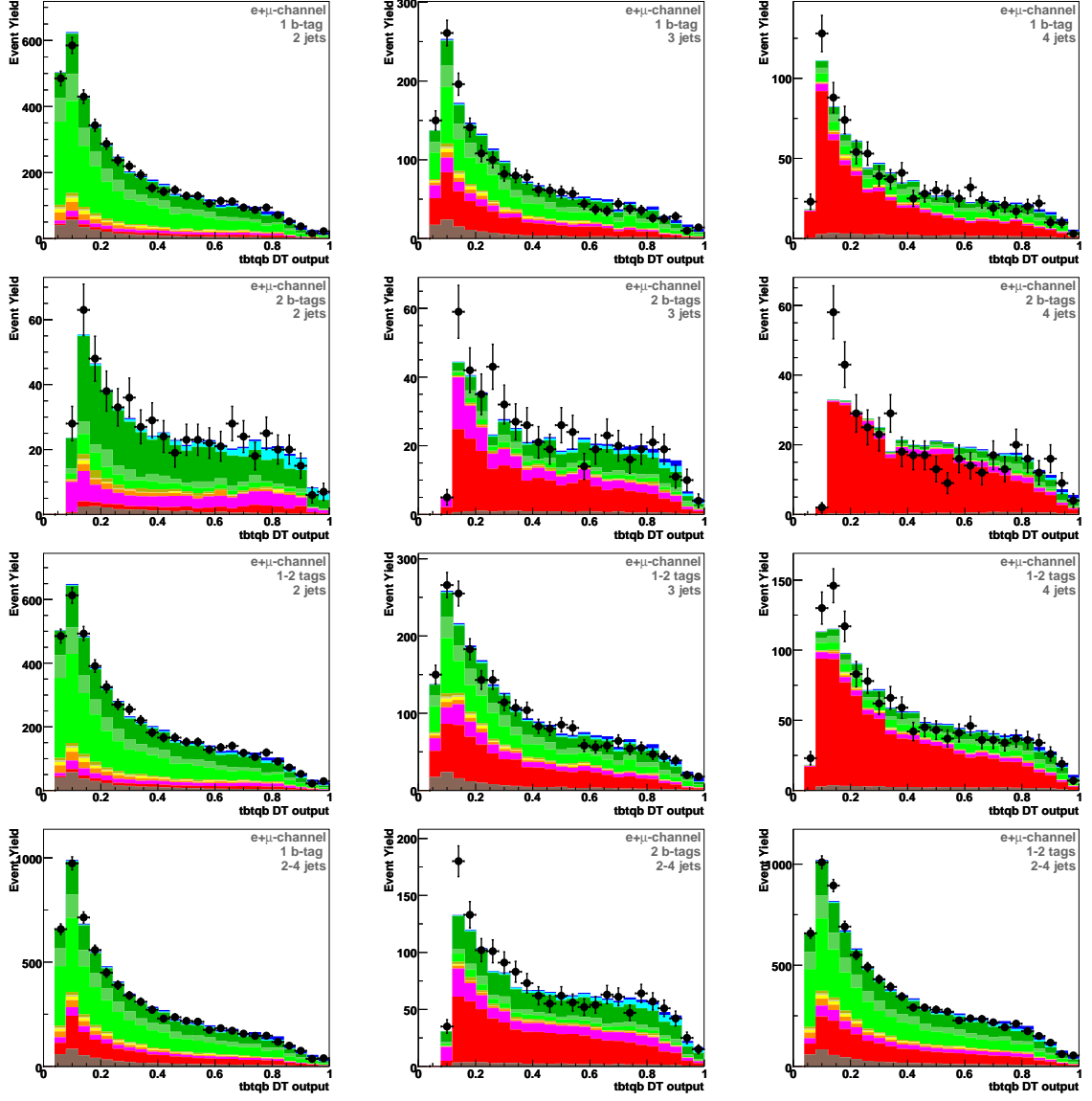


Figure A.1: Boosted decision tree outputs for the $s+t$ channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.

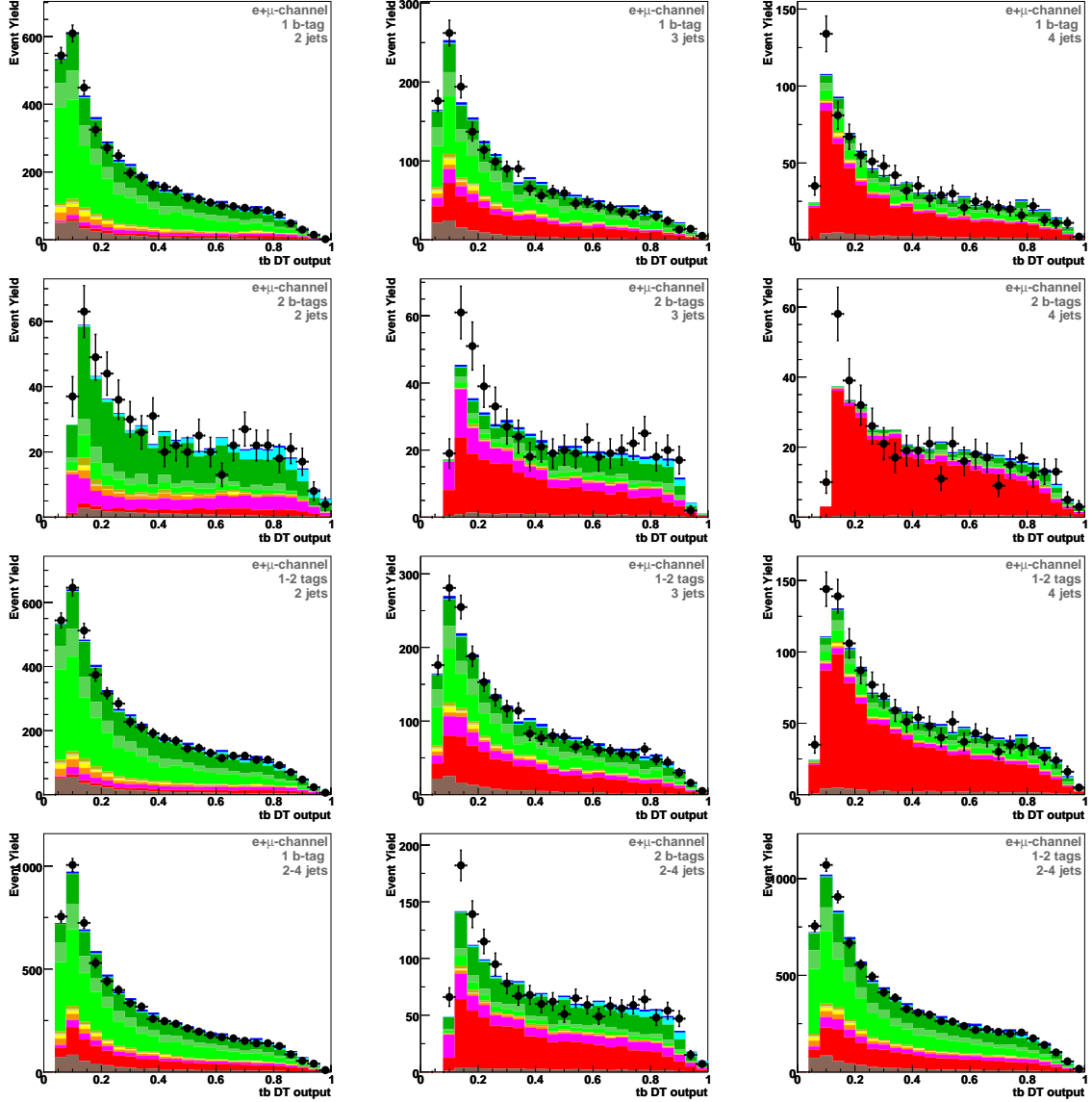


Figure A.2: Boosted decision tree outputs for the s channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.

