MEASUREMENT OF TOP QUARK MASS IN THE ALL HADRONIC CHANNEL IN $\sqrt{s} = 1.96$ TeV, $p\bar{p}$ COLLISIONS AT DØ

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

David Wai Kui Lam,

Michael D. Hildreth, Director

Graduate Program in Physics
Notre Dame, Indiana

April 2008

© Copyright by

David Wai Kui Lam

2008

All Rights Reserved

MEASUREMENT OF TOP QUARK MASS IN THE ALL HADRONIC CHANNEL IN $\sqrt{s} = 1.96$ TeV, $p\bar{p}$ COLLISIONS AT DØ

Abstract

by

David Wai Kui Lam

A measeurement of the top quark mass in proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV using $1040 fb^{-1}$ of data collected in DØ detector at Fermilab is presented. This analysis focuses on the all-hadronic decay mode of the top quark and therefore only events with six or more calorimeter jets in the final state are considered. The measured top masses in the six jets and seven or more jets events are:

$$m_{top}^{6j} = 187.3^{+6.4}_{-5.7}(\text{stat.})^{+3.2}_{-4.0}(\text{syst.})\ \text{GeV}/c^2 \tag{1}$$

$$m_{top}^{7+j} = 173.8^{+9.4}_{-7.6}(\text{stat.})^{+3.1}_{-7.4}(\text{syst.})\ \text{GeV}/c^2 \tag{2}$$

The results from both cases can be combined to give:

$$m_{top} = 183.6^{+5.3}_{-4.8}(\text{stat.})^{+3.2}_{-4.6}(\text{syst.})\ \text{GeV}/c^2 \tag{3}$$

Dedication

To my father and mother for their continuing support and encouragement.

CONTENTS

FIGURES vii

TABLES xvii

ACKNOWLEDGMENTS xix

SYMBOLS xxi

CHAPTER 1: INTRODUCTION 1
1.1 Standard Model 1
1.1.1 Quantum Chromodynamics 2
1.1.2 Electroweak interaction 4
1.2 Top Quark Production at the Tevatron 6
1.3 Top Quark decay 6
1.4 Top mass definition 11
1.5 Top relation to Higgs 13

CHAPTER 2: EXPERIMENTAL APPARATUS 19
2.1 Fermilab Tevatron 19
2.1.1 Pre-Accelerator 19
2.1.2 Antiproton Source 21
2.1.3 Tevatron 22
2.2 DØ detector 22
2.2.1 DØ coordinate system 24
2.2.2 The Silicon Microstrip Tracker, SMT 25
2.2.3 The Central Fiber Tracker, CFT 25
2.2.4 The Solenoid 26
2.2.5 Preshower Detector 28
2.2.6 The Calorimeters 28
2.2.7 The Muon System 30
2.2.8 The DØ Trigger system 32

CHAPTER 3: OBJECT IDENTIFICATION 36
3.1 Tracks 36
3.2 Primary Vertex 37
3.3 Muons 39
3.4 Jets 39
3.4.1 The Cone Cluster Algorithm 40
3.4.2 Infrared safety requirement 41
3.4.3 Jet Merging and Splitting 41
3.4.4 Jet quality requirement 42
3.4.5 Jet Energy Scale 44
3.4.6 Muon Correction in JES 46
3.4.7 Jet Energy Resolution 46
3.5 Electrons 53
3.6 b-tagging 53
3.6.1 Counting Signed Impact Parameter CSIP 54
3.6.2 Jet Lifetime Impact Parameter JLIP 54
3.6.3 Secondary Vertex Tagging SVT 56
3.6.4 Neural Network Tagging NN 56

CHAPTER 4: EVENT SELECTION, WEIGHTS AND TRIGGERS . . . 63
4.1 Data samples 63
4.1.1 Data Quality requirements 63
4.2 Trigger details and simulation 64
4.2.1 Trigger Efficiencies 68
4.2.2 L1 Trigger Turn-ons 69
4.2.3 L2 Trigger Turn-ons 70
4.2.4 L3 Trigger Turn-ons 70
4.2.5 L4 Trigger 77
4.2.6 Trigger simulation in Monte Carlo simulation 79
4.3 Luminosity re-weighting 79
4.4 Event Selection 82
4.4.1 Primary Vertex Position 82
4.4.2 Lepton veto 82
4.4.3 Jet quality and p_T cuts 83
4.4.4 b-tagging requirement 84

CHAPTER 5: SIGNAL AND BACKGROUND STUDY 87
5.1 MC Jet Energy Scale 88
5.1.1 Sample dependent correction 89
5.2 Jet Combinatorics 89
5.2.1 best chi-square selection 91

5.2.2 jet combination 92
5.2.3 Mass distribution 103
5.3 Template fitting 107
5.3.1 Signal template 107
5.4 Background Modeling 132
5.5 Verifying the Background Model 135
5.6 Background template fitting 136

CHAPTER 6: DEVELOPMENT of DISCRIMINANT 144
6.1 Topological variables 144
6.1.1 H_T and Centrality 144
6.1.2 Momentum tensor, Sphericity and Aplanarity 145
6.1.3 Fox-Wolfram moments 146
6.1.4 dR_{bb}, dR_{max}, dY_{01}, dY_{02}, dY_{12} and $cos\theta^*$ 147
6.1.5 $gluonM$ 147
6.1.6 M_{bb}/M_3 and p_{Tn}/M_4 147
6.2 Variables Correlation 148
6.3 Discriminant methods 150
6.3.1 Likelihood 151
6.3.2 Fisher Discriminant 152
6.3.3 Artificial Neural Network 155
6.3.4 Evaluation of the discriminant 157
6.4 Purity Extraction and mass bias 160

CHAPTER 7: LIKELIHOOD METHOD AND ENSEMBLE TEST 171
7.1 Likelihood method 171
7.2 Ensemble test 172
7.2.1 2-D Likelihood test 174
7.2.2 1-D Ensemble test 178
7.2.3 Toy statistics experiment 187

CHAPTER 8: SYSTEMATICS AND STATISTICAL STUDIES 192
8.1 Likelihood Cut 192
8.2 Jet Energy Scale 197
8.3 Signal Purity 200
8.4 Parton distribution function 203
8.5 b-fragmentation 205
8.6 Signal Modeling 206
8.7 Background Modeling 209
8.8 Signal fitting 215
8.9 Background fitting 219
8.10 Statistical Error 229

CHAPTER 9: RESULTS AND OUTLOOK 236

APPENDIX A: Trigger Properties 240
A.1 3JT Triggers 240
A.2 4JT Triggers 244

BIBLIOGRAPHY 249

FIGURES

1.1 Feynman diagrams of leading order top quarks pair production at hadron colliders generated using MadGraph [16]. Graph 1 represents quark annilhilation and Graph 2 and 3 represent gluon fusion. 7

1.2 $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV as a function of top mass m. 8

1.3 Feynman Diagrams of the three different decay channels of $t\bar{t}$ . . . 10

1.4 Virtual top quark one loop contributions to the W and Z mass . . 14

1.5 Virtual Higgs boson one loop contributions to the W mass 15

1.6 Theoretical constraint of the Higgs mass in SM and MSSM using the W and top pole mass [28]. The blue ellipse is the 68% confidence level from the combined Tevatron measurements as of Spring 2008 17

1.7 World average of the top mass mesurement in different decay channels from CDF and DØ in March 2008 [29]. 18

2.1 Schematic view of the Tevatron located in Fermilab, Batavia, IL. Different components including the LINAC, Booster, pbar source, Main Injector and Tevatron are shown. 20

2.2 Cut through view of the DØ detector. 23

2.3 3D layout of the Silicon Microstrip Tracker (SMT) 26

2.4 The central tracking region of the DØ detector 27

2.5 A cut-off view of the central and end cap calorimeters 29

2.6 The schematics of a calorimeter cell 30

2.7 Schematics of the arrangement of calorimeter cells as towers of different pseudo-rapidity 31

2.8 Expanded view of the muon drift tubes in the A, B and C layers of the muon system 33

2.9 DØ Level 1 and 2 trigger architecture. The arrows indicates the flow direction of the information 34

3.1 Illustration of the effect of infrared radiation in jet algorithm. The picture on the left shows the original two jets. Introduction of radiation merges the 2 jets together. 42

3.2 Offset energy for Rcone=0.5 jets for different primary vertex multiplicities, as a function of jet pseudorapidity from the center of the detector 47

3.3 Illustration of MPF method. It makes use of the good photon p_T resolution of together with the missing E_T to extract the hadronic jet response. 48

3.4 JES MPF response for the MC events (on the left) and for the data events (on the right) using jet cones size of 0.5 49

3.5 JES Eta dependent response for the MC events (on the left) and the data events (on the right) using jet cones size of 0.5 49

3.6 JES Showering in different η_{det} regions for jet cones size of 0.5 . . 50

3.7 JES combined correction factor in different η_{det} regions for jet cones size of 0.5 51

3.8 JES uncertainties in different η_{det} regions for jet cones size of 0.5 . 52

3.9 JLIP tagger probability for data jets and light quark, c-quark and b-quark jets from Monte Carlo events. 55

3.10 Illustration of tagging a jet due to the presence of a secondary vertex within its jet cone. 57

3.11 Tagging efficiency and TRF of b-quark jets with different p_T (left) and different η (right) using 3 different NN cuts: $NN > 0.5$ (old Loose), $NN > 0.65$ (Medium) and $NN > 0.775$ (Tight). The red curve shows the tagging efficiency and the green curve shows the TRF. 59

3.12 Tagging efficiency and TRF of c-quark jets with different p_T (left) and different η (right) using 3 different NN cuts: $NN > 0.5$ (old Loose), $NN > 0.65$ (Medium) and $NN > 0.775$ (Tight). 60

4.1 L2 $H_T > 50$ GeV turn-on curves in data JES corrected H_T (JES correction applied to individual jets in data as described in section 3.4, then $H_T = \sum_i^{jet} p_{Ti}$) for different trigger version v9-v12 71

4.2 L2 $H_T > 90$ GeV turn-on curves in offline JES corrected H_T for different trigger version v9-v12 72

4.3 L2 $H_T > 70$ GeV and $H_T > 75$ GeV turn-on curves in offline JES corrected H_T for different trigger version v13 and v14 73

4.4 L3 $CFTVertex > 20$ GeV turn-on curves in offline primary vertex z position, for different trigger version v12-v14 74

4.5 L3 HT120 turn-on curves in offline JES corrected Jet E_T, for trigger version v13a and v13b 75

4.6 $P(BID[NN, ntag]|IP)$ in trigger version v14a. The three different lines correspond to events having ≥1, ≥2 and ≥3 b-ID NN tags. The horizontal axis shows the different b-ID NN operating points. 78

4.7 L4 $p_T > 15$ GeV turn-ons using JES corrected jet p_T for different detector regions: CC ($|\eta_{det}| < 0.8$), ICR ($0.8 < |\eta_{det}| < 1.2$) and EC ($1.2 < |\eta_{det}| < 2.6$). The red points correspond to JES with muon corretion. 80

4.8 L4 $p_T > 20$ GeV turn-ons using JES corrected jet p_T for different detector regions: CC ($|\eta_{det}| < 0.8$), ICR ($0.8 < |\eta_{det}| < 1.2$) and EC ($1.2 < |\eta_{det}| < 2.6$). The red points correspond to JES with muon corretion. 81

5.1 Jet energy correction factor for $t\bar{t}$ events, for b-quark jets, c-quark jets and light quark jets in the 4 η_{det} regions. 90

5.2 bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 93

5.3 jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 94

5.4 bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. 95

5.5 jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. 96

5.6 bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. 97

5.7 jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. 98

5.8 bjj mass difference between the 2 top candidates from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. 99

5.9 jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 100

5.10 Angular Separation (dR) between the MC parton and the closest RECO jet using MC ttbar samples. 102

5.11 Reconstructed bjj mass vs jj mass distribution from best chi2 jet combination in MC $mass_{top} = 170$ GeV signal sample 105

5.12 Reconstructed bjj Dmass vs jj mass distribution from best chi2 jet combination in MC $mass_{top} = 170$ GeV signal sample 106
5.13 bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 108
5.14 bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. 109
5.15 bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. 110
5.16 bjj Dmass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. 111
5.17 bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 112
5.18 bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. 113
5.19 bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. 114
5.20 bjj mass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. 115
5.21 jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. 116
5.22 jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. 117
5.23 jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. 118
5.24 jj mass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. 119
5.25 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65. 121
5.26 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65. 121
5.27 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2. 122

5.28 Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65. The line shows the function fit and the points are the original histograms from MC ttbar events. 123

5.29 Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65. The line shows the function fit and the points are the original histograms from MC ttbar events. 124

5.30 Template functions for different MC top mass with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2. The line shows the function fit and the points are the original histograms from MC ttbar events. 125

5.31 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in exactly 6 jets, 2 BIDNN > 0.65. 126

5.32 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in exactly 6 jets, 2 BIDNN 0.2-0.65. 127

5.33 Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in 7+ jets, 2 BIDNN > 0.2. 128

5.34 Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65. The line shows the function fit and the points are the original histograms from MC ttbar events. 129

5.35 Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65. The line shows the function fit and the points are the original histograms from MC ttbar events. 130

5.36 Template functions for different MC top mass with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2. The line shows the function fit and the points are the original histograms from MC ttbar events. 131

5.37 Illustration of 6-jet background generation. The softest jet from the 6-jet events is inserted into the 5-jet event if the 5th jet have similar p_T. 134

5.38 1st Jet p_T comparison between the 5-jet data events and the "4+1" background events. 136

5.39 2nd Jet p_T comparison between the 5-jet data events and the "4+1" background events. 137
5.40 3rd Jet p_T comparison between the 5-jet data events and the "4+1" background events. 137
5.41 4th Jet p_T comparison between the 5-jet data events and the "4+1" background events. 138
5.42 5th Jet p_T comparison between the 5-jet data events and the "4+1" background events. 138
5.43 Maximum angular separation (dR) among all the jets between the 5-jet data events and the "4+1" background events. The red cross points are the "4+1" background generated following the procedure mentioned in this section. 139
5.44 Minimum angular separation (dR) among all the jets between the 5-jet data events and the "4+1" background events. 140
5.45 Maximum Rapidity difference (dY) among all the jets between the 5-jet data events and the "4+1" background events. 140
5.46 Rapidity difference (dY) of the leading and 2nd leading jets between the 5-jet data events and the "4+1" background events. 141
5.47 mass(bjj) of all 3-jet combinations comparison between the 5-jet data events and the "4+1" background events. 141
5.48 mass(jj) of 2-jet (un-tagged only) combinations comparison between the 5-jet data events and the "4+1" background events. 142
5.49 Spline fit of $Dmass(bjj)$ using background model fits in 3 different BID NN regions: Exactly 6 jets, 2 BIDNN > 0.65, 2 BINNN 0.2-065, 7+ jets, 2 BIDNN > 0.2. 143

6.1 Topological variables λ_2, λ_3, $cos\theta^*$, FWM_2, FWM_3, FWM_4 between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65 149
6.2 Topological variables $FWM_5 - FWM_9$ and $dRmax$ between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65150
6.3 Topological variables dR_{bb}, dY_{01}, dY_{02}, dY_{12}, M_{bb}/M_3 and p_T0/M_4 between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65 151
6.4 Topological variables p_T0/M_4 - p_T5/M_4 and gluonM between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65152

6.5 Correlation Matrix between different candidate variables, using MC $t\bar{t}$ events with $m_{top} = 170$ GeV, with 6 jets, 2 b-jets with BIDNN > 0.65 153

6.6 Correlation Matrix between different candidate variables, using Background events from the background model, with 6 jets, 2 b-jets with BIDNN > 0.65 154

6.7 An example of neural network being used as a discriminant. . . . 156

6.8 An example of a rejected combination of variables for the discriminant. 158

6.9 Trained discriminant output (normalized) for signal and background from different methods. The signal samples are MC $t\bar{t}$ $m_{top} =$ 170 GeV, exactly 6 jets, 2 b-jets with BIDNN > 0.65. 159

6.10 Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, exactly 6 jets, 2 b-jets with BIDNN > 0.65. 161

6.11 Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, exactly 6 jets, 2 b-jets with BIDNN $0.2 - 0.65$. 162

6.12 Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, 7+ jets, 2 b-jets with BIDNN > 0.2. 163

6.13 Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV. 164

6.14 Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN $0.2 - 0.65$ case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV. 165

6.15 Fitted candidate purity using the 4 different discriminant methods for 7+ jets, 2 b-jets with BIDNN > 0.2 case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV. 166

6.16 Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 140$ GeV. 168

6.17 Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 200$ GeV. 169

6.18 Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are combined MC $t\bar{t}$ with $m_{top} = 155$ to 185 GeV. 170

7.1 Log Likelihood surface as a function of m_{top} and $f_{correct}$. This is a pseudo-experiment with signal purity of 18% 175

7.2 Extracted m_{top} from log likelihood using ensembles of different input m_{top} with standard JES. Different colors corresponds to different input top mass as shown, and each ensemble has 1000 pseudo-experiment. 176

7.3 Mean extracted m_{top} versus input m_{top} using standard JES. 177

7.4 Pull mass distribution from different input m_{top} (160-200 GeV) after calibration using standard JES. 179

7.5 Ensemble test results with calibration applied, using standard JES and $+/-1\sigma$ JES as well. The top left plot compares the resulting (mean extracted - input m_{top}) from 3 different JES. The rest of of the plots shows the mass pull values from 3 different JES. 180

7.6 Ensemble test mass error with calibration applied using standard JES. Different plots correspond to different input m_{top} from 160 to 200 GeV. 181

7.7 Ensemble test $f_{correct}$ with calibration applied using standard JES. 182

7.8 Log Likelihood surface as a function of m_{top}. 183

7.9 Mean extracted m_{top} versus input m_{top} using standard JES. 184

7.10 Ensemble test results with calibration applied in the 1D likelihood, using standard JES and $+/-1\sigma$ JES as well. 185

7.11 Ensemble test mass error with calibration applied using standard JES and 1D likelihood. 186

7.12 Width of mass pull in toy experiment with signal purity = 0.20 and 500 events from top left to bottom right are from different Ensemble size. 188

7.13 Width of mass pull in toy experiment with signal purity = 0.20 and 300 events from top left to bottom right are from different Ensemble size. 190

7.14 Width of mass pull in toy experiment with signal purity = 1.00 and 100 events from top left to bottom right are from different Ensemble size. 191

8.1 Ensemble mass error vs Fisher discriminant cut before mass calibration, for different input top mass in exactly 6 jets, 2 BIDNN > 0.65 case. 194

8.2 Ensemble mass error vs Fisher discriminant cut after mass calibration, for different input top mass in exactly 6 jets, 2 BIDNN > 0.65 case. 195

8.3 Ensemble mass error vs Fisher discriminant cut after mass calibration, for different input top mass in 7 or more jets, 2 BIDNN > 0.2 case. 196

8.4 Ensemble test results with mass calibration applied for standard, $+/-1\sigma$ JES in exactly 6 jets, 2 BIDNN > 0.65 case. 198

8.5 Ensemble test results with mass calibration applied for standard, $+/-1\sigma$ JES in 7 or more jets, 2 BIDNN > 0.2 case. 199

8.6 Ensemble test results with mass calibration applied for central, $+/-1\sigma$ signal purity in exactly 6 jets, 2 BIDNN > 0.65 case. . . 201

8.7 Ensemble test results with mass calibration applied for central, $+/-1\sigma$ signal purity in 7 or more jets, 2 BIDNN > 0.2 case. . . 202

8.8 Ensemble test results with mass calibration applied between PDF CTEQ6.5M and CTEQ6L1 in exactly 6 jets, 2 BIDNN > 0.65 case. 204

8.9 Ensemble test results with mass calibration applied between PDF CTEQ6.5M and CTEQ6L1 in 7 or more jets, 2 BIDNN > 0.2 case. 205

8.10 Ensemble test results with mass calibration between standard and $+1\sigma$ b-fragmentation in exactly 6 jets, 2 BIDNN > 0.65 case. . . 207

8.11 Ensemble test results with mass calibration between standard and $+1\sigma$ b-fragmentation in 7 or more jets, 2 BIDNN > 0.2 case. . . 208

8.12 Ensemble test results with mass calibration applied between Pythia and ALPGEN MC $t\bar{t}$ in exactly 6 jets, 2 BIDNN > 0.65 case. . . 210

8.13 Ensemble test results with mass calibration applied between PDF Pythia and ALPGEN $t\bar{t}$ in 7 or more jets, 2 BIDNN > 0.2 case. . 211

8.14 Background distributions comparison between 4+x and 5+x background, for jet 1-6 p_T. 212

8.15 Background distributions comparison between 4+x and 5+x background, for jet 7-8 p_T, dR_{max} (max. angular separation), dR_{min}, dY_{max} (max rapidity difference) and dY_{01} (between 2 leading jets). 213

8.16 Background distributions comparison between 4+x and 5+x background, for M_{all} (invariant mass of all jets), M_{bb} (between 2 b-jets), H_T and λ_3 (momentum tensor eigenvalue). 214

8.17 $Dmass(bjj)$ distribution comparison between 4+x and 5+x background, in 3 different BID NN regions: Exactly 6 jets, 2 BIDNN > 0.65, 2 BINNN 0.2-065, 7+ jets, 2 BIDNN > 0.2. 216

8.18 Ensemble test results with mass calibration applied between 4+x and 5+x background model in exactly 6 jets, 2 BIDNN > 0.65 case. 217

8.19 Ensemble test results with mass calibration applied between PDF 4+x and 5+x background model in 7 or more jets, 2 BIDNN > 0.2 case. 218

8.20 Alternate signal template fits for $m_{top} = 160 - 200$ GeV in exactly 6 jets, 2 BIDNN > 0.65 case. 220

8.21 Alternate signal template fits for $m_{top} = 160 - 200$ GeV in 7 or more jets, 2 BIDNN > 0.2 case. 221

8.22 Ensemble test results with mass calibration between standard and alternate signal fit in exactly 6 jets, 2 BIDNN > 0.65 case. 222

8.23 Ensemble test results with mass calibration between standard and alternate signal fit in 7 or more jets, 2 BIDNN > 0.2 case. 223

8.24 Alternate Background template fits for the exactly 6 jets, 2 BIDNN > 0.65 case, 6 jets, 2 BIDNN 0.2-0.65 case, and 7 or more jets, 2 BIDNN > 0.2 case. 225

8.25 Ensemble test results with mass calibration between Spline and alternate background fit in exactly 6 jets, 2 BIDNN > 0.65 case. 226

8.26 Ensemble test results with mass calibration between Spline and alternate background fit in 7 or more jets, 2 BIDNN > 0.2 case. . 227

8.27 Mass error comparison with and without Fisher discriminant cut, in exactly 6 jets, 2 BIDNN > 0.65 case. 230

8.28 Mass error comparison with and without Fisher discriminant cut, in 7 or more jets, 2 BIDNN > 0.2 case. 231

8.29 Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 6 jets, 2 BIDNN > 0.65 case. 232

8.30 Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 7 jets, 2 BIDNN > 0.2 case. . 233

8.31 Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 6 jets, 2 BIDNN > 0.65 case with Fisher Discriminant > -0.1. 234

8.32 Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 7 jets, 2 BIDNN > 0.2 case. with Fisher Discriminant > -0.1. 235

TABLES

1.1 The Standard Model Fermions 3

1.2 The Standard Model Bosons 5

3.1 NN input variables ranked in order of separation power 58

3.2 Fraction of 2 b-tagged jets having different quark flavors in the exactly 6 jets $t\bar{t}$ sample. bb: 2 b-quarks, cc: 2 c-quarks, bc: 1 b and 1 c quarks, bq: 1 b-quark and 1 lighter quark, ll: 2 lighter quarks 61

3.3 Fraction of 2 b-tagged jets having different quark flavors in the 7 or more jets $t\bar{t}$ sample. bb: 2 b-quarks, cc: 2 c-quarks, bc: 1 b and 1 c quarks, bq: 1 b-quark and 1 lighter quark, ll: 2 lighter quarks 62

4.1 The different run ranges used defined in terms of run numbers from trigger list v8-v12 65

4.2 The different run ranges used defined in terms of run numbers in trigger version v13-v14 66

4.3 The luminosity and trigger choices for the different run ranges. The luminosity shown is the integrated recorded luminosity after data quality selection 67

4.4 $P(IP)$ term from 4JT triggered data events, 6 offline jets $p_T > 15$ GeV 76

4.5 Summary of event selection cuts 84

4.6 Selection cut efficiencies of the MC $t\bar{t}$ and data events. Inclusive $t\bar{t}$ indicates all decay modes. Culmulative efficiencies are the efficiencies with respect to the first row before any cuts 86

5.1 Mass resolution values used in jet combination chi-square 92

5.2 Percentage of the best chi-square jet triplets having the correct/wrong/not-matched jet combinations 103

6.1 Fitted candidate purity using Fisher discriminant 168

8.1 Effects of Fisher discriminant cut on the signal purity and number of events in the candidate samples 193
8.2 Summary of systematic errors in the 2 BIDNN regions without Fisher Discriminant cut. 228

ACKNOWLEDGMENTS

First of all I owe my special thanks to my thesis advisor Prof. Michael Douglas Hildreth. I really appreicate the degree of freedom given to my research in particle physics. Thank you for providing all the dedication for helping me through the years of physics research, the thesis writing process and the preparation for my defense.

I would also like to thank for the DØ top all hadronic group leader Michael Begel as well, who continuously provides challenges and helps me understand all the different aspects (and loop-holes) in this analysis. I also appreciate all the help from the all hadronic fanatics as well, including Jeroen Hegeman, Hendrik Hoeth and Karel Soustruznik, for walking all the difficult steps toward the goal of the analysis. I would also like to thank Amnon Harel and Elizaveta Shabalina for contributing ideas for this analysis.

Thanks also goes out to the Jet Energy Scale group, especially group convenor Aurelio Juste and Alexander Kupco. I would also like to thank my Jet Energy Scale collaborators: Mikko Voutilainen, Jiri Kvita, Andres Tanasijczuk and Christophe Ochando who devoted lots of time for the improvement of the Jet Energy Scale.

In addition I would want to thank Jadzia Warchol, Drew Alton, Abid Patwa and George Ginther for their patience. They gave me a better understanding of the pieces behind the DØ detector and helped me through the various projects

on the detector. I would also like to thank my fellow DØ colleages such as Peter, Tammy, Dimitri and Jyotsna, for helping me through the long hours in the lab.

Finally, I would like to thank my siblings: Eleyne, Siu Ling and April, for all their continuous encouragement and for supplying good food for my basic survival. Last but not least, I would really appreciate for the help (and the push) from Selma, without you this thesis would never be possible.

SYMBOLS

p_T	transverse momentum
c	speed of light
α_s	strong coupling constant
m	mass
m_{top}, m_t	mass of the top quark
e	elementary charge
E	energy
H_T	sum of transverse momentum
η	pseudo-rapidity
η_{det}	pseudo-rapidity with respect to detector coordinates
$\mathcal{D}, LLH$	Discriminant

CHAPTER 1

INTRODUCTION

Working on science is like working on a jigsaw puzzle. It takes many different small pieces to build up a model to describe our world. This is especially true in high energy physics because in a theoretical sense, the standard model is actually a combination of the models which describe the four fundamental forces and their interactions. On the experimental side, a typical high energy physics experiments is getting so complicated nowadays that it is necessary to have a collaboration of several hundred scientists with different expertise in order to carry out the science.

One of the important pieces in the standard model is the top quark. The massive nature of the top quark compared to the other quarks makes it an interesting candidate to investigate. We shall start with a introduction of the Standard model and then illustrate the properties of the top quark. Then the motivation of measuring the mass of the top quark will be explained.

1.1 Standard Model

It is now a widely accepted concept that there are four fundamental forces (interactions): Electromagnetic, Strong, Weak and Gravity. These four forces yield all the physical phenomena observed in Nature. The Standard Model of

high energy physics is a theoretical model which explains how matter interacts. According to the standard model [1], [2], the world is made up of matter particles and force carriers (which are also particles). Matter particles are all fermions with spin 1/2, and the force carriers are all bosons with integral values of spin. Further more, the matter particles can be categorized into quarks and leptons, which can be further categorized into three different families (or generations).

Table 1.1 shows the list of matter particles together with their known properties. Each particle in Table 1.1 is associated with an anti-particle partner of exactly the same mass, opposite electric charge, color charge and weak charge.

The first generation particles make up the bulk of the observable matter of the universe. Due to the nature of the strong force, quarks have to exist as bound states, typically either a quark anti-quark pair (meson) or three quarks (baryon). For example a proton is a baryon of bound state uud, and the π^+ is a meson with a bound state of $u\bar{d}$. The protons, neutrons can then combine with electrons to form the bulk of the visible universe.

Discovered in 1995 [3], [4], the top quark is the last member of the quark family necessary for the completion of the standard model. Different properties of the top quark have been measured and the main purpose of this thesis is to improve on one of them: the mass measurement of the top quark.

1.1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the quantum theory which describes the strong force, or interactions due to color charges. It can be described by a $SU(3)_c$ field [5] and [6] which is generated by three different color charges: red(r), green(g), blue(b). Each quark carries a color charge and the interaction between quarks is

TABLE 1.1

The Standard Model Fermions

Generation	Particle	Name	Mass(MeV/c^2)	Electric Charge(e)
		Quarks		
1	d	Down	$4.0 \sim 8.0$	-1/3
	u	Up	$1.5 \sim 4.0$	2/3
2	s	Strange	95 ± 25	-1/3
	c	Charm	1250 ± 90	2/3
3	b	Bottom	4250 ± 150	-1/3
	t	Top	$172,600 \pm 1,400$	2/3
		Leptons		
1	e	Electron	0.511	-1
	ν_e	Electron Neutrino	$< 15eV$	0
2	μ	Muon	105	-1
	ν_μ	Muon Neutrino	< 0.17	0
3	τ	Tau	1777	-1
	ν_τ	Tau Neutrino	< 24	0

mediated by eight massless gauge bosons, called gluons. Since gluons themselves carry color charges also, they couple to quarks as well as other gluons. This causes the strong coupling constant to decrease as momentum transfer μ increases [7], [8]:

$$\alpha_s = \frac{4\pi}{(11 - \frac{2n_f}{3}) ln(\frac{\mu^2}{\Lambda^2_{QCD}})} \tag{1.1}$$

In the above equation, n_f is the number of quarks with masses less than μ. This illustrates the idea of "asymptotic freedom" in QCD. As the the energy scale of the interaction increases (which corresponds to the decrease in length scale), the coupling constant vanishes. This create an "anti-screening" effect that quarks which are close together don't "feel" the strong force. If one attempts to separate a quark from such confinement, more energy is required to overcome the increasing coupling constant, and eventually there is enough energy to create another quark-antiquark pair from the vacuum. Therefore for the quarks to be stable they need to be in a color neutral state, which means they have to be color singlet as: $r\bar{r}$, $g\bar{g}$, $b\bar{b}$ or color triplet: rgb or $\bar{r}\bar{g}\bar{b}$.

1.1.2 Electroweak interaction

The electromagnetic interaction between electrically charged particles is described by Quantum Electrodynamics (QED) [9], which involves the photon as the gauge boson. In contrast, the weak interactions involves the 3 other gauge bosons: W^+, W^- and Z^0 as the force carriers. It turns out that the electromagnetic and weak interactions can be unified into electroweak interactions described by an $SU(2)_L \times U(1)_Y$ gauge group [2], with four massless gauge bosons as the generators. By adding the a scalar Higgs field to the electroweak Lagrangian, choosing a particular minimum of the Higgs field simultaneously breaks the symmetry, and

produces Goldstone bosons [12]. The Goldstone bosons will then interact with the massless gauge bosons to yield 3 massive vector bosons and one massless photon. This is the famous Higgs mechanism [10] and [11]. It also introduces a new massive particle: the Higgs boson. Currently, this is the only particle in the Standard model that has not yet been observed experimentally.

The theory predicts the relations between the boson masses and the electroweak coupling constant as well [13]:

$$\frac{m_W^2}{m_Z^2} = 1 - sin^2\theta_W \tag{1.2}$$

where θ_W is the weak mixing angle which determines the respective coupling strengths between the SU(2) and U(1) groups in the electroweak Lagrangian. Table 1.1.2 summarizes the properties of gauge bosons in the Standard Model.

TABLE 1.2

The Standard Model Bosons

Particle	Name	Force	Mass(GeV/c^2)	Electric Charge(e)
γ	Photon	Electromagnetic	0	0
g	Gluon	Strong	0	0
$W^{\pm}$	W boson	Weak	80.403 ± 0.029	± 1
Z^0	Z boson	Weak	91.1876 ± 0.0021	0

1.2 Top Quark Production at the Tevatron

As of 2007, the only location that is artificially producing top quarks is the Tevatron. The Tevatron is a proton-antiproton accelerator/collider and is able to reach a center-of-mass energy of 1.96 TeV. The dominant production of top quarks is through the strong interaction. The leading order (LO) [14] Feynman diagrams are shown in Figure 1.1. The $t\bar{t}$ quark pair is produced through quark-antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$) and gluon-gluon fusion ($g\bar{g} \rightarrow t\bar{t}$).

At the Tevatron, the $t\bar{t}$ cross section is estimated to be of order of a few pico-barns using perturbative QCD upto Next-to-Next-to-Leading-Order (NNLO) [22]. As shown in Figure 1.2, the cross section also depends on the top mass as well. Roughly 85% of $t\bar{t}$ pairs are produced through quark annilhilation and 15% of $t\bar{t}$ pairs through gluon fusion [15].

1.3 Top Quark decay

The produced top quarks cannot be detected directly, but rather indirectly through their decay products. This is possible due to flavor mixing between different quark states, as described by the CKM matrix [19], [20]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.3}$$

The squared magnitude of each term in the matrix corresponds to the probability of transition of the quark from one flavor to another in weak decay. Note that the off-diagonal elements of the matrix are non-zero, which allows cross-generational interactions of the the quarks. Based on the current assumption that

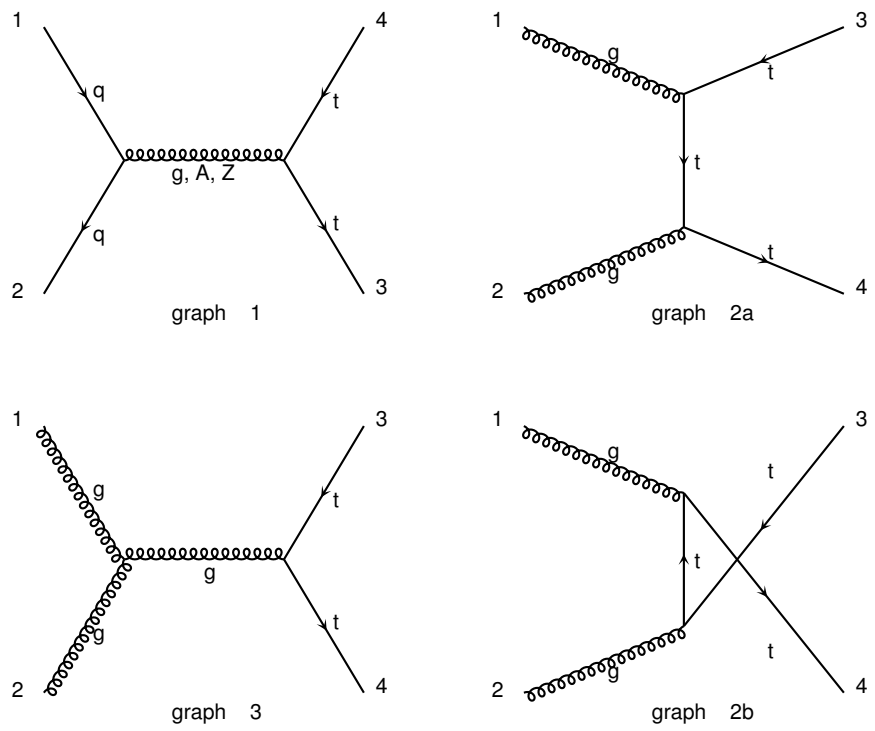

Figure 1.1. Feynman diagrams of leading order top quarks pair production at hadron colliders generated using MadGraph [16]. Graph 1 represents quark annilhilation and Graph 2 and 3 represent gluon fusion.

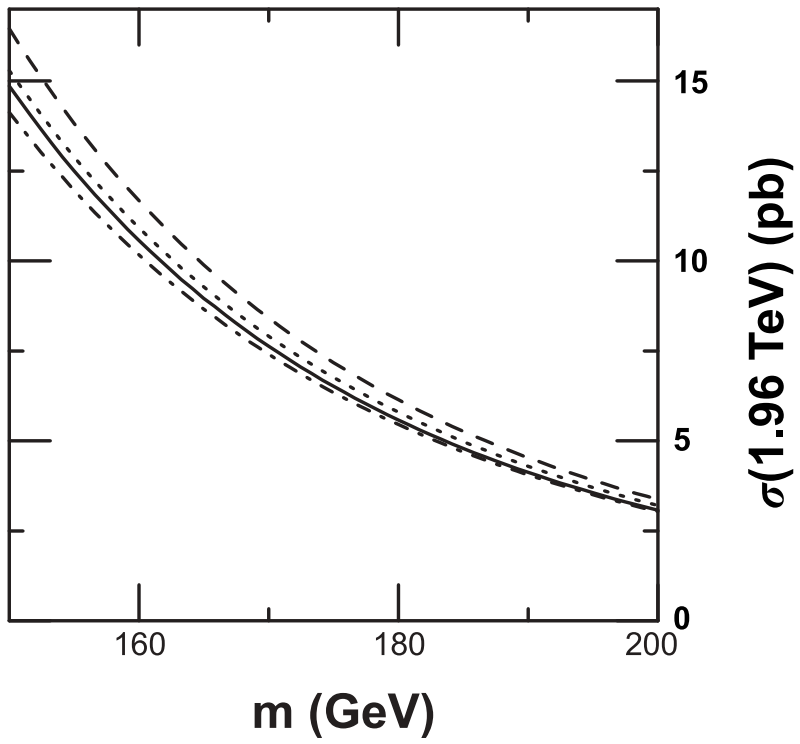

Figure 1.2. $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV as a function of top mass m. The solid line represents the NLO calulation adapted from reference [17]. The approximated NNLO results [22] shown are single-particle-inclusive (dashed), pair-invariant-mass (dot-dashed) and averaged (dotted).

there are only three generations of quarks, the matrix is 3x3 and unitary. This leads to the following constraint [18]:

$$0.9989 < |V_{tb}| < 0.9993 \tag{1.4}$$

This implies a top quark decays almost 100% of the time to a b quark and a W boson. If the mass the b quark and the width of the W boson are neglected, the top decay width can be predicted as [21]:

$$\Gamma(t \to W_b) = \frac{G_F m_t^3}{8\pi\sqrt{2}}|V_{tb}|^2 \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \tag{1.5}$$

Where G_F is the Fermi coupling constant and α_s is the strong coupling constant. For a top mass of 170 GeV the top decay width is approximately 1.5 GeV. This is larger than the strong interaction energy scale $\Lambda_{QCD} = 200$ MeV, which indicates the lifetime of the top quark is less than the QCD timescale. Therefore most of the time the top quark pair decays before it has enough time to form a $t\bar{t}$ bound state or t hadrons.

The resulting W boson can either decay hadronically or leptonically as given by the respective branching ratios. Therefore in $t\bar{t}$ decay, there can be three different decay channels: The dilepton channel is the case where both W bosons decay to leptons. The semi-leptonic channel is the case where one W decays to leptons and the other one decays to quarks. The all-hadronic channel corresponds to the case where both W bosons decay to quarks, and is the channel investigated in this thesis. The three channels are illustrated in the Feynman diagrams in 1.3 below.

In all three channels there are two b quarks. Each b-quark has a longer lifetime,

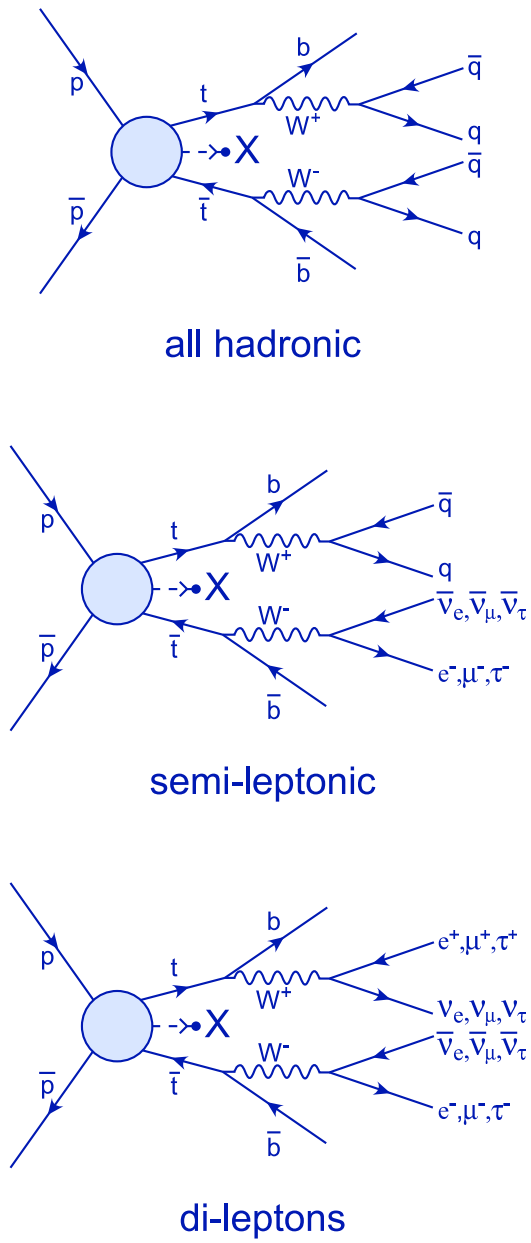

Figure 1.3. Feynman Diagrams of the three different decay channels of $t\bar{t}$

which gives enough time for hadronization and results in a particle "jet" as the decay product. Also a b quark may decay to a lepton such as muon or electrons through weak decay. Such a long lifetime and the occasional presence of a lepton within the particle "jet" serves as a handle to identify b-quarks (a technique known as b-tagging).

1.4 Top mass definition

In quantum field theory (QFT) there are different definitions of mass, and it is important to figure out what is being measured in the experiment. A natural choice is the renormalized pole mass m_{pole}, which corresponds to the pole position of the top quark propagator:

$$D(\bar{p}) = \frac{i}{\bar{p} - m_{pole} - \Sigma(\bar{p})}. \tag{1.6}$$

Here m_{pole} is the pole mass of the fermion with four-momentum $\bar{p}$, and $\Sigma(\bar{p})$ is the renormalized one-particle-irreducible quark self-energy. However, it can be shown that, due to the effects of QCD, the top quark pole mass is unavoidably ambiguous just like other quark masses, with the mass uncertainty estimated to be [23]:

$$\delta m_{pole} \approx \frac{8\pi}{3\beta_0} e^{-C/2} \Lambda_{QCD} \tag{1.7}$$

Where C is a renormalization-scheme-dependent constant, $C = -5/3$. β_0 is the one-loop QCD beta-function coefficient, $\beta_0 = 11 - (2/3)N_f$. This implies the pole mass ambiguity is around 600 MeV.

Due to the appearance of divergences in perturbative calculations, renormal-

ization schemes are invented to absorb the infinities. One of the special choices of renormalized schemes is the $\overline{MS}$ scheme, where a new renormalized mass (or running mass, $\overline{MS}$ mass) and a renormalized coupling constant arise:

$$m_{pole} = m_{\overline{MS}} + \delta m \tag{1.8}$$

$$\alpha_{pole} = \alpha_{\overline{MS}} + \delta\alpha \tag{1.9}$$

where δm and $\delta\alpha_s$ are the terms introduced to absorb the infinities. Ideally, the $\overline{MS}$ mass (or running mass) is the preferred mass definition because it can be measured with arbitrary accuracy. The name "running mass" also indicates that the mass changes depending on the energy scale μ of the interaction.

It can be shown that for a massive fermion propagator, the pole mass of the fermion has the following relation to the $\overline{MS}$ mass [24]:

$$m_{pole} = m_{\overline{MS}}(\mu)\left(1 + c_1\left(\frac{\alpha_s}{4\pi}\right) + c_2\left(\frac{\alpha_s}{4\pi}\right)^2 + ...\right) \tag{1.10}$$

where the 2nd order coefficients are:

$$c_1 = C_F(4 + 3ln\frac{\mu^2}{M^2}) \tag{1.11}$$

$$\begin{aligned} c_2 = {} & C_F C_A\left(\frac{1111}{24} - 8\zeta_2 - 4I_3(1) + \frac{185}{6}ln\frac{\mu^2}{M^2} + \frac{11}{2}ln^2\frac{\mu^2}{M^2}\right) \\ & - C_F t n_f\left(\frac{71}{6} + 8\zeta_2 + \frac{26}{3}ln\frac{\mu^2}{M^2} + 2ln^2\frac{\mu^2}{M^2}\right) \\ & - C_F^2\left(\frac{121}{8} + 30\zeta_2 + 8I_3(1) + \frac{27}{2}ln\frac{\mu^2}{M^2} + \frac{9}{2}ln^2\frac{\mu^2}{M^2}\right) - 12C_F t(1 - 2\zeta_2) \end{aligned} \tag{1.12}$$

where $C_A = N_c$ as the number of color in QCD and $C_F = (N_c^2 - 1)/(2N_c)$,

$t = 1/2$. $I_3(1)$ is derived from the two-loops self energy. $\zeta_n = \sum_{i=1}^{\infty} 1/n^i$ is the Riemann zeta function, with $\zeta_2 = \pi^2/6$.

The Monte Carlo Event generator used in this experiment is Pythia 6.3, which simulates the physics processes of the $t\bar{t}$ events to Leading Order (LO) in perturbation theory. It takes in the top pole mass as the mass parameter in the simulation and therefore the observed top quark mass in this thesis is the measured pole mass. Although the choice of pole mass implies the intrinsic ambiguity of the mass measurement, as we shall see in this thesis, the attainable mass error from this experiment is still large compared to the ambiguity.

1.5 Top relation to Higgs

Rather than being just another parameter in the SM, the determination of the top quark mass is also beneficial to other particle physics measurement as well. As we shall see, the top quark mass can serve as a candlelight in the search of for the Higgs [26] and for physics beyond the SM.

From the definition of the weak mixing angle θ_W above, performing a tree level calculation using the pole mass of the W, m_W, will yield its relation to θ_W:

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F sin^2\theta_W} \tag{1.13}$$

where $\alpha = 1/137.03599976(50)$ is the electroweak coupling constant, $G_F = 1.16637(1) \times 10^{-5}$ GeV^{-2} is the Fermi constant extracted from the muon lifetime, $m_Z = 91.1876(21)$ GeV is the Z mass determined from e^+e^- annihilation. Adding the virtual top quark one-loop contribution as shown in the Feynman diagram in Figure 1.4 would result in an additional correction factor Δr in the W mass:

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F sin^2\theta_W(1-\Delta r)} \tag{1.14}$$

where the correction term Δr is a function of the top quark pole mass:

$$(\Delta r)_{top} \approx -\frac{3G_F m_{top}^2}{8\sqrt{2}\pi^2 tan^2\theta_W} \tag{1.15}$$

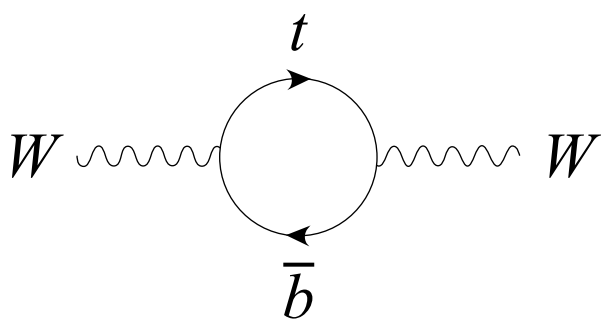

Figure 1.4. Virtual top quark one loop contributions to the W and Z mass

Similarly, the Higgs boson also contributes to Δr via the one-loop diagrams in Figure 1.5:

$$(\Delta r)_{Higgs} \approx -\frac{11G_F m_Z^2 cos^2\theta_W}{24\sqrt{2}\pi^2} ln\frac{m_h^2}{m_Z^2} \tag{1.16}$$

The above equations illustrate the inter-dependence between the W mass, top mass and the Higgs mass. Therefore, the range of possible Higgs mass is limited depending on the precision of the W and top mass we measure. Figure 1.6 shows

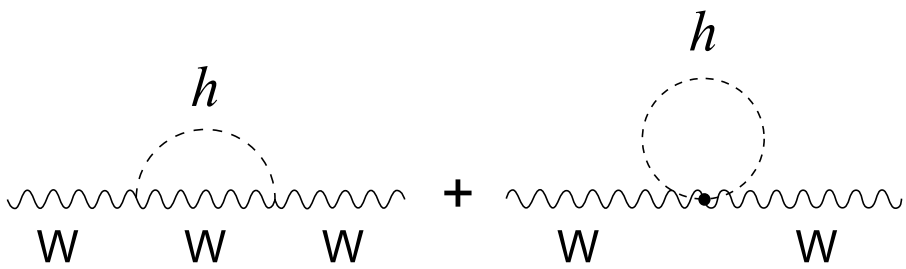

Figure 1.5. Virtual Higgs boson one loop contributions to the W mass

the predicted range of the Higgs mass versus the W and top mass, together with an experimental measurement of the W and top mass from the Tevatron as of Spring 2008.

In addition to the SM Higgs, it is possible to probe the possibility of physics beyond the SM such as the Minimal Supersymmetric Standard Model (MSSM) [27]. There are two doublets of complex Higgs fields in the MSSM which gives rise to five Higgs particles; three of them are electrically neutral. The mass of the lightest scalar SM-like Higgs boson, m_h, is limited to be less than the Z mass:

$$m_h^2 \leq m_Z^2 cos^2 2\beta \tag{1.17}$$

where the angle β is a parameter of the Higgs doublet model and controls the coupling strength between the Higgs particles and the top and bottom quarks. Similar to the one-loop corrections of the W mass, radiative corrections due to the couplings to the top quark result in an increase to the Higgs mass [25]:

$$m_h^2 \leq m_Z^2 cos^2 2\beta + \frac{3}{\pi^2}\frac{m_{top}^4 sin^4\beta}{v^2} ln \frac{m_{\widetilde{top}}}{m_{top}} \tag{1.18}$$

which depends strongly on the top quark mass m_{top}, and is logarithmic in the mass of the top squark, $m_{\widetilde{top}}$, the supersymmetric partner of the top quark [28]. The allowable mass range of the MSSM Higgs is shown in Figure 1.6 as well. It seems like the current measurements from Tevatron favors the MSSM Higgs rather than the SM, at least at the 68% confidence level.

Since the discovery of top quark in 1995, numerous mass measurements of the top quark have been carried at the CDF and DØ experiments in Fermilab. A summary of the top mass measured from different decay channels is shown in Figure 1.7. As of March 2008, the combined world average of the top mass is $m_{top} = 172.6 \pm 1.4$ GeV [29].

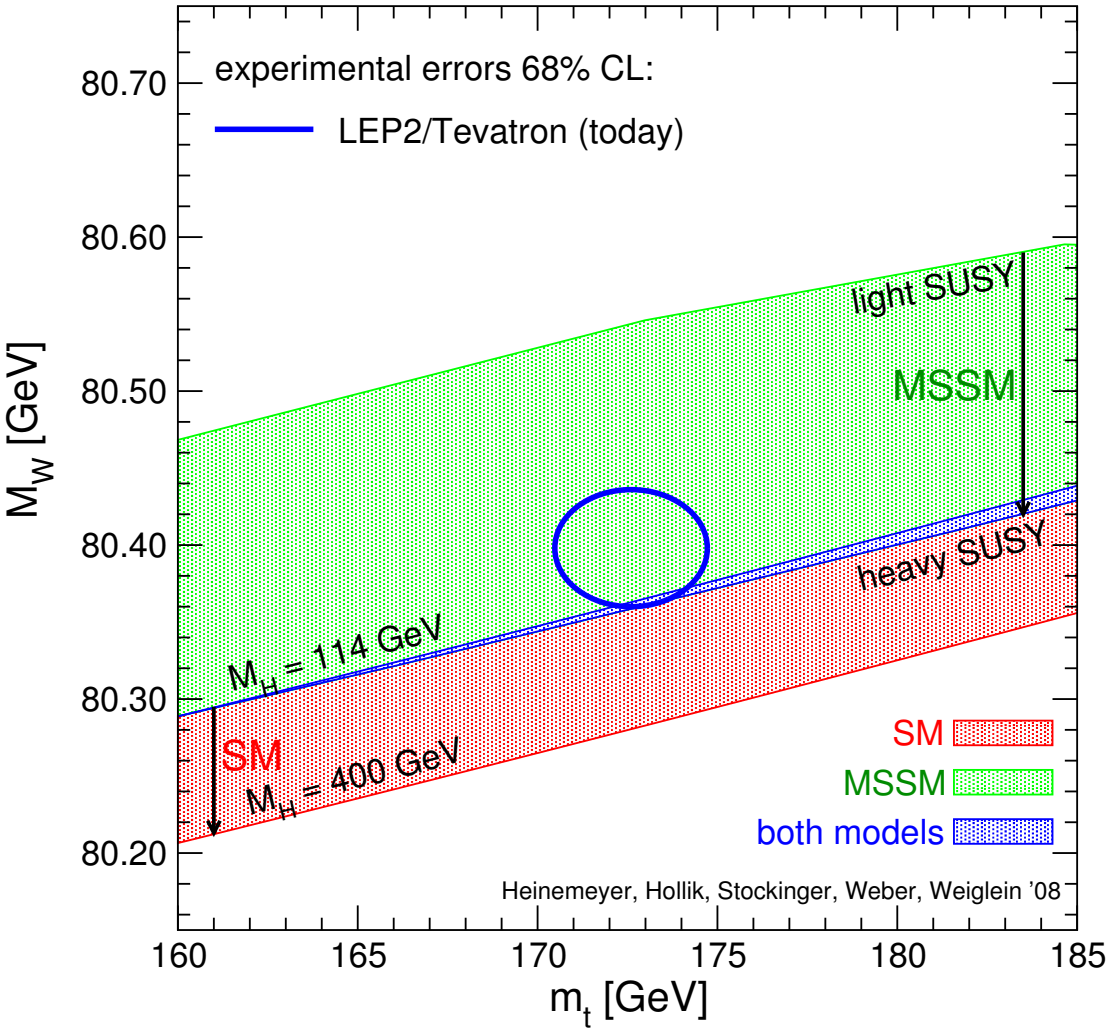

Figure 1.6. Theoretical constraint of the Higgs mass in SM and MSSM using the W and top pole mass [28]. The blue ellipse is the 68% confidence level from the combined Tevatron measurements as of Spring 2008

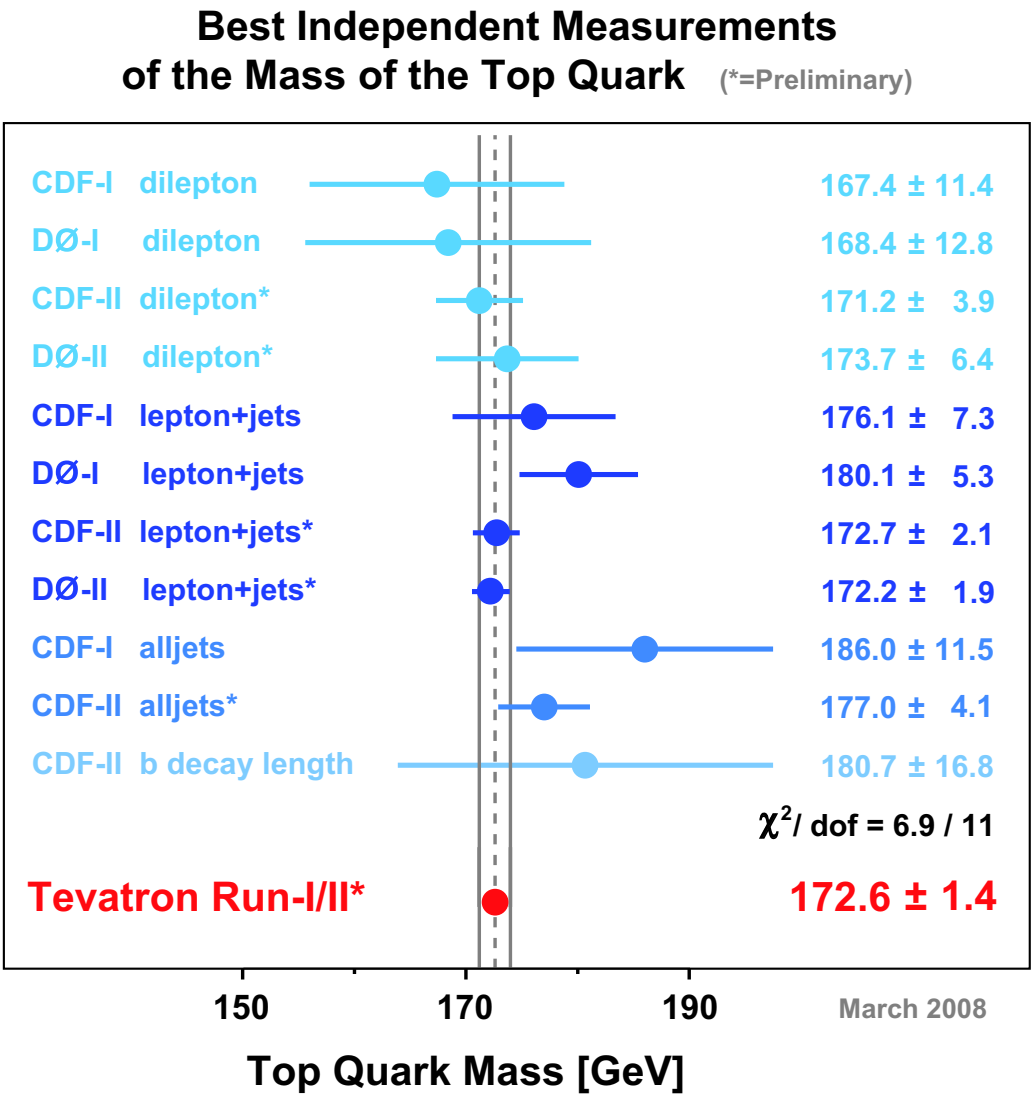

Figure 1.7. World average of the top mass mesurement in different decay channels from CDF and DØ in March 2008 [29].

CHAPTER 2

EXPERIMENTAL APPARATUS

The Fermilab Tevatron is the world's largest proton-antiproton (ppbar) collider and is able to collide particles with the highest center-of-mass energy of 1.96 TeV as of 2007. Currently this synchrotron is also the only laboratory on Earth which can produce top quarks, which will then be detected by the two detectors (CDF and DØ) where the collisions of particles take place. This experiment is based on the results from the DØ detector [30]. The DØ detector is designed to identify different particles emerging from the collisions, including particle jets, electrons, muons, and also the presence of neutrinos indirectly.

2.1 Fermilab Tevatron

A general schematic of the Fermilab Tevatron is shown in Figure 2.1.

The different components of the Fermilab Tevatron will be described briefly in this chapter. For a more detailed desciption please refer to [31].

2.1.1 Pre-Accelerator

The protons used in the collisions are first produced by turning hydrogen gas (H^2) into hydrogen ions (H^-) of 18 keV by the Cockroft-Walton pre-accelerator. The H^- ions are produced by a Magnetron Surface Plasma Source (SPS), where a

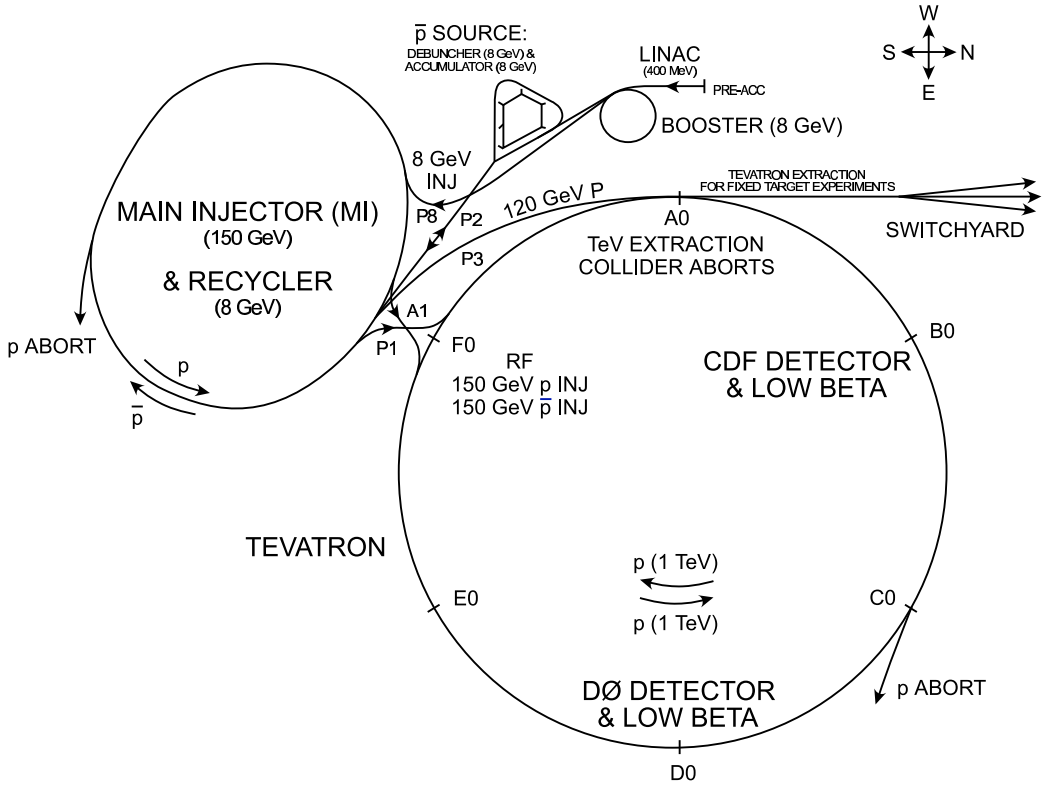

Figure 2.1. Schematic view of the Tevatron located in Fermilab, Batavia, IL. Different components including the LINAC, Booster, pbar source, Main Injector and Tevatron are shown.

constant magnetic field in the magnetron causes the charged particles to circulate in the cavity. The electric field between anode and cathode together with the cesium gas help separate the hydrogen gas into ions. Negatively charged H^-, once they are produced, will be accelerated through the extractor plate and enter the Cockroft-Walton. The Cockroft-Walton converts an AC source to a DC source, which further accelerate the H^- to 750 keV.

The Linac, or linear accelerator, a consists of Alvarez drift-tubes and a series of RF accelerating cavities. Using an alternating electric field, the drift-tubes can accelerate the beam to 116 MeV, while the Side-Coupled cavities accelerate it further to 400 MeV. The H^- ions are then sent through a carbon foil to have the electrons stripped off.

The proton beam is then injected into the Booster synchrotron ring, which is approximately 500 meters in circumference. The synchrotron is a cyclic machine consists of a series of bending magnets, which both confine and accelerate the particles circulating in the ring. The Booster accelerates the protons to 8 GeV and then injects the particles to the Main Injector.

Fermilab's Main Injector is a synchrotron accelerator with a circumference of around 3.3 kilometers. It contains series of dipole magnets for bending of the particle beam, together with quadrupole magnets to focus the beam. The main injector is capable of accelerating both protons and antiprotons to 150 GeV.

2.1.2 Antiproton Source

To generate antiprotons, the 120 GeV protons from the main injector are directed to strike a nickel target and are collected using a liquid lithium "lens" – a cylindrical lithium conductor operated with a focusing magnetic field of 740 T/m.

For every 10^6 incident protons onto the nickel target only about 20 antiprotons are created. The antiprotons then enter the debuncher, where a process called stochastic cooling is applied to minimize the momentum spread and the transverse oscillations of the antiproton beam.

The antiprotons are then transferred to a second storage ring called the accumulator, where further cooling is applied to the antiprotons. Both the debuncher and accumulator reside in the same tunnel of 520m in circumference. Once the antiprotons in the accumulator reaches $80 - 200 \times 10^{10}$ the antiprotons are then fed into the main-injector and further accelerated to 150 GeV.

2.1.3 Tevatron

When there are enough anti-protons in the main injector, the protons and antiprotons are transferred to the Tevatron. The Tevatron is a synchrotron accelerator of nearly 2km in diameter, and consists of a series of superconducting magnets which produce a magnetic field of 4.2 Tesla. The protons and antiprotons circulate the Tevatron in opposite directions and are organized in 3 super-bunches of 12 bunches each. Both beams are then accelerated to an energy of 980 GeV. Finally the beams are focused using quadrupole magnets, having the beam halos removed by a process called "scraping", and brought into the beam crossing points at CDF and DØ .

2.2 DØ detector

The DØ detector [30] is a multipurpose detector designed to study the phenomena arises from the $p\bar{p}$ collisions in the Tevatron. It consists of layers of different sub-detectors and designated to cover a large fraction of 4π solid angle around

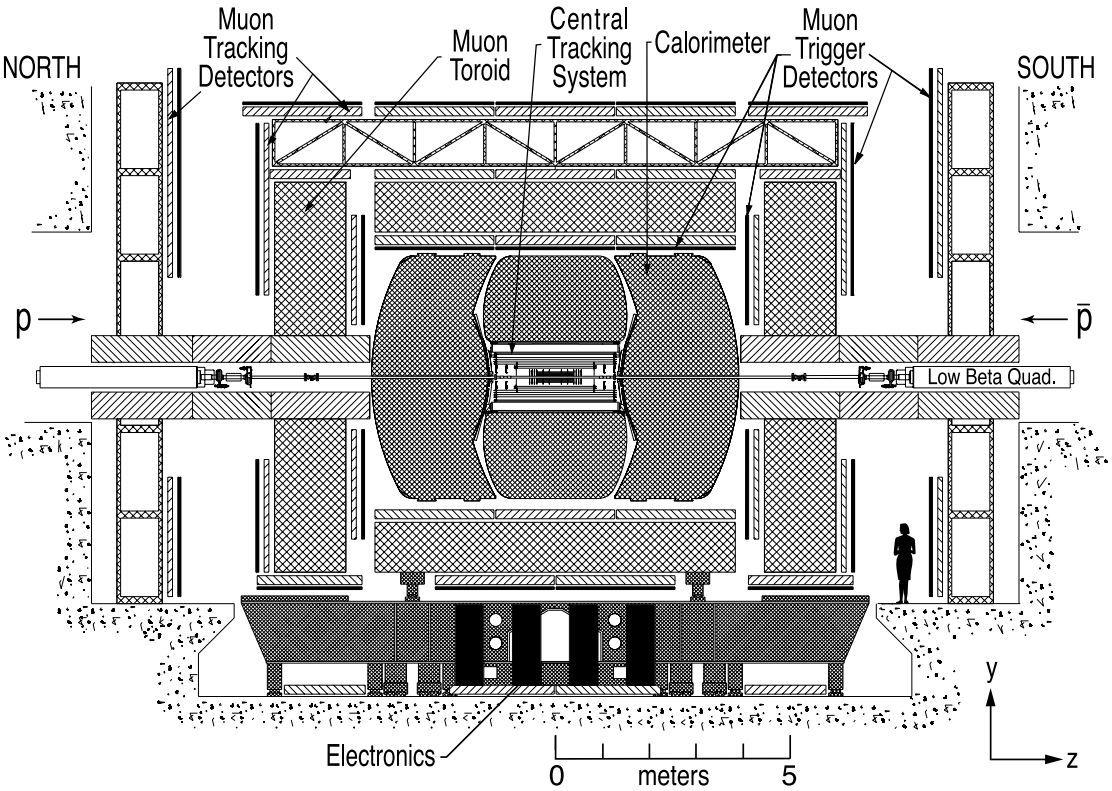

Figure 2.2. Cut through view of the DØ detector.

the collision point. Figure 2.2 shows a cut through view of the DØ detector. The inner most layer includes the central tracker system which is designated to record the trajectory of charged particles. The tracker is surrounded by a solenoid magnet which generates a $2T$ magnetic field. Outside of the solenoid is the preshower scintillating detector and a sampling calorimeter, which records the energy of the particles as they interact with the calorimeter materials. The outermost layer of the detector is the muon detector which is designed to detect the escaping muons.

2.2.1 DØ coordinate system

The DØ experiment employs the usual right-handed coordinate system and this shall be used throughout this document. The origin is chosen as the center of the detector and the positive z-axis is chosen as the direction of the incident protons in the beam line. The y-axis is chosen as vertically pointing upward. Then using cylindrical coordinates one can define the radial distance (r) and the azimuthal angle (ϕ) as:

$$(r, \phi) = \left(\sqrt{x^2 + y^2}, tan^{-1}\frac{y}{x}\right) \tag{2.1}$$

In hadron colliders it is common to use rapidity y or pseudo-rapidity η instead of the polar angle θ (defined as the angle with respect to the z-axis):

$$y = \frac{1}{2}ln\left(\frac{E + p_z}{E - p_z}\right) \tag{2.2}$$

$$\eta = -ln\left(tan\frac{\theta}{2}\right) \tag{2.3}$$

In the ultra-relativistic limit ($i.e., m/E \rightarrow 0$), η approaches true rapidity y. Thus the pseudorapidity is a zero-mass approximation of the real rapidity of highly relativistic particles. Usually small $|\eta|$ values are referred as "central" and large $|\eta|$ values are regarded as "forward". Also, it is common to use dR as a measure of the angular separation:

$$dR = \sqrt{d\eta^2 + d\phi^2} \tag{2.4}$$

2.2.2 The Silicon Microstrip Tracker, SMT

The SMT [32] is the detector which is closest to the beam pipe. The high density of silicon strips in the SMT can provide axial hit resolution of $10\mu m$ and the z hit resolution within $35\mu m$, which allows high resolution track finding and vertex identification. The active SMT detector (strip) consists of p-n junction diode made from doped silicon, and are operated at a reverse bias. The p-n junctions form a depleted zone where there are no free electrons in the conduction band. Therefore when a charge particle passed through the silicon it induces a charge separation within the material and creates a electron-hole pair. The charged pair is then separated by the biasing electric field and is collected in the capacitor for read-out.

Figure 2.3 shows the layout the SMT detector. It consists of silicon wafers in the form of barrel and disk modules. There are six barrel modules, each having four cylindrical layers of 12 cm long. Each of the four layers is made up of two sublayers of overlapping silicon sensors to cover the whole ϕ range.

Inter-spaced between the barrels are the 12 F-discs. These disks have an outer radius of 10cm. The silicon wafers on the discs are double-sided. The F-discs consists of 12 wedges overlapping to cover the transverse $(x - y)$ plane as well. In addition, there are 4 larger H-discs which provide high $|\eta|$ region coverage and they are located at $|z| = 100.4$ and 121.0 cm. The H-discs are made of 24 overlapping wedges and have an inner radius of 9.5 cm and outer radius of 26 cm.

2.2.3 The Central Fiber Tracker, CFT

The CFT [33] measures the position of the charged particles using scintillating fibers. An ionizing particle excites the material in the fiber to scintillate,

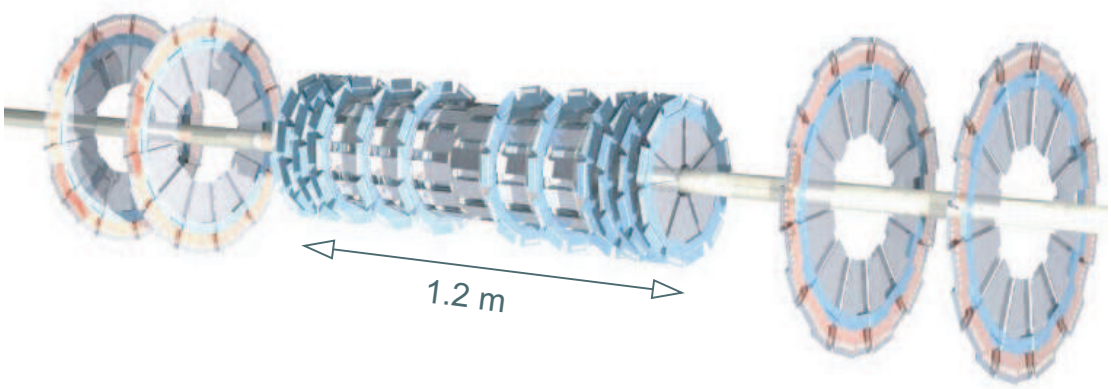

Figure 2.3. 3D layout of the Silicon Microstrip Tracker (SMT)

which emits 530 nm photons that are detected by the visible light photon counters (VLPC).

As can be seen from figure 2.4, there are 8 superlayers of the CFT, each consists of an axial layer (along the z axis) and one stereo layer (± 3 degrees tilted). The η coverage range of the CFT is $|\eta| < 2.0$, and the CFT has an inner radius of 19.5cm and extends to 51.5cm. The two innermost layers are made out of fibers of 166cm long, and the rest are made out of 252 cm long fibers. Each fiber has a diameter of 835 μm and provide a resolution of the charged particle within 100 μm.

2.2.4 The Solenoid

The momentum of charged particles can be determined in the presence of a magnetic field, which generates the Lorentz force $\vec{F} = q\vec{v} \times \vec{B}$ and causes the helical trajectory of the charged particles. The superconducting solenoid [34] is built using super-conducting CuNbTi multi-filamentary cable strands that are

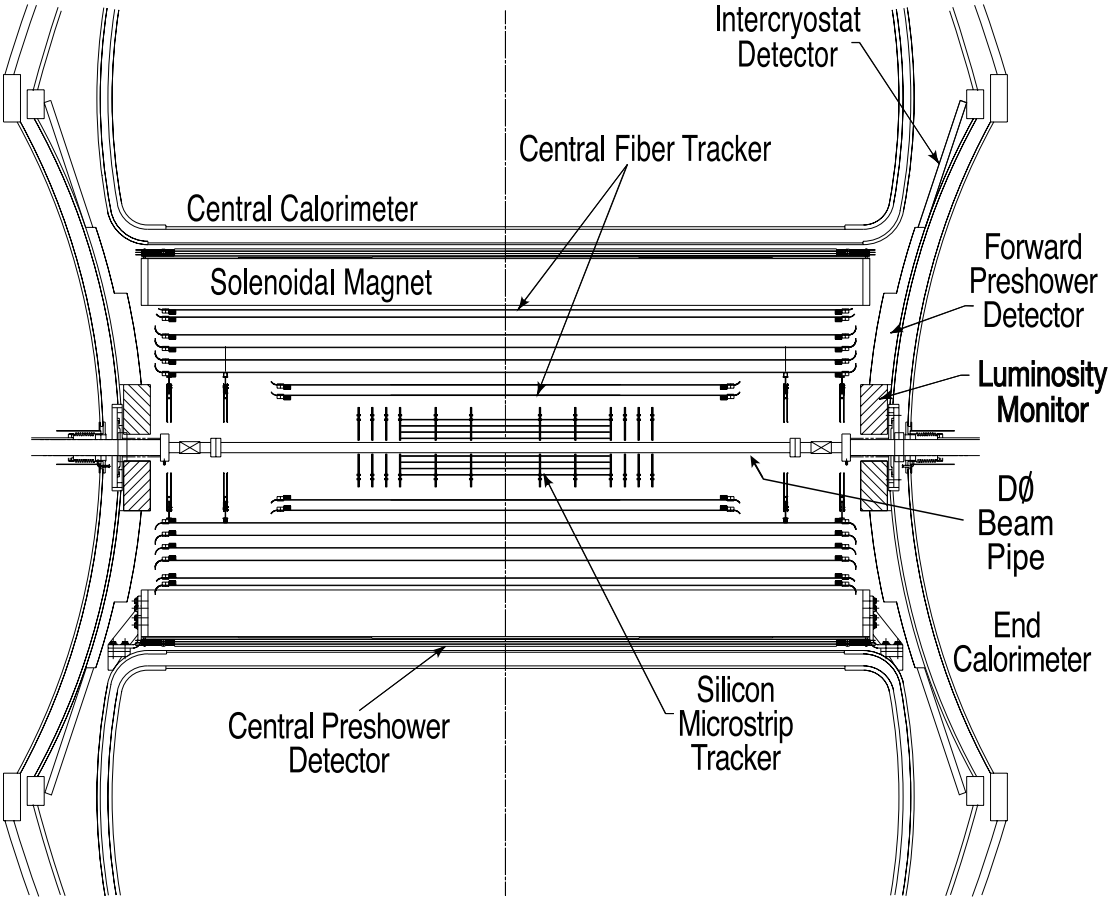

Figure 2.4. The central tracking region of the DØ detector

stabilized by aluminum, and cooled by liquid Helium at a temperature of 4K. The solenoid is 2.73 m in length and 1.42 in diameter. It generates a nearly uniform magnetic field of 2T within the tracker volume using a current of 4749A.