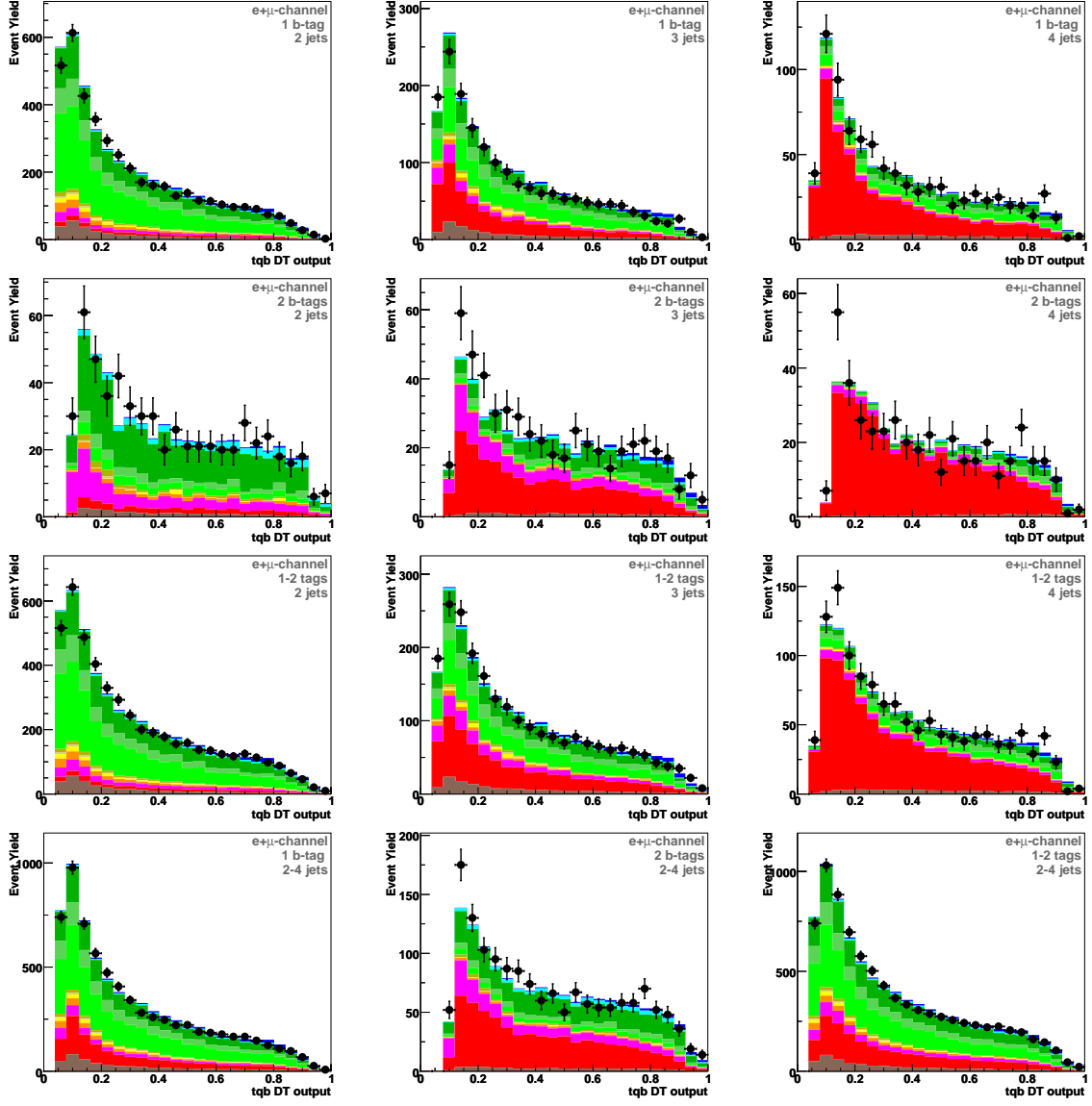


Figure A.3: Boosted decision tree outputs for the t channel. The first row consists of the once tagged 2,3 and 4 jet samples, the second row the twice tagged 2,3 and 4 jet samples, the third row the 1-2 tagged 2,3 and 4 jet samples and the last row consists of the once tagged 2-4 jets combined sample, the twice tagged 2-4 jets combined sample and the combined sample of 1-2 tags and 2-4 jets.

Appendix B

Residual Systematic Errors

B.1 Normalization Uncertainties

Tables B.1- B.6 show the effects of the systematic uncertainties (Section 6.3.1) for the six jet and b tag categories used in the analysis. The values listed here are used as inputs to the BDT discriminants that are used to measure the single top cross sections. It should be noted that systematic samples which have a minus sign are considered anti-correlated with other samples in the same row. This is a constraint implemented in this analysis to preserve normalization.

UNCERTAINTIES FOR CHANNELS WITH TWO JETS

SINGLE TAG

Source	Uncertainties %												
	$t\bar{l}l$	$t\bar{l}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	$t\bar{b}$	t	$t\bar{b} + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cross section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	−0.4	−0.4	—	—	—	1.0	—	—	—	—	0.5	0.5	0.5
Lepton ID	2.9	2.9	—	—	—	2.4	2.4	2.6	2.7	—	2.8	2.8	2.8
Jet fragmentation and higher order	1.6	1.6	—	—	—	—	—	—	—	—	0.7	0.7	0.7
ISR/FSR	4.3	4.3	—	—	—	—	—	—	—	—	0.8	0.8	0.8
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	7.1	5.6	5.6	5.1	6.2	6.2	6.4	5.8	5.7	—	6.6	5.9	6.1
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
W+Jet and QCD	—	—	1.8	1.8	1.8	—	—	—	—	−34.6	—	—	—

Table B.1: Residual normalization systematic uncertainties: two jets with 1- b -tag. These values are used as input parameters to the BDT analysis.

DOUBLE TAG

Source	Uncertainties %												
	$t\bar{l}l$	$t\bar{l}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	tb	t	$tb + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cros section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	−0.4	−0.4	—	—	—	1.0	—	—	—	—	0.5	0.5	0.5
Lepton ID	2.9	2.8	—	—	—	2.3	2.4	3.2	2.7	—	2.8	2.8	2.8
Jet fragmentation and higher order	1.6	1.6	—	—	—	—	—	—	—	—	0.7	0.7	0.7
ISR/FSR	4.3	4.3	—	—	—	—	—	—	—	—	0.8	0.8	0.8
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	3.5	3.2	3.8	3.7	4.4	3.8	3.4	3.6	3.4	—	3.5	3.9	3.6
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
$W + \text{Jet and QCD}$	—	—	1.8	1.8	1.8	—	—	—	—	−34.9	—	—	—

Table B.2: Residual normalization systematic uncertainties: two jets with 2- b -tags. These values are used as input parameters to the BDT analysis.

UNCERTAINTIES FOR CHANNELS WITH THREE JETS

SINGLE TAG

Source	Uncertainties %												
	$t\bar{t}ll$	$t\bar{t}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	tb	t	$tb + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cross section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	0.0	0.0	—	—	—	0.6	—	—	—	—	0.8	0.8	0.8
Lepton ID	2.8	2.8	—	—	—	2.5	2.5	2.7	2.7	—	2.7	2.8	2.8
Jet fragmentation and higher order	1.7	1.7	—	—	—	—	—	—	—	—	3.7	3.7	3.7
ISR/FSR	2.2	2.2	—	—	—	—	—	—	—	—	—5.9	—5.9	—5.9
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	13.3	8.6	7.9	8.0	8.8	9.4	7.7	14.8	8.7	—	9.3	8.6	8.9
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
W +Jet and QCD	—	—	1.8	1.8	1.8	—	—	—	—	—36.2	—	—	—

Table B.3: Residual normalization systematic uncertainties: three jets with 1- b -tag. These values are used as input parameters to the BDT analysis.

DOUBLE TAG

Source	Uncertainties %												
	$t\bar{t}ll$	$t\bar{t}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	tb	t	$tb + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cross section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	0.0	0.0	—	—	—	0.6	—	—	—	—	0.8	0.8	0.8
Lepton ID	2.8	2.8	—	—	—	2.5	2.4	2.2	2.7	—	2.8	2.8	2.8
Jet fragmentation and higher order	1.7	1.7	—	—	—	—	—	—	—	—	3.7	3.7	3.7
ISR/FSR	2.2	2.2	—	—	—	—	—	—	—	—	—5.9	—5.9	—5.9
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	9.3	5.9	6.8	6.6	6.7	6.9	7.2	6.3	8.4	—	6.3	6.4	6.4
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
W +Jet and QCD	—	—	1.8	1.8	1.8	—	—	—	—	—36.0	—	—	—

Table B.4: Residual normalization systematic uncertainties: three jets with 2- b -tags. These values are used as input parameters to the BDT analysis.

UNCERTAINTIES FOR CHANNELS WITH FOUR JETS

SINGLE TAG

Source	Uncertainties %												
	$t\bar{l}l$	$t\bar{l}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	tb	t	$tb + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cross section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	0.3	0.3	—	—	—	0.8	—	—	—	—	0.9	0.9	0.9
Lepton ID	2.8	2.8	—	—	—	2.6	2.5	2.4	2.7	—	2.7	2.8	2.8
Jet fragmentation and higher order	-7.0	-7.0	—	—	—	—	—	—	—	—	4.7	4.7	4.7
ISR/FSR	0.8	0.8	—	—	—	—	—	—	—	—	-10.9	-10.9	-10.9
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	15.9	11.8	14.9	9.2	14.4	10.6	13.4	10.7	9.5	—	12.6	12.1	12.2
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
W +Jet and QCD	—	—	1.8	1.8	1.8	—	—	—	—	-33.8	—	—	—

Table B.5: Residual normalization systematic uncertainties: four jets with 1- b -tag. These values are used as input parameters to the BDT analysis.

DOUBLE TAG

Source	Uncertainties %												
	$t\bar{l}l$	$t\bar{l}lj$	$Wb\bar{b}$	$Wc\bar{c}$	Wlp	$Zb\bar{b}$	$Zc\bar{c}$	Zlp	dibosons	QCD	tb	t	$tb + t$
Luminosity	6.1	6.1	—	—	—	6.1	6.1	6.1	6.1	—	6.1	6.1	6.1
Cross section	9.0	9.0	—	—	—	3.3	3.3	3.3	7.0	—	3.8	5.3	4.8
Branching fraction	1.5	1.5	—	—	—	—	—	—	—	—	1.5	1.5	1.5
PDF	—	—	—	—	—	—	—	—	—	—	2.0	2.0	2.0
Triggers	5.0	5.0	—	—	—	5.0	5.0	5.0	5.0	—	5.0	5.0	5.0
Luminosity	1.0	1.0	—	—	—	1.0	1.0	1.0	1.0	—	1.0	1.0	1.0
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Primary Vertex	1.4	1.4	—	—	—	1.4	1.4	1.4	1.4	—	1.4	1.4	1.4
Reweighting	—	—	—	—	—	—	—	—	—	—	—	—	—
Color Reconnection	1.0	1.0	—	—	—	—	—	—	—	—	1.0	1.0	1.0
b/light Jet Response	0.3	0.3	—	—	—	0.8	—	—	—	—	0.9	0.9	0.9
Lepton ID	2.8	2.8	—	—	—	2.6	2.6	2.4	2.7	—	2.8	2.8	2.8
Jet fragmentation and higher order	-7.0	-7.0	—	—	—	—	—	—	—	—	4.7	4.7	4.7
ISR/FSR	0.8	0.8	—	—	—	—	—	—	—	—	-10.9	-10.9	-10.9
b-jet fragmentation	2.0	2.0	—	—	—	2.0	—	—	—	—	2.0	2.0	2.0
Taggability	11.7	8.2	9.8	6.4	8.8	9.7	18.4	21.1	7.7	—	8.6	14.1	11.7
λ_{HF}^W	—	—	12.0	12.0	—	—	—	—	—	—	—	—	—
λ_{HF}^Z	—	—	—	—	—	12.0	12.0	—	—	—	—	—	—
W +Jet and QCD	—	—	1.8	1.8	1.8	—	—	—	—	-34.3	—	—	—

Table B.6: Residual normalization systematic uncertainties: four jets with 2- b -tags. These values are used as input parameters to the BDT analysis.

B.2 Shape and Normalization Uncertainties

The systematic uncertainties that change both the normalization and the shape of the kinematic distributions are: jet reconstruction and identification, (JetID), jet energy scale (JES), jet energy resolution (JER), primary vertex confirmation (VC), V +jets angular corrections (AC), b tag scale factors (BTag). These systematics are divided into two groups. The first being corrections that are applied before event selection and the other after event selection. The JetID, JES, JER, and VC corrections are all applied before event selection and change the amount of events passing selection, event weight and jet multiplicity. The AC and BTag corrections are applied after selection and only change event weight.

Figures [B.1-B.4] show the effects of each of these uncertainties on the boosted decision tree output discriminants for $s+t$, s and t channel analyses. The AC systematic is not included because the change in shape is very small.