2.2.5 Preshower Detector

The scintillating fiber preshower detectors are located between the solenoid and the central calorimeter. There is a central preshower detector (CPS) [35] which covers $|\eta| < 1.3$ and two forward preshower detectors (FPS) [36] which cover $1.5 < |\eta| < 2.5$. The CPS has 3 layers of triangular scintillating fiber strips interlocked together, and the FPS has 4 layers. This arrangement is used to cover most of the calorimeter space and therefore improve the identification and the energy measurement of photons and electrons.

2.2.6 The Calorimeters

The DØ Calorimeter [37] is designed to capture the electrons, photons and hadronic jets from $p\bar{p}$ collisions. It employs high density material of sufficient thickness so that particles will interact and deposit all of their energy inside the detector through a cascade of lower-energy particles. The electrons and photons will go through bremsstrahlung and pair production to create a shower of subsequently lower energy electrons and photons. Hadrons interact with the nuclei of the absorbing material. This can generate more intricate showers containing not only electrons, photons and hadrons, but also occasionally muons and neutrinos as well. Nevertheless, the principle behind the hadronic calorimeter is similar to the EM (electromagnetic) calorimeter.

Figure 2.5 shows the details of the calorimeter. The Central Calorimeter

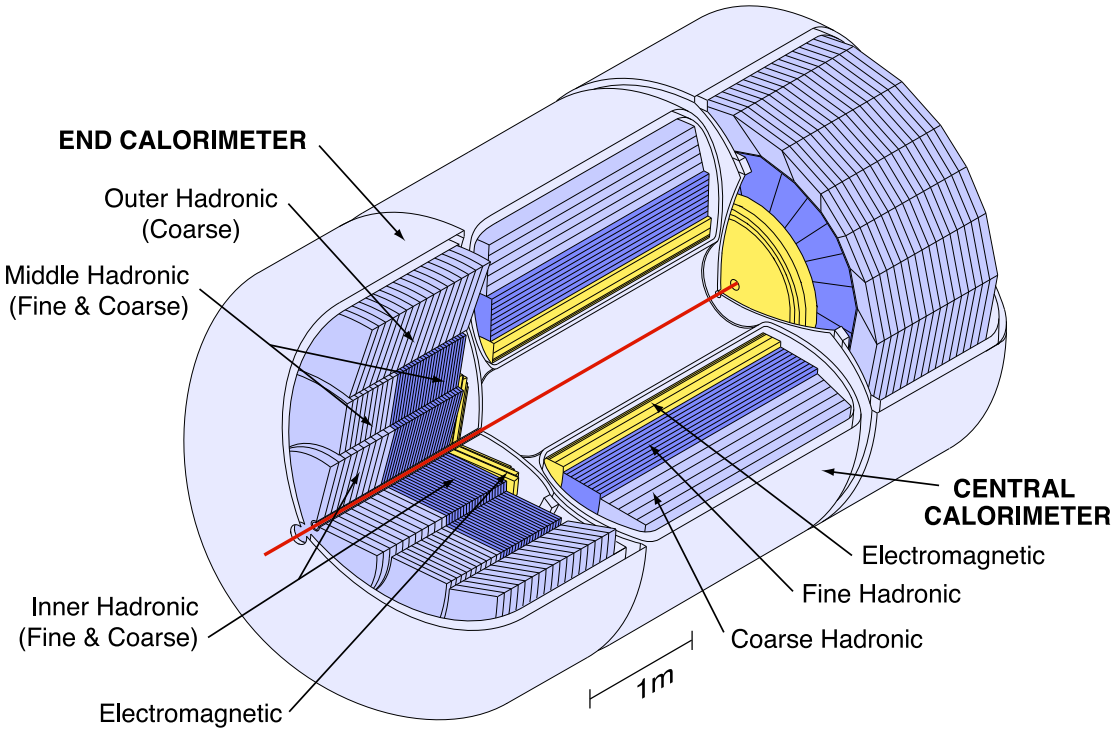

Figure 2.5. A cut-off view of the central and end cap calorimeters

(CC) covers $|\eta| < 1.1$ the two End Calorimeter (EC) covers $1.3 < |\eta| < 4.0$. Each calorimeter is broken up into two sections: EM and hadronic. The hadronic section is further separated into a fine and coarse based on the size of the cells.

A typical calorimeter cell unit is shown in Figure 2.6. An absorber plate containing heavy material of large M (mass number) is used to initiate particle showers. In the EM calorimeter 3-4mm of depleted uranium is used as the absorber. In the fine hadronic calorimeter the absorber plates are 6 mm of uranium/niobium alloy, and in the coarse hadronic calorimeter 46.5 mm of copper or stainless steel is used. A signal board covered with resistive coating is charged to a potential of 2kV to detect the ionizing particles produced by the shower. The 2-3 mm gaps between the absorber and signal board are filled with liquid argon.

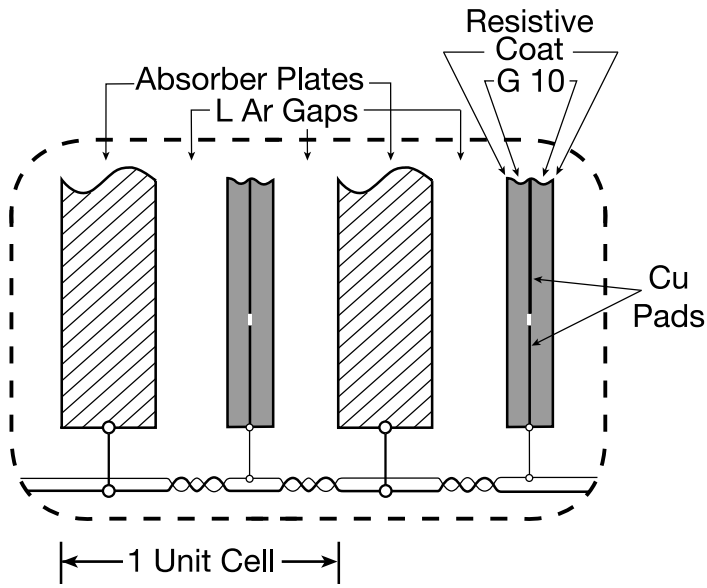

Figure 2.6. The schematics of a calorimeter cell

The calorimeter cells in either CC or EC are arranged in pseudo-projective towers, as shown in figure 2.7. The tower size is designed generally having $\Delta\phi = 0.1, \Delta\eta = 0.1$. The calorimeter resolution is $\approx 5 - 7\%$ for electrons having energy above 20 GeV. For hadrons the resolution is worse: approximately 30% for jets.

2.2.7 The Muon System

Most of the particles are captured by the calorimeter except neutrinos (which cannot be detected due to the tiny interaction cross section) and muons. The outer-most layer of the detector is therefore designed to identify muons. The muon system [38] contains 3 different layers (A, B and C) of scintillators and drift tubes, where the scintillators and the drift tubes provide position and timing measurements. A toroidal magnetic field of 1.8T is embedded between the first 2 layers of the muon system, thus allowing the momentum of the muon to be

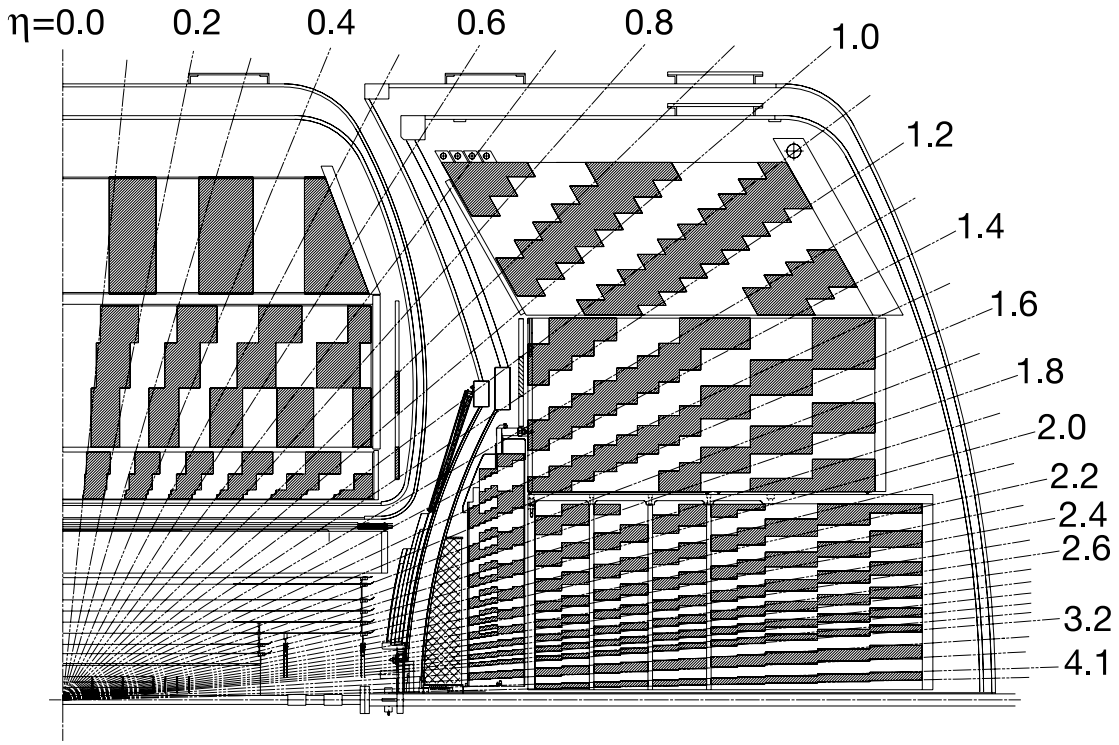

Figure 2.7. Schematics of the arrangement of calorimeter cells as towers of different pseudo-rapidity

determined.

The proportional drift tubes (PDT) and mini drift tubes identify the muons through the charge ionized when the muons interact with the gas inside the tube. At the center of the tube there is an anode wire held at 4.7kV and the wall of the tubes are cathode pads held at 2.3kV. Therefore the time it takes for the ionized charge to "drift" to the wire can be measured, which has a maximum of $500\mu m$

The scintillators are used for precision measurement and timing of the arrival of the muons as well, which can be used to reject cosmic muons. Each scintillator is one half inch think and made of wave-shifting fiber. Figure 2.8 shows the layout of the drift tubes in the muon system. The layout of the scintillators is almost identical. The central muon system has covers $|\eta| < 1$ and the forward system covers $1 < |\eta| < 2$. There is no coverage at the bottom of the A layer in the central system $225 < |\phi| < 310$ as it is occupied by the calorimeter support structure.

2.2.8 The DØ Trigger system

There are an average of 2.5 million $p\bar{p}$ interactions occurring every second at DØ . Given that a typical event can yield about 250 kilobytes of data, it is impossible to record 600 Gigabytes of data per second given the economic scale. In addition, the dominant processes in $p\bar{p}$ collisions are inelastic scattering processes and only a fraction of the events are interesting enough for further investigation. The trigger system is designed to pick out the interesting events.

There are 3 levels of triggers in the DØ trigger system, with each higher level given more time allowing more sophisticated mechanism for a trigger decision. The Level 1 (L1) [39] and Level 2 (L2) [40] architecture is shown in Figure 2.9.

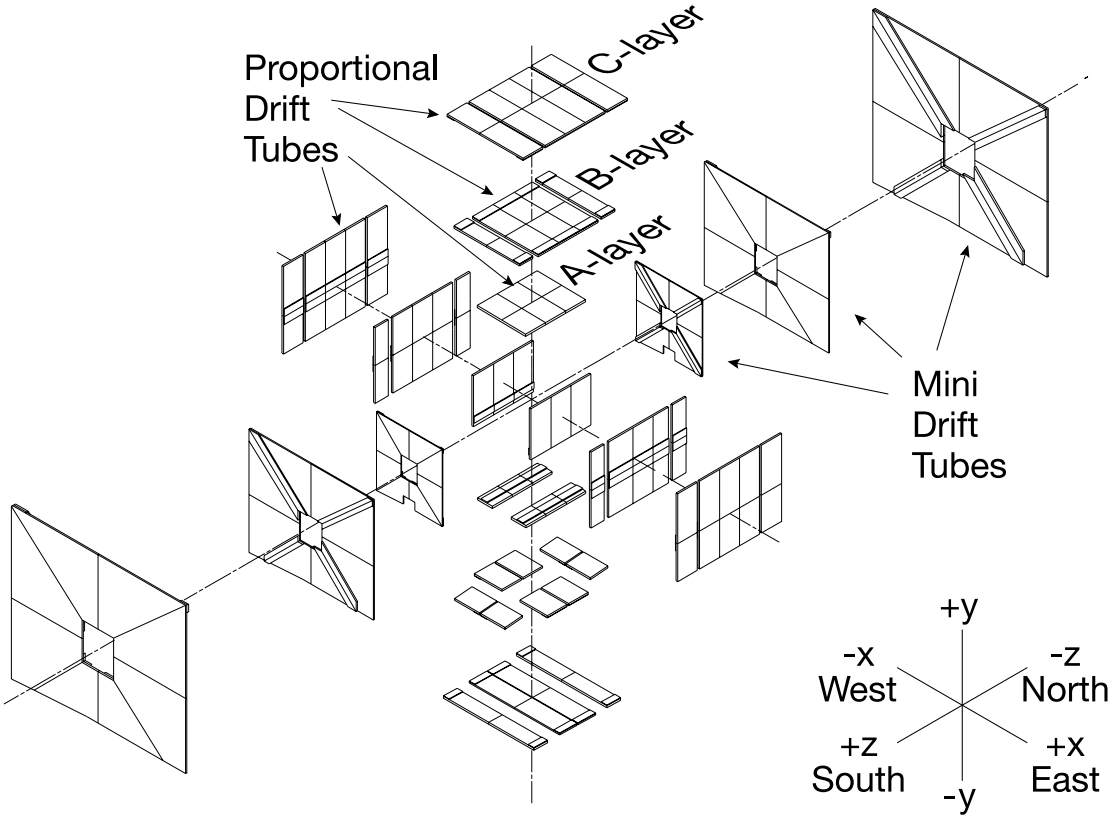

Figure 2.8. Expanded view of the muon drift tubes in the A, B and C layers of the muon system

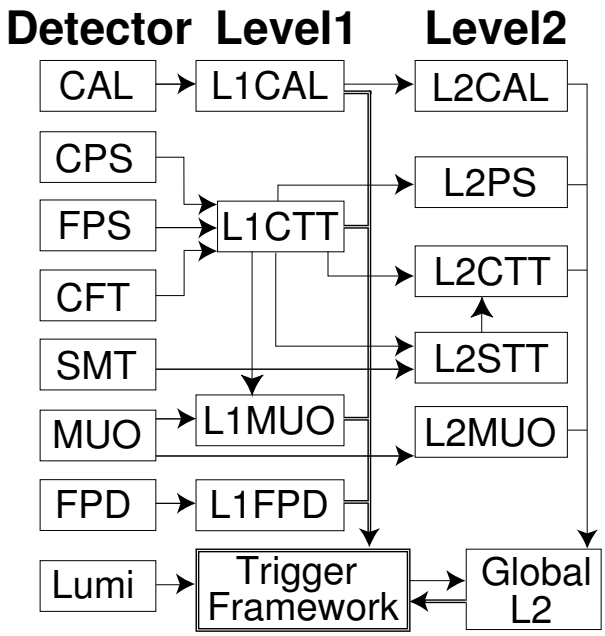

Figure 2.9. DØ Level 1 and 2 trigger architecture. The arrows indicates the flow direction of the information

The L1 system is completely implemented in hardware for high speed decisions. Each L1 trigger processes the input from the corresponding sub-detector, and then collects the result and passes it to the L1 framework (L1FW). The L1FW makes the global decision to either accept or reject the event. L1 can reduce the event rate by a factor of more than 1000 – from 2.5MHz to around 1 kHz.

If an event is accepted by L1, then it would be digitized and stored temporarily in one of the 16 event buffers in L2. The event is then reconstructed with better precision and the L2 trigger can make a decision by combining results from different subdetectors. The L2 trigger is also hardware based and is designed to reduce the event rate further to around 1 kHz.

The Level 3 (L3) trigger [41] is a software-based trigger running on a farm of Linux PCs. Each event is partially reconstructed in even better precision within

a single node to determine whether it passes the acceptance. The L3 system is also capable of making decisions using more complicated variables such as event missing transverse energy and b-tagging probability. L3 has a large rejection factor of around 20, which further reduces the event rate to around 50 Hz – which is the rate that the data is written on tape.

CHAPTER 3

OBJECT IDENTIFICATION

As the generated particles travel through the different parts of the detector, they leave different signatures behind. The signals are then digitized and readout from the sophisticated electronics and reconstructed by the D0reco software. Different algorithms are implemented in D0reco to identified the various objects such as tracks, jets, and the primary vertex which are approximation of their physical counter-parts, and will be explained in detail below.

3.1 Tracks

As charged particles propagate through the tracker, they interact with the detector materials and deposit in hits in different layers. Those hits can then be combined to form tracks which resemble the original particle trajectories. Two algorithms, HTF [42] and AA [43], are implemented for track reconstruction.

The HTF (Histogramming Track Finder) makes use of the uniform magnetic field within the tracker. First consider only the x-y plane. For any hits in the tracker, one can assign arbitrary values of curvature ρ to the hit. Fixing the value of ρ also fixes the value of angle ϕ_0 where the track originated from the beamspot. Therefore each hit in the x-y plane can be transformed to a line in the (ρ, ϕ_0) plane. Given that a track has constant curvature when propagated under

a uniform magnetic field. The hits that belong to the same track will therefore overlap at the same point in the (ρ, ϕ_0) plane and thus the corresponding track can be identified. Similarly, a second transformation which transform a hit in (r, z) to a line in $(z_0, dz/dr)$ can help to fix the z-coordinate at its closest approach to the beamspot as well.

The AA (Alternate Algorithm) starts with 3 clusters of hits in the SMT layers, where the first hit can come from any barrel or an F disk. The second hit has to be within an azimuthal window of 0.08 radians from the first hit with respect to the beam spot. The 3 hits combined have to yield a track with track $p_T \geq 180 MeV$ (which correspond to the curvature $\rho \leq 1/30\ cm$). Also the impact parameter of the track should be less than 2.5 cm with respect to the beam spot, and the χ^2 of the track hypothesis has to be greater than 16. The resulting track hypothesis will then be extended to the next outer layer to pick out the hits from the expected locations. The searches for the successive hits will continue until the track $\chi^2 > 16$, or when there are 3 conescutive misses in a row, or when the limit of the tracker is reached.

Given the list of track candidates from both algorithms, Kalman filters [44] are then used to filter and further improve the tracks [45]. The filter also tries to add hits to the track candidate to optimize the track parameters. Similar tracks can also be merged together if they are essentially pointing to the same track.

3.2 Primary Vertex

The primary vertex (PV) is the interaction point of the $p\bar{p}$ collision and therefore is the originating point of the tracks. There can be multiple primary vertices in a single event as there can be multiple $p\bar{p}$ collisions. There can be secondary

vertices (SVX) as well which result from the decay of long-lived particles such as B, K or D hadrons. Secondary vertices will be described in more detail in section 3.6.3.

A vertexing algorithm [46] is developed to separate the primary vertex from other vertices such those produced in minimum-bias interactions or decays of long-lived particles. Beginning with the list of tracks having $p_T > 0.5$ GeV and at least 2 SMT hits, the tracks are grouped to form a z-cluster if the separation in z between the track and the cluster is within 2 cm.

The tracks in the z-cluster are then fitted to a common vertex using the Kalman filter Algorithm [47], and the track contributes the largest χ^2 to the fit is then removed. The Kalman filter fit is repeated until the total vertex χ^2/dof of the fit is smaller than 10.

Then the remaining tracks in each z-cluster are sorted in order of the their distance of closest approach (DCA) from the fitted vertex position above. Only those tracks having DCA significance (defined as $|DCA|/\sigma_{DCA}$) less than 5.0 are selected. An iteration procedure using the Kalman filter algorithm is then applied to the selected tracks, where each track is assigned a weight given its χ^2 contribution to the fitted vertex:

$$w_i = \frac{1}{1 + e^{(\chi_i^2 - \chi_c^2)/2T}} \tag{3.1}$$

The weight is taken in the form of the Fermi function, where χ_i^2 is the Chi-squared contribution of the track to the fitted vertex, and χ_c^2 is the cutoff value. T is a parameter which determines the sharpness of the function. The Kalman filter is then recomputed again using these track weights to refit the vertices. Those tracks with a weight smaller than 10^{-6} are removed from the fit, and the weights

will be recomputed and Kalman filter applied again until the weights become stable.

Finally given the list of fitted vertices with associated tracks, the best primary vertex is chosen as the one most likely from the hard scattering $p\bar{p}$ collision. The decision is based on the track $log_{10}p_T$ because minimum-bias events tend to have smaller momentum transfer and therefore have tracks with lower p_T. The probabilities from each track $log_{10}p_T$ are then combined for each vertex candidate to form a minimum bias probability. The vertex having the lowest minimum bias probability is then chosen as the primary vertex.

3.3 Muons

Muons [52] can be identified using the tracks from the central tracker together with the muon detector. First the algorithm combines the hits from the muon scintillators and wire chambers to form local muons. Here, the momentum of the muon can be inferred as well due to the presence of the toroidal magnetic field. The Local muon tracks will then be matched with central tracks to form global muons. The global muon will have a longer trajectory and therefore provides a more accurate measurement of the track momentum; for this reason most analyses use only global muons. Also the selected muons are required to have a scintillator hit time within 10ns of the bunch crossing time, in order to suppress the background cosmic muons.

3.4 Jets

As strongly interacting partons (such as quarks and gluons) emerge from hard scattering interactions, they will evolve into streams of hadrons and travel ap-

proximately in the same direction as the original partons. The hadrons interact with the calorimeter to generate clusters of energy, which can then be identified as jets. The charged hadrons can also leave bundles of tracks in the tracker as well and can be identified as track-jets.

Many different algorithms are developed to reconstruct jets from the calorimeter [48]. In DØ Run-II the preferred algorithm is the improved legacy cone algorithm, which will be discussed in details below.

3.4.1 The Cone Cluster Algorithm

Given the presence of various types of noise in the calorimeter, a noise removing algorithm is first applied to the calorimeter cells. The algorithm removes isolated calorimeter cells which have no signal-like neighbor (which implies the cell hit is likely due to noise) and those cells which have negative energy. Certainly the cells within the jets are less likely to be rejected and they are likely not isolated and have neighboring cell hits.

The calorimeter cells having the same "direction", or same detector coordinates (ϕ_{det}, η_{det}) are combined to form calorimeter towers. The tower's transverse energy $E_T = Esin\theta$ is then calculated using the sum of energy E of the cells in the tower together with the polar angle θ between the tower and the beam using the primary vertex as the origin.

The cone cluster algorithm starts with a list of "seeds", which consists of all the calorimeter towers having $E_T > 0.5$ GeV and the sum of transverse energy $\Sigma E_T > 1.0$ GeV within a cone of radius $R = 0.3$. Then the energy within a cone radius of R=0.5 around each seed is calculated, together with a new cone center using the E_T weighted tower:

$$(\phi_{center}, \eta_{center})_{new} = \frac{1}{\sum_{dR_{cell}<0.5} E_T} \left(\sum_{dR_{cell}<0.5} E_T \phi_{cell}, \sum_{dR_{cell}<0.5} E_T \eta_{cell} \right) \quad (3.2)$$

The above calculations are then iterated until the cone is stable (when the difference in the cone center between iterations becomes negligible). Each stable cone found is called proto jet.

3.4.2 Infrared safety requirement

Occasionally, the presence of low energy radiation can affect the cone algorithm and results in unstable results. An example is illustrated in figure 3.1 where a lower energy radiation exists between two higher energy jets. Such lower energy (infrared) radiation causes the unwanted merging of the two jets.

To avoid unwanted merging of the two jets, additional seeds called midpoints are added between the found proto-jets. These midpoints are E_T weighted centers between pairs of proto-jets, lie within a distance of $\Delta R < 2.0R_{cone}$, and are intended to pick up the low energy towers between jets.

3.4.3 Jet Merging and Splitting

With the list of proto-jets found, duplicate proto-jets having the same axis are removed. Algorithms are also developed to merge and split jets depending on the fraction f of the transverse energy in the sharing towers between 2 neighboring jets. If $f > 0.5$ then the two jets containing the sharing towers are merged to form a single jet and the jet center/energy will be recalculated. If $f < 0.5$ then the overlapped towers will be split between 2 neighboring jets based on their proximity to the jet center. The merging/splitting algorithms will be repeated until all the

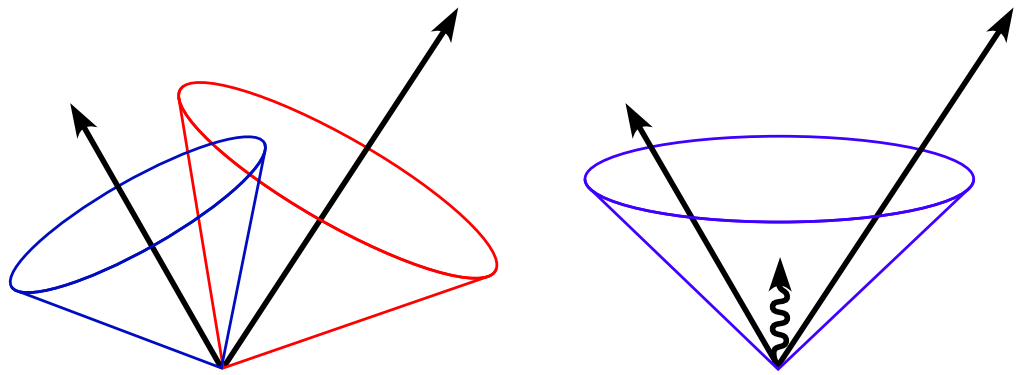

Figure 3.1. Illustration of the effect of infrared radiation in jet algorithm. The picture on the left shows the original two jets. Introduction of radiation merges the 2 jets together.

proto-jets have stable axes. The surviving proto-jets form the final list of Jets and all the jets with $E_T > 6$ GeV are kept.

3.4.4 Jet quality requirement

In order to reduce the amount of false/fake jets due to calorimeter noise, the jets used for physics analysis also need to satisfy a number of quality cuts [49]. The cuts also serve to separate the electrons and photons from hadronic jets.

EMF (EM fraction): Based on the design of the calorimeter, hadronic jets deposit a large fraction of their energy in the hadronic layers. By contrast, electrons and photons deposit a large amount of energy in the EM layers. Thus the EM fraction is simply:

$$EMF = \frac{p_T^{EM}}{p_T^{total}} \tag{3.3}$$

CHF (Coarse hadronic fraction): Due to the relatively large size of the readout cells in the coarse hadronic layer, this part of the calorimeter is most sensitive to noise. It is unlikely that a jet will deposit most of its energy in the coarse hadronic section. By cutting on the fraction of CHF one can remove the jets more likely due to noise:

$$CHF = \frac{p_T^{\text{coarse hadronic}}}{p_T^{total}} \tag{3.4}$$

N90: As the jet is constructed from calorimeter towers, N90 is the number of towers containing 90% of the total energy within the jet. Electons/photons and noisy towers tend to deposit most of the energy within a single tower, which allows them to be separated from jets.

Level-1 Confirmation: Coherent noise from the precision readout system can create fake jets. The fake jets can be reduced by comparing with the jet p_T with the energy found by the L1 trigger system, as the L1 electronic readout chain is different.

$$L1_{ratio} = \frac{p_T^{\text{Level 1 readout}}}{p_T^{\text{precision readout}}} \tag{3.5}$$

The jet quality cuts using above variables are optimized and need to satisfy:

- CHF 0.4, or
- CHF 0.44, and $|\eta_{det}| < 0.8$ (central region), or
- CHF 0.6, and $0.85 < |\eta_{det}| < 1.25$ (ICR region), or
- CHF 0.46, and $1.5 < |\eta_{det}| < 2.5$ (endcap excluding forward regions).

Also the jet needs to satisfy EMF < 0.95 and:

- EMF > 0.05, or
- EMF > 0.03, and $1.1 < |\eta_{det}| < 1.4$ (Gap between central and endcap), or
- EMF > 0.04, and $|\eta_{det}| > 2.5$ (endcap).
- $L1_{ratio} > 0.5$, or
- $L1_{ratio} > 0.35$ and $p_T < 15$ GeV and $|\eta_{det}| > 1.4$ (Endcap), or
- $L1_{ratio} > 0.1$ and $p_T < 15$ GeV and $|\eta_{det}| > 3.0$ (forward), or
- $L1_{ratio} > 0.2$ and $p_T \geq 15$ GeV and $|\eta_{det}| > 3.0$ (forward)/

Only Jets passed these cuts are chosen for further analysis.

3.4.5 Jet Energy Scale

Even after quality cuts, the jet constructed from the calorimeter only contains energy seen by the calorimeter only, and it certainly doesn't match the original parton energy. There are numerous factors in the detector which contribute to the differences: There is additional absorbing material in front of the calorimeter, and also various gaps exist within the calorimeter. The cone algorithm has a fixed size cone and may miss the energy contribution outside the cone. The goal of the Jet Energy Scale is to correct the measured jet energy from the calorimeter such that, the corrected jet energy, on average, is equal to the parton jet energy before any interaction with the detector.

The corrected Jet energy is given by:

$$E_{jet}^{corr} = \frac{E_{jet}^{meas} - E_O}{R_{jet} S_{jet}} \tag{3.6}$$

There are additional bias correction factors necessary for more accurate correction, and they are described in more details in the DØ JES note [50].

E_{jet}^{meas} is the measured jet energy using the jet finding algorithm above.

E_O is the offset energy which is not associated with the hard scattering process of the jet. It includes the contributions from both the electronic noise and the uranium decay of the detector. Other contributions are the additional $p\bar{p}$ interactions in the same bunch crossing (pile-up). Such energy can be shown as a function of both the jet detector pseudorapidity (η_{jet}^{det}) and the number of primary vertex n_{PV} in the event. In Figure 3.2 the dependence of the offset energy on η_{jet}^{det} and n_{PV} is shown.

R_{jet} is the energy response of the calorimeter. There is energy lost in the material of the detector and also the calorimeter response is non-uniform especially in the ICR region. The energy response is first measured in the central region of the calorimeter (CC) using the MPF (Missing E_T Projection Fraction) method. Using back-to-back events with one photon and one jet, one can determine the response from the missing E_T of the event and the energy of the photon (its response is measured in high precision using $Z \rightarrow ee$ events). Figure 3.3 shows the illustration of the MPF method and figure 3.4 shows the MPF response in data. In addition, the non-uniformity of the response in different detector eta regions need to be corrected, especially in the ICR region. This is done by looking at response in different η^{det} relative to the CC, and is shown in figure 3.5.

S_{jet} is the correction due to the development of showers in the calorimeter. It corrects for the energy that is deposited outside of the jet cone from the

particles that belong to the particle jet, and also corrects for the energy deposited inside the jet cone from the particles not belonging to the particle jet. As shown in figure 3.6 the showering correction depends mostly on η^{det} and mildly on jet energy.

The various corrections are then combined to give the corrected jet energy. Figure 3.7 shows the combined JES correction and figure 3.8 shows the relative uncertainties of the JES for jets of 0.5 cone size. As will be shown in later sections, the error due to JES is a major source of uncertainty in this analysis.

3.4.6 Muon Correction in JES

Occasionally a muon may exist within a Jet. Since muon only interacts weakly with the calorimeter material, its response is very different compare to hadrons and therefore muon energy contribution is needed to be corrected separately. A muon is considered to belongs to a jet if it lies within the jet cone. The muon correction will then use the momentum of the muon from the muon system/tracker, and add the energy of the muon to the jet. However, it is shown that this correction behaves well if the energy correction is only applied to muon p_T up to 60 GeV [51]. For muons with p_T above 60 GeV the correction will be calculated assuming the muon p_T is 60 GeV. Such muon correction in JES, or called JESMU, will be applied throughout this anaylsis (except for the triggers in section 4.2).

3.4.7 Jet Energy Resolution

The jet resolution is determined separately for jets $p_T < 50$ GeV (using photon+jets events) and $p_T < 50$ GeV (using dijets events). Using the asymmetry variable:

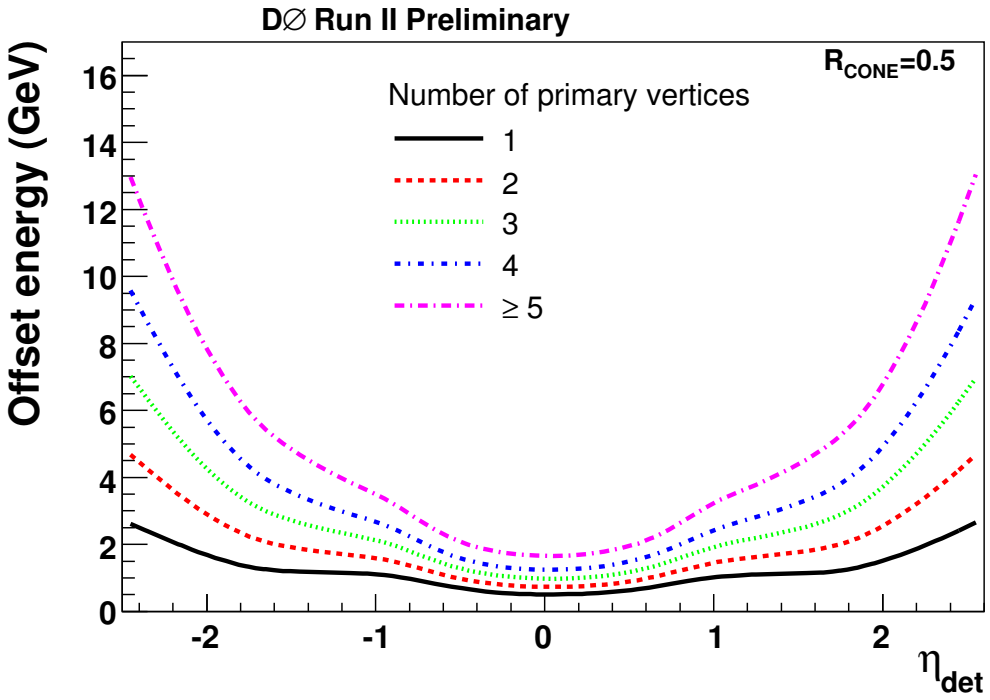

Figure 3.2. Offset energy for Rcone=0.5 jets for different primary vertex multiplicities, as a function of jet pseudorapidity from the center of the detector

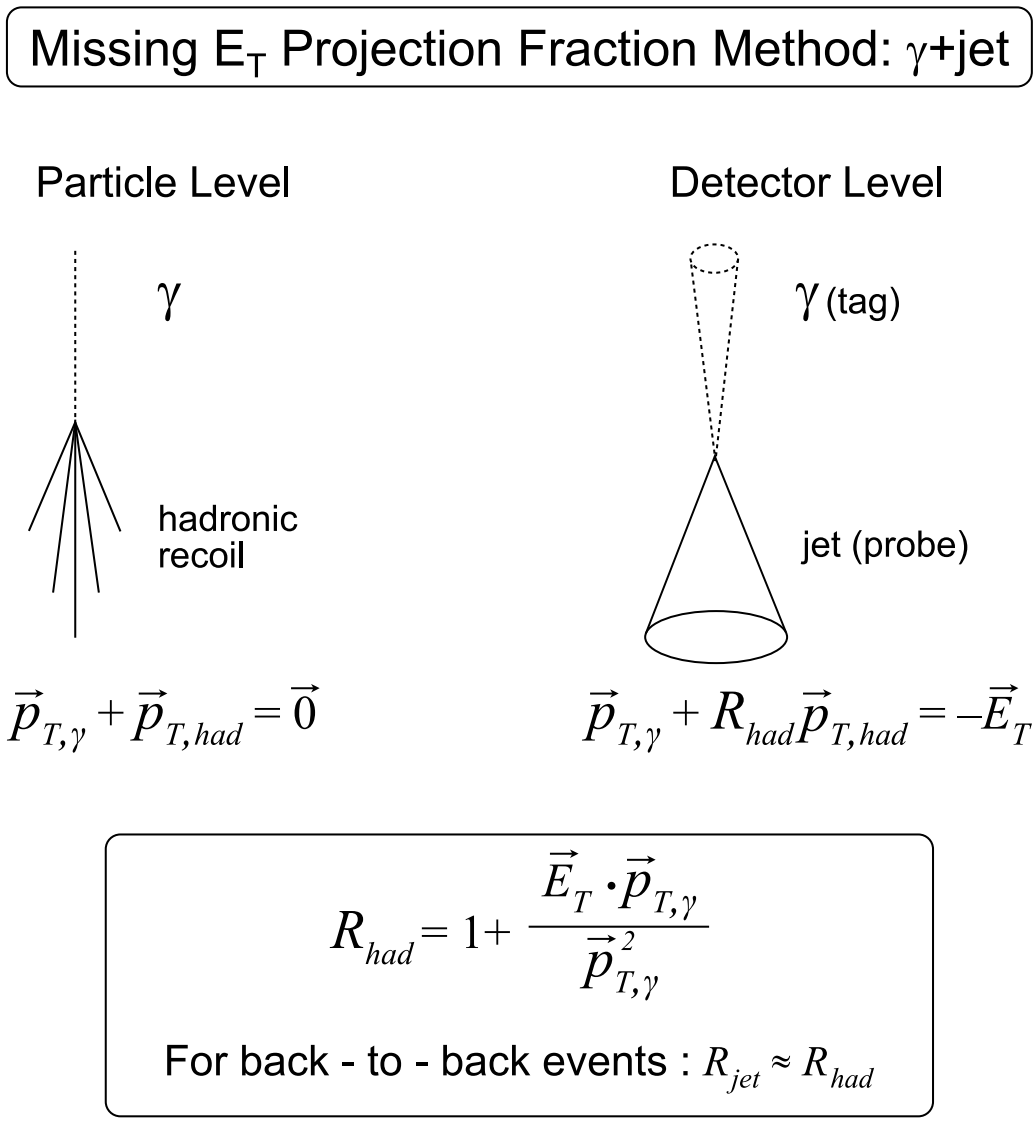

Figure 3.3. Illustration of MPF method. It makes use of the good photon p_T resolution of together with the missing E_T to extract the hadronic jet response.

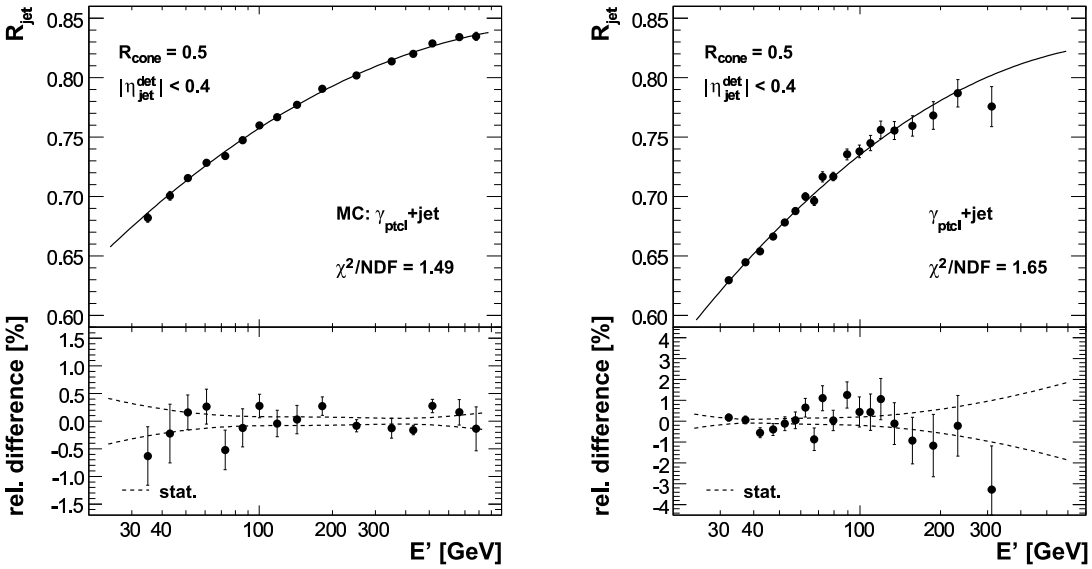

Figure 3.4. JES MPF response for the MC events (on the left) and for the data events (on the right) using jet cones size of 0.5

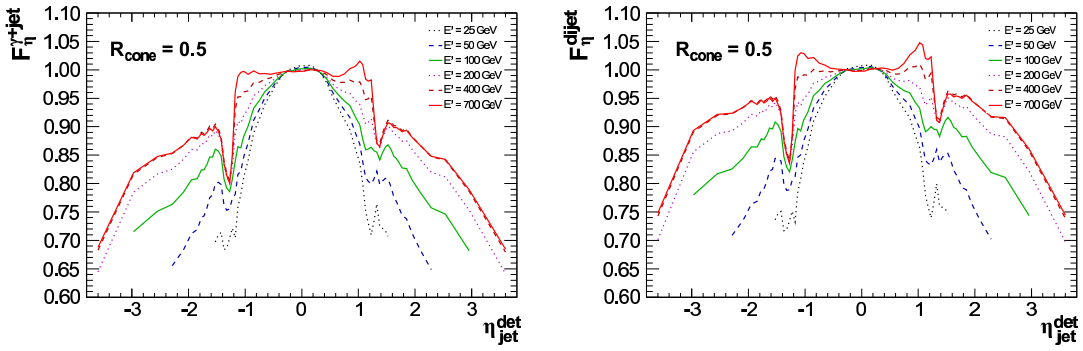

Figure 3.5. JES Eta dependent response for the MC events (on the left) and the data events (on the right) using jet cones size of 0.5

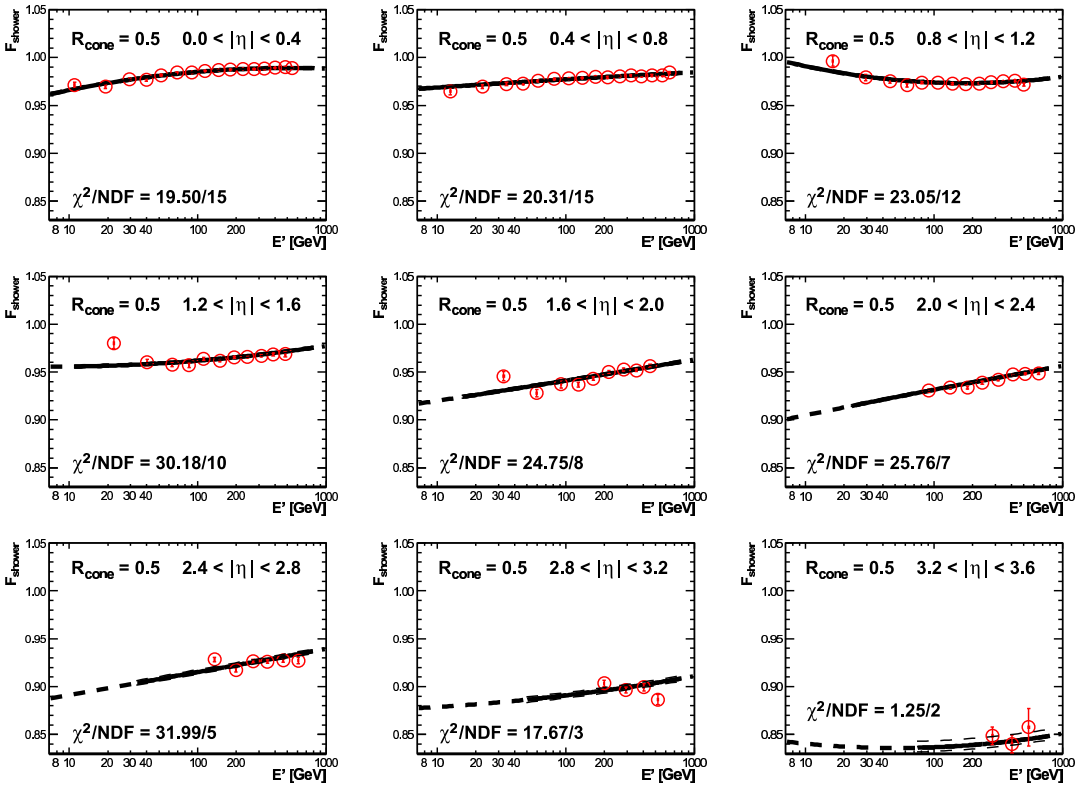

Figure 3.6. JES Showering in different η_{det} regions for jet cones size of 0.5

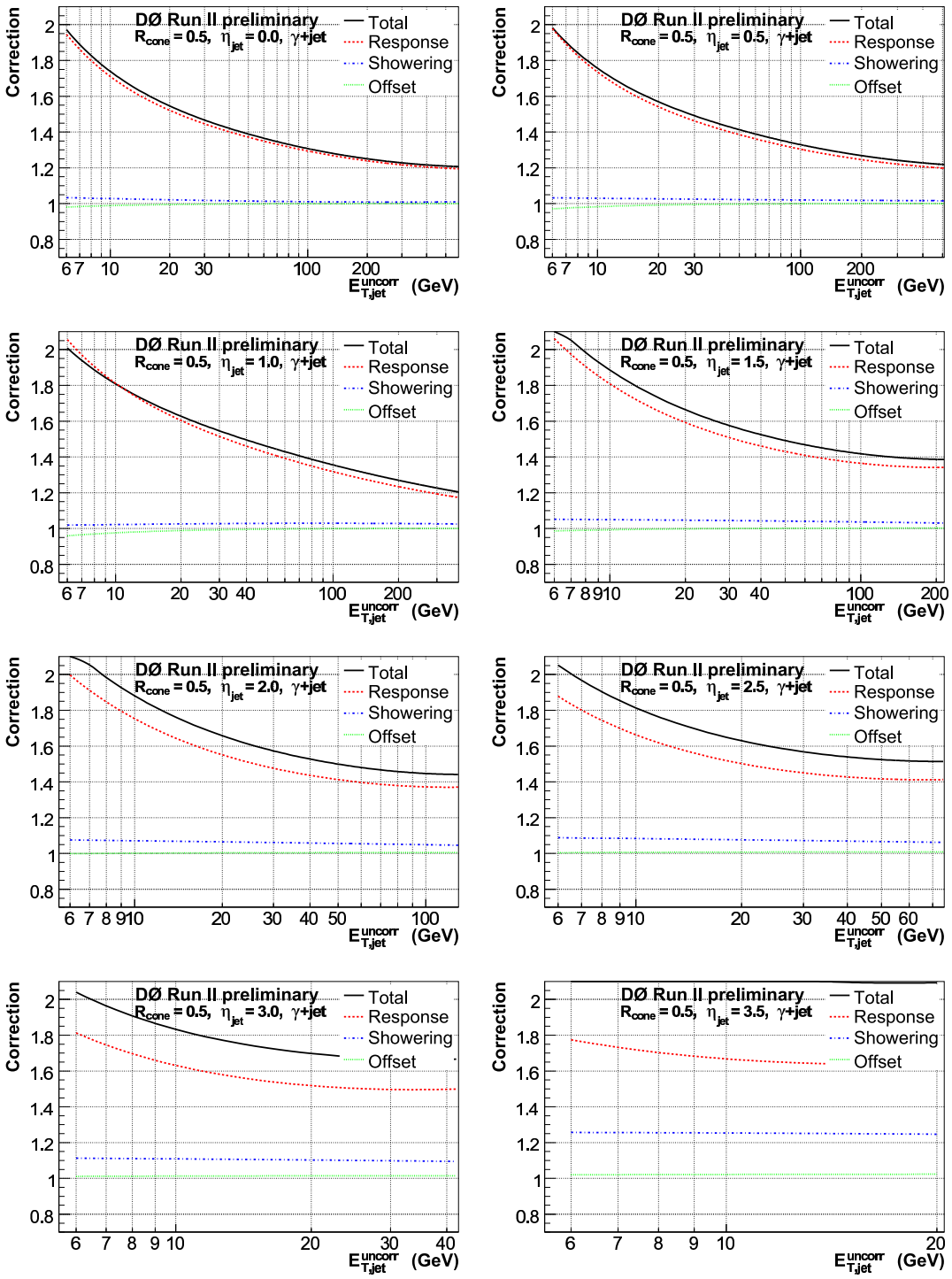

Figure 3.7. JES combined correction factor in different η_{det} regions for jet cones size of 0.5

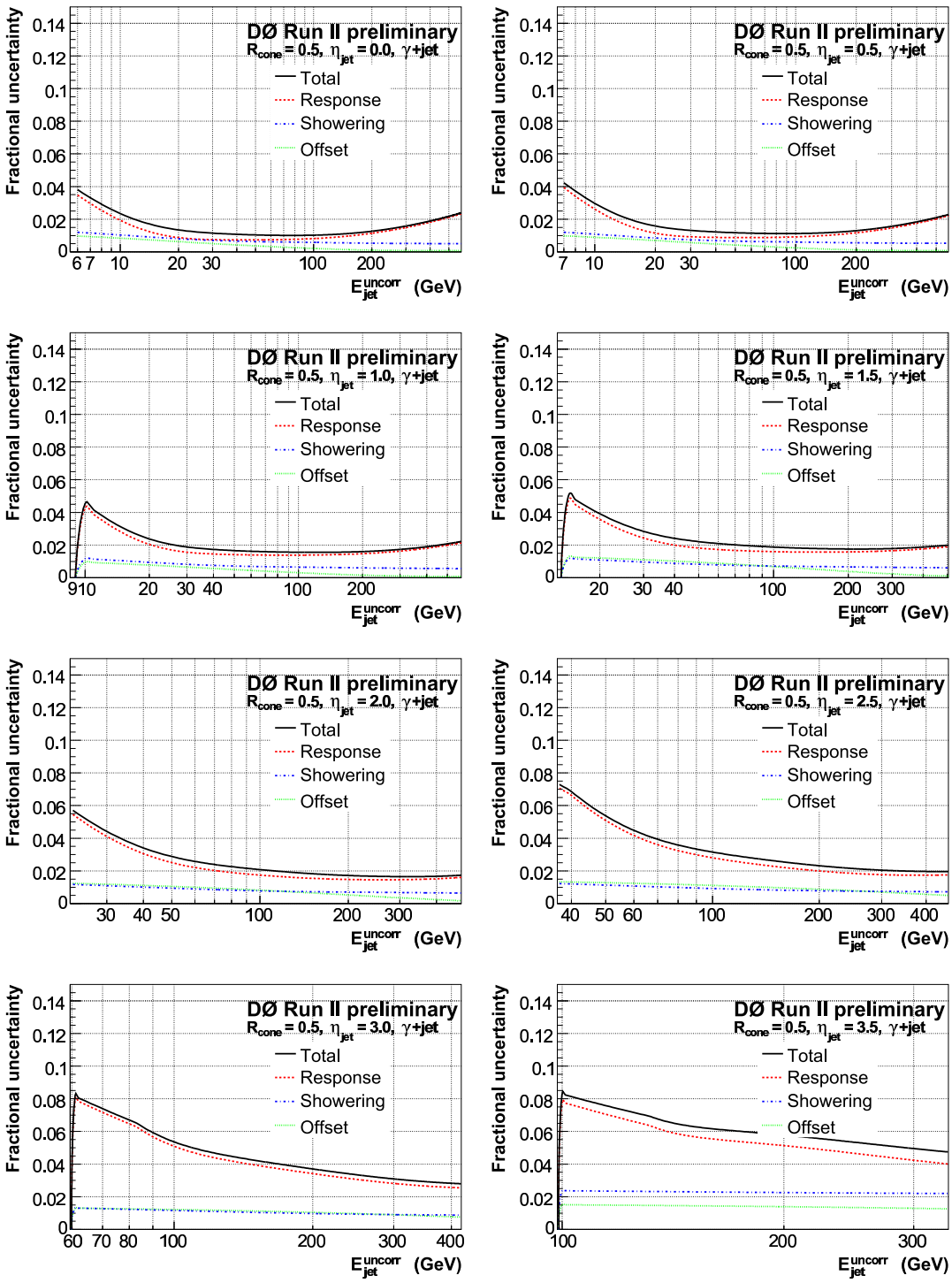

Figure 3.8. JES uncertainties in different η_{det} regions for jet cones size of 0.5

$$A = \frac{|p_T^1 - p_T^2|}{p_T^1 + p_T^2} \tag{3.7}$$

The jet p_T resolution is directly related to resolution of the measured asymmetry resolution:

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{2}\sigma_A \tag{3.8}$$

3.5 Electrons

Electrons can be identified using the same cone algorithm as in the jet reconstruction and the tracks found from the central tracker. The cluster due to an electron is relatively narrow in the electromagnetic calorimeter and a cone size of 0.2 is used for the cone algorithm. The EMF of the jet electron candidate is required to be greater than 90%. Also, the EM cluster is required to be matched with a track within $\Delta\eta < 0.05$ and $\Delta\phi < 0.05$ to suppress the background from jet production [53].

3.6 b-tagging

From the Standard Model and current experimental results of the CKM Matrix, the top quark decays almost 100% to a b-quark and a W-boson, which can further decay to 2 lighter quarks. Due to the heavy mass of the b-quark, the decay process is slightly different from the lighter quarks. A b-quark can also decay weakly to a lighter quark and a W boson, which in 1/3 of the time will further decay in a lepton which results in the presence of a lepton within the jet. Another feature of a b hadron decay is that a b-quark has a relatively long lifetime. The relativistic boost of the b hadron allows it to travel a measurable distance (around

a few mm) within the detector. This results in a secondary vertex displaced from the production point of the b-quark (primary vertex) and therefore provides a hint for the presence of the b-quark. This study will focus mainly on the b-tagging techniques using the long life-time of the b-quark.

3.6.1 Counting Signed Impact Parameter CSIP

Given that the the b-quark may travel a short distance before it decays, the direction of the track from its decay products may be slightly deflected from the jet axis. This results in a measurable distance, or distance of closest approach (DCA) when the track is extrapolated back to the proximity of the primary vertex.

Note that one can assigned a positive or negative sign to the DCA depending on whether the extrapolated track crosses the jet axis in front of or behind the primary vertex. One can define a good observable: signed(IP) by looking at the significance of the DCA: $signed(IP) = signed(DCA)/\sigma(DCA)$, where the denominator is the error of the DCA. In general the tracks from the b-quark jets have larger $|DCA|$ compare to the lighter quark jets. This allows one to tag the b-jet by counting the number of tracks in the jet cone having $|IP|$ above certain values.

3.6.2 Jet Lifetime Impact Parameter JLIP

Rather than counting the number of tracks satisfying different impact parameters cuts, one can instead use the DCA significance to develop a probability that the tracks inside the jets originate from the primary vertex. The probabilities of jets from b or c quarks are expected to have a peak at very low values, while the jets from lighter quarks are essentially flat, as shown in figure 3.9. By cutting on

the probability one can separate the b-jets from the lighter quark jets.

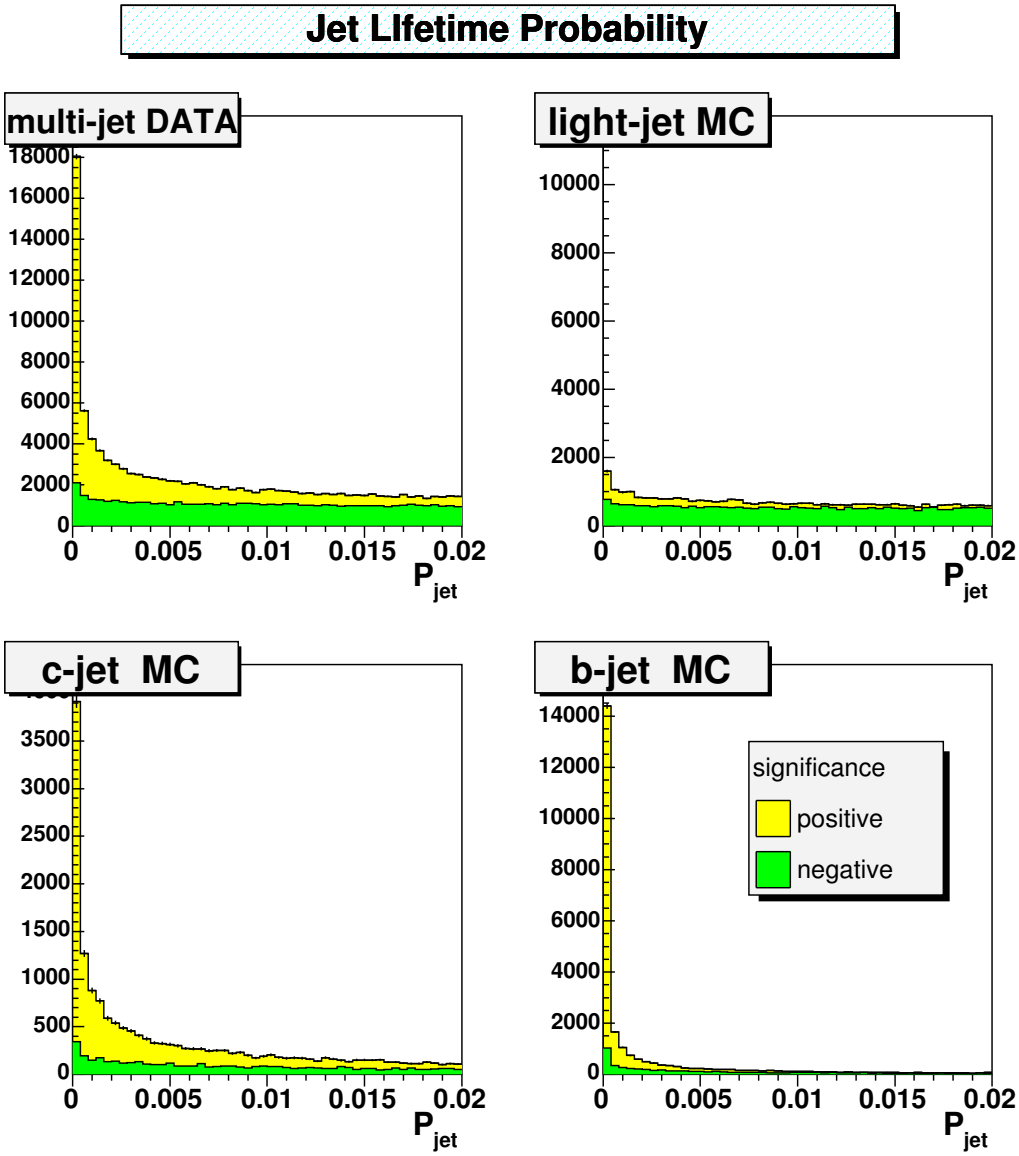

Figure 3.9. JLIP tagger probability for data jets and light quark, c-quark and b-quark jets from Monte Carlo events.

3.6.3 Secondary Vertex Tagging SVT

Since the b-quarks tends to decay at a distance from the primary vertex, the tracks from the decay tends to converge to form a secondary vertex which is located at a distance from the primary vertex.

To identify secondary vertices [54], [55], first the reconstructed tracks are group together as a track jet if the tracks are within a certain cone size. Then within each track jet, tracks with high impact parameter (IP), at least one hit in the SMT, and a transverse momentum of $p_T > 1$ GeV are selected. There are 3 different operating points of SVT, where the tight and medium working points require $IP > 3.5$, and the loose working point requires $IP > 3.0$. Those track jets that have 2 or more selected tracks are then processed by the Kalman Filter (similar to the one used for primary vertex reconstruction) and identified as a secondary vertex if the fit is good.

One can then define the decay length significance of the secondary vertex as $S(L_{xy}) = L_{xy}/\sigma(L_{xy})$, where L_{xy} is the decay length. Those tags with decay length $L_{xy} < 0$ are called negative tags and those with $L_{xy} > 0$ are called positive tags. If a jet has a secondary vertex within $dR(\text{jet}, \text{secondary vertex}) < 0.4$, and the decay length significance satisfies $|S(L_{xy}) > 5|$ for the loose working point ($|S(L_{xy}) > 6|$ for medium and $|S(L_{xy}) > 7|$ for tight), the jet is considered SVT tagged. Figure 3.10 shows an example of a jet tagged with the presence of secondary vertex within its jet cone.

3.6.4 Neural Network Tagging NN

It is possible to combine the results from the CSIP, JLIP and SVT methods into a single output discriminant [56]. This involves the use of an artificial Neural

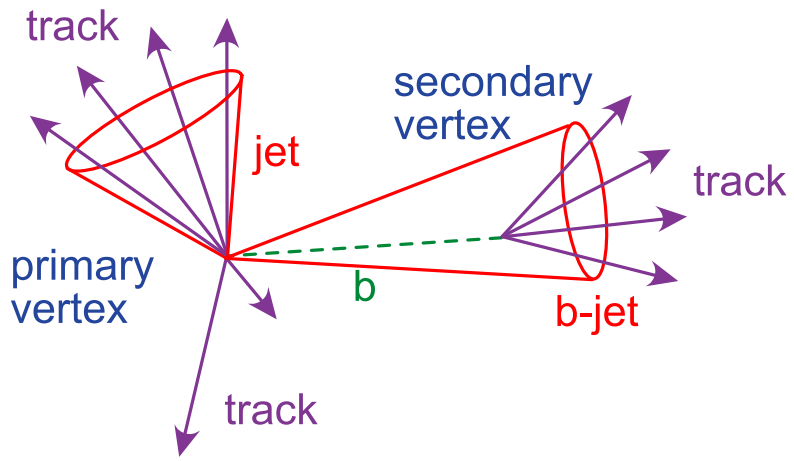

Figure 3.10. Illustration of tagging a jet due to the presence of a secondary vertex within its jet cone.

Network (NN) which combines several different variables, as shown in Table 3.6.4.

The output of the NN is a discriminant value between 0 and 1, where most of b-quark jets have a NN value close to 1. The neural network used consists of an input layer of 7 nodes, 2 hidden layers of 24 and 1 output nodes, and an output layer of 1 node. The different NN values correspond to different operating points. The Neural Network was trained using MC $Z \rightarrow b\bar{b}$ events and $t\bar{t}$ events.

Figure 3.11 shows the efficiency and the tagged rate function (TRF) of tagging b-quark jets with different p_T and different η based on 3 different NN cuts: $NN > 0.5$ (old Loose), $NN > 0.65$ (Medium) and $NN > 0.775$ (Tight). The difference between efficiency and TRF is that the TRF includes the extra scale which accounts for the differences between MC simulation and data. Figure 3.12 shows the efficiency of tagging c-quark jets instead. The tagging efficiency of a b-jet is around 60% to 70% and for a c-jet is around 15% to 20%. Also as jet η increases and approachs the end of coverage of the central tracking region, the

TABLE 3.1

NN input variables ranked in order of separation power

Rank	Variable	Description
1	SVT_{SL} DLS	Decay Length Significance of Secondary Vertex (SV)
2	CSIP Comb	Weighted combination of the tracks' IP Significances
3	JLIP Prob	Probability that the jet originates from the PV
4	$SVT_{SL}\chi^2_{dof}$	Chi square per degree of freedom from SV
5	$SVT_L N_{Tracks}$	Number of tracks associated with SV
6	SVT_{SL} Mass	Mass of the SV
7	SVT_{SL} Num	Number of SV found in the jet

tagging efficiency decreases.

To see how the NN tag performs in the all-hadronic $t\bar{t}$ MC events Table 3.6.4 and Table 3.6.4 shows the tagging fraction of different flavors (b, c and lighter quarks) with 2 NN-tagged jets at different operating points, for the case of 6 jets and 7+ jets respectively. The reason for the separation into exactly 6 jets and 7+ jets is because both the multijets mass distribution and the jet combinationrics are quite different between the 2 cases, as shall be covered in more details in Chapter 5.

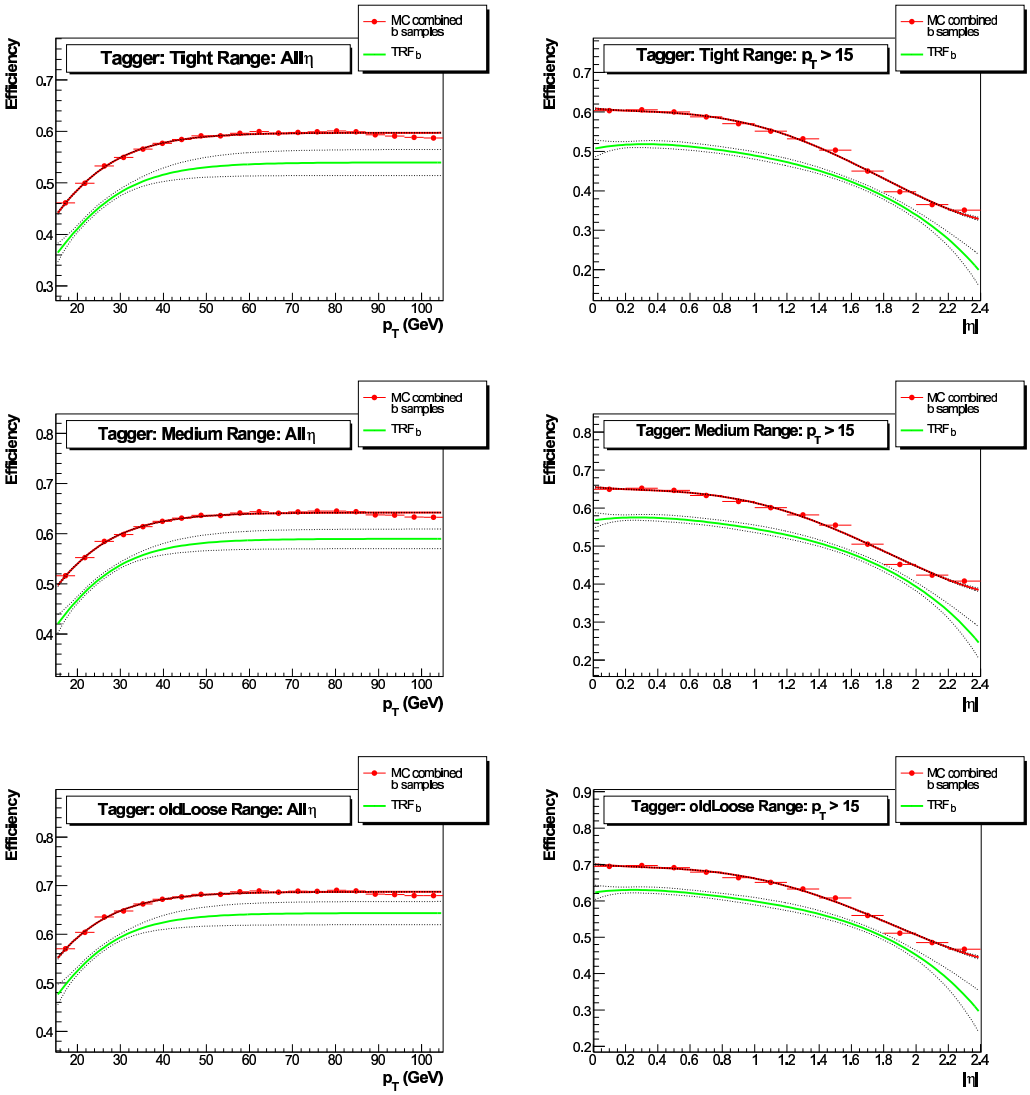

Figure 3.11. Tagging efficiency and TRF of b-quark jets with different p_T (left) and different η (right) using 3 different NN cuts: $NN > 0.5$ (old Loose), $NN > 0.65$ (Medium) and $NN > 0.775$ (Tight). The red curve shows the tagging efficiency and the green curve shows the TRF.

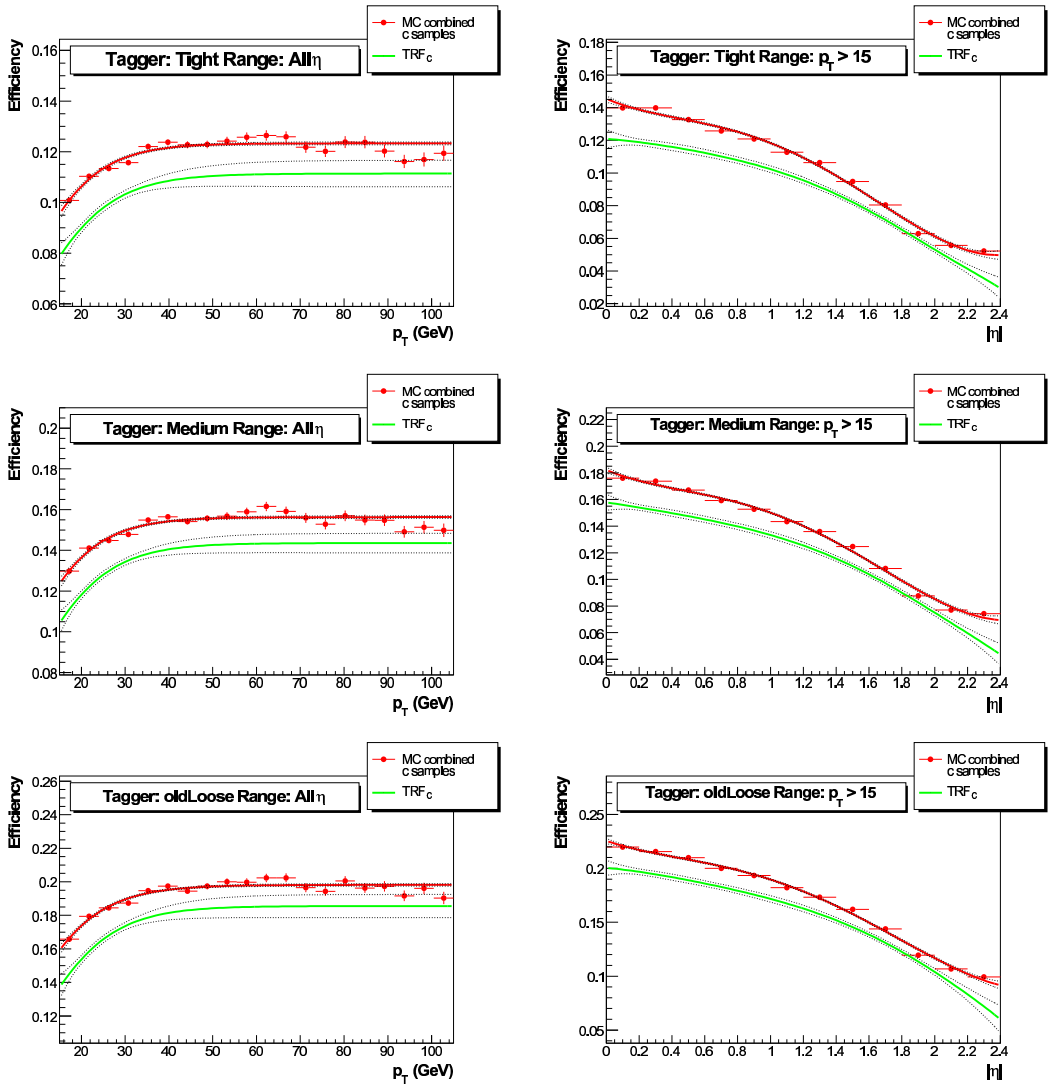

Figure 3.12. Tagging efficiency and TRF of c-quark jets with different p_T (left) and different η (right) using 3 different NN cuts: $NN > 0.5$ (old Loose), $NN > 0.65$ (Medium) and $NN > 0.775$ (Tight).

TABLE 3.2

Fraction of 2 b-tagged jets having different quark flavors in the exactly 6 jets $t\bar{t}$ sample. bb: 2 b-quarks, cc: 2 c-quarks, bc: 1 b and 1 c quarks, bq: 1 b-quark and 1 lighter quark, ll: 2 lighter quarks

NN output range (2 jets)	fraction bb	bc	cc	bl	ll
BIDNN > 0.1	0.597	0.237	0.010	0.134	0.022
BIDNN > 0.15	0.635	0.232	0.008	0.108	0.017
BIDNN > 0.2	0.668	0.226	0.007	0.086	0.013
BIDNN > 0.25	0.694	0.219	0.006	0.070	0.011
BIDNN > 0.325	0.724	0.207	0.006	0.056	0.008
BIDNN > 0.45	0.762	0.189	0.004	0.039	0.006
BIDNN > 0.5	0.774	0.183	0.003	0.034	0.005
BIDNN > 0.65	0.804	0.165	0.002	0.024	0.004
BIDNN > 0.775	0.831	0.148	0.002	0.016	0.003
BIDNN > 0.85	0.849	0.134	0.001	0.013	0.003
BIDNN > 0.9	0.867	0.120	0.001	0.009	0.003
BIDNN > 0.925	0.877	0.113	0.001	0.008	0.002
BIDNN 0.65-0.925	0.679	0.259	0.005	0.050	0.007
BIDNN 0.45-0.65	0.518	0.340	0.011	0.116	0.015
BIDNN 0.2-0.45	0.329	0.375	0.015	0.243	0.038
BIDNN 0.2-0.65	0.401	0.357	0.015	0.195	0.032
BIDNN 0.1-0.2	0.172	0.326	0.018	0.415	0.068

TABLE 3.3

Fraction of 2 b-tagged jets having different quark flavors in the 7 or more jets $t\bar{t}$ sample. bb: 2 b-quarks, cc: 2 c-quarks, bc: 1 b and 1 c quarks, bq: 1 b-quark and 1 lighter quark, ll: 2 lighter quarks

NN output range (2 jets)	fraction bb	bc	cc	bl	ll
BIDNN > 0.1	0.557	0.248	0.009	0.160	0.026
BIDNN > 0.15	0.597	0.245	0.008	0.130	0.020
BIDNN > 0.2	0.633	0.240	0.007	0.105	0.015
BIDNN > 0.25	0.664	0.231	0.006	0.087	0.012
BIDNN > 0.325	0.698	0.219	0.005	0.069	0.009
BIDNN > 0.45	0.740	0.202	0.005	0.047	0.007
BIDNN > 0.5	0.754	0.193	0.005	0.041	0.006
BIDNN > 0.65	0.787	0.174	0.003	0.030	0.005
BIDNN > 0.775	0.819	0.154	0.002	0.021	0.004
BIDNN > 0.85	0.841	0.139	0.002	0.015	0.003
BIDNN > 0.9	0.856	0.128	0.002	0.012	0.003
BIDNN > 0.925	0.867	0.117	0.002	0.011	0.002
BIDNN 0.65-0.925	0.658	0.268	0.006	0.058	0.009
BIDNN 0.45-0.65	0.478	0.367	0.013	0.129	0.013
BIDNN 0.2-0.45	0.297	0.376	0.010	0.278	0.038
BIDNN 0.2-0.65	0.360	0.371	0.012	0.226	0.031
BIDNN 0.1-0.2	0.134	0.311	0.020	0.474	0.062

CHAPTER 4

EVENT SELECTION, WEIGHTS AND TRIGGERS

4.1 Data samples

The DØ detector went through a major upgrade between 1996-2001 after the completion of Run-I. With numerous improvements in the detector components, data acquisition electronics and supporting software, data-taking restarted in late 2001 as Run-II. This analysis will be based on the events from the Run-IIa dataset which begins in July 2002 and ends in February 2006.

4.1.1 Data Quality requirements

Since the DØ detector is such a complicated machine, there are occasions that certain components of the detector are not performing as they should and the resulting events are not suitable for physics analysis. Therefore a number of quality requirements are established to ensure the full operation of all the detector components in the dataset.

The quality criteria are evaluated for each Run, where a Run is a data-taking period which usually lasts between two to four hours. The detector configuration remains constant within each Run. The performance of each sub-detector is evaluated in each Run and is marked as "good", "reasonable" or "bad", where "bad"

usually indicate the sub-detector is not working properly or completely switched off.