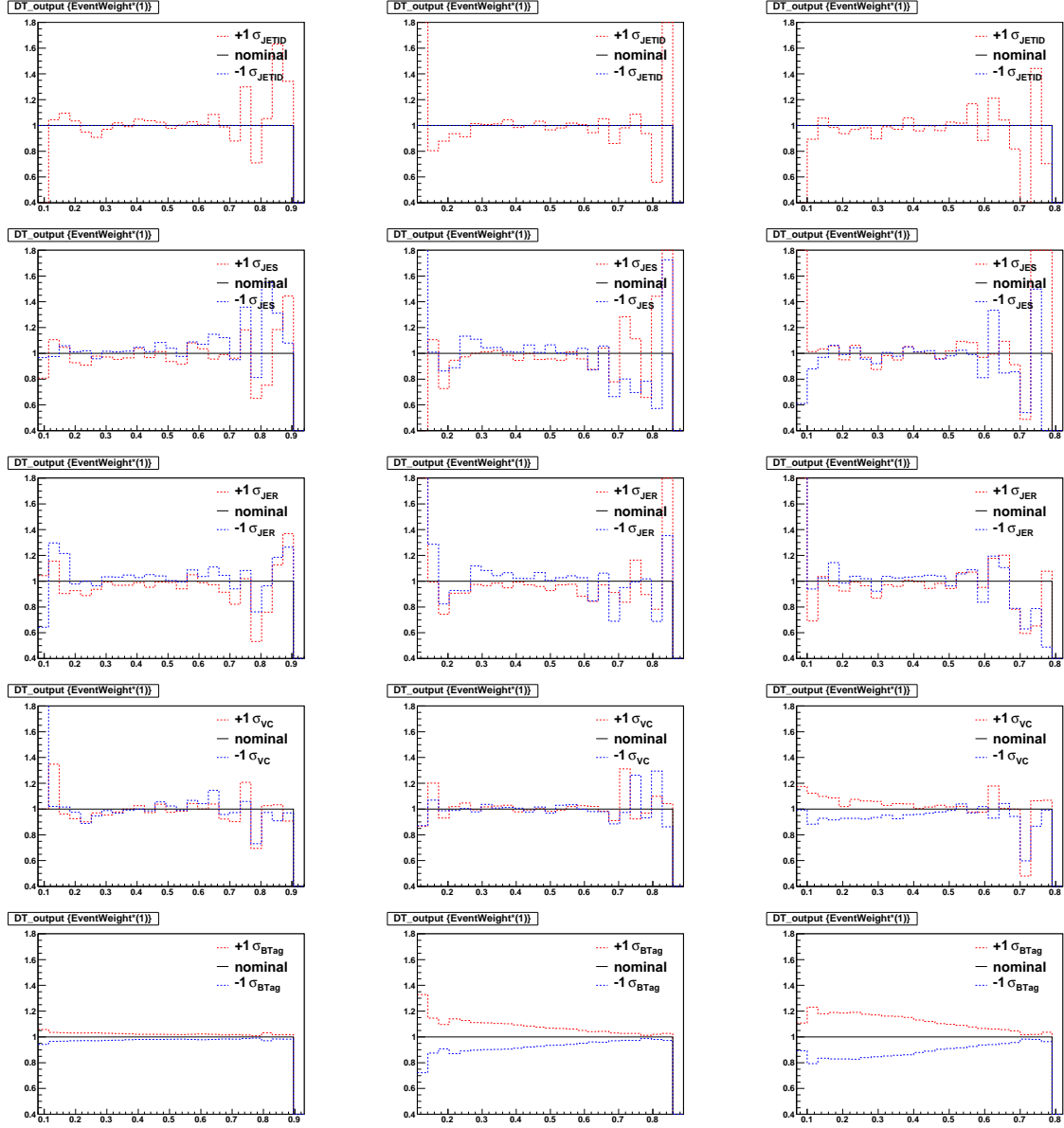


Figure B.1: Combined $s+t$ channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

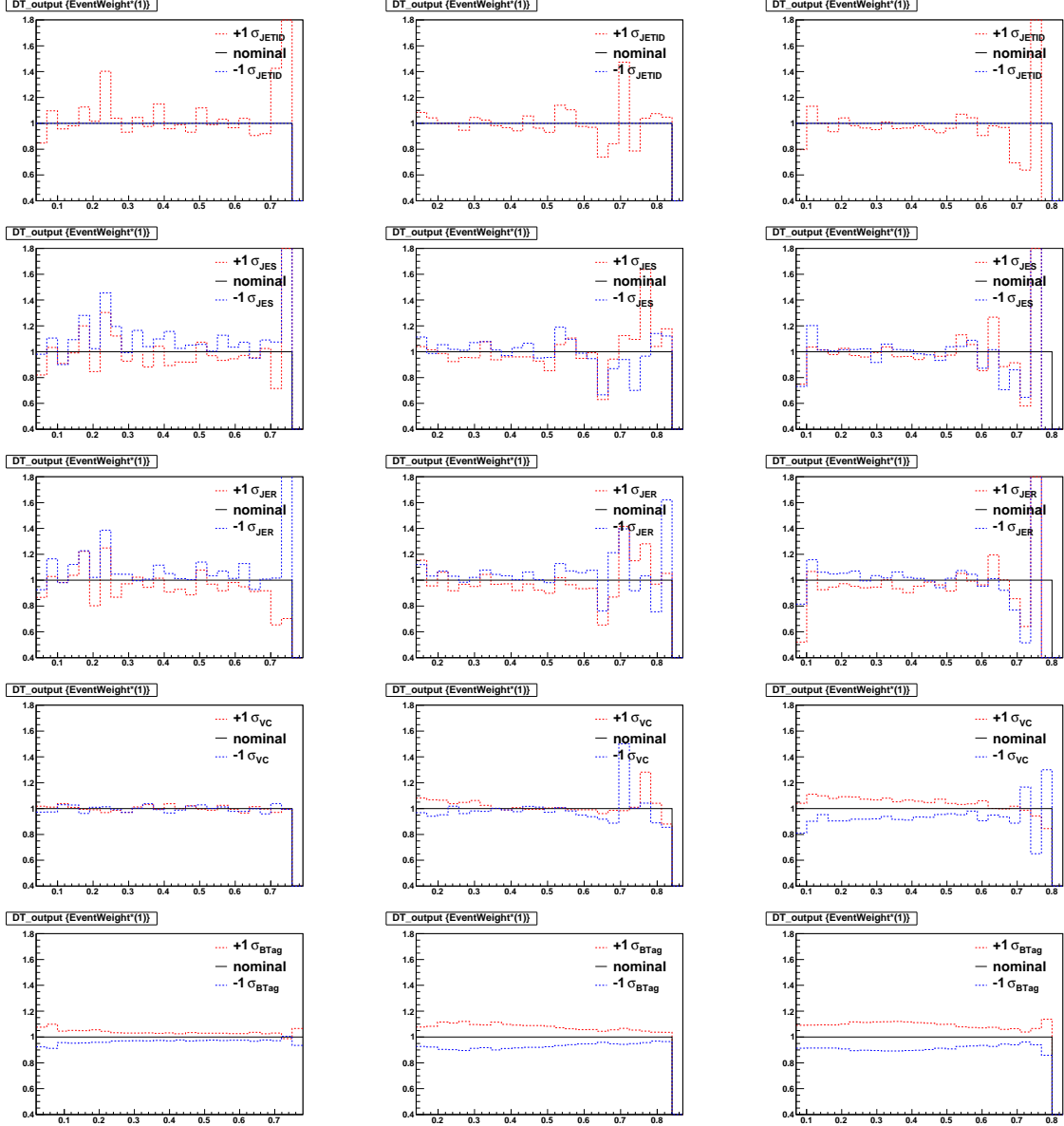


Figure B.2: Combined $s + t$ channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

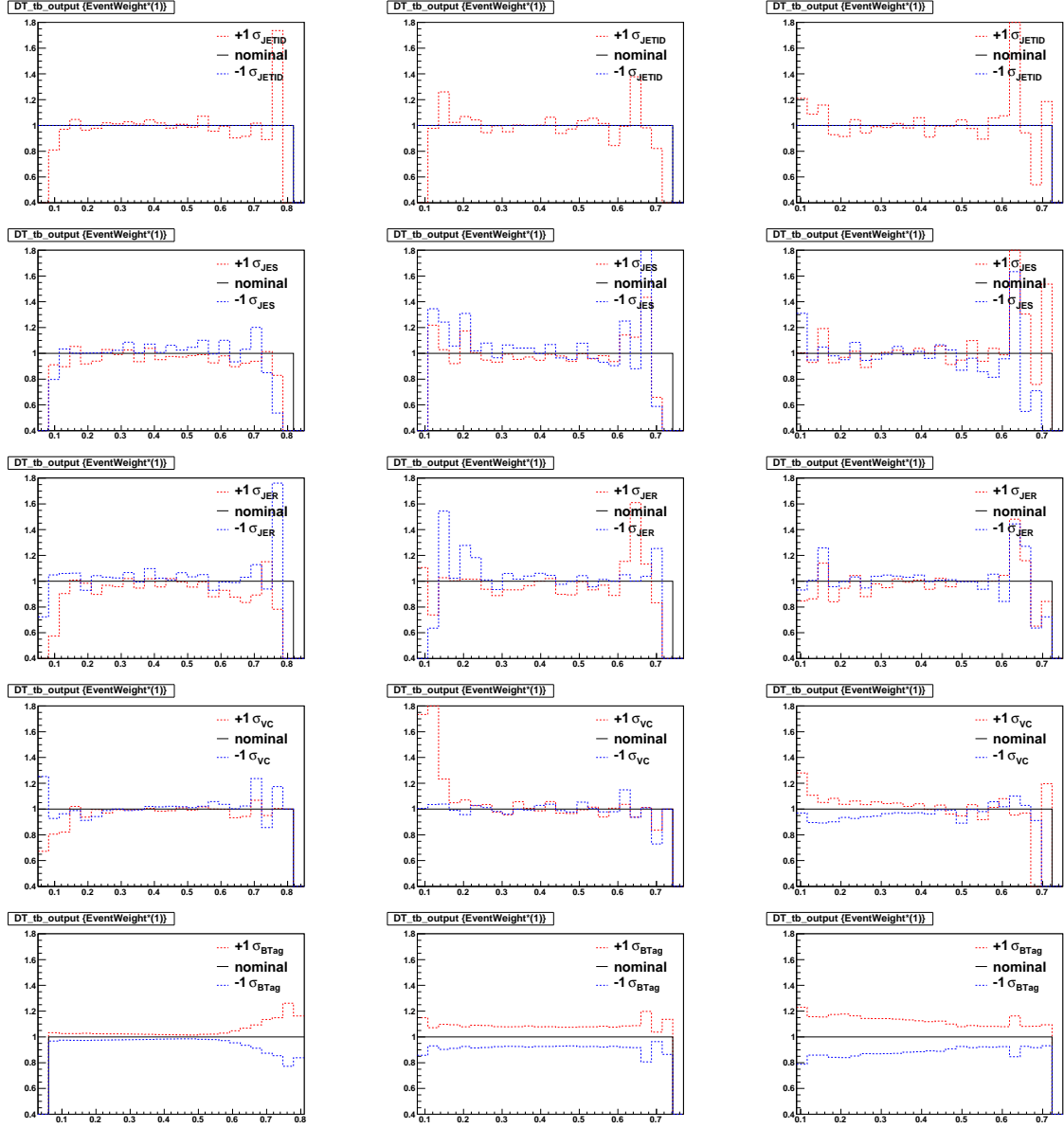


Figure B.3: Combined s channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

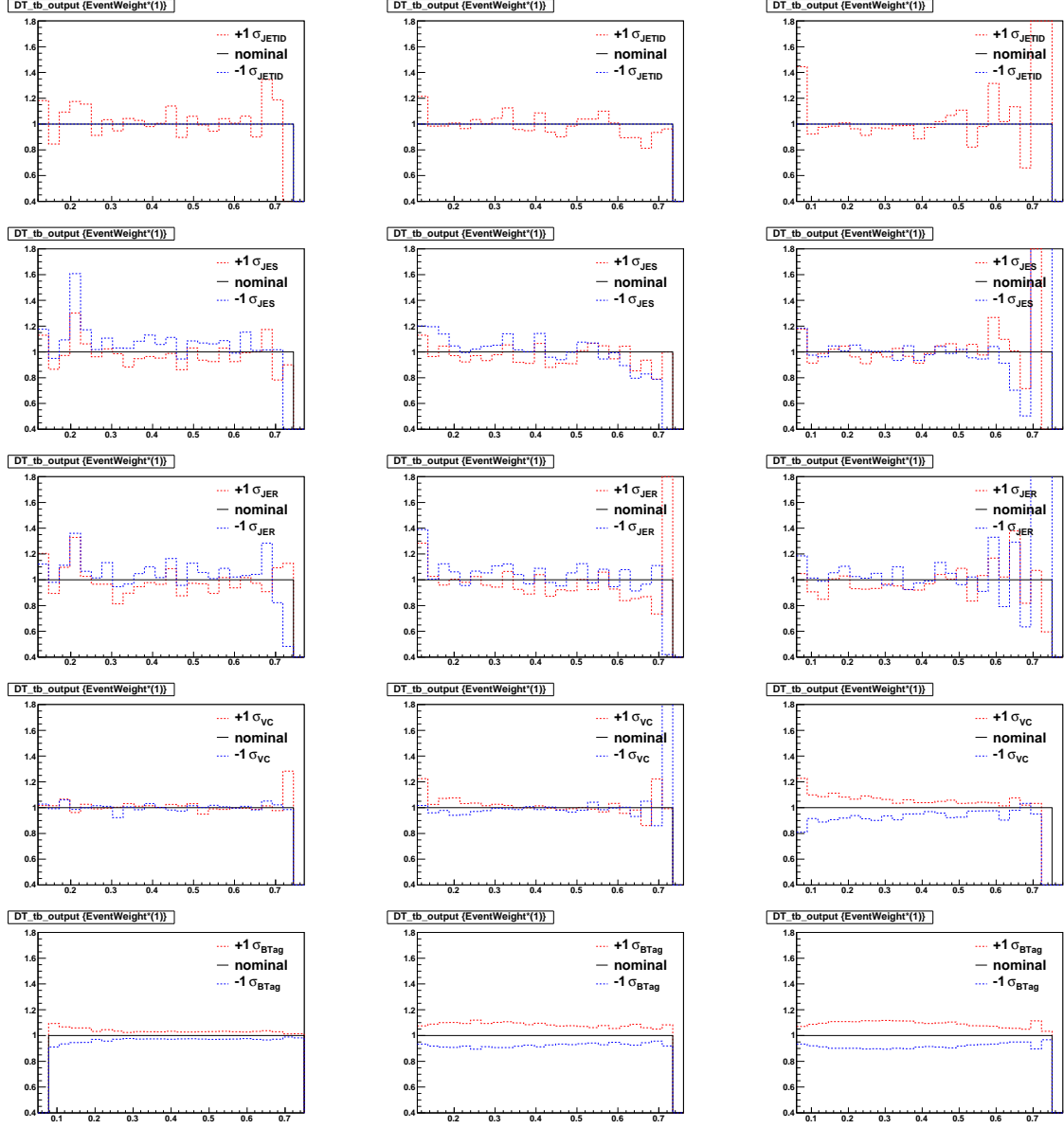


Figure B.4: Combined s channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