In addition, the Calorimeter has a more refined data quality criteria which are based on luminosity blocks. A luminosity block is a much shorter period of time (around one minute) for the calculation of the integrated luminosity. Certain noise problems in the calorimeter (such as the firing of a ring of energy around ϕ) arise randomly and lasts only a fraction of a Run. Therefore rejecting bad luminosity blocks ensures good performance while preserving some good data.

4.2 Trigger details and simulation

As mentioned in the detector chapter, the trigger system of DØ only picks out interesting events. Even after the trigger selection, the amount of data is still large and it is necessary for one to pick out the triggered events that are relevant to the final state of analysis that one is looking for. In the hadronic decay of $t\bar{t}$ pairs the final state consists of 6 or more jets. Therefore it is natural to look at the events passing the multijet triggers only.

The two sets of the triggers that are particularly interesting are the 3JT and 4JT triggers, which require three and four jets in the events respectively. The implementations of the triggers during Run-IIa are different depending on the trigger version, since constant improvement of the trigger system is introduced by the trigger experts.

The detailed definitions of the 3JT and 4JT triggers chosen for this analysis are shown in the Appendix section. The corresponding run ranges in each trigger version.are shown in table (table 4.1 and 4.2).

TABLE 4.1

The different run ranges used defined in terms of run numbers from trigger list v8-v12

Run range	Trigger version [from	to)	Run numbers [from-to)
v8	v8.00	v9.00	160582-167016
v9 - v10	v9.00	v11.00	167019-174803
v11	v11.00	v12.00	174845-178018
			178080-178097
			178104-178560
			178645-178722
v12	v12.00	v13.03	178019-178071
			178097-178104
			178618-178620
			178722-194567
			194595-194598

TABLE 4.2

The different run ranges used defined in terms of run numbers in trigger version v13-v14

Run range	Trigger version [from	to)	Run numbers [from-to)
v13a	v13.03	v13.30	194567-194574
			194662-202025
v13b	v13.30	v14.00	202025-207217
			207229-207258
			207288-207343
			207350-207681
			207749-208146
			208246-208247
			208444-208501
			209102-209105
v14a	v14.00	v14.80	207217-207223
			207279-207280
			207343-207346
			207719-207741
			208168-208207
			208276-208435
			208512-209100
			209147-211483
v14b	v14.80	v15.00	211513-215671

TABLE 4.3

The luminosity and trigger choices for the different run ranges. The luminosity shown is the integrated recorded luminosity after data quality selection

Run range	Trigger	$\mathcal{L}(pb^{-1})$
v8	4JT	24.54
v9	3JT OR 4JT	24.07
v10	3JT OR 4JT	10.22
v11	3JT OR 4JT	64.68
v12	3JT OR 4JT	210.23
v13a	3JT OR 4JT	51.72
v13b	3JT OR 4JT	322.12
v14a	3JT OR 4JT	189.83
v14b	3JT OR 4JT	143.55
Total		1040.96

Since both 3JT and 4JT triggers are preferred for this analysis, by ORing the triggers we are able to increase the acceptance of data events. Table 4.3 shows the integrated luminosity by combining the 3JT and 4JT triggers in each trigger version except v8.

4.2.1 Trigger Efficiencies

Given that all the data events need to pass the trigger selection in order to be stored, we need to investigate the triggers in more details and understand their effects on the signal of interest. This involves the extraction of trigger efficiencies from the data and combining them into a single efficiency per event.

There are 3 levels of triggers as mentioned before in the detector section, together with an additional L4 trigger which will be discussed later in this section. The total probability of an event will then be:

$$P_{evt} = P(L1) \times P(L2|L1) \times P(L3|L1L2) \times P(L4) \tag{4.1}$$

With the additional ORing of the 2 different triggers, the calculation will be more involved:

$$\begin{aligned} P_{evt}(3JT \cup 4JT) = & P(L1, 3JT \cup 4JT) \times P(L2, 3JT \cup 4JT|L1) \times \\ & P(L3, 3JT \cup 4JT|L1L2) \times P(L4) \end{aligned} \tag{4.2}$$

where the OR operation of each trigger level is:

$$P(Ln, 3JT \cup 4JT) = P(Ln, 3JT) + P(Ln, 4JT) - P(Ln|3JT \cap 4JT) \tag{4.3}$$

The above calculations need to be done separately for each trigger version, and then weighted by the recorded luminosity of each trigger version in table 4.3:

$$P_{evt} = \frac{\sum_{v8-v14} P_{\text{trig ver.}} \mathcal{L}_{\text{trig ver.}}}{\sum_{v8-v14} \mathcal{L}_{\text{trig ver.}}} \tag{4.4}$$

The trigger probability is essential for the modeling of the signal events, as will be shown at the end of this section.

4.2.2 L1 Trigger Turn-ons

The Level 1 portion of the of the 3JT and 4JT triggers are multi-jet p_T triggers. The single jet trigger turn-on curve has been determined by the jet ID group. It can be shown that out of N jets in an event, the probability of at least one jet satisfying the trigger condition is:

$$P(\geq 1\text{jet L1 } p_T > x) = 1 - P^0 = 1 - \prod_{i=1}^{N}(1 - P_i(x)) \tag{4.5}$$

Where $P_i(x)$ is the probability of the ith jet having L1 $p_T > 5$ GeV. The condition will be a little bit more complicated for multi-jet triggers. For example in the CJT(3,5) trigger, which requires at least 3 jets with L1 $p_T > 5$ GeV is:

$$\begin{aligned} P(\geq 3\text{jet L1 } p_T > 5) &= 1 - P^0 - P^1 - P^2 \\ &= 1 - \prod_{i=1}^{N}(1 - P_i(5)) - \sum_{i=1}^{N} P_i(5) \prod_{j\neq i}^{N}(1 - P_j(5)) \\ &- \sum_{i=1}^{N} P_i(5) \sum_{j=1}^{N} P_j(5) \prod_{k\neq i,j}^{N}(1 - P_k(5)) \end{aligned} \tag{4.6}$$

Similar calculations of the multijet triggers will be used in the L2 and L3 triggers as well.

4.2.3 L2 Trigger Turn-ons

The Level 2 portion of the of the 3JT and 4JT triggers consists of the jet p_T triggers and the L2HT triggers. Figure 4.1 and 4.2 show the L2 HT50 and HT90 turn-on curves respectively from trigger versions v9-12. Figure 4.3 shows the L2 HT70 turn-on curves for trigger version v13 and ETA26HT75 for trigger version v14.

4.2.4 L3 Trigger Turn-ons

As usual the Level 3 portion of the 3JT and 4JT triggers both contain some jet triggers as well. The resulting multiple jet terms can also be calculated using single jet turn-ons as well.

In addition, there exists a CFT Vertex term in 3JT where it requires the z position of the primary vertex to be within 35 cm of the detector origin. Such triggers were used from trigger version 12 onwards. Figure 4.4 shows the turn-on curve from trigger versions v12-14, as a function of offline (reconstructed) z position of primary vertex. It can be seen that the turn-ons have very sharp rise around $|z_{pv}| = 35cm$.

Also in trigger version v13 there is a L3 HT term in 4JT triggers, which requires the L3 HT (sum of L3 Jet E_T for E_T greater than 9 GeV, and Jet $|det\eta| < 3.6$.) Figure 4.5 shows the turn-on curve from trigger versions v13, as a function of offline JES corrected Jet HT (sum of JES corrected Jet $E_T > 15$ GeV).

The 3JT trigger also contains a L3 IP component in trigger version v13 and v14. However this term cannot be obtained directly as it is correlated with b-tagging results of the jets. Since we are using BID Neural Network tagger, we can calculate the chance of IP term fired in some BID Neural Network regions using

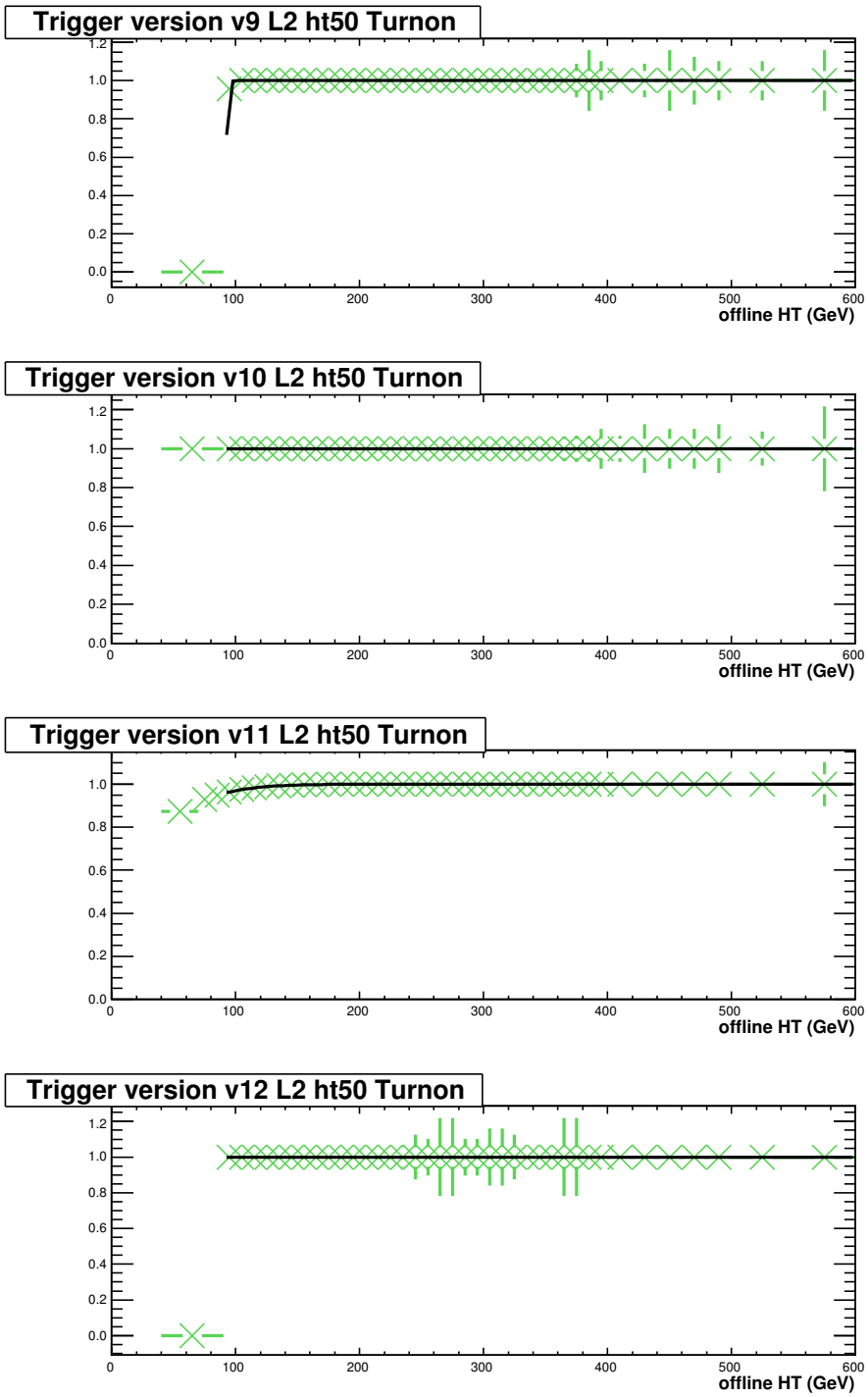

Figure 4.1. L2 $H_T > 50$ GeV turn-on curves in data JES corrected H_T (JES correction applied to individual jets in data as described in section 3.4, then $H_T = \sum_i^{jet} p_{Ti}$) for different trigger version v9-v12

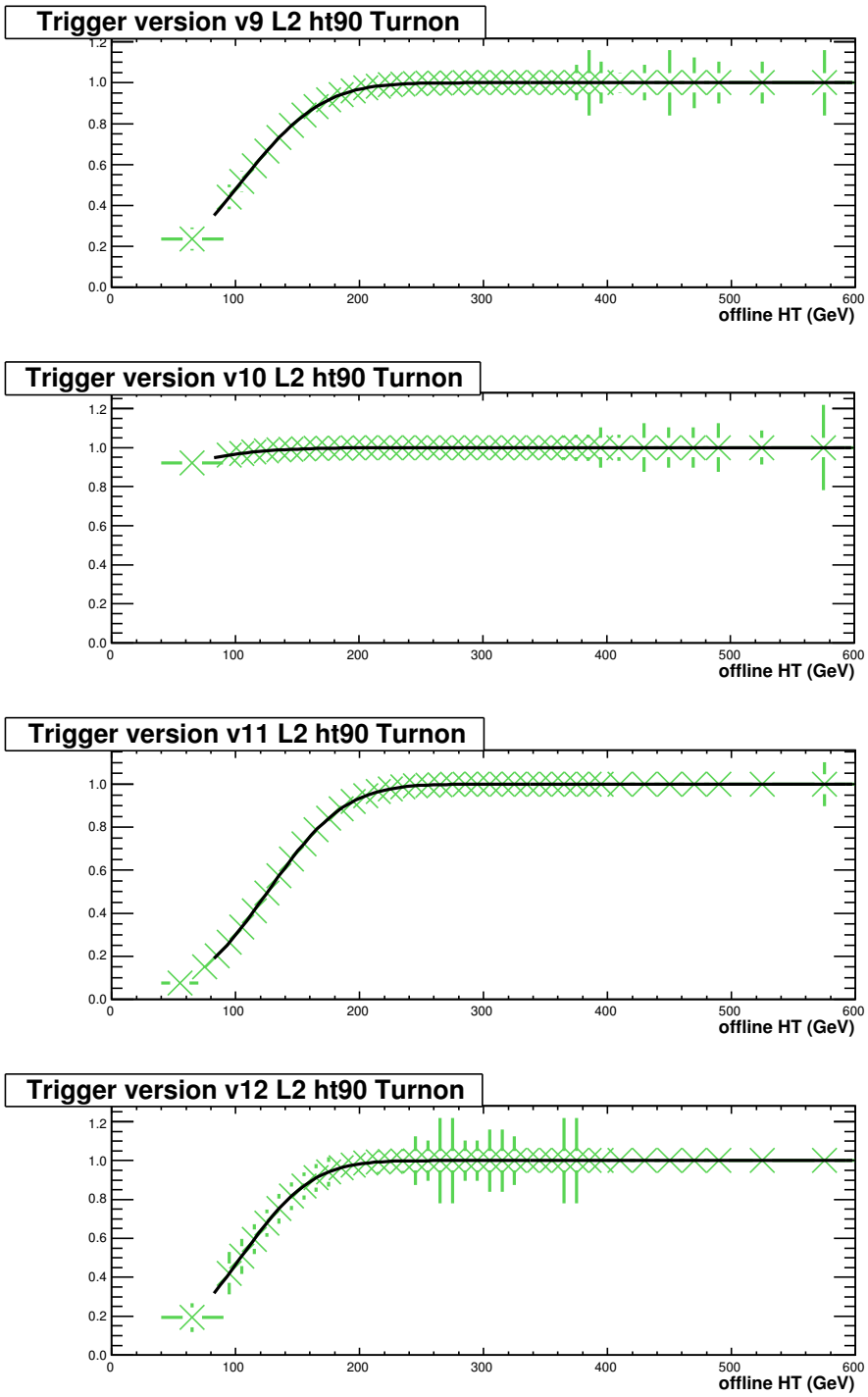

Figure 4.2. L2 $H_T > 90$ GeV turn-on curves in offline JES corrected H_T for different trigger version v9-v12

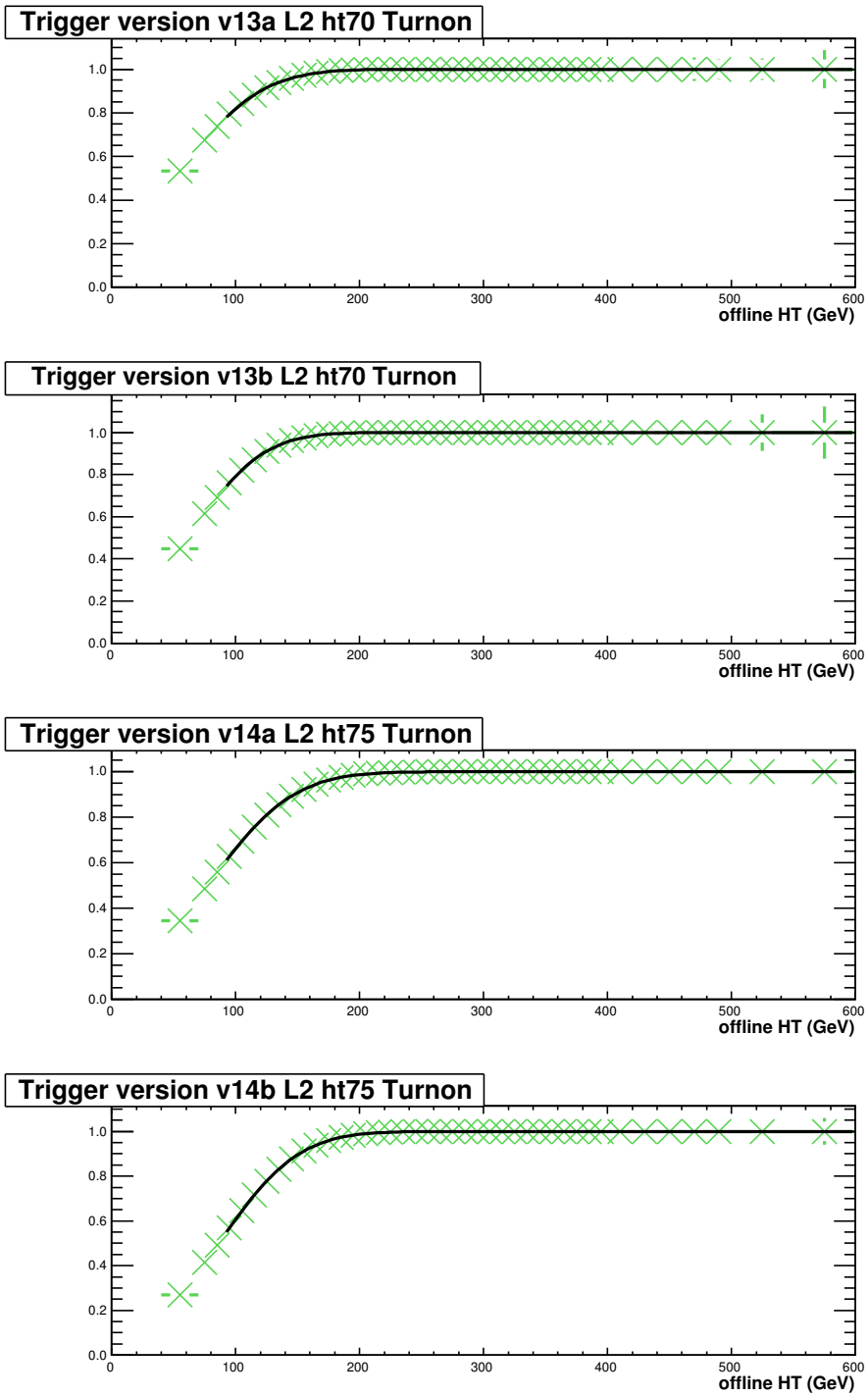

Figure 4.3. L2 $H_T > 70$ GeV and $H_T > 75$ GeV turn-on curves in offline JES corrected H_T for different trigger version v13 and v14

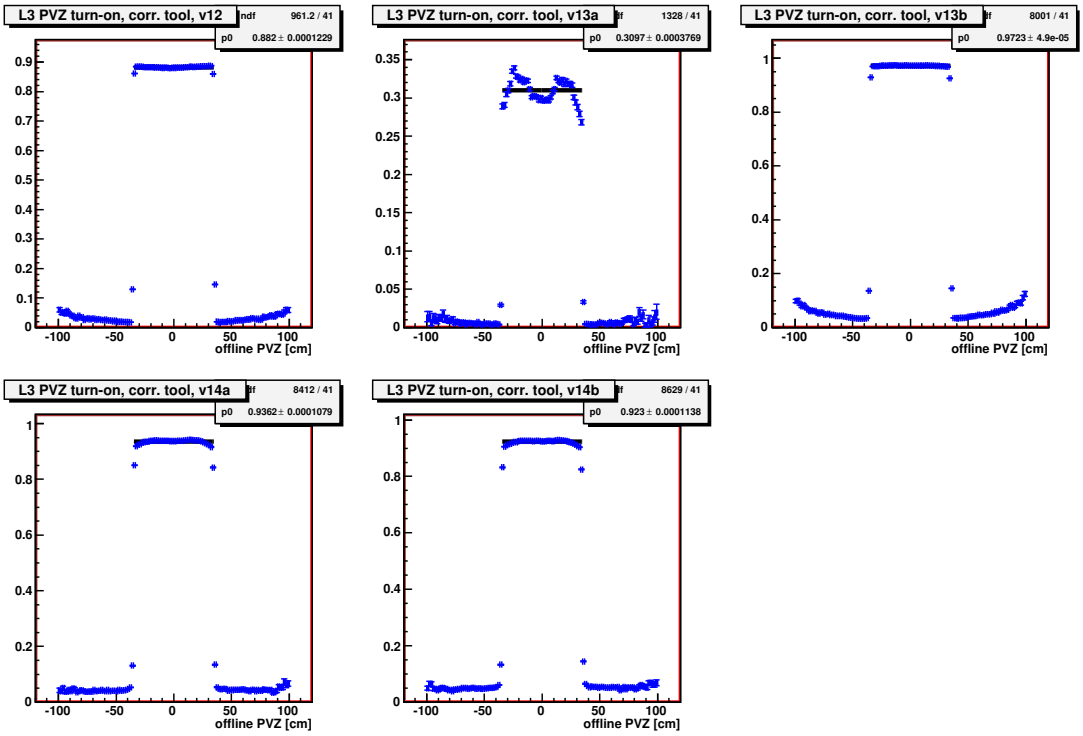

Figure 4.4. L3 $CFTVertex > 20$ GeV turn-on curves in offline primary vertex z position, for different trigger version v12-v14

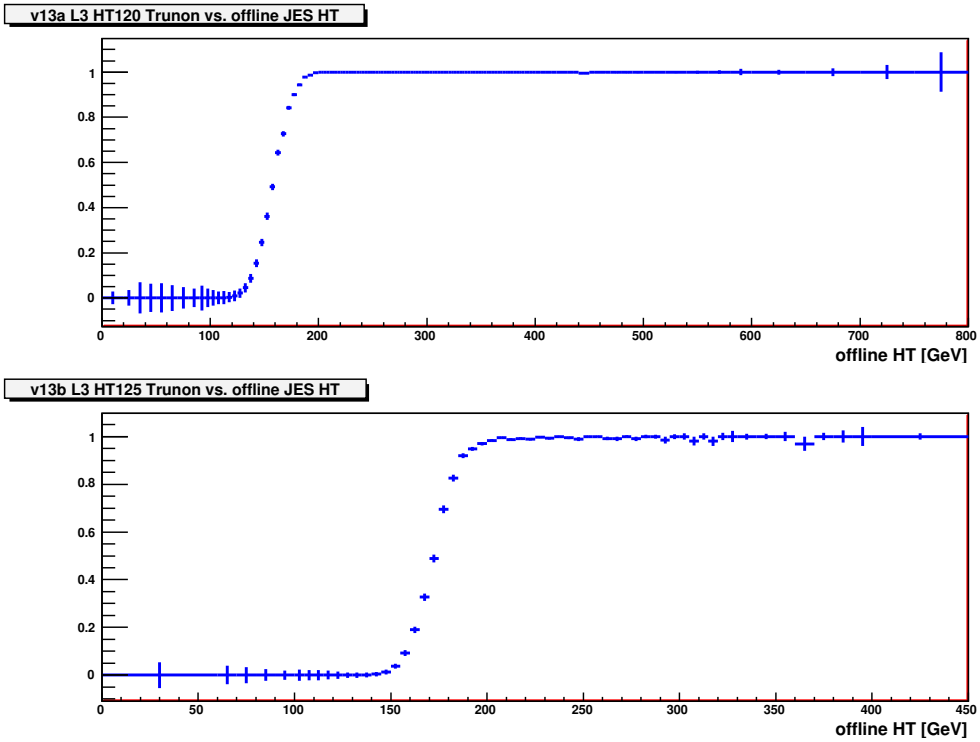

Figure 4.5. L3 HT120 turn-on curves in offline JES corrected Jet E_T, for trigger version v13a and v13b

TABLE 4.4

$P(IP)$ term from 4JT triggered data events, 6 offline jets $p_T > 15$ GeV

Trigger Version	$P(IP)$	Error
v13a	0.2249	0.0031
v13b	0.2656	0.0015
v14	0.2390	0.0019

Bayes' law:

$$P(IP|BID[NN, ntag]) = \frac{P(BID[NN, ntag]|IP)}{P(BID[NN, ntag])} P(IP) \tag{4.7}$$

where $P(IP)$ is the probability of a event having the L3 the IP term fired regardless of the neural network b-tagging results. This is estimated by looking at the ratio of 4JT triggered data events with 6 offline jets $p_T > 15$ GeV, $|\eta_{det}| < 2.6$, and all taggable.

$$P(IP) = \frac{(0.0 < IP_{L3} < 0.05) \wedge (L3online) \wedge (4JT)}{(L3online) \wedge (4JT)} \tag{4.8}$$

Table 4.2.4 shows the $P(IP)$ results from real data events.

Then the $P(BID[NN, ntag]|IP)$ term is estimated from the real data events having different number of b-ID tags and different b-ID NN values, as a fraction of 3JT triggered events (which have the IP term fired). Again the events need to satisfy 6 offline jets $p_T > 15$ GeV, $|\eta_{det}| < 2.6$ and all taggable.

Figure 4.6 shows an example plot $P(BID[NN, ntag]|IP)$ in different b-ID

NN cases in trigger version v14a. Since we have chosen the requirement of 2 b-ID NN tags for our analysis, starting with b-ID NN> 0.2, therefore we shall use the red histogram value at "BIDNN0.2" in the plots.

Finally for the $P(BID[NN, ntag])$ term, this can be estimated using the MC b-ID NN TRF (Tag rate function). This is explained in more details in the b-ID section of this note. Again, we are mainly interested in value of the operating point 2 b-ID tagged jets with NN> 0.2.

4.2.5 L4 Trigger

In addition to the hardware triggers, the 3JET SKIM criteria are also applied to the data to pick those events with high jet multiplicity. This skim is implemented in software and requires those offline jets having detector eta within 2.6 satisfy the following p_T cuts:

Leading p_T Jet: $p_T > 25$ GeV

2nd p_T Jet: $p_T > 15$ GeV

3rd p_T Jet: $p_T > 15$ GeV

To investigate the effects of the skim on MC events, it is convenient to treat the 3JET skim as an additional level of trigger "L4". Therefore the efficiency of the skim cut can be treated the same way as the trigger.

Figure 4.7 shows the JES corrected Jet p_T turn-on curves for L4 $p_T > 20$ GeV comparing JES corrected p_T and JESMU (JES with muon correction) p_T. Figure 4.7 shows the turn-on curves for L4 $p_T > 15$ GeV. Notice that turn-on curves using JESMU p_T do not have "sharper" turn-ons (i.e., they approach the plateau at a slower rate) compared to the JES turn-on curves. This is expected as JESMU introduces extra correction and smearing to the turn-on curves Therefore the JES

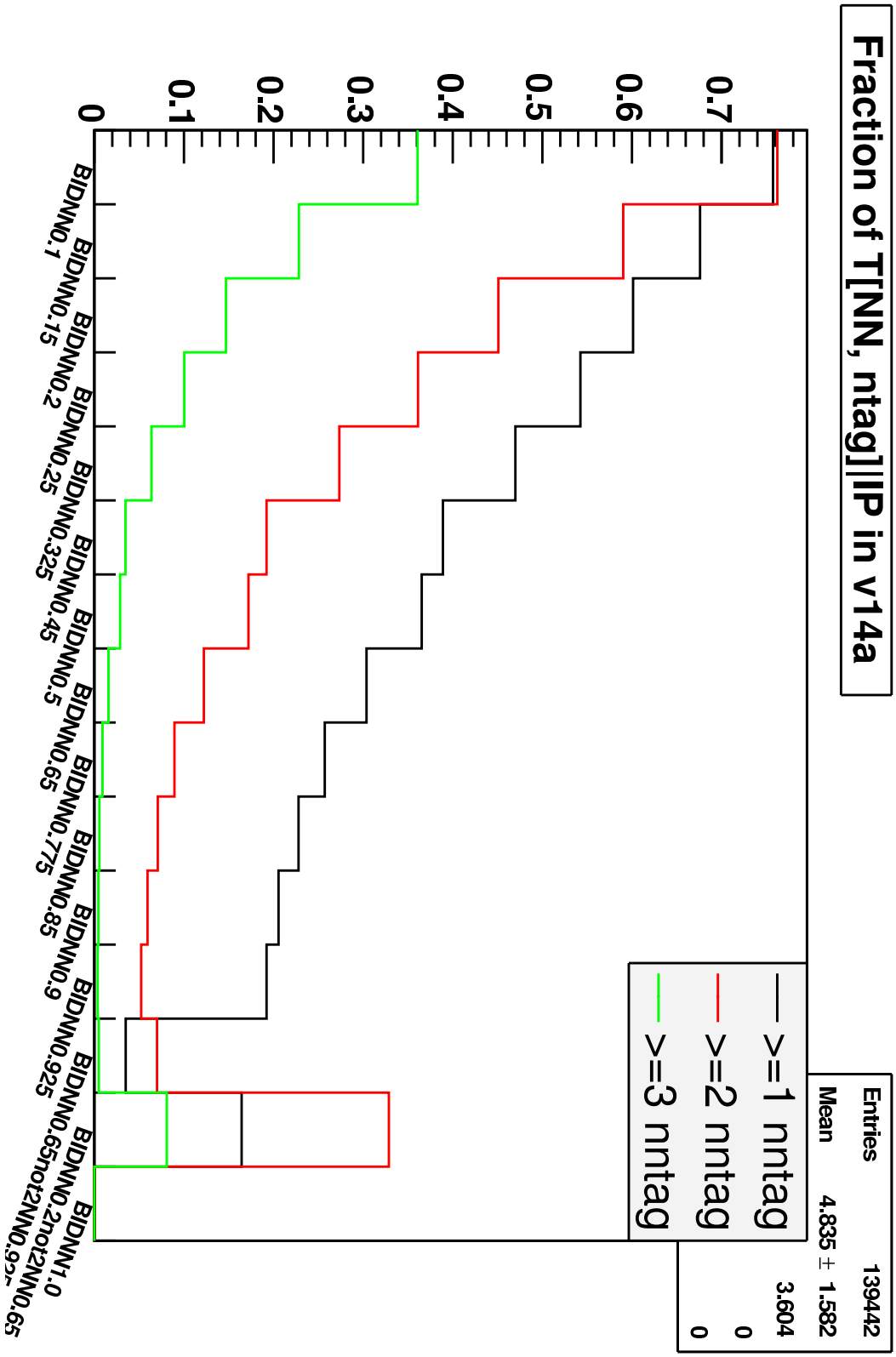

Figure 4.6. $P(BID[NN, ntag]|IP)$ in trigger version v14a. The three different lines correspond to events having ≥1, ≥2 and ≥3 b-ID NN tags. The horizontal axis shows the different b-ID NN operating points.

turn-on curves are preferred over the JESMU turn-on curves for trigger probability extraction.

4.2.6 Trigger simulation in Monte Carlo simulation

All the data events have been filtered through the trigger system and therefore have their phase space modified by the specific requirements of the trigger selections. This causes what is commonly referred to as trigger bias. It is therefore necessary to replicate the effect of the trigger system on the Monte Carlo events such that their phase space is modified to the same extent. Since all the trigger turn-ons have already been derived, one can then take the MC event and feed in the physics objects to obtain the probabilities of each trigger term. Then we can combine the individual trigger probabilities using the above recipe and get a single event probability. Then a random number between 0 and 1 is drawn and compared with the event probability to determine if the event is accepted or rejected.

4.3 Luminosity re-weighting

The real data events taken by DØ have different luminosity. Events of higher luminosities have larger number of hard scattering processes per event on average, which results in a larger number of primary vertices within the event. This changes the event characteristics and therefore such effects should be replicated in the Monte Carlo events as well. This can be done by extracting the luminosity profile of the data and Monte Carlo events. Then relative scaling factors in each luminosity bin can be determined and applied to Monte Carlo events within that bin. The details of implementation is discussed in Ref. [57].

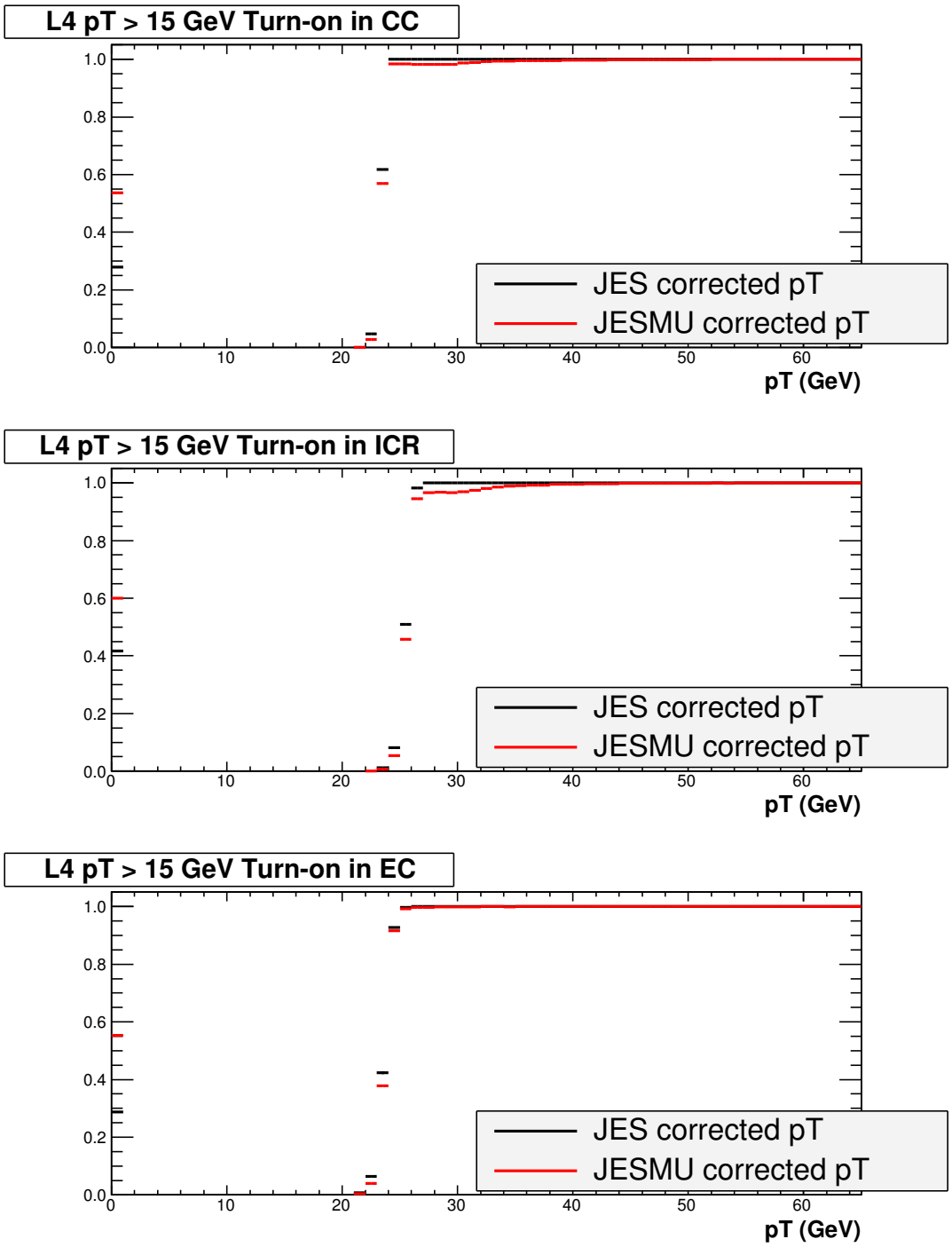

Figure 4.7. L4 $p_T > 15$ GeV turn-ons using JES corrected jet p_T for different detector regions: CC ($|\eta_{det}| < 0.8$), ICR ($0.8 < |\eta_{det}| < 1.2$) and EC ($1.2 < |\eta_{det}| < 2.6$). The red points correspond to JES with muon corretion.

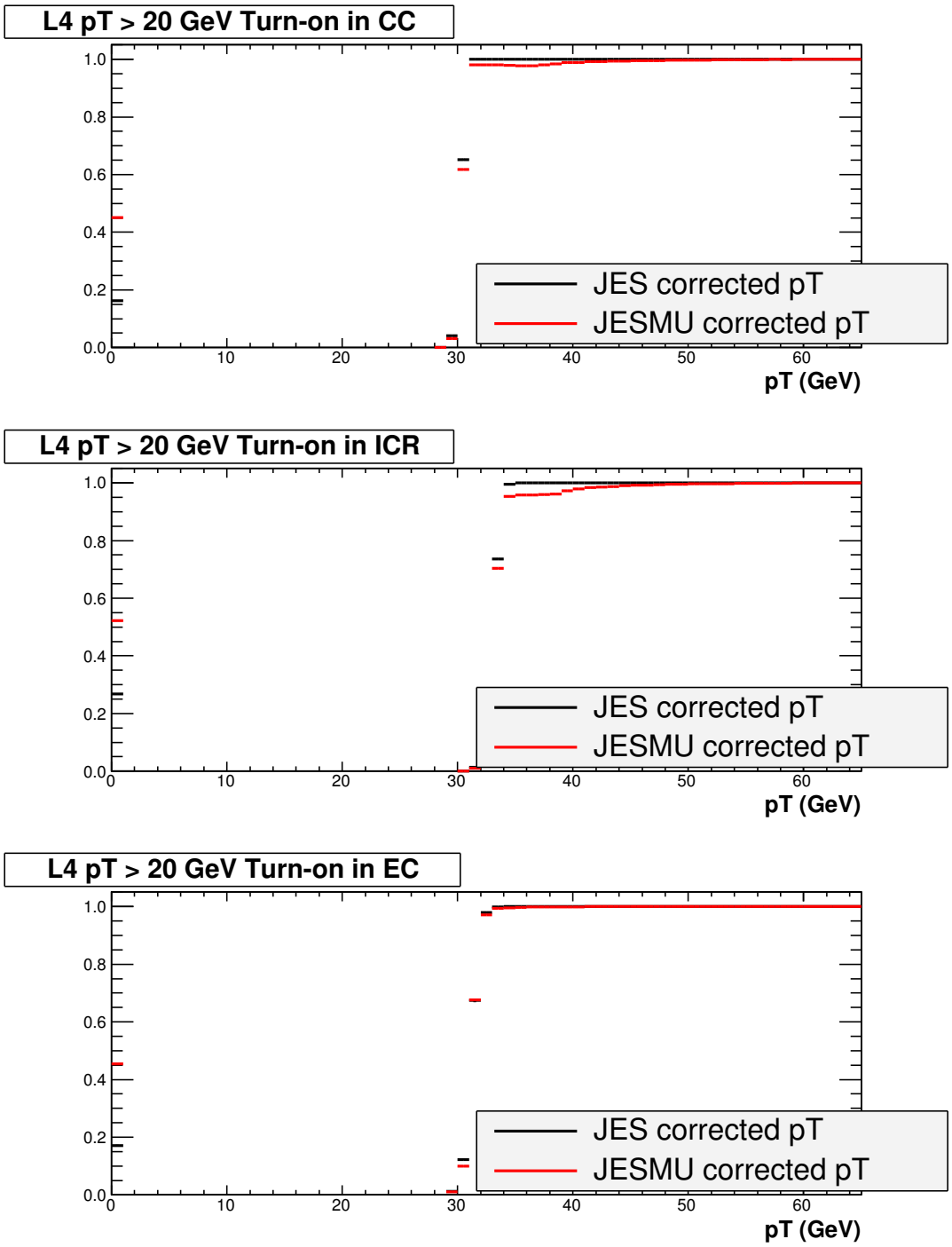

Figure 4.8. L4 $p_T > 20$ GeV turn-ons using JES corrected jet p_T for different detector regions: CC ($|\eta_{det}| < 0.8$), ICR ($0.8 < |\eta_{det}| < 1.2$) and EC ($1.2 < |\eta_{det}| < 2.6$). The red points correspond to JES with muon corretion.

4.4 Event Selection

After the trigger system and the skimming cuts, the data events are still dominated by background events. Therefore additional event selection cuts are required to increase the signal fraction of the event. The selection are based on the properties of the all hadronic $t\bar{t}$ events as well as the QCD background events, and will be covered in details below.

4.4.1 Primary Vertex Position

From the L3 PVZ turn-ons figure 4.4, it is shown that the plateau of the turn-ons are $|z| < 35cm$. Therefore it is safe to require the primary vertex z position to be within $\pm 34cm$.

4.4.2 Lepton veto

In order to be orthogonal to the other leptonic decay modes of $t\bar{t}$ events, one need to reject event with a tight isolated electron satisfying:

- Calorimeter isolation < 0.15
- Electromagnetic fraction (emf) < 0.9
- H-matrix $7 < 50$
- E/p track match chi-squared > 0.0
- Electron Likelihood > 0.85

Events with one or more isolated muons satisfying the following will be rejected as well:

- muon has a global track match together with converged local fit
- >= 1 A layer scintillator hit and >= 2 A layer drift tube hit
- >= 1 BC layer scintillator hit and >= 3 BC layer drift tube hit
- Calorimeter $E_T < 2.5$ GeV within cone $(0.1 < dR < 0.4)$
- Track $p_T < 2.5$ GeV/c within cone $(dR < 0.5)$

4.4.3 Jet quality and p_T cuts

Given that there are 6 jets in the final state of the all-hadronic $t\bar{t}$ decay, it is natural to require the event to have at least 6 jets. Also due to the background from multiple interactions, all 6 jets have to satisfy track multiplicity and momentum of the taggability requirement and be within $|\eta_{det}| < 2.5$. Jets outside of this eta range are not covered by the tracking system and therefore have a sharply lower taggability. Requiring all 6 jets to be taggable ensure the jets are coming from the same primary vertex.

Given the jet finding algorithm and the jet p_T threshold in the algorithm, the jets are required to satisfy $p_T > 15$ GeV. In addition, in $t\bar{t}$ events the final quarks are produced from heavy objects such as the top quark and W bosons, and therefore the resulting jets are more energetic compare to the jets from background events (which are dominantly QCD multijet background as will be discussed in chapter 5). The 3 leading p_T jets are required to have $p_T > 40$ GeV to increase the signal fraction in the data.

TABLE 4.5

Summary of event selection cuts

lepton veto	no isolated μ/e
6 or more jets	
all jets taggable	
$\lvert z_{PV} \rvert$	$< 34cm$
jet $\lvert \eta_D \rvert$	< 2.5
jet p_T 1-3	> 40 GeV/c
jet p_T 4+	> 15 GeV/c
2 NN tagged jets p_T	> 40 GeV

4.4.4 b-tagging requirement

As each $t\bar{t}$ event yields two b-jets, therefore one can require two jets to be tagged to reduce the background. The best choice of the tagger being used is the BID Neural Network (NN) tagger. Also as both of the b-quarks from top quarks decay tend to have high jet p_T, therefore it is required the b-jets satisfy $p_T > 40$ GeV/c as well.

Since the Neural Network tagger gives the likelihood of a jet to be a b-jet from a scale of 0 to 1, it allows separation of the data events with different characteristics into different BIDNN regions. In this analysis the samples will split into 3 different regions: 6 jets 2 BIDNN > 0.65, 6 jets 2 BIDNN 0.2-0.65 and 7+ jets, 2 BIDNN > 0.2, which will be explained in more detail in Chapter 5.

The summary of the selection cuts is shown in table 4.4.4.

To illustrate how the selection cuts perform on the MC events and data, Table 4.4.4 shows the cut efficiencies and the culmulative efficiencies of the selection cuts. It can be seen that there are lots of background events in the data events and the selection cuts reduce a significant portion of background.

TABLE 4.6: Selection cut efficiencies of the MC $t\bar{t}$ and data events. Inclusive $t\bar{t}$ indicates all decay modes. Culmulative efficiencies are the efficiencies with respect to the first row before any cuts

	Inclusive $t\bar{t}$			All hadronic $t\bar{t}$			Data		
Selection	Total	Eff.	Cul. Eff.	Total	Eff.	Cul Eff.	Total	Eff.	Cul Eff.
Initial	754042	1.000	1.000	344645	1.000	1.000			
>= 6 jets $p_T > 15$, $\|\eta_{det}\| < 2.5$	171984	0.228	0.228	146268	0.424	0.424			
Primary vertex	144277	0.839	0.191	122751	0.839	0.356			
Trigger (3JT or 4JT)	123136	0.853	0.163	107468	0.876	0.312	1737862	1.000	1.000
lepton veto	119521	0.971	0.159	107437	1.000	0.312	1737573	1.000	1.000
all jets taggable	76604	0.641	0.102	71200	0.670	0.209	571908	0.329	0.329
>= 2 jets $p_T > 15$, BIDNN> 0.2	50856	0.664	0.067	47892	0.665	0.139	77647	0.136	0.045
>= 1 jet $p_T > 40$	50841	1.000	0.067	47879	1.000	0.139	75259	0.969	0.043
>= 2 jets $p_T > 40$	50454	0.992	0.067	47539	0.993	0.138	61381	0.816	0.035
>= 3 jets $p_T > 40$	45593	0.904	0.060	43168	0.908	0.125	29021	0.473	0.017
>= 2 jets $p_T > 40$, BIDNN> 0.2	31076	0.682	0.041	29434	0.682	0.085	12073	0.416	0.007

CHAPTER 5

SIGNAL AND BACKGROUND STUDY

As a means to predict the outcome of the experiment and to check the consistency of the experiment, simulated events are generated using the Monte Carlo method. This is especially important for the mass analysis as it allows one to study the mass dependent effects of the top quark and develop better methods to extract the desired signal.

The Monte Carlo event simulation consists of two parts. The first part simulates the hard scattering of the $p\bar{p}$ using PYTHIA 6.3 [58] or Alpgen 1.2 [59] Monte Carlo Event Generators. The parton distribution function (PDF) used to model the proton and anti-proton quark/gluon content is CTEQ6.1M [72]. The subsequent production, fragmentation and decay of the partons are then simulated using PYTHIA 6.3. In this study, events of $p\bar{p} \rightarrow t\bar{t} \rightarrow bW^{+}\bar{b}W^{-}$ with m_{top} ranging from 110 GeV to 550 GeV are generated. The generated samples are inclusive in the sense that the different decay modes of the W are included, which provides all-leptonic, semi-leptonic and all-hadronic $t\bar{t}$ events.

With the parton level physics process generated, the second step performs the simulation of the DØ detector. This is done using the software d0gstar, which is based on GEANT 3 [60]. The d0gstar simulates the interaction of the various particles with the material inside the detector – including scattering, pair production and deposit of energy in different materials. In addition, the occasion of having

more than one $p\bar{p}$ interaction is also simulated. This is done by overlapping minimum biased events (which includes low p_T QCD interactions, single and double diffractions) on top of the simulated events. The detector output from d0gstar is then collected and combined by the d0sim software which simulates the signal outputs of the detector.

The resulting simulated events have the same format as real data coming from the detector, which will then go through the same event reconstruction and object identification as described in chapter 3.

5.1 MC Jet Energy Scale

The JES correction discussed in chapter 3.4.3 is designed to be applied on both MC and data and is referred to as the absolute JES. However, such a correction is insufficient for Monte Carlo events as it doesn't include the additional smearing effect of the jet energy due to the imperfect detector simulation. The resulting mismatch can be observed from the p_T imbalance, ΔS, which measures the difference between the photon and the jet in the simulated γ+jet events:

$$\Delta S = \frac{p_T^{jet} - p_T^{\gamma}}{p_T^{\gamma}} \tag{5.1}$$

An additional JES correction recipe, named Shifting, Smearing and Removing (SSR) is developed to apply on the jets of MC simulated event [61]. The jet p_T of the simulated jets is first oversmeared using a smearing factor, which is randomly picked from a Gaussian distribution of mean 0 and a width depending on the particle jet p_T. Then the jet p_T is shifted and the jet is removed if the corrected jet p_T is below 15 GeV. Here is the corrected energy of the simulated jet:

$$p_T^{new} = p_T^{RECO} + p_T^{ptcl}[shift(p_T^{ptcl}) + Gauss(0, smear(p_T^{ptcl}))] \tag{5.2}$$

The jet energy is then scaled by the fraction p_T^{new}/p_T^{RECO} such that the direction of the jet remains the same. The resultant SSR correction can be referred as the relative JES correction between MC and data. In this thesis such correction will be referred to as JES Smeared.

5.1.1 Sample dependent correction

Given that the JES is derived from photon+jet events, the derived correction may be slightly different from those derived using $t\bar{t}$ events. Therefore it is preferred to adjust the JES correction factor correspondingly. This can be done by looking at the ratio of MC jet response between photon+jets and $t\bar{t}$ events in different η^{det} regions and jet flavor. Figure 5.1 shows the fitted correction of Jet energy in $t\bar{t}$ MC events.

5.2 Jet Combinatorics

In the all-jets channels of top decay 6 jets are produced. Each combination of 3 jets can yield a measurement of the top mass. For example in the case of exactly 6 jets there are $6!/3! = 120$ different ways to associate the reconstructed jets with the partons in calculation of the top mass.

For top and W mass measurements the order of the 2 jets associated with the W decay does not matter, which reduces the number of combinations to $6*5*3 = 90$. The number of jet combinations further reduces to $2*3 = 6$ in the case of exactly 6 jets with 2 of them b-tagged. For the case of 7 jets there are 30 different jet combinations in the case of 2 b-tagged jets, and 90 jet combinations with 3 b-

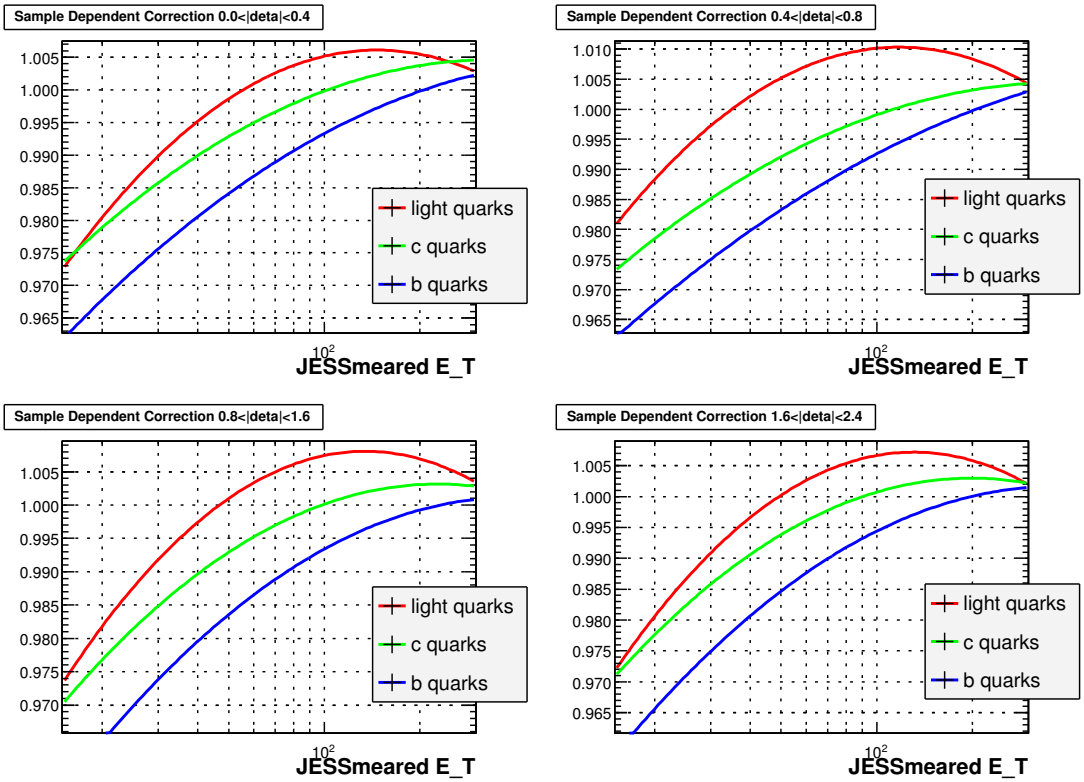

Figure 5.1. Jet energy correction factor for $t\bar{t}$ events, for b-quark jets, c-quark jets and light quark jets in the 4 η_{det} regions.

tagged jets. The purpose of this section is to present a method of jet combination selection for a better top mass measurement.

5.2.1 best chi-square selection

For each jet combination, one can reconstruct an empirical W and top pair by treating the reconstructed jet as a parton. Based on the SM, the top quark pair produced in our experiment should have the same mass. Also, the W pair has the same mass value already measured in other experiments. Therefore it is natural to construct a chi square operator to pick out the jet combination which minimizes the difference between the two measurements of top mass, as well as difference between the measured W mass and the standard value:

$$\begin{aligned}\chi^2 =& \frac{(mass(b_1 j_1 j_2) - mass(b_2 j_3 j_4))^2}{\sigma_t^2} + \\ & \frac{(mass(j_1 j_2) - 80.4 \text{ GeV})^2 + (mass(b_1 j_1 j_2) - 80.4 \text{ GeV})^2}{\sigma_W^2}\end{aligned} \tag{5.3}$$

Where σ_t and σ_W are the resolution of top and W mass differences respectively. Such quantities are derived from MC events using reconstructed jets. Figure 5.2, 5.4, 5.6 and 5.8 show the top mass differences in 4 different BIDNN regions: 6 jets 2NN¿0.925, 6 jets 2NN 0.65-0.925, 6 jets 2NN 0.2-0.65 and 7+ jet 2NN¿0.2. Similarly, Figure 5.3, 5.5, 5.7 and 5.9 show the W mass differences in 4 different BIDNN regions: 6 jets 2NN¿0.925, 6 jets 2NN 0.65-0.925, 6 jets 2NN 0.2-0.65 and 7+ jet 2NN¿0.2. From the plots we can see that the mass resolution is pretty much independent of BIDNN cuts and correct combinations, as long as all RECO jets are matched to daughter partons of the top quark. The resolution terms are then taken from the width of the Gaussian fit and are shown in table 5.1

Sample	σ_t(GeV)	σ_W(GeV)
Exactly 6 jets	25.5	13.2
7+ jets	16.0	9.2

TABLE 5.1

Mass resolution values used in jet combination chi-square

5.2.2 jet combination

The best chi-square selection yields a combination of 2 jet triplets. To evaluate the performance on such a selection it would be useful to see how often the b-tagged jet in each triplet matches the b quark from top, and to see if two other jets in the triplet also match the 2 lighter quarks decay from the same top quark as the b jet in the same triplet.

First we need to define what is meant by matching. A RECO jet is considered "matched" to a MC parton if it satisfies the condition:

$$\begin{aligned} dR(jet_{RECO}, parton_{MC}) < 0.5; \\ dR = \sqrt{d\eta^2 + d\phi^2} \end{aligned} \tag{5.4}$$

The choice of 0.5 in the cut is based on 2 considerations. Figure 5.10 shows the angular separation (dR) between the MC parton and its closest RECO jet (if it exists), using MC ttbar samples. It can be seen that the majority of the b-tagged jets are within $dR < 0.5$ of the originating b-quark. Only a tiny fraction of the b-tagged jets are out of such a range and therefore it is sufficient to pick 0.5 as

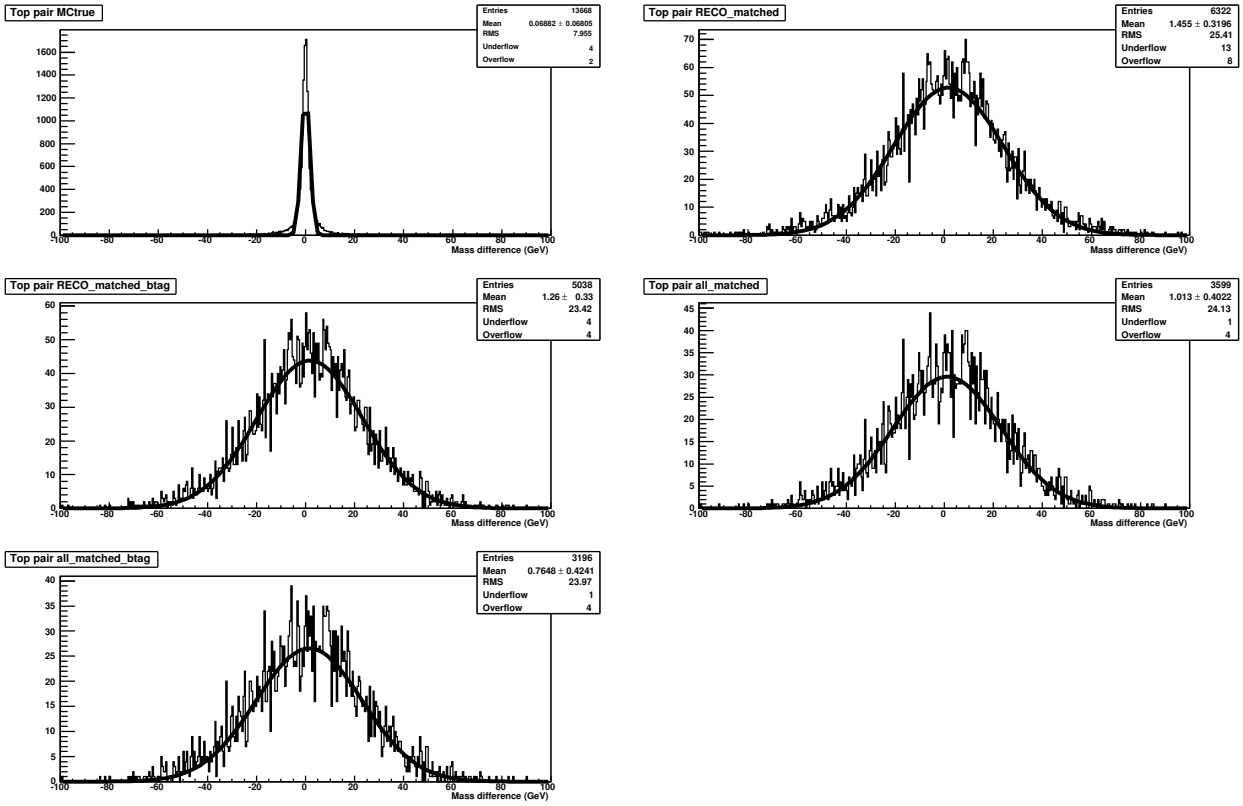

Figure 5.2. bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

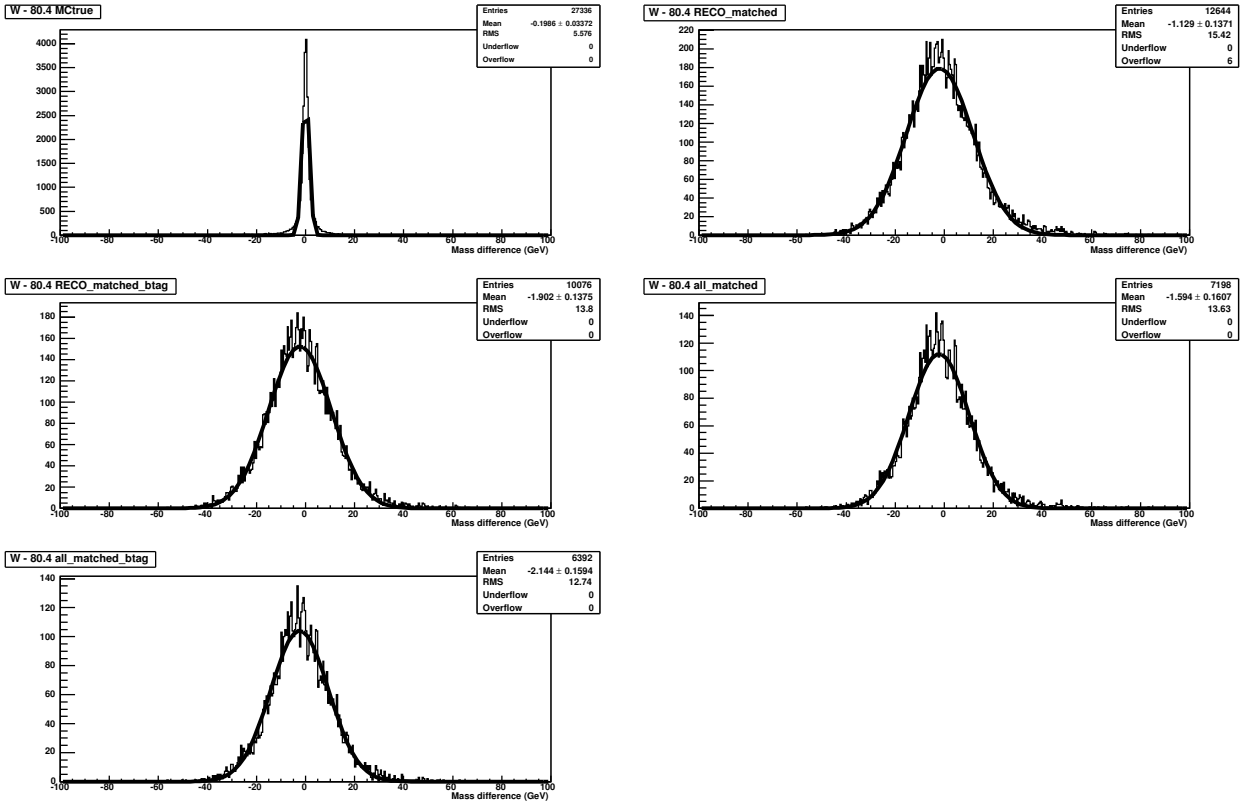

Figure 5.3. jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

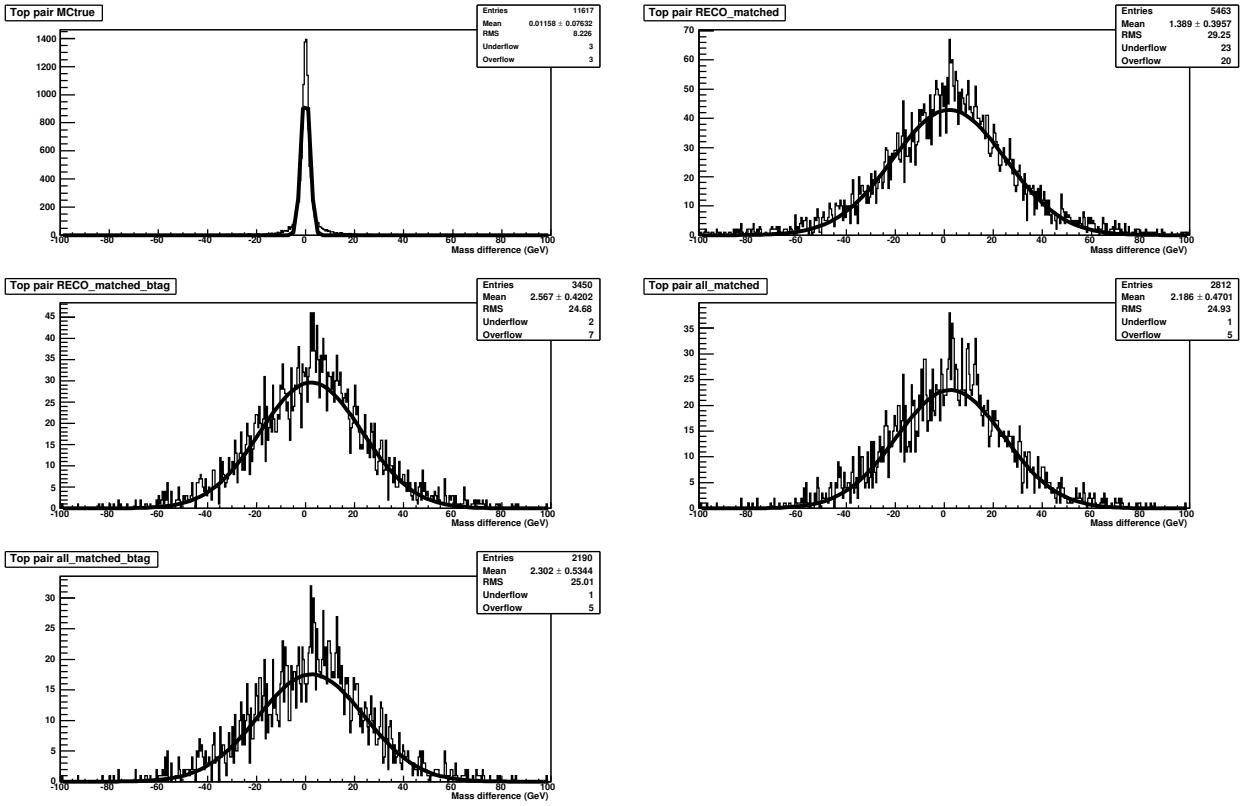

Figure 5.4. bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

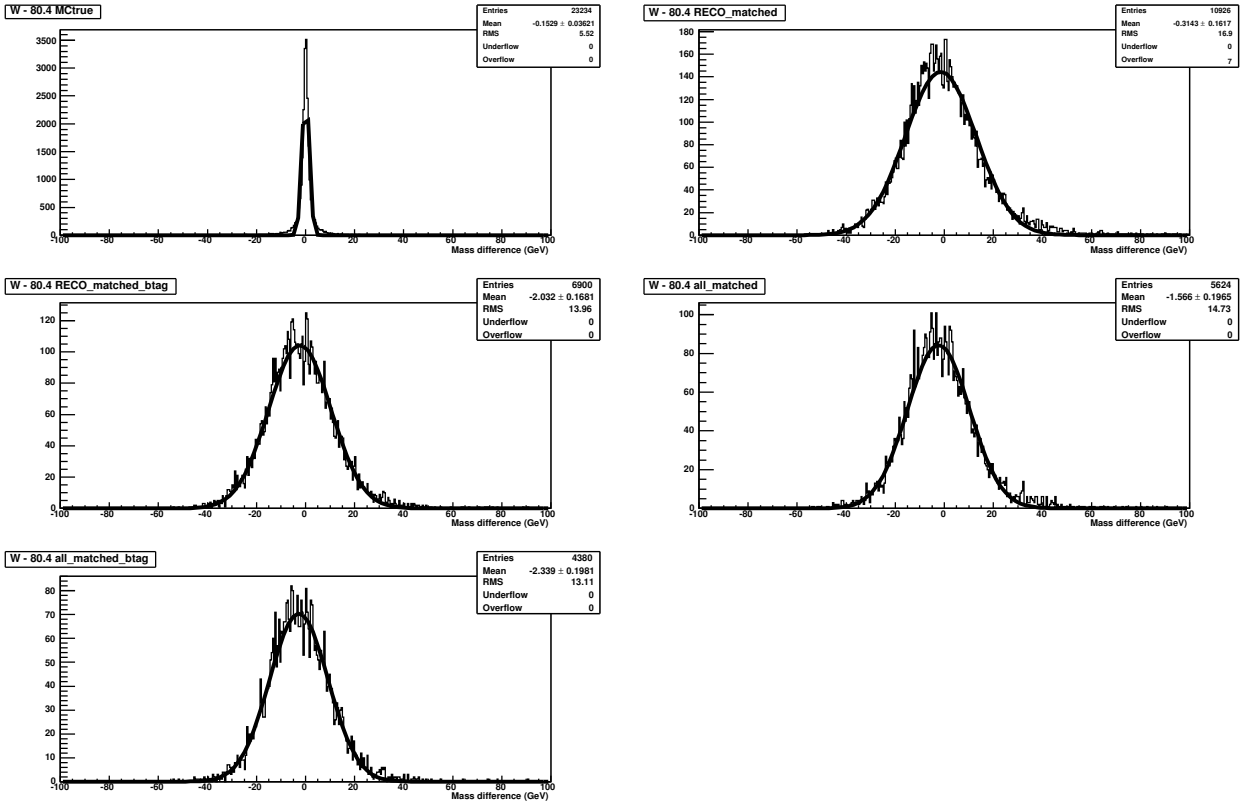

Figure 5.5. jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

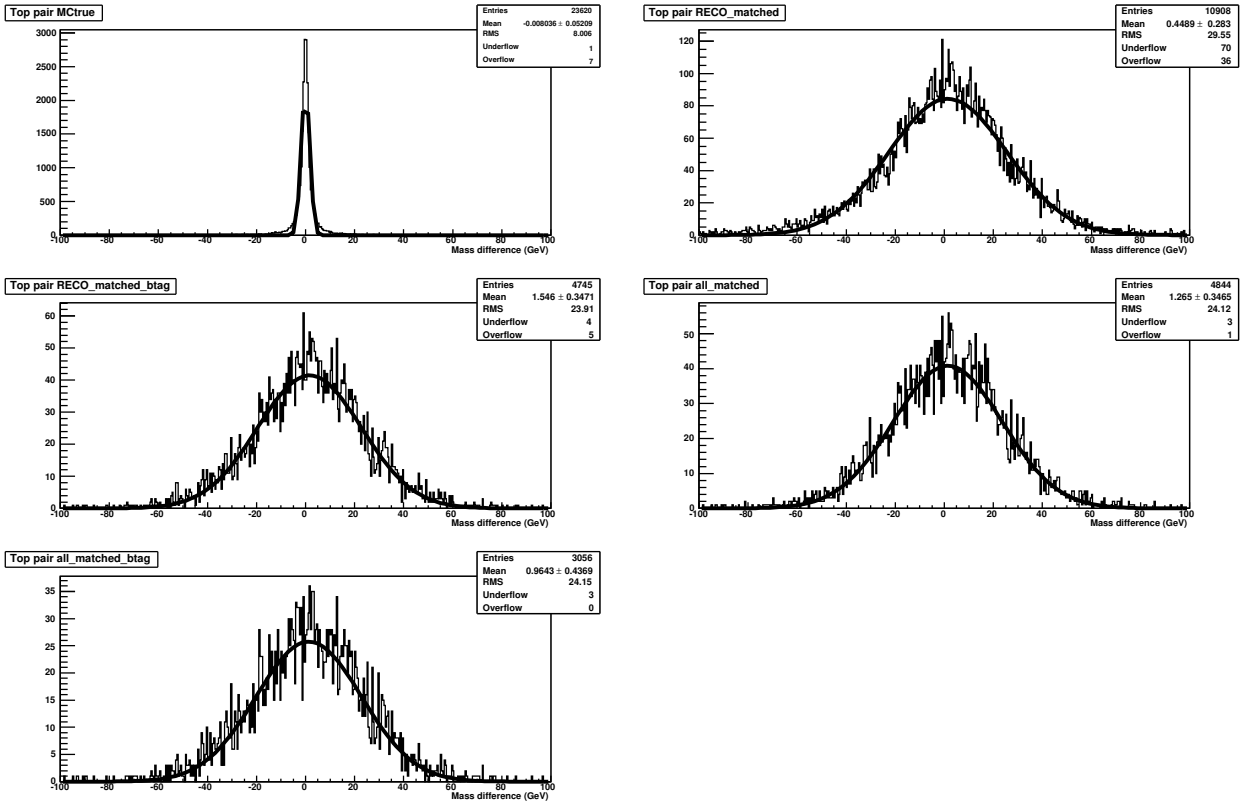

Figure 5.6. bjj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

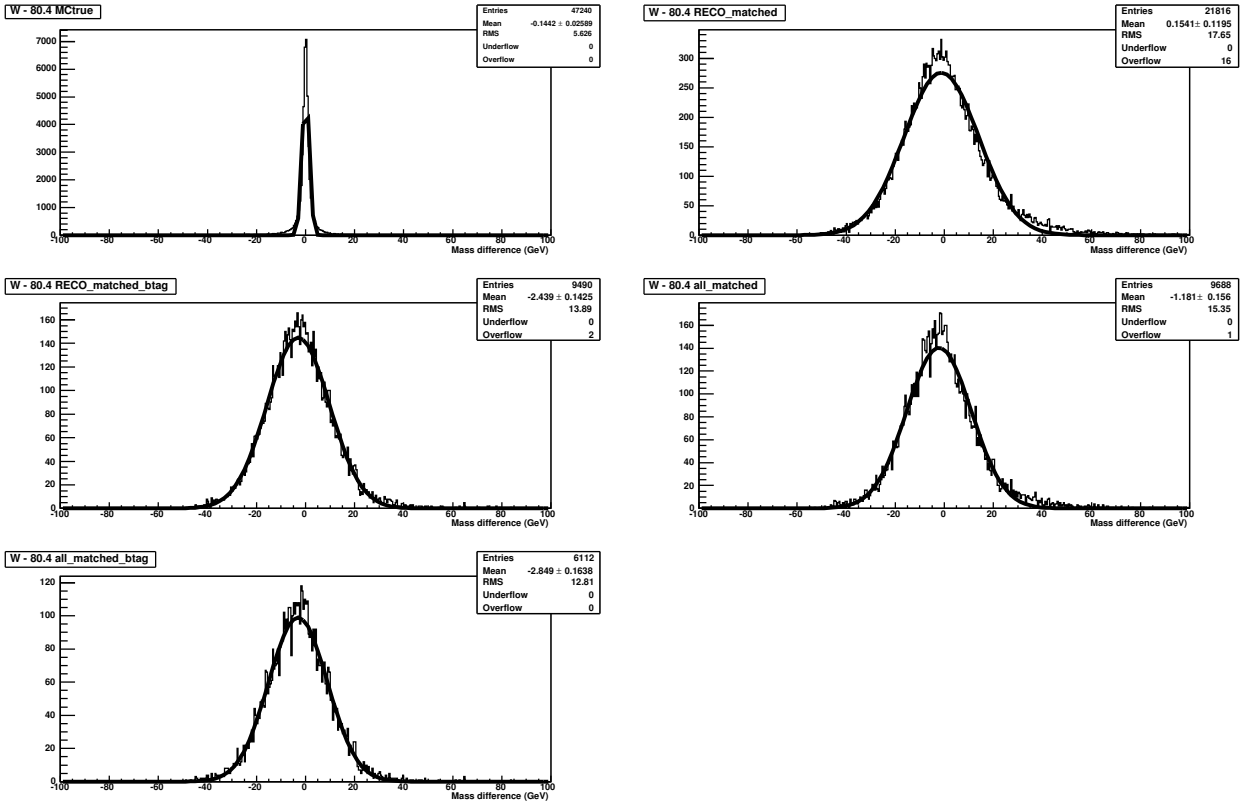

Figure 5.7. jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

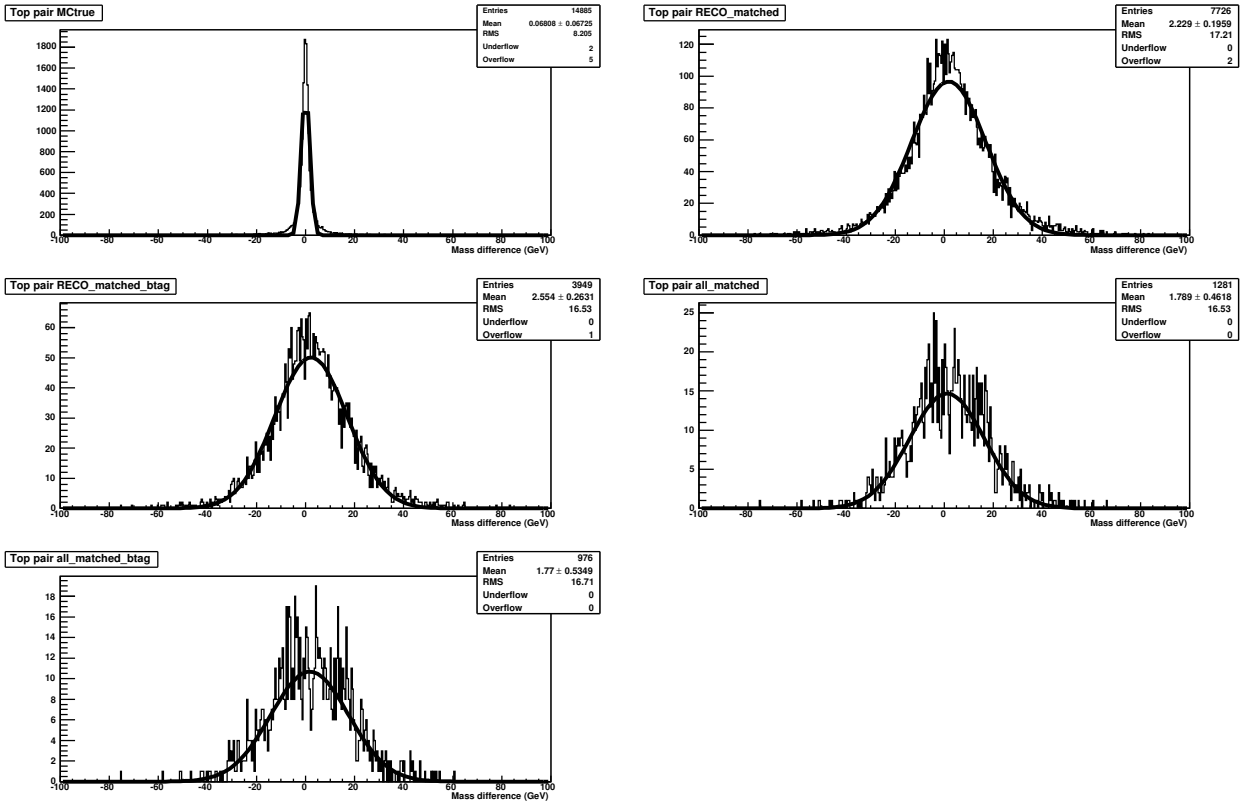

Figure 5.8. bjj mass difference between the 2 top candidates from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

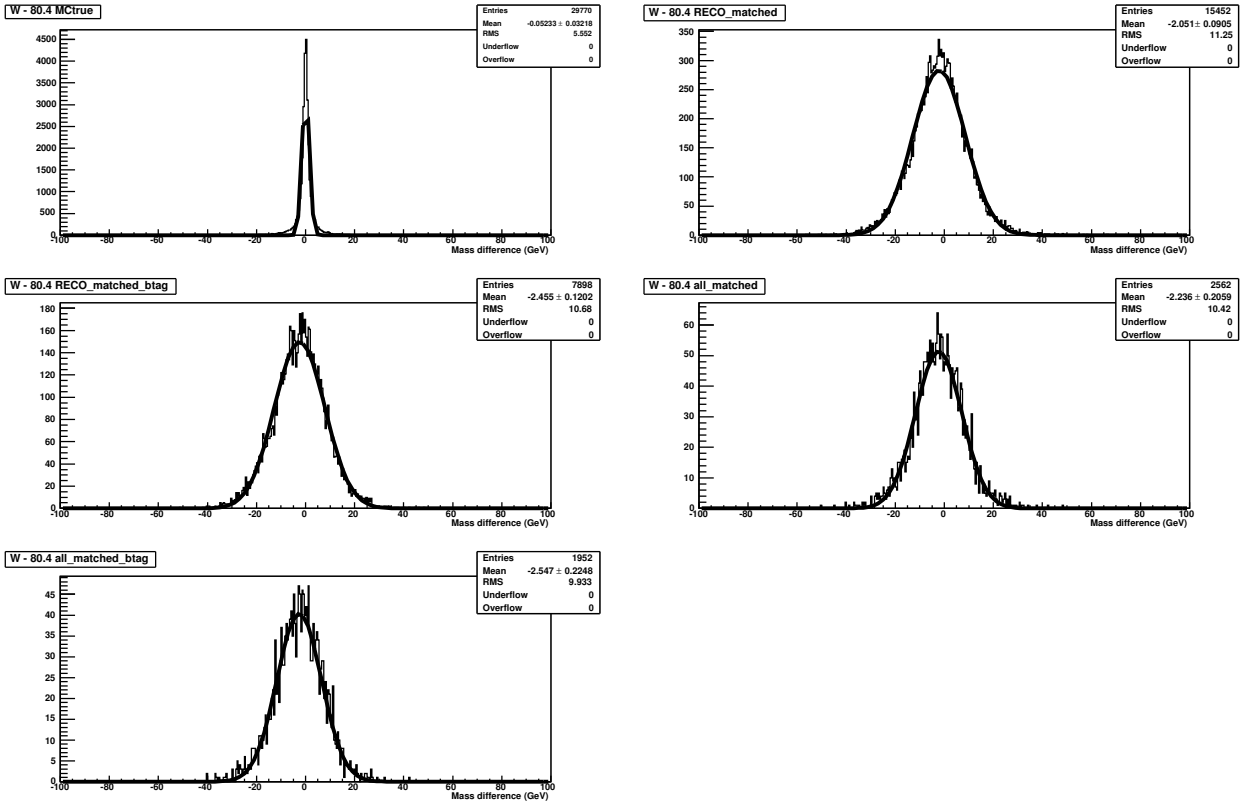

Figure 5.9. jj mass difference between the 2 top candidates from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top left plot shows the mass difference using MC partons. The rest show the mass difference using reconstructed RECO jets, where the top left is all RECO jets have a matching MC parton from top decay, and middle left is the same with the b-tagged jets matched to the b-quarks. The middle right is the RECO jets triplets are matched correctly to the top quark triplets and the bottom left is the same with the correctly matched b-tagged jets

the cut.

For light quarks, the second consideration comes in: the JCCB jets used in this analysis have a cone size of 0.5. There is no overlap between the 2 RECO jet cones if their angular separation is more than 0.5. The same requirement of $dR < 0.5$ is also applied for lighter quark as well. The losses due to the longer tail in the dR distribution for light quarks don't have a major impact in this analysis.

Given the matching definition, now we can see how often the jet triplets from the best chi square combination correctly match the 3 quarks from the top decay, considering only those events where all 6 quarks from the top decay have a matching RECO jet. Comparing the best chi square triplet with the matched jet triples can yield 4 different outcomes:

bjj correct, jj correct: b-tagged jet and the 2 other jets matched correctly to the b-quark and lighter quarks from the W decay.

bjj correct, jj wrong: The 3 jets to the top are matched correctly but the 2 RECO jets (either one or both) don't match lighter quarks from the W decay.

bjj wrong, jj correct: b-tagged jet is not matched correctly to the b-quark, but the other jets match lighter quarks from the W decay.

bjj, jj wrong: The 3 jets in triplet don't match to the top decay quarks and the 2 lighter jets don't match the lighter quarks from the W decay.

Table 5.2 shows the percentage of the above 4 cases for the jet triplets in MC ttbar $m_t = 170$ GeV events (each events have 2 top triplets), together with those

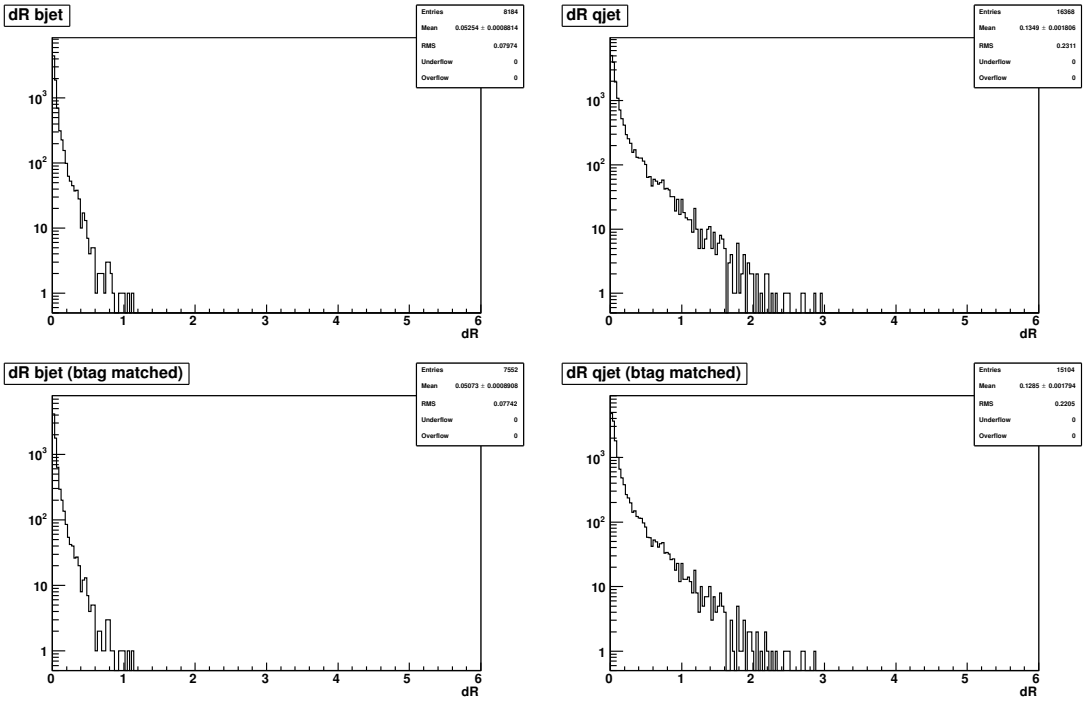

Figure 5.10. Angular Separation (dR) between the MC parton and the closest RECO jet using MC ttbar samples. Top left shows dR between MC b quarks and the closest RECO jet, and the bottom left shows the same plot with additional requirement that the RECO jet is also b-tagged. Top Right shows dR between lighter quarks (from W decay) and the closest RECO jet and the bottom right shows the same plot with additional requirement that the b-tagged jets in the same events are matched correctly to the b-quarks.

BID NN regions	bjj corr. jj corr.	bjj corr. jj wrong	bjj wrong jj corr.	bjj wrong jj wrong	not all matched
6j, $NN > 0.65$	28.1%	2.52%	10.2%	38.3%	20.8%
6j, $0.65 > NN > 0.2$	19.9%	5.42%	9.41%	43.8%	21.5%
7+j, $NN > 0.2$	14.0%	2.07%	7.86%	55.5%	20.6%

TABLE 5.2

Percentage of the best chi-square jet triplets having the correct/wrong/not-matched jet combinations

events that don't have all 6 quarks have matching jets. Notice that the results are separated into different B-ID Neural Networks output cuts to illustrate the dependence on the BID NN values as well.

From the table it can be inferred that b-tagged jets with higher BID NN have a higher chance of getting the correct jet combinations. This is expected as higher BID NN corresponds to higher probability of getting a real b-jet and less likely originating from the lighter quarks from the W, therefore less likely to be mismatched. Also higher jet multiplicities in the 7+ jet case decrease the chances of getting the correct combinations as the extra jets add confusion to the jet selection.

5.2.3 Mass distribution

It would be interesting to see how the correct/wrong jet combinations changes the reconstructed top/W mass. The three jets (bjj) invariant mass corresponds

to the top mass and the 2 not b-tagged jets (jj) correspond to the W mass. Since the bjj is mass heavily dependent on the jj mass, it would interesting to remove the correlation with the reconstructed W mass by using the quantity Dmass(bjj) instead:

$$Dmass(bjj) = mass(bjj) - mass(jj) + 80.423 \text{ GeV} \tag{5.5}$$

Such a quantity replaces the reconstructed W mass contribution with the current standard W mass measurement. As can be seen in Figure 5.12, the correlation between Dmass(bjj) and mass(jj) is significantly less compared to Figure 5.11, and results in a narrower mass peak in $Dmass(bjj)$.

Figures 5.13, 5.14, 5.15 and 5.16 show the comparison of $Dmass(bjj)$ between parton matched/not matched events, and between different correct/wrong jet combinations in 4 different BIDNN regions. From the plots it can be observed that the events with all partons matched yield a narrower peak than the events without partons matched, as expected. The mass peaks between the two cases are not shifted away too much from either the case including all events, or the case where the correct jet combinations are always found. Since our mass measurement will depend on the $Dmass(bjj)$ distribution, this shows that the mass bias due to parton matching will be small.

When looking at the the correct/wrong jet combination mass distributions, the mass peaks between the 4 cases have noticeable differences. As expected, the correct bjj and jj combinations give the best mass reconstruction and therefore the narrowest peak. The correct bjj and wrong jj combination case shifts considerably to lower mass, and the wrong bjj and correct jj case shifts toward higher mass. We also observe that the correct bjj and jj combinations case match very well with

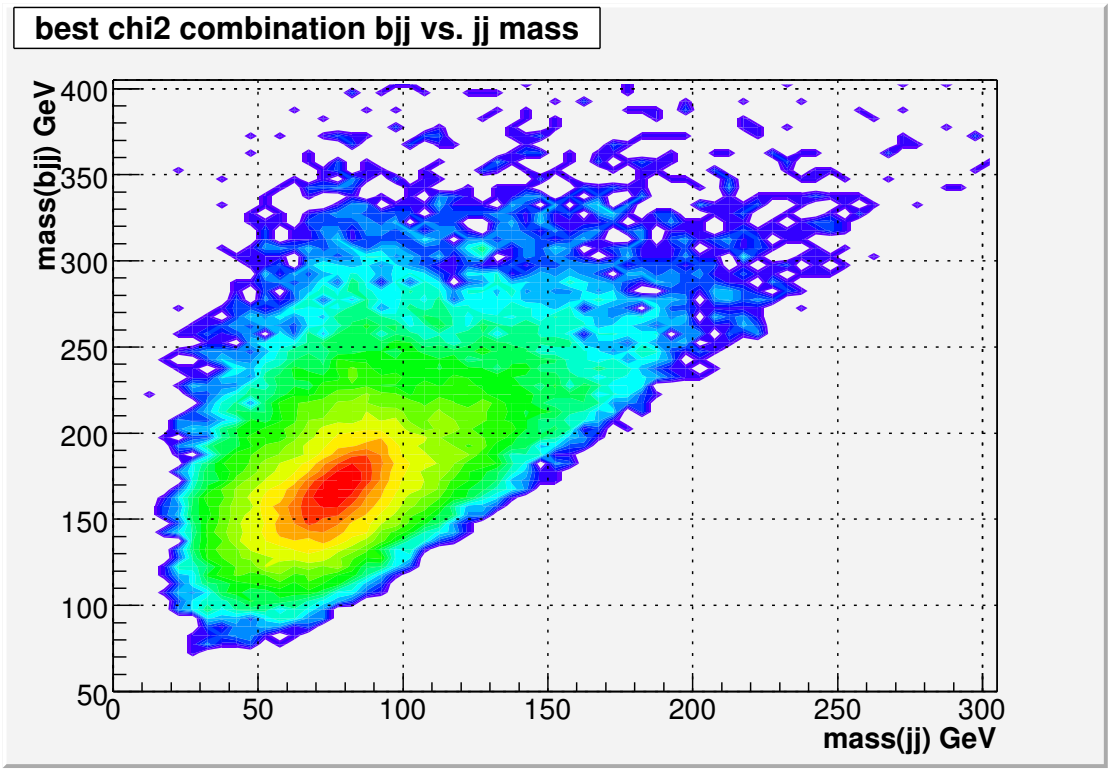

Figure 5.11. Reconstructed bjj mass vs jj mass distribution from best chi2 jet combination in MC $mass_{top} = 170$ GeV signal sample

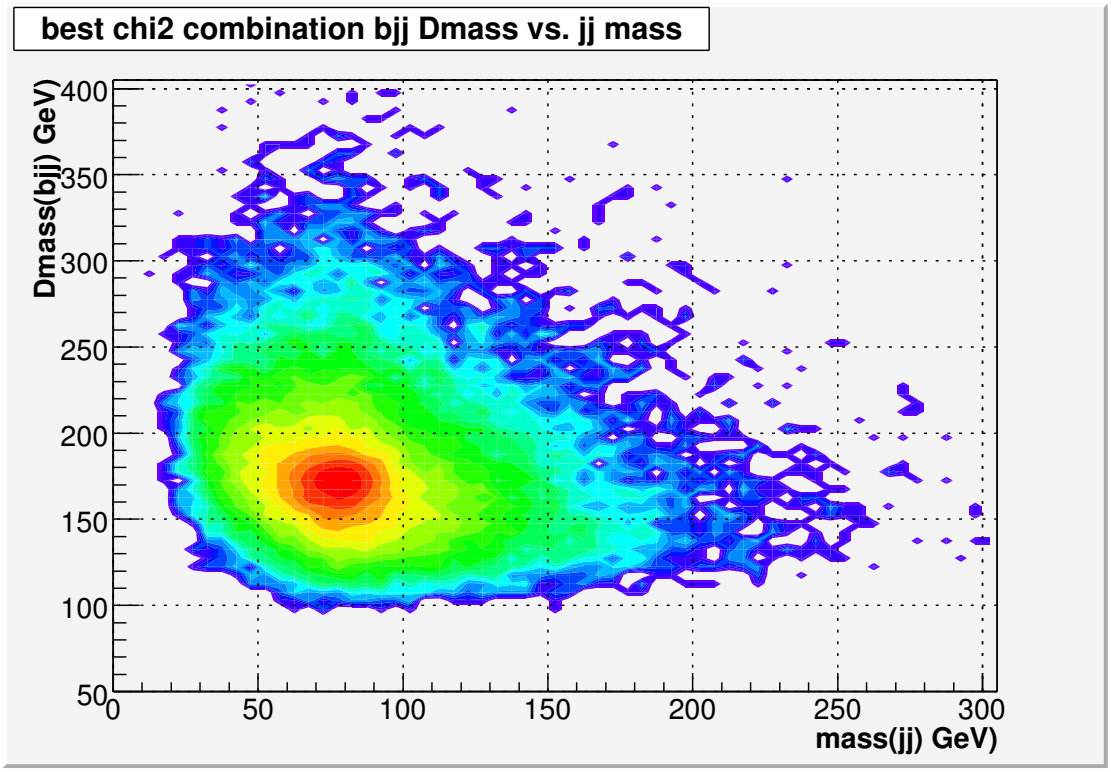

Figure 5.12. Reconstructed bjj Dmass vs jj mass distribution from best chi2 jet combination in MC $mass_{top} = 170$ GeV signal sample

the ideal case when the correct combination can be found every time. This shows that the correct combination can be used as a separate template versus everything else (parton matched wrong combinations plus not-parton-matched events), as the correct bjj and jj combinations gives the correct mass peak and width.

Similarly, figure 5.17, 5.18, 5.19 and 5.20 shows the comparison of $mass(bjj)$ in 4 different BIDNN regions. It can be seen that the $mass(bjj)$ distributions are smeared out compared to the $Dmass(bjj)$ distributions in all cases. This is expected as the inclusion of $mass(jj)$ adds extra uncertainties. Therefore, $Dmass(bjj)$ is preferred over $mass(bjj)$ as the estimator for the top mass measurement.

Figures 5.21, 5.22, 5.23 and 5.24 show the comparison of $mass(jj)$ in 4 different BIDNN regions. Again the mass bias between the parton matched and not-parton-matched events is small. The cases with wrong jj combinations result in a significant shift of the mass peak compared to the correct jj combination case.

5.3 Template fitting

5.3.1 Signal template

From the previous section it was shown that $Dmass(bjj)$ is a better choice in measuring the top mass and will be used as a signal template when looking at the data candidates. However, the presence of statistical fluctuations in the histogram will translate to fluctuations in top mass since it is used to model the mass peak. By replacing the template with analytic functions one can reduce the statistical fluctuations. The resulting fit can also be parametrized as a continuous function of input top mass as well, which gives a better estimate of the measured mass and error compared to the histograms of discrete MC top mass (as will be

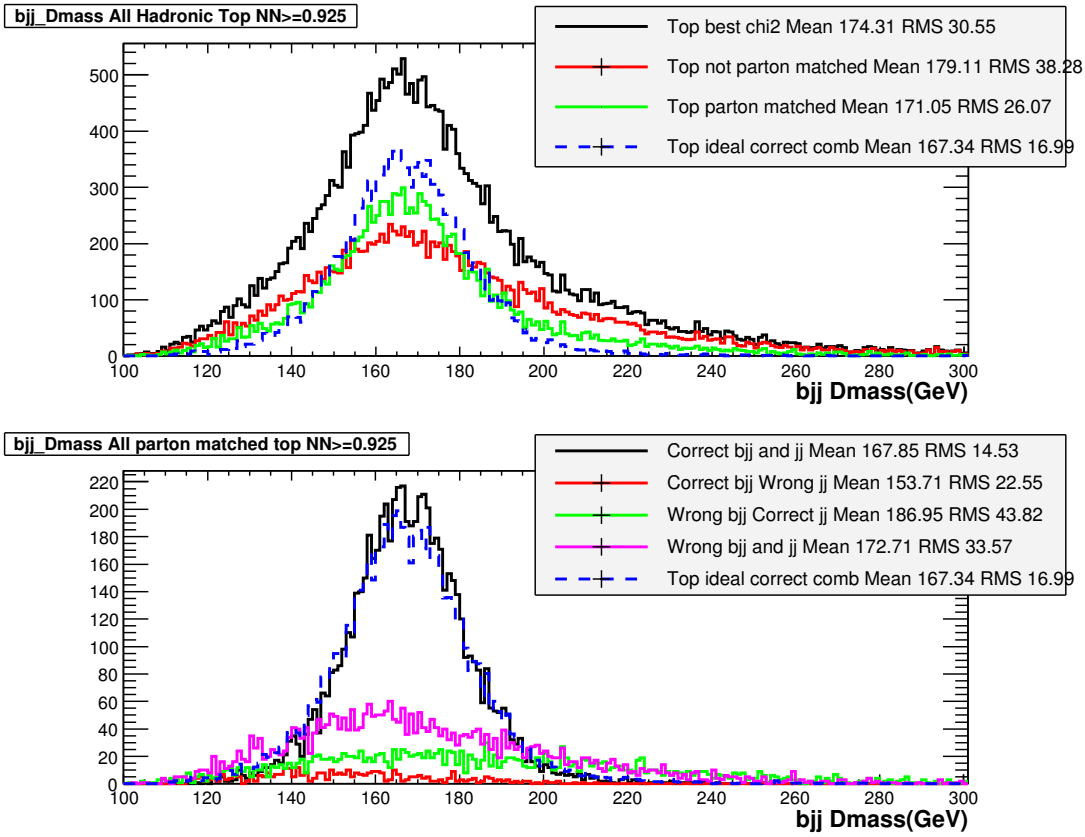

Figure 5.13. bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

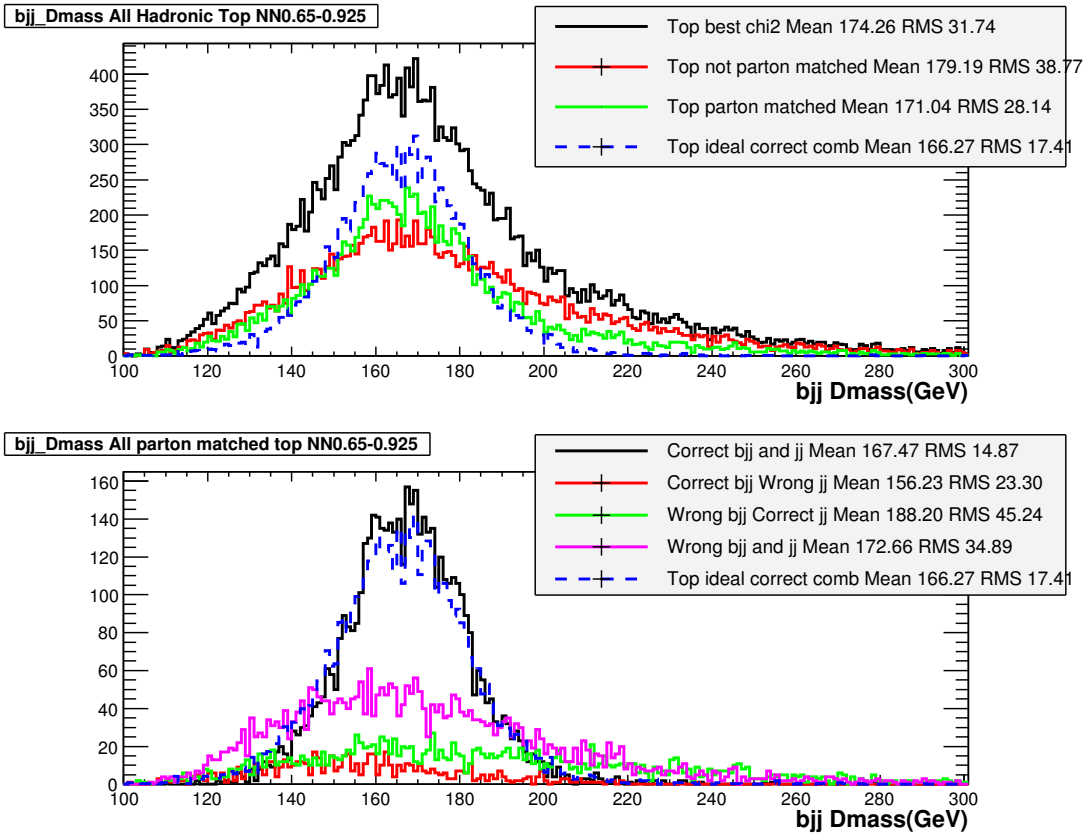

Figure 5.14. bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

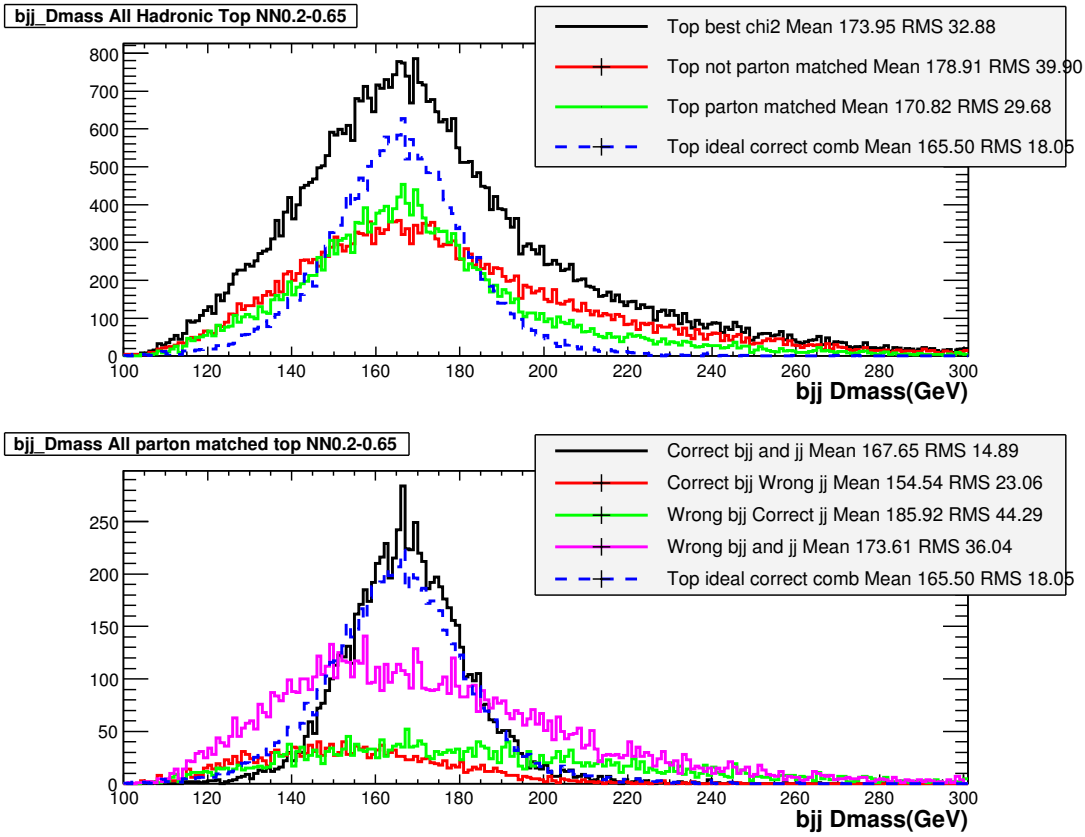

Figure 5.15. bjj Dmass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

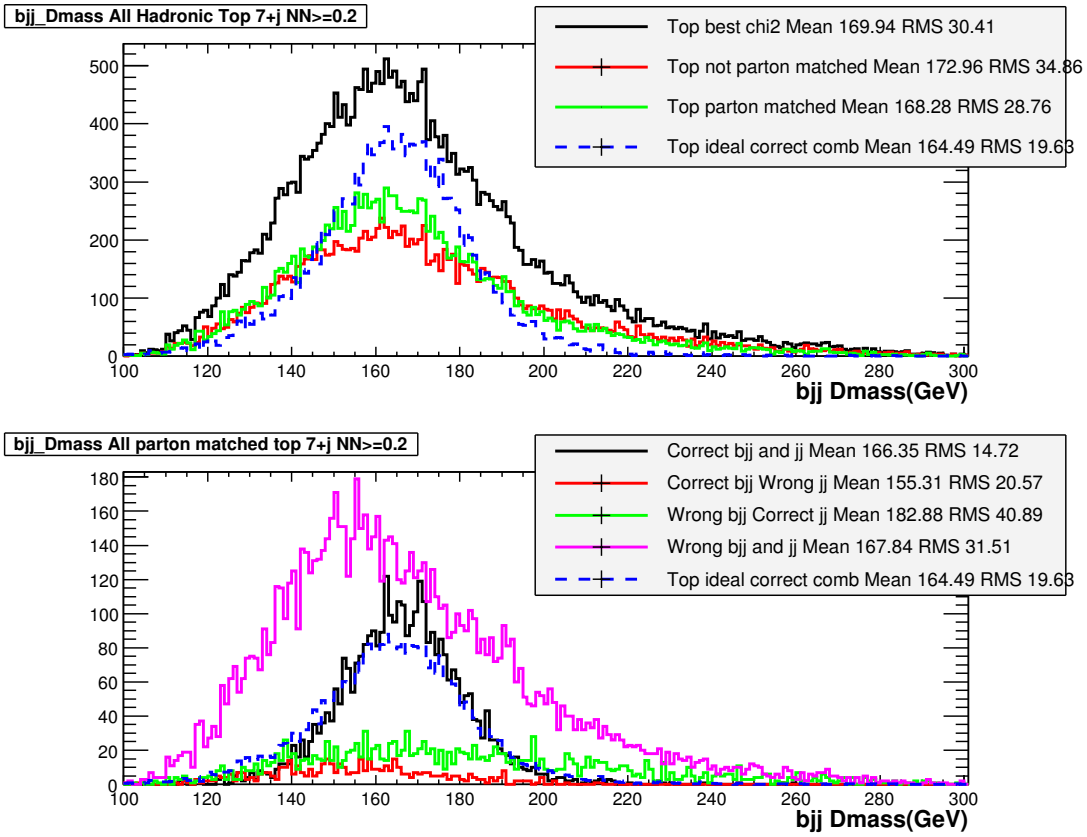

Figure 5.16. bjj Dmass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

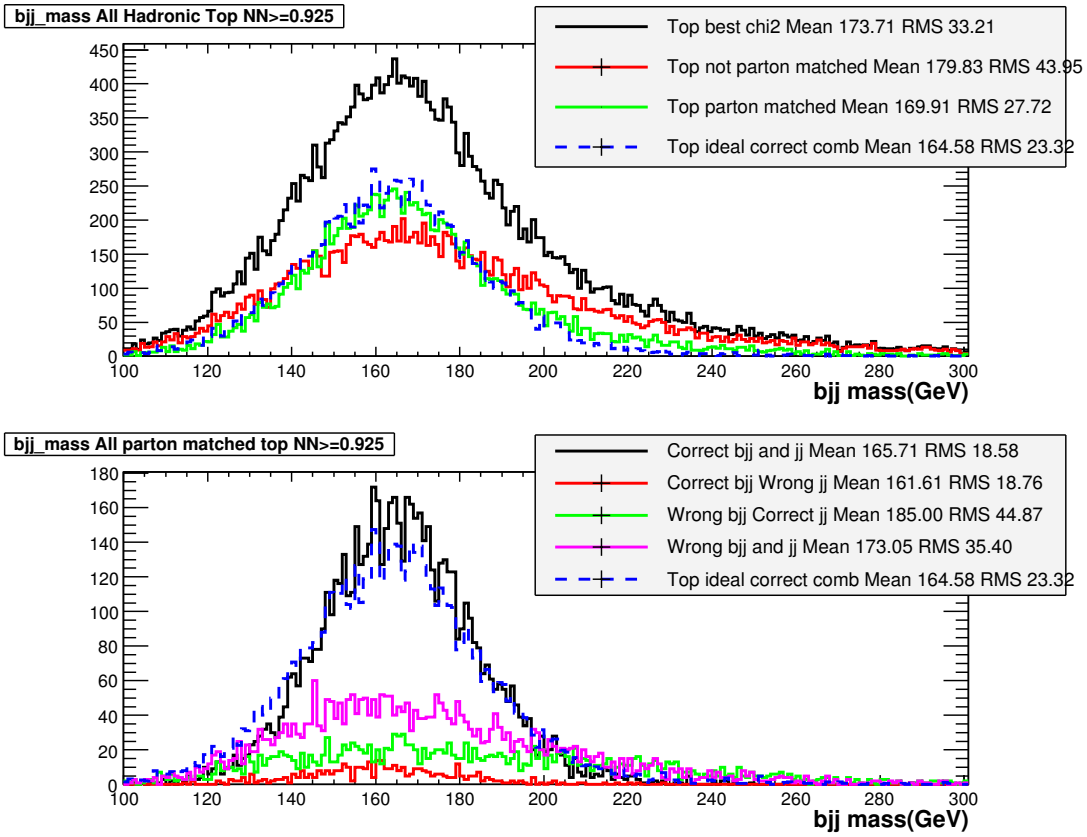

Figure 5.17. bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

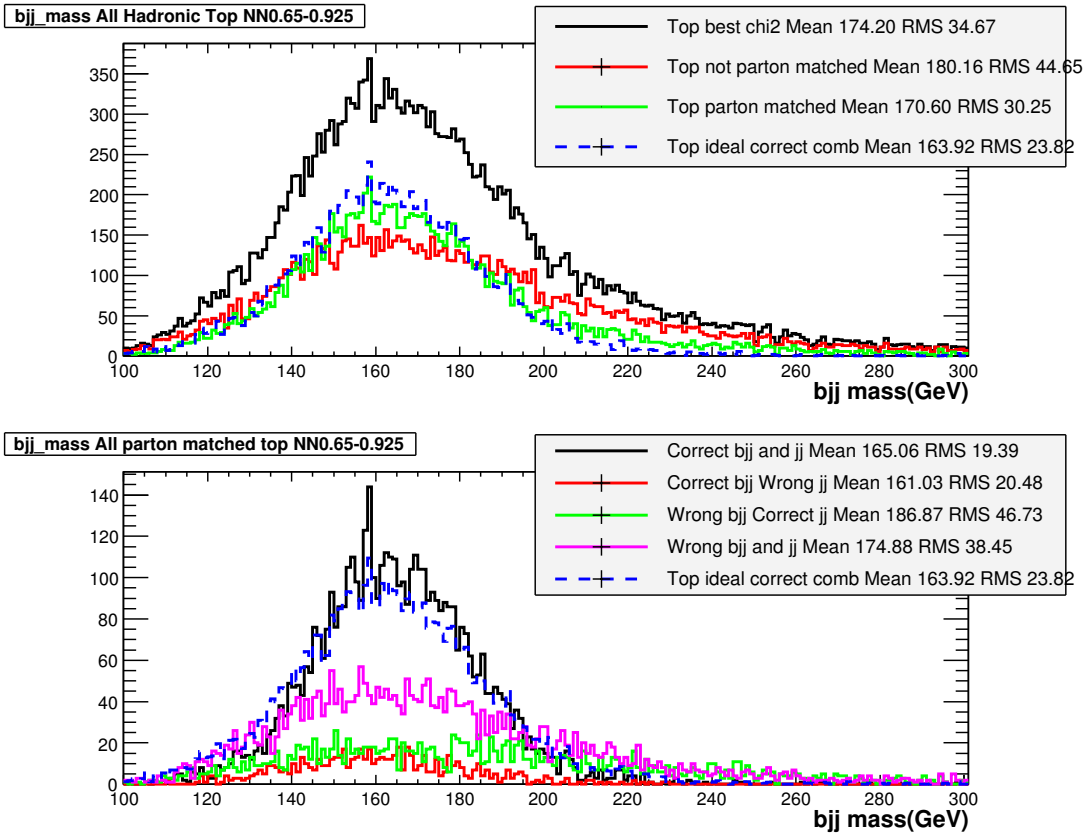

Figure 5.18. bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

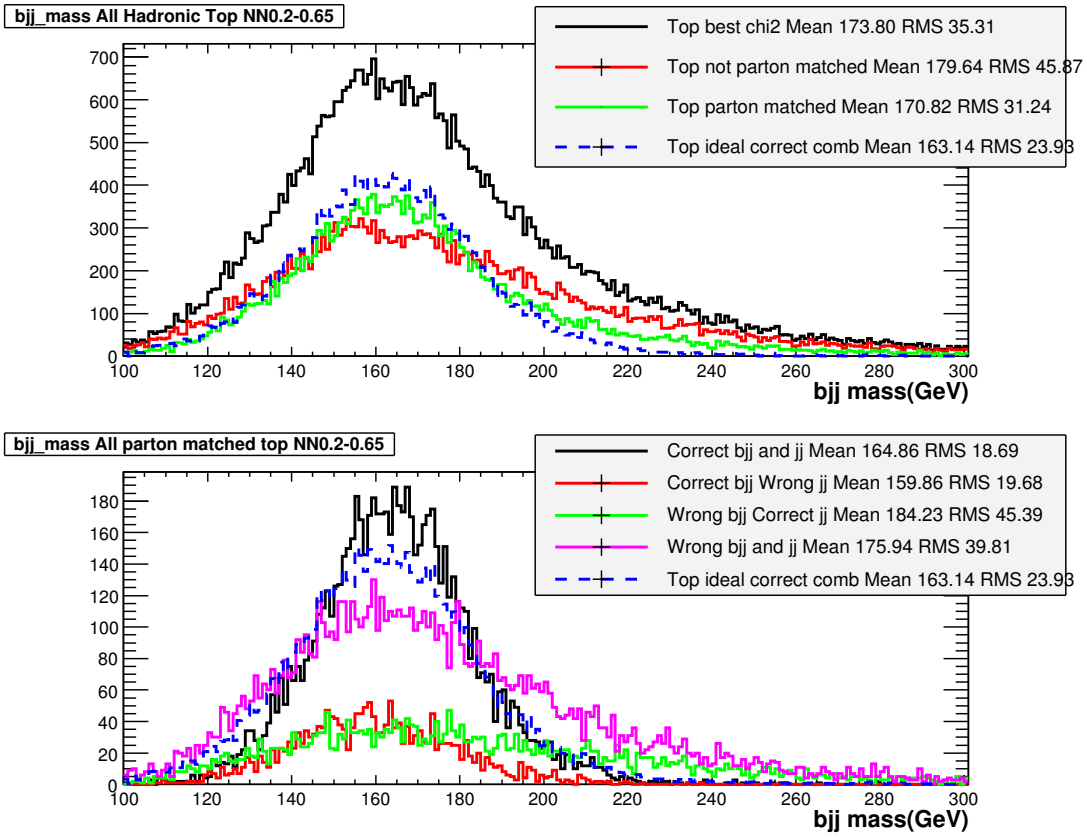

Figure 5.19. bjj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

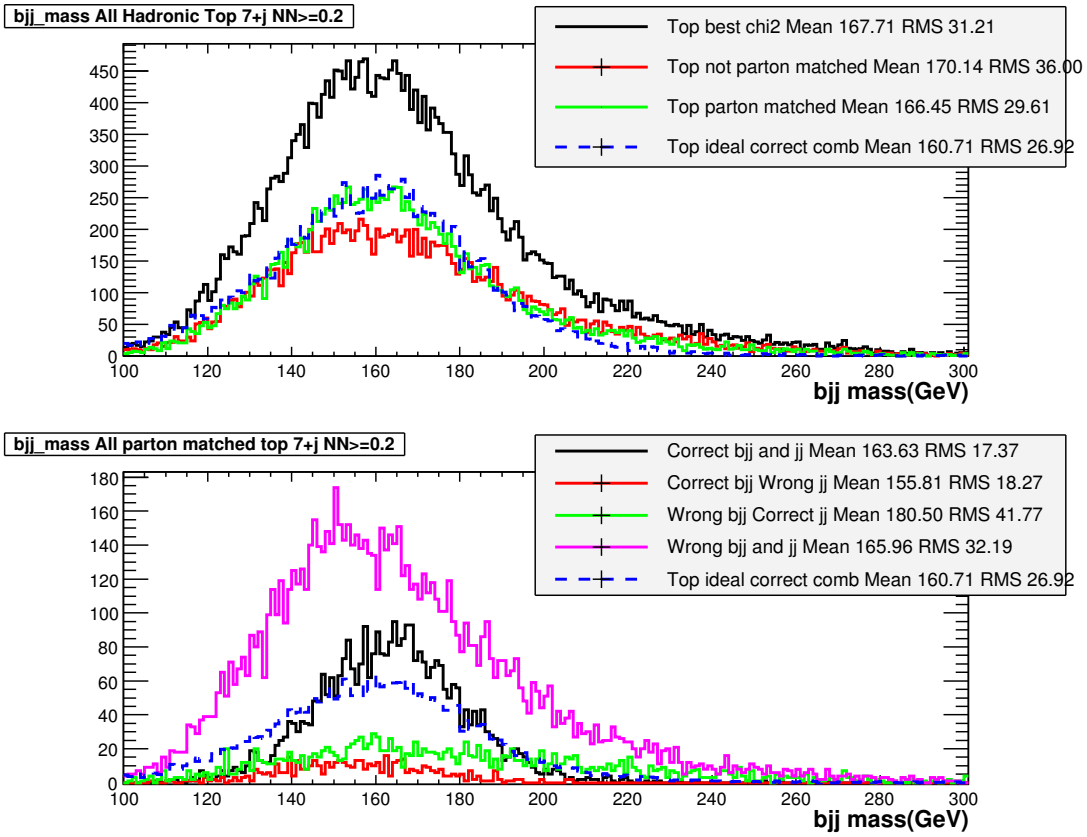

Figure 5.20. bjj mass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

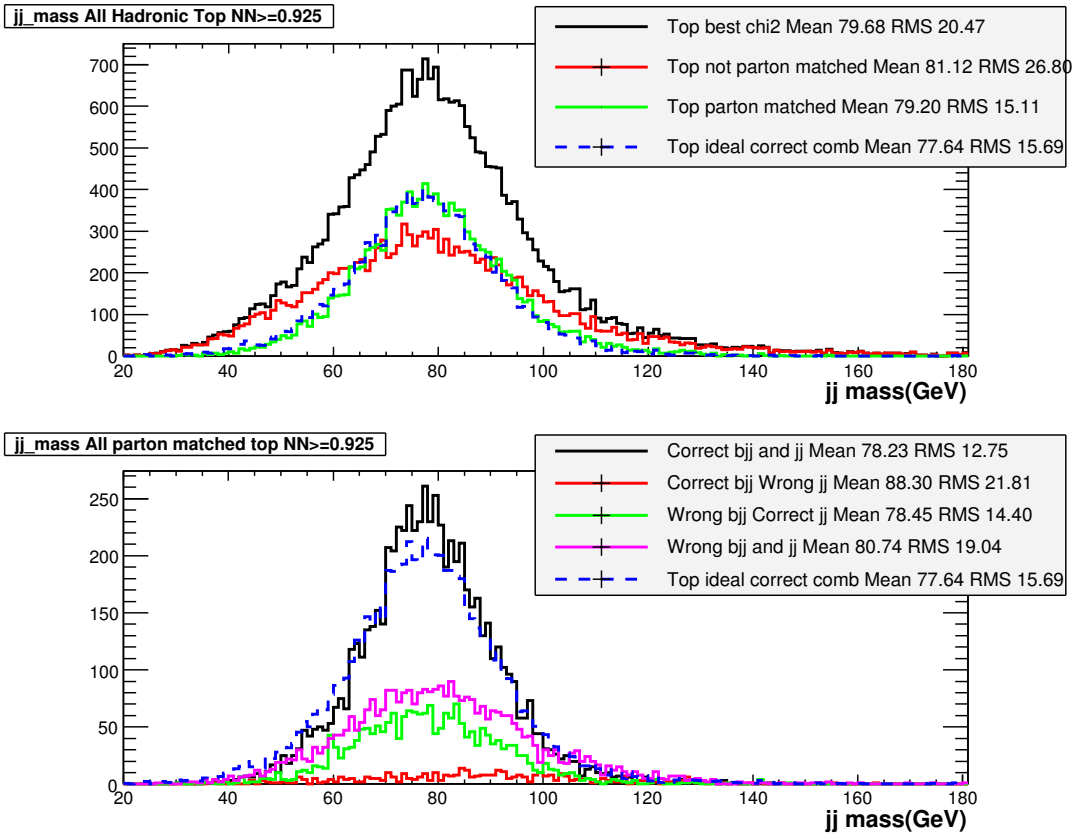

Figure 5.21. jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN > 0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

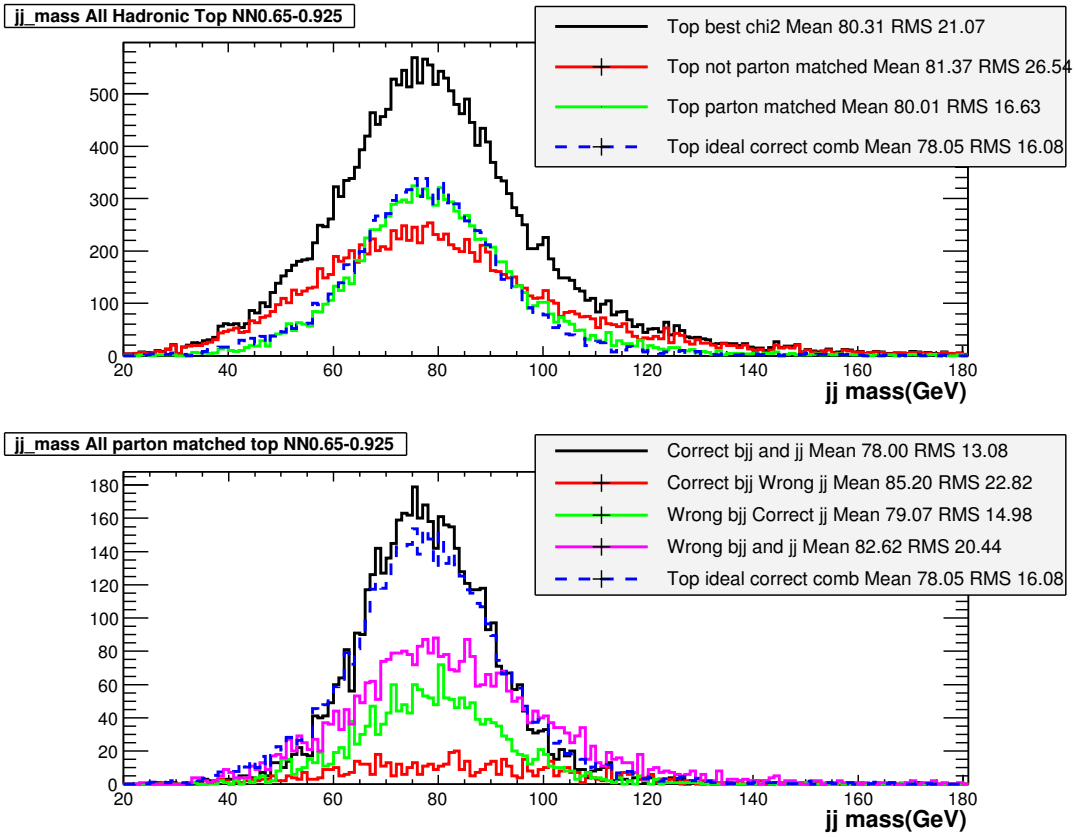

Figure 5.22. jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.65-0.925. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

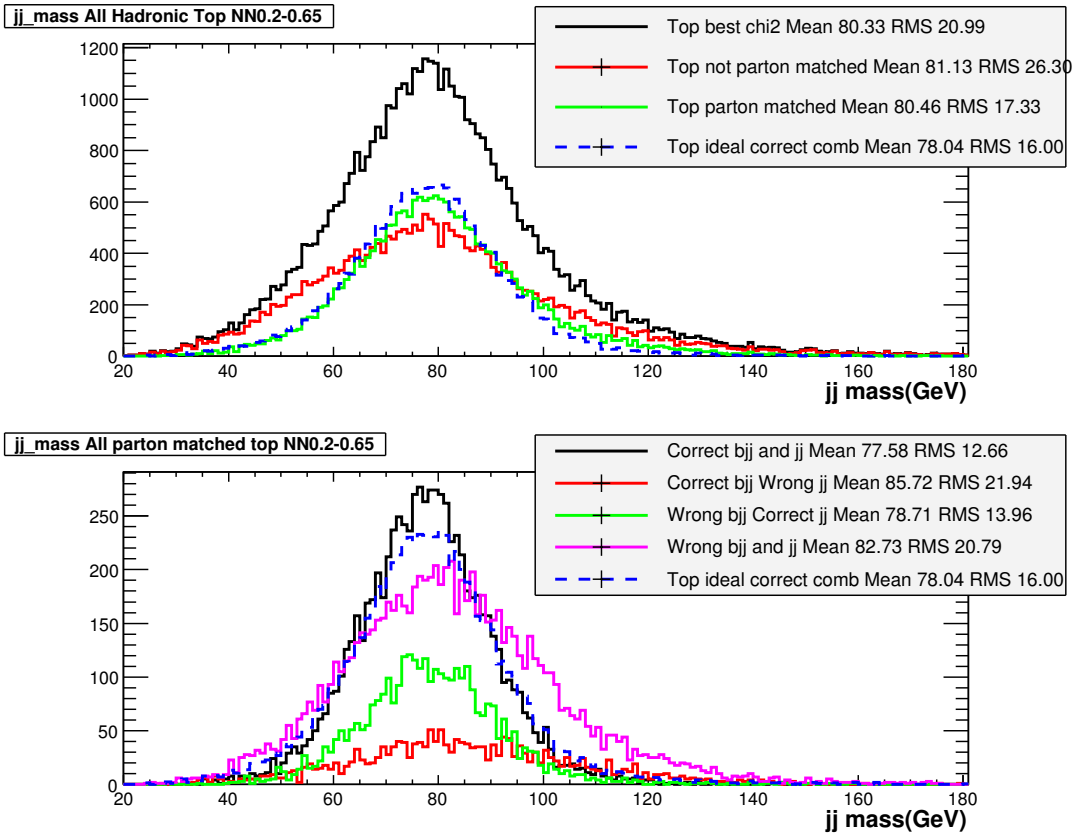

Figure 5.23. jj mass from the best chi2 jet combination in exactly 6 jets, 2 BIDNN 0.2-0.65. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

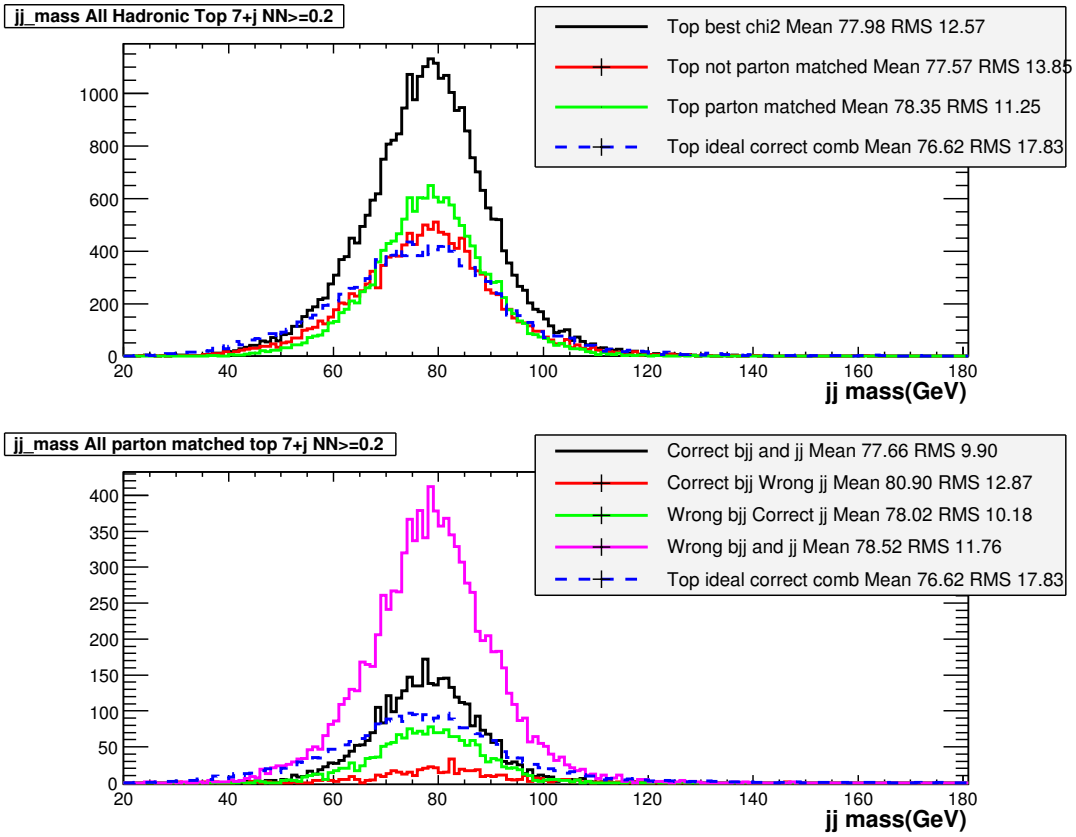

Figure 5.24. jj mass from the best chi2 jet combination in 7+ jets, 2 BIDNN > 0.2. The top plot compares between the events having RECO jets matched to top decay partons and those without, using MC ttbar all hadronic events. The bottom compares between correct/wrong jet combinations in the case where all partons have matching RECO jets. The blue dotted histogram shows the ideal case if the correct jet combinations is found in every event (it is normalized to parton-matched events in the top plot, and normalized to bjj and jj correct case in the bottom plot).

shown in the section of ensemble tests).

First the MC signal of correct combinations can be nicely fitted using a simple Gaussian distribution:

$$fit_{\text{signal,correct bjj/jj}}(Dmass(bjj); p_0...p_2) = \frac{p_0}{N} * exp\left(-\frac{(Dmass(bjj) - p_1)^2}{2 * p_2^2}\right) \tag{5.6}$$

where each parameter: $p_i(m_{top}) = a_i + b_i * m_{top}$ is a function of input top mass. Such second parameterizations are obtained by fitting the fitted parameters of the templates in MC ttbar samples. Figure 5.25, 5.26 and 5.27 shows the linear fits of the parameters extracted from correct bjj/jj jet combinations in different BID NN regions.

Using this second parameterization of parameters, the resulting template is then a function of $Dmass(bjj)$ and top mass. Figure 5.28, 5.29 and 5.30 shows the templates fit functions (lines) compared with the original template histograms (points) for correct bjj/jj combinations in different BID NN regions. We can see the template fits agree well with the histograms and have much smaller fluctuations.

The wrong jet combinations and not-parton-matched distributions have more complicated shapes. Investigations show that they can be nicely fitted by a log-normal function:

$$fit_{background}(Dmass(bjj); p_0...p_3) = p_0 * LogNormal(Dmass(bjj), p_1, p_2, p_3) \tag{5.7}$$

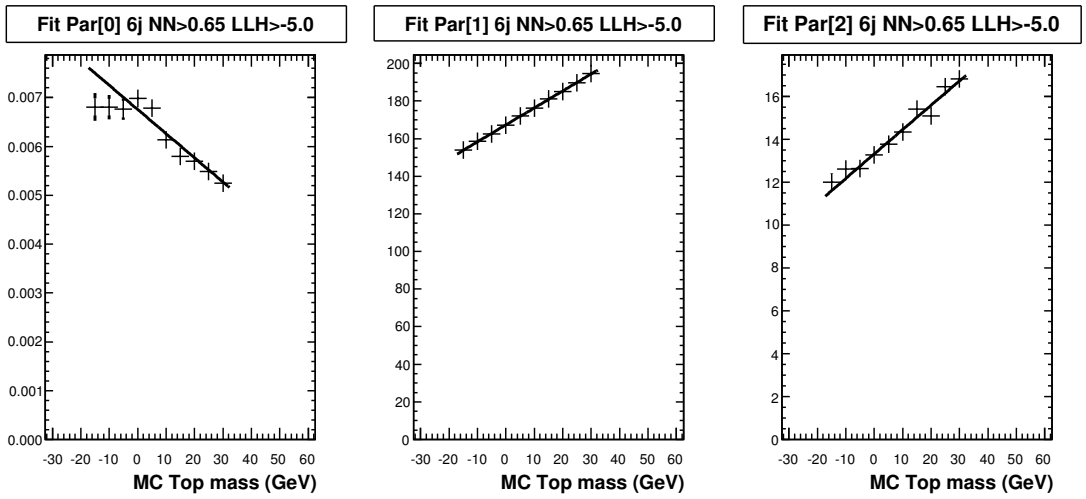

Figure 5.25. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65.

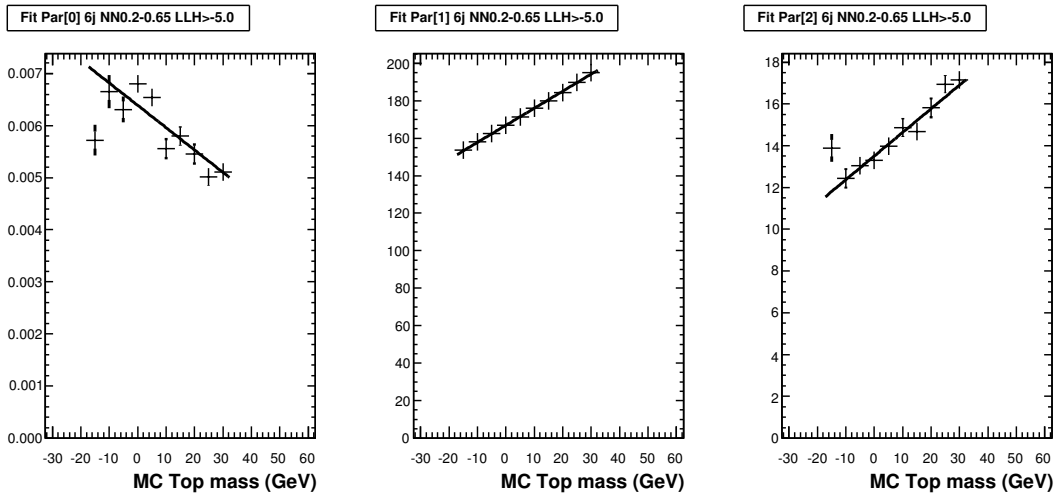

Figure 5.26. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65.

where the LogNormal function is defined as:

$$LogNormal(x, c_1, c_2, c_3) = \frac{e^{-(ln((x-c_2)/c_3))^2/(2c_1^2)}}{(x-c_2)c_1\sqrt{2\pi}} \tag{5.8}$$

Again each parameter p_i is a function of input top mass. Figure 5.31, 5.32 and 5.33 shows the linear fits of the parameters extracted from combining wrong jet combinations and not-parton-matched events in different BID NN regions.

Figure 5.34, 5.35 and 5.36 shows the templates fit functions (lines) compared with the original template histograms (points) for wrong jet combinations plus not-parton-matched events in different BID NN regions.

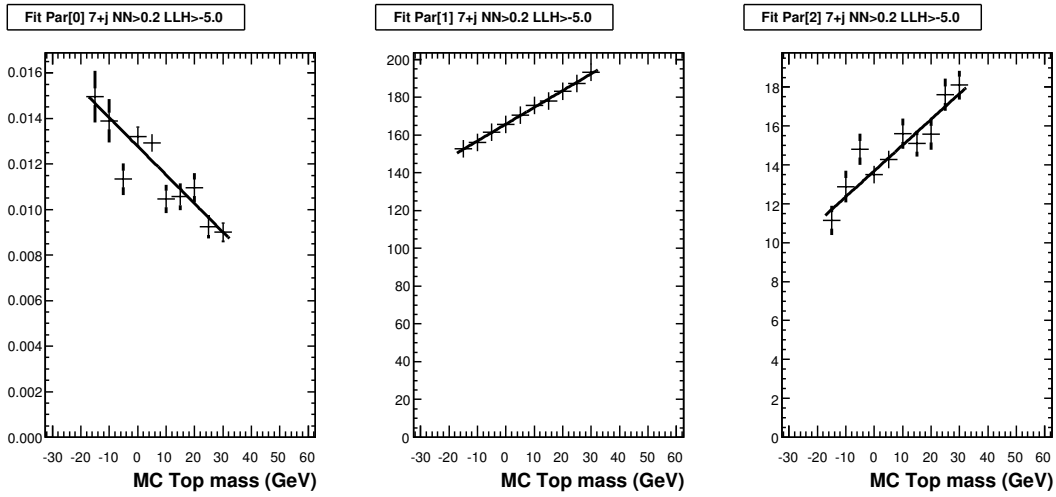

Figure 5.27. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2.

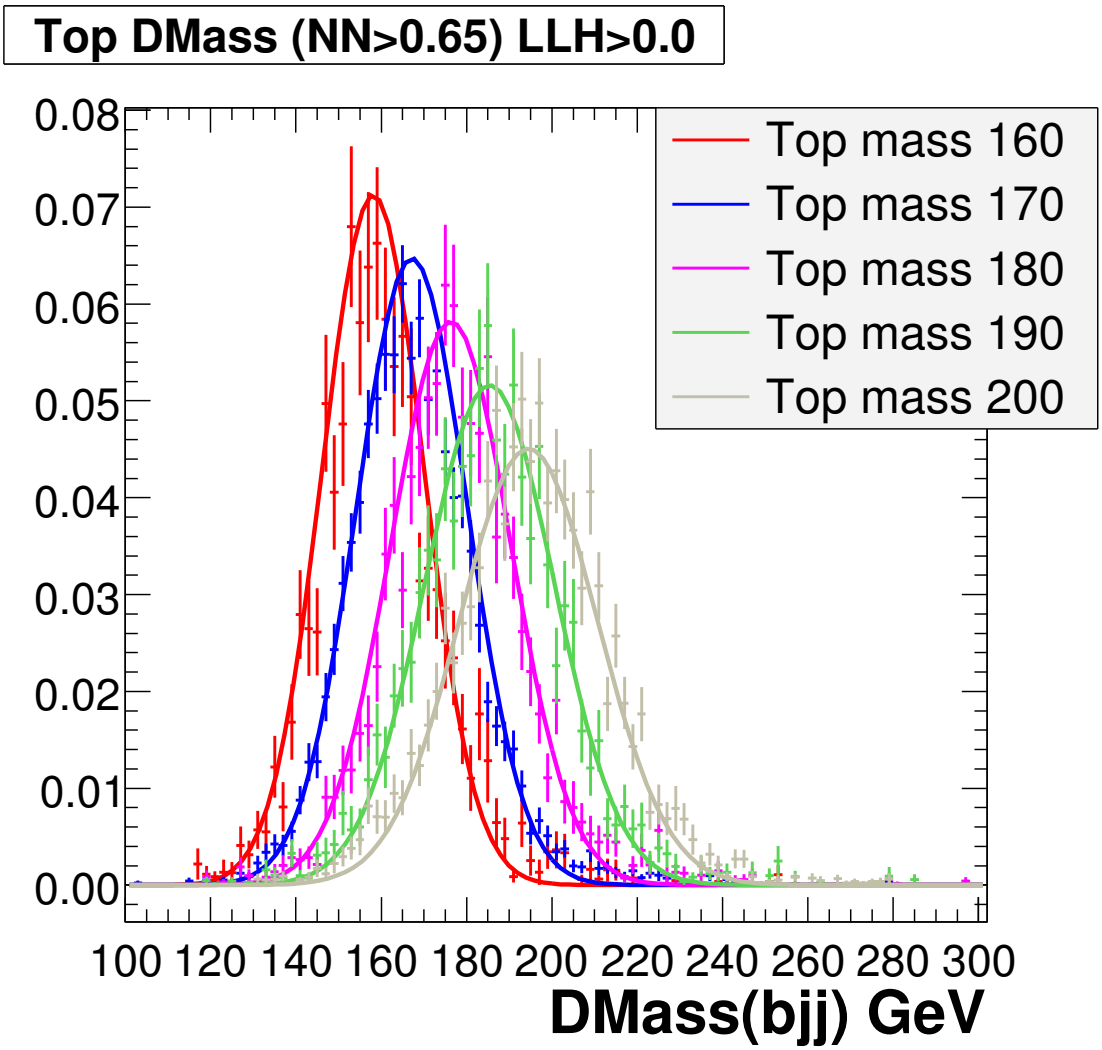

Figure 5.28. Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65. The line shows the function fit and the points are the original histograms from MC ttbar events.

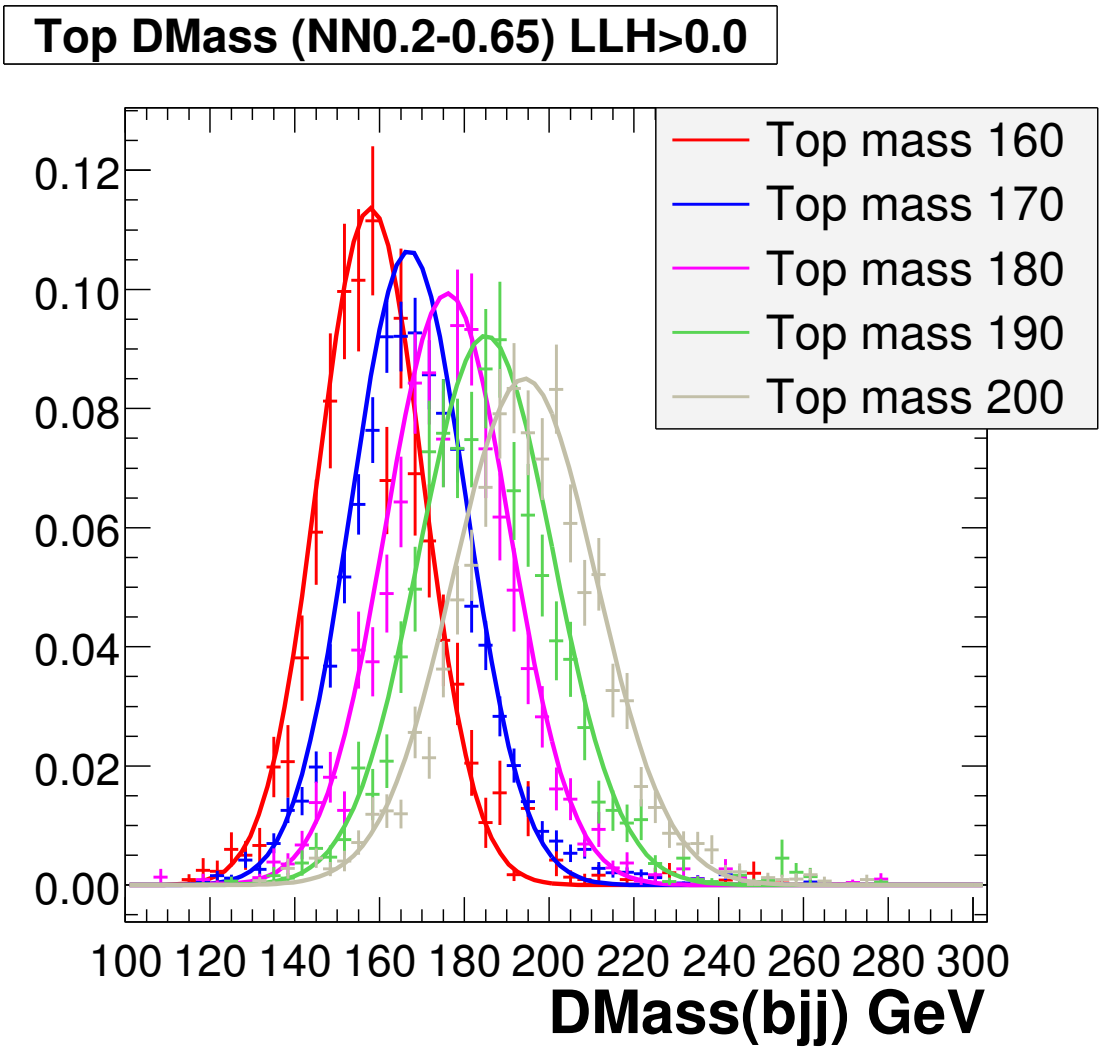

Figure 5.29. Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65. The line shows the function fit and the points are the original histograms from MC ttbar events.

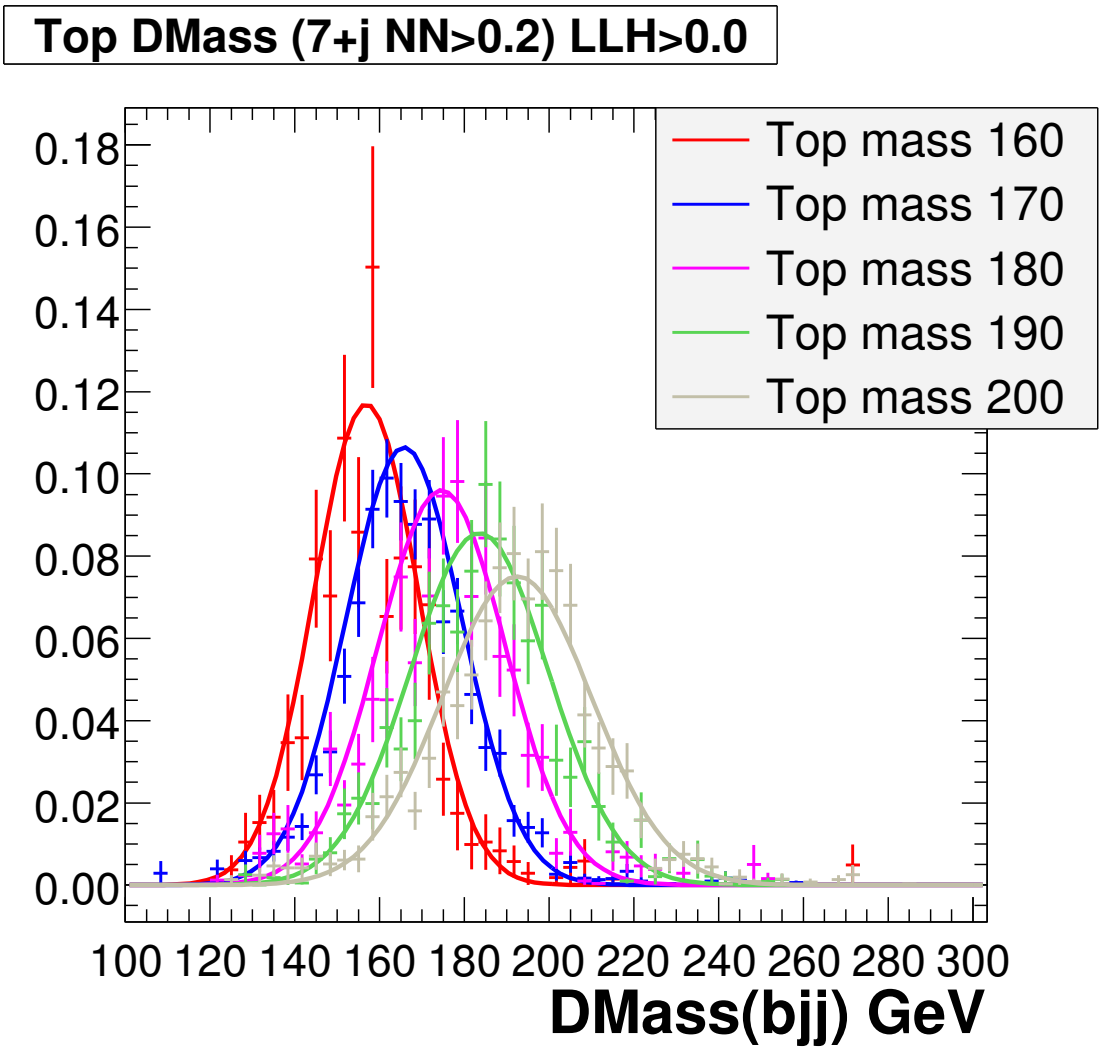

Figure 5.30. Template functions for different MC top mass with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2. The line shows the function fit and the points are the original histograms from MC ttbar events.

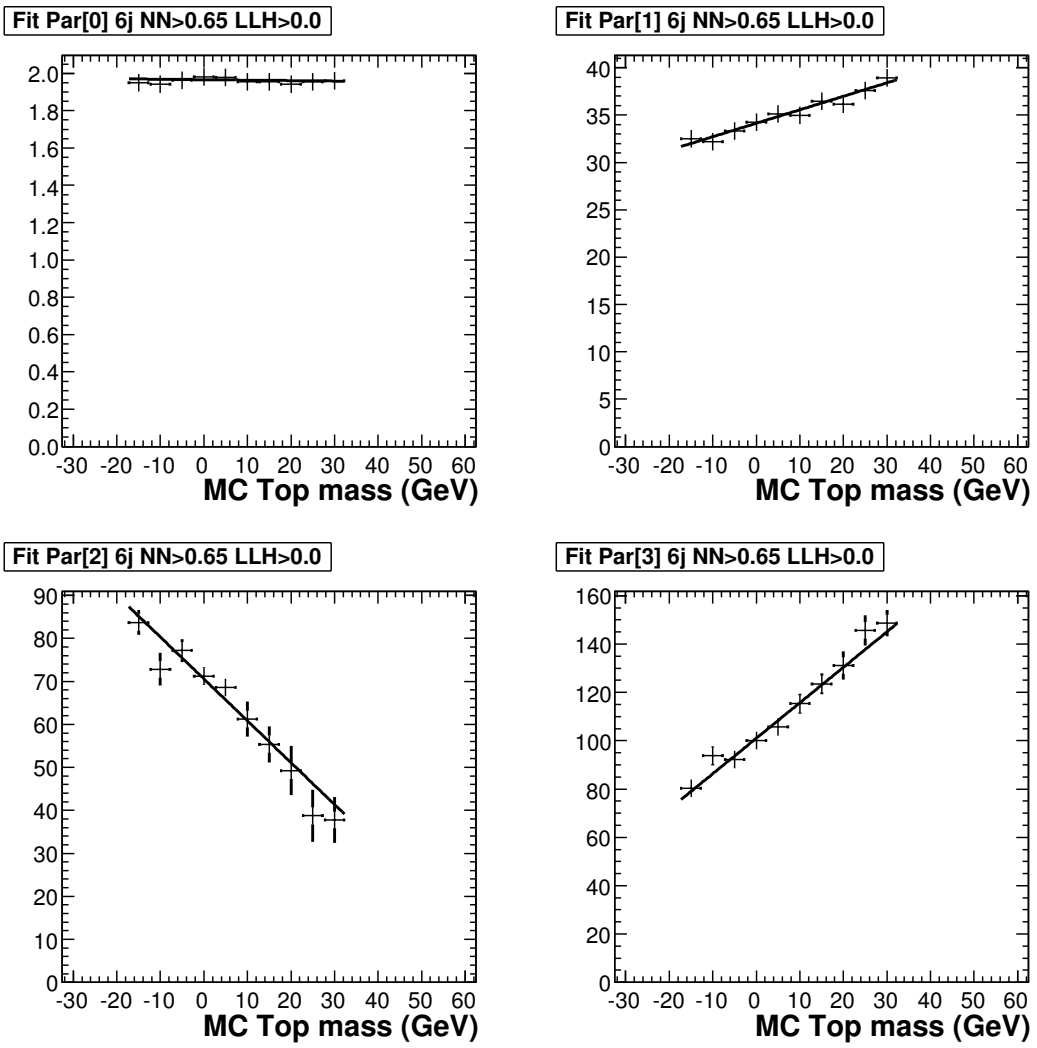

Figure 5.31. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in exactly 6 jets, 2 BIDNN > 0.65.