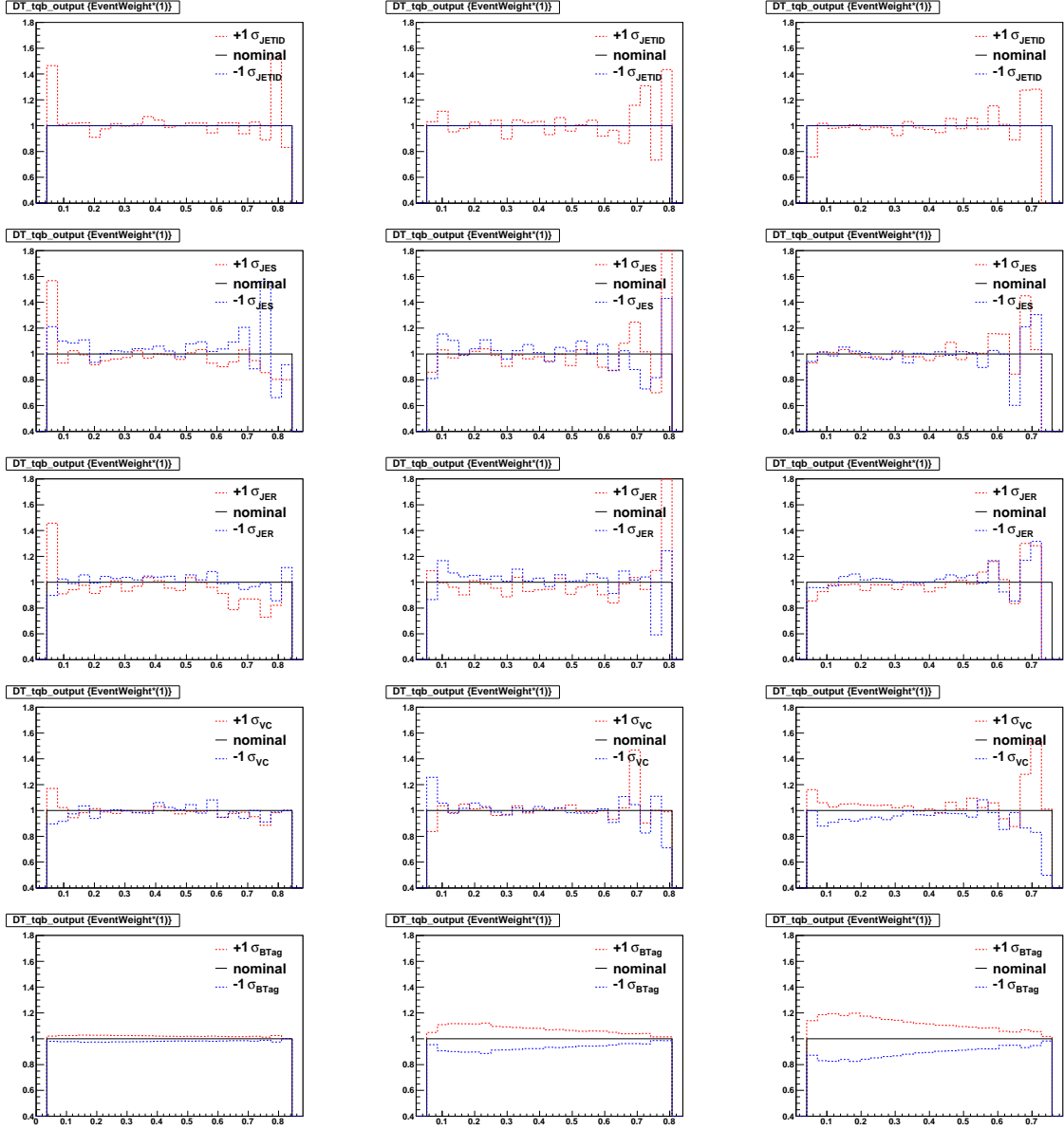


Figure B.5: Combined t channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a single b tag. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

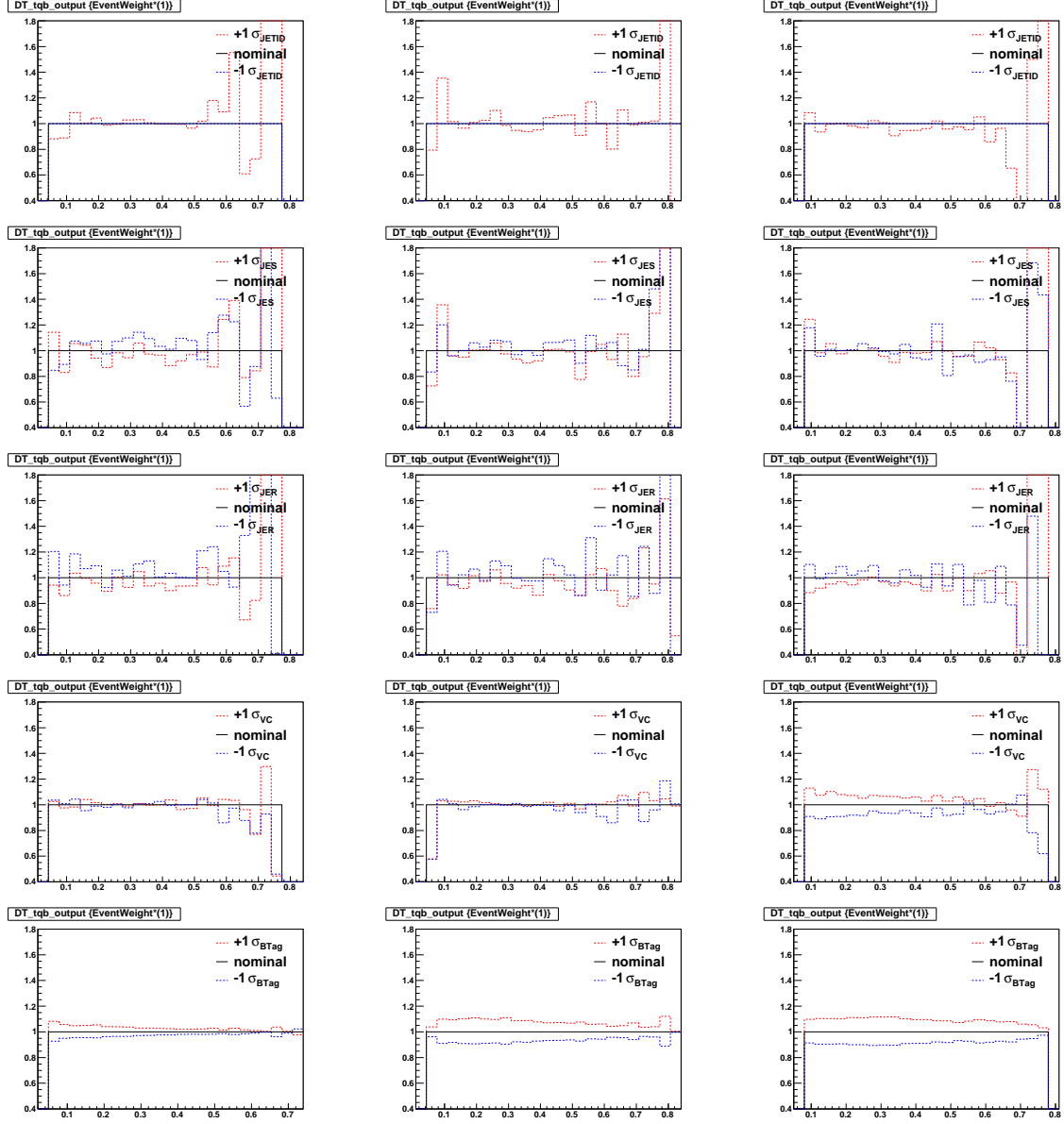


Figure B.6: Combined t channel discriminant, showing the effects of a ± 1 fluctuation in each systematic. The plots are for events with a two b tags. The left column is for the 2 jet events, the center column is for 3 jet events and the right column is for 4 jet events.