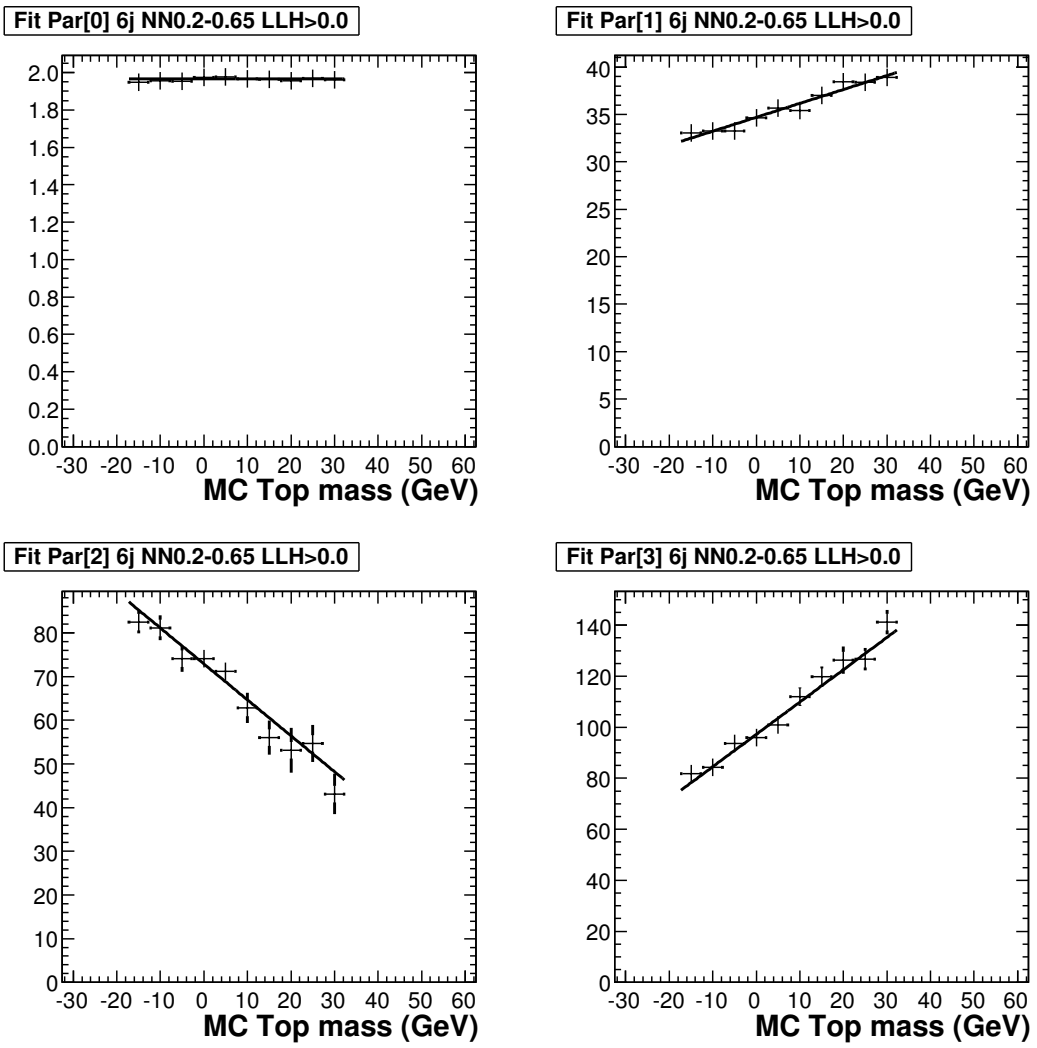

Figure 5.32. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in exactly 6 jets, 2 BIDNN 0.2-0.65.

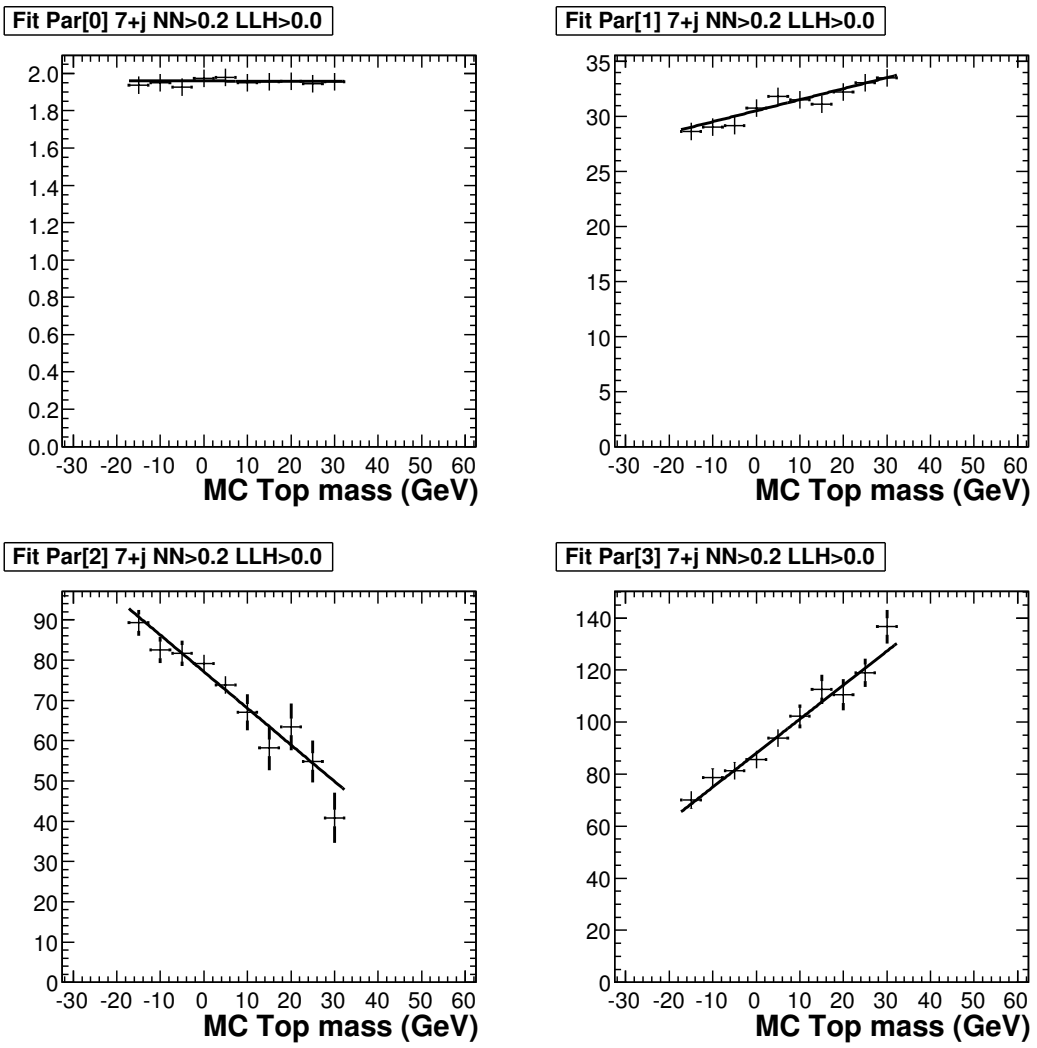

Figure 5.33. Second parameterization of parameters as a function of top mass, from the fit $Dmass(bjj)$ with combined wrong jet combinations and not-parton-matched events in 7+ jets, 2 BIDNN > 0.2.

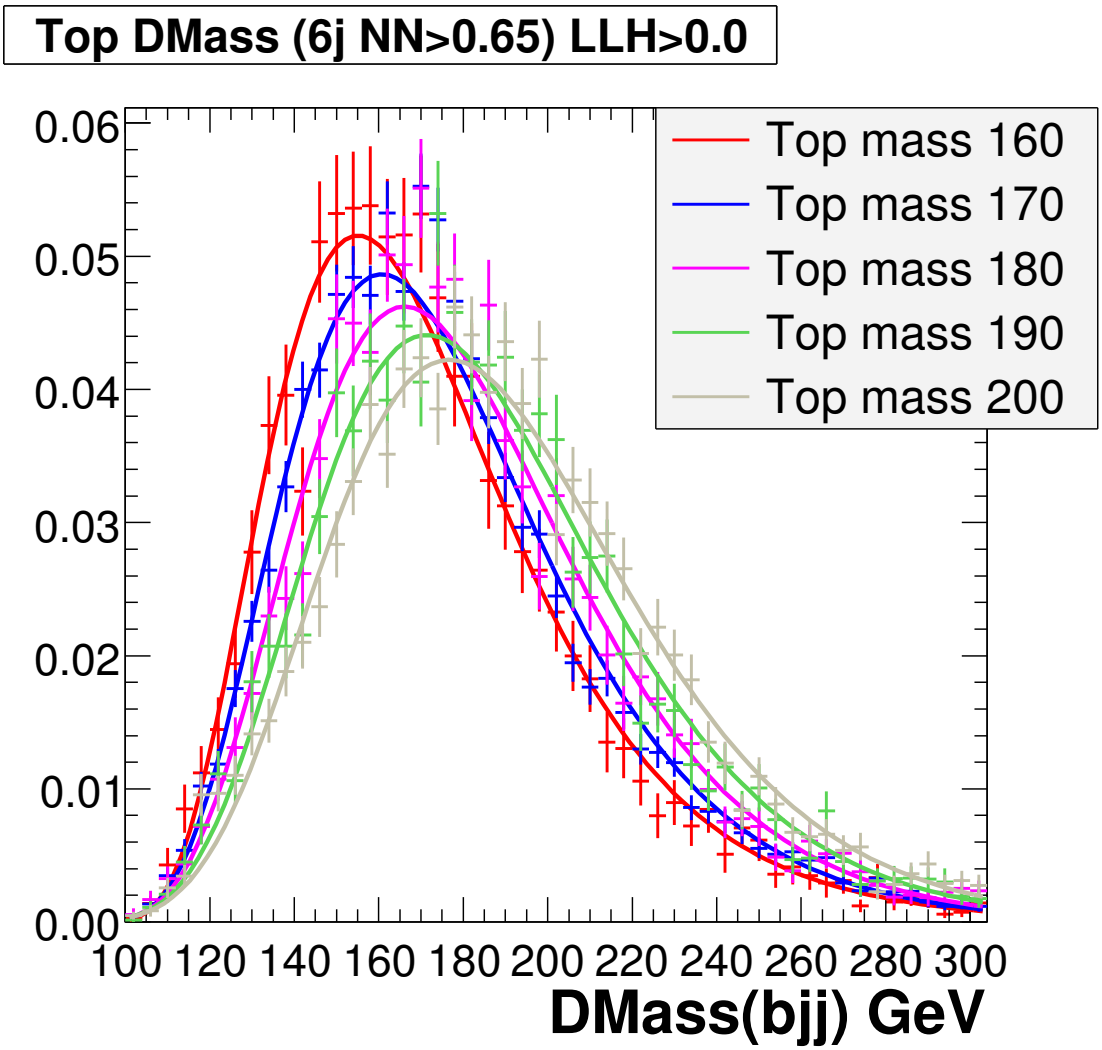

Figure 5.34. Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN > 0.65. The line shows the function fit and the points are the original histograms from MC ttbar events.

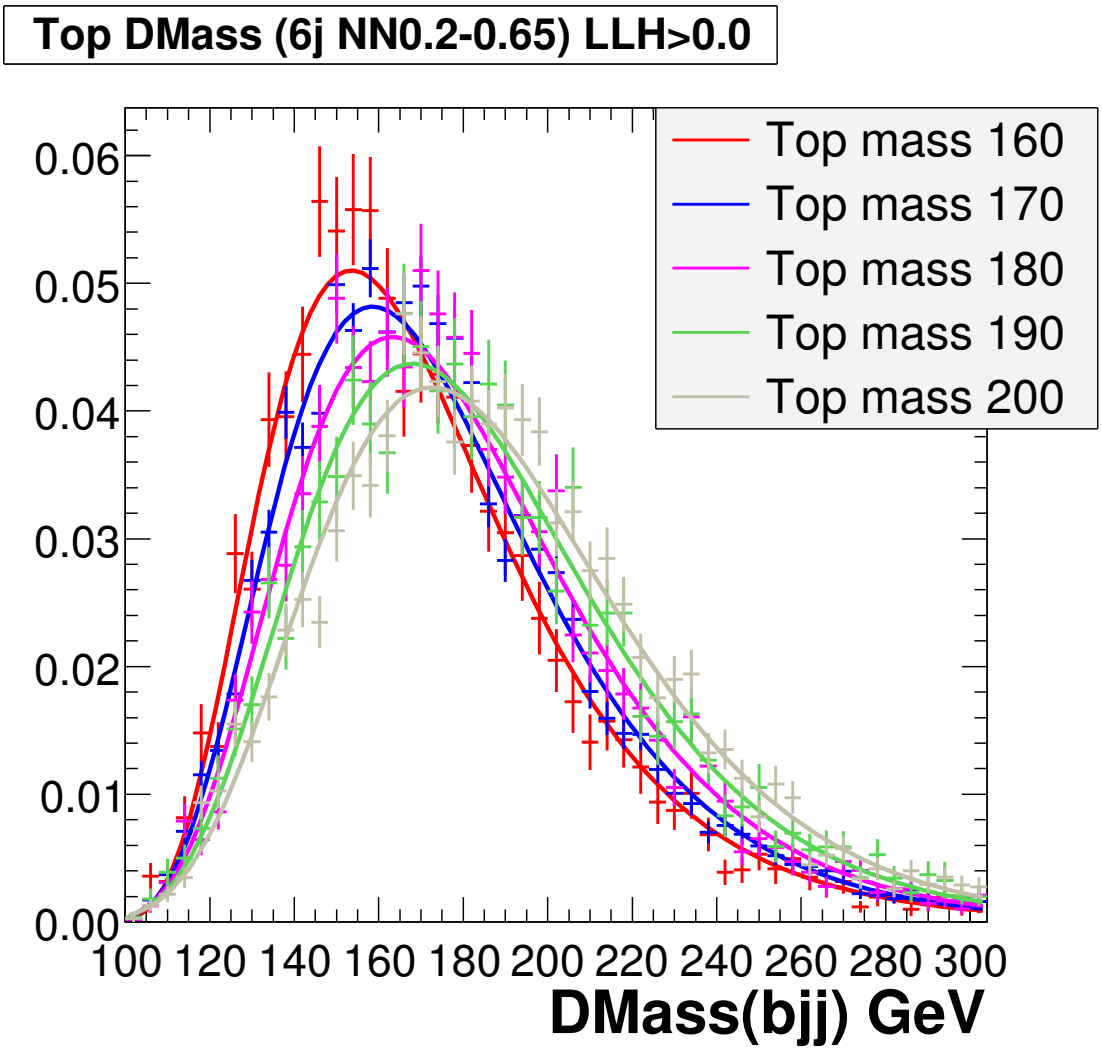

Figure 5.35. Template functions for different MC top mass with correct bjj/jj jet combinations in exactly 6 jets, 2 BIDNN 0.2-0.65. The line shows the function fit and the points are the original histograms from MC ttbar events.

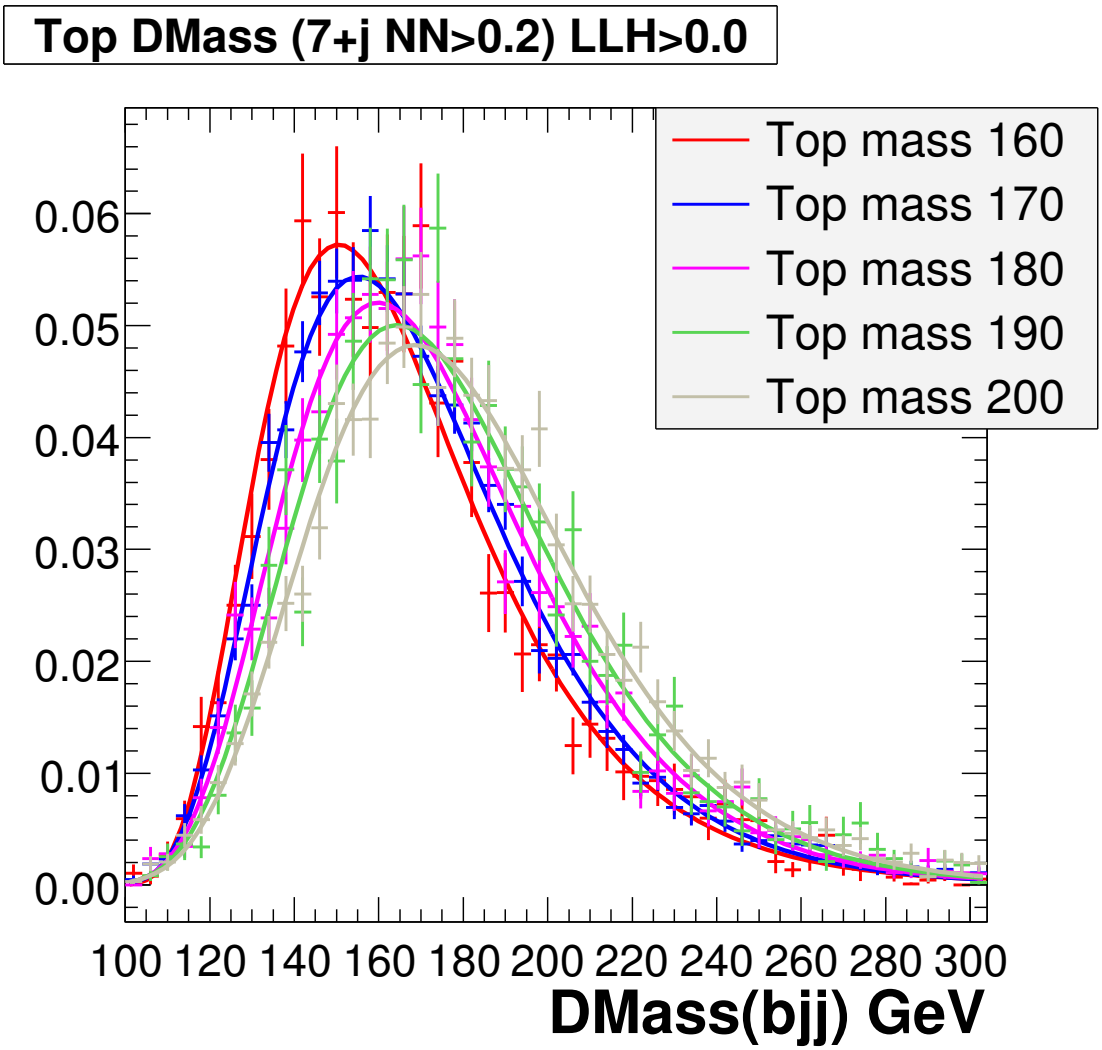

Figure 5.36. Template functions for different MC top mass with correct bjj/jj jet combinations in 7+ jets, 2 BIDNN > 0.2. The line shows the function fit and the points are the original histograms from MC ttbar events.

5.4 Background Modeling

The dominant background in the all hadronic channel is QCD-multijet events. Although W boson plus 4 extra jets events also exist and have a similar cross section as the $t\bar{t}$ events, the requirement of the presence of a b-quark results in a smaller accepted cross section (about 15% compared to $t\bar{t}$). The requirement of 2 b-tagged jets in our event selection further reduces the background portion of W plus 4 jets events.

Although QCD multijet events can be simulated using Monte Carlo generators, the high multiplicity and high energies of the jets selected in this analysis represent a exotic corner of phase space that must be treated with care. For example, parton shower generators such as Pythia use leading order matrix elements to simulate the initial partons, then evolve these using splitting functions to generate additional hard radiation. Given that neither the appropriate evolution scale nor the absolute value of α_s is known that well [62] for these high-multiplicity events, relying on these generators to accurately predict the multi-jet background would result in huge theoretical or modelling uncertainties. On the other hand, a direct calculation of the 6-jet matrix element does not exist due to the huge number of potential Feynman diagrams. Therefore, we shall try to model the background using data events.

It is not possible to simply extract the background events directly from the 6-jets data samples, since from Monte Carlo studies it is known that there are always some $t\bar{t}$ events left in the 6-jets no matter what selection cut is being applied, or whether the jets are b-tagged or not.

Another proposal of starting with untagged 6-jets data events and then assigning b-tags to 2 of the jets is not feasible either. This is because there exist

correlations between the 2 b-jets in hard-scattering events that are not possible to replicate with this procedure. In addition, the difference in the phase space of the jets between b-tagged and untagged events are quite different. Randomly assigning jets as b-tagged does not reproduce the energy spectra, for example, of the b-tagged jets.

Instead, consider the 5-jet events with 2 b-tagged jets. The majority of the 5-jets events are QCD multijets and are free of signal contamination, since the $t\bar{t}$ hadronic signal has at least 6 jets in the final state. Therefore one can add one (or more) additional soft jet(s) to the 5-jets data samples to simulate 6-jets (or higher jet multiplicity) QCD multijet events. This is similar to the FSR (final state radiation) simulation in the Pythia generator. Also as the 5-jet samples are already b-tagged, the existing correlations between the b-jets are conserved.

When adding extra jets to the 5-jet events (acceptor), one need to make sure its phase space is similar the 6-jet events (donor). This is done by requiring the 5th jet p_T between the 5-jets and 6-jets events are the same within a narrow range (< 1 GeV). Also the jet that is being inserted into the acceptor event should not overlap with the existing jets in the 5 jets event. Figure 5.37 shows an illustration of the 6-jet background generation.

In addition, sampling techniques are required for better matching of the phase space between the 5-jets event and the first 5 jets of the 6-jets events. A 6 jet multijet QCD event can be treated as having 5 hard-scattering jets plus 1 soft radiative jet. Under the same viewpoint a 5-jet event can be viewed as 4 hard-scattering jets plus 1 soft radiative jet. The last hard-scattering jet in both cases results from similar QCD processes and therefore should have similar jet p_T characteristics as well. Therefore one can first start with the 4th jet p_T distribution

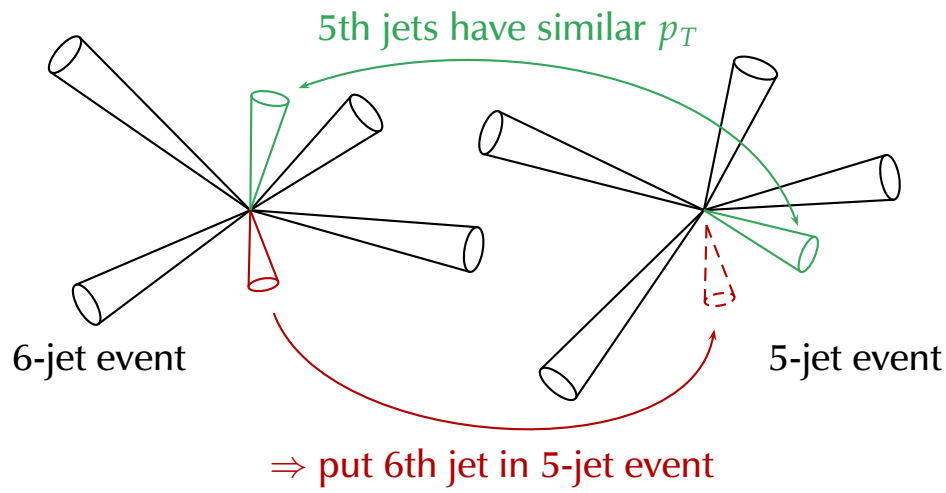

Figure 5.37. Illustration of 6-jet background generation. The softest jet from the 6-jet events is inserted into the 5-jet event if the 5th jet have similar p_T.

in 5-jet events and select an acceptable 6-jet donor event based on the matching of its 5th jet p_T.

The following shows the detailed flow of the background generation algorithm:

- Rank the jet order in decreasing jet p_T.
- Get the 4th jet p_T distribution from all 5-jet samples, scale the distribution to give sampling probablity $prob_{samp}(p_T)$ (where each p_T bin has a maximum probability of 1).
- Start with a 5-jet (acceptor) event.
- Randomly pick a 6-jet (donor) event satisfying 5th jet $\Delta p_T(donor, acceptor) <$ 1 GeV/c.
- Throw a random real number in $[0, 1)$. Reject the 6th (and up) jet(s) from

the donor sample if the random number is larger than the sampling probability using the 5th jet of the donor: $prob_{samp}(jet5p_T)$.

- Reject the 6th (and up) jet(s) from the donor sample if any of the donor jets would overlap ($dR < 0.5$) with any jets in the acceptor event, or if any of the donor jet has a p_T higher than the softest acceptor jet.

- Insert the 6th (and up) jet(s) into the acceptor event.

Given the limited quantities of 5-jet acceptor events, to generate sufficient background events one needs to reuse the softest jet in the 6th jet donor events. This can be done by randomly rotating the *phi* angle of the jet to be inserted into the acceptor. Certainly, the jet will be rejected if its angular separation with any of the jets within the acceptor event is less than 0.5.

The above algorithm can be applied to generate a 6 (or more) jets background sample by adding 2 (or more) jets to the 4-jet data events as well. In this case the 2 (or more) jets are rotated as a whole in the *phi* angle when they are inserted into the 4-jet event, in order to preserve the correlation between soft jets.

5.5 Verifying the Background Model

There are various ways to confirm the events generated by the algorithm are actually background-like. Given that the 5-jet data events are essentially QCD multijet events, one can use that as a reference and compare with the background events generated by adding 1 jet to the 4-jet events – "4+1" sample. For verfication in of the phase space one can compare the different jet p_T distributions. Figure 5.38 to 5.42 shows the 1st to 5th jet p_T comparison, from which it can be seen the generated background events have a good match with the data. Figure 5.43

and 5.44 shows the comparison of the maximum and minimum angular separation among all the jets. Figure 5.45 and 5.46 shows the comparison of the maximum and minimum rapidity differences among all the jets. Finally figure 5.47 and 5.48 shows the comparison of the mass(bjj) from all 3-jet combinations (1 tagged plus 2 untagged jets) and the mass(jj) from all 2-jet combinations (2 untagged jets).

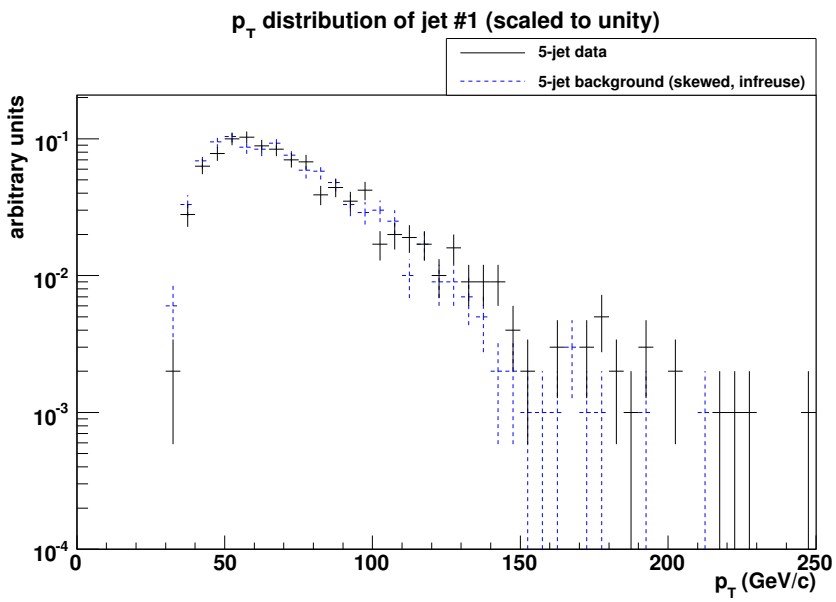

Figure 5.38. 1st Jet p_T comparison between the 5-jet data events and the "4+1" background events.

5.6 Background template fitting

With the background model in hand, the background $Dmass(bjj)$ distribution can then be obtained and used for likelihood fitting. However since the majority of

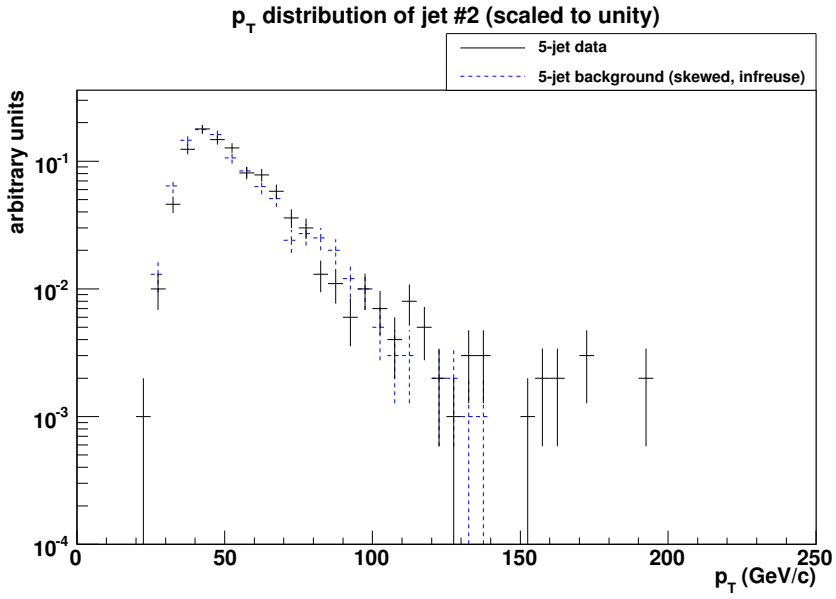

Figure 5.39. 2nd Jet p_T comparison between the 5-jet data events and the "4+1" background events.

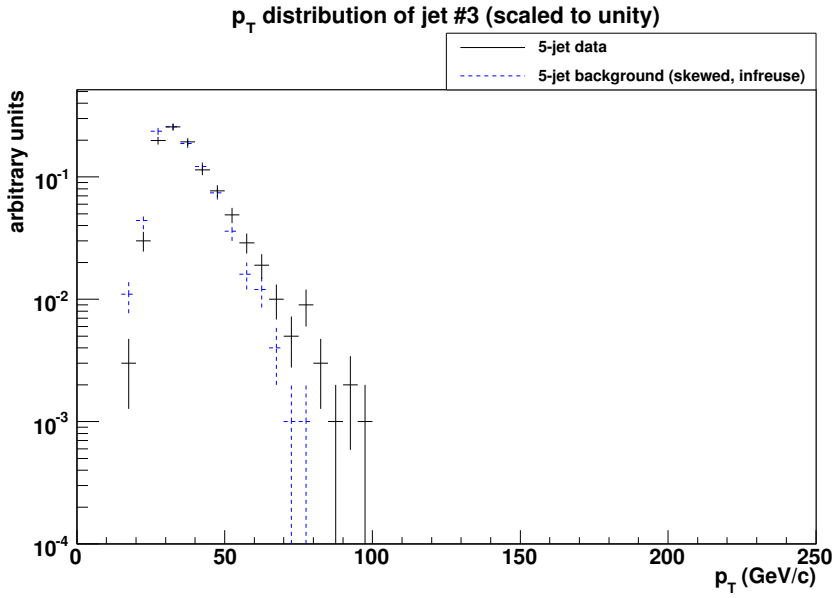

Figure 5.40. 3rd Jet p_T comparison between the 5-jet data events and the "4+1" background events.

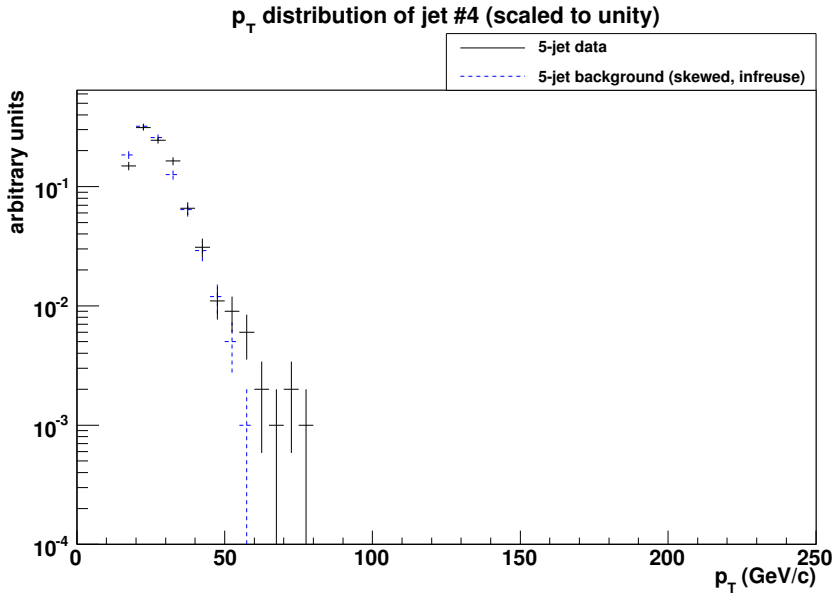

Figure 5.41. 4th Jet p_T comparison between the 5-jet data events and the "4+1" background events.

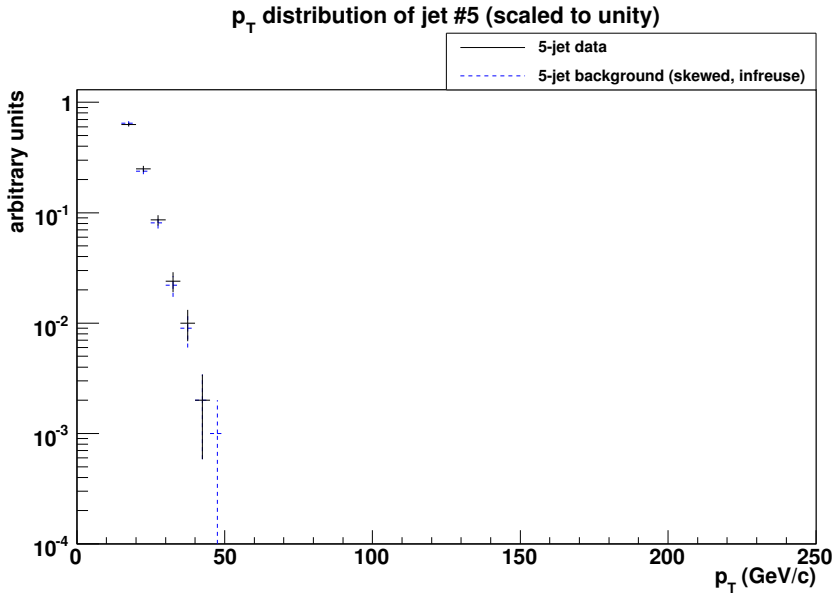

Figure 5.42. 5th Jet p_T comparison between the 5-jet data events and the "4+1" background events.

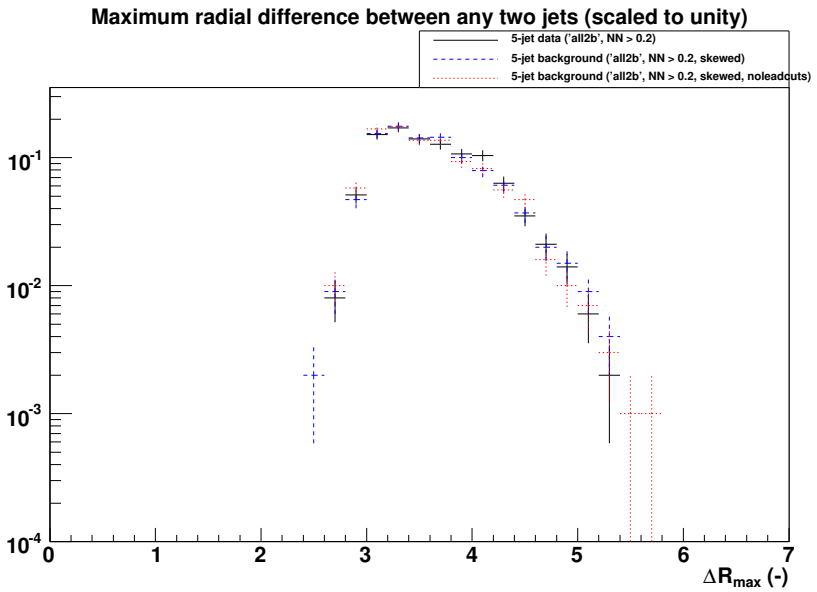

Figure 5.43. Maximum angular separation (dR) among all the jets between the 5-jet data events and the "4+1" background events. The red cross points are the "4+1" background generated following the procedure mentioned in this section.

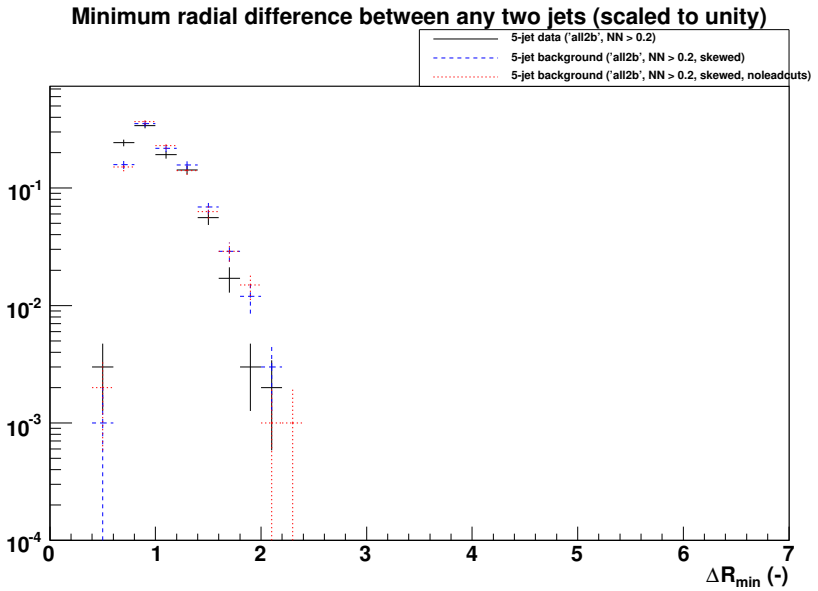

Figure 5.44. Minimum angular separation (dR) among all the jets between the 5-jet data events and the "4+1" background events.

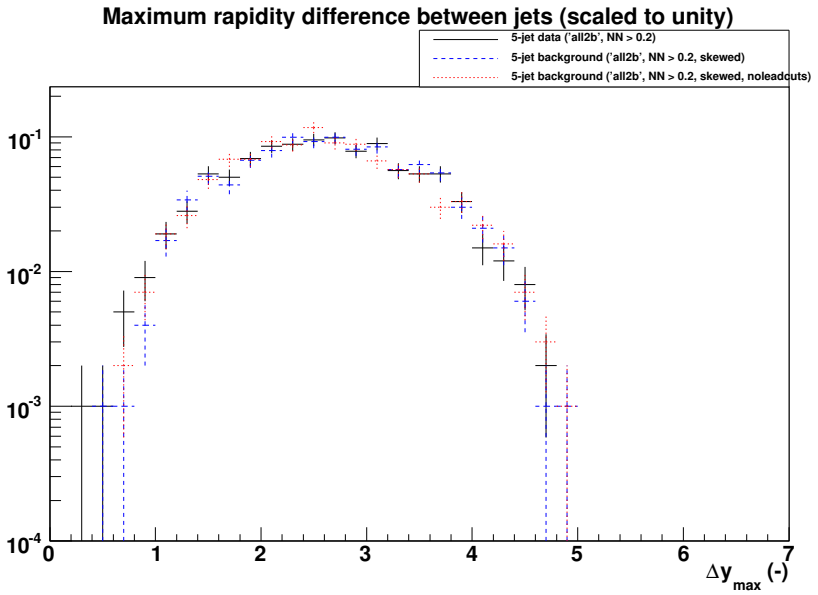

Figure 5.45. Maximum Rapidity difference (dY) among all the jets between the 5-jet data events and the "4+1" background events.

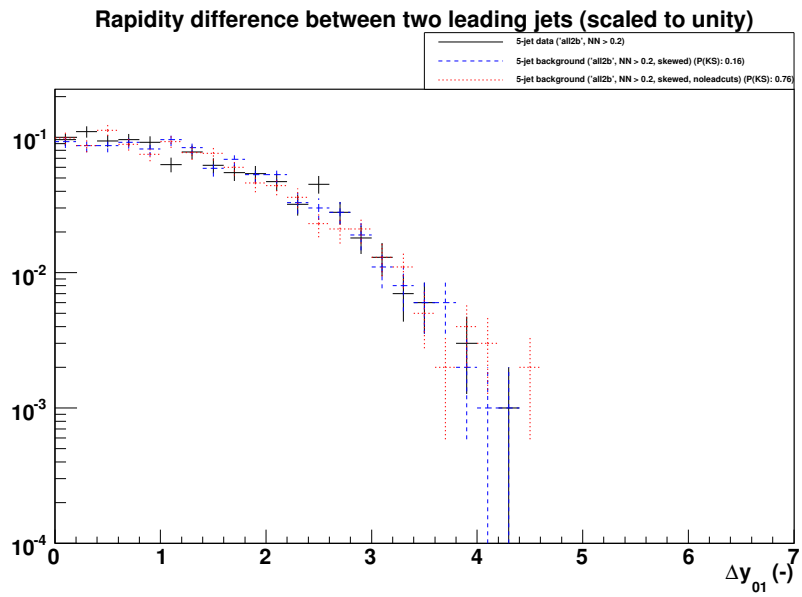

Figure 5.46. Rapidity difference (dY) of the leading and 2nd leading jets between the 5-jet data events and the "4+1" background events.

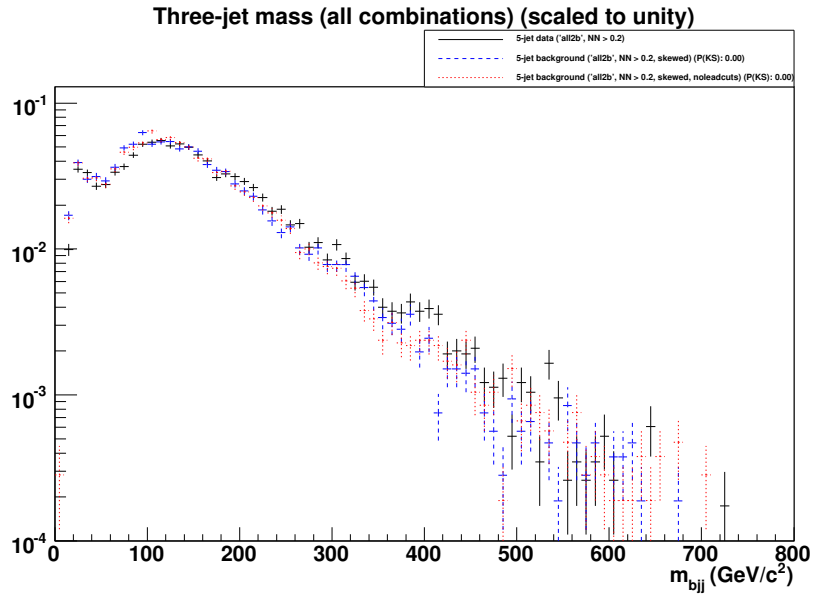

Figure 5.47. mass(bjj) of all 3-jet combinations comparison between the 5-jet data events and the "4+1" background events.

the candidate events are background (from Likelihood fit results), the fitted mass result is very sensitive to the background $Dmass(bjj)$ shape. Fit using simple analytic functions cannot yield a satisfactory background template. As it turns out, Splines, which are piece-wise polynomial functions, can provide a good fit to the shape while removing the statistical fluctuations. Figure 5.49 is the Spline fitted template using background model in different BID NN regions. Polynomials of order 3 are used in the spline fit.

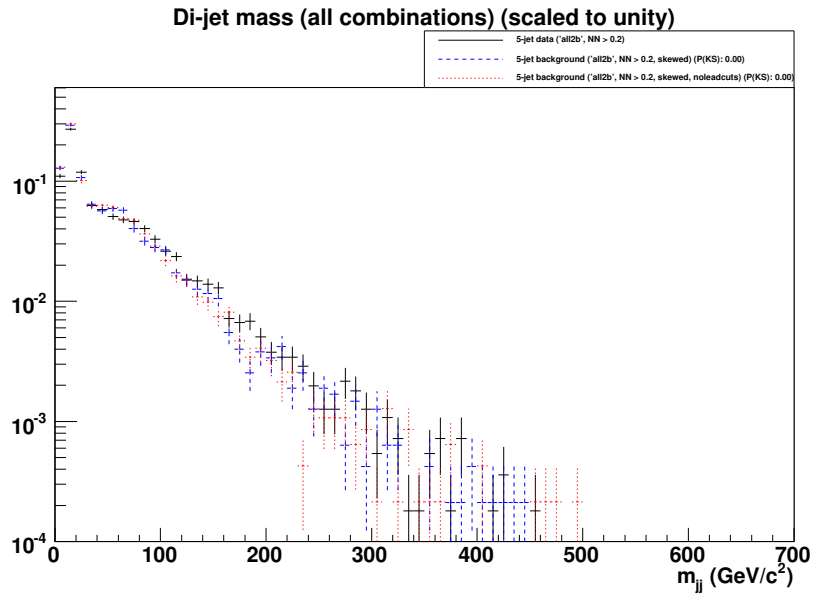

Figure 5.48. mass(jj) of 2-jet (un-tagged only) combinations comparison between the 5-jet data events and the "4+1" background events.

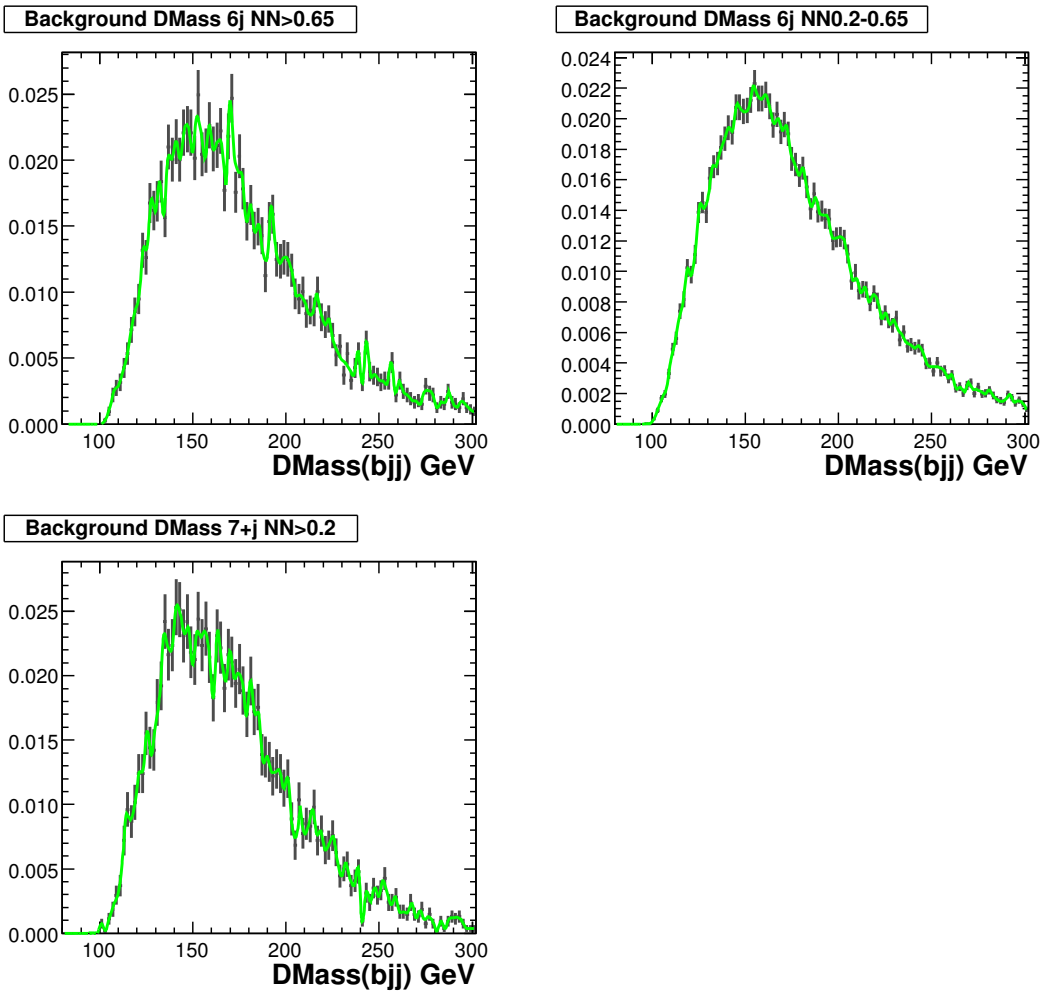

Figure 5.49. Spline fit of $Dmass(bjj)$ using background model fits in 3 different BID NN regions: Exactly 6 jets, 2 BIDNN > 0.65, 2 BINNN 0.2-065, 7+ jets, 2 BIDNN > 0.2.

CHAPTER 6

DEVELOPMENT of DISCRIMINANT

The trigger system, together with the offline 3JET skim and event selections have already reduced the overwhelming background from the data sample. However, there is still a significant amount of background events the data candidates. Due to the overlap in phase space it is impossible to completely separate signal and background. However the differences in the topological properties between signal and background allow one to develop a discriminant. Such a discriminant can provide a measurement of the signal purity of the candidate events and maximize the signal to background ratio.

6.1 Topological variables

Here are the candidate topological variables to be tested in the development of the topological discriminant.

6.1.1 H_T and Centrality

The scalar sum of the momenta and transverse momenta of all jets (with $p_T > 15$ GeV) are defined as H and H_T:

$$H = \sum_i |\vec{p}_i| \tag{6.1}$$

$$H_T = \sum_i p_{T_i} \tag{6.2}$$

$$Centrality = \frac{H_T}{H} \tag{6.3}$$

H_T and H are 2 powerful variables to discriminate between $t\bar{t}$ and QCD events. However they are strongly correlated to the top mass and therefore not suitable to be included in the discriminant.

Centrality is defined as the ratio H_T/H. This variable is less correlated with the top mass.

6.1.2 Momentum tensor, Sphericity and Aplanarity

A quadratic momentum tensor of the event can be defined as [63]:

$$M^{\alpha,\beta} = \frac{\sum^i p_i^\alpha p_i^\beta}{\sum^i |\vec{p}_i|^2} \tag{6.4}$$

where $\vec{p}_i$ is the three-momentum of the i-th jet in the final state, and α, β are the x-, y- and z components of the momentum. The three eigenvalues from the tensor are ordered to give:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \tag{6.5}$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{6.6}$$

The sphericity and aplanarity are defined as:

$$Sphericity = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{6.7}$$

$$Aplanarity = \frac{3}{2}\lambda_3 \tag{6.8}$$

The sphericity lies between zero and 1. For two jets in back-to-back events the sphericity is closer to zero, whilst for spherical events such as $t\bar{t}$ the value is closer to 1. The aplanarity is zero if all the jets in the event lie on the same plane. Again jets from $t\bar{t}$ events are more spherically distributed and therefore the events have larger aplanarity.

Due to the correlation between Sphericity and Aplanarity, the eigenvalues are used directly instead for likelihood development.

6.1.3 Fox-Wolfram moments

In 1979 Fox and Wolfram devised a complete set of spherically symmetric "event shape" variables to describe the final states of electron-positron collisions [64], [65]. These l-th Fox-Wolfram moment FWM_l is defined as

$$FWM_l = \left(\frac{4\pi}{2l+1}\right) \sum_{m=-l}^{+l} \left| \sum_{i}^{ptcl} Y_l^m(\Omega_i) \frac{|\vec{p}_i|}{E_{\text{tot}}} \right|^2 \tag{6.9}$$

where $\vec{p}_i$ is the three-momentum of the i-th particle in the final state (which are jets in our case). $\Omega_i = (\theta_i, \phi_i)$ shows the direction of $\vec{p}_i$, E_{tot} is the total energy of all final state particles and Y_l^m are the spherical harmonics. All FWM_l fall into the range $[0, 1]$. Conservation of momentum in the final state ($\sum_i \vec{p}_i = \vec{0}$) gives $H_1 = 0$.

The combination of the spherical harmonics Y_l^m with the directions of the final

state particles $\vec{p}_i$ gives a unique relation between the spatial shape of the final state and the values for FWM_l..

For perfect back-to-back di-jets events $FWM_l = 0$ (1) for all even (odd) l. Perfectly spherical events will lead to all FWM_l being zero for all $l > 0$.

6.1.4 dR_{bb}, dR_{max}, dY_{01}, dY_{02}, dY_{12} and $cos\theta^*$

The angular separation dR of the two b-tagged jets is useful, where $dR = \sqrt{d\eta^2 + d\phi^2}$. In addition the rapidity difference between the two b-tagged jets: $dY_{bb} = |Y_{b1} - Y_{b2}|$ is also useful.

The rapidity difference among the 1st, 2nd and 3rd leading p_T jet: dY_{01}, dY_{02} and dY_{12} are also useful, together with the angle between the leading p_T jet and the beam line ($cos\theta^*$) as well.

6.1.5 $gluonM$

In QCD events, sometimes a gluon can split to yield 2 additional gluons or quarks, thus faking a top all-hadronic 6 (or more) jets event. Since the original gluon itself is massless, the invariant mass of the two jets from gluon splitting is expected to be small as well. Therefore one can define $gluonM$ by looking at the minimum invariant mass among all the 2 jets combinations:

$$gluonM = min\{M_{ij}\};\ i \neq j \text{ are any two jets in the event} \tag{6.10}$$

6.1.6 M_{bb}/M_3 and p_{Tn}/M_4

The invariant mass of the 2 two b-tagged jets can be calculated by first summing the 4-vectors of the two b-tagged jets:

$$p_{bb} = p_{b1} + p_{b2} = (E_{b1} + E_{b2}, \vec{p}_{b1} + \vec{p}_{b2}) \tag{6.11}$$

$$M_{bb} = \sqrt{E_{bb}^2 - \vec{p}_{bb}^2} \tag{6.12}$$

However, M_{bb} by itself is correlated to the top mass. It is possible to reduce the correlation between M_{bb} and top mass by using the ratio M_{bb}/M_3 instead, where M_3 is the invariant mass of the leading 3 jets in the event:

$$p_{1-3} = (E_{1-3}, \vec{p}_{1-3}) = \left(\sum_{i=1,2,3} E_i, \sum_{i}^{jets} \vec{p}_i \right) \tag{6.13}$$

$$M_{1-3} = \sqrt{E_{1-3}^2 - \vec{p}_{1-3}^2} \tag{6.14}$$

Similarly, the jet p_T for each jet is useful. However the jet p_T also needs to be divided by the invariant mass of the leading 4 jets (M_4) in order to reduce the correlation with the top mass.

Figure 6.1 to 6.4 shows the distributions of the above topological variables between signal (MC $t\bar{t}$ m_{top} = 170 GeV) and background samples, where their discriminating power can be observed.

6.2 Variables Correlation

It is useful to check whether the list of the candidate variables are correlated with each other and it is desirable to pick the set of variables that are not so correlated with each other. Although some discriminants such as a Neural Network would take care of the correlation, other methods such as a simple likelihood will

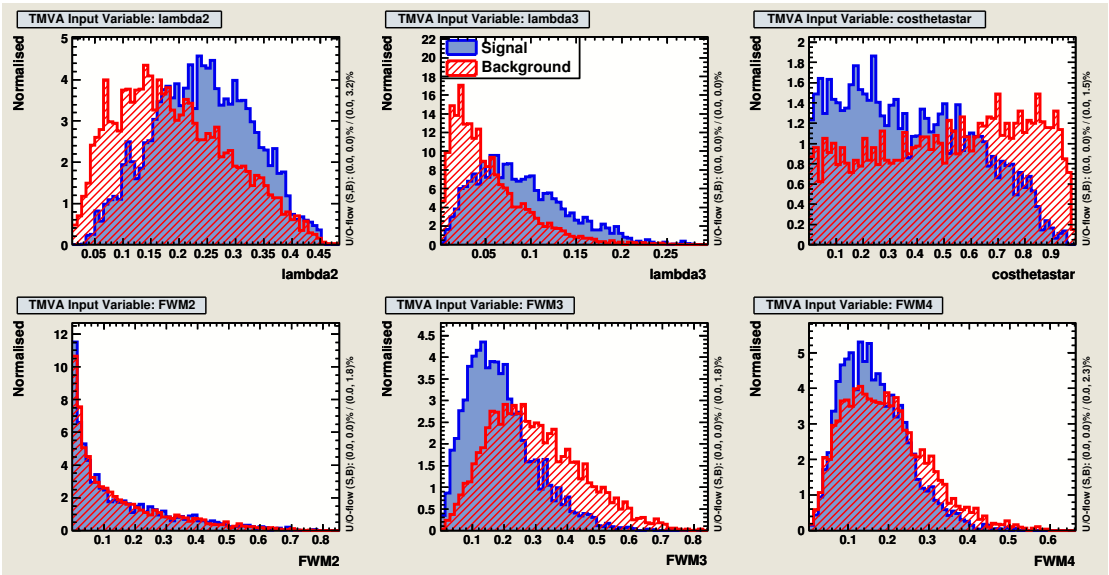

Figure 6.1. Topological variables λ_2, λ_3, $cos\theta^*$, FWM_2, FWM_3, FWM_4 between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65

have its performance affected by the correlation of the variables.

Figure 6.5 shows the correlations between different candidate variables, together with the invariant mass variables, for MC $t\bar{t}$ $m_{top} = 170$ GeV events. Similarly, Figure 6.6 shows the correlations of the variables in the background events.

From the plots it can be seen that FWM_3 is correlated with λ_2, λ_3, dY_{01} as well as the scaled $p_T n/M_4$ variables. Similarly the first 3 leading p_T jets are quite correlated to the rapidity difference among those 3 jets as well, as they point to the same jets. Among those variables which are correlated, it is preferred to use only 1 or 2 of them for the least correlation within the final discriminant.

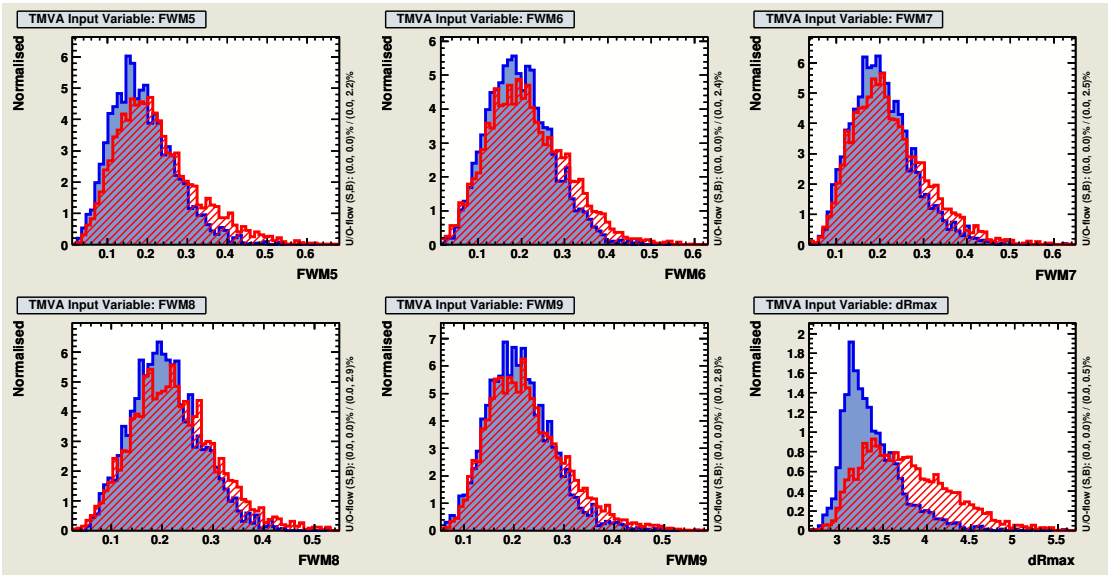

Figure 6.2. Topological variables $FWM_5 - FWM_9$ and $dRmax$ between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65

6.3 Discriminant methods

Here we present the different discriminant methods that will be investigated in this study. The discriminant methods are already implemented in the software package TMVA [66], which is employed in this study.

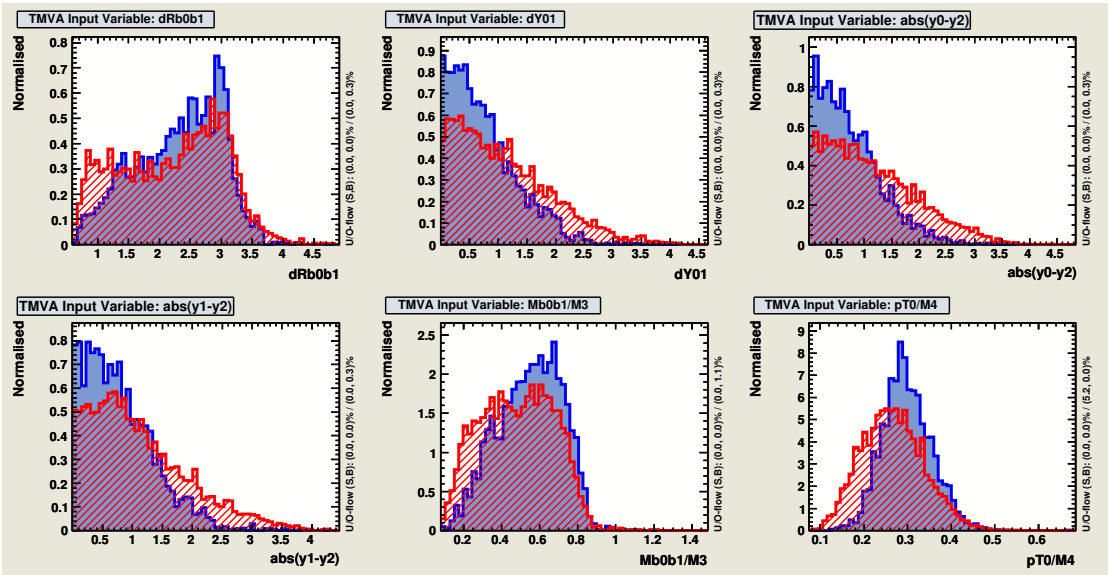

Figure 6.3. Topological variables dR_{bb}, dY_{01}, dY_{02}, dY_{12}, M_{bb}/M_3 and p_T0/M_4 between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65

6.3.1 Likelihood

A simple likelihood discriminant, $\mathfrak{D}$ can defined as:

$$\begin{aligned}\mathfrak{D} &= \frac{S(x_1, x_2, ..., x_i)}{S(x_1, x_2, ..., x_i) + B(x_1, x_2, ..., x_i)} \\ &\approx \frac{\prod_i S_i}{\prod_i S_i + \prod_i B_i} = \frac{\prod_i S_i/B_i}{\prod_i S_i/B_i + 1} \\ &= \frac{exp\left(\prod_i ln(S_i/B_i)\right)}{exp\left(\prod_i ln(S_i/B_i)\right) + 1}\end{aligned} \tag{6.15}$$

Each S_i and B_i are the normalized signal and background distributions of the topological variable i. The second line of the equation uses an approximation that neglects the correlation between variables. This is the reason why we want to pick the set of variables that have small correlations among themselves. For better

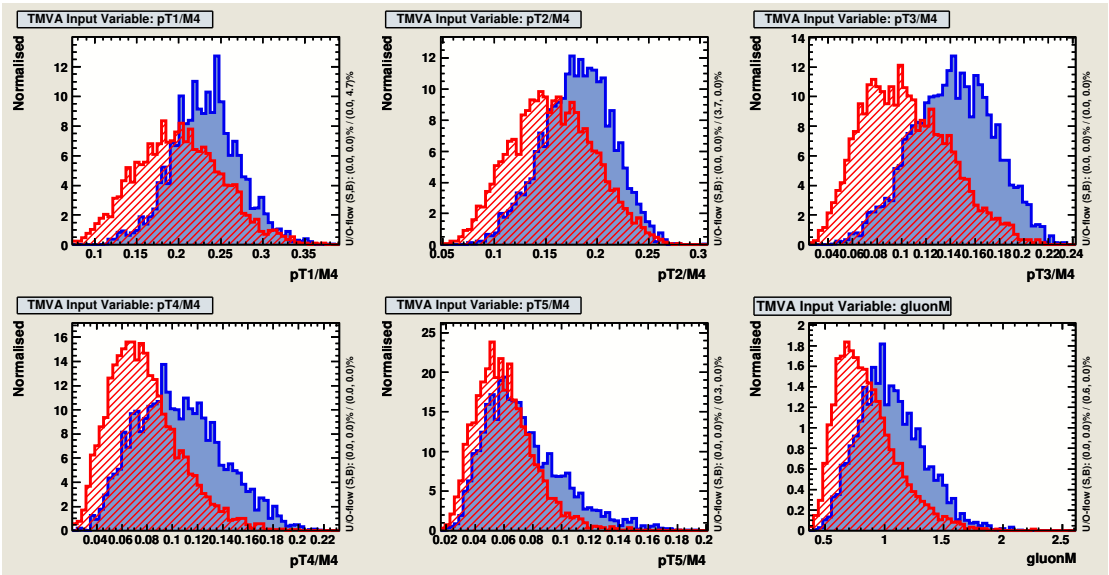

Figure 6.4. Topological variables p_T0/M_4 - p_T5/M_4 and gluonM between signal and background events for exactly 6 jets, 2 b-jets with BIDNN > 0.65

results, polynomial fits of $ln(S_i/B_i)$ (rather than the actual distribution itself) are used in the likelihood.

In fact, there exist different methods to optimize the likelihood by transforming the variables. For example, one of the methods is Principle Component Analysis (PCA) [67], which involves the eigenvalue decomposition of the covariance matrix to find the set of linear transformed variables that are the least correlated to each other. Then the transformed variables replace the original variables in the likelihood equations above to calculate the discriminant.

6.3.2 Fisher Discriminant

Another good discriminant is the Fisher discriminant (also called a Linear discriminant analysis, LDA). This method is closely related to linear regression

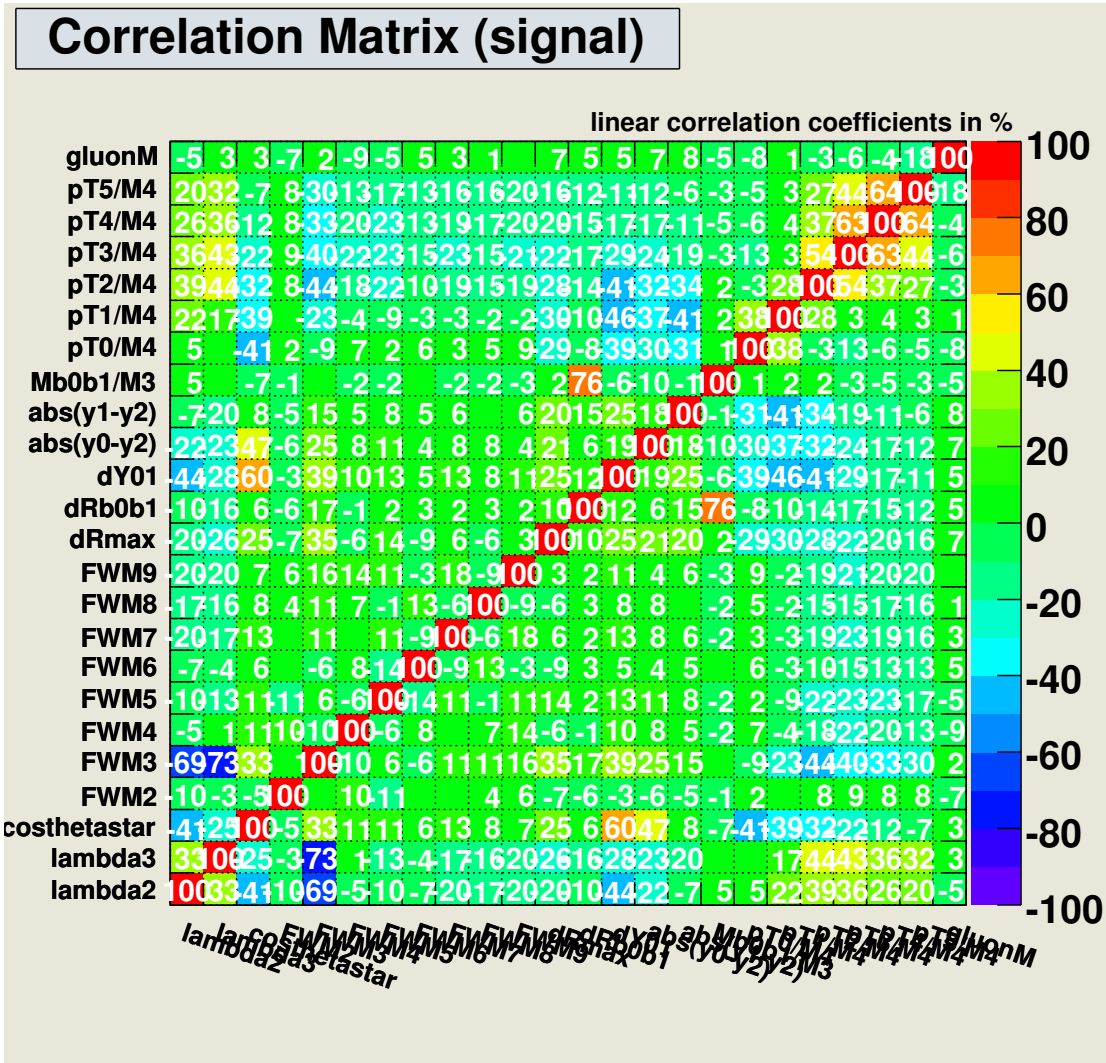

Figure 6.5. Correlation Matrix between different candidate variables, using MC $t\bar{t}$ events with $m_{top} = 170$ GeV, with 6 jets, 2 b-jets with BIDNN > 0.65

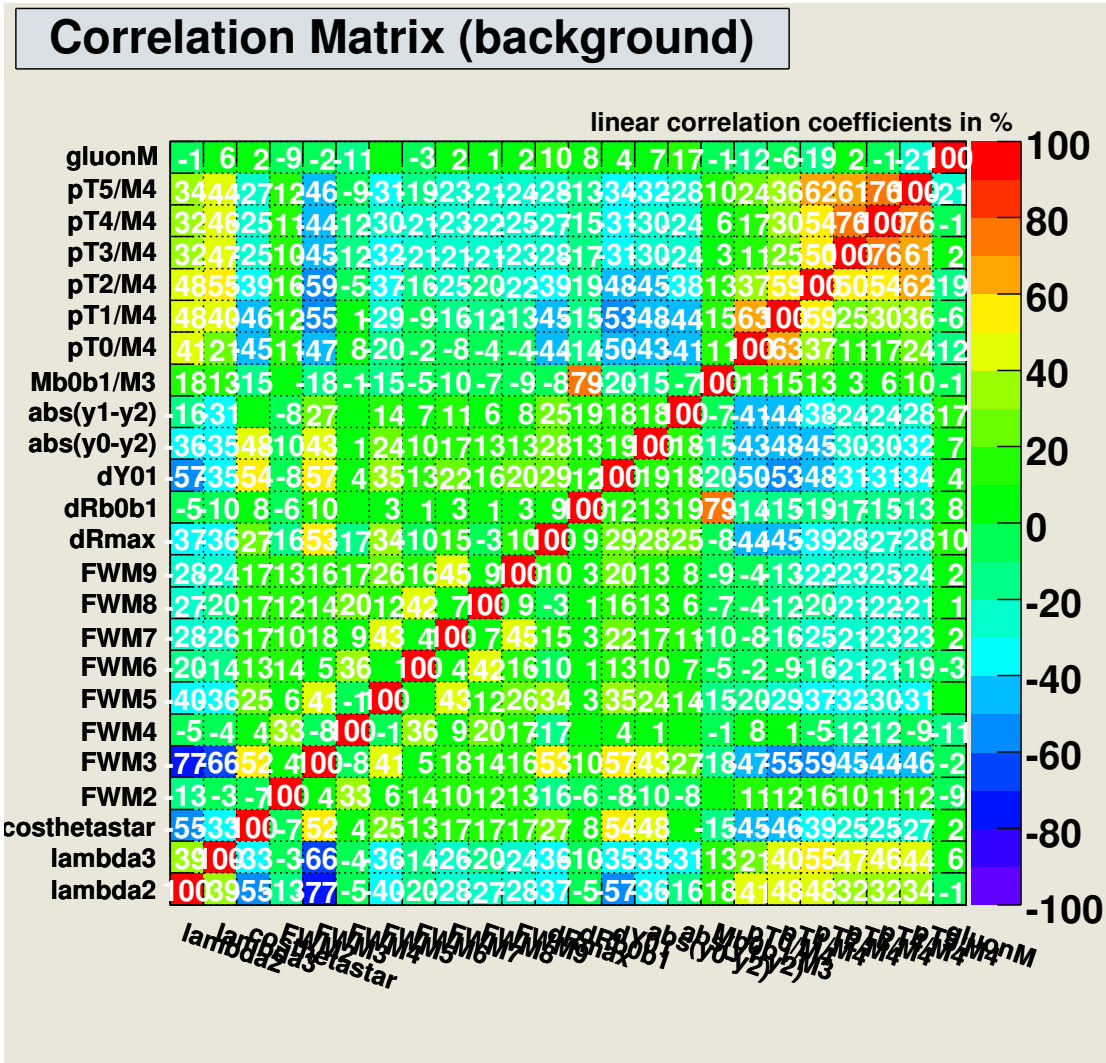

Figure 6.6. Correlation Matrix between different candidate variables, using Background events from the background model, with 6 jets, 2 b-jets with BIDNN > 0.65

in multiple dimensions. Suppose the vectors of observables $\vec{x}$ from signal and background samples have means $\vec{\mu}_{sig}$ and $\vec{\mu}_{bkg}$, and covariances Σ_{sig} and Σ_{bkg}. One can define the separation between signal and background by looking at the ratio of the variances "between" the signal and background and "within" the signal and background:

$$\begin{aligned} S &= \frac{\sigma^2_{between}}{\sigma^2_{within}} \\ &= \frac{(\vec{w} \cdot \vec{\mu}_{sig} - \vec{w} \cdot \vec{\mu}_{bkg})^2}{\vec{w}^T \Sigma_{sig} \vec{w} \vec{w}^T \Sigma_{bkg} \vec{w}} \end{aligned} \tag{6.16}$$

Therefore maximizing the quantity S above will give the maximum separation between the signal and background samples. It can be shown that the vector $\vec{w}$ maximizes separation between signal and background:

$$\vec{w} = (\Sigma_{sig} + \Sigma_{bkg})^{-1}(\vec{\mu}_{sig} - \vec{\mu}_{bkg}) \tag{6.17}$$

Notice that the above maximizing vector is found using training samples of signal and background. The resulting discriminant for any event with observable $\vec{x}$ is then simply given by:

$$D = \vec{w} \cdot \vec{x} \tag{6.18}$$

6.3.3 Artificial Neural Network

An artificial Neural Network (NN) is a computational model which mimics a biological neural network [68], [69]. It is an adaptive system which "learns" from the input data and generates the corresponding output. A typical neural

network consists of an input layer, one (or several) hidden layers, and an output layer. Each layer consists a number of nodes (which are akin to the neurons in a biological sense). The nodes in different layers are interconnected. Figure 6.7 shows an example of the neural network used as a discriminant.

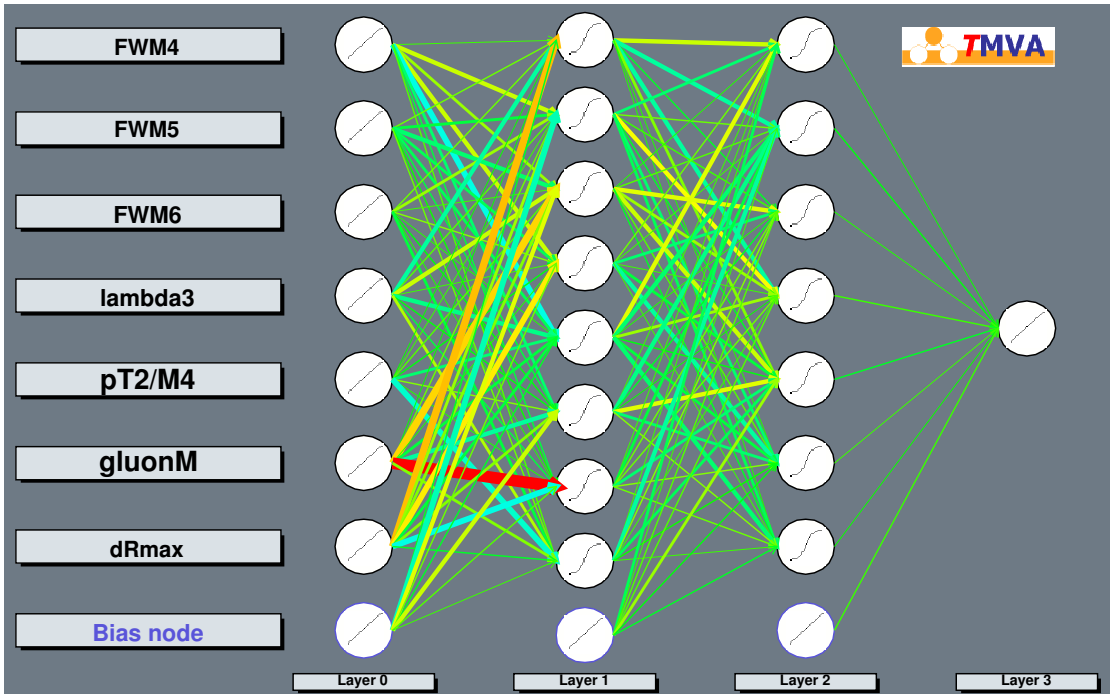

Figure 6.7. An example of neural network being used as a discriminant. Layer 0 corresponds to the input layer and each node corresponds to a candidate variable input. Layer 1 and 2 are the hidden layer which stores the learning results. Layer 3 is the output layer which gives the output value.

Neural Networks need to be trained before using, which involves repeatly feeding equal amounts of signal and background events to the Neural Network to

"learn". After each pass of events the neural network adjusts the parameters of interconnections between the neurons based on the error of the output, using training algorithms such as a back propagation algorithm. The training is done repeatly until the parameters within the NN become stable (usually using 500 training cycles). Both the number of input nodes (number of input variables) and the output node (1 discriminant) are fixed. However the number of hidden layers and the number of nodes within each hidden layer can vary to give different results. After some trial and error it turns out one can get good performance with 2 hidden layers, each having N+1 nodes (where N is the number of input nodes, or variables), and this is what is being used in this analysis.

6.3.4 Evaluation of the discriminant

Only a subset of topological variables mentioned in section 6.1 are required for the discriminant. Numerous different combinations have been tried in order to find the set that gives stable results. For example, the resulting discriminant distribution of both signal and background should be smooth enough without strange looking peaks. Also, for the likelihood case, the background discriminant distribution should not have a spike at $\mathfrak{D} \approx 1$, and similarly the signal discriminant should not have a spike at $\mathfrak{D} \approx 0$ either.

Out of all the combinations considered, the chosen combination of variables should satisfy the following:

1. They have minimal correlation with each other
2. They exhibit good separation between signal and background
3. The simulated distributions match the observed ones in the data

Note that a set of variables that optimizes signal vs. background separation can be defined, but would not satisfy conditions 1 and 3 above. For example, a particular set of variables involving p_{T4}/M_4 and p_{T5}/M_4 gives very good signal efficiency, however the simulated distribution of the event variables, as shown in Figure 6.8, doesn't match well with the observed distribution from the data. This would introduce large systematic modelling errors, so we have chosen a more conservative set of topological variables:

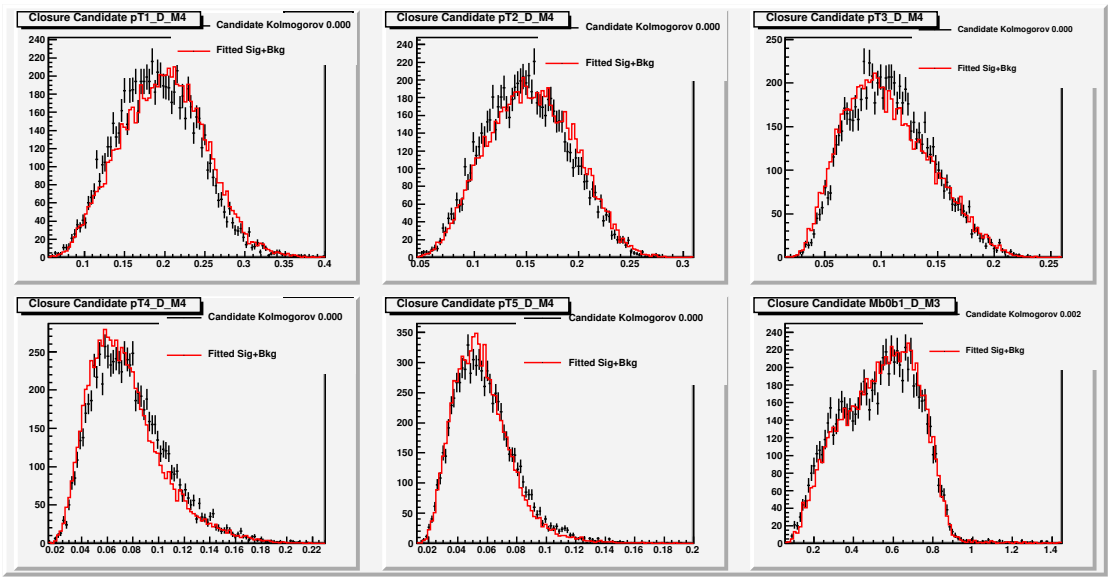

Figure 6.8. An example of a rejected combination of variables for the discriminant. When comparing between simulated distributions (red) and observed distributions (black) of the event variables, the KS (Kolmogorov-Smirnov) probability shows the matches are poor.

$$FWM_4, FWM_5, FWM_6, \lambda_3, p_{T2}/M_4, gluonM, dR_{max} \tag{6.19}$$

An example of the discriminant output from the 4 different discriminants is shown in 6.9.

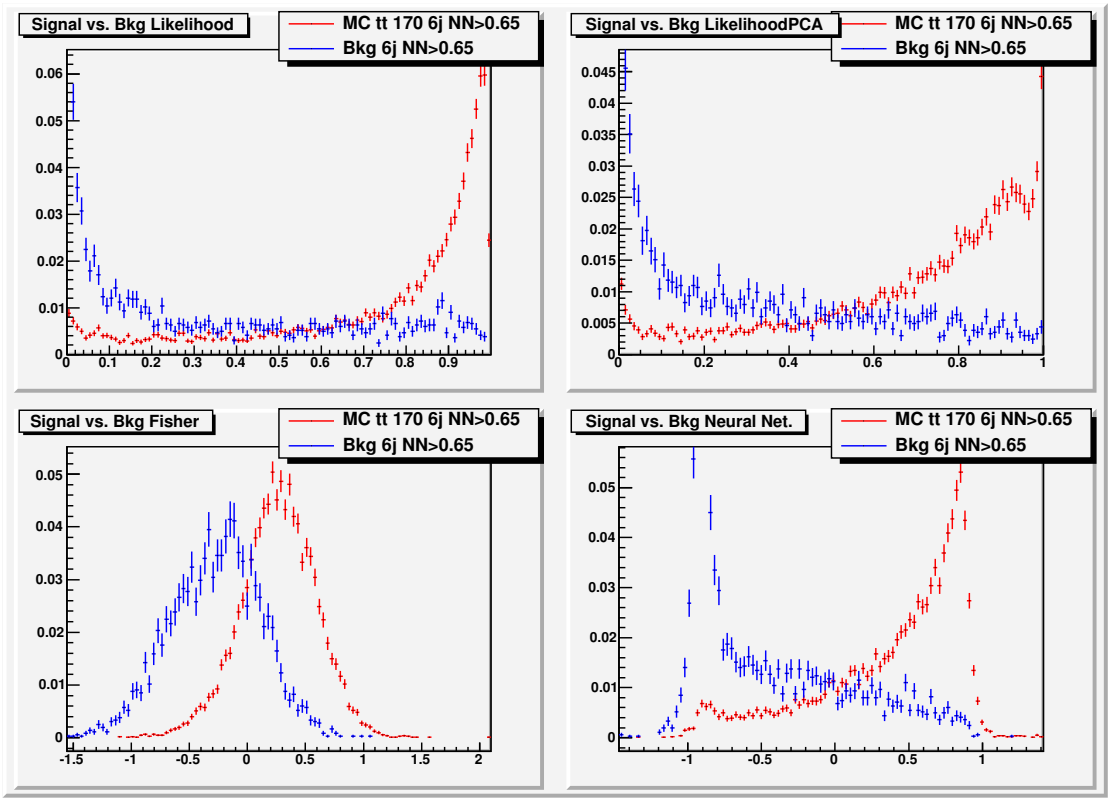

Figure 6.9. Trained discriminant output (normalized) for signal and background from different methods. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, exactly 6 jets, 2 b-jets with BIDNN > 0.65.

To evaluate the output of the discriminant, one feeds another set of signal and

background events (preferably a different set of events from the training events) to the trained method, and then observes the signal efficiency versus the background rejection rate. Figure 6.10, 6.11 and 6.12 shows the plots for the 3 different BID regions: 6 jets $BIDNN > 0.65$, 6 jets $BIDNN 0.2 - 0.65$ and 7+ jets $BIDNN > 0.2$, all using MC $m_{top} = 170$ GeV as signal events. The closer the curve moves towards the point of signal efficiency of 1 and background rejection of 1, the more powerful the discriminant is. From the plots it can be seen that both Fischer and the Neural network (MLP NN) have comparable performance, with the 2 Likelihood methods slightly lagging behind.

6.4 Purity Extraction and mass bias

After the discriminant has been trained, it can then be applied to the data candidate events to obtain a distribution of the discriminant. The resulting distributions can then be fitted with the signal and background discriminant distributions to obtain a measurement of the signal purity. Figure 6.13, 6.14 6.15 shows the fitted purity for the 3 different BID regions: 6 jets $BIDNN > 0.65$, 6 jets $BIDNN 0.2 - 0.65$ and 7+ jets $BIDNN > 0.2$.

From these plots, one can see that the purity values are fairly consistent among the 4 different discriminant methods, except that the purity from the simple Likelihood is lower in the 6 jets $BIDNN 0.2 - 0.65$ and 7+ jets $BIDNN > 0.2$ case.

Note that using MC $t\bar{t}$ $m_{top} = 170$ GeV as signal events may generate results biased towards the same mass value. As a cross check one should apply the trained discriminants on MC signals of different top mass, and then re-fit to extract the candidate purity to see if they are compatible. Figure 6.16 and 6.16 shows the fitted purity using m_{top} of 140 and 200 GeV in the 6 jets $BIDNN > 0.65$ case.

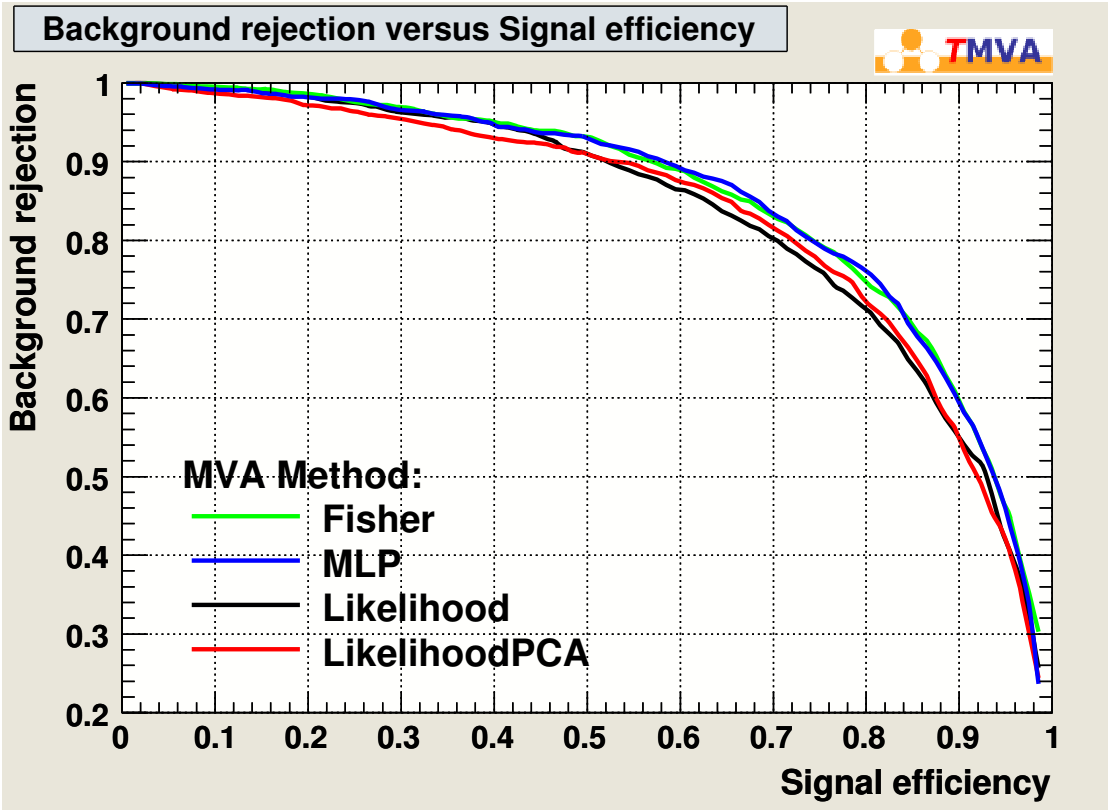

Figure 6.10. Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ m_{top} = 170 GeV, exactly 6 jets, 2 b-jets with BIDNN > 0.65.

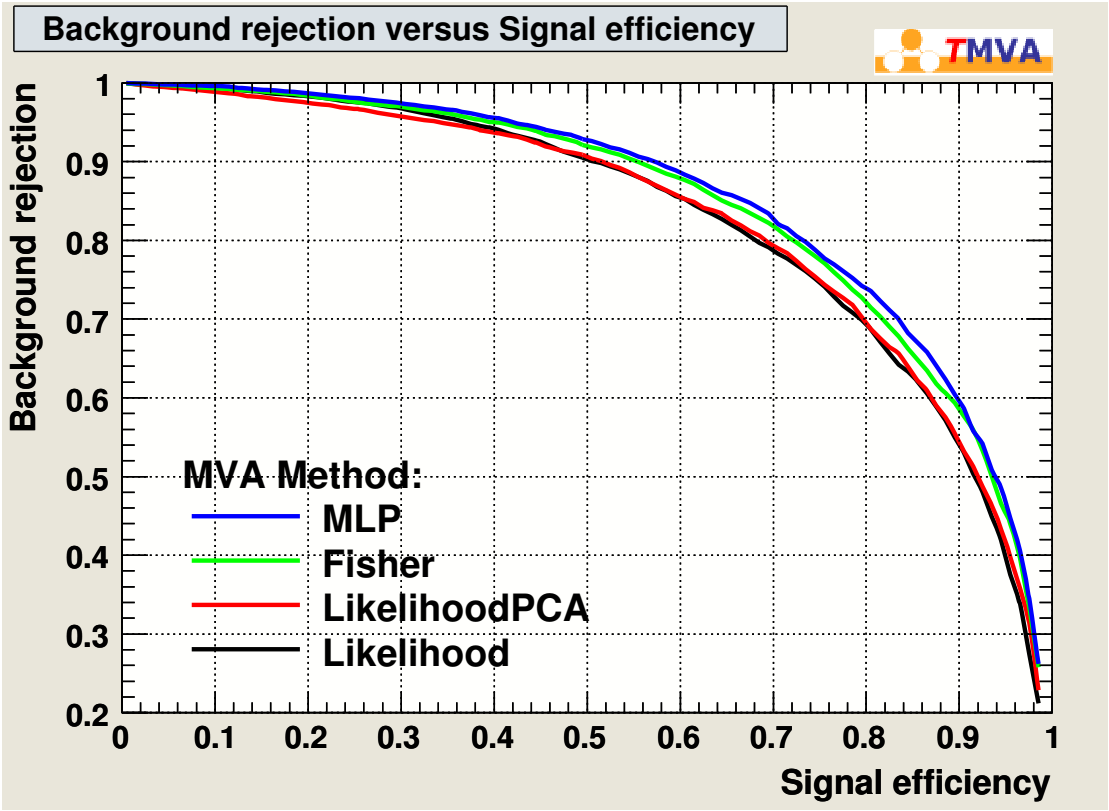

Figure 6.11. Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, exactly 6 jets, 2 b-jets with BIDNN $0.2 - 0.65$.

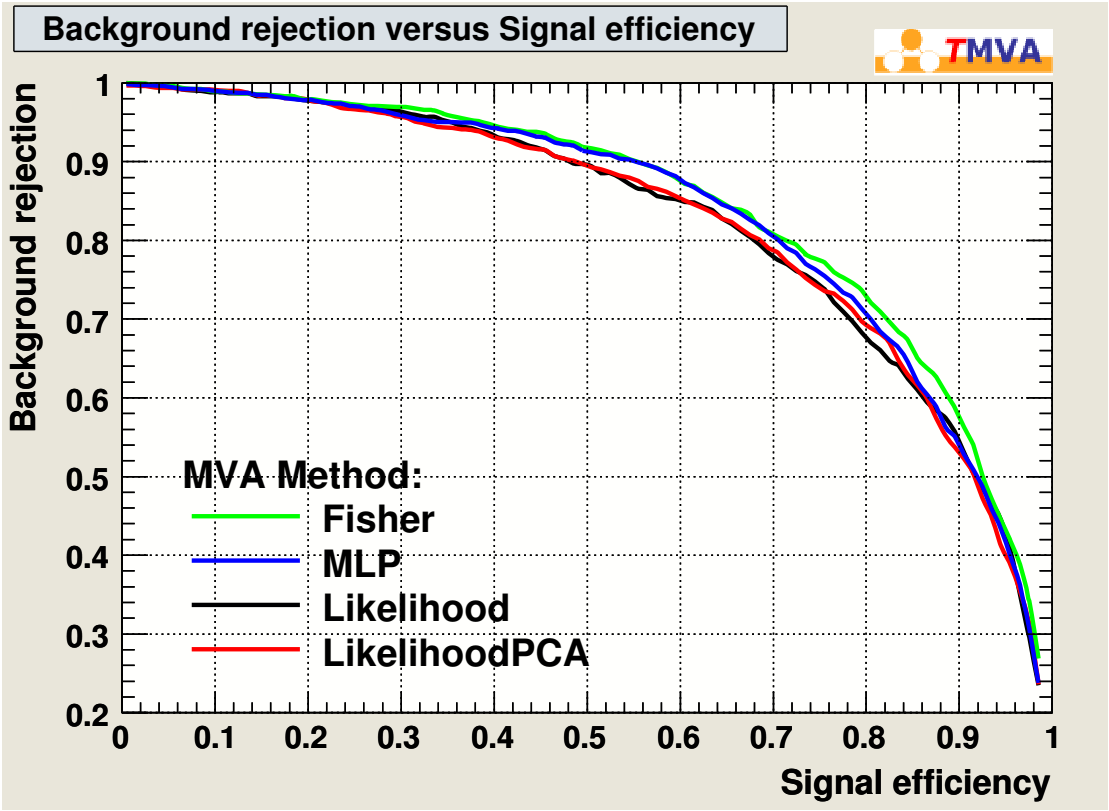

Figure 6.12. Background rejection rate vs. Signal efficiency for different trained discriminant. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV, 7+ jets, 2 b-jets with BIDNN > 0.2.

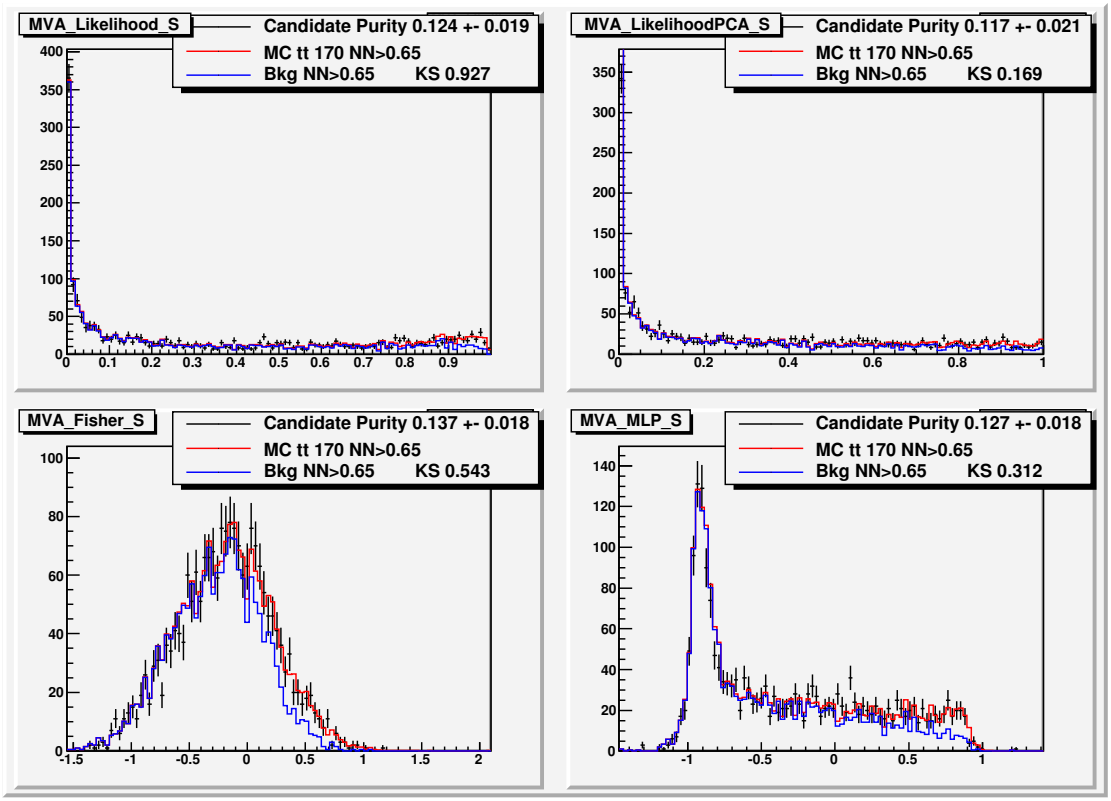

Figure 6.13. Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV.

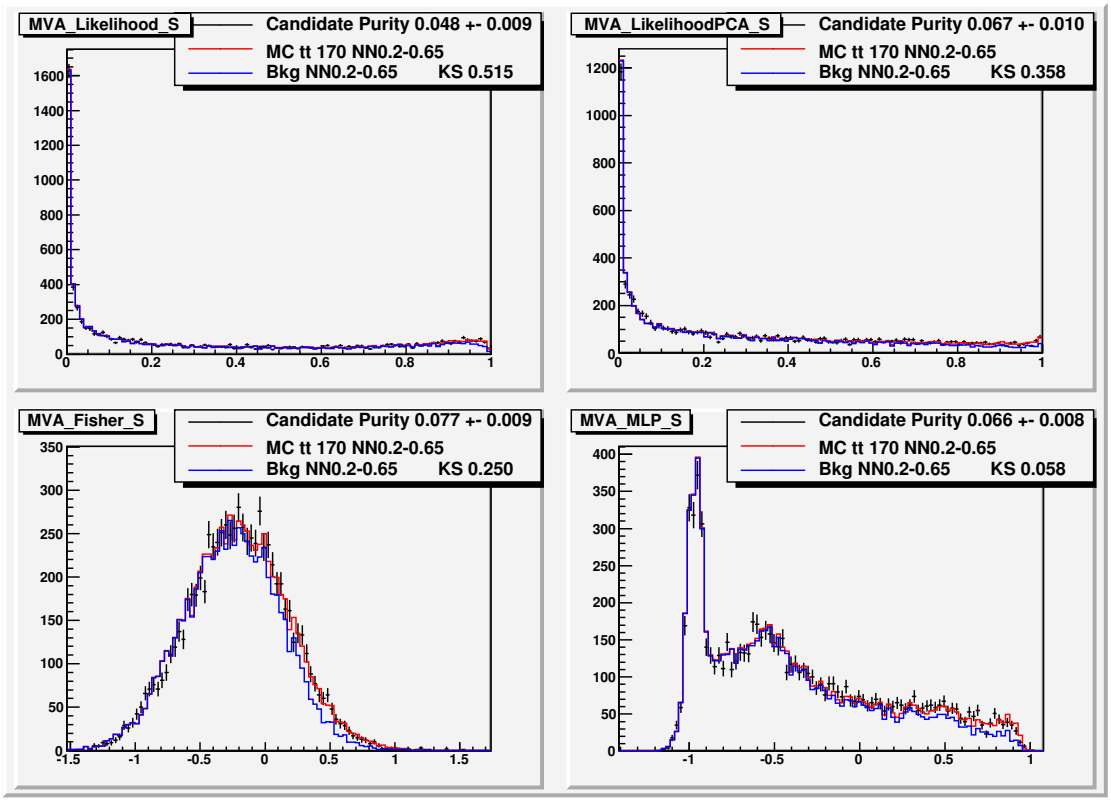

Figure 6.14. Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN $0.2 - 0.65$ case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV.

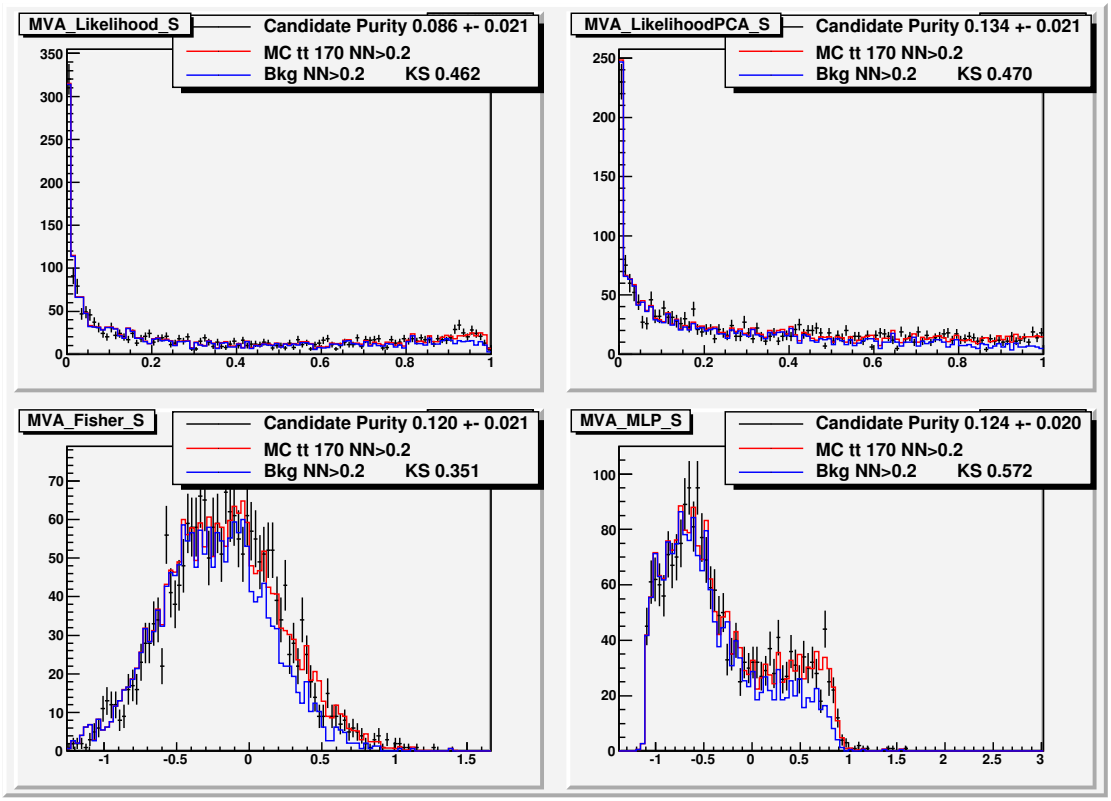

Figure 6.15. Fitted candidate purity using the 4 different discriminant methods for 7+ jets, 2 b-jets with BIDNN > 0.2 case. The signal samples are MC $t\bar{t}$ $m_{top} = 170$ GeV.

Here one can see that the purity values from m_{top} of 140 GeV are higher than the m_{top} of 170 GeV and the purity values from m_{top} of 200 GeV are lower. This is expected because a signal of low m_{top} tends to look more like background (recall the mass templates from the previous chapter), which results in a higher purity as a larger fraction of background will look more like signal. It is exactly the opposite in the case of higher m_{top}. Nevertheless, the purity extracted from both cases are consistent with the purity from m_{top} of 170 GeV within the MC and background statistics.

As an additional check, one can combine MC $t\bar{t}$ events of different m_{top} as signal and see how the result changes. Figure 6.18 shows the fitted purity using a signal consisting of equal fractions of MC $t\bar{t}$ events from $m_{top} = 155 - 185$ GeV events. As can be seen from the plots the results are very close to the $m_{top} = 170$ GeV case.

Given the above results, the Fisher and Neural Network methods are preferred over the Likelihood methods because they have better signal vs. background efficiencies. However the Neural Network discriminant distributions are not entirely smooth and have small peaks between large signal and background peaks. Therefore the Fisher discriminant is the chosen method and table 6.4 summarizes the fitted candidate purity.

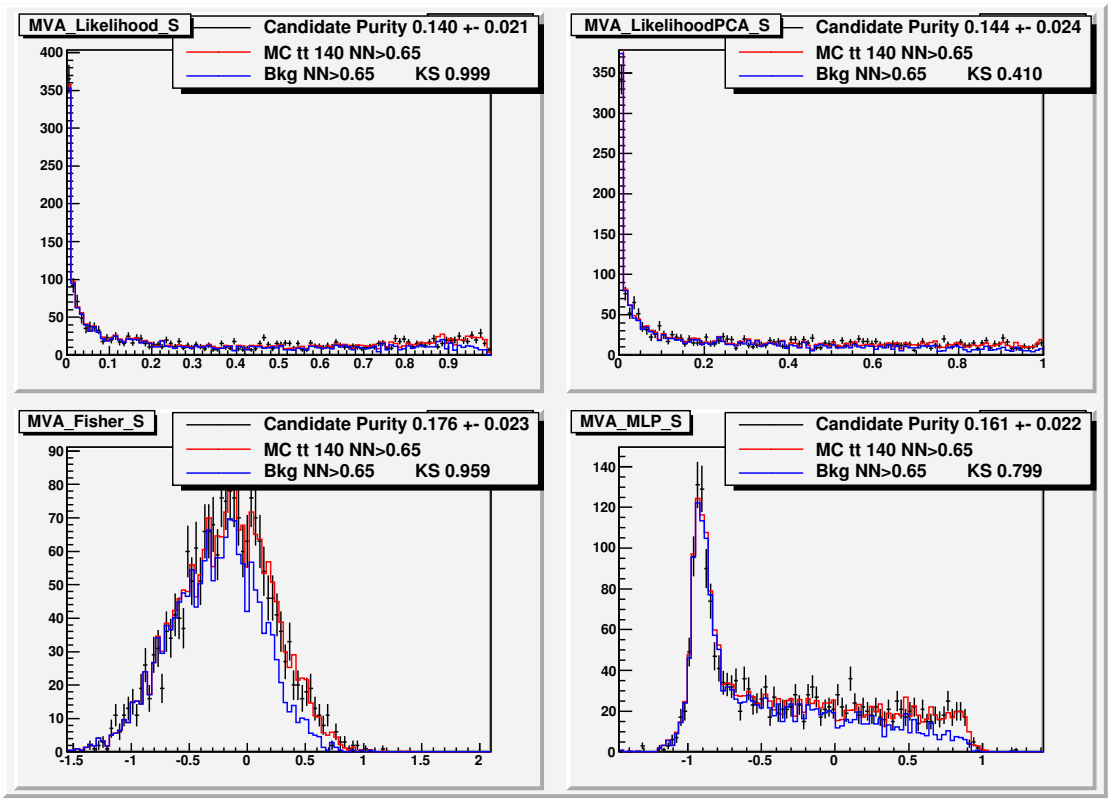

Figure 6.16. Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 140$ GeV.

TABLE 6.1

Fitted candidate purity using Fisher discriminant

Sample	candidate purity
6 jets BIDNN > 0.65	0.137 ± 0.018
6 jets BIDNN $0.2 - 0.65$	0.077 ± 0.009
7+ jets BIDNN > 0.2	0.120 ± 0.021

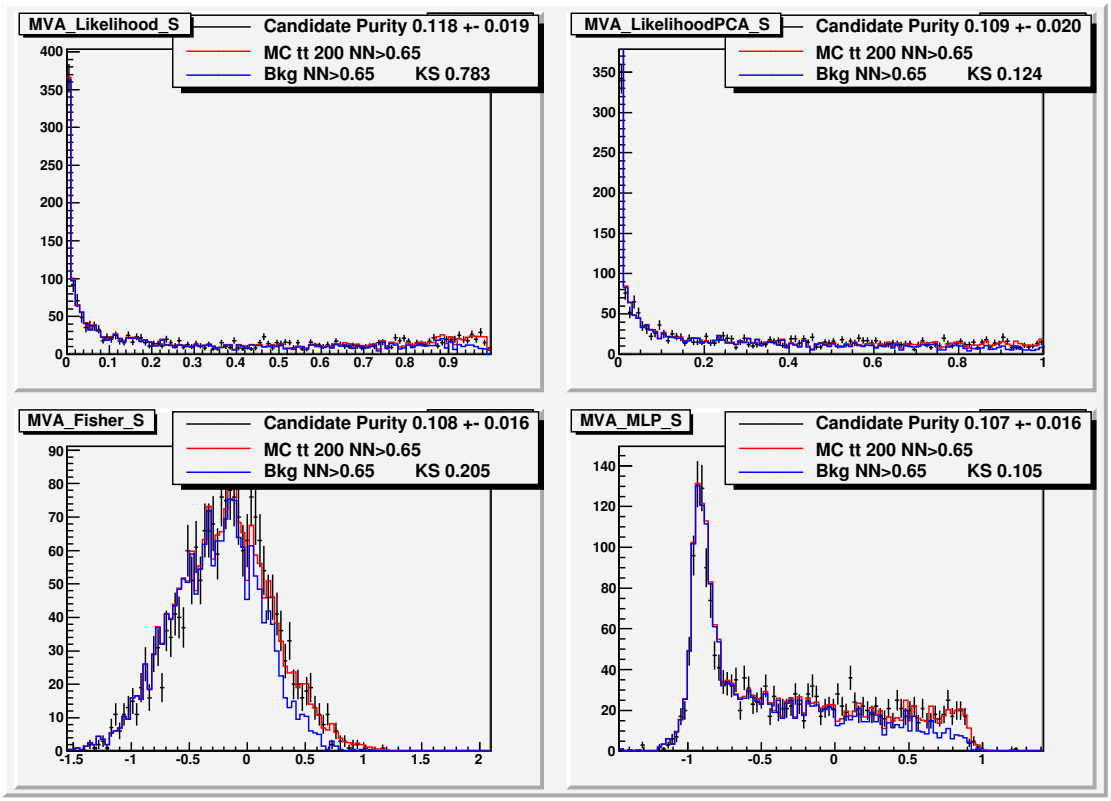

Figure 6.17. Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are MC $t\bar{t}$ $m_{top} = 200$ GeV.

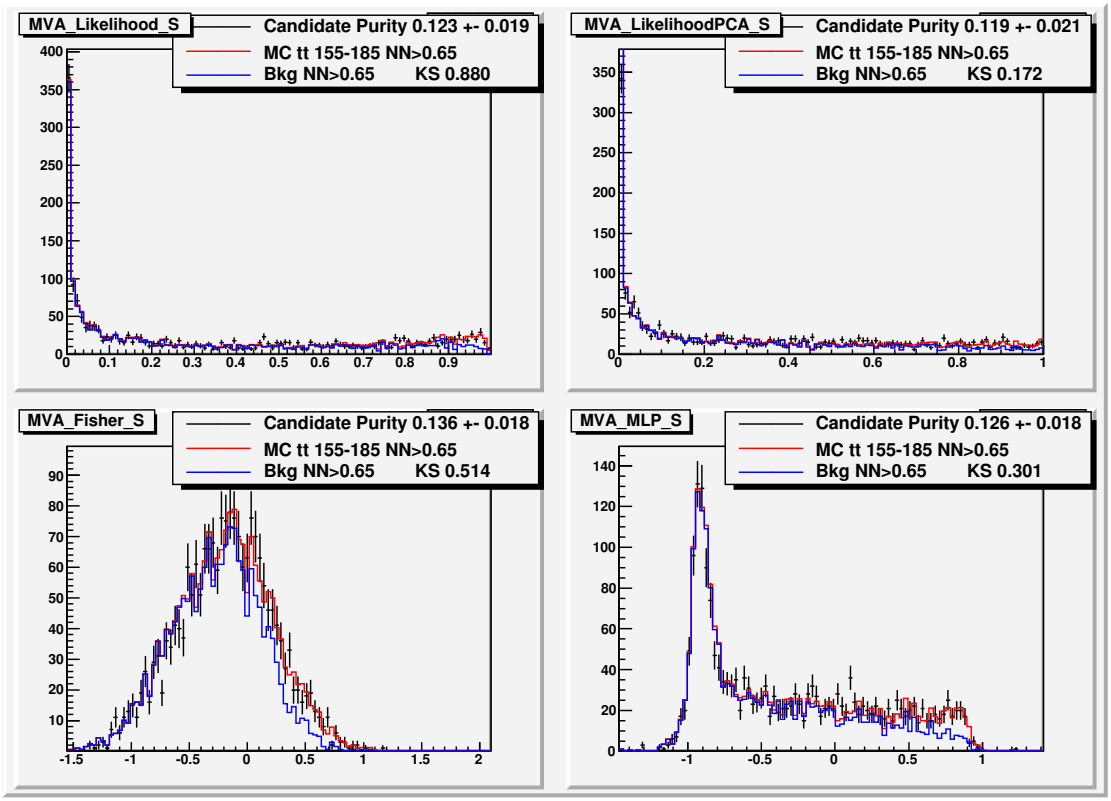

Figure 6.18. Fitted candidate purity using the 4 different discriminant methods for exactly 6 jets, 2 b-jets with BIDNN > 0.65 case. The signal samples are combined MC $t\bar{t}$ with $m_{top} = 155$ to 185 GeV.

CHAPTER 7

LIKELIHOOD METHOD AND ENSEMBLE TEST

7.1 Likelihood method

Given the expected signal and background templates in the all-jets channel, together with the expected purity from the likelihood fit of the candidate events, the top mass in the all hadronic channel can now be extracted. Since the signal templates are continuous functions of the top mass, this allows the derivation of a event-by-event likelihood to extract the most probable top mass.

Since the candidate sample consists of 3 different processes: signal events with correct bjj/jj jet combinations, signal events with wrong jet combinations or not parton matched, and background events, the probability of a event with $Dmass(bjj)$ and m_{top} is therefore:

$$
\begin{aligned}
P_{evt}(Dmass(bjj), purity_{signal}; m_{top}, f_{correct}) = \\
& purity_{signal} * f_{correct} * Sig_{correct}(Dmass(bjj); m_{top}) \\
& + purity_{signal} * (1 - f_{correct}) * Sig_{wrong}(Dmass(bjj); m_{top}) \\
& + (1 - purity_{signal}) * Bkg(Dmass(bjj))
\end{aligned} \tag{7.1}
$$

Here $purity_{signal}$ is the signal purity in the candidate sample from the likelihood

fit. $f_{correct}$ is the fraction of correct bjj/jj jet combinations within the signal process. $Sig_{correct}$ and Sig_{wrong} are the signal template functions for correct and wrong (with not parton matched events) jet combinations respectively. Bkg is the background Spline function fit.

A candidate sample consists of a certain number of events, each providing two measurements of $Dmass(bjj)$. The probabilities for each measurement of $Dmass(bjj)$ are then multiplied together to generate the Likelihood. It is natural to represent the results as $-Ln(Likelihood)$ because this allow results from different events to be simply added. The most probable mass is then simply the minimum point of $-Ln(Likelihood)$, and the mass error is simply the the mass value where $-Ln(Likelihood_{min}) + 1$.

$$\begin{aligned} -ln(L(m_{top}, f_{correct})) &= -ln\left(\prod_{evt} P_{evt}(Dmass(bjj), purity_{signal}; m_{top}, f_{correct})\right) \\ &= \sum_{evt} -ln(P_{evt}(Dmass(bjj), purity_{signal}; m_{top}, f_{correct})) \end{aligned} \tag{7.2}$$

Notice that the fraction of correct jet combinations in the signal can also be allowed to vary. We shall experiment with both fixing $f_{correct}$ to the value from MC (0.248 if $dR(parton, jet) < 0.5$ is required for parton matching) and allow it to float to see if the mass measurements differ.

7.2 Ensemble test

Before applying the template function to the data candidates, it is necessary to test the validity and the stability of the template function to see if it will give

the correct answer. This can be accomplished by performing an ensemble test - which is basically simulating the same experiments many times.

An ensemble is basically a collection of pseudo-experiments. A pseudo-experiment can be carried out by combining randomly picked simulated signal and background events. To coincide with the real experiment, the same selection cuts are applied. For a more realistic simulation, the number of total events in each pseudo-experiment is not constant within the ensemble, but the signal and background portion are allowed to vary separately corresponding to a Poisson probability distribution:

$$p(N_{sig}; N_{cand} * purity_{signal}) = \frac{e^{N_{cand} * purity_{signal}} (N_{cand} * purity_{signal})^{N}_{sig}}{N_{sig}!} \tag{7.3}$$

$$p(N_{bkg}; N_{cand} * (1 - purity_{signal})) = \frac{e^{N_{cand} * (1 - purity_{signal})} (N_{cand} * (1 - purity_{signal}))^{N}_{bkg}}{N_{bkg}!} \tag{7.4}$$

Where N_{evt} is the number of candidate events. N_{sig} and N_{bkg} are the fluctuating number of signal and background events in each pseudo-experiment.

Performing Ensemble tests is the equivalent of repeating the experiment for a certain number of times (in this case 1000) and seeing how the results are distributed statistically. In addition, there may exist systematic shifts of the results and therefore the result of the ensemble test can be used as calibration of the method.

In this study MC ttbar events of different m_{top} ranging from 155-200 GeV and JES Smeared $0, \pm 1\sigma$ are combined with the events from the background model to

generate the ensembles. Again they are mixed according the fitted signal purity and the expected fraction of correct jet combinations.

7.2.1 2-D Likelihood test

First, allowing both m_{top} and $f_{correct}$ to vary results in a 2-dimensional likelihood surface as shown in Figure 7.1. The minimum point of the surface is located to give the measurement of m_{top} and $f_{correct}$. Since the derivation of +1 in likelihood corresponds to 1 standard derivation, the mass error can be obtained by finding the mass value corresponds to the minimum point +1 along the mass axis, and similarly for the error in $f_{correct}$.

The extracted results of all the pseudo-experiments within an ensemble should yield a normal distribution according to the central limit theorem. Figure 7.2 shows the distribution of the extracted mass value from ensembles of different top mass.

The mean of the m_{top} distribution should be calibrated to yield the input m_{top}. Figure 7.3 shows the mean of the extracted m_{top} versus the input m_{top} using standard JES, together with polynomial fits of different orders. As shown in the figure it turns out a linear fit is sufficient for calibration. The ensemble test will then be needed to rerun, applying the inverse fit function (which is the calibration) on the mass value and error from each pseudo-experiment.

Besides the mean value of the mass, the error of the mass returned from individual pseudo-experiments should also agree with the width of the mass distribution. This can be checked by looking at the pull distribution:

$$x_{pull}(i\text{th experiment}) = \frac{x(i\text{th experiment}) - x_{mean}}{\delta x(i\text{th experiment})} \tag{7.5}$$

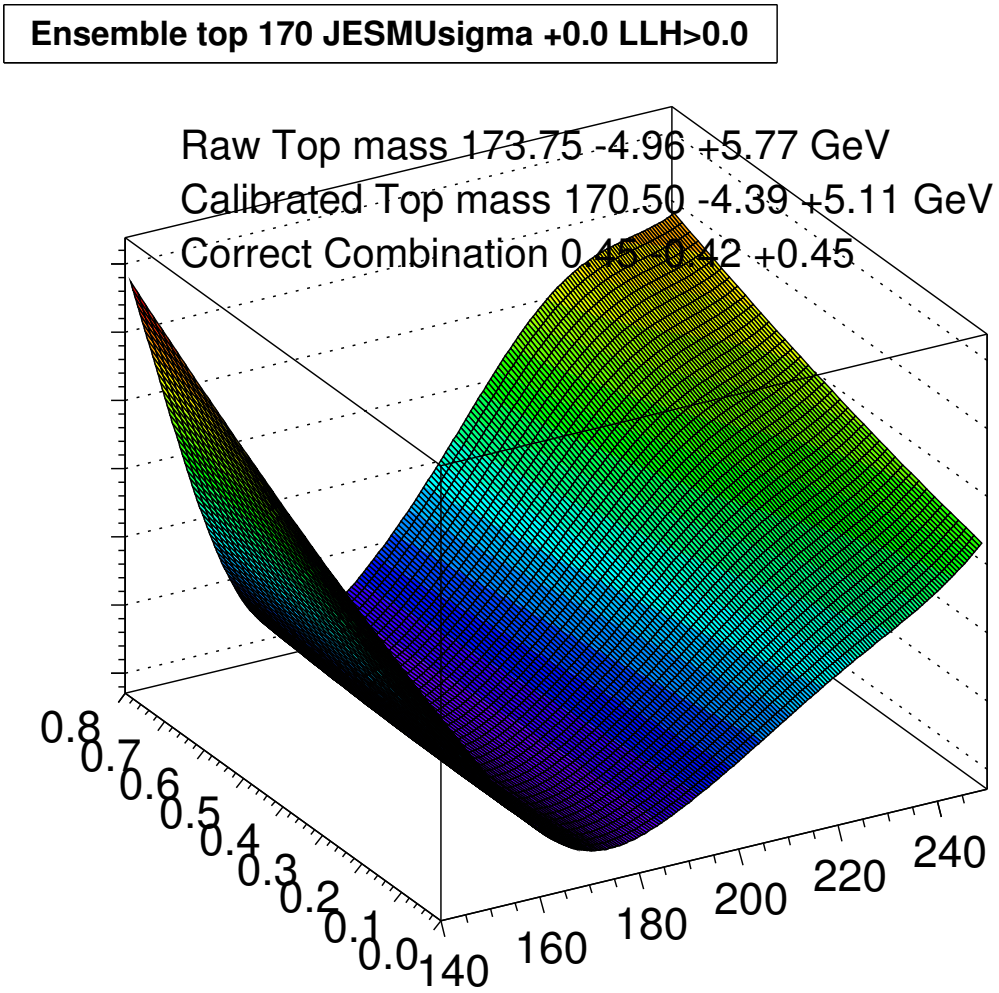

Figure 7.1. Log Likelihood surface as a function of m_{top} and $f_{correct}$. This is a pseudo-experiment with signal purity of 18% by combining MC $m_{top} = 170$ GeV, standard JES and background events. The shown m_{top} and $f_{correct}$ are obtained from the minimum point of the Likelihood

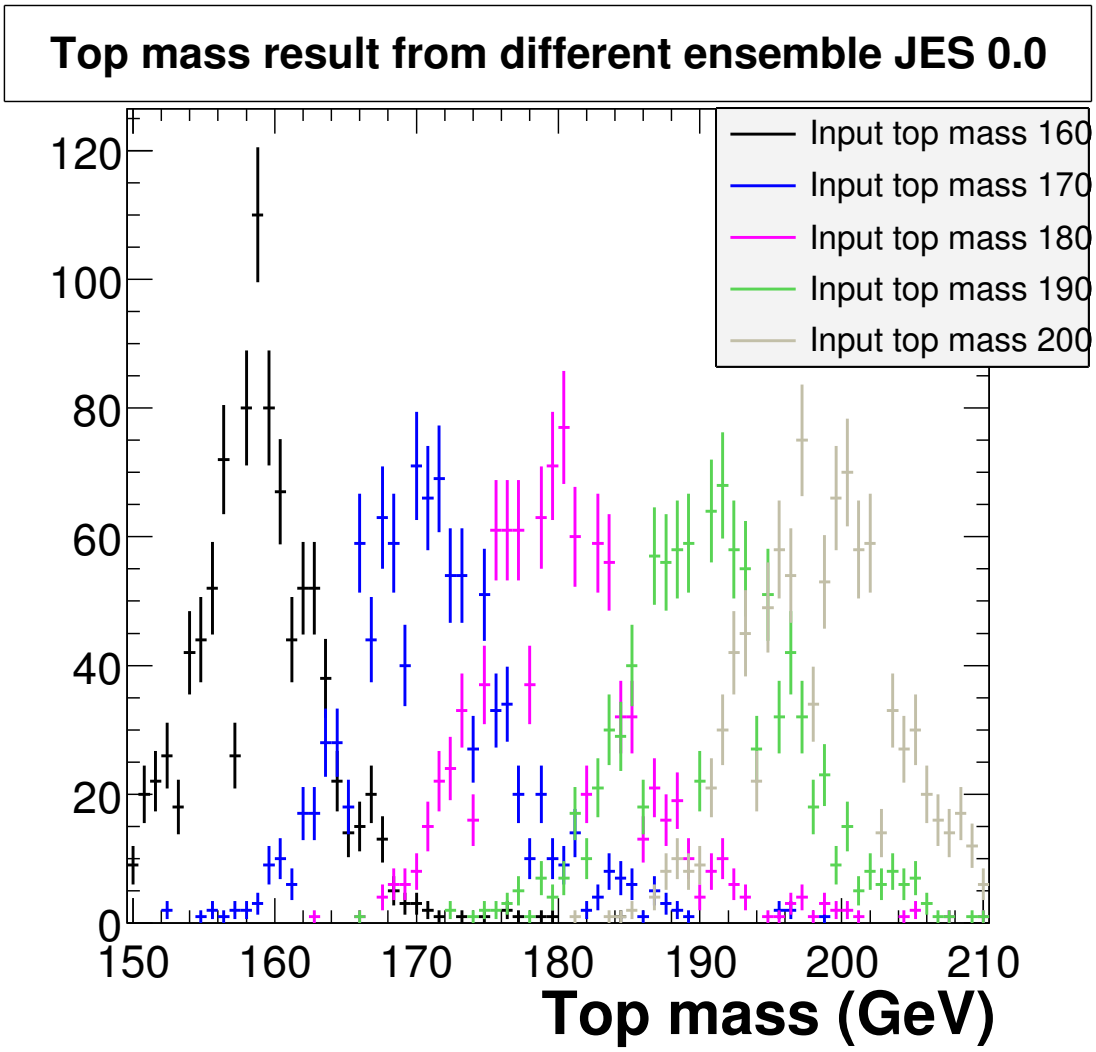

Figure 7.2. Extracted m_{top} from log likelihood using ensembles of different input m_{top} with standard JES. Different colors corresponds to different input top mass as shown, and each ensemble has 1000 pseudo-experiment.

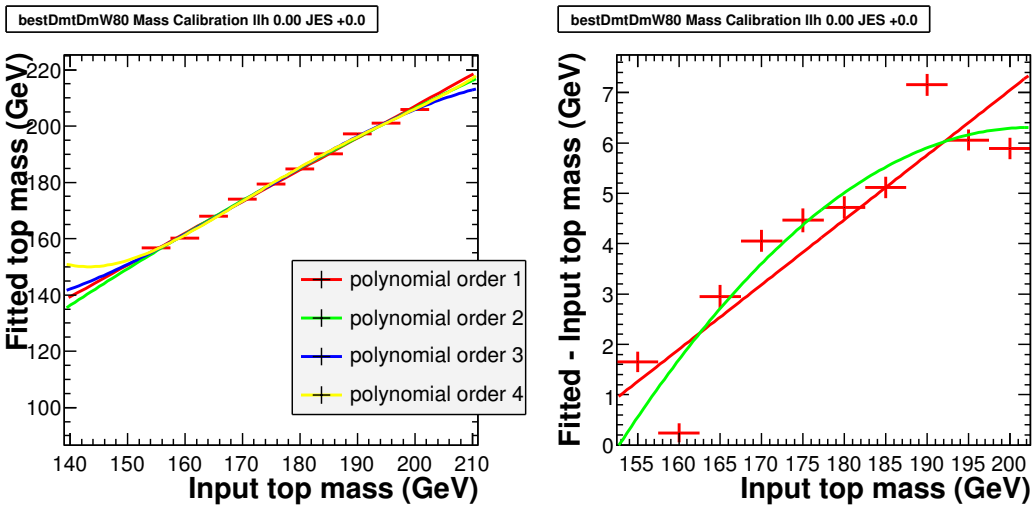

Figure 7.3. Mean extracted m_{top} versus input m_{top} using standard JES. Different color lines correspond to polynomial fits from order 2 to order 5. In the ideal case the line has slope 1 with no offset. The plot on the right shows (Mean extracted - input m_{top}) instead, which really shows the derivation of the extracted mass from input mass. It can seen that a linear fit (red line) is sufficient to bring the extracted mass back to input mass

Where $\delta x(i\text{th experiment})$ is the error of the variable x from the i-th pseudo-experiment . Ideally the pull distribution shall coincide with a standard normal distribution of $\sigma = 1$, which shows the mass error from each pseudo-experiment is correct. Figure 7.4 shows the pull mass distributions from the 2D ensemble test and they are all normally distributed.

As mentioned, the ensembles test will need to be repeated again with calibration applied. Also, ensembles of different JES will also be tested using the same calibration. Figure 7.5 compares the calibrated extracted mass in 3 different JES. It can be shown that the difference due to different JES is around 0.8 GeV/c^2 and will be used as systematic error. The mass pull value from the different JES is a little bit larger than 1 (around 1.1-1.2). This is likely affected by the low signal purity of our sample as will be investigated further in a later section about toy statistic experiments.

In addition, the mass error plot for the standard JES case is shown in figure 7.6. This gives an estimate on the size of statistical error expected.

In addition to the mass value, the fraction of correct bjj/jj jet combination, or $f_{correct}$, can also be extract from the 2D likelihood. Figure 7.7 shows the $f_{correct}$ values from the standard JES with different m_{top}. The values are restricted between 0.0 to 0.8. From the plot it is shown that due to the low signal purity of the sample it is not possible to obtain a measurement of $f_{correct}$ from the data. Therefore the $f_{correct}$ is fixed from now on using the MC values in table 5.2.

7.2.2 1-D Ensemble test

By fixing $f_{correct}$ one can obtain a simpler 1D likelihood in m_{top}. Figure 7.8 shows the resulting 1D log likelihood curve in m_{top}. Again the minimum point of

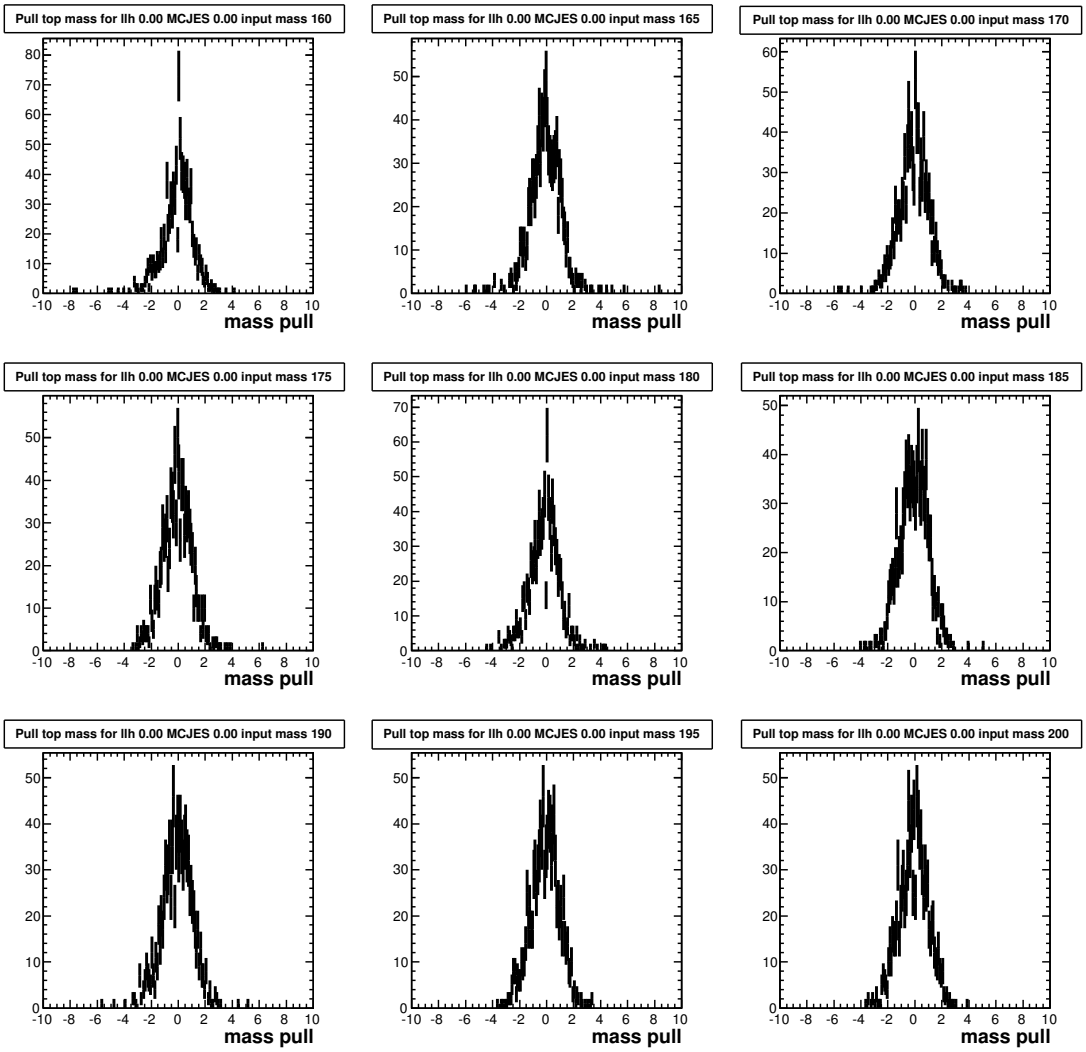

Figure 7.4. Pull mass distribution from different input m_{top} (160-200 GeV) after calibration using standard JES.

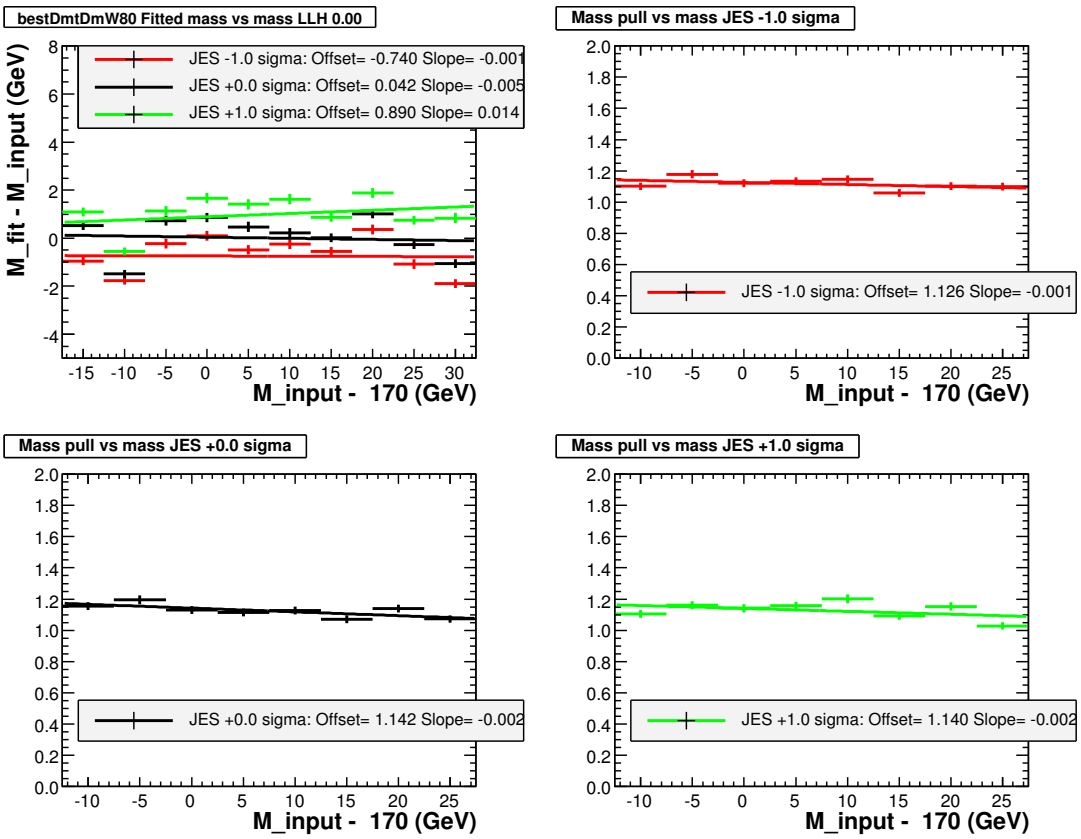

Figure 7.5. Ensemble test results with calibration applied, using standard JES and $+/-1\sigma$ JES as well. The top left plot compares the resulting (mean extracted - input m_{top}) from 3 different JES. The rest of of the plots shows the mass pull values from 3 different JES.

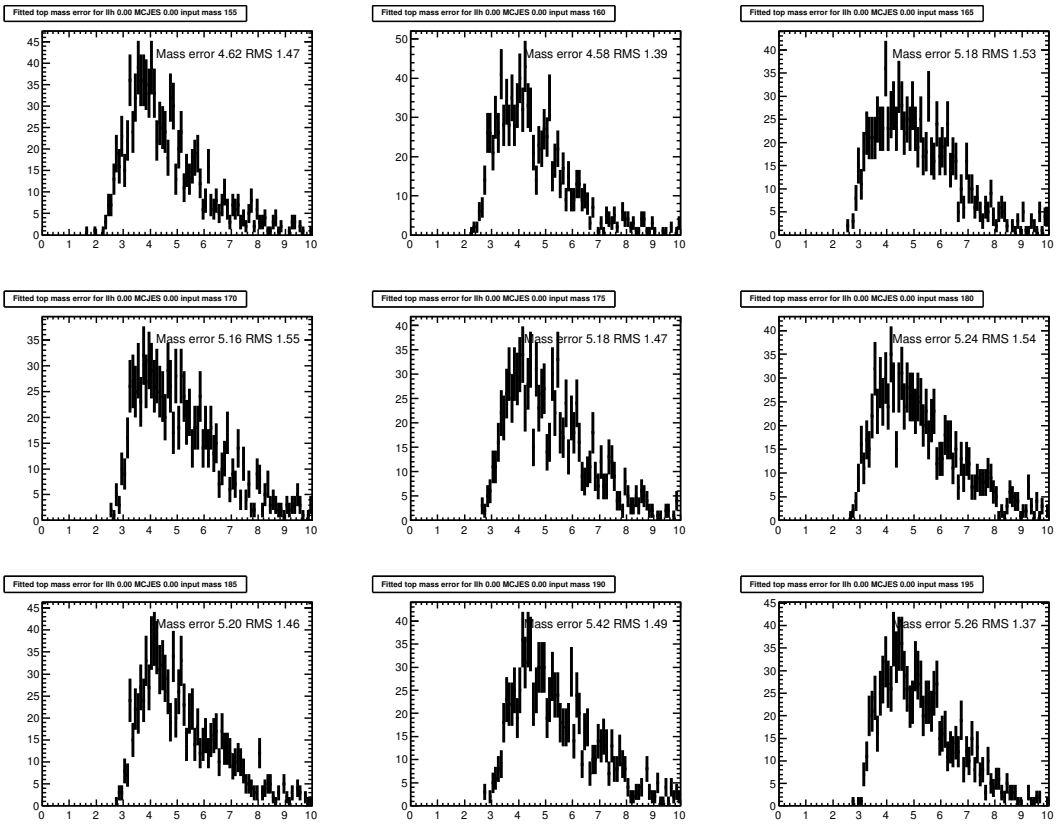

Figure 7.6. Ensemble test mass error with calibration applied using standard JES. Different plots correspond to different input m_{top} from 160 to 200 GeV.

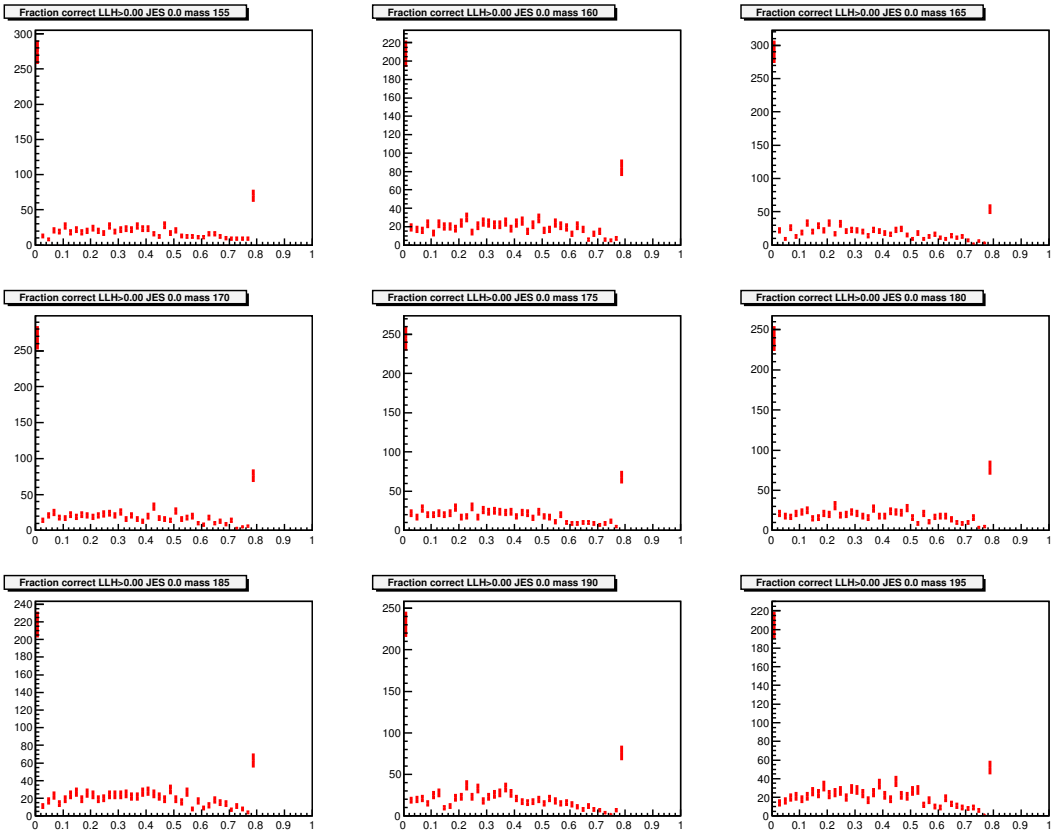

Figure 7.7. Ensemble test $f_{correct}$ with calibration applied using standard JES. Different plots correspond to different input m_{top} from 155 to 195 GeV. Note the spike around 0.0 and 0.8 is simply because there is no local minimum $f_{correct}$ found in the likelihood, and thus those 2 points should be discarded.

the likelihood curve is located to give the measurement of m_{top}.

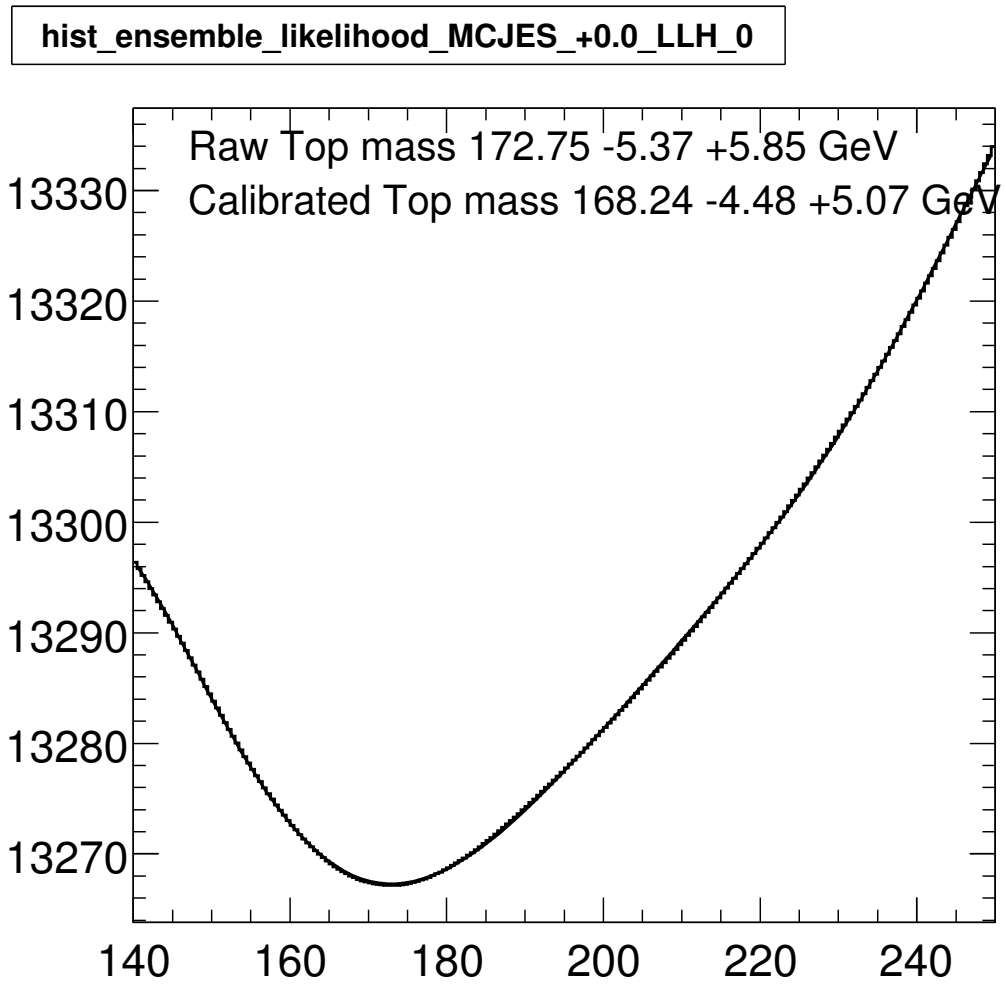

Figure 7.8. Log Likelihood surface as a function of m_{top}. This is a pseudo-experiment with signal purity of 18% by combining MC $m_{top} = 170$ GeV, standard JES and background events. $f_{correct}$ is fixed to the MC value. The shown m_{top} is obtained from the minimum point of the Likelihood

The exact same procedure is used to calibrated the mass in the 1D case. Figure 7.9 shows the mean of the extracted m_{top} versus the input m_{top} using standard JES and a 1D likelihood. It turns out a quadratic fit works better for calibration in this case.

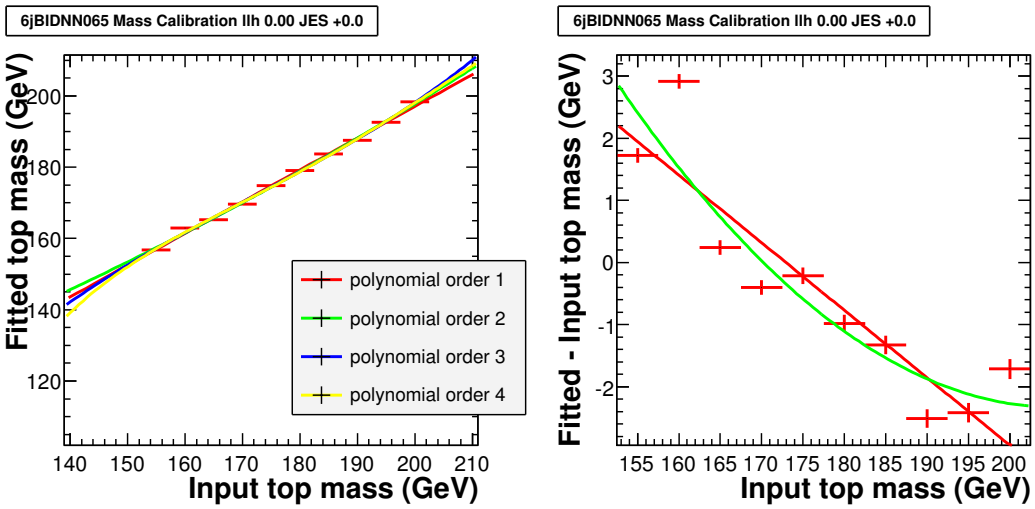

Figure 7.9. Mean extracted m_{top} versus input m_{top} using standard JES. Different color lines correspond to polynomial fits from order 2 to order 5. In the ideal case the line has slope 1 with no offset. The plot on the right shows (Mean extracted - input m_{top}) instead, which really shows the derivation of the extracted mass from input mass. It can seen that a quadratic fit (green line) works better to bring the extracted mass back to the input mass

The results from the repeated ensembles tests comparing ensembles of 3 different JES are shown in Figure 7.10. It can be shown that the difference due to different JES is around 0.75 GeV/c^2 and is slightly smaller than the 2D case.

Similarly the mass pull value from the different JES is a little bit larger than 1 (around 1.1-1.2) as in the 2D case.

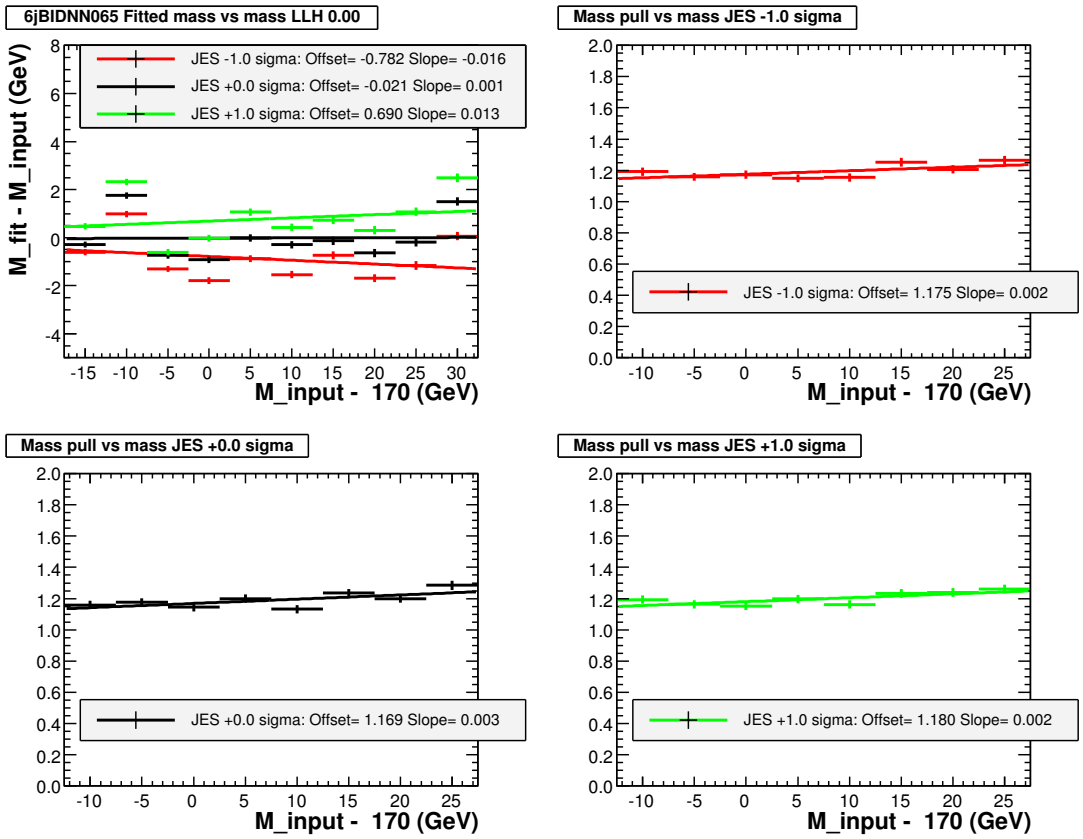

Figure 7.10. Ensemble test results with calibration applied in the 1D likelihood, using standard JES and $+/-1\sigma$ JES as well. The top left plot compares the resulting (mean extracted - input m_{top}) from 3 different JES. The rest of of the plots shows the mass pull values from 3 different JES.

The mass error plot for the standard JES case is shown in figure 7.11. Compared with the 2D likelihood case both the mean and width of the mass error

distributions from the 1D likelihood are slightly smaller. This is expected because in the 1D case there is one less degree of freedom and therefore the mass values are more constrained, hence a smaller error.