Bibliography

- [1] K. Nakamura *et al.* (Particle Data Group), “The Review of Particle Physics” J. Phys. G **37**, 075021 (2010).
- [2] Wikipedia. Standard model wikipedia, the free encyclopedia, 2011. [Online; accessed 16-April-2011].
- [3] M. Kobayashi, T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction”, Progress of Theoretical Physics **49** (2): 652, (1973).
- [4] S. W. Herb *et al.*, “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions”, Phys. Rev. Lett. **39**,252(1977).
- [5] S. Abachi *et al.*($D\bar{O}$ Collaboration), “Observation of the Top Quark”, Phys. Rev. Lett. **74**, 2632(1995); F. Abe *et al.*(CDF Collaboration), “Observation of the Top Quark Production in $p\bar{p}$ Collisions”, Phys. Rev. Lett. **74**,2626(1995).
- [6] V. Avazov *et al.*($D\bar{O}$ Collaboration), “Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}=1.96$ TeV,” FERMILAB-PUB-10/544-E, Submitted to Phys. Rev. D (2010).
- [7] A. Quandt. “Top quark physics at hadron colliders”. The European Journal, C48:835, (2006).

- [8] C. Amsler *et al.*, (Particle Data Group), Review of particle physics, Phys. Lett. B 667, 1 (2008).
- [9] A. Heinson, 2011. <http://www-d0.fnal.gov/heinson/>
- [10] V.M. Abazov *et al.*(*DØ* Collaboration), “Observation of Single Top Quark Production,” Phys. Rev. Lett. **103**, 092001 (2009).
- [11] N. Kidonakis, “Single Top Quark Production at the Fermilab Tevatron: Threshold Resummation and Finite-Order Soft Gluon Corrections,” Phys. Rev. D **74**, 114012 (2006).
- [12] A. Heinson, *et al.*, “Single top quarks at the Fermilab Tevatron”, Phys. Rev. D 56, 3114- 3128 (1997).
- [13] G. Mahlon and S.J.Parke. “Improved spin basis for angular correlation studies in single top quark production at the Tevatron”. Phys. Rev., D55:72497254, 1997.
- [14] E. E. Boos and A. V. Sherstnev. “Spin effects in processes of single top quark production at hadron colliders”. Phys. Lett., B534:97105, 2002.
- [15] T. M. P. Tait and C. P. Yuan, “Single top quark production as a window to physics beyond the standard model”, Phys. Rev. D 63, 1014018 (2000).
- [16] “Linac Rookie Book,” version 2.1 (2006).
http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/LINAC_v2.pdf
- [17] “Booster Rookie Book,” version 4.1 (2009).
http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/Booster_V4.1.pdf

- [18] “Main Injector Rookie Book,” version 1.1 (2003).
[http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/
Main_Injector_1.1.pdf](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/Main_Injector_1.1.pdf)
- [19] “Antiproton Source Rookie Book,” version 2.1 (2010).
[http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/
Pbar_2.1.pdf](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/Pbar_2.1.pdf)
- [20] “Recycler Rookie Book,” version 1.42 (2010).
[http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/
Recycler_RB_v1.42.pdf](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/Recycler_RB_v1.42.pdf)
- [21] “Tevatron Rookie Book,” version 2.3 (2009).
[http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/
Tevatron_2.3.pdf](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks/Tevatron_2.3.pdf)
- [22] S. Abachi *et al.* ($D\bar{O}$ Collaboration), “The $D\bar{O}$ Detector”, Nucl. Inst. Methods **A338**, 185 (1994).
- [23] S. Abachi *et al.* ($D\bar{O}$ Collaboration), “The $D\bar{O}$ Upgrade: The Detector and its Physics”, Fermilab Technical Report Fermilab-Pub-96/357-E (1996).
- [24] B. Olivier, U. Bassler, G. Bernardi, “NADA: A New Event by Event Hot Cell Killer,” D Note 3687, (2000).
- [25] J. Vilimant *et al.*, “Technical Description of the T42 Algorithm for Calorimeter Noise Suppression,” D Note 4146, (2003).
- [26] V. M. Abazov *et al.* ($D\bar{O}$ Collaboration), “The Upgraded $D\bar{O}$ Detector”, Nucl. Inst. Methods **A565**, 463 (2006).

- [27] V.M. Abazov *et al.*, “The muon system of the Run II $D\bar{O}$ detector,” Nucl. Instrum. Meth., **A552** 372-398 (2005).
- [28] A. Khanov, “HTF: histogramming method of finding tracks. The algorithm description,” $D\bar{O}$ Note 3778, (2000).
- [29] G. Borissov. Ordering a chaos or .. technical details of aa tracking All D Meeting. 2003. http://www-d0.fnal.gov/global_tracking/talks/
- [30] H. Greenlee, “The D Kalman track fit,” $D\bar{O}$ Note 4303, (2004).
- [31] A. Schwartzman, C. Tully, “Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting,” $D\bar{O}$ Note 4918, (2005).
- [32] A. Schwartzman and M. Narain. Probabilistic primary vertex selection. $D\bar{O}$ Note, **4042**, 2002.
- [33] J. Hays *et al.*, Single Electron Efficiencies in p17 Data and Monte-Carlo, $D\bar{O}$ Note 5025 (2006).
- [34] M. Narain, Electron identification in the $D\bar{O}$ detector, Fermilab Meeting: DPF 92, Batavia, FERMILAB-CONF-93-054-E. (1992).
- [35] J. Kozminski *et al.*, “The Electron Likelihood in p14,” $D\bar{O}$ Note 4449, (2003).
- [36] P. Calfayan *et al.*, Muon identification certification for p17 data. $D\bar{O}$ Note, **5157**, 2006.
- [37] S. Cho *et al.*, Muon identification certification for p20 data. $D\bar{O}$ Note, **5824**, 2008.
- [38] G. Blazey, “Run II Jet Physics,” $D\bar{O}$ Note 3750, (2000).

- [39] E. Bustato, B. Andreui, “Jet Algorithms in the D Run II Software: Description and Users Guide,” *DØ Note* 4457, (2004).
- [40] A. Harel, “Jet ID Optimization,” *DØ Note* 4919, (2006).
- [41] T. Scanlon, “Development and optimization on MC of a Neural Network tool to identify b-jets,” *DØ Note* 4889, (2005).
- [42] S. Bala *et al.*, “Single Top Quark Production in 5.4 fb⁻¹ of Data Signal and Background Modeling and Event Selection”, *DØ Note* 6099, (2011).
- [43] R. Brun and F. Carminati, “GEANT: Detector Description and Simulation Tool,” CERN Program Library Long Writeup, Report No. W5013 (1993).
- [44] Y. Fisyak, J. Womersley, “*DØ*star *DØ* GEANT Simulation of the Total Apparatus Response”, *DØ internal note* 3191.
- [45] E.E. Boos *et al.*, “Method for Simulating Electroweak Top-Quark Production Events in the NLO Approximation: SingleTop Generator,” *Phys. Atom. Nucl.* **69**, 1317 (2006). We use version 4.2p1.
- [46] J. Pumplin *et al.*, “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis,” *J. High Energy Phys.* **0207**, 012 (2002). We used versions CTEQ6M (signals) and CTEQ6L1 (backgrounds).
- [47] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 Physics and Manual,” *J. High Energy Phys.* **0608**, 026 (2006). We used version 6.409.
- [48] S. Jadach *et al.*, “The Tau Decay Library TAUOLA: version 2.4,” *Comput. Phys. Commun.* **76**, 361 (1993). We used version 2.5.