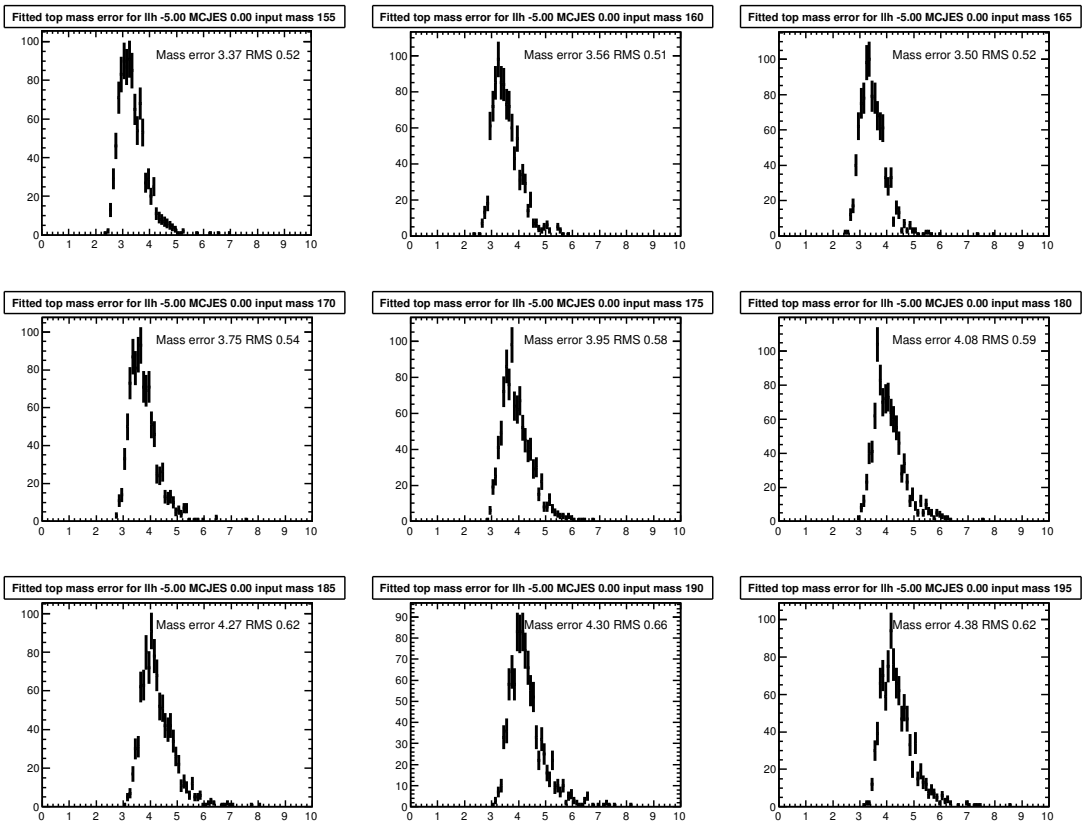

Figure 7.11. Ensemble test mass error with calibration applied using standard JES and 1D likelihood. Different plots correspond to different input m_{top} from 160 to 200 GeV.

7.2.3 Toy statistics experiment

As can be seen from the 2D Ensemble test results above, the width of the mass pull distribution fluctuates around 1.2-1.4, even when the mean mass values show excellent agreement with the expected value. Ideally the width of the mass pull should be close to 1 and therefore it is necessary to understand such discrepancies. Here a "toy" statistics experiment will be performed. Assuming an imaginary mass measuring experiment where the signal distribution is a 1-D Gaussian distribution with the mean as the expected mass value, and the background is a fixed LogNormal distribution. Then similar to the ensemble test above, pseudo-experiments are generated by combining randomly picked signal and background "events" from the corresponding probability distribution. The template function will simply be the sum of the LogNormal and Gaussian distribution, which will then being used to fit the pseudo-experiment to extract the mass.

The idea of the this statistics experiment is to see how the width of the mass pull changes with different numbers of ensembles, and different numbers of pseudo-experiments within an ensemble. First, assuming a fixed number of 500 pseudo-experiments within each ensemble, Figure 7.12, shows the width of mass pull for signal purity 0.20.

As can be seen, with ensemble size of 50 and 100 the width of the mass pull is fluctuating. In fact the fluctuations can be more severe if there are even less statistics as the number of pseudo-experiments per ensemble decreases. Therefore it is more desirable to use an ensemble size of 300 or 500 to obtain stable results. With a fixed number of 500 pseudo-experiments, the width of mass pull is within error to 1.

However, when the size of pseudo-experiments decreases, such reduction of

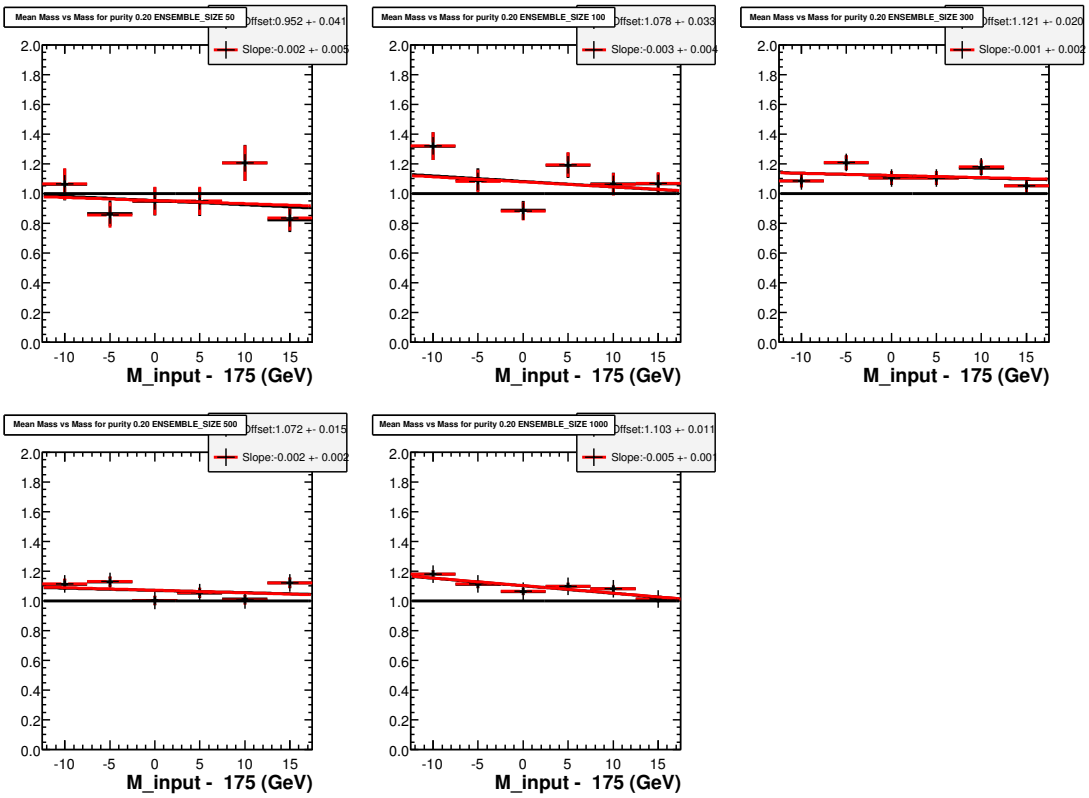

Figure 7.12. Width of mass pull in toy experiment with signal purity = 0.20 and 500 events from top left to bottom right are from different Ensemble size. The black solid line is the linear fit and the dashed line is the ideal value.

statistics per mass bin results in an increases. Further reduction of signal purity will also enhance the increases in the width of mass pull. Figure 7.13 show the width of mass pull for signal purity 0.20 with only 300 pseudo-experiments. The width of the mass pull increases to around 1.35. As the number of pseudo-experiments is further decreased to 100 as in figure 7.14, such inscrease of width of mass pull occurs even for a pure signal case (purity of 1.00)

In conclusion, these statistical tests illustrates that if there are not enough statistics per mass bin in the ensemble test, the width of the mass pull will increase correspondingly. In previous section the widths of the mass pull from ensemble tests are around 1.2, which indicate the low signal purity of the samples we are dealing with in this analysis.

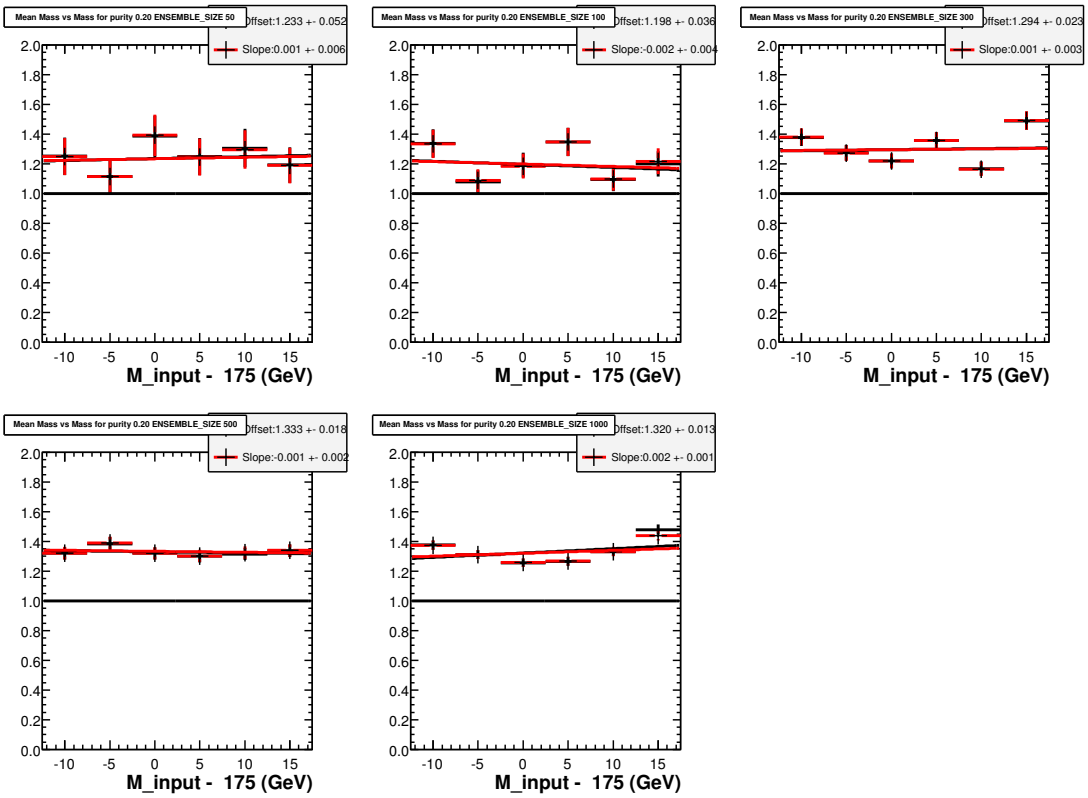

Figure 7.13. Width of mass pull in toy experiment with signal purity = 0.20 and 300 events from top left to bottom right are from different Ensemble size.] The black solid line is the linear fit and the dashed line is the ideal value.

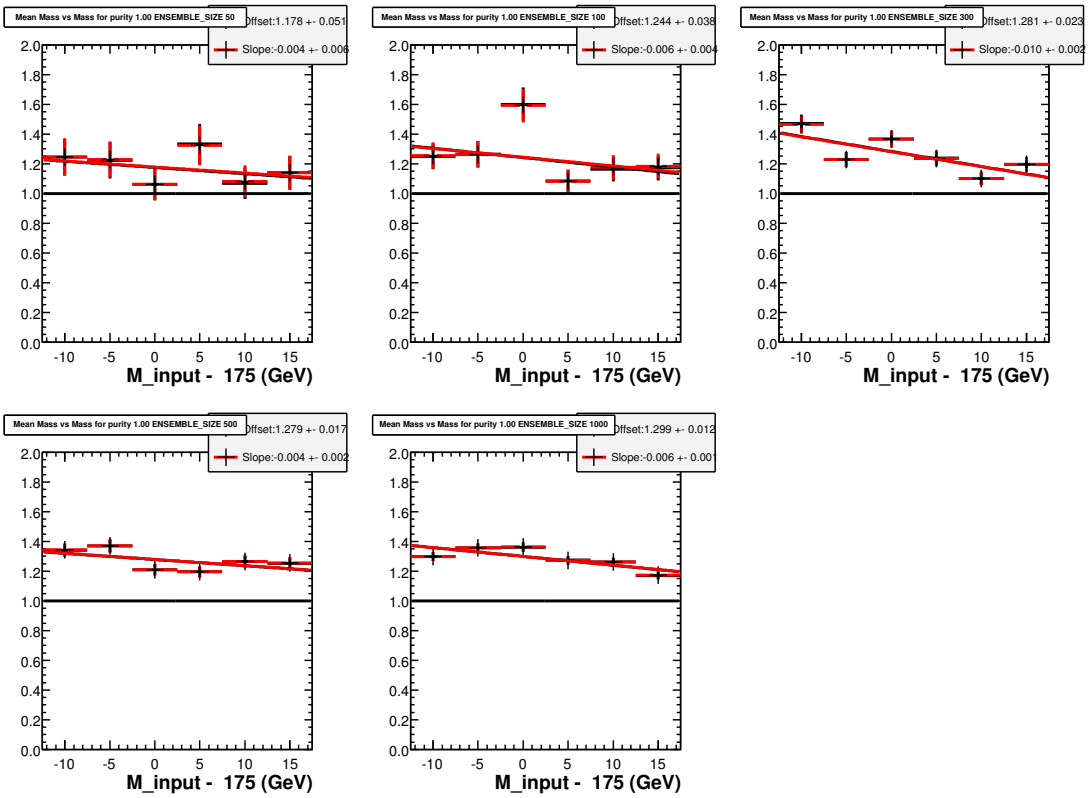

Figure 7.14. Width of mass pull in toy experiment with signal purity = 1.00 and 100 events from top left to bottom right are from different Ensemble size. The black solid line is the linear fit and the dashed line is the ideal value.

CHAPTER 8

SYSTEMATICS AND STATISTICAL STUDIES

For any scientific experiments, error determination is a necessary step to estimate the validity of the experimental results. Given the complexity of this analysis, there are a handful of systematics sources which needed to be addressed and will be discussed in details below. As one shall see, most of the systematics investigations involve the use of the ensemble test.

8.1 Likelihood Cut

Before we investigate the error, it is also interesting to see whether the discriminant developed in Chapter 6 is helpful in improving our top mass measurement. As shown in the likelihood plots, increasing the likelihood cuts increases the signal purity of the samples. However, such cuts will also decrease the number of events in the sample as well, as shown in table 8.1.

To determine which discriminant (LLH) cut gives the best answer, one can look at the size of the expected statistical error on the mass. This statistical error can be estimated by running ensemble tests with different discriminant cuts.

Figure 8.1 shows the statistical mass errors dependence on the Fisher discriminant cut, for the exactly 6 jets, 2 BIDNN > 0.65 case before mass calibration. From the plots it looks like a cut of Fisher discriminant around -0.1 will give the

TABLE 8.1

Effects of Fisher discriminant cut on the signal purity and number of events in the candidate samples

	6 jets, 2 BIDNN > 0.65		7+ jets, 2 BIDNN > 0.2	
Fisher Discriminant cut	Purity	Events	Purity	Events
No cut	0.137	2041	0.120	2117
LLH > -0.9	0.144	1940	0.124	2047
LLH > -0.7	0.156	1803	0.133	1925
LLH > -0.5	0.180	1574	0.152	1681
LLH > -0.3	0.219	1245	0.189	1298
LLH > -0.1	0.292	857	0.247	897
LLH > 0.0	0.340	677	0.297	706
LLH > 0.1	0.404	480	0.343	522
LLH > 0.2	0.474	341	0.404	356
LLH > 0.5	0.677	96	0.564	99

best statistical mass error. However, after mass calibration, the Fisher discriminant cut of -0.1 no longer represents the minimum point of statistical mass errors as shown in figure 8.2.

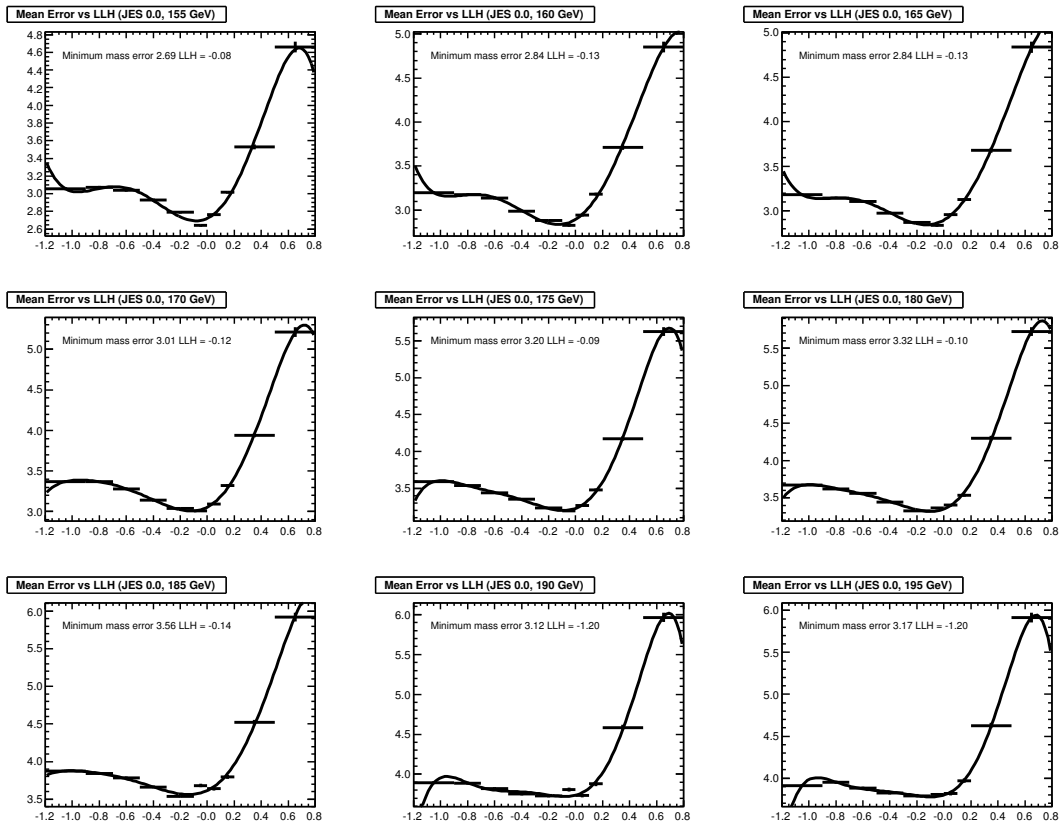

Figure 8.1. Ensemble mass error vs Fisher discriminant cut before mass calibration, for different input top mass in exactly 6 jets, 2 BIDNN > 0.65 case. The line shows the polynomial fit to the points.

Figure 8.3 shows the same feature of the disappearance of the minimum in the

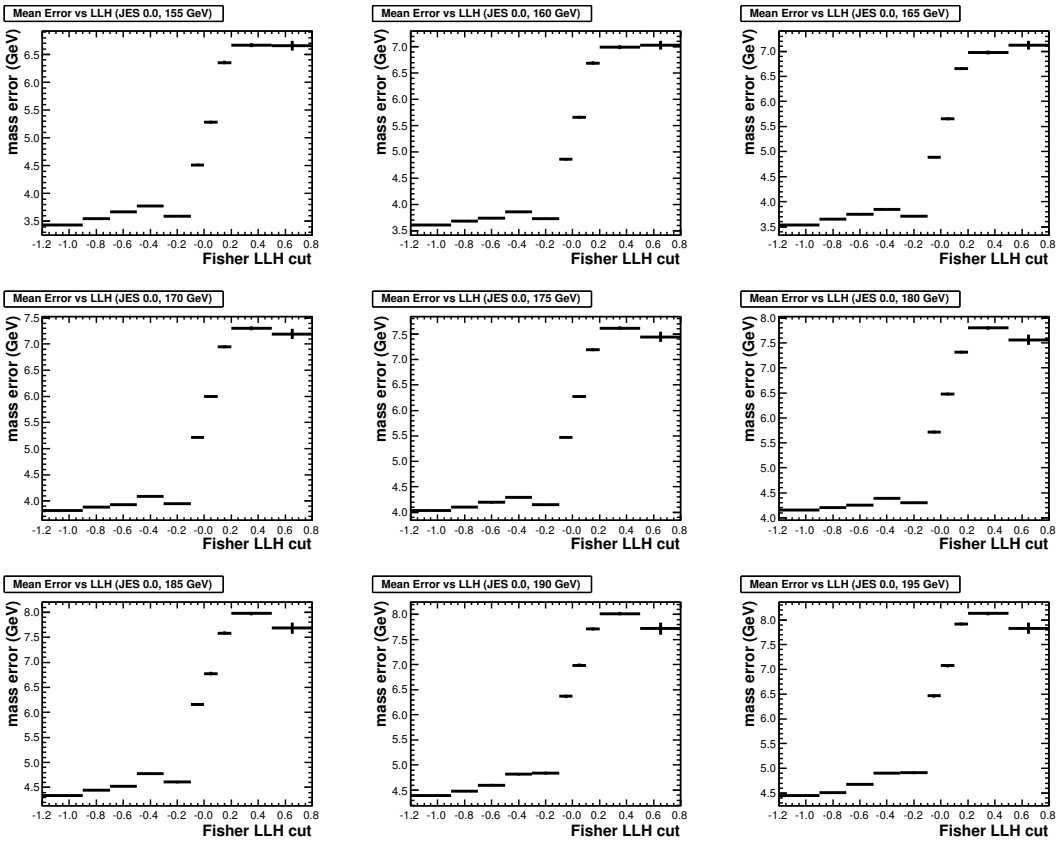

Figure 8.2. Ensemble mass error vs Fisher discriminant cut after mass calibration, for different input top mass in exactly 6 jets, 2 BIDNN > 0.65 case.

7 or more jets, 2 BIDNN > 0.2 case. From the results it can be inferred that a cut on Fisher discriminant is unlikely to reduce the mass uncertainty. The statistics of the whole distributions turns out to be more powerful. Nevertheless, the case of having Fischer discriminant cut of $>= -0.1$ will be shown alongside with the case of no cut for comparisons.

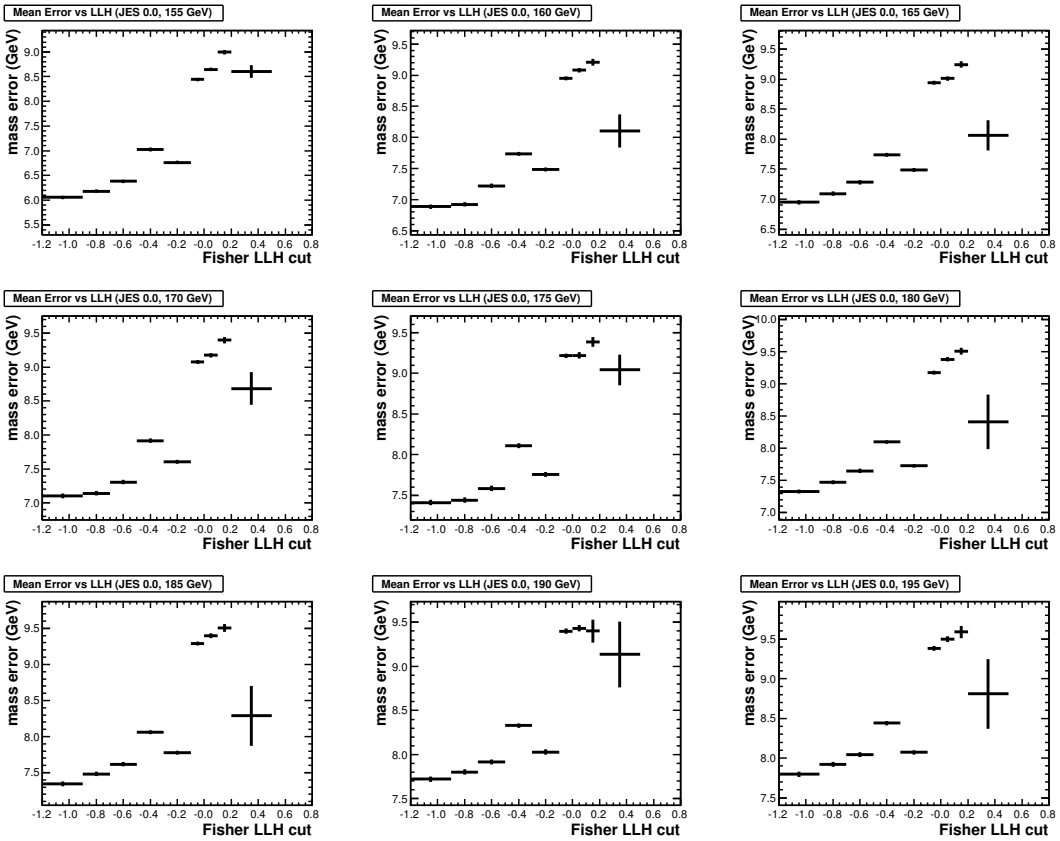

Figure 8.3. Ensemble mass error vs Fisher discriminant cut after mass calibration, for different input top mass in 7 or more jets, 2 BIDNN > 0.2 case.

8.2 Jet Energy Scale

The Jet Energy Scale (JES) is one major source of systematic uncertainties in the all hadronic channel because all final state objects of the channel are jets, and each jet energy is corrected by JES. As indicated in Chapter 3.4.5, JES is determined with great precision with uncertainties down to 1 % in the CC region and down to 2 % in the EC region.

To determine the uncertainty of the JES, one varies the JES for the signal Monte Carlo events by ± 1 standard deviations. They are then combined with background events to generate ensembles. The corresponding mass will then be extracted from an ensemble test and will be compared with the case using standard JES correction. Note that the JES of the background events doesn't need to be shifted because the background events are constructed using jets from data and should have the same JES correction as data.

Figure 8.4 shows the results from the 3 different JES for the exactly 6 jets, 2 BIDNN > 0.65 case. Figure 8.5 shows similar results for the 7 or more jets, 2 BIDNN > 0.2 case. From the fits it can be seen that the error due to different JES is $^{+0.75}_{-0.77}$ GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and $^{+1.13}_{-1.28}$ GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case.

In section 3.4.5 the uncertainty of JES is shown to be $\approx 1\%$ per jet in the central calorimeter (CC) region. Since the top mass is reconstructed by combining 3 jets, the estimated uncertainty is therefore $1\%/sqrt3 \approx 0.57\%$. This agrees with the uncertainties above since the top mass is ≈ 170 GeV would give uncertainty of 1 GeV/170 GeV $\approx 0.59\%$.

Additional uncertainties may exist in the b-tagged jets because the current Jet Energy Scale does not have a separate correction for the b-jets, and the long life-

time, more energetic b-quark jets may require special attention. However, given that there are 6 jets in the all-hadronic final state the uncertainty due to the b-jets are likely to be smeared out. In addition, the sample dependent correction in section 5.1.1 introduces extra correction for b-jets, c-jets and lighter quark jets separately, which further reduces the additional uncertainties due to the b-jets.

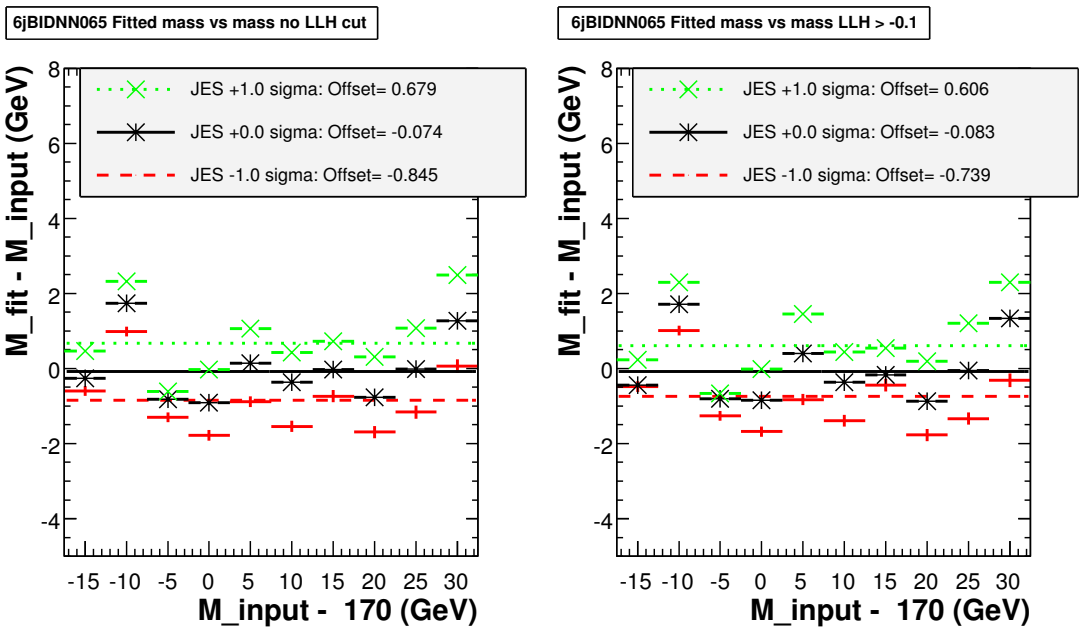

Figure 8.4. Ensemble test results with mass calibration applied for standard, $+/-1\sigma$ JES in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

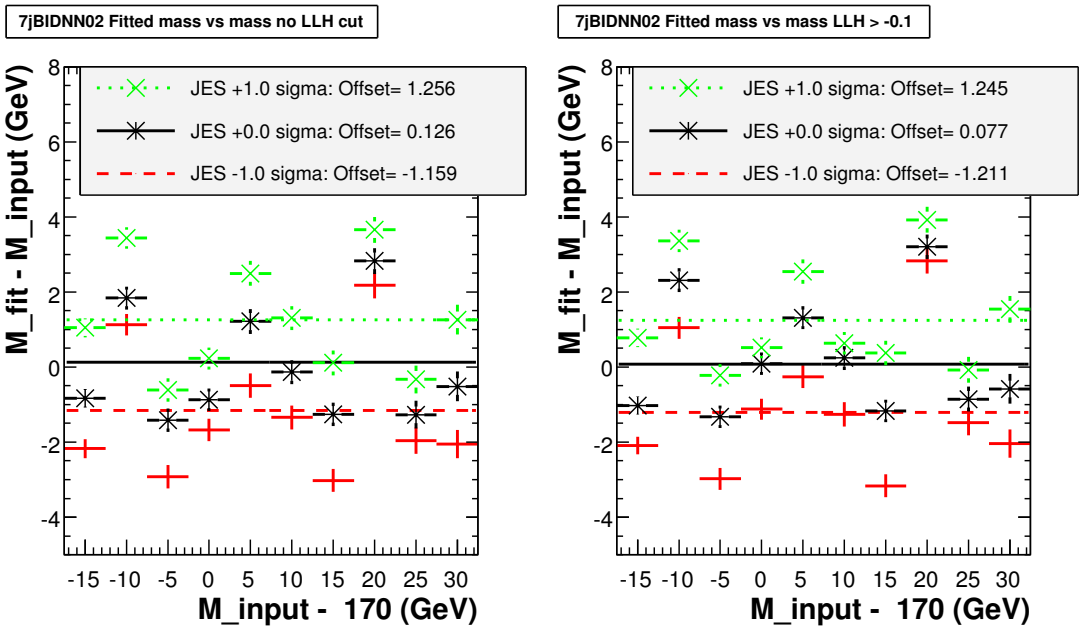

Figure 8.5. Ensemble test results with mass calibration applied for standard, $+/-1\sigma$ JES in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

8.3 Signal Purity

Given the uncertainties of signal purity as shown in table 6.4, such uncertainties will also contribute to the top quark mass as well. This can be observed with the ensemble tests by shifting the purity up and down by 1σ. Figure 8.6 shows the results using different signal purity for the exactly 6 jets, 2 BIDNN > 0.65 case. Figure 8.7 shows similar results for the 7 or more jets, 2 BIDNN > 0.2 case.

Using linear fits the uncertainties due to signal purity are $+0.59 + 0.097m_{top}$ and $-0.03 - 0.003m_{top}$ GeV$/c^2$ for the exactly 6 jets, 2 BIDNN > 0.65 case. For for the 7+ jets, 2 BIDNN > 0.2 case the up and down uncertainties are: $+0.39 + 0.230m_{top}$ and $-0.02 + 0.021m_{top}$ GeV$/c^2$. Notice that without the discriminant cut, increasing the signal purity only has very small effects on the result, but decreasing the signal purity will result in a noticeable shift. This is expected because in the case of low purity, there should be less signal present. The signal template peak is then required to shift further away from the background peak, which results in the higher mass value extracted.

In contrast, for the cases with discriminant cut, the uncertainties in mass are larger in both cases of the signal purity shift. This is due to the reduction in the number of events after the discriminant cuts, which makes the top mass more sensitive to the shift in signal purity, even when the signal purity is high. Also the discriminant cut removes the background events that don't look like signal, which then makes the remaining background events harder to be distinguished from signal events.

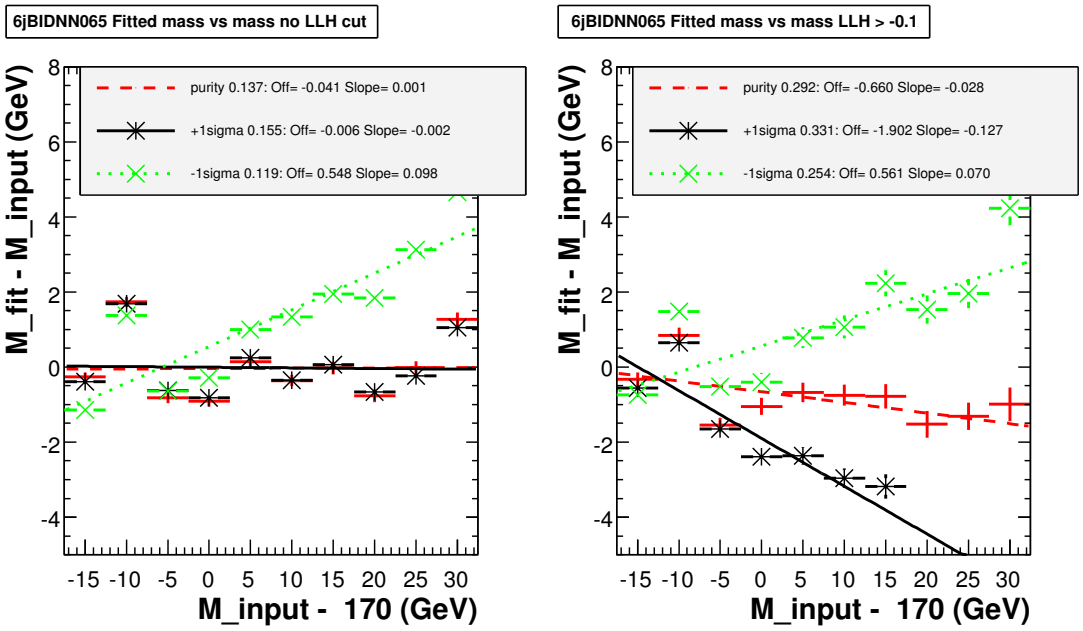

Figure 8.6. Ensemble test results with mass calibration applied for central, $+/-1\sigma$ signal purity in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

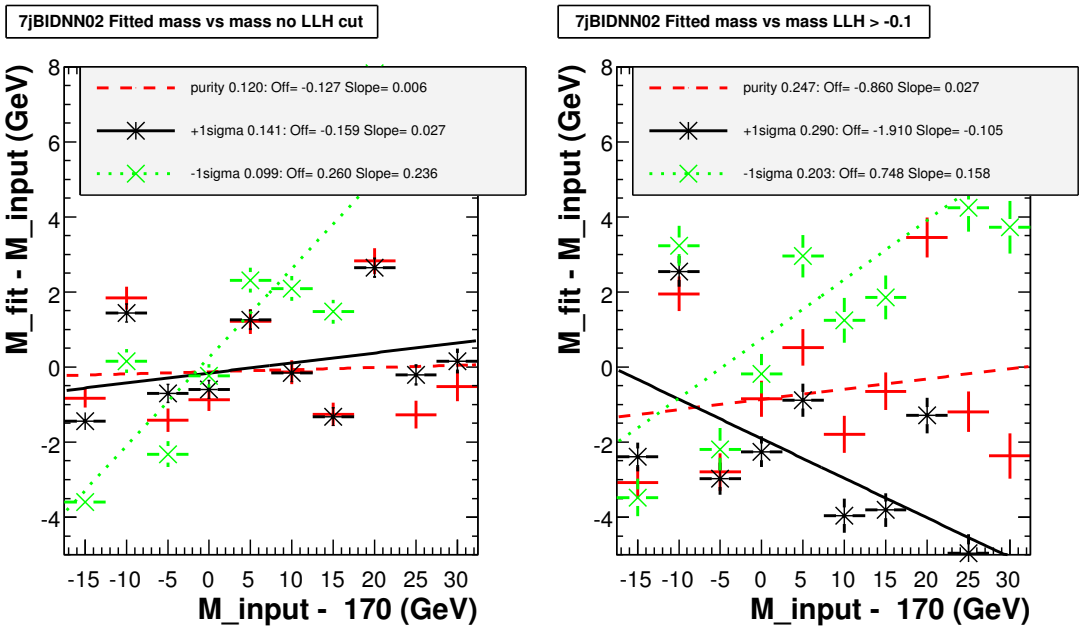

Figure 8.7. Ensemble test results with mass calibration applied for central, $+/-1\sigma$ signal purity in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

8.4 Parton distribution function

The top quarks from the experiment are generated from the collision of protons and anti-protons. Both protons and anti-protons are composite objects which contain not only the $uu\bar{d}$ ($\bar{u}\bar{u}d$ for anti-proton) quarks, but also a sea of virtual quark-anti-quark pairs as well as gluons. The probability density of a Parton type i within a proton is described by the Parton distribution function (PDF) $f(x_i, \mu_F^2)$ [70], [71] , where x_i is the longitudinal momentum fraction of Parton i within the proton and μ_F^2 is the factorization scale.

Given the PDF, the top quark production can then be calculated by considering the different Parton combinations available inside the proton to generate a $t\bar{t}$ pair:

$$\sigma(p\bar{p} \to t\bar{t} + X) = \sum_{a,b} \int dx_a \int dx_b f_a^p(x_a, \mu_F^2) f_b^{\bar{p}}(x_b, \mu_F^2) \sigma(p_a, p_b \to p_t p_{\bar{t}}) \quad (8.1)$$

Notice that the PDF will affect the Parton contributions to $t\bar{t}$ production, modifying the different regions of phase space that contribute to the interaction. Therefore the variation of the PDF will propagate to the kinematics of the produced $t\bar{t}$ pair as well and introduce uncertainties to the top mass as well.

To estimate the uncertainty introduced by the PDF, we compare the signal results using CTEQ6.5M [72] (which is based next-to-leading order perturbative QCD) and CTEQ6L1 [73] (which involves leading order only). The PDFs are applied to the Monte Carlo events and introduce a weight to each event based on their modification to the phase space. Then signal ensembles are generated using those weights and the top mass will be extracted. Figure 8.8 shows the results of different PDF for the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.9 shows

results for the 7 or more jets, 2 BIDNN > 0.2 case.

The uncertainty for PDF can be estimated as ± 0.07 GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and ± 0.09 GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case. Since the large mass of top quark requires the partons involved in the hard scattering to be sufficiently energetic and carries larger fraction of the proton momentum (x_i). The uncertainty in PDF itself gets larger for smaller values of x, therefore the $t\bar{t}$ production is relatively unaffected and the corresponding uncertainty is small. Since the fraction of $t\bar{t}$ pairs produced in gg fusion is only 15%, this further reduces sensitivity to the gluon PDF uncertainties.

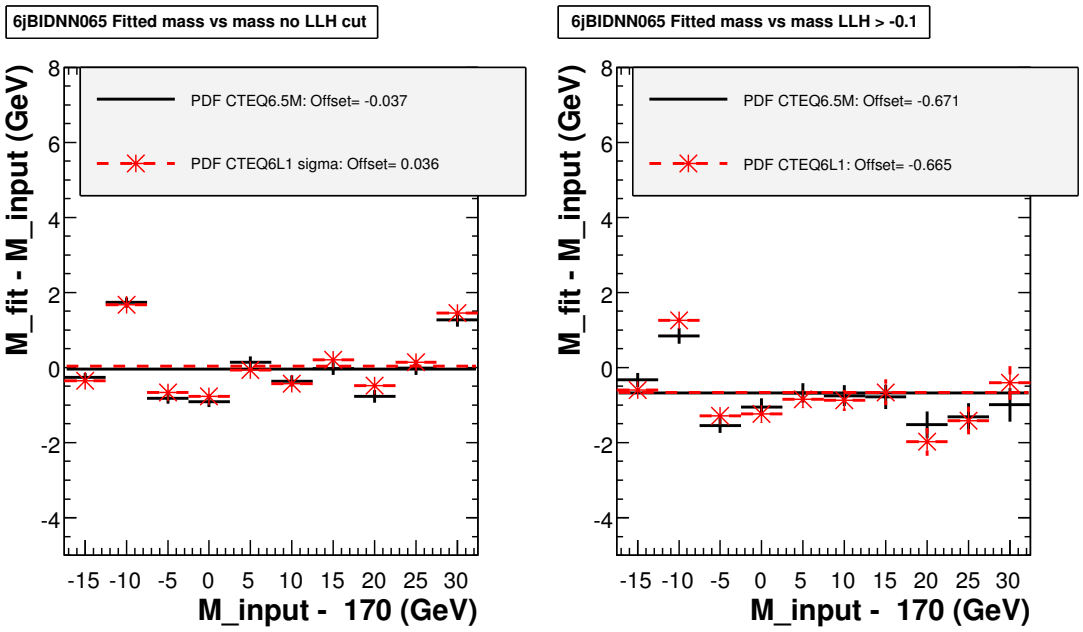

Figure 8.8. Ensemble test results with mass calibration applied between PDF CTEQ6.5M and CTEQ6L1 in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

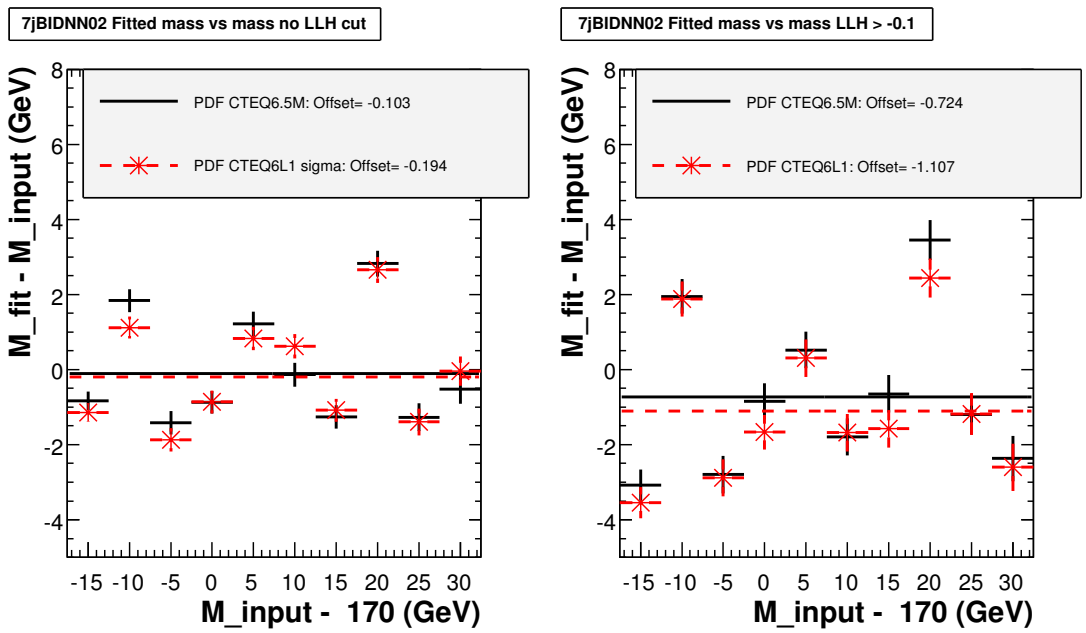

Figure 8.9. Ensemble test results with mass calibration applied between PDF CTEQ6.5M and CTEQ6L1 in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

8.5 b-fragmentation

Perturbative QCD cannot account for the hadronization (fragmentation) of quarks due to its occurence in the low energy range where the strong coupling constant α_s becomes large. Non-perturbative QCD models are required instead and they are tuned to fit with experiment results.

For top decay it is important to have the correct fragmentation model for b-quarks as it is a significant source of uncertainty [74]. In this analysis the string

model from Bowler [75] with Peterson [76] fragmentation is being used together with the tuning to the LEP/SLD [77] experimental data. The b-fragmentation function is then applied to Monte Carlo events as event weights based on the b-jet kinematics.

To evaluate the uncertainty, one can shift the fragmentation function (since the function itself is a fit to the data anyways) by 1 sigma in a certain direction, then re-weight the Monte Carlo events and perform the ensemble test to see how it is different from the central value.

Figure 8.10 shows the results of standard b-fragmentation and $+1\sigma$ for the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.11 shows results for the 7 or more jets, 2 BIDNN > 0.2 case.

It is safe to assume the uncertainty for b-fragmentation to be symmetric, and is found to be ± 0.01 GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and ± 0.05 GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case. The uncertainties due to b-fragmentation are small because there are 6 jets in all-hadronic events which smears out the effect of b-jets.

8.6 Signal Modeling

The signal Monte Carlo events in this analysis are generated by Pythia, which builds the events based on Parton showering. In contrast, one can use a matrix element based generator such as ALPGEN to simulate the signal. The difference between the two cases can be treated as the systematics in signal modeling.

Figure 8.12 compares between Pythia and Alpgen generator for the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.13 shows results for the 7 or more jets, 2 BIDNN > 0.2 case. Currently there are only 2 different top mass ALPGEN $t\bar{t}$

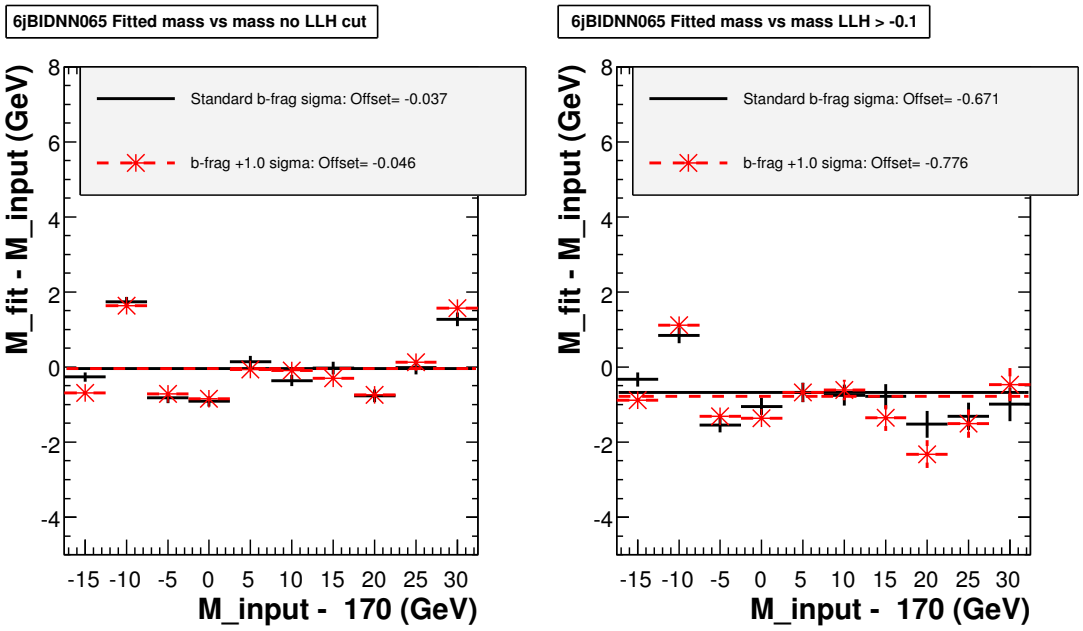

Figure 8.10. Ensemble test results with mass calibration between standard and $+1\sigma$ b-fragmentation in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

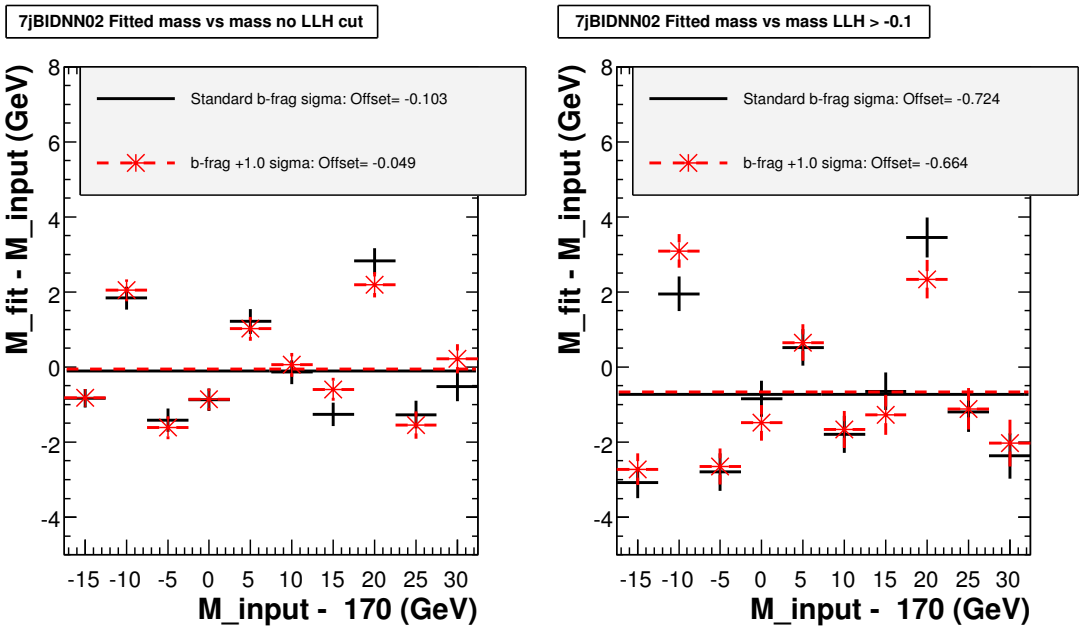

Figure 8.11. Ensemble test results with mass calibration between standard and $+1\sigma$ b-fragmentation in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

samples available, the uncertainty is taken as the average of the absolute difference between the Pythia and ALPGEN fitted mass, and can be treated as a symmetric error. The uncertainty due to signal modeling is found to be ± 2.56 GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and ± 1.84 GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case.

8.7 Background Modeling

As mentioned in section 5.4, one can also build a 6-jets background sample by adding 2 soft jets to 4-jets data events. Such "4+x" background samples can be used to estimate the uncertainty of the background model by comparing with the standard "5+x" background samples for the 6-jets case. Similarly for 7-or-more jets case, one can compare between the "5+x" and "4+x" background samples to estimate the uncertainties.

Before employment, one should check the distributions between the 4+x and 5+x background samples to see if they are compatible. Figure 8.14 shows the comparison of the Jet 1-6 p_T distributions between the 5+x and 4+x background. Figure 8.15 shows the comparison of the Jet 7-8 p_T, maximum/minimum angular separation (dR) among jets, maximum rapidity difference among jets and between the 2 leading jets. Figure 8.16 shows the comparison of the invariant mass of all jets, 2 b-jets, H_T and λ_3. The jet 4-8 p_T distributions are expected to be somewhat different. This is because the 4+x background is generated from 4-jets donor samples, which can only match (by sampling) the donor jet p_T distributions up to the 4th jet (This is also another reason why p_T beyond 4th jet should not be used for the discriminant). The rest of the variables match quite well between the 2 background samples.

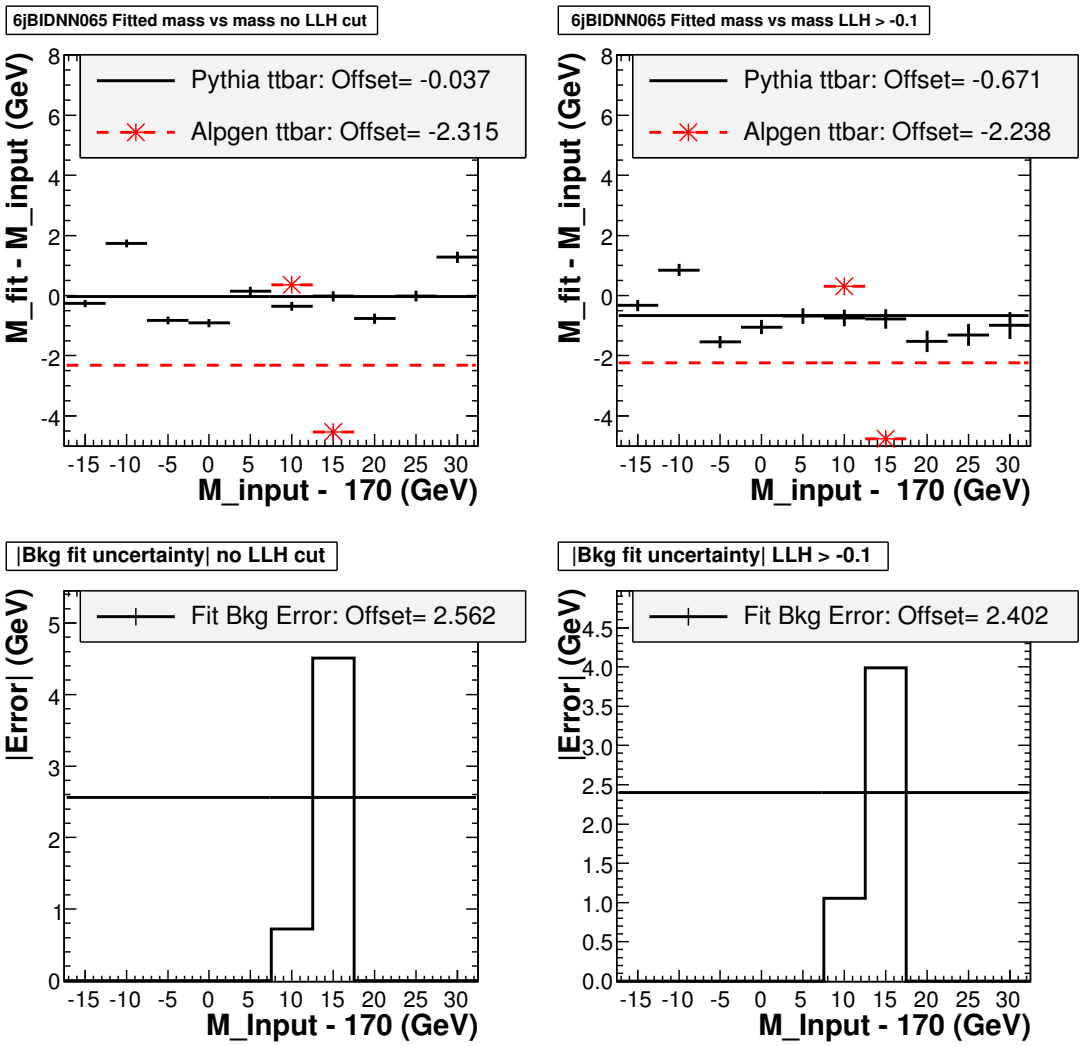

Figure 8.12. Ensemble test results with mass calibration applied between PDF Pythia and ALPGEN MC $t\bar{t}$ in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1. The bottom plots show the absolute difference between the Pythia and ALPGEN MC samples, which gives the fit systematics.

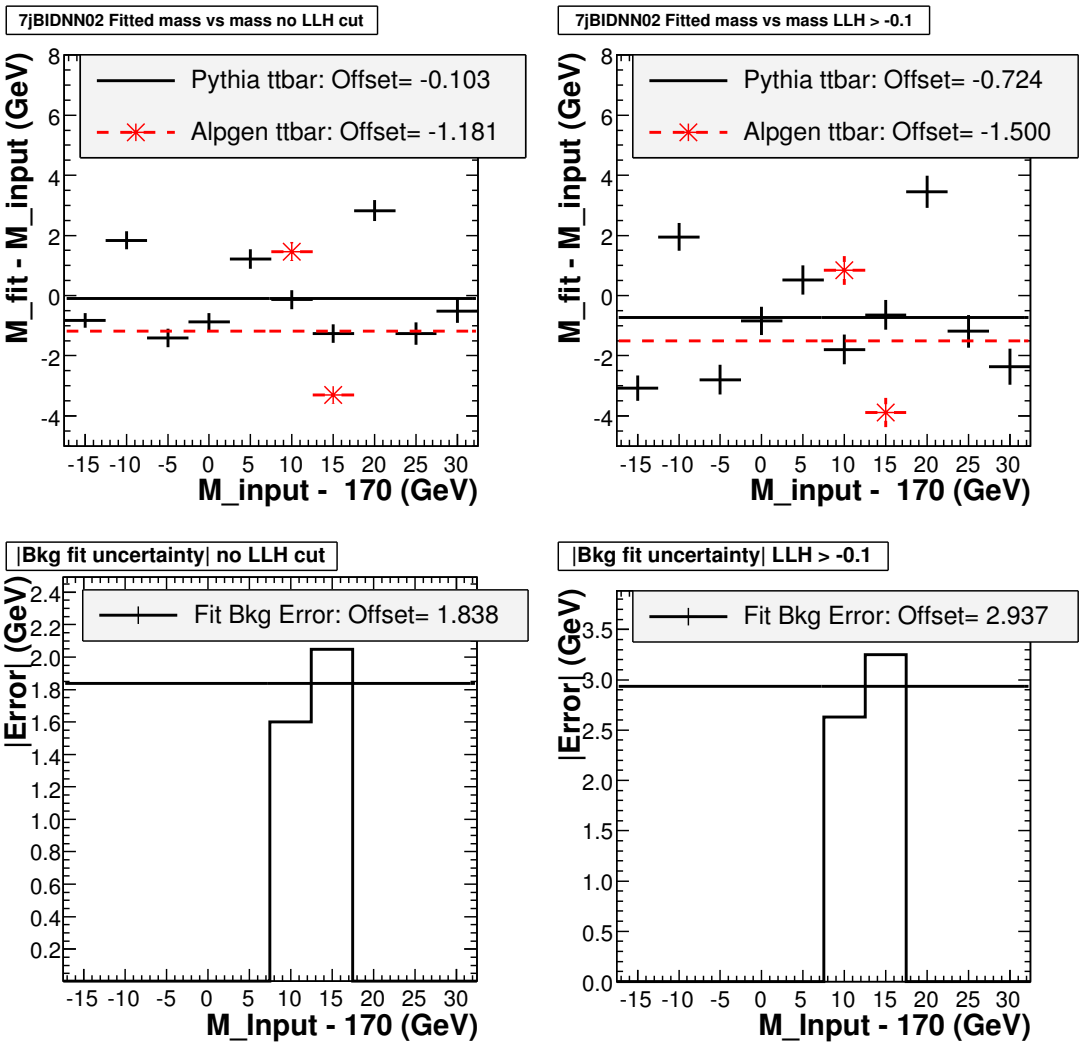

Figure 8.13. Ensemble test results with mass calibration applied between PDF Pythia and ALPGEN $t\bar{t}$ in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1. The bottom plots show the absolute difference between the Pythia and ALPGEN MC samples, which gives the fit systematics.

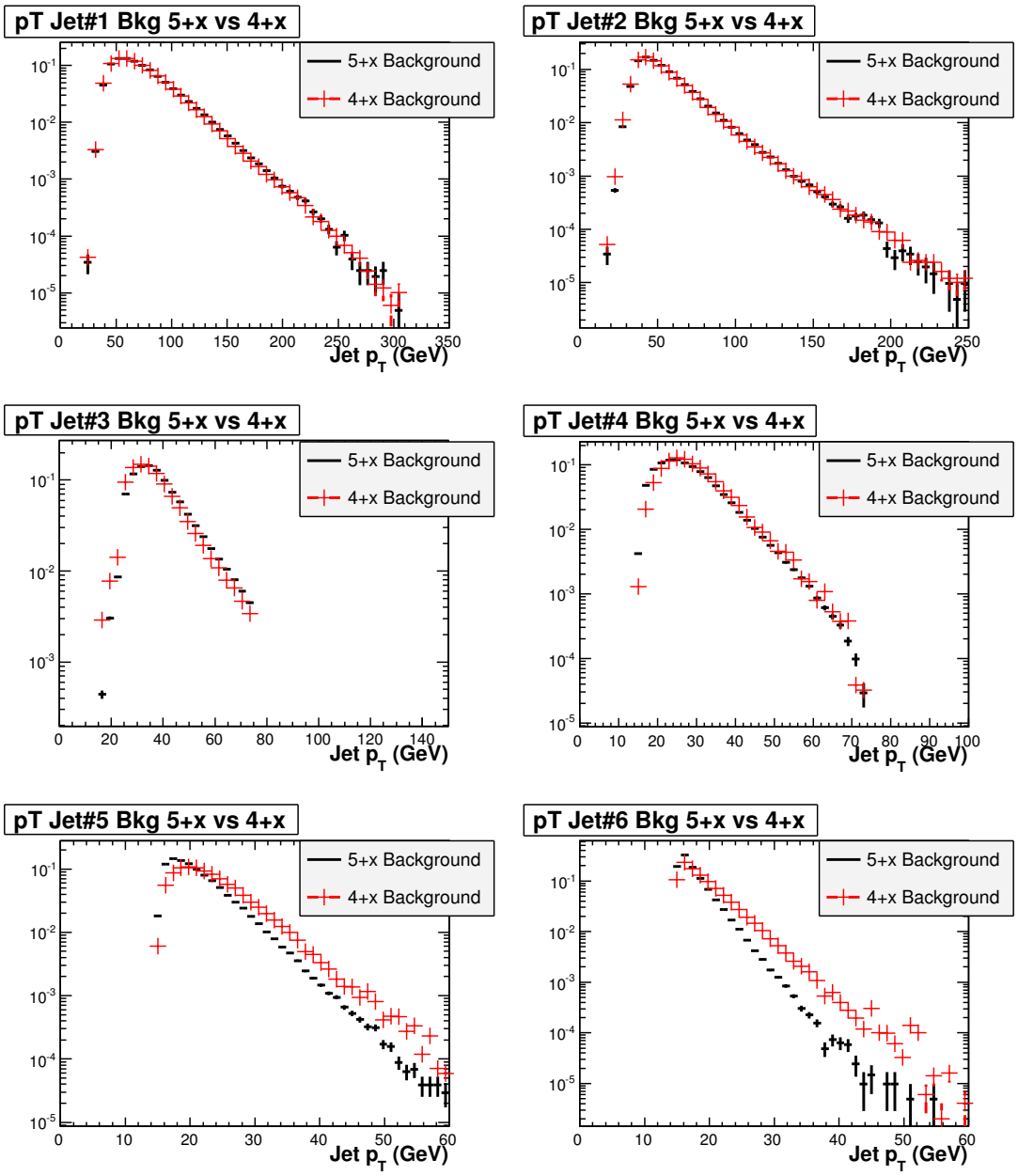

Figure 8.14. Background distributions comparison between 4+x and 5+x background, which includes jet 1-6 p_T. All distributions are normalized to one.

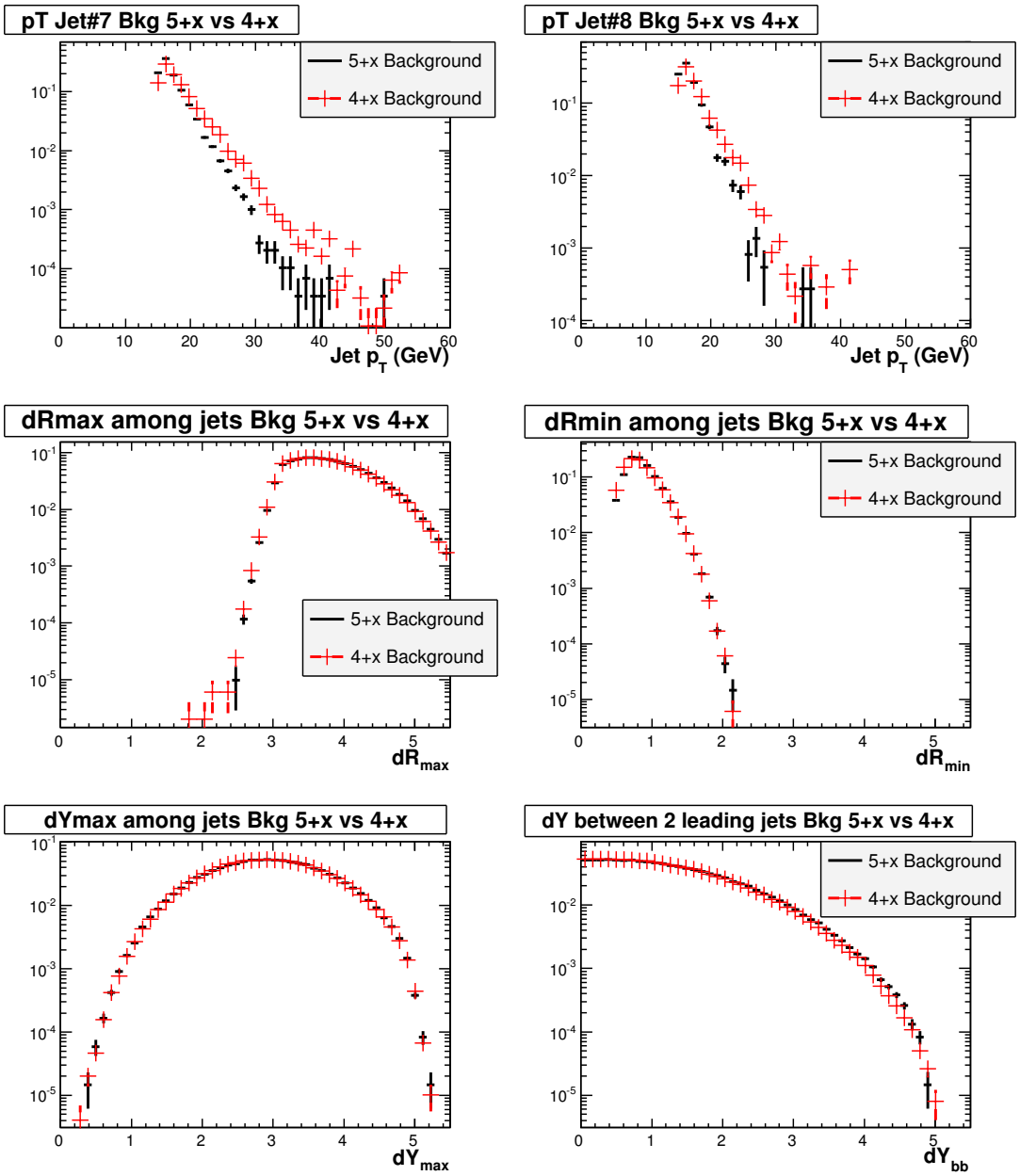

Figure 8.15. Background distributions comparison between 4+x and 5+x background, for jet 7-8 p_T, dR_{max} (max. angular separation), dR_{min}, dY_{max} (max rapidity difference) and dY_{01} (between 2 leading jets). All distributions are normalized to one.

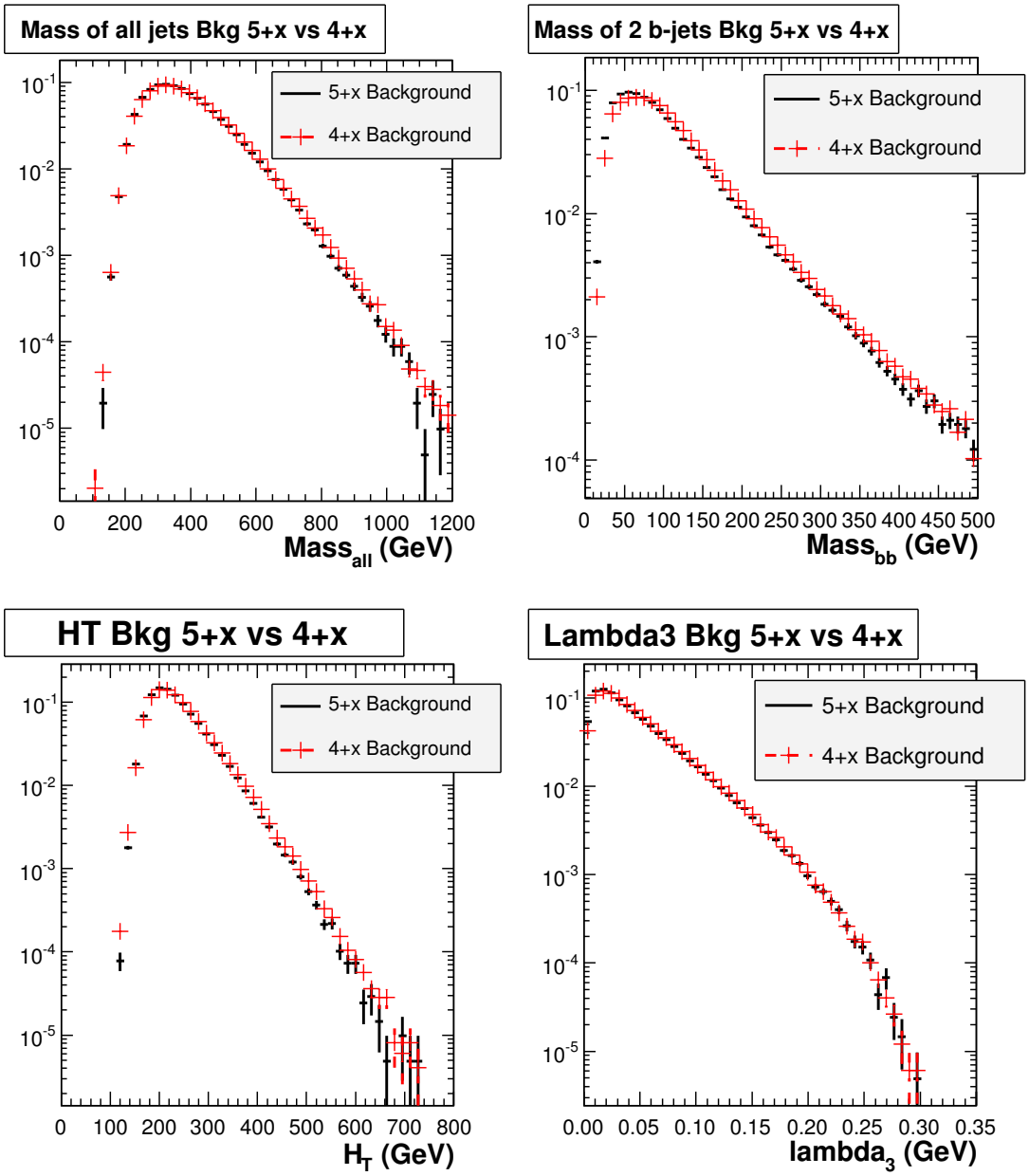

Figure 8.16. Background distributions comparison between 4+x and 5+x background, for M_{all} (invariant mass of all jets), M_{bb} (between 2 b-jets), H_T and λ_3 (momentum tensor eigenvalue).] All distributions are normalized to one.

In addition, figure 8.17 shows the Dmass distribution between the 4+x and 5+x background samples. In most cases the Dmass distributions agree within errors between the two backgrounds. Notice that there are some larger differences around the peak in the 7+ jets, 2 BIDNN > 0.2 case. Therefore it is anticipated that the 7+ jets case will have larger background model systematic error compared to 6 jets case.

Figure 8.18 shows the fitted mass comparison between 4+x and 5+x background for the exactly 6 jets, 2 BIDNN > 0.65 case. Figure 8.19 shows a similar comparison for the 7 or more jets, 2 BIDNN > 0.2 case. From these figures it can be seen that the fitted mass values using the 4+x background are lower than those from the 5+x background events. The background model uncertainty is therefore taken as a one sided error, which is -2.52 GeV$/c^2$ for the exactly 6 jets, 2 BIDNN > 0.65 case, and -6.66 GeV$/c^2$ for the 7+ jets, 2 BIDNN > 0.2 case. As expected, the background model uncertainty is the dominant systematic error.

8.8 Signal fitting

The signal template involves separately fitting the correct jet combinations with a Gaussian distribution, and fitting the wrong jet combinations part using a Log-Normal function. Such a choice of signal template function may not be optimal and it is useful to see how alternate fits change the answer. An alternate fit we use will be using 2 Log-Normal functions fitting the correct and wrong jet combinations components together as shown below:

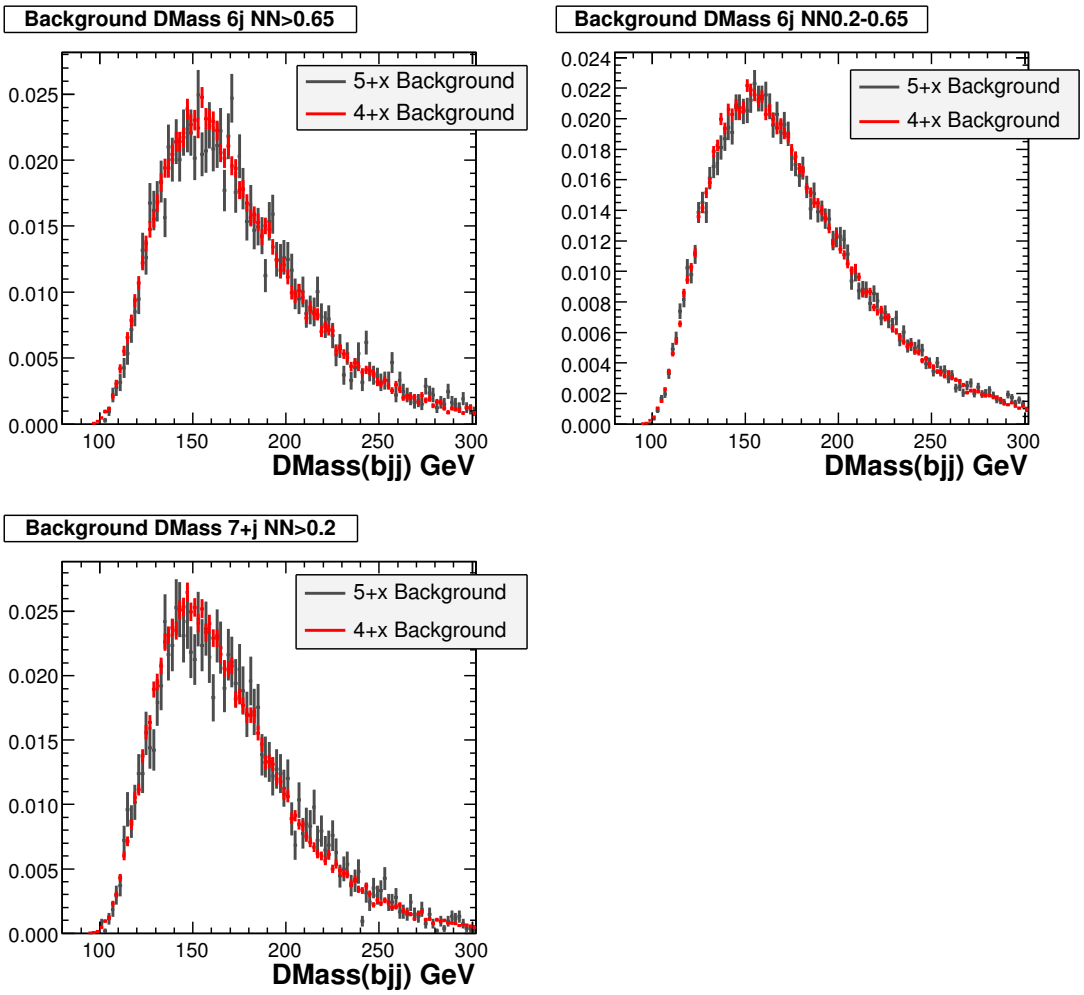

Figure 8.17. $Dmass(bjj)$ distribution comparison between 4+x and 5+x background, in 3 different BID NN regions: Exactly 6 jets, 2 BIDNN > 0.65, 2 BINNN 0.2-065, 7+ jets, 2 BIDNN > 0.2.

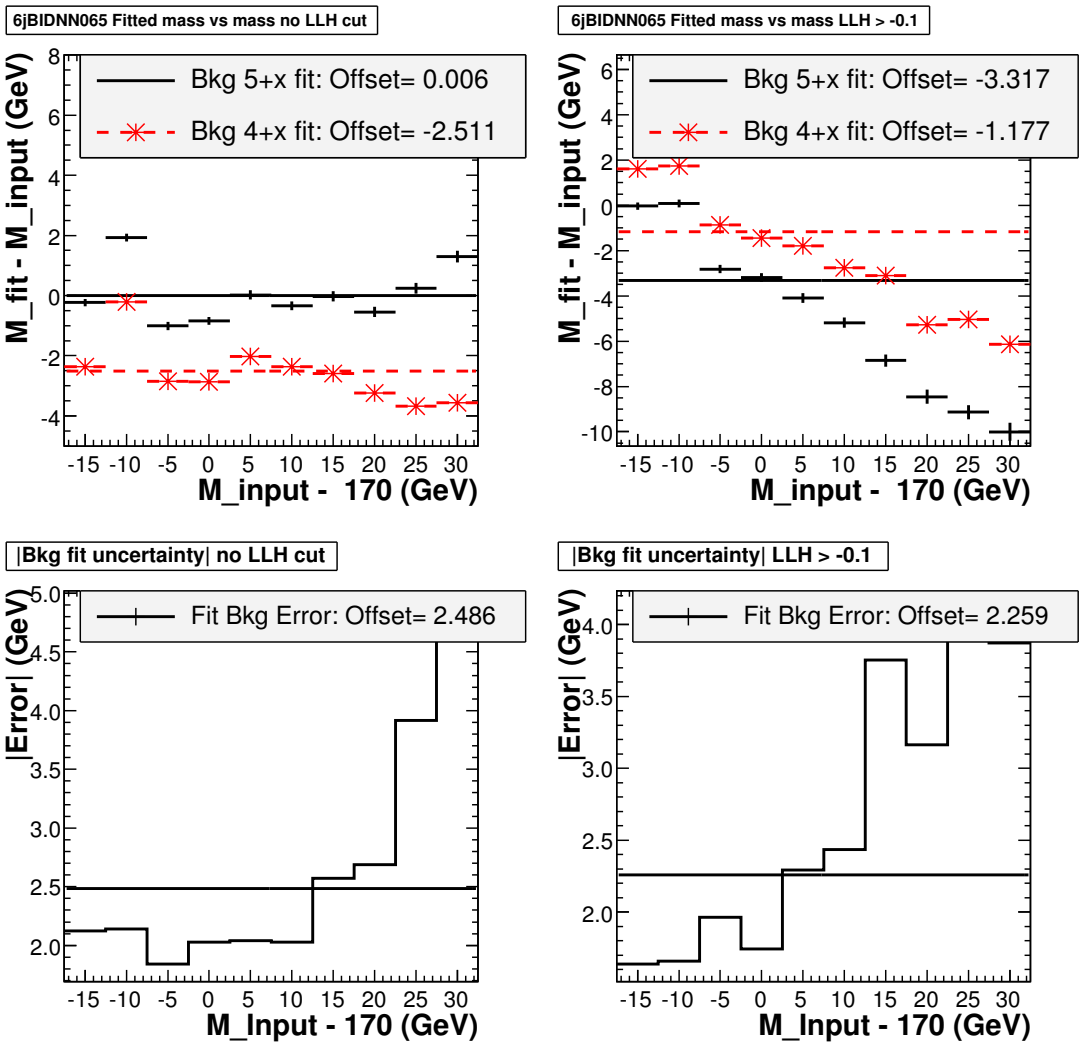

Figure 8.18. Ensemble test results with mass calibration applied between PDF 4+x and 5+x background model in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1. The bottom plots show the absolute difference between the 4+x and 5+x background, which gives the fit systematics.

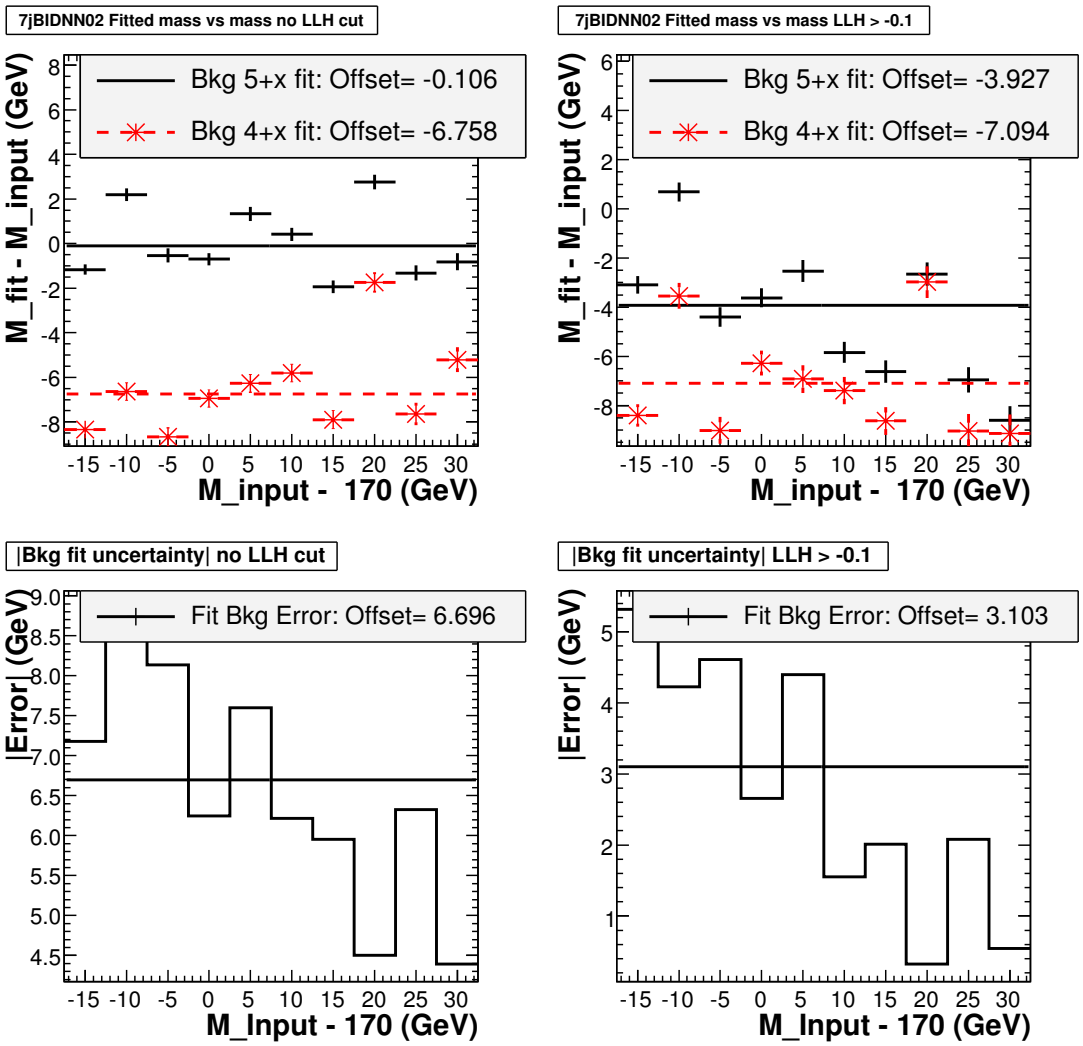

Figure 8.19. Ensemble test results with mass calibration applied between PDF 4+x and 5+x background model in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1. The bottom plots show the absolute difference between the 4+x and 5+x background, which gives the fit systematics.

$$
\begin{aligned}
& fit_{signal}(Dmass(bjj); p_0 ... p_7) = \\
& \qquad p_0 * LogNormal(Dmass(bjj), p_1, p_2, p_3) + \\
& \qquad p_4 * LogNormal(Dmass(bjj), p_5, p_6, p_7)
\end{aligned} \tag{8.2}
$$

Figure 8.20 shows the signal template alternate fits using the above function for the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.21 shows the fits for the 7 or more jets, 2 BIDNN > 0.2 case. The fit results look satisfactory for the signal templates.

The alternate fits are then calibrated separately and then applied to the ensemble test to extract the mass measurement. Figure 8.22 compares between standard and alternate signal fits the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.23 shows results for the 7 or more jets, 2 BIDNN > 0.2 case.

By looking at the average absolute difference between the standard signal fit and the 2 Log-Normal alternate fit, the uncertainty due to signal fitting is found to be ± 1.13 GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and ± 0.29 GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case. It can be observed that the mass measurement method is quite sensitive to the template fit being used.

8.9 Background fitting

Similar to the case for Signal template, alternate fits are used for the background as well for the estimation of the background fitting error. The alternate fit uses the combination of a Log-Normal function and a Gaussian distribution:

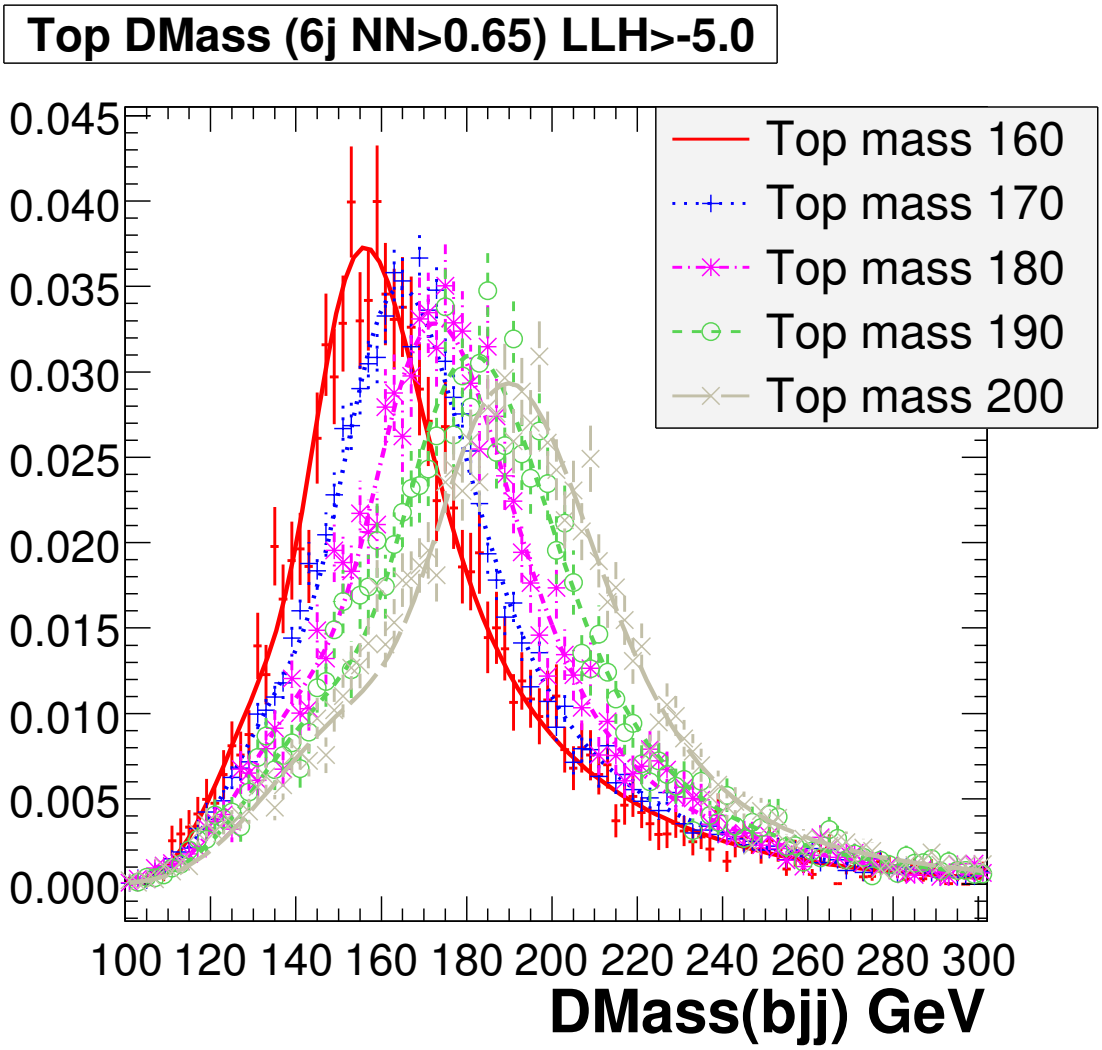

Figure 8.20. Alternate signal template fits for $m_{top} = 160 - 200$ GeV in exactly 6 jets, 2 BIDNN > 0.65 case.

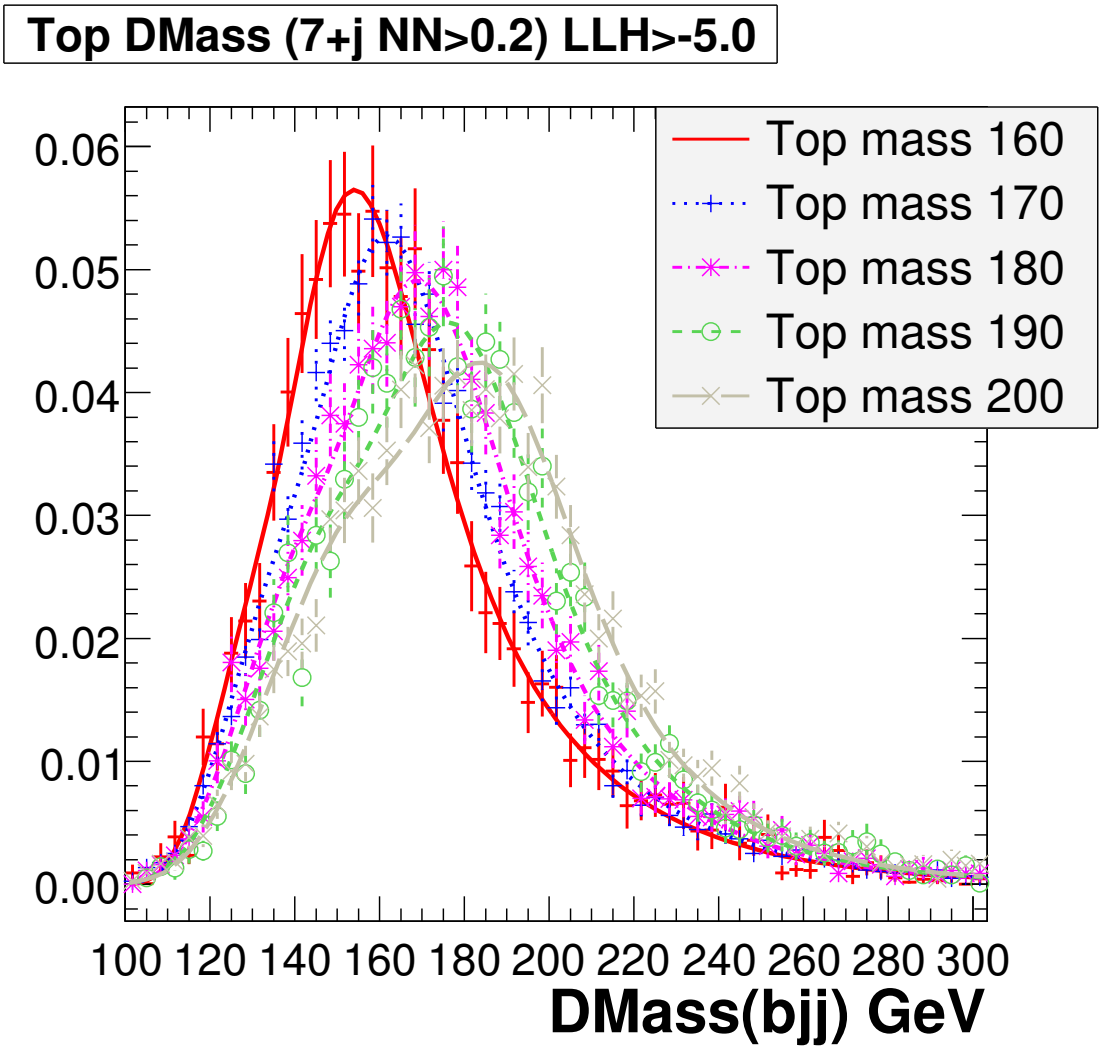

Figure 8.21. Alternate signal template fits for $m_{top} = 160 - 200$ GeV in 7 or more jets, 2 BIDNN > 0.2 case.

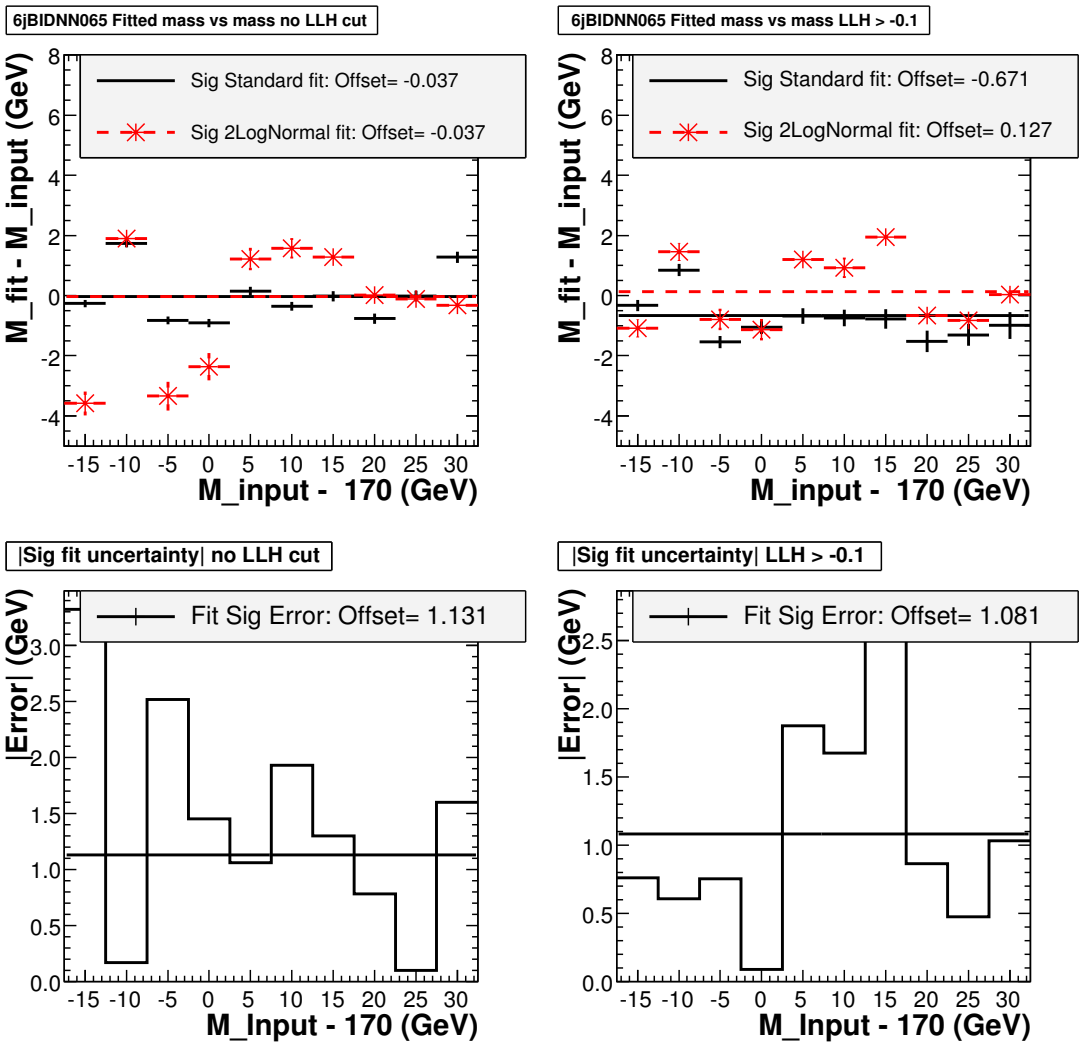

Figure 8.22. Ensemble test results with mass calibration between standard and alternate signal fit in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1. The bottom plots show the absolute difference between the standard and alternate signal template, which gives the fit systematics.

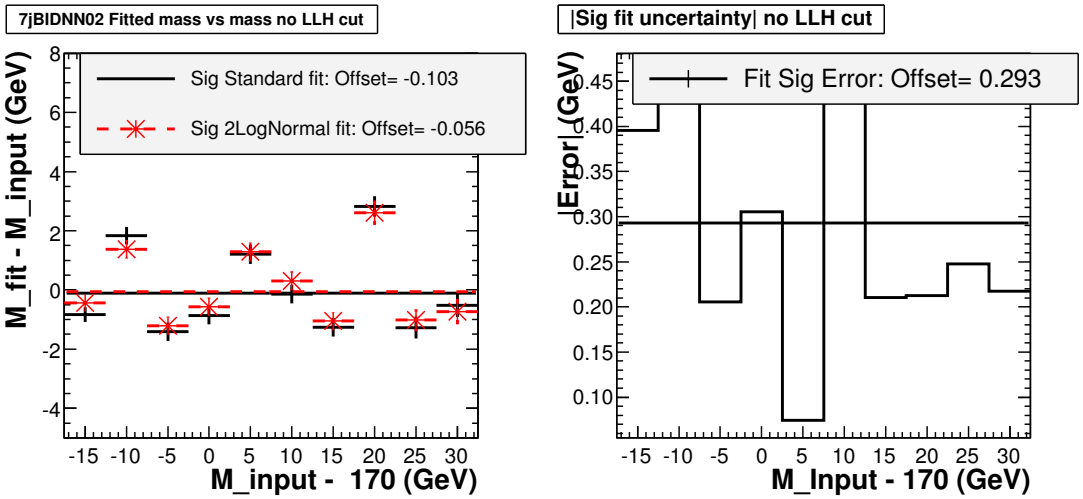

Figure 8.23. Ensemble test results with mass calibration between standard and alternate signal fit in 7 or more jets, 2 BIDNN > 0.2 case. Only the case with no discriminant cut is shown. The bottom plots show the absolute difference between the standard and alternate signal template, which gives the fit systematics.

$$
\begin{aligned}
&fit_{bkg}(Dmass(bjj); p_0...p_6) = \\
&\quad p_0 * LogNormal(Dmass(bjj), p_1, p_2, p_3) + \\
&\quad \frac{p_4}{N} * exp\left(-\frac{(Dmass(bjj) - p_5)^2}{2 * p_6^2}\right)
\end{aligned} \tag{8.3}
$$

Figure 8.24 shows the background template alternate fits using the above function for the 3 different BIDNN regions.

The alternate fits are then calibrated separately and then applied to the ensemble test to extract the mass measurement. Figure 8.25 compares between Spline and Log-Normal+Gaussian background fits the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.26 shows results for the 7 or more jets, 2 BIDNN > 0.2 case.

By looking at the average absolute difference between the standard signal fit and the 2 Log-Normal alternate fit, the uncertainty due to signal fitting is found to be ± 0.96 GeV/c^2 for the exactly 6 jets, 2 BIDNN > 0.65 case, and ± 1.60 GeV/c^2 for the 7+ jets, 2 BIDNN > 0.2 case.

Note that when comparing between Spline fits and alternate fits for the background events, the ensembles test shows that the fitted mass results using the Spline fit have less fluctuations compared to the case using the alternate fits. This indicates the Spline fit is preferred for background template fitting.

Table 8.9 summarize the systematic errors investigated in this section.

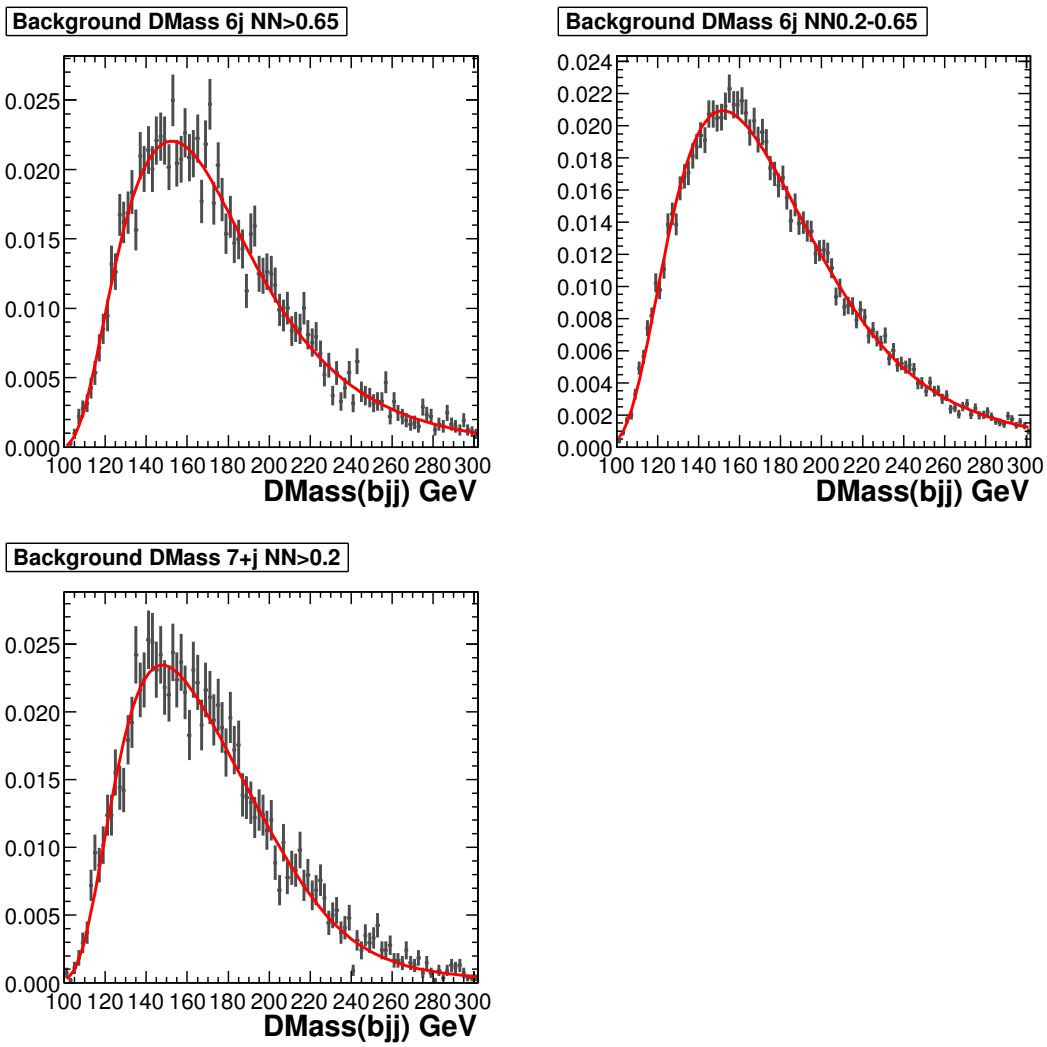

Figure 8.24. Alternate Background template fits for the exactly 6 jets, 2 BIDNN > 0.65 case, 6 jets, 2 BIDNN 0.2-0.65 case, and 7 or more jets, 2 BIDNN > 0.2 case.

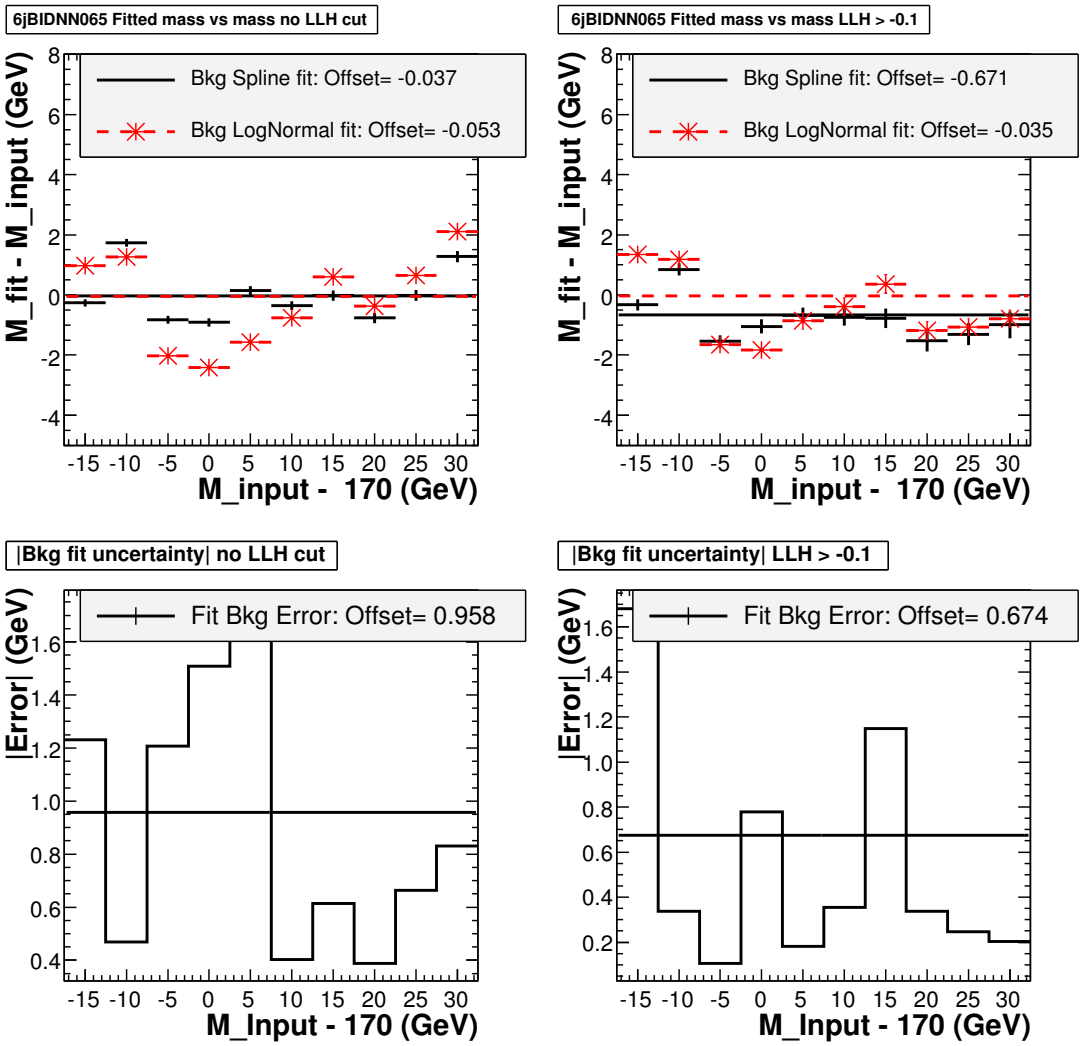

Figure 8.25. Ensemble test results with mass calibration between Spline and alternate background fit in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

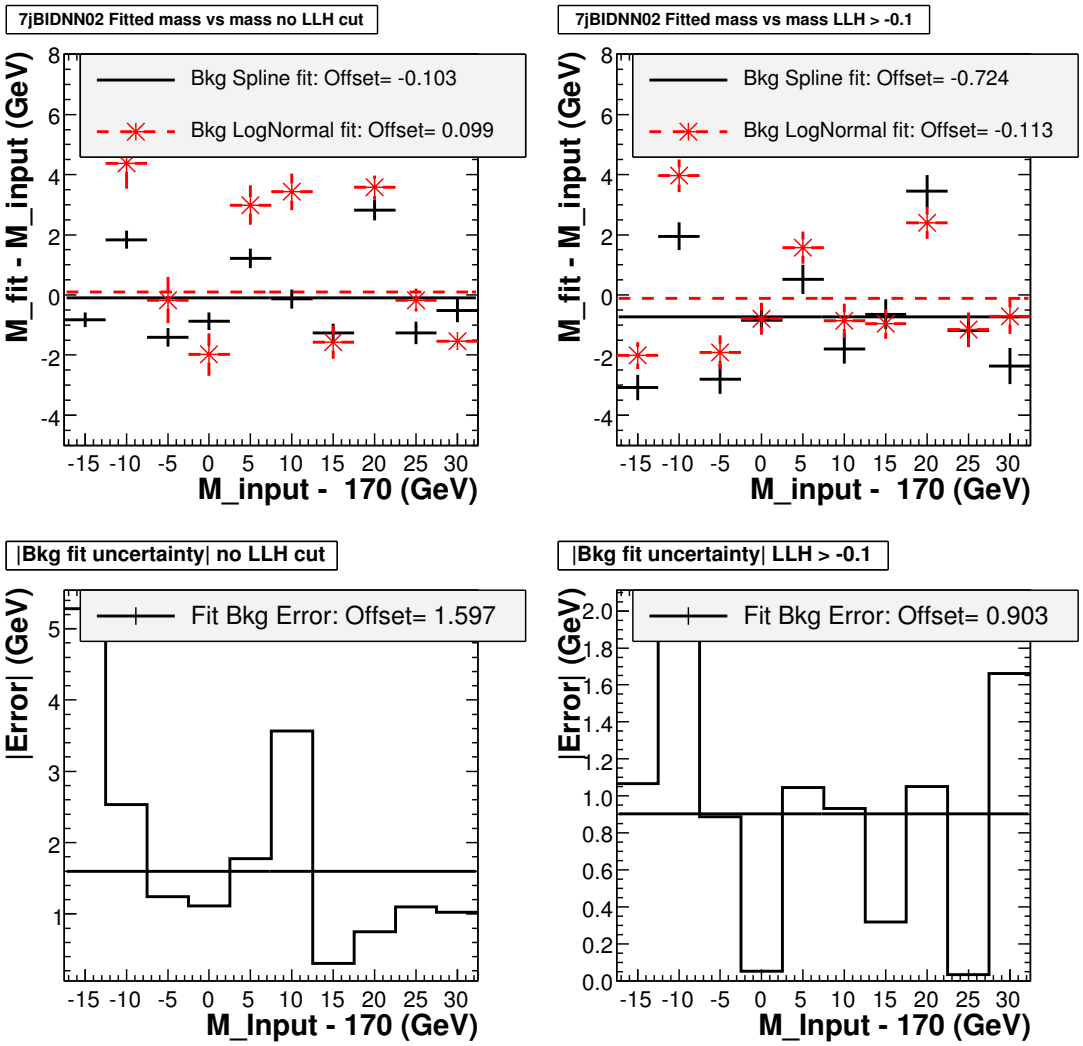

Figure 8.26. Ensemble test results with mass calibration between Spline and alternate background fit in 7 or more jets, 2 BIDNN > 0.2 case. The left plot shows the case with no discriminant cut, and the right plot shows the case with discriminant cut ≥ -0.1.

TABLE 8.2

Summary of systematic errors in the 2 BIDNN regions without Fisher Discriminant cut.

Source of Error	6 jets, 2 BIDNN > 0.65		7+ jets, 2 BIDNN > 0.2	
Jet Energy Scale	+0.75	-0.77	+1.13	-1.28
PDF	+0.07	-0.07	+0.09	-0.09
b-fragmentation	+0.01	-0.01	+0.05	-0.05
Physics Model	+2.56	-2.56	+1.84	-1.84
Signal purity	+1.05	-0.02	+1.54	-0.21
Signal fit	+1.13	-1.13	+0.29	-0.29
Bkg. Model (jet multiplicity)		-2.52		-6.66
Bkg. Model (fit)	+0.96	-0.96	+1.60	-1.60
Total	+3.23	-3.96	+3.11	-7.35

8.10 Statistical Error

For an overview of the expected statistical errors for the mass measurement, figure 8.27 compares the mass error with and without Fisher discriminant cut in the exactly 6 jets, 2 BIDNN > 0.65 case and figure 8.28 shows the 7 or more jets, 2 BIDNN > 0.2 case. In the plots the mass errors distributions with the Fisher discriminant is always larger than the case without the cut, which is another proof that that it is better to not apply any Fisher discriminant cut.

Finally we can look at the candidate events and extract the top mass. The signal purity in the exactly 6 jets, 2 BIDNN 0.2-0.65 region is too low and therefore it is not being used to extract a measurement of top mass.

The top mass will be extracted separately in the exactly 6 jets, 2 BIDNN > 0.65 case and the 7 or more jets, 2 BIDNN > 0.2 cases, each using 2 independent background samples ("5+x" and "4+x").

Figure 8.29 and 8.30 shows the fitted top mass from the candidate events using "5+x" background samples.

For illustration of the presence of $t\bar{t}$ signal in the candidate, Figure 8.31 and 8.32 shows the fitted top mass from the candidate events with Fisher discriminant cut of -0.1. As expected, the top mass error is larger with Fisher discriminant.

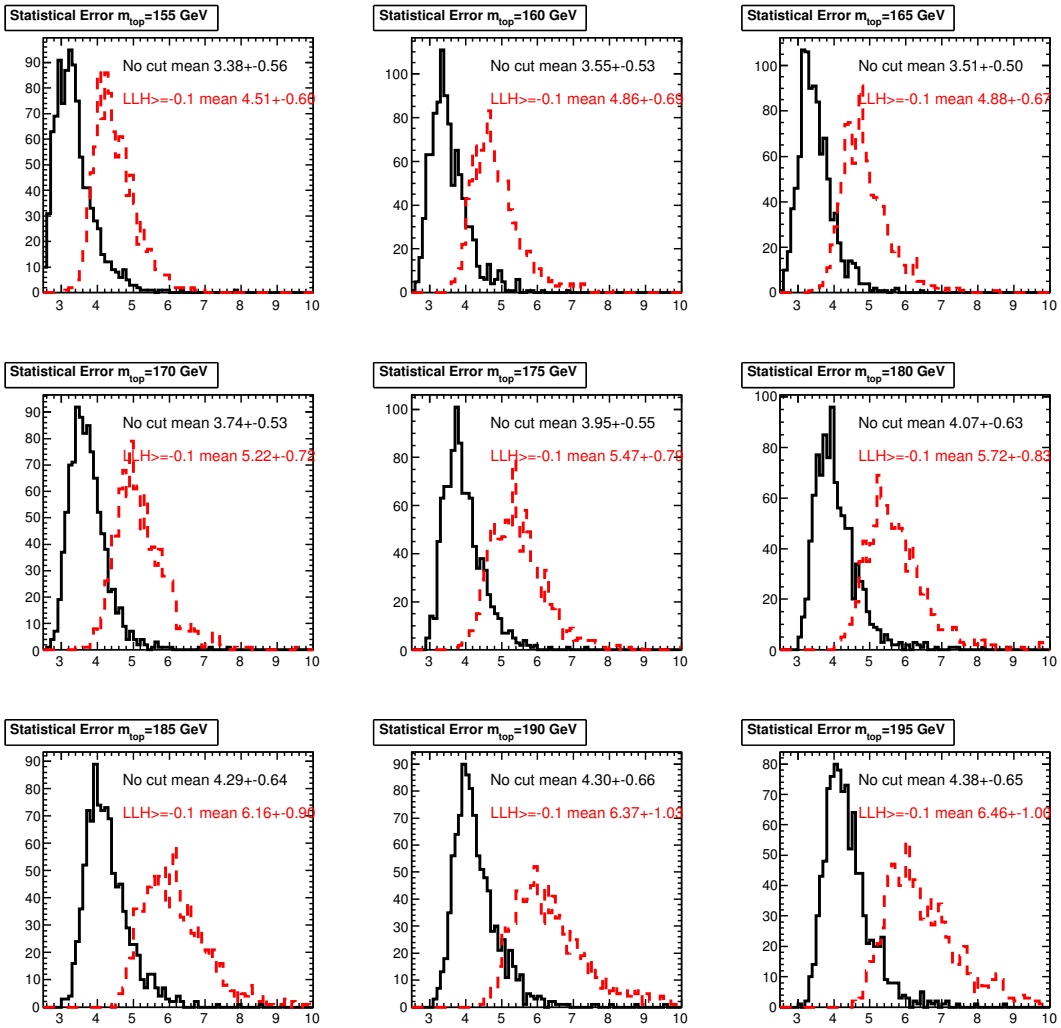

Figure 8.27. Mass error comparison with and without Fisher discriminant cut, in exactly 6 jets, 2 BIDNN > 0.65 case. For different input top mass the mass error distributions with discriminant cut ≥ -0.1 are always higher.

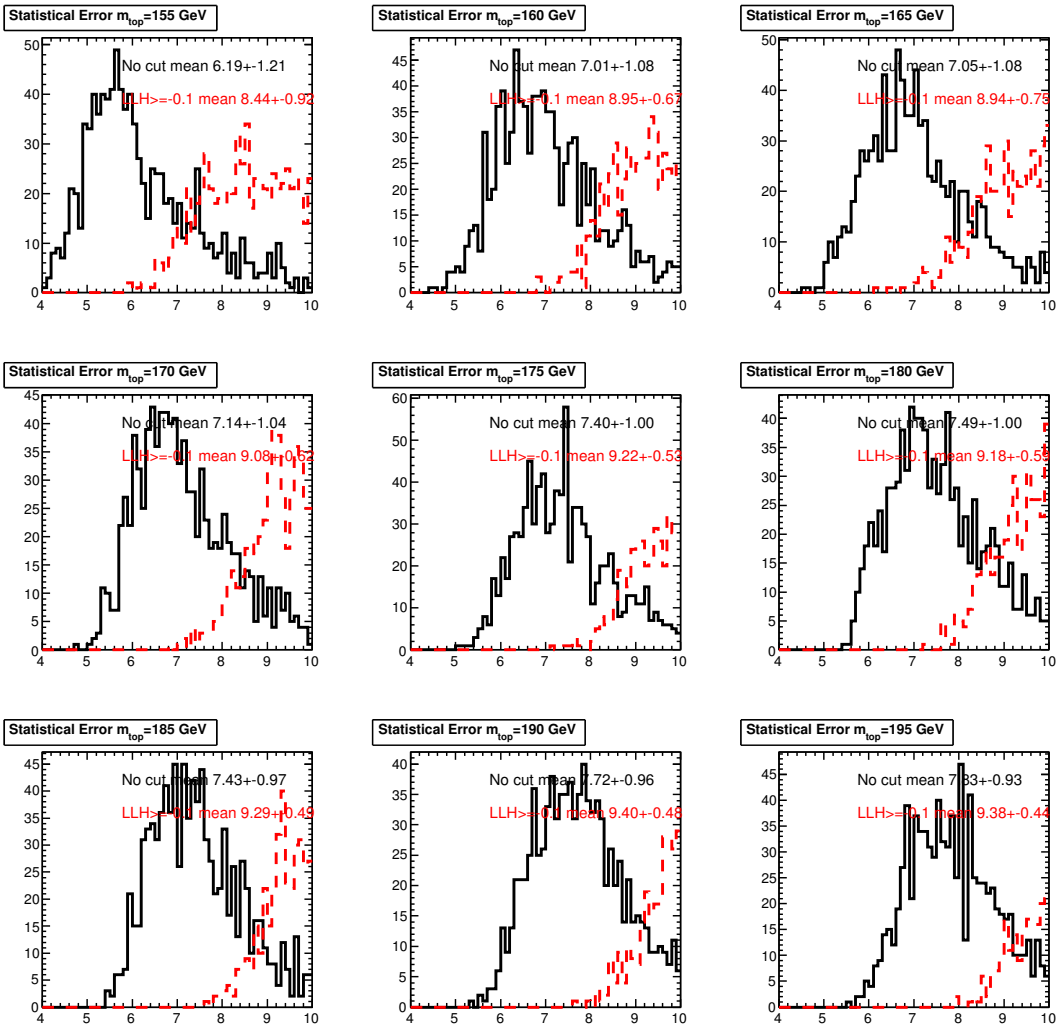

Figure 8.28. Mass error comparison with and without Fisher discriminant cut, in 7 or more jets, 2 BIDNN > 0.2 case. For different input top mass the mass error distributions with discriminant cut ≥ -0.1 are always higher.

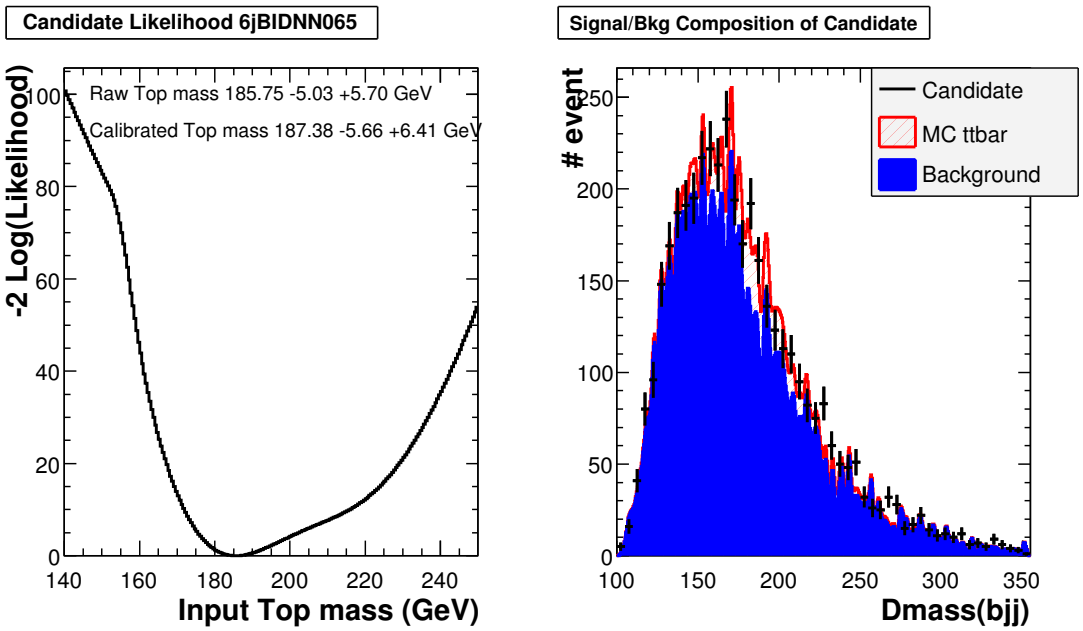

Figure 8.29. Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 6 jets, 2 BIDNN > 0.65 case. The left plot shows the Log(Likelihood) curve which gives the top mass. The right plot shows the signal/background composition according to the fit.

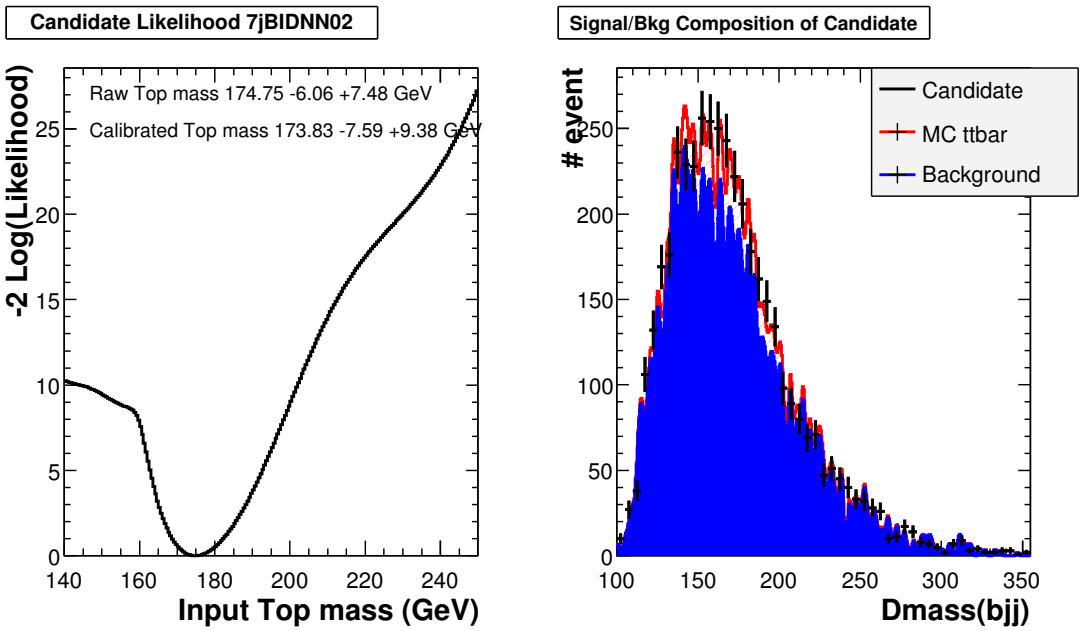

Figure 8.30. Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 7 jets, 2 BIDNN > 0.2 case. The left plot shows the Log(Likelihood) curve which gives the top mass. The right plot shows the signal/background composition according to the fit.

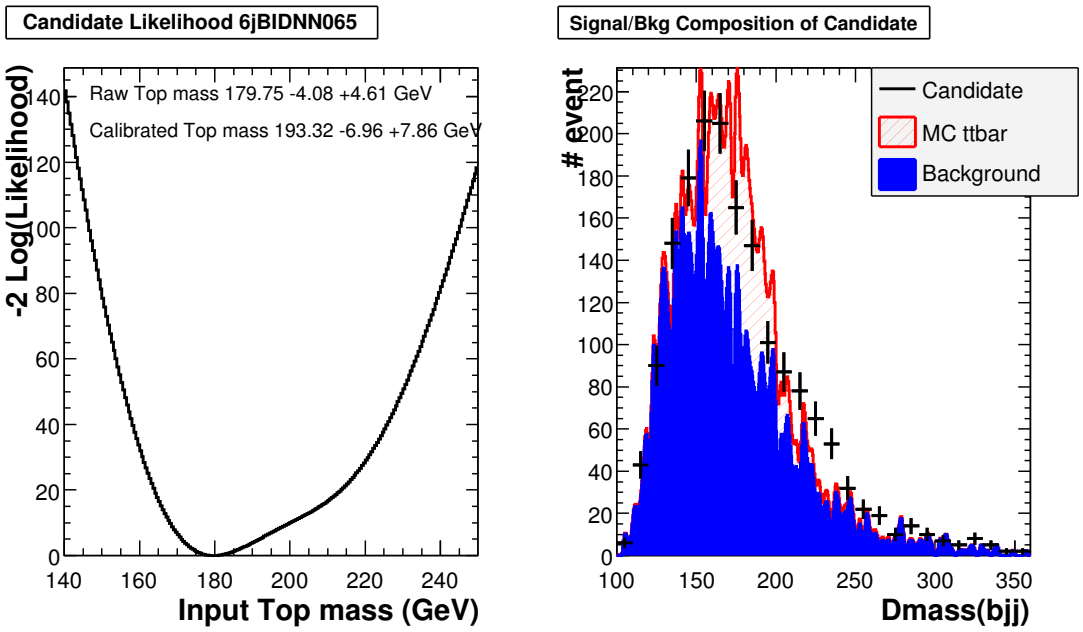

Figure 8.31. Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 6 jets, 2 BIDNN $>$ 0.65 case with Fisher Discriminant > -0.1. The left plot shows the Log(Likelihood) curve which gives the top mass. The right plot shows the signal/background composition according to the fit.

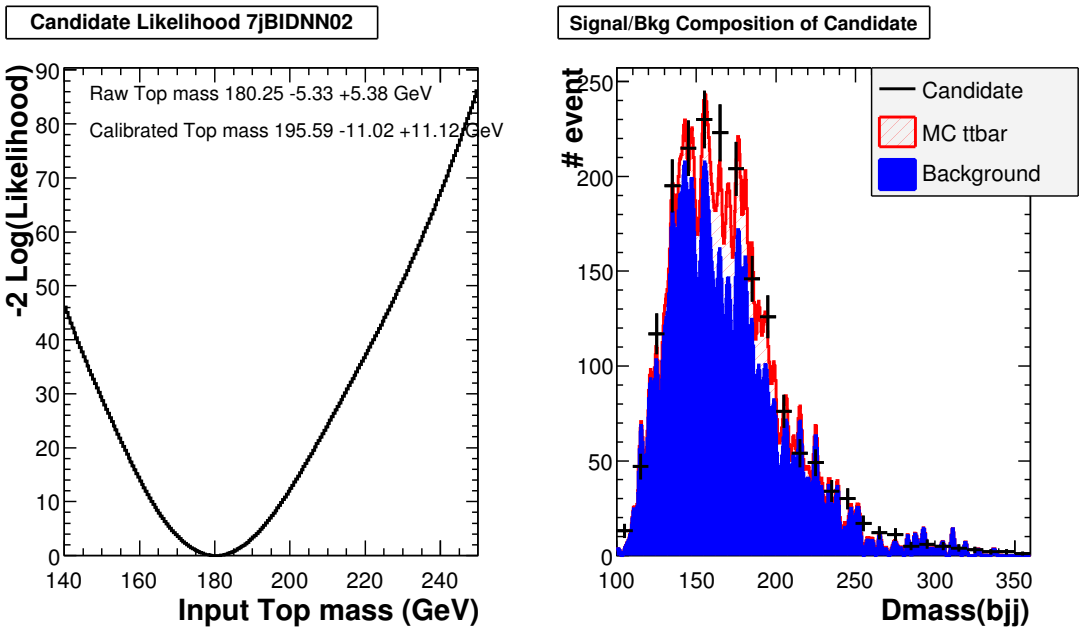

Figure 8.32. Mass Likelihood and fitted composition of the candidate events using "5+x" background in exactly 7 jets, 2 BIDNN $>$ 0.2 case with Fisher Discriminant > -0.1. The left plot shows the Log(Likelihood) curve which gives the top mass. The right plot shows the signal/background composition according to the fit.

CHAPTER 9

RESULTS AND OUTLOOK

With the selection of events in hand, together with all the analysis methods fixed and the calibration of the method performed, the top mass from the all-hadronic channel can be extracted. The exactly 6 jets, 2 BIDNN > 0.65 tagged jets region gives:

$$m_{top} = 187.3^{+6.4}_{-5.7}(\text{stat.})^{+3.2}_{-4.0}(\text{syst.})\ \text{GeV}/c^2 \tag{9.1}$$

The 7 or more jets, 2 BIDNN > 0.2 tagged jets region gives:

$$m_{top} = 173.8^{+9.4}_{-7.6}(\text{stat.})^{+3.1}_{-7.4}(\text{syst.})\ \text{GeV}/c^2 \tag{9.2}$$

Both of them can then be combined to give:

$$m_{top} = 183.6^{+5.3}_{-4.8}(\text{stat.})^{+3.2}_{-4.6}(\text{syst.})\ \text{GeV}/c^2 \tag{9.3}$$

Currently the best top mass measurement in the all-hadronic channel is from CDF [78], which gives the top mass as:
$m_{top} = 177.0 \pm 3.6(\text{stat.+JES}) \pm 1.6(\text{syst.})\ \text{GeV}/c^2$. Here, we will give a brief comparison of the analyses. First the CDF analysis has a smaller statistical error than the result presented here, because CDF has almost double amount of data

($1.9 fb^{-1}$). Their statistical error is also smaller because they included data samples with 1 b-tagged jet as well. Its template method for mass fitting is similar to this analysis, except that CDF uses a 2-D template which requires the calibration in Jet Energy Scale as well. Still, the uncertainties in the JES between the two analysis are compatible. The background model from the CDF is also data-driven. However, they extrapolate a background dominated 4-jets sample into the signal region using tag-rate parameterizations to estimate the tagging probability of the events. Since a detailed explanation of how the tagging probability is used to generate the background samples has not been presented, it is not possible to determine which background model behaves better. Our fitting and background model errors are, however, much larger.

Given that the current world average of the top mass measurement is $m_{top} = 172.6 \pm 1.4$ GeV/c^2 [29], the top mass measured here only contributes a small fraction to the world average. Nevertheless, the results are compatible with other previous measurements in other channels in DØ and CDF (Figure 1.7). Notice that in general the top all-hadronic mass tends to be the largest, then followed by the lepton+jets measurements, and with dilepton mass measurements being the smallest, which may suggest the inadequacy of the physics/background model involved in the top analysis, or other systematic effects that shift the mass. This could be something specific to the all-jets final state, such as color reconnection of some other final state interaction [80], [81], [**?**].

There is no previous published Run-II top mass measurement in the all-hadronic channel from DØ . So for comparison we pick the top mass measurement in the lepton+jets channel in DØ using the same Run-IIa data set [79], which observes $m_{top} = 170.5 \pm 1.8(\text{stat}) \pm 1.6(\text{JES}) \pm 1.2(\text{syst.})$ GeV/c^2. The analysis

methods between that and our analysis are similar except that for signal modeling the l+jets analysis performs a matrix element calculation, and the background samples the l+jets uses are from Monte Carlo simulation. The l+jets analysis has smaller statistical errors compared to our analysis; this is because the decay products contain 4 jets + 1 lepton which have much fewer background events in the data sample. The JES error is comparable to our analysis. The systematic errors of the l+jets are smaller because of the simpler final state, fewer combinatorics to form top candidates, and fewer background events. Also it is free of fitting errors because it involves a more precise (but very time-consuming) matrix element calculation.

The dominating uncertainty in our result is statistical error, which is due to the limited statistics available in DØ Run-IIa and the fact that signal-background separation is very difficult. As data-taking in DØ Run-IIb is already underway, several improvements can be made if the same analysis is carried out again. Run-IIb is expected to have a total of 7 times more data than Run-IIa, which shall reduce the statistical error and also improve the accuracy of the background model (since the number of background events generated depends on the size of the data). The increased amount of data also provides a motivation to increase the Monte Carlo $t\bar{t}$ statistics accordingly, which will then improve the signal template fit as well. Also the trigger list in Run-IIb is more stable and therefore the trigger efficiencies can be modeled more accurately for the Monte Carlo samples. The Run-IIb also has an additional layer 0 within the SMT, which improves the b-tagging efficiencies. Given that there will be $\approx 8fb^{-1}$ of data by the end of Run II, the statistical uncertainty of the top mass from the all-hadronic channel is anticipated to be 5 GeV/$\sqrt{8}$ = 1.77 GeV. The systematic uncertainty is expected

to improve as well as can be seen from the lepton+jets mesurements [29], where the systematic error reduces by a factor around 0.8 when the data size doubles from $1fb^{-1}$ to $2fb^{-1}$. Therefore in the all-hadronic case the systematic error is anticipated to be $3.9 * 0.8^3 \approx 2.0$ GeV, which would combine with the statistical error to give a mass uncertainty of ≈ 2.7 GeV.

APPENDIX A

Trigger Properties

Here are detailed explanation of the 3JT and 4JT triggers used in my analysis. This includes the names and requirement of each L1, L2 and L3 terms involved.

A.1 3JT Triggers

v8	3JT15
L1	CJT(3,7)
	3 L1 cal trigger towers with $E_T > 7$ GeV
L2	none
L3	3JT15
	3 L3 jets with $E_T > 15$ GeV and $\lvert\eta_{det}\rvert < 3$

v9-v10	3JT15
L1	CJT(4,5)
	4 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT50
	3 L2 jets with $E_T > 8$ GeV and $H_T > 50$ GeV
	(using L2 jets with $E_T > 5$ GeV)
L3	3JT15
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3$

v11	3JT15_PV
L1	CJT(3,5)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT50
	3 L2 jets with $E_T > 8$ GeV and $H_T > 50$ GeV
	(using L2 jets with $E_T > 5$ GeV)
L3	3JT15_2JT25_PV
	3 L3 jets with $E_T > 15$ GeV + $\|\eta_{det}\| < 3$
	2 L3 jets with $E_T > 25$ GeV + $\|\eta_{det}\| < 3$ and primary vertex info

v12	3J15_2J25_PVZ
L1	CJT(3,5)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT50
	3 L2 jets with $E_T > 8$ GeV and $H_T > 50$ GeV
	(using L2 jets with $E_T > 5$ GeV)
L3	3JT15_2JT25_PVZ
	3 L3 jets with $E_T > 15$ GeV + $\|\eta_{det}\| < 3$,
	2 L3 jets with $E_T > 25$ GeV + $\|\eta_{det}\| < 3$, primary vertex $\|Z\| < 35$ cm

v13a	JT2_3JT15L_IP_VX
L1	CJT(3,4eta2.4)CJT(2,5)_ncu - v13.03
	3 L1 cal trigger towers with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.4$,
	2 L1 cal trigger rowers with $E_T > 5$ GeV
L2	3JT6_HT70
	3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (jets with $E_T > 6$ GeV)
L3	3JT15_2JT25_IP_PVZ
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$,
	>= 1 jet with b prob. > 0.05 and primary vertex $\|Z\| < 35$ cm

v13b	JT2_3JT15L_IP_VX
L1	CJT(3,5)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT6_HT70
	3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (jets with $E_T > 6$ GeV)
L3	3JT15_2JT25_IP_PVZ
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$,
	$>= 1$ jet with b prob. > 0.05 and primary vertex $\|Z\| < 35$ cm

v14a	JT2_3JT15L_IP_VX
L1	CJT(3,5)CJT(3,4eta2.6)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV, 3 L1 cal trigger tower
	with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.6$
L2	3JT6_ETA26HT75
	3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV
	(jets with $E_T > 6$ GeV and $\|\eta_{det}\| < 2.6$)
L3	3JT15_2JT25_IP_PVZ
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$,
	$>= 1$ jet with b prob. > 0.05 and primary vertex $\|Z\| < 35$ cm

v14b	JT2_3JT15L_IP_VX
L1	CJT(3,5)CJT(3,4eta2.6)CJT(1,7eta1.8)_ncu
	3 L1 cal trigger towers with $E_T > 5$,
	3 L1 cal trigger tower with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.6$,
	1 L1 cal trigger tower with $E_T > 7$ GeV and $\|\eta_{det}\| < 1.8$
L2	3JT6_ETA26HT75
	3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV
	(jets with $E_T > 6$ GeV and $\|\eta_{det}\| < 2.6$)
L3	3JT15_2JT25_IP_PVZ
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$,
	$>= 1jet$ with b prob. > 0.05 and primary vertex $\|Z\| < 35$ cm

A.2 4JT Triggers

v8	4JT10
L1	CJT(4,5)
	4 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8
	3 L2 jets with $E_T > 8$ GeV
L3	4JT10
	4 L3 jets with $E_T > 10$ GeV and $\|\eta_{det}\| < 3$

v9-v10	4JT10
L1	CJT(4,5)
	4 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT90
	3 L2 jets with $E_T > 8$ GeV and $H_T > 90$ GeV (jets with $E_T > 5$ GeV)
L3	4JT10_2JT20
	4 L3 jets with $E_T > 10$ GeV and $\|\eta_{det}\| < 3$,
	2 L3 jets with $E_T > 20$ GeV and $\|\eta_{det}\| < 3$

v11	4JT10
L1	CJT(4,5)_ncu
	4 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT90
	3 L2 jets with $E_T > 8$ GeV and $H_T > 90$ GeV (jets with $E_T > 5$ GeV)
L3	4JT10_2JT20
	4 L3 jets with $E_T > 10$ GeV and $\|\eta_{det}\| < 3$,
	2 L3 jets with $E_T > 20$ GeV and $\|\eta_{det}\| < 3$

v12	4JT12
L1	CJT(3,5)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT8_HT50
	3 L3 jets with $E_T > 8$ GeV and $H_T > 50$ GeV (jets with $E_T > 5$ GeV)
L3	4JT12_3JT15_2JT25
	4 L3 jets with $E_T > 12$ GeV and $\|\eta_{det}\| < 3.6$,
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$

v13a	JT2_4JT12L_HT
L1	CJT(3,4eta2.4)CJT(2,5)_ncu
	3 L1 cal trigger towers with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.4$,
	2 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT6_HT70
	3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (jets with $E_T > 6$ GeV)
L3	4JT12_3JT15_2JT25_HT120
	4 L3 jets with $E_T > 12$ GeV and $\|\eta_{det}\| < 3.6$,
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$ and $H_T > 120$ GeV
	(L3 jets with $E_T > 9$ GeV and $\|\eta_{det}\| < 3.6$)

v13b	JT2_4JT12L_HT
L1	CJT(3,5)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV
L2	3JT6_HT70
	3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (l2 jets with $E_T > 6$ GeV)
L3	4JT12_3JT15_2JT25_HT125
	4 L3 jets with $E_T > 12$ GeV and $\|\eta_{det}\| < 3.6$,
	3 L3 jets with $E_T > 15$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 25$ GeV and $\|\eta_{det}\| < 3.6$ and $H_T > 125$ GeV
	(L3 jets with $E_T > 9$ GeV and $\|\eta_{det}\| < 3.6$)

v14a	JT2_4JT20
L1	CJT(3,5)CJT(3,4eta2.6)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV,
	3 L1 cal trigger towers with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.6$
L2	3JT6_ETA26HT75
	3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV
	(jets with $E_T > 6$ GeV and $\|\eta_{det}\| < 2.6$)
L3	4JT20_2JT25
	4 L3 jets with $E_T > 10$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 20$ GeV and $\|\eta_{det}\| < 3.6$

v14b	JT2_4JT20
L1	CJT(3,5)CJT(3,4eta2.6)CJT(1,7eta1.8)_ncu
	3 L1 cal trigger towers with $E_T > 5$ GeV, 3 L1 cal trigger
	towers with $E_T > 4$ GeV and $\|\eta_{det}\| < 2.6$, 1 L1 cal trigger tower
	with $E_T > 7$ GeV and $\|\eta_{det}\| < 1.8$
L2	3JT6_ETA26HT75
	3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV
	(jets with $E_T > 6$ GeV and $\|\eta_{det}\| < 2.6$)
L3	4JT20_2JT25
	4 L3 jets with $E_T > 10$ GeV and $\|\eta_{det}\| < 3.6$,
	2 L3 jets with $E_T > 20$ GeV and $\|\eta_{det}\| < 3.6$

BIBLIOGRAPHY

1. D. Griffiths, *Introduction to Elementary Particles*, John Wiley & Sons (1987) ISBN 0471-60386-4.

2. S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19 (1967), 1264-1266.
   S.L. Glashow, Nuclear Phys. 22 (1961), 579.
   A. Salam and J.C. Ward, Phys. Lett. 13 (1964), 168.

3. S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. 74 (1995), 2632-2637, hep-ex/9503003.

4. F. Abe et al., *Obersvation of top quark production in $p\bar{p}$ collisions*, Phys. Rev. Lett. 74 (1995), 2626-2631, hep-ex/0612052.

5. O.W. Greenberg, *Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons*, Phys. Rev. 135 (1964), B1447-1450.

6. M.Y. Han and Y. Nambu, Phys. Rev. 139 (1965), B1006.

7. D.J. Gross and F. Wilczek, *Ultraviolet Behaviour of Nonabelian Gauge Theories*, Phys. Rev. Lett. 30 (1973), 1343-1346.

8. H.D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. 30 (1973), 1346.

9. J. Schwinger, *Selected Papers on Quantum Electrodynamics*, Dover (1958), ISBN 0486-60444-6.

10. P.W. Higgs, *Broken Symmetries and the Masses of Gauge Fields*, Phys. Rev. Lett. 13 (1964), 508.

11. F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Bosons*, Phys. Rev. Lett. 13 (1964), 321.

12. J. Goldstone, A. Salam and S. Weinberg, *Broken Symmetries*, Phys. Rev. 127 (1962), 965.

13. A. Sirlin, *Radiative Corrections in the $SU(2)_L \times U(1)$ theory: A simple renormalization framework*, Phys. Rev. D22 (1980), 971-981.

14. M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press (1995), ISBN 0201-50397-2.

15. S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, hep-ph/08041476, Apr. 2008.

16. J. Alwall et. al. *MadGraph/MadEvent v4: The New Web Generation*, JHEP 0709, 028 (2007), hep-ph/07062334.

17. J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, JHEP 0207, 012 (2002).

18. R.M. Barnett et. al., *Review of Particle Properties*, Phys. Rev. D54 (1996), 1.

19. N. Cabibbo, Phys. Rev. Lett, 10, (1963), 531.

20. M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49 (1973), 652.

21. M. Jezabek and J.H. Kuhn, Nucl. Phys. B314, 1 (1989)

22. N. Kidonakis and R. Vogt, *Top quark production at the Tevatron at NNLO*, hep-ph/0309045, Sep. 2003.

23. Martin C. Smith and Scott S. Willenbrock, *Top-quark pole mass*, Phys. Rev. Lett. 79 (1997), 3825, hep-ph/9622329.

24. J. Fleischer, F. Jegerlehner, O.V. Tarasov and O.L. Veretin, *Two-loop QCD corrections of the massive fermion propagator*, Nucl. Phys. B539 (1999), 671-690, hep-ph/9803493.

25. N. Polonsky and S. Su, *More Corrections to the Higgs Mass in Supersymmetry*, Phys. Lett. B508 (2001), 103-108, hep-ph/0010113, and references within.

26. S. Willenbrock, *The Standard Model and the Top Quark*, hep-ph/0211067.

27. R. Arnowitt and P. Nath, *Supersymmetry and sugergravity: Phenomenology and Grand Unification*, hep-ph/9309277, Sep. 1993.

28. S. Heinemeyer, W. Hollik, D. Stockinger, A.M. Weber and G. Weiglein, *Precise Prediction for M_W in MSSM*, JHEP 0608 (2006), 052, hep-ph/0604147.

29. Tevatron Electroweak Working Group, *A Combination of CDF and DØ Results on the Mass of the Top Quark*, hep-ex/08031683, Mar. 2008, and references within.

30. V. M. Abazov et al., *The upgraded DØ detector*, Nucl. Intrum. Meth. A565 (2006), 463-537.

31. *Fermilab's Chain of Accelerators*, `http://www-bd.fnal.gov/public/chain.html`. Fermilab Beams Division, *Run II Handbook*, `http://www-bd.fnal.gov/runII/index.html`.

32. E. Kajifasz (for the DØ Collaboration), *DØ Silicon Microstrip Tracker for RunIIa*, Nucl. Intrum. Meth. A511 (2003), 16-19.

33. *The DØ Upgraded Central Fibre Tracker, Technical Design Report*, `http://d0server1.fnal.gov/projects/SciFi/cfthome.html`.

34. J. Brzezniak et. al., *Conceptual Design of a 2 Tesla Superconducting Solenoid for the Fermilab DØ Detector Upgrade*, FERMILAB-TM-1886, DØ Note 2167 (1994).

35. *Design Report of Central Preshower Detector for the DØ Upgrade*, (1996)

36. A. Gordeev et. al., *Technical Design Report of the Forward Preshower Detector for the DØ Upgrade*, DØ note 3445 (1998).

37. J. Kotcher (for the DØ Collaboration), *Design, Performance and Upgrade of the DØ Calorimeter*, FERMILAB-Conf-95/007-E, DØ note 2417 (1995).

38. J. Bulter et. al., *The DØ Muon System Upgrade*, DØ note 2780 (1996).

39. M. Abolins et al., *DØ Run II Level 1 Trigger Framework Technical Design Report*, DØ note 328 (1998).

40. D. Edmunds et al., *Technical Design Report for the Level 2 Global Processor*, DØ note 3402 1998.

41. A. Boehnlein et al., *Description of DØ L3 Trigger software components*, DØ note 3630 1999.

42. A. Khanov, *HTF: histogramming method for finding tracks; the algorithm description*, DØ note 3778 (2000).

43. G. Borissov, *Technical Details of AA Tracking*, presented at ALL DØ metting, Februrary 28, 2003.

44. R.E. Kalman, *A new Approach to Linear Filtering and Prediction Problems*, Tansaction of the ASME - Journal of Basic Engineering, Vol. 82 (1960), 35-45. R.E. Kalman and R.S. Bucy, *New Results in Linear Filtering and Prediction Theory*, Tansaction of the ASME - Journal of Basic Engineering, Vol. 83 (1961), 95-107.

45. H. Greenlee, *The DØ kalman track fit*, DØ note 4303 (2003).

46. M. Narain et al., *Primary Vertex Selection*, DØ note 3906 (2001).

47. A. Schwartzman and M. Narain, *Probabilistic Primary Vertex Selection*, DØ note 4042 (2002).

48. E. Busato and B. Andrieu, *Jet Algorithms in the DØ Run II Software: Description and User's Guide*, DØ note 4457 (2004).

49. A. Harel, *Jet ID Optimization*, DØ note 4919 (2006).

50. JES Group, *Jet Energy Scale Determination at DØ Run II (final p17 version)*, DØ note 5382 (2007).

51. Amnon Harel, *Capping the JES muon corrections*, DØ note 5563 (2008).

52. C. Clemente et. al., *The MuonID Certification for p14*, DØ note 4350 (2004).

53. F. Beaudette and J.F. Grivaz, *The Road Method*, DØ note 3976 (2002).

54. A. Schwartzman, M. Narain, *Secondary Vertex Reconstruction using the Kalman Filter*, DØ note 3908 (2004).

55. A. Schwartzman, M. Narain, *Secondary Vertex b tagging using the Kalman Filter Algorithm*, DØ note 3909 (2004).

56. M. Anastasiaie, S. Robinson, T. Scanlon, *Performance of the NN b-tagging Tool on p17 Data*, DØ note 5213 (2007).

57. J. Hegeman, *Luminosity determination and reweighting of Monte Carlo overlay luminoisty for p17 hadronic top analyses* DØ note 5561 (2008).

58. T. Sjostrand et al, *PYTHIA 6.3 Physics and manual*, hep-ph/0308153.

59. M. L. Mangano, M. Moretti, F. Piccinini, R.Pittau and A. D. Polosa, *ALPGEN, a generation for hard multiparton processes in hadronic collisions*, hep-ph/0206193 (2002).

60. R. Brun et al., *Geant3*, CERN-DD/EE/81-1 (1987).

61. N. Makovac and J. F. Grivaz, *Shifting, Smearing and Removing Simulated Jets*, DØ note 4914 (2005).

62. S. Bethke, *Determination of the QCD Coupling α_s*, J. Phys. G26 (2000) R27, hep-ex/0004021.

63. J. D. Bjorken and S. J. Brodsky, *Statistical model for electron-positron annihilation into hadrons*, Phys. Rev. D1 (1970), 1416-1420.

64. G. C. Fox and S. Wolfram, *Observables for the analysis of event shapes in e^+e^- annihilation and other processes*, Phys. Rev. Lett. 41 (1978), 1581.

65. G. C. Fox and S. Wolfram, *Test for planar events in e^+e^- annihilation*, Phys. Lett. B82 (1979), 134.

66. A. Hocker et. al, *TMVA - Toolkit for Multivariate Data Analysis*, CERN-OPEN-2007-007, physics/0703039.

67. I. T. Joliffe, *Priniciple Component Analysis, 2nd ed.*, Springer, NY (2002).

68. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan Publishing, NY (1994).

69. D. Patterson, *Artificial Neural Networks*, Prentice Hall, Singapore (1996).

70. J. D. Bjorken and E. A. Paschos, *Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon*, Phys. Rev. 185, 1975-1982 (1969).

71. M. Gluck, E. Reya and A. Vogt, *Dynamical Parton Distributions Revisited*, Jun. 1998, hep-ph/9806404.

72. D. Stump, J. Huston, J. Pumplin, W.K. Tung, H.L. Lai, S. Kuhlmann and J. F. Owens, *Inclusive Jet Production, Parton Distributions, and the Search for New Physics* , JHEP, 0310 (2003), 046, hep-ph/0303013.

73. J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP, 0207 (2002) 012, hep-ph/0201195.

74. G. Corcella and A.D. *Mitov, Bottom Quark Fragmentation in Top Quark Decay*, Oct. 2001, hep-ph/0110319.

75. M.G. Bowler, *e^+e^- production of heavy quarks in the string model*,

76. C. Peterson et. al., *Scaling Violations in Inclusive e^+e^- Annihilation Spectra*, Phys. Rev. D27 (1983) 105.

77. ALEPH Collaboration, Phys. Lett. B512 (2001) 30.
DELPHI Collaboration, Z. Phys. C73 (1996) 11.
OPAL Collaboration, Eur. Phys. J. C29 (2003) 463.
SLD Collaboration, Phys. Rev. D 65 (2001) 092006.

78. CDF Collaboration, *Measurement of the top mass with in situ jet energy scale calibration in the all-hadronic channel using the Template Method with 1.9 fb^{-1}*, Jan. 2008.

79. DØ Collaboration, *Measurement of the top quark mass with the matrix element method using the lepton+jets 1 fb^{-1} data set*, DØ note 5362 (2007).

80. G. Gustafson, U. Petterson and P.M. Zerwas, Phys. Lett. B209 (1988) 90.

81. T. Sjostrand and V.A. Khoze, *Does the W Mass Reconstruction Survive QCD Effects?* Phys. Rev. Lett. 72 (1994) 28-31, hep-ph/9310276.

82. P. Skands and D. Wicke, *Non-perturbative QCD Effects and the Top Mass at the Tevatron*, hep-ph/0703081.

This document was prepared & typeset with LaTeX 2_ε, *and formatted with* NDdiss2_ε *classfile (v3.0[2005/07/27]) provided by Sameer Vijay.*