- [49] D.J. Lange, “The EvtGen Particle Decay Simulation Package,” Nucl. Instrum. Meth. **A 462**, 152 (2001). We used a $D\bar{O}$ version similar to version 00-14-05.
- [50] M.L. Mangano *et al.*, “ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions,” J. High Energy Phys. **0307**, 001 (2003). We used ALPGEN version 2.11.
- [51] Combination of CDF and $D\bar{O}$ Results on the Mass of the Top Quark, arXiv:hep-ex/0603039 (2006).
- [52] The Single Top Working Group, “Single Top Quark Production in 2.3 fb^{-1} of Data — Signal and Background Modeling and Event Selection”, $D\bar{O}$ Note 5810 (2011)
- [53] S. Bala *et al.*, “Measurement of the Single Top Production Cross Section in 5.4 fb^{-1} of Data”, $D\bar{O}$ Note 6128, (2011).
- [54] H. Schellman, “The Longitudinal Shape of the Luminous Region at $D\bar{O}$,” $D\bar{O}$ Note 5142, (2006); H. Schellman, “Run IIb Longitudinal Beam Shape,” $D\bar{O}$ Note 5540, (2007).
- [55] L. Wang, S. Eno, M. Sanders and J. Zhu, ”Measurement of the Shape of the Boson Transverse Momentum Distribution using $p\bar{p} \rightarrow Z/\gamma^* \rightarrow ee$ Events at the Tevatron Collider”, $D\bar{O}$ Note 5187 (2006).
- [56] J. Hays *et al.*, “Single Electron Efficiencies in p17 Data and Monte-Carlo using p18.05.00 d0correct,” $D\bar{O}$ Note 5105 (2006)
- [57] O. Atramentov *et al.*, “Electron and Photon Identification with p20 data,” $D\bar{O}$ Note 5761 (2008).

- [58] S. Hche *et al.*, “Matching Parton Showers and Matrix Elements,” hep-ph/0602031.
- [59] D. Gillberg, “Heavy Flavour Removal and Determination of Weighting Factors for ALPGEN W+jets Monte Carlo,” *DØ* Note 5129 (2006).
- [60] N. Makovec and J.-F. Grivaz, “Shifting, Smearing and Removing Simulated Jets”, Internal *DØ* Note 4914 (2005).
- [61] J. BackusMayes, “b-Tagging Studies in $ZH \rightarrow \mu\mu bb$ ” Presentation in b-ID Meeting *DØ* Collaboration Week 12 May, (2010).
- [62] C.E. Gerber, A. Juste, G.J. Otero y Garzon, A. Tanasijczuk, “Studies on the Instrumental Background in electron+jets+MET Final State Selections using p17 Monte Carlo,” *DØ* Note 5955, (2009).
- [63] T. Andeen *et al.*, “The *DØ* experiments integrated luminosity for Tevatron Run IIa,” FERMILAB-TM-2365 (2007)
- [64] S. Klimenko, J. Konigsberg, T.M. Liss, FERMILAB-FN-0741 (2003).
- [65] N. Kidonakis, “Single Top Quark Production at the Fermilab Tevatron: Threshold Resummation and Finite-Order Soft Gluon Corrections,” *Phys. Rev. D* **74**, 114012 (2006).
- [66] Sven Moch and Peter Uwer, “Theoretical status and prospects for top-quark pair production at hadron colliders,” *Phys. Rev. D* **78**, 034003 (2008).
- [67] J. Hays *et al.*, “Single Electron Efficiencies in p17 Data and Monte-Carlo using p18.05.00 d0correct,” *DØ* Note 5105, (2006); O. Atramentov *et al.*, ”Electron and Photon Identification with p20 data,”

- DØ* Note 5761, (2008). NOTE: We use the “re-certified EMID scale factors for p17, Winter and Summer 2009 datasets” provided by the EMID group. https://plone4.fnal.gov/P1/DØWiki/object-id/emid/emcert/Redo_Winter2009Summer2009/description_redo_winter09summer09
- [68] T. Gadfort *et al.*, “Muon Identification Certification for p17 Data,” *DØ* Note 5157, (2006). Oleg Brandt *et al.*, “Muon Identification Certification for the Summer 2009 Extended Dataset (Run IIb-1 and -2)” *DØ* Note 6025 (2010).
- [69] Frederic Deliot, Alexander Grohsjean, Christian Schwanenberger, Elizaveta Shabalina, Michael Wang, Daniel Wicke, Zhenyu Ye “Systematic Uncertainties in Top Quark Measurements,” *DØ* Note 6024, (2010).
- [70] Y. Peters, M. Begel, K. Hamacher, and D. Wicke, “Reweighting of the Fragmentation Function for the *DØ* Monte Carlo,” *DØ* Note 5325, (2007).
- [71] A. Harel, “An estimate of the data over MC, *b* over light jet response uncertainty for RunIIa JES,” *DØ* Note 5654, (2008).
- [72] Yvonne Peters, Frederic Deliot, Christian Schwanenberger, Elizaveta Shabalina, “Measurement of the $t\bar{t}$ Cross Section using Kinematic and *b*-tagging Information in the Lepton+Jets Channel with 5.3 fb^{-1} ,” *DØ* Note 6109, (2010).
- [73] M. Voutilainen, “Jet p_T resolution for Run IIa final JES (v7.2) with dijet J4S jet correction.” *DØ* internal note 5499.
- [74] JES Group, “Jet Energy Scale Determination at *DØ* Run II”, *DØ* internal note 5382.

- [75] Morgan, James N. and John A. Sonquist, “Problems in the Analysis of Survey Data, and a Proposal”, *Journal of the American Statistical Association*, 58:415-435 (1963)
- [76] L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen, “Classification and Regression Trees”, Wadsworth, Stamford, (1984).
- [77] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer system sciences*, **55**:119, 1997.
- [78] B.P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor. “Boosted decision trees as an alternative to artificial neural networks for particle identification”. *Nucl. Instr. and Meth. A*, **543**:577, 2005.
- [79] H.-J. Yang, B.P. Roe, and J. Zhu. “Studies of boosted decision trees for mini-boone particle identification”. *Nucl. Instr. and Meth. A*, **555**:370, 2005.
- [80] The Single Top Working Group. Observation of single top quark production in 2.3 fb^{-1} of data using boosted decision trees. *DØ Note*, **5811**, 2009.
- [81] V.M. Abazov *et al.* (*DØ Collaboration*), “Evidence for production of single top quarks”, *Phys. Rev. D* 78, 012005 (2008).
- [82] S. Jain, H. Prosper and R. Schwienhorst, “Statistical Methods Implemented in the Package `top_statistics`,” *DØ Note* 5817 (2008).
- [83] S. Jain *et al.*, “Computing Limits Using a Bayesian Approach in the Package `top_statistics`,” *DØ Note* 5123 (2006).
- [84] “Model-independent measurement of t-Channel Single Top Quark Production,” *DØ Note* 5954.

- [85] T. Aaltonen *et al.*, “First Observation of Electroweak Single Top Quark Production,” arXiv:0903.0885 [hep-ex].
- [86] G. Jain, K. Crammer, E. Gross and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” 50 arXiv:1007.1727v2 (2010).
- [87] R.M. Neal, “Bayesian Learning of Neural Networks”, (Springer-Verlag, New York, 1996).
- [88] K. O. Stanley and R. Miikkulainen, Evolving Neural Networks through Augmenting Topologies. *Evolutionary Computation* 10(2): 99-127